

A Priori Probability and Localized Observers.

Matthew J. Donald

**The Cavendish Laboratory, JJ Thomson Avenue,
Cambridge CB3 0HE, Great Britain.**

e-mail: mjd1014@cam.ac.uk

web site: <http://people.bss.phy.cam.ac.uk/~mjd1014>

November 1991

Revised: July 1992

Appears: Foundations of Physics 22, 1111-1172 (1992)

Abstract. A physical and mathematical framework for the analysis of probabilities in quantum theory is proposed and developed. One purpose is to surmount the problem, crucial to any reconciliation between quantum theory and space-time physics, of requiring instantaneous "wave-packet collapse" across the entire universe. The physical starting point is the idea of an observer as an entity, localized in space-time, for whom any physical system can be described at any moment, by a set of (not necessarily pure) quantum states compatible with his observations of the system at that moment. The mathematical starting point is the theory of local algebras from axiomatic relativistic quantum field theory. A function defining the a priori probability of mistaking one local state for another is analysed. This function is shown to possess a broad range of appropriate properties and to be uniquely defined by a selection of them. Through a general model for observations, it is argued that the probabilities defined here are as compatible with experiment as the probabilities of conventional interpretations of quantum mechanics but are more likely to be compatible, not only with modern developments in mathematical physics, but also with a complete and consistent theory of measurement.

CONTENTS

1. **Introduction and synopsis.**
 2. **What are the fundamental concepts for interpretations of universally valid quantum theories?**
 3. **Mathematical background.**
 4. **Postulates about observers.**
 5. **The a priori probability function and measurement theory.**
 6. **A priori probability and observations.**
 7. **Properties of a priori probability.**
 8. **The definition of the a priori probability function.**
 9. **A priori probability for a succession of collapses.**
 10. **Elementary models.**
 11. **Consistency.**
- References**

1. Introduction and synopsis.

There seems to be a wide gap between the entities traditionally seen as central in the interpretation of quantum mechanics (particles, eigenvalues, wave-functions) and the von Neumann algebras associated with localized regions of space-time, which are focused on by mathematical physicists interested in the conceptual foundations of relativistic quantum field theories. The purposes of this paper are; first, to draw attention to this gap; second, to discuss a function measuring relative probabilities between local algebra states which is intended to be a tool for interpretations based on field theory; and third, to develop a framework for defining the a priori probability of a localized observer with a given quantum-mechanically specified life history up to a given time observing a given set of local quantum states. Most of the paper should be accessible to physicists unfamiliar with the mathematics of local algebras.

The characterization of “wave-packet collapse” might well be thought of as the primary problem of the foundations of quantum theory. We understand very well how to model many physical situations, at a given moment, by an appropriate wave-function. The problem is that the wave-function appropriate at one moment appears to change abruptly whenever an act of measurement or an observation occurs. There is no widely-accepted detailed theory of such acts.

In this paper, as in Donald (1990), we shall consider a more general idea of “collapse” as being some process, to be characterized, which results in the discontinuous replacement of one quantum state by another, whether or not these states are pure. This process must, of course, be distinguished from Hamiltonian time propagation, which is continuous, and, indeed, does not cause any change in state if we work, as we always shall here, in the Heisenberg picture. Working in a local space-time region, we are not given wave-functions, but rather states in the operator algebra sense, which correspond, in general, to mixed-state density matrices. Local information is not sufficient to distinguish whether, at the global level, a given state is pure or not.

In the conventional approach to quantum theory, where we work globally, we are given an initial wave-function ψ and a choice $(\varphi_n)_{n=1}^N$ of wave-functions for the result of a collapse. It is then claimed that the a priori probability of the collapse to φ_n is given by the squared amplitude $|\langle \varphi_n | \psi \rangle|^2$. The most fundamental unsolved problems in this approach lie in giving a detailed specification of the sequence $(\varphi_n)_{n=1}^N$ and of the times at which collapse occurs. Of course, the $(\varphi_n)_{n=1}^N$ are usually taken to be the eigenvectors of some “measured operator”. The difficulty lies in identifying this operator without ambiguity, given a real physical measuring apparatus. It is also far from easy to see how such an approach can be made compatible with relativity theory, because of the implicit assumption of a globally simultaneous collapse. The desirability of working locally comes partly from this problem and partly from the manifest fact that observers and observations are intrinsically localized.

In this paper, at the simplest level, the aim is to develop the possibility of considering an initial state ρ to be given on some local region specified by a set of operators \mathcal{B} , and to give an appropriate definition for a function $\text{app}_{\mathcal{B}}(\sigma | \rho)$ to measure the a priori probability of collapse occurring to a new state σ . I have previously considered, in Donald (1986, 1987a), the mathematics of this definition, but, here, I emphasize its physical motivation. By itself, however, the function $\text{app}_{\mathcal{B}}(\sigma | \rho)$ is merely a tool, albeit one which might be useful to any interpreter of quantum mechanics wishing to work with density matrices rather than wave functions. The really challenging question is how that tool might be used in progress towards an interpretation of quantum mechanics based on the idea of locally defined states. This paper, therefore, is mainly concerned with building a framework for such an interpretation. Here too there is a problem of detailed specification, but, I believe, that the framework proposed here is more likely to permit the construction of such a specification than is the framework of conventional quantum mechanics. Indeed, I began such a construction in Donald (1990), based on the idea of the physical structure of “an observer” (or equivalently, of “a consciousness”), as being constituted by an information processor operating through a localized family of abstractly definable quantum switches. The definition

of a quantum switch and a demonstration that such switches exist in the human brain are given in Donald (1990). As an interpretation, even the combination of that paper and this will still be incomplete. This paper gives an analysis of observation processes which allows most of the problems of detailed specification to be passed back to the characterization of the observer. However, only an outline description of an observer is given here. I intend to continue this work in a further paper, already in progress, which will look at observers more closely. Although the ideas in this paper and in Donald (1990) were originally developed together, this paper may be read independently of its predecessor and of its sequel. I hope, in fact, that the material in much of this paper is sufficiently general that it might be possible to use it to develop more than one route towards a complete interpretation of quantum theory.

It is clearly not possible to develop interpretations of quantum theory from universally accepted first principles; comparing at each stage with all the alternatives. Instead, one must choose the first principles which one wishes to use as the fundamental ingredients of an interpretation. One should try to make those ingredients explicit in the form of premises which cannot be totally superseded without the abandonment of the interpretation. Then one can begin to evaluate the consequences of these premises. Disagreements about premises are inevitable, but they tend not to be as fruitful as discussions of their consequences. In this paper, I wish to develop techniques for the calculation of a priori probability, in the context of the premise that quantum field theory is universally valid and the premise that observers are individual, localized, animate repositories of information gained about the universe which are describable in quantum mechanical terms.

The emphasis on individual observers is a direct reflection of the idea that in relativity theory we assign separate proper times to separate observers. This leads to a many-worlds type theory in which, at each moment, each observer assigns his own separate quantum state to the “reality” with which he is interacting. In such a theory, compatibility with relativity is not a problem essentially because “collapse” is an alteration of the state currently assigned by a given observer to a given system rather than a message sent to or from that system. For the same reason, the violation of Bell’s inequality is also not a problem. One must, of course, answer questions about the relationship between the observations of different observers, but, as we shall see, this is no more difficult than giving a general account of observation. The problem which is difficult, on the other hand, is the development and understanding of a coherent formalism for such a theory. This paper forms part of a serious attack on that problem. Without alternative proposals reconciling relativity and quantum theory, merely to dismiss such a many-worlds theory as being philosophically unacceptable is, in my opinion, the least fruitful of all the potential ways of disagreeing with this paper. For myself, I see it as being quite natural to allow each observer his own independent objective existence, his own independent observations, and his own independent a priori probabilities.

The development of these vague premises into a detailed interpretation will come in the form of “postulates” or “hypotheses”. These will convert the premises into specific formalism; providing a model with consequences which may be ascertained

and argued over. They will be open to modification and refinement as the theory develops. In this paper, as in Donald (1990), I shall only be presenting a partial set of postulates. While this is dangerous, since over-all consistency is one of the hardest goals for an interpretation to achieve, it should, nevertheless, be permitted, in view of the difficulty of the problem and in order to encourage variant developments. Widespread dissatisfaction has been expressed about every interpretation of quantum theory proposed in the course of the last sixty years. I suspect that any successful interpretation will have to introduce so many difficult ideas as to be, in its entirety, well-nigh incomprehensible at first sight. It seems to me, therefore, that it is worthwhile trying to discuss separately possible postulates for potential interpretations.

Despite its incompleteness, it may well appear that this paper rapidly becomes swamped by a welter of abstract technical detail when the postulates finally appear in section 4 and thereafter. This is regrettable, although I believe that it is also inevitable. Much of the abstraction arises because we shall be working not with individual quantum states but with set of states. An analogous abstraction allowed tremendous progress in probability theory when the introduction of measure theory required attention to pass from individual events to sets of events. Of course, in the present context, there is a philosophical question about the desirability of ascribing ontological priority to sets of states: Is it desirable to describe an observer as existing as a set of sequences of density matrices? This question is left for the reader to ponder. However, even leaving aside the claim that locally defined neighbourhoods of states are more natural entities within relativistic quantum field theories than globally defined eigenstates of specified operators, it may be noted that traditional measurement theory is at a dead end, and so any theory allowing greater flexibility should be explored.

Synopsis In section 2, it is proposed that the local quantum state is the natural fundamental entity in quantum theory, the idea of working with sets of such states is motivated, seven different notions of quantum probability are introduced, and the importance of correlations between different observables is stressed. Section 3 discusses the mathematical background necessary for the analysis of local states and gives a formal definition of correlation. Attention is drawn to an important but troublesome property of local algebras without which the mathematics which we shall need, might be considerably more straightforward. In order to indicate how a complete interpretation of quantum theory might be based on the concepts central to this paper, some fairly broad postulates about observers are put forward in section 4. Section 5 presents a general model for an observation in terms of a state decomposition with negligible interference effects. This imposes an important property on the function $\text{app}_{\mathcal{B}}(\sigma | \rho)$. Section 6 extends the model of section 5 to the context of the postulates of section 4. In section 7, properties for $\text{app}_{\mathcal{B}}(\sigma | \rho)$ are proposed which are appropriate for a definition of a priori probability and which include the property imposed in section 5. Much of this work involves interpreting the notion of a quantum state. In section 8, a selection of the properties is shown to be sufficient to provide a unique definition which satisfies all the other properties. Section 9 discusses the mathematics of multiple-observation probabilities, section 10 sketches mathematically elementary

models of some of the structures introduced, and the [final](#) section considers the various consistency issues which have arisen.

2. What are the fundamental concepts for interpretations of universally valid quantum theories?

Sometimes, quantum mechanics can seem no more than a disparate collection of calculational recipes. For example, we can calculate the frequency-dependent susceptibility of a family of atoms in a cavity, we can calculate electron-positron annihilation cross-sections, and we can relate the zero-temperature energy gap in weak-coupling superconductivity to the transition temperature. The glory of quantum mechanics, however, is the unity it brings to such disparate topics. It is not just that the methods for calculations are often similar; involving perturbation theory and phenomenological, approximate Hamiltonians. It is also the suspicion that there is a true universal Hamiltonian in some wonderful super-theory of everything, and that all our calculations are valid approximations for different situations in that universal theory.

Of course, it is not known whether footballs, planets, and steam engines are solely governed by quantum mechanical laws. The Copenhagen interpretation of quantum theory demands the existence of a “classical regime” in which macroscopic objects can be described using Newtonian mechanics. On the other hand, [Everett \(1957\)](#), developed the many-worlds interpretation, which he referred to as “The Theory of the Universal Wave Function”, precisely in order to be able to apply quantum theory to the entire universe. In this paper the universal validity of quantum theory will also be postulated. However, the framework of the development here will be different. In constructing a universal quantum theory, Everett conceived a brilliant intuition about the nature of observers which will be discussed below. He expounded his intuition in the framework of elementary quantum mechanics. I shall use Everett’s intuition, but I shall try to expound it in a framework compatible with quantum field theory.

In most interpretations of quantum mechanics, the concept of the individual particle is fundamental. In interacting relativistic quantum field theories, however, local particle number becomes indefinite, because of the possibility of particle (or particle-antiparticle) creation. Only in scattering theories are particle states clearly definable. In other words, either we deal with an idealized infinite-volume theory, or we work locally and accept that the concept of “an individual particle” is as “classical” in relativistic quantum theory as the concept of “a particle with precisely defined position” in non-relativistic quantum theory. Indeed, in a way, the former concept is even more “classical”, since, charge quantum numbers apart, it is not even plausible that it is possible to specify, for example by appropriate renormalizations or dressing transformations, operators unambiguously “measuring” the expected number of particles within a given region. The mathematical construction of quantum field theory models in two-dimensional space-time has provided models in which this sort of problem could be made explicit ([Glimm and Jaffe \(1971, 1972, 1979\)](#)).

It is a first principle in this paper that the quantum measurement problem should be discussed and solved in terms of localized entities – entities that exist and are defined within bounded regions of space-time. In particular, observers and experiments are assumed to be localized in this sense. In order to define local entities, one must be able to define “local quantum states”. Fortunately, a mathematical theory of such states already exists. This theory will be reviewed briefly below. Of course, the idea of the “local quantum state” is natural, and does not need any deep mathematics for an initial physical perspective. From such a perspective, we would say that the quantum state in a space-time region Λ (by convention, a region is an open bounded connected set) is simply the state that one assigns to the objects within Λ . In view of quantum measurement problems, however, this statement does require considerable elucidation and modification. In particular, the word “one” presumably refers to an observer and must require specification of that observer and his time frame. Also, since we have assumed that Λ is extended in time (and this may well be necessary in relativistic field theories (Streater and Wightman (1964) §3.1)), we should expect that it is possible for a process analogous to “wave-packet collapse” to require the assignment of more than one state to Λ . This paper provides tools for attacking these difficulties.

Despite being bounded, the space-time regions that we shall be working with are to be thought of as macroscopic; perhaps, for example, big enough to contain an animate observer over an extended time period. This implies, in particular, the consideration of multiple observations and the assignment to the region of a sequence of several or many states. The theory for single observations is considerably simpler than that for multiple observations. Much of the mathematics of the single observation theory is presented in Donald (1986, 1987a). This paper provides a broader physical framework for that mathematics and also tackles the multiple observation theory.

As already mentioned, a local state is to be thought of as a density matrix, which is, in general, mixed; rather than as a pure state wave-function. There are several grounds for wishing to base interpretations on density matrices rather than on wave-functions. Of course, quantum theory often provides wave-functions as initial models for a given physical system, but it is elementary to use these to construct density matrices. For a first ground, at the mathematical level, it is worth mentioning for the experts, that on a type III von Neumann algebra (the sort relevant for quantum statistical mechanics and local quantum field theory), there is no such thing as a pure normal state. Secondly, at the physical level, we shall be dealing with the best guess that a given observer can make for the current state of a given macroscopic subsystem of the universe. It is certainly more natural for him to assign thermodynamical properties, for example, entropy and temperature, to such subsystems than to try to assign, for example, an exact eigenstate of some, not necessarily well-defined, energy operator. Finally, at the formal level, it is often claimed that one can simply replace a density matrix by an element of its eigenexpansion. However, this is difficult to justify because the relationship between the two is both ambiguous (e.g. $\frac{1}{2}|\psi_1\rangle\langle\psi_1| + \frac{1}{2}|\psi_2\rangle\langle\psi_2| = \frac{1}{4}|\psi_1 + \psi_2\rangle\langle\psi_1 + \psi_2| + \frac{1}{4}|\psi_1 - \psi_2\rangle\langle\psi_1 - \psi_2|$ for orthogonal ψ_1 and ψ_2) and unstable (e.g. $(\frac{1}{2} + \varepsilon)|\psi_1\rangle\langle\psi_1| + (\frac{1}{2} - \varepsilon)|\psi_2\rangle\langle\psi_2|$ is arbitrarily close to

($\frac{1}{4} + \varepsilon$) $|\psi_1 + \psi_2\rangle\langle\psi_1 + \psi_2| + (\frac{1}{4} - \varepsilon)|\psi_1 - \psi_2\rangle\langle\psi_1 - \psi_2|$ for sufficiently small $\varepsilon > 0$). The assumption that this relationship is easily dealt with is one of the major problems with, for example, Everett's formalism. By always working at the density matrix level, we can give a much more sophisticated analysis and a generalization of this relationship; thus, allowing, for example, for imprecision. This is one of the central tasks of this paper.

Our starting point, nevertheless, is that for a density matrix

$$\rho = \sum_{i=1}^{\infty} r_i |\psi_i\rangle\langle\psi_i| \quad (2.1)$$

where $(\psi_i)_{i=1}^{\infty}$ is an orthonormal basis of \mathcal{H} and $0 \leq r_i \leq 1$ with $\sum_{i=1}^{\infty} r_i = 1$, r_i should be the a priori probability of going from ρ to $|\psi_i\rangle\langle\psi_i|$ when such a collapse is warranted. A decomposition of the form (2.1) will be recognized as a decomposition "without interference effects". Justifying such a decomposition for "the true physical state" in appropriate circumstances has long been seen by many to be a solution to the problems of measurement theory. However, while it seems to be widely accepted that, in as far as interference effects can often be physically negligible, such decompositions can often be plausible models for physical situations, it also seems that justification from first principles has been lacking or unsatisfactory. The validity of such a decomposition depends on an appeal to non-unitary, irreversible time propagation or to some sort of "coarse-graining". In this paper, coarse-graining comes from the restriction to localized regions, and the localization of observers is a first principle. The realization that coarse-graining arises naturally from the incomplete, non-universal, structure of an observer is precisely the brilliant intuition which I ascribed above to Everett.

No less than seven different notions of quantum probability arise in this paper or are in common usage. The first notion is that of the a priori probability of observation of a subsystem occupying a given set of states by an observer with a given life history up to a particular time. This notion is presented in section 4. The second notion is the function $\text{app}_{\mathcal{B}}(\sigma | \rho)$ measuring the a priori probability of the collapse of the state ρ on \mathcal{B} to the state σ . The third notion involves taking a density matrix as a mixture of components weighted by probabilities, as exemplified by (2.1). $\text{app}_{\mathcal{B}}(\sigma | \rho)$ will be constructed as a generalization of this notion. The fourth notion is the idea of the expected value of a projection P in a state ρ . Recall that if $\rho = \sum_{i=1}^{\infty} r_i |\psi_i\rangle\langle\psi_i|$ is a state and $P = \sum_{n=1}^{\infty} |\varphi_n\rangle\langle\varphi_n|$ is a projection, for suitable (orthonormal) wavefunctions ψ_i and φ_n , then this expected value is defined as

$$\rho(P) = \sum_{n,i} r_i |\langle\psi_i|\varphi_n\rangle|^2. \quad (2.2)$$

The left hand side of (2.2) uses an alternative notation for $\text{tr}(\rho P)$, which is basic to the mathematical analysis of states on von Neumann algebras and which will be used throughout this paper (see 3.3). Notion four is also a generalization of notion three, since r_i is the expected value of the projection $|\psi_i\rangle\langle\psi_i|$ in the state given by (2.1). This notion is fundamental to textbook quantum theory; being at the heart of such calculated quantities as scattering cross sections and decay rates

where it often appears in the guise of the fifth notion which is that of probability as amplitude squared. Indeed, whenever this arises as a “transition probability” in textbook calculations, it can always be interpreted as a special case of notion four.

The sixth notion may be referred to as “relative frequency”, or, more generally, as “statistical probability”. A particle lifetime, for example, equivalent to a probability per unit time for decay, is not directly observable by a single measurement on one particle. Instead, statistics have to be gathered from many measurements. The quoted “observed” lifetime is found by an analysis of these statistics. Statistical probability itself comes in many different varieties, depending not only on the analysis made of a given set of measurements, but also on the particular methods of observation. The empirical justification of conventional quantum theory lies in the observed agreement between these statistical probabilities and theoretical calculations of probabilities of the third, fourth, and fifth kinds. In section 6, it will be argued that this empirical evidence yields just as strong a justification for a theory based on the first notion.

Ultimately, each of these six notions depend on, and are consistent with, the seventh, and fundamental, notion of the “typical” observer. The typical observer observes a world in which, when suitably applied, the first six notions are almost always consistent. The typical observer also lies at the heart of the justification of classical probability theory, because it is only for the typical observer that relative frequencies measured over long periods usually come very close to predicted probabilities. The typical observer may occasionally win a lottery, but cannot win every lottery he enters. The deepest idea in this paper is a definition of a priori probability for the families of sets of states occupied by individual observers. Typical observers are those for which this a priori probability is relatively high. In quantum theory, as in classical theory, any method of calculating probabilities will be justified in as far as it can be argued that typical observers will tend to observe relative frequencies which agree, in the long run, with those calculations. The first five notions of quantum probability described [above](#) provide tools with which, in appropriate circumstances, such calculations can be made.

By choosing appropriate local quantum states for a given physical situation, it is possible to model the “different situations” which distinguish between, for example, regions where there are lasers, and laser theory applies, and regions where there are superconductors, and superconductivity theory applies. Alternatively, appropriate local quantum states might be chosen to be approximate scattering states and could, say, represent “close to” five “reasonably well-separated” particles initially at “around” five given places and with “roughly” five given momenta. Elementary quantum mechanics tells us how to define “wave-packets” and more sophisticated theories build on the same intuition. There is, however, a central dilemma here, exemplified by the fact that a wave-packet is neither a position nor a momentum eigenstate (of course, it could not be so, because none such exist within the space of normable wave-functions). The dilemma is that, while a range of approximating wave-packets is possible, with limits set only by the uncertainty principle, there is no definitive theory for choosing within that range. The essence of the problem of state collapse is that we cannot merely assume that some state within the range is given to us by an

ultimately deterministic dynamics – we need to understand the processes by which the states that we do see come to be seen.

In the standard interpretation of quantum mechanics, the resolution of this dilemma is supposed, somehow, to lie in the correct description of the measurement process and its equivalent, the state preparation process. Any measurement is supposed, somehow, to be a measurement of some given, predetermined, operator with a given discrete spectrum. The association between measuring apparatus and measured operator, however, is seldom made explicit, and there seem to be no generally applicable techniques for defining such an association. It is not clear, for example, that a bubble chamber experiment is not just as much a measurement of bubble positions as of elementary particle positions.

As a first step in reaction to this state of affairs, we shall defocus; allowing that there is a wide variety of operators that one is, to some approximation, measuring. We shall, therefore, consider a measuring apparatus or an observation as describable not in terms of eigenvectors of a given operator, but merely in terms of a given family of neighbourhoods of local quantum states. For example, in the traditional approach, the results of an energy measurement will be the eigenvalues $\{E_i : i \in I\}$ of a prescribed Hamiltonian H . These correspond to density matrices $\{\sigma_i : i \in I\}$ such that $\sigma_i((H - E_i)^2) = 0$. The methods introduced in this paper would allow the results of an approximate energy measurement, or a measurement in which H is an approximate Hamiltonian, to correspond to sets of density matrices – say,

$$\{S_i : i \in I\} \text{ where } S_i = \{\sigma : \sigma((H - E_i)^2) < \varepsilon\} \text{ for some } \varepsilon > 0. \quad (2.3)$$

While an appropriate formalism for any sort of approximate measurement could be developed from this paper, the focus here will be on allowing for small variations in the states assigned by conventional measurement theory. Indeed, we shall focus on unobservably small variations, because these will still be enough to avoid the assumption, for example, that a particle must occupy an exactly specified wave-packet. It is possible to allow such variations because there is no need to try to associate definitive structure with each measuring apparatus in a theory which gives the cardinal role to the observer. The measurement made by the apparatus need be specified only to the extent that the observer's observations of the apparatus requires the apparatus states to be defined. Ultimately, the “observed apparatus” in which we shall be most interested and to which these remarks will still apply, is the observer's own brain.

The distinction between different neighbourhoods of states is made by the expected values which these states assign to various operators. These operators could include any of the possible operators which we might traditionally think that we were measuring. Thus, in a bubble chamber experiment, the results are distinguished both by the original elementary particle positions and by the bubble positions, because the entire purpose of the apparatus is to correlate these variables. More generally, the purpose of quantum mechanical experiments is to correlate variables linked to microscopic objects with variables linked to macroscopic objects. Bohr saw such macroscopic variables as constituting a “classical regime”. In my opinion, it is their observed behaviour which is most important, rather than any direct relation to Newtonian mechanics. A macroscopic variable is one which is always seen by observers to

take definite values in such a way that different values are only seen to be taken on states which differ extensively throughout macroscopic space-time regions. Explaining the existence of such variables is a major task for an interpretation of quantum theory; one which will need a considerable input from statistical mechanics as well as a theory of observers.

Macroscopic variables are always extensively correlated with other macroscopic variables. At a descriptive level, any one of a wide variety of correlated macroscopic variables will suffice to distinguish the different possible results of an experiment. The a priori probability of a given result will be determined, to a good approximation, by the initial expectation value of any of a wide variety of projection operators correlated to these macroscopic variables. The compatibility between these various approximations suggests the possibility of replacing textbook quantum mechanics by an alternative theory without violating the compatibility between theory and observation. We shall give this a more formal treatment in sections 5 and 6. Macroscopic situations can be of great complexity. For example, one can observe not only an experimental result, but also the reaction of a colleague to that result. The fact that different observers will, in general, agree between themselves about the result of an experiment, is reflected, at the theoretical level, in correlations between appropriate macroscopic variables. Thus a theory which deals adequately with correlations will automatically be a theory in which the observations of different observers appear to be compatible.

The range of correlated macroscopic variables makes it difficult to choose definitive observables for each particular experiment. Ultimately, however, the purpose of any experiment is to cause changes in the brain of the observer. For this reason, it is possible to reduce all observations to neural observations. Indeed, this is the only way that I can see of finding “simple”, “natural” descriptions, sufficient to specify arbitrary observations. Moreover, I think it is then necessary to use abstract definitions and to give an abstract definition of an observer as an information processor. All entities satisfying such a definition are to be allowed equivalent status as possible manifestations of the observer, each with its own a priori probability. This led in Donald (1990) to the introduction of a “quantum switch” as an information processing primitive. Abstractly specified neighbourhoods of states were used in that definition partly in order to allow for perturbations and yield a “structurally stable” theory and partly in order to deal with the fact that there do seem to be many alternative sets of equivalent quantum switches with which a human brain can be seen as functioning as an information processor. In section 4, I introduce a preliminary model, compatible with that work, of an observer as a neighbourhood of sequences of quantum states with correlations expressed on a distinguished set of observables. In the sequel to this paper, the ideas in section 4 and in Donald (1990) will be developed to give a full abstract model of an observer. Leaving the details of all this to one side, however, there remains plenty to do in this paper in finding a method for calculating probabilities in such a context.

3. Mathematical Background.

At an elementary level, a quantum mechanical system S is described by a Hilbert space \mathcal{H}_S of wave functions defined in terms of variables appropriate to that system. If the Hilbert space for the entire universe is taken to be \mathcal{H} , then it is assumed that there is another system S' describing the rest of the universe such that \mathcal{H} can be written as a tensor product: $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_{S'}$. Observables on the system S are defined as operators on \mathcal{H}_S , that is as elements of $\mathcal{B}(\mathcal{H}_S)$ – the space of all bounded operators on \mathcal{H}_S . For simplicity of language, I shall use the word “observable” to refer to any bounded operator, not just to a Hermitian operator. Also, I shall not distinguish between an operator $A \in \mathcal{B}(\mathcal{H}_S)$ and the operator $A \otimes 1_{S'} \in \mathcal{B}(\mathcal{H})$, where $1_{S'}$ is the identity operator on $\mathcal{H}_{S'}$. The states of S are defined to be the density matrices of \mathcal{H}_S .

A mathematically more sophisticated approach allows for more general sets of operators to correspond to the observables of S . This approach says that to S corresponds a von Neumann algebra \mathcal{A}_S , and that physical states of S correspond to normal von Neumann algebra states on \mathcal{A}_S . It will always be assumed here that there is a global Hilbert space \mathcal{H} describing the entire universe. Every subsystem von Neumann algebra \mathcal{A} will then be a subset of $\mathcal{B}(\mathcal{H})$.

Definition

3.1) Given $\mathcal{B} \subset \mathcal{B}(\mathcal{H})$, we define $\mathcal{B}' = \{A \in \mathcal{B}(\mathcal{H}) : [A, B] = 0 \text{ for all } B \in \mathcal{B}\}$ and $\mathcal{B}'' = (\mathcal{B}')'$.

3.2) A von Neumann algebra \mathcal{A} is a subset of $\mathcal{B}(\mathcal{H})$ such that

- i) $\lambda_1 A_1 + \lambda_2 A_2$, $A_1 A_2$, and $A_1^* \in \mathcal{A}$ for all $\lambda_1, \lambda_2 \in \mathbb{C}$, $A_1, A_2 \in \mathcal{A}$,
- ii) $\mathcal{A}'' = \mathcal{A}$.

3.3) Mathematicians define a normal state ρ on a von Neumann algebra \mathcal{A} to be a σ -weakly continuous, positive, linear function on \mathcal{A} such that $\rho(1) = 1$ where 1 is the identity operator. There is an equivalent but less sophisticated definition. First, a density matrix $\omega = \sum_{i=1}^{\infty} p_i |\psi_i\rangle\langle\psi_i|$ on \mathcal{H} is identified with a function on $\mathcal{B}(\mathcal{H})$ by writing $\omega(A) = \sum_{i=1}^{\infty} p_i \langle\psi_i|A|\psi_i\rangle$ for $A \in \mathcal{B}(\mathcal{H})$. Then a normal state ρ on a subsystem corresponding to an algebra \mathcal{A} is defined to be a function on \mathcal{A} which is the restriction of some (not necessarily unique) density matrix ρ' in the sense that $\rho(A) = \rho'(A)$ for all $A \in \mathcal{A}$.

Von Neumann algebras are useful tools in quantum statistical mechanics (Bratteli & Robinson (1979, 1981)), and they are essential for describing local properties in quantum field theory. It is therefore necessary to be able to develop our ideas at the von Neumann algebra level. While this will be done in the sequel, physicists who are not familiar with the notation will not lose any of the fundamental content of the paper by reading throughout “density matrix” for “von Neumann algebra state” and “set of all bounded operators on a Hilbert space defining a subsystem” for “von Neumann algebra”. Attention, however, must be paid to the relationship between subsystems and regions of space-time.

According to the theory of local algebras in quantum field theory, to each region Λ of space-time there is associated a von Neumann algebra $\mathcal{A}(\Lambda)$ consisting of the set

of all observables defining properties within that region. We shall refer to a state on $\mathcal{A}(\Lambda)$ as being a state in the region Λ . The theory of local algebras has been most completely developed in the context of Wightman fields, but there seems little doubt that it can be extended to cover gauge field theories (Seiler (1982)). In my opinion, analogous structures will also exist in the context of the quantization of gravity. Various properties have been postulated for local algebras (Haag and Schroer (1962), Haag and Kastler (1964), Driessler, Summers, and Wichmann (1986), Buchholz and Wichmann (1986)). Two which will be relevant below are:

3.4) If Λ_1 and Λ_2 are spacelike separated then $[A_1, A_2] = 0$ for all $A_1 \in \mathcal{A}(\Lambda_1)$ and $A_2 \in \mathcal{A}(\Lambda_2)$.

3.5) If Λ_1 and Λ_2 are spacelike separated by a strictly positive distance, and ρ_1 and ρ_2 are arbitrary states on $\mathcal{A}(\Lambda_1)$ and $\mathcal{A}(\Lambda_2)$ then, there exists a state ρ on $\mathcal{B}(\mathcal{H})$ with $\rho|_{\mathcal{A}(\Lambda_1)} = \rho_1$ and $\rho|_{\mathcal{A}(\Lambda_2)} = \rho_2$.

(3.5) is an expression of the physical independence of strictly spacelike separated regions. It is related, but not equivalent, to (3.4) (Haag and Kastler (1964), Roos (1970), Ekstein (1969, Appendix C), De Facio and Taylor (1973), Buchholz, D'Antoni, and Fredenhagen (1987)).

$\mathcal{A}(\Lambda)$ is the set of all operators that could conceivably be measured within the region Λ . For many purposes this set might appear to be much too large. However, the $\mathcal{A}(\Lambda)$ are the only fundamental sets of observables supplied by quantum field theory, and, in my view, they are the only fundamental sets of observables on which to base interpretations of quantum theory. Fundamental sets of observables are called for, in particular, in the task, mentioned in section 2, of finding abstract definitions for observers.

At the elementary level, if we have two subsystems S_1 and S_2 then we have, correspondingly, two Hilbert spaces \mathcal{H}_1 and \mathcal{H}_2 , and two sets of observables $\mathcal{B}(\mathcal{H}_1)$ and $\mathcal{B}(\mathcal{H}_2)$. The observables $\mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2)$ of the combined system on $\mathcal{H}_1 \otimes \mathcal{H}_2$ do not consist just of the union of $\mathcal{B}(\mathcal{H}_1) \otimes 1_2$ and $1_1 \otimes \mathcal{B}(\mathcal{H}_2)$. There are also correlations between S_1 and S_2 . Similarly, for two space-time regions Λ_1 and Λ_2 , $\mathcal{A}(\Lambda_1 \cup \Lambda_2)$ is, in general, strictly larger than $\mathcal{A}(\Lambda_1) \cup \mathcal{A}(\Lambda_2)$. In what follows, we shall be considering sets of observables like, for example, $\mathcal{B} = \cup_{n=1}^N \mathcal{A}(\Lambda_n) \cup \mathcal{C}$ for some set $\{\Lambda_n : n = 1 \dots, N\}$ of relevant space-time regions and some subset \mathcal{C} of $\mathcal{A}(\cup_{n=1}^N \Lambda_n)$ expressing a physically relevant selection of the correlations between the Λ_n . In general, \mathcal{B} will not be a von Neumann algebra because it will not be true for every pair $B_1, B_2 \in \mathcal{B}$ that we also have $B_1 B_2 \in \mathcal{B}$. Physically, we shall be dealing with sets of subsystems, rather than with a single all-embracing subsystem. This leads to mathematical complications, so an explanation of why it would be inappropriate to use $\mathcal{A}(\cup_{n=1}^N \Lambda_n)$ is called for.

Definition

3.6) A double cone $C(x, y)$ is a region in space-time, defined by two points x and y such that $x - y$ is timelike and future directed, which takes the form

$$C(x, y) = \{z : x - z \text{ and } z - y \text{ are timelike and future directed}\}.$$

3.7) The causal shadow Λ^{sh} of a region Λ is defined as

$$\Lambda^{sh} = \{x : \text{every future directed timelike path from } x \text{ meets } \Lambda\} \\ \cup \{y : \text{every past directed timelike path from } y \text{ meets } \Lambda\}.$$

In terms of (3.7), it is very natural to assume that our local algebras should possess the property that, for any region Λ ,

$$\mathcal{A}(\Lambda) = \mathcal{A}(\Lambda^{sh}). \quad (3.8)$$

This property, proposed by Haag and Schroer (1962), amounts to the assumption of determinism for quantum field theory, in the sense that the Cauchy problem is well-posed. If (3.8) holds and we know the state of the world in the region Λ , then we also know the state of the world in the causal shadow of Λ .

Haag and Schroer exhibited a certain Wightman field theory (a generalised free field) which did not have property (3.8). This failure, which should lead us to discard such fields as being unphysical, has been examined further by Garber (1975). On the other hand, property (3.8) has been proved for ordinary free fields and for $P(\varphi)_2$ and Y_2 fields (see Glimm and Jaffe (1971, 1972)).

The following property is closely related, but takes a different angle on the casual shadow:

Property 3.9 Suppose that $x, y \in \Lambda$ with $x - y$ timelike and future directed, and suppose that Λ contains a neighbourhood of some timelike path from y to x . Then $\mathcal{A}(C(x, y)) \subset \mathcal{A}(\Lambda)$, where $C(x, y)$ is given by (3.6).

It has been proved by Borchers (1961) and Araki (1963) that this property holds for all Wightman fields. I shall assume that it holds for all quantum field theories.

Property 3.9 is troublesome: Suppose that local algebras do provide the fundamental variables for describing physical systems. Consider a small physical object on the surface of the Earth. Suppose that we study that object for three seconds, and that we learn its state precisely over that entire period. If we think of that state as being a state on the local algebra of the space-time region swept out by the object, then according to property 3.9, we would know the precise state of everything from the Earth to the Moon at some time within that interval. This is essentially because the amount of information and degree of precision being required is absurd. Events on the Moon affect the precise correlations between two events on the Earth separated by a time interval of the appropriate length. Indeed, property 3.9 says that so many correlation variables can be generated that correlations of Earth events can reveal Moon events. This would appear to destroy the potential of localization for providing a mechanism for coarse-graining. The problem is to loosen the force of the requirement that a local state be known precisely over an extended period, while retaining a theory compatible with relativistic quantum field theory and definable in abstract terms. One way of doing this is to consider states of physical objects to be defined on non-algebras. We shall throw away information about all except a specified subset of the correlations between time-like separated events. Simplicity of language will be maintained by introducing the following, to some extent non-standard, definition:

Definition 3.10 A normal state σ on a subset $\mathcal{B} \subset \mathcal{B}(\mathcal{H})$ is the restriction of a density matrix σ' on \mathcal{H} to \mathcal{B} .

As noted above, this definition agrees with the standard definition of a normal von Neumann algebra state when \mathcal{B} is a von Neumann algebra. If $\mathcal{B} \subset \mathcal{A}(\Lambda)$ for some

bounded space-time region Λ , then the set of states defined by 3.10 is exactly the same as the set of restrictions to \mathcal{B} of pure normal states on \mathcal{H} . This is a consequence of the fact that the vacuum may be assumed to be a separating vector for the algebra $\mathcal{A}(\Lambda)$ (Streater and Wightman (1964, theorem 4.3), Driessler, Summers, and Wichmann (1986), Strătilă and Zsidó (1979, §5.24)).

When appropriate, mathematical readers should extend definition 3.10 and consider also restrictions of non-normal states. Non-normal states are states with inferior continuity properties which are related to normal states very much as irrational numbers are to rational numbers. They arise, for example, in mathematically natural limiting processes. The existence of such states is largely irrelevant to the physics of this paper because the formalism will allow us to impose whatever properties we find physically necessary on the sets of local states we want to use. For physicists, it is much more important to bear in mind that “state” here means “density matrix” rather than “wave function”. At the technical level, however, when I need to restrict myself to normal states, I shall do so explicitly. In particular, the supremum in 8.4 and all related suprema are intended to be over all states rather than just over normal states (cf. the remarks following 7.6 and Donald (1986, example 6.6) and (1987a, lemma 4.3)).

Remark 3.11 It is an immediate consequence of 3.10 that every state on a subset \mathcal{B} does have extensions to $\mathcal{B}(\mathcal{H})$. These extensions will usually not be unique. It will often be convenient in the sequel to allow ambiguity in notation by not distinguishing between a set C of states on \mathcal{B} and the set C' of states such that $\sigma'|_{\mathcal{B}} \in C$. (For a function f on $\mathcal{B}(\mathcal{H})$, we write $f|_{\mathcal{B}}$ – the restriction of f to \mathcal{B} – to denote the function on \mathcal{B} defined by $(f|_{\mathcal{B}})(B) = f(B)$ for $B \in \mathcal{B}$.)

Remark 3.12 A zealous logician would now notice that definition 3.10 really refers to the set of states on $\mathcal{B}(\mathcal{H})$ given by $\{\sigma' : \sigma'|_{\mathcal{B}} = \sigma\}$, and would suggest that everything below be rephrased entirely in terms of sets of states on $\mathcal{B}(\mathcal{H})$. This would be possible and we shall need to return to this idea in section 7, but the sets referred to by definition 3.10 are of such central significance that it is useful to adopt a special notation for them. As these sets do define (single-valued) functions on \mathcal{B} , the notation is not inappropriate.

Definition 3.13 As we shall only be interested in the set \mathcal{B} through the states definable on it, it will always be possible to replace \mathcal{B} by the largest set with the same states. This will be denoted by $c(\mathcal{B})$ and is the norm closure of the linear span of $\{1\} \cup \mathcal{B} \cup \{B^* : B \in \mathcal{B}\}$. $c(\mathcal{B})$ is the largest set to which every state on \mathcal{B} has a unique extension in the sense that for every pair of states σ and ρ on $\mathcal{B}(\mathcal{H})$ if $\sigma|_{\mathcal{B}} = \rho|_{\mathcal{B}}$ then $\sigma|_{c(\mathcal{B})} = \rho|_{c(\mathcal{B})}$.

Throughout, this paper focuses on states rather than on operators. Traditionally, measurement theory has focused on operators and the assignment to them of definite values, but I think that this may have been a mistake; resulting in a narrow theory which is viewed as irrelevant by most physicists. The versatility of quantum mechanics is demonstrated in the provision of state descriptions for every physical situation.

The measurement problem arises because such descriptions do not remain valid over extended time periods.

The idea of correlation is central to this paper. This idea is given a mathematical translation by the following lemma and definition, which demonstrate in particular that correlation is state dependent.

Lemma 3.14 *Let ρ be a normal state on a von Neumann algebra \mathcal{A} . Let $P, Q \in \mathcal{A}$ be projections and $R = P \wedge Q$. Suppose that $\rho(P) > 0$, $\rho(Q) > 0$. Then the following are equivalent:*

- (i) $\rho(P) = \rho(Q) = \rho(PQ)$.
- (ii) $P\rho P/\rho(P) = Q\rho Q/\rho(Q)$.
- (iii) $\rho(P - R) + \rho(Q - R) = 0$.

proof Let $s(\rho)$ be the support projection of ρ , which, by definition, is the smallest projection P such that $\rho(P) = 1$. We shall use the fact that $\rho(AA^*) = 0 \iff s(\rho)A = 0$ (Strătilă and Zsidó (1979, §5.15)).

$$\begin{aligned}
(i) &\Rightarrow \rho(QP) = \overline{\rho(PQ)} = \overline{\rho(P)} = \rho(P) \Rightarrow \rho((P - Q)^2) = 0 \Rightarrow s(\rho)(P - Q) = 0 \\
&\Rightarrow P\rho P = Ps(\rho)\rho s(\rho)P = Qs(\rho)\rho s(\rho)Q = Q\rho Q \Rightarrow (ii) \\
&\Rightarrow \rho(PQP) = \rho(P) \Rightarrow 0 = \rho((P - PQ)(P - QP)) \Rightarrow s(\rho)(P - PQ) = 0 \\
&\Rightarrow s(\rho)P = s(\rho)PQ = s(\rho)PQP = s(\rho)PQPQ = \dots = s(\rho)(PQ)^n = s(\rho)R \\
&\hspace{15em} (\text{since } R = s - \lim_{n \rightarrow \infty} (PQ)^n) \\
&\Rightarrow \rho(P - R) = 0. \quad (iii) \text{ follows by exchanging } P \text{ and } Q.
\end{aligned}$$

Suppose (iii). Since $P - R \geq 0$, we have $s(\rho)P = s(\rho)R = s(\rho)Q = s(\rho)PQ \Rightarrow (i)$. ■

Definition 3.15 We shall say that projections P and Q are exactly correlated by the state ρ if the conditions of lemma 3.14 hold. More loosely, we shall say that P and Q are correlated by ρ when the conditions hold to an approximation which is adequate for a particular context.

For an example of the type of situation in which correlations arise in this paper, suppose that we are given two von Neumann algebras \mathcal{A}_1 and \mathcal{A}_2 . Suppose that an observer knows the state on \mathcal{A}_1 and that he is observing the subsystem defined on \mathcal{A}_2 . Suppose that $P_1 \in \mathcal{A}_1$ and $P_2 \in \mathcal{A}_2$ are projections. In order to discover whether P_1 and P_2 are exactly correlated in a given state it will be sufficient (by 3.14(i)) to specify that state on P_1 , P_2 , and P_1P_2 . It may not be sufficient simply to specify the state on $\mathcal{A}_1 \cup \mathcal{A}_2$. To express these pair correlations, it is appropriate to use the set

$$c\{A_1A_2 : A_1 \in \mathcal{A}_1 \text{ and } A_2 \in \mathcal{A}_2\} \quad (3.16)$$

where c is the closure operation defined by (3.13). In general below (for example, (4.2), (4.3)), we shall be considering unions of sets of the form (3.16). These unions could possibly be taken instead over components corresponding to the von Neumann algebra generated by \mathcal{A}_1 and \mathcal{A}_2 , without running into problems with property 3.9. However, I have chosen, as a general principle, always to look for minimal sets of operators on which to define states. If \mathcal{A}_1 and \mathcal{A}_2 are mutually commuting algebras,

then the von Neumann algebra generated by \mathcal{A}_1 and \mathcal{A}_2 is the weak closure of the set given by (3.16).

4. Postulates about observers.

Postulate One Quantum theory is the correct theory for all forms of matter and applies to macroscopic systems as well as to microscopic ones.

Postulate Two For any given observer, any physical system can be described, at any time, by a neighbourhood of quantum states, which is the set of states compatible with his observations of the system at that time. A function can be defined on such a neighbourhood to measure the a priori probability, or likelihood, of any state or set of states, within the neighbourhood given the observer's prior knowledge.

The task in this paper is to find an appropriate definition for “a priori probability” in this postulate. Note that technically a neighbourhood of a state σ is a set containing an open set containing σ . The word is used here simply to indicate that there will always be some variations of any given state which the observer will be unable to perceive. The precise nature of these variations, however, may be quite subtle as will be explained in example 9.8.

Postulate Three An observer O , existing between times t_0 and T , is localized in a succession $\{\Lambda(t) : t \in [t_0, T]\}$ of space-time regions. O is characterized by the neighbourhoods of quantum states which he assigns to the physical systems $\mathcal{A}(\Lambda(t))$ for $t \in [t_0, T]$. There is a distinguished subset $\Gamma(T)$ of $\{A : A \in \mathcal{A}(\Lambda(t)), t \in [t_0, T]\}$ with respect to which O defines correlations.

The simplest version of this postulate would involve a single space-time region; taking $\Lambda(t)$ to be independent of t . However, in view of the problem raised by property 3.9, it is necessary to restrict the temporal extension of the algebras on which we define the observer's states. Thus, for each t , $\Lambda(t)$ will only be extended over time sufficiently to allow the momentary properties of the observer, at time t , to be defined. The time co-ordinate here is simply an observer-dependent parameter. It may be assumed to be an approximate local proper time.

Postulate three assigns to an observer space-time regions which may be thought of as containing his body or his brain, and insists that he is characterized by the states which he assigns to those regions. A generalization of postulate three would allow for several “body parts”. Another generalization would involve neighbourhoods of geometrical structures. Indeed, this is probably necessary in order to make the theory “structurally stable”. It is also necessary to consider variations in the sets $\Gamma(T)$. These generalizations will largely be ignored in this paper since they introduce additional levels of complexity without essentially modifying the central concepts. For example, it is straightforward to allow for both small and large variations in the geometrical structures and in the $\Gamma(T)$ within the general process of finding entities of highest a priori probability. Large variations would be called for in the project of considering as equivalent all entities acting as equivalent information processors of the required nature.

In Donald (1990), I have proposed that observers are composed of elementary body parts which function as quantum switches, and I have given a characterization of the states of a quantum switch on a family of local algebras of the form $\{\mathcal{A}(\Lambda(t)) : t \in [t_0, T]\}$. This proposal allows one specific implementation of postulate three. In this implementation, with the notation of Donald (1990) hypothesis V, $\Gamma(T)$ will be the set of time translates of the switch projections P and Q :

$$\Gamma(T) = \{\tau_{y(t_k)}(P), \tau_{y(t_k)}(Q) : k = 1, \dots, K\}. \quad (4.1)$$

Postulate Four There is a finite sequence of times $t_0 < t_1 < \dots < t_M = T$ and a neighbourhood \mathcal{N}_T of M -component sequences of states such that for any sequence $(\sigma_m)_{m=1}^M \in \mathcal{N}_T$, O views σ_M as a possible present state and the sequence $(\sigma_m)_{m=1}^{M-1}$ as comprising a possible past. In this sequence, σ_m is the state which O sees himself as having occupied in the time interval $[t_{m-1}, t_m]$. Thus $(\sigma_m)_{m=1}^M \in \mathcal{N}_T$ implies $(\rho_m)_{m=1}^M \in \mathcal{N}_T$ for all sequences $(\rho_m)_{m=1}^M$ such that, for each $m \leq M$, $\rho_m(A) = \sigma_m(A)$ for all $A \in \mathcal{A}(\Lambda(t))$, $t \in [t_{m-1}, t_m]$. The states σ_m are all taken to be states, in the sense of 3.10, on

$$\mathcal{B}_M = c\{AC : A \in \mathcal{A}(\Lambda(t)), t \in [t_0, T], C \in \mathcal{C}_M\} \quad (4.2)$$

where \mathcal{C}_M is the von Neumann algebra generated by $\Gamma(T)$ (compare (3.16)).

The finiteness of the sequence in this postulate is necessary if we are to maintain the idea that Hamiltonian evolution is fundamental with quantum measurement providing merely occasional interruptions to reset that evolution. This finiteness also allows us to evade the “quantum Zeno paradox” to which we shall return in section 9. The change in state from σ_{m-1} to σ_m is a “collapse”. Collapse is required to occur because the observer, by his nature, can only occupy certain types of state. This means that \mathcal{N}_T has to be defined in such a way as to forbid superpositions, or mixtures, of O observing different events (distinguishable by O himself), in order to reflect the fact that O only ever sees himself as observing a single reality. A choice of the neighbourhood \mathcal{N}_T corresponds to a choice of a single possible reality history seen by O . In many worlds language, “different worlds” would correspond to different neighbourhoods – each different sequence within a single neighbourhood corresponds to a different possible structure through which O may observe one single “world”.

Postulates three and four introduce an arrow of time and propose that by time t_m , O is gaining information about the set \mathcal{B}_m . The fact that, at time T , O considers all of the members of \mathcal{N}_T as being sequences of states on \mathcal{B}_M is possible in view of 3.11 and will be required both so that “collapse” may be between states defined on a common domain, and, more importantly, as a reflection of the fact that \mathcal{B}_M is the set of operators of which O is then aware. The use of extensions of states from \mathcal{B}_m to \mathcal{B}_M is unrestricted, because, in calculating probabilities, we shall always take suprema over all extensions.

Postulate Five There is an initial quantum state ω describing the universe prior to the existence of any observer.

ω is far removed, by a long sequence of collapses, from the state of the universe that we see. Underlying this work is a picture of observers as information processors

of a particular type of structure. That structure defines the possible neighbourhoods \mathcal{N}_T of postulate [four](#). The aim here is to define an a priori probability for each such neighbourhood given a particular initial state ω . It must be required that this can be done in such a way that the only observers which exist with relatively high a priori probability are those which are processing information giving them an accurate picture of reality and that we ourselves are such observers. This implies that ω should be such that the most likely sequence of information-carrying collapses would result in states which model the flesh and blood brains of apparently evolved life.

Of course, there is an unconventional viewpoint involved in postulating that collapse is observer-dependent. Thus, in the framework of postulate [four](#), the collapses implicit in the usual idea of the past must be seen as being made indirectly by each separate observer as he learns about the apparent consequences of such a past. For example, consider an observer discovering a fossil. The conventional view would be that the fossil must essentially have been fixed from not long after the moment of death. The suggestion here is that the observer in looking at a rock surface is now collapsing out one single image from all the possible images that that surface might reflect. Only a priori probabilities for those images were fixed just before the observer studied the surface, and even they depended on the observer.

This is related to Everett's answer to Einstein's difficulty with the idea of a mouse bringing about drastic changes in the universe simply by looking at it ([Everett \(1957\)](#) p.116) – if it is the mouse and not the universe which changes then the universe has a state which always remains uncollapsed. That state is ω . As far as I can see, despite the change of viewpoint, the speculations of cosmologists are still appropriate in identifying ω . In particular, on a space-time region around what we see as the solar system over a recent period, ω will approximate a 2.7°K thermal equilibrium state.

Postulate Six The a priori probability of O existing at time T in a specific present state σ_M on \mathcal{B}_M with a specific past sequence $(\sigma_m)_{m=1}^{M-1}$ of states on \mathcal{B}_M is a function $\text{app}_{\mathcal{B}_M}((\sigma_m)_{m=1}^M | \omega)$ of \mathcal{B}_M , of $(\sigma_m)_{m=1}^M$, and of $\omega|_{\mathcal{B}_M}$.

Postulate Seven $\text{app}_{\mathcal{B}_M}((\sigma_m)_{m=1}^M | \omega) = \prod_{m=1}^M \text{app}_{\mathcal{B}_M}(\sigma_m | \sigma_{m-1})$ where we define $\sigma_0 = \omega|_{\mathcal{B}_M}$, and where $\text{app}_{\mathcal{B}}(\sigma | \rho)$ is a function of states σ and ρ on \mathcal{B} .

This postulate expresses the intuition that, for the observer O at time T , aware of the set \mathcal{B}_M , the world starts in the state $\omega|_{\mathcal{B}_M}$, and then undergoes a finite succession of “collapses” through a sequence of states $\sigma_1, \sigma_2, \dots, \sigma_M$ on \mathcal{B}_M . Postulate [seven](#) requires that the collapse from σ_{m-1} to σ_m depends only on those two states. It introduces the function $\text{app}_{\mathcal{B}}(\sigma | \rho)$ of just two states which is to be interpreted as measuring the a priori probability of the collapse of the state ρ on \mathcal{B} to the state σ .

Perhaps the most striking difference, at the conceptual level, between postulates [four](#) to [seven](#) and the conventional framework in physics, is that O is described, at a given moment, not just by the state which he has reached at that moment but by the family of states which he has passed through. The conventional framework requires that past information is used to specify a unique global state at the present, but this is not appropriate in the present framework because we are working locally and with

neighbourhoods of states. As we shall see, even within such neighbourhoods, there may not be any states of maximal a priori probability, and, even when there are, they need not be unique. This underscores the proposal of postulate [three](#) that it is the entire set \mathcal{N}_T which characterizes O .

As an aside, I note that one of the fundamental long-term goals of my work, expressed in part in [Donald \(1990\)](#) and in this paper, is to identify the (objective) physical substrate of consciousness. Part of the idea behind postulate [four](#) is that, even at a given moment, that physical substrate is not just the instantaneous state of a brain but instead involves the history of that brain. For example, we do not understand the meaning of a word like “blue” because it can make us remember clear skies – that would merely be a functional definition of consciousness – but because the meaning for us is, in part, our previous experience of those skies. More precisely, the neural structures brought into play by hearing the word “blue” are closely linked to structures formed on sunny days. The meaning of those structures lies in the past. Consciousness reads, for itself, the present neural excitement through being all its previous patterns of excitement.

The [next](#) postulate demonstrates how the probabilities of individual sequences of states are to be combined in order to give a total a priori probability for \mathcal{N}_T . This postulate will be given in a form which is more general than is strictly necessary. It defines, in terms of \mathcal{N}_T and the subsidiary function $\text{app}_{\mathcal{B}}(\sigma | \rho)$, an a priori probability for O to observe a given physical system occupying a given set of states. At the fundamental level, it would be sufficient to define the a priori probability of O observing \mathcal{N}_T , indeed, ultimately, we only need an appropriate relationship between the [sixth](#) and the [seventh](#) notions of probability given in section 2. The more general definition introduced by postulate [eight](#) will provide a model making that relationship plausible and will be useful in motivating the definition of $\text{app}_{\mathcal{B}}(\sigma | \rho)$. While the work of subsequent sections will confirm the broad consistency of the general version of postulate [eight](#) with other methods of calculation in quantum mechanics, at the deepest level, when applied to arbitrary sets of states, it is only a tool, like [notions 3, 4, and 5](#) of section 2. There are many circumstances in which the definition provides a useful tool, but there may be other circumstances in which it is not useful. In view of these remarks, the theory of multiple observations given here will be restricted to the succession of observations by the observer of his own structure. This succession will extend over his entire life to date.

Two versions of postulate [eight](#) will be given. The [preliminary](#) version expresses the fundamental physical idea of a succession of collapses to states of maximal a priori probability given the state already reached. The [second](#) version allows for the fact that, while we can always find states of close to maximal a priori probability, there may be no states for which the maxima are actually attained. In this version, we look for (infinite) sequences of elements of \mathcal{N}_T , (these elements are themselves, of course, finite sequences), which come arbitrarily close to maximal a priori probability at each successive moment of collapse.

Maximizing a priori probability will be a central technique used throughout this paper. This might seem strange, since it is obvious that it is not only the most likely

events which happen. The idea is that the physical structure of a given observer observing a given event can be manifested by many different but equivalent forms – in particular, by any one of the elements of a single set \mathcal{N}_T . The a priori probability of that observer is then measured by a series of steps each maximizing a priori probability over equivalent and, as far as the observer is concerned, indistinguishable structures. No restriction is placed on the different events which the observer may come to find himself observing, which will correspond to different \mathcal{N}_T , beyond the restrictions imposed by his own nature. These are to be expressed through the definition of the \mathcal{N}_T . (It is because its statement does not take full account of these restrictions that postulate [eight](#) is “over-general”.)

Postulate Eight (preliminary version) Suppose that, at time T , O observes a physical system defined on a set of observables \mathcal{B}_S .

$$\text{Define } \mathcal{B}_{SM} = c(\mathcal{B}_M \cup \{BC : B \in \mathcal{B}_S, C \in \mathcal{C}_M\}). \quad (4.3)$$

For $m = 1, \dots, M$, define sets of initial sequences

$$\mathcal{N}^m = \{(\sigma_i)_{i=1}^m : \exists (\sigma_i)_{i=m+1}^M \text{ with } (\sigma_i)_{i=1}^M \in \mathcal{N}_T\},$$

where, in view of [3.11](#), these are considered to be sets of sequences of states on \mathcal{B}_{SM} . Define, by induction on m , the following maximal a priori probabilities and corresponding sets. Start with

$$\begin{aligned} \text{app}^0(\mathcal{N}_T, \mathcal{B}_{SM}, 1, \omega) &= \sup\{\text{app}_{\mathcal{B}_{SM}}(\sigma | \omega) : \sigma \in \mathcal{N}^1\} \quad \text{and} \\ \tilde{\mathcal{N}}_0^1(\mathcal{B}_{SM}) &= \{\sigma \in \mathcal{N}^1 : \text{app}_{\mathcal{B}_{SM}}(\sigma | \omega) = \text{app}^0(\mathcal{N}_T, \mathcal{B}_{SM}, 1, \omega)\}. \end{aligned}$$

Then, for $1 < m + 1 \leq M$, set $\text{app}^0(\mathcal{N}_T, \mathcal{B}_{SM}, m + 1, \omega) = 0$ and $\tilde{\mathcal{N}}_0^{m+1}(\mathcal{B}_{SM}) = \emptyset$ if $\tilde{\mathcal{N}}_0^m(\mathcal{B}_{SM}) = \emptyset$ and if not set

$$\begin{aligned} \text{app}^0(\mathcal{N}_T, \mathcal{B}_{SM}, m + 1, \omega) &= \sup\{\text{app}_{\mathcal{B}_{SM}}((\sigma_i)_{i=1}^{m+1} | \omega) : (\sigma_i)_{i=1}^{m+1} \in \mathcal{N}^{m+1} \\ &\quad \text{and } (\sigma_i)_{i=1}^m \in \tilde{\mathcal{N}}_0^m(\mathcal{B}_{SM})\} \quad \text{and} \\ \tilde{\mathcal{N}}_0^{m+1}(\mathcal{B}_{SM}) &= \{(\sigma_i)_{i=1}^{m+1} \in \mathcal{N}^{m+1} : (\sigma_i)_{i=1}^m \in \tilde{\mathcal{N}}_0^m(\mathcal{B}_{SM}) \\ &\quad \text{and } \text{app}_{\mathcal{B}_{SM}}((\sigma_i)_{i=1}^{m+1} | \omega) = \text{app}^0(\mathcal{N}_T, \mathcal{B}_{SM}, m + 1, \omega)\}. \end{aligned} \quad (4.4)$$

(This says that $\tilde{\mathcal{N}}_0^{m+1}$ is the set of partial sequences with maximal a priori probability given an initial sequence in $\tilde{\mathcal{N}}_0^m$.)

Then, the a priori probability of O , at time T , observing the subsystem to occupy a set of states C is defined by

$$\begin{aligned} \text{app}^0(O, T, \mathcal{B}_S, C | \omega) &= 0 \quad \text{if } \tilde{\mathcal{N}}_0^M(\mathcal{B}_{SM}) = \emptyset \quad \text{and otherwise by} \\ \text{app}^0(O, T, \mathcal{B}_S, C | \omega) &= \sup\{\text{app}_{\mathcal{B}_{SM}}((\sigma_m)_{m=1}^{M+1} | \omega) : (\sigma_m)_{m=1}^M \in \tilde{\mathcal{N}}_0^M(\mathcal{B}_{SM}), \text{ and } \sigma_{M+1} \in C\}. \end{aligned} \quad (4.5)$$

Postulate Eight (final version) Suppose that, at time T , O observes a physical system defined on a set of observables \mathcal{B}_S .

Define \mathcal{B}_{SM} and \mathcal{N}^m as above.

Define, by induction on m , the following a priori probabilities. Start once again with

$$\text{app}(\mathcal{N}_T, \mathcal{B}_{SM}, 1, \omega) = \sup\{\text{app}_{\mathcal{B}_{SM}}(\sigma | \omega) : \sigma \in \mathcal{N}^1\}.$$

Then, for $1 < m + 1 \leq M$, set

$$\begin{aligned} \text{app}(\mathcal{N}_T, \mathcal{B}_{SM}, m + 1, \omega) &= \sup\{\limsup_{n \rightarrow \infty} \text{app}_{\mathcal{B}_{SM}}((\sigma_i^n)_{i=1}^{m+1} | \omega) \\ &\quad : ((\sigma_i^n)_{i=1}^{m+1})_{n \geq 1} \text{ is a sequence of elements of } \mathcal{N}^{m+1} \\ &\quad \text{and, for } 1 \leq k \leq m, \text{app}_{\mathcal{B}_{SM}}((\sigma_i^n)_{i=1}^k | \omega) \rightarrow \text{app}(\mathcal{N}_T, \mathcal{B}_{SM}, k, \omega)\}. \end{aligned} \quad (4.6)$$

(This says that the partial sequences $(\sigma_i^n)_{i=1}^k$ approach the successively maximal a priori probabilities as $n \rightarrow \infty$.)

Then, the a priori probability of O , at time T , observing the subsystem to occupy a set of states C is defined by

$$\begin{aligned} \text{app}(O, T, \mathcal{B}_S, C | \omega) &= \sup\{\limsup_{n \rightarrow \infty} \text{app}_{\mathcal{B}_{SM}}((\sigma_m^n)_{m=1}^{M+1} | \omega) : \\ &\quad ((\sigma_m^n)_{m=1}^M)_{n \geq 1} \subset \mathcal{N}^M, (\sigma_{M+1}^n)_{n \geq 1} \subset C, \text{ and, for } 1 \leq k \leq M, \\ &\quad \text{app}_{\mathcal{B}_{SM}}((\sigma_i^n)_{i=1}^k | \omega) \rightarrow \text{app}(\mathcal{N}_T, \mathcal{B}_{SM}, k, \omega)\}. \end{aligned} \quad (4.7)$$

The mathematics of this postulate is explored in section 9. The inductive definition expressed by (4.6) imposes a natural causal framework on O 's present description of his past as a developing sequence of collapses. In particular, note that $\text{app}(\mathcal{N}_T, \mathcal{B}_{SM}, m, \omega)$ does not depend on the specific set of states C which O will observe on \mathcal{B}_{SM} . There is no particular sequence of states which we can point to as the sequence which O actually occupies or observes. However, for $\delta > 0$, it is useful to define

$$\begin{aligned} \tilde{\mathcal{N}}_\delta^m(\mathcal{B}_{SM}) &= \{(\sigma_i)_{i=1}^m \in \mathcal{N}^m : \text{app}_{\mathcal{B}_{SM}}((\sigma_i)_{i=1}^k | \omega) \geq \text{app}(\mathcal{N}_T, \mathcal{B}_{SM}, k, \omega) - \delta, \\ &\quad \text{for } k = 1 \dots, m\} \text{ and} \\ \tilde{\mathcal{N}}_\delta^{M+1}(\mathcal{B}_{SM}, C) &= \{(\sigma_m)_{m=1}^{M+1} : (\sigma_m)_{m=1}^M \in \tilde{\mathcal{N}}_\delta^M(\mathcal{B}_{SM}), \sigma_{M+1} \in C, \text{ and} \\ &\quad \text{app}_{\mathcal{B}_{SM}}((\sigma_m)_{m=1}^{M+1} | \omega) \geq \text{app}(O, T, \mathcal{B}_S, C | \omega) - \delta\}. \end{aligned} \quad (4.8)$$

For sufficiently small δ , elements of $\tilde{\mathcal{N}}_\delta^m(\mathcal{B}_{SM})$ will successively have close to maximal a priori probability (lemma 9.10). However, these sets do depend on δ , on m , on M , and on S . The validity of the approach taken to quantum mechanics in this paper depends on the claim that, for δ sufficiently small and for suitable systems S , elements of $\tilde{\mathcal{N}}_\delta^m(\mathcal{B}_{SM})$ will be sequences of states close to those which would be assigned, for the appropriate time intervals, to O or to S in conventional quantum theory. This claim is fundamental and we shall return to it in section 6. It implies, for example,

that the various dependencies of $\tilde{\mathcal{N}}_\delta^m(\mathcal{B}_{SM})$ are not problematical. It will only be justified in as far as it is demonstrated that it leads to a consistent and plausible theory.

Considerable philosophical heresy is inherent in postulate [eight](#). It defines a theory of the world as it is observed which, although postulating the existence of a “real world” outside the observer, also claims that how that world is likely to be seen depends essentially on the structure of the individual observer. The “real” “fixed” “external” world enters through the state ω , defined by postulate [five](#), through the Hamiltonian time propagation implicit in our continual reliance on the Heisenberg picture, and through the use of an observer-independent a priori probability. The observer dependence enters through the sets \mathcal{N}_T and \mathcal{B}_M . A full definition of these sets will also make objective the class of physical structures through which an observer may be embodied.

Even when all these definitions are given, we shall not have a full elucidation of the meaning of the defined term “a priori probability”. Philosophers have written voluminously about the meaning of conventional (measure-theoretic) probability theory and without reaching a consensus (see [Cohen \(1989, chapter II\)](#) for a review). The probability introduced by postulate [eight](#) is not even conventional because, for example, it does not define a measure on the space of sets of states. In particular, the a priori probability of a union of sets of states is not, in general, the sum of the separate a priori probabilities (see property [G](#) of section [7](#) and lemma [8.9](#)). Indeed, as anticipated in postulate [two](#), if the set C consists of a single state, [\(4.7\)](#) will be meaningful and may well not vanish. It is for this reason that it is necessary to introduce a postulate showing how to combine probabilities of individual sequences of states, rather than expecting to be able to define general a priori probabilities by sums of terms of the form $\text{app}_{\mathcal{B}_M}((\sigma_m)_{m=1}^M | \omega)$. Thus, philosophical questions both old and new are raised by this paper. Perhaps some of the old questions are made more interesting by providing an alternative framework in which they can be discussed. In this paper, these questions will not be addressed directly. The goal is a method of calculating objective probabilities for individual observations. The meaning of “a priori probability” will be left at the intuitive (and circular) level of “a number measuring (relative) likelihood”. The next two sections provide examples of circumstances in which this intuition is satisfactory.

The numbers defined by [\(4.7\)](#) are trivially Lorentz invariant, by property [D](#) of section [7](#), if the underlying quantum field theory also is. This is an important step in establishing the compatibility of the proposed theory with special relativity. Indeed, given that collapse is observer dependent, it only remains to note that the relation of the observer to space-time geometry is supposed, through postulates [three](#) and [four](#), to be explicit, so that the causal framework in which collapse appears to occur can be made specific. One way in which relativity might lead to difficulties in conceivable variants of the theory proposed here stems from the fact that the number given by [\(4.7\)](#) will depend on the ordering of the underlying collapses. This might be a problem, particularly with the generalization of postulate [three](#) to allow for several body parts, if some of those collapses refer to mutually simultaneous changes in structure (for

example, “switchings” in the sense of Donald (1990)). For a single observer, we can deal with this dependence, either by extending the supremum in (4.7) over all reorderings of such mutually simultaneous changes, or by assuming a fixed ordering as part of the observer’s fundamental structure. However, this ordering problem does seem to me to rule out any attempt to extend the theory to include all observations made by all observers within a single structure analogous to \mathcal{N}_T .

Analogous remarks might be made about the conventional interpretation of quantum mechanics. The Lorentz invariance of probability amplitudes is an automatic consequence of a unitary representation of Lorentz transformations but observer-independent global collapse gives rise to an intractable ordering problem. In the conventional interpretation, global collapse also requires global specification of the eigenfunctions of each measured operator – or equivalently global specification of the operator itself. This once again emphasizes how difficult the specification problem is for the conventional interpretation.

5. The a priori probability function and measurement theory.

(4.7) introduces a function measuring the a priori probability of observation of a given set of states by an observer with a given life history up to a particular time. Through postulate seven, that function is defined in terms of the subsidiary function $\text{app}_{\mathcal{B}}(\sigma | \rho)$. In this section, we link the values of this latter function to conventional measurement theory, first, by requiring that $\text{app}_{\mathcal{B}}(\sigma | \rho)$ should generalize the notion that the weight of a component in a density matrix is the a priori probability of that component, and, second, by examining the link between that notion and measurement theory. The requirement of the first stage is expressed by the following property, which we shall be able to impose on $\text{app}_{\mathcal{B}}(\sigma | \rho)$:

5.1) Suppose that $\rho = p_a \rho_a + (1 - p_a) \rho_d$ for $0 \leq p_a \leq 1$ and ρ, ρ_a , and ρ_d states on some set \mathcal{B} . Suppose that there exists a projection $Q_a \in \mathcal{B}$ such that $\rho_a(Q_a) \geq 1 - \varepsilon$ and $\rho_d(Q_a) \leq \varepsilon$ for some $\varepsilon \in [0, \frac{1}{2}]$. Then

$$p_a \leq \text{app}_{\mathcal{B}}(\rho_a | \rho) \leq p_a - 3\varepsilon \log \varepsilon.$$

This implies that $\text{app}_{\mathcal{B}}(\rho_a | \rho)$ is close to p_a when ε is sufficiently small and ρ_a and ρ_d are, in an obvious sense, close to disjoint.

For the second stage, we introduce:

5.2) **A General Model of an Observation.** Suppose that an observer O observes the outcome of a measurement on a macroscopic system defined by a set of observables \mathcal{B}_S . Let C be the set of states on \mathcal{B}_S compatible with the observations of O just before he learns the outcome of the measurement and let C_a be the set of states on \mathcal{B}_S compatible with his observations when he has discovered that the outcome is a . Suppose that according to conventional quantum mechanics, outcome a has probability p_a and corresponds to a state σ_a on \mathcal{B}_S . Then

A) there is a projection Q_a in \mathcal{B}_S such that $\sigma_a(Q_a) = 1$, and, for some small $\varepsilon \geq 0$, all the most likely states in C will belong to $\{\rho : p_a - \varepsilon \leq \rho(Q_a) \leq p_a + \varepsilon\}$ and all the most likely states in C_a will belong to $\{\rho : \rho(Q_a) \geq 1 - \varepsilon\}$.

B) Among the most likely states in C , there are states ρ taking the form

$$\rho = p_a \rho_a + (1 - p_a) \rho_d \tag{5.3}$$

where $\rho_a \in C_a$, $\rho(Q_a) \geq 1 - \varepsilon$, and ρ_d is some state on \mathcal{B}_S with $\rho_d(Q_a) \leq \varepsilon$.

Some of the terms in this model are undefined. The words “macroscopic system” allow for the imposition of suitable conditions on the set \mathcal{B}_S . The scenario to be invoked is of a human experimenter studying \mathcal{B}_S directly with his unaided senses. “Small” will not be defined, but, as will be explained, it is plausible that, in the present context, it can be taken to refer only to unmeasurably small effects. “Most likely” will be interpreted intuitively in this section. The appeal to “most likely states” is part of a consistency argument for the a priori probability being introduced. Postulate [eight](#) permits us to restrict attention to such states. As we are working with a many worlds theory, in which there is no unique state which is the “true” observed state of the world at a given moment, it is necessary to allow for highly unlikely or bizarre states in which, for example, entropy decreases on a macroscopic scale during the experiment. There will be such states in C_a and there will be other highly unlikely states in which the correlations invoked below do not hold.

In as far as it is a state decomposition with negligible interference effects, (5.3) is like (2.1), and so is very much part of the normal focus of measurement theory. For example, [Daneri, Loinger, and Prosperi \(1962\)](#), [Hepp \(1972\)](#), [Whitten-Wolfe and Emch \(1976\)](#), [Machida and Namiki \(1980\)](#), [Araki \(1980, 1986\)](#), and [Zurek \(1981, 1982\)](#), provide a variety of models and arguments all giving plausible evidence that inference effects can be neglected in practice because of the macroscopic nature of measuring devices. These reference mainly rely, or can be read as relying, on coarse-graining and, in particular, on the fact that we cannot have perfect and complete knowledge of the states of macroscopic systems. The aim in this paper is to take a step beyond these arguments and to introduce a formalism within which the limited nature of the knowledge of an observer can be expressed. This will be a formalism appropriate for a universe in which, as far as any observer is concerned, inference effects are indeed negligible, in the appropriate circumstances, but over which, nevertheless, a global, unitary, time propagation still holds sway.

The central task in this section and the [next](#) is to make explicit plausible minimal assumptions under which (5.3) will be valid in the framework of section 4. As we shall see, a theory based on the function app is a theory in which information only becomes definite in circumstances where interference effects can be neglected. The coarse-graining implicit in (5.3) comes from restricting \mathcal{B}_S to be only part of the total local algebra affected by the measurement. This section gives an elementary explanation of how such coarse-graining works as a preliminary to the [next](#) section in which the scenario will be less conventional and much harder to handle explicitly. In particular, in this section we shall consider states on \mathcal{B}_S but not the extensions to \mathcal{B}_{SM} required by postulate [eight](#).

If we put $\varepsilon = 0$ then [5.2A](#) would say that the most likely states in C_a are eigenvectors of Q_a . As it is claimed that ε is negligible, it might be suggested that there is not much difference between the present model and conventional quantum

mechanics – we are considering eigenvectors of a less than completely defined operator. However, the framework is quite different. Any macroscopically observable situation allows for the definition, in many different possible ways, of projections Q_a . Being able to deal in the same way with any of the enormous variety of different projections is a mark of consistency for the present theory, rather than an ambiguity problem, as it is for conventional quantum theory. The reason for this is that in the conventional theory, the fundamental entity is supposed to be exactly one of these projections, but we have not been told which one, while in the present theory, the fundamental entities are to be defined by the structure of an observer, and the measurement model merely allows that observer to make a variety of mutually consistent predictions. Ultimately, also, there is a tremendous difference at the theoretical level between an ε which vanishes and an ε which is unobservably small. Only in the latter case, do we really avoid having to specify definitive observables for each measuring apparatus.

The closeness to conventional quantum mechanics otherwise is, of course, no bad thing, because of the success of that theory. It is a reflection of that success that the model of this section can be claimed to be “general” even although it does not use the full generality of postulate [eight](#), according to which C_a might be any set of states. Nevertheless, the existence of that further generality is important; refuting any suggestion that [\(5.2\)](#) is more than a “general model”. In particular, the further generality provides consistency between postulates [eight](#) and [four](#) and allows for situations like [\(2.3\)](#) to which the present model will apply only indirectly.

We begin the justification of the model, with our assumptions about \mathcal{B}_S . We assume that it is restricted in temporal and spatial extension, in the sense that $\mathcal{B}_S \subset \mathcal{A}(\Lambda_S)$ for some restricted space-time region Λ_S . Λ_S could, for example, be a region containing a photograph, which would constitute the result of the experiment for the observer, at around the moment when he first begins to glance at that photograph. We shall further assume that the result of the experiment could be observed in a different restricted region $\Lambda_{S'}$ which is space-like separated from Λ_S . In the example, $\Lambda_{S'}$ might be a region simultaneous with Λ_S containing light which had reflected from the photograph just prior to Λ_S , or it might be a region containing the photographic negative.

The existence of projections Q_a such that Q_a projects onto the wave functions modelling the outcome a is an essential part of the content of traditional measurement theory. For example, in a bubble chamber experiment, Q_a could be a projection onto an orthonormal basis of wave-functions modelling the scattering particle heading in a given direction, or it could be a projection onto wave-functions of local regions of gas (bubbles) in the chamber along the given path, or it could involve wave-functions of the photographic plate, or of the light moving away from a developed photograph. Q_a could even involve suitably chosen wave-functions in the brain of the observer. These projections are not unambiguously defined, but that is merely a reflection of the fact that the “given result” cannot be precisely defined; except, indeed, as a class of correlated projections. The variation in expected value between these projections in likely states is one of the unmeasurably small effects referred to earlier.

It will be possible to choose Q_a to be in \mathcal{B}_S by the assumption that \mathcal{B}_S is large enough to define a physical system. In this case, the statement that Q_a projects onto the wave functions modelling the outcome a implies that the most likely states in C_a belong to $\{\rho : \rho(Q_a) \geq 1 - \varepsilon\}$ for some small ε , because a state in C_a will not belong to that set only if, in that state, Q_a is not correlated with the projections with which O interacts most directly when making his observations. Such states are highly unlikely. There are two quite different aspects to the claim that, for ρ a state of high a priori probability in C , we must have $\rho(Q_a)$ close to p_a . On the one hand, we are requiring that O have sufficient information about the experimental set-up, whether or not he can translate that information into numbers, to pre-determine the probabilities of the various possible outcomes. On the other hand, we are assuming that conventional calculations correctly provide expectations of appropriate projections in appropriate states. This is a consistent assumption as the relevant calculations only use quantum dynamics prior to collapse.

It is possible to find two of these result-determining projections $Q_a \in \mathcal{B}_S$ and $Q_{a'} \in \mathcal{B}_{S'}$ where $\mathcal{B}_{S'}$ is the commutator of \mathcal{B}_S (see (3.1)). This follows from the assumption about the existence of two space-like separated regions Λ_S and $\Lambda_{S'}$ with $\mathcal{B}_S \subset \mathcal{A}(\Lambda_S)$ in each of which the result of the experiment could be determined. We can choose $Q_{a'} \in \mathcal{A}(\Lambda_{S'})$ and, from (3.4), $\mathcal{A}(\Lambda_{S'}) \subset \mathcal{A}(\Lambda_S)' \subset \mathcal{B}_{S'}$. Among the most likely states in C there will be some which can be extended from \mathcal{B}_S to $\mathcal{A}(\Lambda_S \cup \Lambda_{S'})$ in such a way that they are good models on the whole of that set of the physical situation as O sees it. This is because the same physical causes which O observes giving rise to the state on \mathcal{B}_S act with just the same probability on the whole region $\Lambda_S \cup \Lambda_{S'}$. In such a state ρ , the commuting projections Q_a and $Q_{a'}$ will be correlated, in the sense of 3.15 because we have three projections Q_a , $Q_{a'}$, and $Q_a Q_{a'}$ each of which serves to pick out the observed result with the same pre-determined probability, so that (to order ε) $\rho(Q_a) = \rho(Q_{a'}) = \rho(Q_a Q_{a'}) = p_a$. Then, for $B \in \mathcal{B}_S$,

$$\begin{aligned} \rho(B) &= \rho(Q_{a'} B) + \rho((1 - Q_{a'}) B) \\ &= \rho(Q_{a'} B Q_{a'}) + \rho((1 - Q_{a'}) B (1 - Q_{a'})) \quad (\text{since } Q_{a'} \in \mathcal{B}_{S'}) \\ &= \rho(Q_{a'}) \rho(Q_{a'} B Q_{a'}) / \rho(Q_{a'}) + \rho(1 - Q_{a'}) \rho((1 - Q_{a'}) B (1 - Q_{a'})) / \rho(1 - Q_{a'}) \\ &= p_a \rho_a(B) + (1 - p_a) \rho_d(B) \end{aligned}$$

where ρ_a and ρ_d are defined by

$$\rho_a(B) = \rho(Q_{a'} B Q_{a'}) / \rho(Q_{a'}) \quad \text{and} \quad \rho_d(B) = \rho((1 - Q_{a'}) B (1 - Q_{a'})) / \rho(1 - Q_{a'}).$$

But $\rho(Q_a) = \rho(Q_{a'}) = \rho(Q_a Q_{a'}) = p_a$ implies that $\rho_a(Q_a) = 1$ and that $\rho_d(Q_a) = 0$ (all to order ε). ρ_a is constructed from wave-functions modelling outcome a of the experiment so that $\rho_a \in C_a$.

There is a shortcut to the justification of (5.3). This involves the idea of entropy as a measure of the number of available states. For example, the entropy S of a glass of water can be written as $k \log N$ where k is Boltzmann's constant, and the experimentally determined value for S gives N as roughly $10^{3 \times 10^{25}}$. Identifying S with $k \operatorname{tr}(-\sigma \log \sigma)$ for some density matrix σ , N is a lower bound on the number of orthogonal pure states into which σ decomposes (i.e. writing $\operatorname{tr}(-\sigma \log \sigma) = -\sum_{i=1}^M p_i \log p_i$ in the usual way, we have $\operatorname{tr}(-\sigma \log \sigma) \leq \log M$). σ can be decomposed into two

disjoint states in at least $2^{N-1} - 1$ ways, since this is the number of ways that the set $\{1, \dots, N\}$ can be split into two disjoint non-empty subsets.

Because the measured entropies of macroscopic systems are so large, when interpreted in the way described, it is not implausible to assert that relation (5.3) automatically holds for any real physical measurement. For conventional quantum statistical mechanics, the set \mathcal{B}_S would be taken to be a fixed-time local algebra for the system in question. The argument is that any state of a macroscopic system at normal temperature can be decomposed with an astonishing fineness. According to statistical mechanical theory, the decomposition is into energy eigenstates of the system. Those energy eigenstates however are enormously degenerate, so that even the pure states of the finest decomposition are far from unique. To assert (5.3) is then to claim that one can rebuild some decomposition on the left-hand side into the decomposition required on the right. Thus (5.3) will be a simple consequence of the fact that the state splits in so many different ways and the fact that the required splitting is compatible with everything which has been observed about the system. What makes this argument most plausible is the fact that ρ in equation 5.3 is only defined up to the precision of human observation. This level of precision is one on which the laws of classical thermodynamics appear to be obeyed absolutely regardless of finite volume or time effects. Classical thermodynamics tells us that inference effects are negligible, even if it does not explain why.

The well-known work of [Daneri, Loinger, and Prosperi \(1962\)](#), which uses time-averaging, can be interpreted as attempting an explanation by suggesting that a decomposition like (5.3) happens with high ergodic-theory probability. In this paper, localization, a physically much better-founded method of coarse-graining, is used, and instead of appealing to ergodic theory, we can invoke a direct relationship between entropy, or more precisely negative free energy, and the a priori probability function to be defined (see the discussion following property H of section 7). This will make thermodynamically plausible states more likely than other states, and, in particular, other things being equal, states of higher entropy will tend to have higher a priori probability. The first argument for (5.3) is more satisfactory because of its directness. The argument based on entropy is weaker, but it does at least suggest that the smallness of Boltzmann's constant in practical units may be quite as important in explaining the fact that the macroscopic world appears to behave in a classical fashion as the smallness of Planck's constant. It also provides an upper bound on the number of macroscopically distinguishable results for any given experiment and a quantitative measure of macroscopicness.

6. A priori probability and observations.

In section 2, seven different [notions](#) of quantum probability were mentioned. In this section, we consider the link between the a priori probability defined by postulate [eight](#) and the statistical probability which an observer assigns to the outcome of his next experiment following a long sequence of prior trials. Empirical justification for a theory based on the postulates of section 4 will be provided, by arguing that all the empirical evidence normally used to justify conventional quantum mechanics can be transferred to the framework of these postulates.

The empirical evidence for conventional quantum theory lies in the observed agreement between measured statistical probabilities and theoretically defined probabilities. In order to transfer this evidence, it is only necessary to demonstrate that the new theory defines similar probabilities in similar circumstances. Once this has been done, the choice between the theories can only be made on such grounds as completeness, internal consistency, and aesthetic appeal.

The circumstances in which conventional quantum theory defines measurable probabilities can essentially all be subsumed under the general measurement model proposed in section 5. This, of course, is apparently to ignore the distinction between discrete and continuous observables. However, this is acceptable, since we are modelling only circumstances in which information has become observable at a macroscopic level. At such a level, not only does thermal noise inevitably wash out the finest gradations of an observable property, but also, those gradations which are observable, are observable precisely because there is a difference in the macroscopic properties to which they are correlated. The measurement model can be extended to the context of section 4 by the following postulate:

Postulate Nine Suppose that O observes the outcome of an experiment on a macroscopic system defined by a set of observables \mathcal{B}_S . Suppose that model 5.2 applies, and adopt the notation of that model. Recall definition (4.8). Then, for every sufficiently small positive δ ,

A) every sequence $(\sigma_m)_{m=1}^M \in \tilde{\mathcal{N}}_\delta^M(\mathcal{B}_{SM})$ will satisfy $p_a - \varepsilon \leq \sigma_M(Q_a) \leq p_a + \varepsilon$, and every sequence $(\sigma_m)_{m=1}^{M+1} \in \tilde{\mathcal{N}}_\delta^{M+1}(\mathcal{B}_{SM}, C_a)$ will satisfy $\sigma_{M+1}(Q_a) \geq 1 - \varepsilon$.

B) There is a sequence $(\sigma_m^\delta)_{m=1}^M \in \tilde{\mathcal{N}}_\delta^M(\mathcal{B}_{SM})$ which is such that σ_M^δ takes the form

$$\sigma_M^\delta = p_a \rho_a^\delta + (1 - p_a) \rho_d^\delta \quad (6.1)$$

where $\rho_a^\delta|_{\mathcal{B}_S} \in C_a$, $\rho_a^\delta(Q_a) \geq 1 - \varepsilon$, and ρ_d^δ is some state on \mathcal{B}_{SM} with $\rho_d^\delta(Q_a) \leq \varepsilon$.

Once again, we have the phrase “macroscopic system”, in [this](#) postulate, referring to the need to impose the same conditions on \mathcal{B}_S here as were needed for model 5.2. ε will again be left unspecified. [A](#) gives an explicit interpretation of the phrase “most likely states” used in [5.2A](#). As will be shown in proposition 9.11, it is a consequence of postulate [nine](#) and the definition of `app` to be introduced below that the conditional a priori probability of O observing a is close to p_a in the sense that

$$\text{app}(O, T, \mathcal{B}_S, C_a | \omega) / \text{app}(O, T, \mathcal{B}_S, \Sigma | \omega) \sim p_a \quad (6.2)$$

where Σ is the set of all states on \mathcal{B}_S . (This means that the denominator is the a priori probability of the vacuous experiment in which O observes any result on \mathcal{B}_S .) It follows that, if b is a second possible result of the experiment, with conventionally calculated probability p_b , then the ratio of the a priori probability of O observing a to the a priori probability of O observing b will be close to the conventionally calculated ratio p_a/p_b in the sense that

$$\text{app}(O, T, \mathcal{B}_S, C_a | \omega) / \text{app}(O, T, \mathcal{B}_S, C_b | \omega) \sim p_a/p_b. \quad (6.3)$$

It is also demonstrated in proposition 9.11 that, if we set $\sigma_{M+1}^\delta = \rho_a^\delta$, then, for δ sufficiently small, $\text{app}_{\mathcal{B}_{SM}}((\sigma_m^\delta)_{m=1}^{M+1} | \omega)$ is close to $\text{app}(O, T, \mathcal{B}_S, C_a | \omega)$ and that

$$\text{app}_{\mathcal{B}_{SM}}(\rho_a^\delta | \sigma_M^\delta) \sim \text{app}_{\mathcal{B}_{SM}}(\rho_a^\delta |_{\mathcal{B}_S} | \sigma_M^\delta |_{\mathcal{B}_S}) \sim p_a, \quad (6.4)$$

so that, in this particular circumstance, $\text{app}_{\mathcal{B}_{SM}}(\rho_a^\delta |_{\mathcal{B}_S} | \sigma_M^\delta |_{\mathcal{B}_S})$ is indeed an appropriate “collapse” probability.

(6.2) and (6.3) are empirically justified because they make the same predictions as conventional quantum mechanics. They can therefore be used to annex the empirical evidence which is normally taken to support conventional quantum mechanics. What is more, they seem to me to be sufficient to annex all of that empirical support. My disagreement with conventional quantum mechanics is based on that theory’s inexplicable incompleteness rather than on its predictions.

Unlike postulate eight, postulate nine is not a definition. Indeed, at the deepest level, it is merely a model – in other words, something which is a useful general picture rather than a precise and complete description of any particular experiment. Such a complete description would require a much more detailed analysis of the observer and of observation. Postulate nine makes a hypothesis about the nature of the states most likely to be experienced by O in the given circumstances. In judging the validity of this hypothesis, there are various aspects to be considered. The fundamental claim, already mentioned in section 4, is that we can find a mathematical definition for a priori probability which is appropriate in the sense of yielding likelihoods for quantum states which are consonant with intuition. In other words, a definition which is such that the most likely states modelling a given situation are states which conventional quantum theorists would allow to be assigned to that situation. This will then imply, for example, that, for δ sufficiently small, σ_M^δ can be taken to be a standard quantum state modelling an observer, or his brain, on \mathcal{B}_M and an experimental apparatus on \mathcal{B}_S , just before the observer becomes aware of the result of the experiment. In particular, this is compatible with the analysis in Donald (1990), where the brain was considered from the viewpoint of conventional neurophysiology and conventional statistical physics. It will also imply that σ_M^δ does provide the correlations which we would expect between \mathcal{B}_M and \mathcal{B}_S , and that this holds for any of the variety of possible sets \mathcal{B}_S , compatible with model 5.2. The claim implies part A of postulate nine by underpinning the argument given in section 5 for 5.2A. Another implication, which could fail for pathological choices of \mathcal{B}_S – in particular, were we to take $\mathcal{B}_S = \mathcal{B}(\mathcal{H})$, is that $(\sigma_m^\delta)_{m=1}^M$ has a priori probability close to $\sup\{\text{app}_{\mathcal{B}_M}((\sigma_m)_{m=1}^M | \omega) : (\sigma_m)_{m=1}^M \in \tilde{\mathcal{N}}_\delta^M(\mathcal{B}_M)\}$. This implication reflects the expectation that all the constraints on $(\sigma_m^\delta)_{m=1}^M$ are imposed through \mathcal{B}_M . The considerable attention devoted below to motivating the definition of app is with the purpose of making this fundamental claim plausible.

According to (6.1), certain states of highest a priori probability have a property of being mixed. (6.1) would be justified, if the splitting of ρ in (5.3) could be extended from \mathcal{B}_S to \mathcal{B}_{SM} . By (5.3) and the fundamental claim just made, we can assume that $\sigma_M^\delta |_{\mathcal{B}_S}$ takes the form $p_a \hat{\rho}_a + (1-p_a) \hat{\rho}_d$ for suitable $\hat{\rho}_a$ and $\hat{\rho}_d$. We can also assume that we know the state $\sigma_M^\delta |_{\mathcal{B}_M}$. Even with these assumptions, however, a considerable step is required to reach (6.1). This step depends on a degree of physical independence

between the observer and the observed system. That there is a real problem here may not seem obvious, but it must be remembered that, in general, particularly at the macroscopic level, individual quantum states only provide “realistic” descriptions over short time intervals. All our intuition is based on such intervals because the world always appears as if state collapse was occurring with high frequency. Imagine, for example, smoke rising from a fire and then ask how many collapses, in the generalized sense used throughout this paper, are required to produce a particular observed pattern. When, as now, we are required to consider quantum states over extended time intervals without permitting collapse, we need to be extremely careful about existence. It need not be true that there exists any pair of states ρ_a^δ and ρ_d^δ such that

$$\rho_a^\delta|_{\mathcal{B}_S} = \hat{\rho}_a, \quad \rho_d^\delta|_{\mathcal{B}_S} = \hat{\rho}_d, \quad \text{and} \quad \rho_a^\delta|_{\mathcal{B}_M} = \rho_d^\delta|_{\mathcal{B}_M} = \sigma_M^\delta|_{\mathcal{B}_M}. \quad (6.5)$$

This is because \mathcal{B}_M and \mathcal{B}_S do not commute and are not independent in the sense of (3.5), so that, given a state ρ_1 on \mathcal{B}_M and a state ρ_2 on \mathcal{B}_S , there need not exist any state ρ on \mathcal{B}_{SM} such that $\rho|_{\mathcal{B}_M} = \rho_1$ and $\rho|_{\mathcal{B}_S} = \rho_2$.

Nevertheless, if ρ_1 equals $\sigma_M^\delta|_{\mathcal{B}_M}$; which, by assumption, is a typical state for the body of the observer to have reached just before he becomes aware of the result of the experiment, and if ρ_2 is any state modelling a given result of the experiment, then there does exist such an extension ρ . This extension is constructed simply by imagining circumstances, however unlikely, in which the state modelled by ρ_2 on \mathcal{B}_S is produced deterministically without the knowledge of the observer. In a bubble chamber experiment, for example, it is possible to imagine a duplicitous graduate student slipping an old bubble chamber photograph, or even an artificially constructed one, into the pile marked “output”. The first step in this argument is the assumption that there is an extension $\tilde{\rho}^\delta$ of $\sigma_M^\delta|_{\mathcal{B}_M}$ to a system \mathcal{B}_{MG} such that $\tilde{\rho}^\delta|_{\mathcal{B}_{MG}}$ models not only O but also the student. Even this assumption is not necessarily true, because of the temporal extension of \mathcal{B}_M (see (4.2)), but from both mathematical and physical points of view I certainly find it extremely plausible. The second step is to note that, by construction, ρ_2 is the extension of $\tilde{\rho}^\delta$ to \mathcal{B}_S .

It is consistent with section 5 to assume that $\hat{\rho}_d$ can be constructed as a mixture of states modelling possible results of the observation other than a , so this thought experiment yields states ρ_a^δ and ρ_d^δ which do satisfy (6.5). The mixture $\rho^\delta = p_a\rho_a^\delta + (1 - p_a)\rho_d^\delta$ on \mathcal{B}_{SM} is compatible with all the information available to O just before he observes the outcome of his experiment and is a physically plausible state in the sense that in each separate spatial region it has typical thermodynamic properties. Of course, the fantasy required to discover this state is absurd and there is no suggestion that the graduate student is anything other than part of a mathematical construction. We have no interest in the extensions of ρ^δ away from \mathcal{B}_{SM} to the operators with which the state of the graduate student would be described, were he to exist. Given that states satisfying (6.5) do exist, (6.1) reduces to the hypothesis that they are the most likely such states. This hypothesis can now be taken to be part of our fundamental claim, backed up by (5.3) and the mixing-enhancing property of app which will be given as property H in the next section.

There is a problem with postulate [nine](#) which ought at least to be mentioned. This is that insufficient allowance has perhaps been made for approximation. (6.1) is claimed to hold for a component of a sequence of elements which approach arbitrarily close to a supremum of a priori probabilities. The mathematical structure is such, particularly as app will turn out not to be continuous, that, for example, small changes in \mathcal{B}_M or \mathcal{B}_S could conceivably cause large changes in $\tilde{\mathcal{N}}_\delta^m(\mathcal{B}_{SM})$. I believe, however, that this is a mere mathematical pathology and that structural stability can be given to the theory by allowing variations in the precise definition of the system S and in the geometric structures implicit in the definition of \mathcal{B}_M .

7. Properties of a priori probability.

In this section, properties appropriate to an a priori probability which can play the role required, will be listed and justified. In the [following](#) section, it will be shown that certain of these properties are sufficient to give a unique definition which does indeed possess all the properties listed. In my opinion, the totality of properties satisfied by this definition does make it both suitable and hard to modify. Of course, this section has been written with hindsight in that it does proceed towards a pre-established definition: given by property [B](#) and (8.1 – 8.4). Thus its true purpose is to introduce and justify that definition.

Property A $\text{app}_{\mathcal{B}}((\sigma_m)_{m=1}^M | \rho)$ is a function of a set $\mathcal{B} \subset \mathcal{B}(\mathcal{H})$ and of $(\sigma_m)_{m=1}^M$ and ρ – restrictions of states to \mathcal{B} .

This property is consistent in statement with postulate [six](#), but to be fully consistent in notation we should earlier have written $\text{app}_{\mathcal{B}_M}((\sigma_m)_{m=1}^M | \omega|_{\mathcal{B}_M})$ rather than $\text{app}_{\mathcal{B}_M}((\sigma_m)_{m=1}^M | \omega)$.

Although fundamental to the definition of app, property [A](#) is probably the most difficult property to motivate. In essence, it is a postulate about how the global state ω influences a localized observer. Mathematically, it may be noted that, since measuring a priori probability involves the comparison of states, it is desirable to define those states on a single set. There are two obvious alternatives. One of these, requiring that app depends on a given, fixed, globally-defined extension of each σ_m , contradicts the idea of locality. This alternative takes us straight back to the Copenhagen interpretation and to all its problems. The other alternative permits the set on which σ_m is defined to increase with m . As will be shown in section [9](#), under this alternative, ω would not necessarily influence σ_m at all. More generally, the a priori probability of $(\sigma_m)_{m=1}^M$ could equal one without σ_M being equal to ω wherever defined.

Property B $\text{app}_{\mathcal{B}}((\sigma_m)_{m=1}^M | \rho) = \prod_{m=1}^M \text{app}_{\mathcal{B}}(\sigma_m | \sigma_{m-1})$ where we define $\sigma_0 = \rho$.

This property is postulate [seven](#). It allows us to confine attention in the rest of this section and in the [next](#) to the function $\text{app}_{\mathcal{B}}(\sigma | \rho)$ of just two states.

Property C

- (i) $0 \leq \text{app}_{\mathcal{B}}(\sigma | \rho) \leq 1$ for all σ and ρ .
- (ii) $\text{app}_{\mathcal{B}}(\sigma | \rho) = 1$ if and only if $\sigma = \rho$.

(i) is necessary if app is to be a probability in any sense. However, in general, app will not be a probability in the more technical sense of defining a measure. (ii) requires that a priori probability is always lost by a non-trivial collapse.

Property D Let $U \in \mathcal{B}(\mathcal{H})$ be unitary, and define $\tau : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$ by $\tau(B) = UBU^*$. Then $\text{app}_{\mathcal{B}}(\sigma | \rho) = \text{app}_{\tau^{-1}(\mathcal{B})}(\sigma \circ \tau | \rho \circ \tau)$.

This is an obviously necessary isomorphism invariance.

Property E Suppose that $\rho = p_a \rho_a + (1 - p_a) \rho_d$ for $0 \leq p_a \leq 1$ and suppose that there exists a projection $Q_a \in \mathcal{B}$ such that $\rho_a(Q_a) \geq 1 - \varepsilon$, $\rho_d(Q_a) \leq \varepsilon$ for some $\varepsilon \geq 0$. Then, for ε sufficiently small, $\text{app}_{\mathcal{B}}(\rho_a | \rho) \sim p_a$.

This is the essence of (5.1)

Properties C, D, and E are not sufficient, by themselves, to yield a unique definition for $\text{app}_{\mathcal{B}}(\sigma | \rho)$. In order to proceed further, it is necessary to discover a general interpretation for the function being sought. I have little doubt but that the only way to achieve this is to work backwards; first finding a suitable definition and then inventing an interpretation for it. If the interpretation can subsequently be shown to lead uniquely to the definition that is important justification. I have already (Donald (1986)) published one version of this process. In the remainder of this section, I aim to revise and deepen the interpretative analysis sketched in that paper. This will be done, in part, by arguing for a modified, although equivalent, set of defining properties; in part, by motivating a broad set of mutually compatible properties, rather than seeking a minimal set of axioms; and, in part, by the relevance of the rest of this paper. Readers who are prepared to take mathematical proofs on trust will be able to read this paper independently of Donald(1986). The process of giving meaning to $\text{app}_{\mathcal{B}}(\sigma | \rho)$ will largely be the process of interpreting the notion of a quantum state.

Generalizing (6.4) suggests that $\text{app}_{\mathcal{B}}(\sigma | \rho)$ should be interpreted as the a priori probability that an observer, observing a subsystem defined by a set of observables \mathcal{B} , sees the state σ on \mathcal{B} as the outcome of an experiment when ρ was the state on \mathcal{B} just prior to that outcome being ascertained. I have not proposed this generalization as a postulate. This is because I am unable to specify circumstances more general than those given in postulate nine under which (6.4) can be justified. Nevertheless, the general interpretation is compatible with all the ideas of sections 4, 5, and 6.

An alternative way of looking at this generalization of (6.4) emphasizes that, from the point of view expressed in this paper, “collapse” is not a physical process external to, and independent of, the observer. Instead, “collapse” is required to occur because the observer can only occupy certain types of state (section 4). In other words, collapse is, one might say, a mistake which the observer makes about the state of the world because he is physically incapable of seeing its true state. He cannot experience a cat as being both alive and dead, so he experiences it as being either one or the other. $\text{app}_{\mathcal{B}}(\sigma | \rho)$ is to be interpreted as a measure of his falsity of vision in the following sense:

7.1) $\text{app}_{\mathcal{B}}(\sigma | \rho)$ is the probability, per unit trial of the information in \mathcal{B} , of being able to mistake the state of the world on \mathcal{B} for σ , despite the fact that it is actually ρ .

In Donald (1986), I used this interpretation to motivate four axioms which provide the complete definition of $\text{app}_{\mathcal{B}}(\sigma | \rho)$. The first of these axioms derived app on an arbitrary set \mathcal{B} from app on $\mathcal{B}(\mathcal{H})$. This axiom appears below as (7.5) in consequence of property G. The two axioms in Donald (1986) which defined $\text{app}_{\mathcal{B}(\mathcal{H})}$ for finite dimensional \mathcal{H} are replaced here by properties C, D, E, and I, which are both more general and more primitive. Finally, there has to be a technical property allowing the extension to infinite dimensional spaces. In this paper this is property K, which is once again more general than axiom IV of Donald (1986).

With the background of sections 4 and 6, it is possible to be conceptually more sophisticated in this paper. Recall from remark 3.12 that a state σ on \mathcal{B} may be interpreted as the set of states on $\mathcal{B}(\mathcal{H})$ given by $\{\sigma' : \sigma'|_{\mathcal{B}} = \sigma\}$. The set \mathcal{B} is the maximal set of operators which the observer can use to distinguish between states. This idea of indistinguishability may be generalized. The neighbourhood \mathcal{N}_T introduced in postulate four may be viewed as a set of sequences of states which O cannot distinguish among. In these terms, let R and S be arbitrary sets of states on $\mathcal{B}(\mathcal{H})$. Consider an observer who collapses from some state in R to some state in S but who cannot distinguish between the states in R or between the states in S . The a priori probability of this collapse will be a function $\text{app}(S | R)$ which may be interpreted as follows:

7.2) For R and S sets of states on $\mathcal{B}(\mathcal{H})$, $\text{app}(S | R)$ is the probability per unit trial by an observer who cannot distinguish between different states in R or between different states in S , of being able to mistake the world for some state in S despite the fact that it is actually some state in R .

In these terms, (7.1) should be restated as:

7.3) $\text{app}_{\mathcal{B}}(\sigma | \rho)$ is the probability, per unit trial of the information in \mathcal{B} , of being able to mistake the state of the world for a state compatible on \mathcal{B} with σ , despite the fact that it is actually some state compatible on \mathcal{B} with ρ .

Property F If $R \subset R'$ and $S \subset S'$ then $\text{app}(S | R) \leq \text{app}(S' | R')$.

This simply says that an observer who can draw finer distinctions is less likely to make mistakes.

Property G

$$\text{app}(S_1 \cup S_2 | R_1 \cup R_2) = \max\{\text{app}(S_1 | R_1), \text{app}(S_1 | R_2), \text{app}(S_2 | R_1), \text{app}(S_2 | R_2)\}.$$

Property F implies that the left hand side is at least as large as the right hand side. Equality should obtain because no more information is given to the observer collapsing from $R_1 \cup R_2$ to $S_1 \cup S_2$ than that the collapse starts either in R_1 or R_2 and finishes either in S_1 or S_2 . Notice that this is only an argument; 7.2 is an interpretation rather than a definition and property G is an axiom rather than a theorem. A simple extension of property G implies, for any pair R, S , that

$$\text{app}(S | R) = \sup\{\text{app}_{\mathcal{B}(\mathcal{H})}(\sigma | \rho) : \sigma \in S, \rho \in R\}. \quad (7.4)$$

In terms of (3.10)–(3.12), (7.4) implies that we should define

$$\text{app}_{\mathcal{B}}(\sigma | \rho) = \sup\{\text{app}_{\mathcal{B}(\mathcal{H})}(\sigma' | \rho') : \sigma'|_{\mathcal{B}} = \sigma, \rho'|_{\mathcal{B}} = \rho\}. \quad (7.5)$$

This allows $\text{app}_{\mathcal{B}}$ to be derived from $\text{app}_{\mathcal{B}(\mathcal{H})}$. As another consequence, also a special case of property F, we have the following “monotonicity” result:

7.6) Let $\mathcal{B}_1 \subset \mathcal{B}_2$ and σ_2, ρ_2 be extensions to \mathcal{B}_2 of states σ_1, ρ_1 on \mathcal{B}_1 .

Then $\text{app}_{\mathcal{B}_1}(\sigma_1 | \rho_1) \geq \text{app}_{\mathcal{B}_2}(\sigma_2 | \rho_2)$.

It must be emphasized that no physical relevance is to be attached to states σ', ρ' on $\mathcal{B}(\mathcal{H})$ at which the supremum in (7.5) is attained. Such states are only guesses based on limited information and (7.3) is only a way of giving meaning to a function depending on such limited information. In particular, (7.3) is not to be thought of as referring to some physical σ -independent global state ρ' about which the observer is making mistakes, because even when such a state exists, the observer has no access to it except through its restriction to \mathcal{B} . This applies specifically to the state ω given by postulate five. The observer only interacts with $\omega|_{\mathcal{B}_M}$ and the set of extensions of that (partial) state is used in (7.5) only as a mathematical equivalent to $\omega|_{\mathcal{B}_M}$. For different σ , we may well need to use different extensions of ρ in calculating the supremum in (7.5). This mathematics is compatible with a many-worlds theory in which separate possible worlds at a given instant are totally uncommunicating modes of experiencing the universe. In the present theory, different worlds may occupy overlapping quantum states. Everett required that different worlds use orthogonal wave-functions, implying a picture of physically different worlds in different physical dimensions. This would correspond to chopping up a single state given globally by ω . I prefer to think of limited information about a single universe made sense of in many ways.

The simplest case to which 7.2, or indeed 7.1, can be applied arises when $\mathcal{B} = \mathcal{Z}$ – a finite dimensional Abelian algebra. In this case, it is possible to interpret the notion of a state on \mathcal{Z} in terms of conventional probability theory or information theory. This yields a complete definition (equation (7.7)) for $\text{app}_{\mathcal{Z}}(\sigma | \rho)$. We review this analysis with the object of extending its elements to the non-Abelian situation:

\mathcal{Z} is generated by a finite sequence $(P_m)_{m=1}^M$ of orthogonal projections with $\sum_{m=1}^M P_m = 1$. A state ρ on \mathcal{Z} corresponds to a probability distribution $(\rho(P_m))_{m=1}^M$ on $\{1, 2, \dots, M\}$ – ($0 \leq \rho(P_m) \leq 1$ and $\sum_{m=1}^M \rho(P_m) = 1$). A trial of the state ρ can be taken to be an observation of a random variable X with values in $\{1, 2, \dots, M\}$ which is distributed according to $\text{Prob}\{X = m\} = \rho(P_m)$. If we perform N trials on ρ , then, when N is large, we shall be justified in mistaking ρ for σ if the frequency distribution of outcomes of those trials corresponds to the distribution defined by σ (i.e. if $X = m$ in $N\sigma(P_m)$ of the trials). The outcome of a trial on ρ is distributed according to the multinomial distribution, so the probability of the given frequency distribution arising can be explicitly calculated. Setting $s_m = \sigma(P_m)$, $r_m = \rho(P_m)$, and making the assumption that each $s_m N$ is an integer, it equals $N! \prod_{m=1}^M r_m^{s_m N} / (s_m N)!$. By Stirling’s formula, the logarithm of this probability is asymptotic to $N\{\sum_{m=1}^M (-\sigma(P_m) \log \sigma(P_m) / \rho(P_m))\}$ (Sanov (1957), Bratteli and

Robinson (1981, pp 425-427)). This indicates that it would be appropriate for app to satisfy

$$\text{app}_{\mathcal{Z}}(\sigma | \rho) = \exp\left\{\sum_{m=1}^M (-\sigma(P_m) \log \sigma(P_m)/\rho(P_m))\right\}. \quad (7.7)$$

This analysis is particularly straightforward because it has been possible to treat states on \mathcal{Z} as single entities rather than as sets. As will be discussed below, it turns out that this is possible for states on any injective von Neumann algebra. On non-algebras, however, states must, in general, be taken to be sets:

Example 7.8 Let $(\psi_i)_{i=1}^4$ be an orthonormal basis for \mathbb{C}^4 and

$$P = |\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2|, \quad Q = |\psi_1\rangle\langle\psi_1| + |\psi_3\rangle\langle\psi_3|.$$

Let $\mathcal{B} = c\{P, Q\}$. A complete mathematical analysis of $\text{app}_{\mathcal{B}}(\sigma | \rho)$ can be given (see Donald (1987a, §5)). Conceptually, we are still working in terms of conventional probability theory because P and Q commute and because, as a matter of fact, we need only consider extensions of states on \mathcal{B} to the Abelian algebra generated by $\{|\psi_i\rangle\langle\psi_i| : i = 1, \dots, 4\}$. The set of states on that algebra compatible with a given state ρ on \mathcal{B} is equivalent to the set of probability distributions on $\{1, 2, 3, 4\}$ given by

$$\{(r_i)_{i=1}^4 : 0 \leq r_i \leq 1, \sum_{i=1}^4 r_i = 1, r_1 + r_2 = \rho(P), r_1 + r_3 = \rho(Q)\}.$$

In these terms, (7.3) does not pose a well-defined classical problem with a unique solution but an inference problem with incomplete information. Specifically, the problem is to estimate a maximal value, as $N \rightarrow \infty$, for

$$\left(\text{Prob}\left(\left|\frac{1}{N}\left(\sum_{n=1}^N X_n^P\right) - \sigma(P)\right| < \frac{1}{2N}, \left|\frac{1}{N}\left(\sum_{n=1}^N X_n^Q\right) - \sigma(Q)\right| < \frac{1}{2N}\right)\right)^{1/N}$$

where $(X_n^P)_{n=1}^N$ (resp. $(X_n^Q)_{n=1}^N$) are independent identically distributed random variables on $\{1, 2, 3, 4\}$ such that $X_n^P(\omega) = 1$ if and only if $\omega \in \{1, 2\}$ (resp. $X_n^Q(\omega) = 1$ iff $\omega \in \{1, 3\}$) and all that is known about the distribution is that it is in the set above – which is equivalent to saying that $E(X_n^P) = \rho(P)$ and $E(X_n^Q) = \rho(Q)$.

Whether it is reasonable to assume properties F and G (or their Abelian analogues) for the solution to such an inference problem may depend on the context in which the problem arises, but, in the present context, it is my opinion that they are appropriate. ■

The analysis of $\text{app}_{\mathcal{B}}$ on sets other than Abelian algebras will focus on concavity properties of a priori probability. By definition 3.10, a convex combination $\sigma = x_1\sigma_1 + x_2\sigma_2$ of two states σ_1 and σ_2 on any set \mathcal{B} is also a state. Here x_1 and x_2 satisfy $0 \leq x_1, x_2 \leq 1$ and $x_1 + x_2 = 1$. Generalizing the standard interpretation, already invoked in section 2, of a density matrix as a mixture of components weighted by probabilities, it is appropriate to assume that an observer given the information in \mathcal{B} will have observed σ if he has observed σ_1 a fraction x_1 of the time and σ_2 a fraction x_2 of the time.

Suppose then that N_1 and N_2 are integers with $x_1 = N_1/(N_1 + N_2)$ and $x_2 = N_2/(N_1 + N_2)$. According to 7.3, $\text{app}_{\mathcal{B}}(\sigma_1 | \rho)^{N_1} \text{app}_{\mathcal{B}}(\sigma_2 | \rho)^{N_2}$ is the probability that an observer given the information in \mathcal{B} will be able to mistake the state of the world for a state compatible with σ_1 in an initial sequence of N_1 trials and for σ_2 in a subsequent sequence of N_2 trials despite the fact that the state is actually some state compatible with ρ . Following the total sequence of $N_1 + N_2$ trials, however, he will be prepared to believe that the state of the world is compatible with $x_1\sigma_1 + x_2\sigma_2$. This justifies the assumption of the following property:

Property H Let σ_1 , σ_2 , and ρ be states on a set \mathcal{B} . Let $x_1, x_2 \in [0, 1]$ with $x_1 + x_2 = 1$. Let $\sigma = x_1\sigma_1 + x_2\sigma_2$. Then

$$\text{app}_{\mathcal{B}}(\sigma | \rho) \geq \text{app}_{\mathcal{B}}(\sigma_1 | \rho)^{x_1} \text{app}_{\mathcal{B}}(\sigma_2 | \rho)^{x_2}.$$

Property H is a mixing-enhancing or entropy increasing property because it implies that the a priori probability of the mixture $x_1\sigma_1 + x_2\sigma_2$ is at least as great as the minimum a priori probability of the component states σ_1 and σ_2 . In fact, the link between app and entropy is much more direct and an alternative interpretation of app in thermodynamic terms is possible (see (Donald (1987b))). For this interpretation, we again consider a pair (S, R) of sets of states on $\mathcal{B}(\mathcal{H})$, or, indeed, on any other algebra on which thermodynamic systems can be defined. We now interpret each state in R as a possible equilibrium state at an arbitrary temperature T – any state will be such for some system. Then $-kT \log \text{app}(S | R)$ is the minimum free energy which would be sufficient to move one of these systems from its equilibrium state $\rho \in R$ to a state $\sigma \in S$.

Thus far, the idea of a “trial” has been left at an entirely intuitive level. In conventional probability theory, this idea can be given an operational definition, but this is not appropriate in the more general context. A trial of a quantum state is to be thought of not as a physical experiment on that state, but rather as an imagined way in which that state may be experienced. It is postulate nine that gives the link between app and experiments. The problem is that, in quantum theory, an experiment, or indeed any kind of acquisition of knowledge, involves the process of “collapse” and the resulting change of state. The experiencing of a state as being the state which one is occupying is different from investigating that state. The image of a trial as a mode of experiencing is introduced as a fundamental image of a quantum state. Ultimately, it is for the reader to consider whether that image is compatible with all the other interpretative structures used in this paper.

On \mathcal{Z} , a single trial gave a choice from the set $\{1, 2, \dots, M\}$ and a series of trials built up a picture of the state. Now we shall work by analogy on a finite-dimensional Hilbert space \mathcal{H} and we shall take $\mathcal{B} = \mathcal{B}(\mathcal{H})$. Suppose then that we are given an orthogonal basis $(\psi_m)_{m=1}^M$ for \mathcal{H} and let $S = \{|\psi_m\rangle\langle\psi_m| : m = 1, \dots, M\}$ be the corresponding set of pure states on $\mathcal{B}(\mathcal{H})$. Then imagine that one of the ways of experiencing a state ρ on $\mathcal{B}(\mathcal{H})$ is by an “ S -trial” which is to give a choice from S . This notion will be shown to be useful by deriving various mutually consistent properties for app from plausible assumptions about what such trials would mean if the notion were useful. Once again, we are not proving theorems here about formally

defined concepts, but trying to construct a consistent circle of ideas. It is because this is all that we are doing that the restriction to finite dimensions is permissible at this stage. This allows us to put various technicalities to one side for the moment. By 7.3, the a priori probability of choosing σ_m in an S -trial should be $\text{app}_{\mathcal{B}(\mathcal{H})}(\sigma_m | \rho)$. A series of S -trials should allow us, if their results have an appropriate distribution, to mistake the state of the world for some state in the linear span of S . These are the only states that can be precisely identified by such a series of trials. This is reflected in the fact (lemma 8.9) that $\sum_{m=1}^M \text{app}_{\mathcal{B}(\mathcal{H})}(\sigma_m | \rho) \leq 1$ with equality only if ρ is in the span of S . $1 - \sum_{m=1}^M \text{app}_{\mathcal{B}(\mathcal{H})}(\sigma_m | \rho)$ may be interpreted as the a priori probability of an S -trial providing the experience that ρ is not in that span.

For example, suppose that the state σ on $\mathcal{B}(\mathcal{H})$ is in the span of S , having the form $\sigma = \sum_{m=1}^M s_m \sigma_m$ where $0 \leq s_m \leq 1$ and $\sum_{m=1}^M s_m = 1$. If, in accordance with our interpretation of (2.1), we interpret σ as being a mixture of the states σ_m with probabilities s_m then we would mistake ρ for σ in a series of N S -trials, with N large, if the result σ_m occurred $s_m N$ times. Thus $(\text{app}_{\mathcal{B}(\mathcal{H})}(\sigma | \rho))^N$ should be asymptotic to the probability that N trials of a random variable X with values in $\{1, \dots, M\} \cup \{\infty\}$ and distribution $\text{Prob}\{X = m\} = \text{app}_{\mathcal{B}(\mathcal{H})}(\sigma_m | \rho)$, $\text{Prob}\{X = \infty\} = 1 - \sum_{m=1}^M \text{app}_{\mathcal{B}(\mathcal{H})}(\sigma_m | \rho)$ has the result $X = m$ in $s_m N$ of the trials. By the same calculation that gave (7.7), this yields

$$\text{app}_{\mathcal{B}(\mathcal{H})}(\sigma | \rho) = \prod_{m=1}^M (\text{app}_{\mathcal{B}(\mathcal{H})}(\sigma_m | \rho) / s_m)^{s_m}. \quad (7.9)$$

Property I Let σ_1 , σ_2 , and ρ be states on a finite-dimensional algebra \mathcal{A} . Let $x_1, x_2 \in [0, 1]$ with $x_1 + x_2 = 1$. Let $\sigma = x_1 \sigma_1 + x_2 \sigma_2$.

Then $\text{app}_{\mathcal{A}}(\sigma_1 | \rho)^{x_1} \text{app}_{\mathcal{A}}(\sigma_2 | \rho)^{x_2} / \text{app}_{\mathcal{A}}(\sigma | \rho)$ is independent of ρ .

For simplicity, we shall only consider the cases of $\mathcal{A} = \mathcal{B}(\mathcal{H})$, for \mathcal{H} a finite-dimensional Hilbert space, and of \mathcal{A} an Abelian algebra. Readers familiar with the structure theory for finite-dimensional algebras (Takesaki (1979, §1.11)) should have no difficulty in generalizing to the wider context. For $\mathcal{A} = \mathcal{B}(\mathcal{H})$, suppose that, for $i = 1, 2$, $\sigma_i = \sum_{m=1}^M s_m^i \sigma_m^i$ for some sets $S^i = \{\sigma_m^i : m = 1, \dots, M\}$ of disjoint pure states. A test for σ may then be made by performing N_1 S^1 -trials followed by N_2 S^2 -trials where N_1 and N_2 are large and $N_1 / (N_1 + N_2) = x_1$. $\text{app}_{\mathcal{B}(\mathcal{H})}(\sigma | \rho)^{N_1 + N_2}$ is to be interpreted as the a priori probability of ρ being mistaken for σ in this test. One set of test results justifying this mistake arises when all the S^1 -trials have results compatible with σ_1 and all the S^2 -trials have results compatible with σ_2 . This set should have a priori probability $\text{app}_{\mathcal{B}(\mathcal{H})}(\sigma_1 | \rho)^{N_1} \text{app}_{\mathcal{B}(\mathcal{H})}(\sigma_2 | \rho)^{N_2}$. The ratio

$$\begin{aligned} & \text{app}_{\mathcal{B}(\mathcal{H})}(\sigma_1 | \rho)^{N_1} \text{app}_{\mathcal{B}(\mathcal{H})}(\sigma_2 | \rho)^{N_2} / \text{app}_{\mathcal{B}(\mathcal{H})}(\sigma | \rho)^{N_1 + N_2} \\ & = (\text{app}_{\mathcal{B}(\mathcal{H})}(\sigma_1 | \rho)^{x_1} \text{app}_{\mathcal{B}(\mathcal{H})}(\sigma_2 | \rho)^{x_2} / \text{app}_{\mathcal{B}(\mathcal{H})}(\sigma | \rho))^N \end{aligned}$$

is to be interpreted as the relative probability of a test, compatible with σ , having a result in this particular set. This relative probability should be independent of ρ because it depends only on the set of possible σ -compatible tests.

For example, if $\sigma_1 = \sigma_2$ then any σ -compatible sequence of trials will be both σ_1 and σ_2 compatible, so that

$$\text{app}_{\mathcal{B}(\mathcal{H})}(\sigma_1 | \rho)^{N_1} \text{app}_{\mathcal{B}(\mathcal{H})}(\sigma_2 | \rho)^{N_2} / \text{app}_{\mathcal{B}(\mathcal{H})}(\sigma | \rho)^{N_1+N_2} = 1.$$

On the other hand, if σ_1 and σ_2 are disjoint in the sense of having orthogonal support projections, then it is possible to choose $S^1 = S^2$. Then any σ -compatible test is some rearrangement of N_1 trials compatible with σ_1 and N_2 trials compatible with σ_2 . We should then have

$$\text{app}_{\mathcal{B}(\mathcal{H})}(\sigma_1 | \rho)^{N_1} \text{app}_{\mathcal{B}(\mathcal{H})}(\sigma_2 | \rho)^{N_2} / \text{app}_{\mathcal{B}(\mathcal{H})}(\sigma | \rho)^{N_1+N_2} \sim N_1!N_2!/(N_1 + N_2)! \quad (7.10)$$

as the right hand side is the reciprocal of the number of such rearrangements.

For $\mathcal{A} = \mathcal{Z}$ – an Abelian algebra, property **I**, like (7.7), is a provable statement concerning the asymptotic distribution of outcomes of repeated trials of a multinomial distribution and the justification just given can be seen to be sound. Indeed, with an obvious extension of the notation introduced in the discussion of (7.7), the core of the justification is the simple fact that, when all the relevant numbers are integers, for $N = N_1 + N_2$ trials on the random variable X , the probability

Prob(For each m , $X = m$ in $N_1 s_m^1$ of the first N_1 trials and in $N_2 s_m^2$ of the last N_2 trials, given that $X = m$ in $(N_1 + N_2)(x_1 s_m^1 + x_2 s_m^2)$ of all the trials)

is independent of the underlying distribution of X . It is, perhaps, remarkable that this independence can be generalized to the non-Abelian situation, when one considers that on $\mathcal{B}(\mathcal{H})$ a state σ may split in many different ways, corresponding to $S^1 = \{|\psi_m^1\rangle\langle\psi_m^1| : m = 1, \dots, M\}$ and $S^2 = \{|\psi_m^2\rangle\langle\psi_m^2| : m = 1, \dots, M\}$ for many different mutually incompatible bases $(\psi_m^1)_{m=1}^M$ and $(\psi_m^2)_{m=1}^M$. Indeed, S -trials can be defined consistently for arbitrary sets of disjoint states.

Property **I** generalizes further, applying for σ_1 , σ_2 , and ρ normal states on any injective von Neumann algebra \mathcal{A} . An analogous justification can be given in this context. In particular, property **I** extends to normal states on $\mathcal{B}(\mathcal{H})$ for \mathcal{H} infinite-dimensional. However, such a generalization may fail if \mathcal{A} is not an algebra; indeed, using Donald (1987a), it can be seen to fail for suitable states when \mathcal{B} is as given in example 7.8. For a non-algebra, the argument for ρ independence fails because, in terms of 7.3 and 7.5, the states on $\mathcal{B}(\mathcal{H})$ compatible with ρ which make σ_1 or σ_2 most likely may necessarily be different from the states which make σ most likely. This means that the ratio $\text{app}_{\mathcal{B}}(\sigma_1 | \rho)^{N_1} \text{app}_{\mathcal{B}}(\sigma_2 | \rho)^{N_2} / \text{app}_{\mathcal{B}}(\sigma | \rho)^{N_1+N_2}$ cannot be interpreted as a relative probability because the trials for the numerator cannot be interpreted as trials on the same unique state extension as the trials for the denominator. On the other hand, for any injective von Neumann algebra \mathcal{A} , there exists a map (a conditional expectation) $\varepsilon : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{A}$ with the property that, for any pair of states on \mathcal{A} , $\text{app}_{\mathcal{A}}(\sigma | \rho) = \text{app}_{\mathcal{B}(\mathcal{H})}(\sigma \circ \varepsilon | \rho \circ \varepsilon)$. This implies that whatever state σ we are collapsing to, we can deal with the unique extension $\rho \circ \varepsilon$ of ρ . This mathematics is compatible with, and may be taken to explain, the common experience that physical subsystems defined on algebras can be treated as complete and independent entities.

The point about property **I**, to be shown in the next section, is that it is sufficient, in combination with properties **C**, **D**, and **E**, to yield a complete definition of app on

any finite-dimensional algebra. This definition is compatible with everything else which has been claimed for app and, in particular, with (7.5), (7.7), (7.9), and (7.10).

(7.5) will give app on an arbitrary set $\mathcal{B} \subset \mathcal{B}(\mathcal{H})$ once we have app on $\mathcal{B}(\mathcal{H})$. Thus to complete our definition, we only need some method for extending from $\mathcal{B}(\mathcal{H})$ with \mathcal{H} finite-dimensional to the infinite-dimensional case. This is largely a technical problem.

The properties already assumed are sufficient to define $\text{app}_{\mathcal{B}(\mathcal{H})}(\sigma | \rho)$ for a large set of pairs σ and ρ even if \mathcal{H} is infinite-dimensional.

For example, if $\sigma = \sum_{j=1}^J s_j |\varphi_j\rangle\langle\varphi_j|$ and $\rho = \sum_{i=1}^I r_i |\psi_i\rangle\langle\psi_i|$ where $(\varphi_j)_{j=1}^J$ and $(\psi_i)_{i=1}^I$ are both contained in a single finite-dimensional subset \mathcal{H}_1 of \mathcal{H} , then, by (7.5), $\text{app}_{\mathcal{B}(\mathcal{H})}(\sigma | \rho) = \text{app}_{\mathcal{B}(\mathcal{H}_1)}(\sigma|_{\mathcal{B}(\mathcal{H}_1)} | \rho|_{\mathcal{B}(\mathcal{H}_1)})$, as the extensions to $\mathcal{B}(\mathcal{H})$ of $\sigma|_{\mathcal{B}(\mathcal{H}_1)}$ and $\rho|_{\mathcal{B}(\mathcal{H}_1)}$ are unique. In fact, the set of pairs (σ, ρ) for which $\text{app}_{\mathcal{B}(\mathcal{H})}(\sigma | \rho)$ is already defined is dense in the w^* -topology. Thus it is possible to extend our definition by imposing a continuity property on app .

Property J Let $((\sigma_\alpha, \rho_\alpha))_{\alpha \in I}$ be a net of pairs of states on some set $\mathcal{B} \subset \mathcal{B}(\mathcal{H})$ with $\sigma_\alpha(B) \rightarrow \sigma(B)$ and $\rho_\alpha(B) \rightarrow \rho(B)$ for all $B \in \mathcal{B}$. Then

$$\text{app}_{\mathcal{B}}(\sigma | \rho) \geq \limsup_{\alpha \in I} \text{app}_{\mathcal{B}}(\sigma_\alpha | \rho_\alpha).$$

Note that a net is a generalization of a sequence (see Reed and Simon (1972, §IV.2)).

Property J is referred to as “ w^* upper semicontinuity”. For a justification, suppose α to parametrize a small perturbation of a physical situation. A circumstance in which $\text{app}_{\mathcal{B}}(\sigma_\alpha | \rho_\alpha) > \text{app}_{\mathcal{B}}(\sigma | \rho)$ for all $\alpha \in I$, in other words in which property J is violated, would mean that one could be arbitrarily close to (σ, ρ) with a given a priori probability but that the limit point itself would be strictly less likely. We then make the assumption that the physical distinguishability of two states is determined entirely by the values those states take on bounded operators. As will be discussed in connection with example 9.8, this assumption may not always be appropriate, but no substitute seems sufficiently general to give a natural alternative to property J. Anyway, under this assumption, however good our technology, there is always a point $(\sigma_\alpha, \rho_\alpha)$ so close to (σ, ρ) that it would be physically impossible to distinguish the two points. This means that, allowing for arbitrarily small perturbations, $\limsup_{\alpha \in I} \text{app}_{\mathcal{B}}(\sigma_\alpha | \rho_\alpha)$ would be a maximal measure of likelihood for a situation which would inevitably be described as (σ, ρ) . That measure, however, is precisely what $\text{app}_{\mathcal{B}}(\sigma | \rho)$ is supposed to describe. If such a situation is never to arise, then property J must hold.

The contrary situation in which we have a net such that $\text{app}_{\mathcal{B}}(\sigma | \rho) > \limsup_{\alpha \in I} \text{app}_{\mathcal{B}}(\sigma_\alpha | \rho_\alpha)$ does arise under the proposed definition. In such a situation, the pair (σ, ρ) is strictly more likely than some arbitrarily close neighbours. This just means that small perturbations from (σ, ρ) on to $((\sigma_\alpha, \rho_\alpha))_{\alpha \in I}$ are unlikely. An example is as follows:

Example 7.11 Let $|\varphi\rangle\langle\varphi|$ and $|\psi\rangle\langle\psi|$ be pure states on $\mathcal{B}(\mathcal{H})$.

$$\begin{aligned} \text{Then } \text{app}_{\mathcal{B}(\mathcal{H})}(|\varphi\rangle\langle\varphi| || \psi\rangle\langle\psi|) &= 1 & \text{if } \varphi = \psi \\ &= 0 & \text{if } \varphi \neq \psi. \end{aligned}$$

This might, at first, seem surprising or even undesirable, but it can be proved just by property **C** and by applying (7.6) and (7.7) to the algebra generated by the projection onto ψ . What it says is that if one has information about all the observables in $\mathcal{B}(\mathcal{H})$ then, although it is possible to mistake a state which is a mixture for one of its components, it is not possible to make any mistake about a pure state. This indicates the importance of the set \mathcal{B} and shows that, under the present theory, coarse-graining is necessary for non-trivial measurement. Once again, we must turn to (2.2) for the role of the squared amplitude $|\langle\psi|\varphi\rangle|^2$.

W^* upper semicontinuity, unlike continuity, is not sufficient by itself to give a unique extension of a definition from a w^* dense subset to the set of all pairs of states. The problem is, for example, the possibility of isolated pairs at which app jumps up. Nevertheless, there is a unique minimal w^* upper semicontinuous extension. For precisely this extension, the supremum in (7.5) can be approximated with arbitrary accuracy by near-by states satisfying constraints like, for example, normality or finite expected energy, which we might wish to impose on physical grounds.

Thus, let \mathcal{D} be a dense vector subspace of \mathcal{H} and define sets of states on $\mathcal{B}(\mathcal{H})$ by $S(\mathcal{D}) = \{\sigma : \sigma = \sum_{j=1}^J s_j |\varphi_j\rangle\langle\varphi_j| \text{ for } J \text{ finite and } \varphi_j \in \mathcal{D}\}$ and on \mathcal{B} by $S_{\mathcal{B}}(\mathcal{D}) = \{\sigma : \sigma = \sigma'|_{\mathcal{B}} \text{ for some } \sigma' \in S(\mathcal{D})\}$.

Property K For any pair (σ, ρ) of states on \mathcal{B} and any dense vector subspace $\mathcal{D} \subset \mathcal{H}$, there exist nets $(\sigma_\alpha)_{\alpha \in I}, (\rho_\alpha)_{\alpha \in I} \subset S_{\mathcal{B}}(\mathcal{D})$ such that $\sigma_\alpha(B) \rightarrow \sigma(B)$ and $\rho_\alpha(B) \rightarrow \rho(B)$ for all $B \in \mathcal{B}$ and such that $\text{app}_{\mathcal{B}}(\sigma_\alpha | \rho_\alpha) \rightarrow \text{app}_{\mathcal{B}}(\sigma | \rho)$.

8. The definition of the a priori probability function.

The unique function satisfying the properties proposed in the [previous](#) section can be defined as follows:

$$\text{app}_{\mathcal{B}}(\sigma | \rho) = \exp\{\text{ent}_{\mathcal{B}}(\sigma | \rho)\} \tag{8.1}$$

where $\text{ent}_{\mathcal{B}}(\sigma | \rho)$ satisfies

$$\begin{aligned} 8.2) \quad \text{ent}_{\mathcal{B}(\mathcal{H})}(\sigma | \rho) &= \sum_{i,j} (-s_j \log s_j + s_j \log r_i + s_j - r_i) |\langle\varphi_j|\psi_i\rangle|^2 \\ &= \text{tr}(-\sigma \log \sigma + \sigma \log \rho) \end{aligned}$$

for $\sigma = \sum_j s_j |\varphi_j\rangle\langle\varphi_j|$ and $\rho = \sum_i r_i |\psi_i\rangle\langle\psi_i|$ normal states on $\mathcal{B}(\mathcal{H})$ expanded in orthonormal eigenvectors.

$$8.3) \quad \text{ent}_{\mathcal{B}(\mathcal{H})}(\sigma | \rho) = \inf\{F(\sigma, \rho) : F \text{ is } w^* \text{ upper semicontinuous, concave, and given by 8.2 for } \sigma \text{ and } \rho \text{ normal}\}.$$

$$8.4) \quad \text{ent}_{\mathcal{B}}(\sigma | \rho) = \sup\{\text{ent}_{\mathcal{B}(\mathcal{H})}(\sigma' | \rho') : \sigma'|_{\mathcal{B}} = \sigma \text{ and } \rho'|_{\mathcal{B}} = \rho\}.$$

$\text{ent}_{\mathcal{B}}(\sigma | \rho)$ is referred to as the relative entropy of σ with respect to ρ on the set \mathcal{B} . This function was introduced in Donald (1986) as a generalization of a widely-studied function defined for \mathcal{B} an algebra by Araki (1976), who, in turn, was generalizing ideas of Umegaki (1962) and Lindblad (1974) about quantum information theory. Two alternative derivations of (8.2) have been given by Petz (preprint) and Hiai and Petz (1991). These authors show that ent on finite-dimensional algebras satisfies two striking mathematical properties either of which could, perhaps, be used as the basis of an interpretation in place of property I. Equation 7.7 is the exponential of the (negative) “cross-entropy”, which has been extensively used for the solution of inference problems in classical probability. A review of this work and an axiomatic derivation are given by Shore and Johnson (1980, 1981).

The mathematical properties of $\text{ent}_{\mathcal{B}}(\sigma | \rho)$ have been studied at length in Donald (1986, 1987a). These papers rely heavily on the work initiated by Araki. Direct proofs, or reference to proofs, can be found there for many of the claims made in the previous section. In proving the remaining claims, in this entirely technical section, familiarity will be assumed with those papers.

proof of Property E

By Donald (1987a, lemma 2.11), $\text{ent}_{\mathcal{B}}(\rho_a | \rho) \geq \log p_a$.

But by (7.6) and (7.7) applied to the algebra generated by Q_a ,

$$\begin{aligned} & \text{ent}_{\mathcal{B}}(\rho_a | \rho) \\ & \leq -\rho_a(Q_a) \log \rho_a(Q_a)/\rho(Q_a) - \rho_a(1 - Q_a) \log \rho_a(1 - Q_a)/\rho(1 - Q_a) \\ & \sim \log p_a. \end{aligned} \tag{8.5}$$

Deriving from (8.5) the upper bound presented in 5.1 is a lengthy but prosaic exercise in mathematical analysis. This will be omitted here as that bound was only presented in order to give an explicit meaning to the symbol \sim in property E. ■

Proposition 8.6 *Properties C, D, E, and I define a unique function (given by (8.1)) on a finite-dimensional algebra \mathcal{A} .*

proof Attention will be restricted to the case of $\mathcal{A} = \mathcal{B}(\mathcal{H})$ for \mathcal{H} finite-dimensional. Other cases may be handled similarly using the structure theory for finite-dimensional algebras (Takesaki (1979, §I.11)).

Let $\mathcal{D}(\mathcal{H})$ be the set of pairs of states on $\mathcal{B}(\mathcal{H})$ on which any function satisfying properties C, D, E, and I agrees with the function given by (8.1). That function does indeed satisfy these properties on all pairs. All that is required is to show that $\mathcal{D}(\mathcal{H})$ is the set of all pairs of states.

Let $A(\sigma | \rho)$ denote an arbitrary function satisfying C, D, E, and I, and let $E(\sigma | \rho) = \log A(\sigma | \rho)$. By taking logarithms and using property C, property I implies that

$$x_1 E(\sigma_1 | \rho) + x_2 E(\sigma_2 | \rho) = E(\sigma | \rho) + x_1 E(\sigma_1 | \sigma) + x_2 E(\sigma_2 | \sigma) \tag{8.7}$$

for all σ_1, σ_2 , and ρ , where $\sigma = x_1\sigma_1 + x_2\sigma_2$.

lemma 8.8 Let $(\psi_n)_{n=1}^N$ be an orthonormal basis for \mathcal{H} and $\sigma = \sum_{i=1}^N s_i |\psi_i\rangle\langle\psi_i|$, $\rho = \sum_{i=1}^N r_i |\psi_i\rangle\langle\psi_i|$. Then $(\sigma, \rho) \in \mathcal{D}(\mathcal{H})$.

proof By property **E**, $E(|\psi_i\rangle\langle\psi_i| | \rho) = \log r_i$. A unique value is then given to $E(\sigma | \rho)$ by induction on (8.7). ■

The proof of Proposition 8.6 now follows Donald (1987a, §3). ■

lemma 8.9 Let \mathcal{A} be a finite-dimensional algebra. Let $(P_m)_{m=1}^M \in \mathcal{A}$ be a sequence of orthogonal projections and $S = \{\sigma_m : m = 1, \dots, M\}$ be a set of states such that $\sigma_m(P_m) = 1$. Then, for any state ρ on \mathcal{A} , $\sum_{m=1}^M \text{app}_{\mathcal{A}}(\sigma_m | \rho) \leq 1$ with equality if and only if ρ is in the linear span of S .

proof Note that Donald (1986) and Araki (1977, theorem 3.6) give (7.9) in its logarithmic form:

$$\text{ent}_{\mathcal{A}}(\sigma | \rho) = \sum_{m=1}^M (s_m \text{ent}_{\mathcal{A}}(\sigma_m | \rho) - s_m \log s_m) \text{ for } \sigma = \sum_{m=1}^M s_m \sigma_m. \quad (8.10)$$

Write $a_m = \text{app}_{\mathcal{A}}(\sigma_m | \rho)$ and $\alpha = \sum_{m=1}^M a_m$. Without loss of generality suppose that $\alpha > 0$. Then, setting $b_m = a_m/\alpha$, (8.10) gives

$$\text{ent}_{\mathcal{A}}\left(\sum_{m=1}^M b_m \sigma_m | \rho\right) = \sum_{m=1}^M (b_m \log a_m - b_m \log b_m) = \log \alpha.$$

By property **C**, $\log \alpha \leq 0$ (which is $\sum_{m=1}^M \text{app}_{\mathcal{A}}(\sigma_m | \rho) \leq 1$) and equality is attained

only if $\rho = \sum_{m=1}^M b_m \sigma_m$.

On the other hand, if $\rho = \sum_{m=1}^M s_m \sigma_m$ for arbitrary $(s_m)_{m=1}^M$ then

$$0 = \text{ent}_{\mathcal{A}}\left(\sum_{m=1}^M s_m \sigma_m | \rho\right) = \sum_{m=1}^M (s_m \log a_m - s_m \log s_m) \leq \sum_{m=1}^M (a_m - s_m),$$

so that $\sum_{m=1}^M a_m \geq 1$. This contradicts what has just been proved unless $\sum_{m=1}^M a_m = 1$, so that equality does hold whenever ρ is in the linear span of S . ■

proof of Property K Let $(\psi_n)_{n=1}^{\infty}$ be an orthonormal basis for \mathcal{D} . This is constructable by Gram-Schmidt orthogonalization. Let P_N be the projection onto the space spanned by $(\psi_n)_{n=1}^N$. Note that $P_N \xrightarrow{s} 1$.

Use axiom four of Donald (1986) to find states σ' and ρ' on $\mathcal{B}(\mathcal{H})$ satisfying $\sigma'|_{\mathcal{B}} = \sigma$ and $\rho'|_{\mathcal{B}} = \rho$ and a net $((\sigma'_\beta, \rho'_\beta))_{\beta \in J}$ of pairs of normal states on $\mathcal{B}(\mathcal{H})$ such that $((\sigma'_\beta, \rho'_\beta))_{\beta \in J} \xrightarrow{w^*} (\sigma', \rho')$ and $\lim_{\beta \in J} \text{app}_{\mathcal{B}(\mathcal{H})}(\sigma'_\beta | \rho'_\beta) = \text{app}_{\mathcal{B}(\mathcal{H})}(\sigma' | \rho') = \text{app}_{\mathcal{B}}(\sigma | \rho)$.

Choose N_β such that $N \geq N_\beta \Rightarrow \sigma'_\beta(P_N) > 0$ and $\rho'_\beta(P_N) > 0$, and, for $N \geq N_\beta$ define $\sigma''_{(N,\beta)} = P_N \sigma'_\beta P_N / \sigma'_\beta(P_N)$, $\rho''_{(N,\beta)} = P_N \rho'_\beta P_N / \rho'_\beta(P_N)$. Note that $\sigma''_{(N,\beta)}$ and $\rho''_{(N,\beta)} \in \mathcal{S}(\mathcal{D})$.

By Donald (1987a, lemma 2.5), as $N \rightarrow \infty$, $(\sigma''_{(N,\beta)}, \rho''_{(N,\beta)}) \xrightarrow{w^*} (\sigma'_\beta, \rho'_\beta)$ and $\text{app}_{\mathcal{B}(\mathcal{H})}(\sigma''_{(N,\beta)} | \rho''_{(N,\beta)}) \rightarrow \text{app}_{\mathcal{B}(\mathcal{H})}(\sigma'_\beta | \rho'_\beta)$.

Now given $\varepsilon > 0$ and a w^* -open neighbourhood U of (σ', ρ') we can choose β such that $(\sigma'_\beta, \rho'_\beta) \in U$ and $|\text{app}_{\mathcal{B}(\mathcal{H})}(\sigma'_\beta | \rho'_\beta) - \text{app}_{\mathcal{B}(\mathcal{H})}(\sigma' | \rho')| < \varepsilon/2$. Then we can choose $N \geq N_\beta$ and $(\sigma''_{(N,\beta)}, \rho''_{(N,\beta)}) \in U$ with

$$|\text{app}_{\mathcal{B}(\mathcal{H})}(\sigma''_{(N,\beta)} | \rho''_{(N,\beta)}) - \text{app}_{\mathcal{B}(\mathcal{H})}(\sigma'_\beta | \rho'_\beta)| < \varepsilon/2.$$

This is sufficient to yield the required property. \blacksquare

9. A priori probability for a succession of collapses.

In this section, mathematics arising from properties A and B of section 7 and from postulate eight, is analysed. The most important result in this section is proposition 9.11 which confirms the claimed consequences of postulate nine. From (8.1), $\prod_{m=1}^M \text{app}_{\mathcal{B}}(\sigma_m | \sigma_{m-1}) = \exp\{\sum_{m=1}^M \text{ent}_{\mathcal{B}}(\sigma_m | \sigma_{m-1})\}$ so we shall also be considering sums of relative entropies.

As far as property A is concerned, we need to compare $\prod_{m=1}^M \text{app}_{\mathcal{B}}(\sigma_m | \sigma_{m-1})$ with fixed \mathcal{B} to $\prod_{m=1}^M \text{app}_{\mathcal{B}_m}(\sigma_m | \sigma_{m-1})$ with an increasing sequence of sets \mathcal{B}_m . A function of the latter form does not give an appropriate definition for a priori probability, because of the role of σ_0 in giving the only input from the global initial state of postulate five (section four). In the extreme case, putting $\mathcal{B}_1 = \{1\}$ removes all input from ω . More generally, note that $\prod_{m=1}^M \text{app}_{\mathcal{B}_m}(\sigma_m | \sigma_{m-1}) = 1$ if $\sigma_m|_{\mathcal{B}_m} = \sigma_{m-1}|_{\mathcal{B}_m}$, which implies only that $\sigma_m|_{\mathcal{B}_1} = \sigma_0|_{\mathcal{B}_1}$ for $m = 1, \dots, M$. $\prod_{m=1}^M \text{app}_{\mathcal{B}}(\sigma_m | \sigma_{m-1}) = 1$ only if $\sigma_m|_{\mathcal{B}} = \sigma_0|_{\mathcal{B}}$ for $m = 1, \dots, M$.

It is immediate from the results for $\text{ent}_{\mathcal{B}}(\sigma | \rho)$ that $\sum_{m=1}^M \text{ent}_{\mathcal{B}}(\sigma_m | \sigma_{m-1})$ has properties of monotonicity, concavity, w^* upper-semicontinuity, non-positivity, non-triviality, the Uhlmann-Lindblad inequality, and Araki's property corresponding to those described in Donald (1986) for $\text{ent}_{\mathcal{B}}(\sigma | \rho)$. The following proposition corresponds to Theorem 4.4 of Donald (1987a) and has a similar proof. We introduce the notation $\Sigma^*(\mathcal{A})^{\times M}$ (resp. $\Sigma_*(\mathcal{A})^{\times M}$) for the set of M -tuples of states (resp. normal states) on an algebra \mathcal{A} .

Proposition 9.1 *Let ω be a faithful normal state on an injective von Neumann algebra \mathcal{A} . Let $K \subset \Sigma^*(\mathcal{A})^{\times M}$ be a w^* -closed convex set with non-empty interior.*

Then there is a unique element $(\tilde{\sigma}(K, \omega)_m)_{m=1}^M \in K$ which attains $\sup\{\text{app}_{\mathcal{B}}((\sigma_m)_{m=1}^M | \omega) : (\sigma_m)_{m=1}^M \in K\}$. $(\tilde{\sigma}(K, \omega))_{m=1}^M \in \Sigma_(\mathcal{A})^{\times M}$.* \blacksquare

This result is very satisfactory mathematically, and may be useful elsewhere, but, in the context of this paper, it is irrelevant in several different ways. In postulate eight, an inductive sequence of suprema is considered rather than a single supremum over the entire set of sequences. Such an inductive sequence is physically more appropriate and helps to avoid the problem to be raised by example 9.3. Under the conditions of

proposition 9.1, Theorem 4.4 of Donald (1987a) can also be applied to show that such an inductive definition will lead to a unique best sequence of states in $K \subset \Sigma^*(\mathcal{A})^{\times M}$. This too is irrelevant here, as because of property 3.9, \mathcal{B}_{S_M} cannot be assumed to be an algebra. Section 5 of Donald (1987a) demonstrates that on non-algebras uniqueness of suprema-attaining states cannot be claimed. Finally, no variant of proposition 9.1 is directly relevant because, as will be explained in example 9.8, the set \mathcal{N}_T of section 4 cannot be assumed to be w^* -closed.

It is essential to the interpretation of $\text{app}_{\mathcal{B}}((\sigma_m)_{m=1}^M | \omega)$ that it should establish a non-trivial correlation between ω and σ_M . That it does is shown by the following extension of the non-triviality property:

Proposition 9.2

$$\sum_{m=1}^M \text{ent}_{\mathcal{B}}(\sigma_m | \sigma_{m-1}) \leq \inf \{ -|\sigma_M(A) - \sigma_0(A)|^2 / (2M\|A\|^2) : A \in \mathcal{B} \}.$$

proof By concavity and property f of Donald (1986), for all $A \in \mathcal{B}$,

$$\begin{aligned} \sum_{m=1}^M \text{ent}_{\mathcal{B}}(\sigma_m | \sigma_{m-1}) &= M \left(\sum_{m=1}^M 1/M \text{ent}_{\mathcal{B}}(\sigma_m | \sigma_{m-1}) \right) \\ &\leq M \text{ent}_{\mathcal{B}}(1/M \sigma_M + 1/M \sum_{m=1}^{M-1} \sigma_m | 1/M \sigma_0 + 1/M \sum_{m=1}^{M-1} \sigma_m) \\ &\leq -M |1/M \sigma_M(A) - 1/M \sigma_0(A)|^2 / (2\|A\|^2). \quad \blacksquare \end{aligned}$$

The M dependence in this result is important. The next example shows that by interpolating specified sequences of arbitrarily many ‘‘collapses’’ between fixed σ_0 and σ_M we can make $\text{app}_{\mathcal{B}}((\sigma_m)_{m=1}^M | \omega)$ arbitrarily small. This is analogous for the present theory to the ‘‘quantum Zeno paradox’’ of the conventional interpretation (see Exner (1985, §2.4)). However, in the present theory the paradox is avoided, partly because M , the number of collapses, is forced to be finite by the natural device of linking it to a sequence of real changes in the brain of the observer, and partly because the inductive definition (4.6) is such as to disallow a sequence of collapses precisely adjusted towards some future goal. In the conventional interpretation, with no unambiguous definition of measurement, there can be no telling how many measurements take place.

example 9.3 Let $\varphi, \psi \in \mathcal{H}$ be orthogonal and normalized vectors.

Let $\sigma_m = \frac{1}{2}(1 + \frac{m}{M})|\varphi\rangle\langle\varphi| + \frac{1}{2}(1 - \frac{m}{M})|\psi\rangle\langle\psi|$, so that $\sigma_0 = \frac{1}{2}|\varphi\rangle\langle\varphi| + \frac{1}{2}|\psi\rangle\langle\psi|$ and $\sigma_M = |\varphi\rangle\langle\varphi|$.

Then $\text{app}_{\mathcal{B}(\mathcal{H})}((\sigma_m)_{m=1}^M | \sigma_0) \rightarrow 1$ as $M \rightarrow \infty$.

proof By (8.2),

$$\sum_{m=1}^M \text{ent}_{\mathcal{B}(\mathcal{H})}(\sigma_m | \sigma_{m-1})$$

$$= -\frac{1}{2} \sum_{m=1}^M \left\{ \left(1 + \frac{m}{M}\right) \log\left(\frac{1 + m/M}{1 + m/M - 1/M}\right) + \left(1 - \frac{m}{M}\right) \log\left(\frac{1 - m/M}{1 - m/M + 1/M}\right) \right\}.$$

Set $x = \frac{m}{M}$ and consider the sum as the integral of a step function on the interval $[0, 1]$. $\log y \leq y - 1$ for $y \geq 0$, so, for $0 \leq x \leq 1$ and $M \geq 2$,

$$\begin{aligned} 0 &\leq M \left\{ (1+x) \log\left(\frac{1+x}{1+x-1/M}\right) + (1-x) \log\left(\frac{1-x}{1-x+1/M}\right) \right\} \\ &\leq \frac{1+x}{1+x-1/M} - \frac{1-x}{1-x+1/M} = \frac{2}{M(1+x-1/M)(1-x+1/M)} \leq 4. \end{aligned}$$

The dominated convergence theorem gives the result. \blacksquare

example 9.4 Let $(\psi_n)_{n \geq 1}$ be an orthonormal basis for \mathcal{H} and let

$$\sigma_2 = \sum_{n=1}^{\infty} 6/n^2 \pi^2 |\psi_n\rangle\langle\psi_n|, \quad \sigma_1 = \sum_{n=1}^{\infty} 90/n^4 \pi^4 |\psi_n\rangle\langle\psi_n|, \quad \sigma_0 = \sum_{n=1}^{\infty} 1/2^n |\psi_n\rangle\langle\psi_n|.$$

Then $\text{app}_{\mathcal{B}(\mathcal{H})}((\sigma_m)_{m=1}^2 | \sigma_0) > 0$ while $\text{app}_{\mathcal{B}(\mathcal{H})}(\sigma_2 | \sigma_0) = 0$. \blacksquare

This should drive home the point that, under the proposed definition of a priori probability, it can be more probable to go from σ_0 to σ_2 via suitable σ_1 rather than directly.

Turn now to the mathematical analysis of postulate [eight](#). The notation will be variously simplified, for example, by writing $\text{app}(k)$ (resp. $\text{app}^0(k)$) for $\text{app}(\mathcal{N}_T, \mathcal{B}_{SM}, k, \omega)$ (resp. $\text{app}^0(\mathcal{N}_T, \mathcal{B}_{SM}, k, \omega)$) and by not mentioning \mathcal{B}_{SM} . It will always be assumed that $\mathcal{N}^M \neq \emptyset$.

example 9.5 Let $\mathcal{N}^3 = \{(\sigma_i^n)_{i=1}^3 : n \geq 1\} \cup \{(\rho_i)_{i=1}^3\}$ where the states ω , ρ_i , and σ_i^n are chosen to satisfy $\text{app}(\sigma_1^n | \omega) = \frac{1}{2} - 1/2^n$, $\text{app}(\rho_1 | \omega) = \frac{1}{2}$, $\text{app}(\sigma_2^n | \sigma_1^n) = \frac{1}{2}$, $\text{app}(\rho_2 | \rho_1) = \frac{1}{4}$, $\text{app}(\sigma_3^n | \sigma_2^n) = \frac{1}{8}$, and $\text{app}(\rho_3 | \rho_2) = \frac{1}{2}$.

Then $\text{app}(1) = \text{app}^0(1) = \frac{1}{2}$, $\text{app}(2) = \frac{1}{4} > \frac{1}{8} = \text{app}^0(2)$, and $\text{app}(3) = \frac{1}{32} < \frac{1}{16} = \text{app}^0(3)$. \blacksquare

This example no doubt uses an entirely unphysical choice for \mathcal{N}^3 . Although the preliminary version of postulate [eight](#) does yield a non-trivial definition, that definition is unsatisfactory. At stage 2, the sequence $(\rho_i)_{i=1}^2$ is comparatively unlikely. Because of this the likelihood of $(\rho_i)_{i=1}^3$ at the next stage should be irrelevant.

lemma 9.6 For $\delta > 0$ and $1 \leq m \leq M$, $\tilde{\mathcal{N}}_\delta^m$ is not empty and there exists a sequence $((\sigma_i^n)_{i=1}^M)_{n \geq 1} \subset \mathcal{N}^M$ such that $\text{app}((\sigma_i^n)_{i=1}^k | \omega) \rightarrow \text{app}(k)$ for $k = 1, \dots, M$.

proof There certainly exists a sequence $((\sigma_i^{1,n})_{i=1}^M)_{n \geq 1} \subset \mathcal{N}^M$ such that $\text{app}(\sigma_1^{1,n} | \omega) \rightarrow \text{app}(1)$.

Suppose then, that for some $m < M$ there exists a sequence $((\sigma_i^{m,n})_{i=1}^M)_{n \geq 1} \subset \mathcal{N}^M$ such that $\text{app}((\sigma_i^{m,n})_{i=1}^k | \omega) \rightarrow \text{app}(k)$ for $k = 1, \dots, m$.

It follows that $\text{app}(m+1)$ is well-defined and that, for $\delta > 0$, $\tilde{\mathcal{N}}_\delta^m \neq \emptyset$.

For $N \geq 1$, choose a sequence $((\rho_i^{N,n})_{i=1}^M)_{n \geq 1} \subset \mathcal{N}^M$ such that for $k = 1, \dots, m$, $\text{app}((\rho_i^{N,n})_{i=1}^k | \omega) \rightarrow \text{app}(k)$ and such that

$\limsup_{n \rightarrow \infty} \text{app}((\rho_i^{N,n})_{i=1}^{m+1} | \omega) \geq \text{app}(m+1) - 1/(2N)$. Then there exists N_0 such

that $|\text{app}((\rho_i^{N, N_0})_{i=1}^k | \omega) - \text{app}(k)| \leq 1/N$ for $k = 1, \dots, m+1$. Set $(\sigma_i^{m+1, N})_{i=1}^M = (\rho_i^{N, N_0})_{i=1}^M$. Then $((\sigma_i^{m+1, n})_{i=1}^M)_{n \geq 1} \subset \mathcal{N}^M$ and $\text{app}((\sigma_i^{m+1, n})_{i=1}^k | \omega) \rightarrow \text{app}(k)$ for $k = 1, \dots, m+1$.

The lemma follows by induction. ■

lemma 9.7 *If \mathcal{N}^M is w^* -closed then $\text{app}^0(m) = \text{app}(m)$ for $1 \leq m \leq M$.*

proof Let $((\sigma_i^n)_{i=1}^M)_{n \geq 1} \subset \mathcal{N}^M$ be a sequence, as given in lemma 9.6, such that $\text{app}((\sigma_i^n)_{i=1}^k | \omega) \rightarrow \text{app}(k)$ for $k = 1, \dots, M$. Let $((\sigma_i^\alpha)_{i=1}^M)_{\alpha \in I}$ be a w^* -convergent subnet with $(\sigma_i^\alpha)_{i=1}^M \rightarrow (\tilde{\sigma}_i)_{i=1}^M$. $(\tilde{\sigma}_i)_{i=1}^M \in \mathcal{N}^M$ and so $(\tilde{\sigma}_i)_{i=1}^m \in \mathcal{N}^m$ for $1 \leq m < M$. By w^* upper-semicontinuity, $\text{app}((\tilde{\sigma}_i)_{i=1}^k | \omega) \geq \text{app}(k)$ for $k = 1, \dots, M$.

Clearly, $\text{app}(\tilde{\sigma}_1 | \omega) = \text{app}(1) = \text{app}^0(1)$. Suppose, for some $m < M$, that $\text{app}((\tilde{\sigma}_i)_{i=1}^k | \omega) = \text{app}(k) = \text{app}^0(k)$ for $k = 1, \dots, m$.

The definitions of app and app^0 then give

$$\text{app}(m+1) \geq \text{app}^0(m+1) \geq \text{app}((\tilde{\sigma}_i)_{i=1}^{m+1} | \omega),$$

and, as $\text{app}((\tilde{\sigma}_i)_{i=1}^{m+1} | \omega) \geq \text{app}(m+1)$, an inductive proof is complete. ■

This lemma is valuable in relating the two versions of postulate eight, and it may be useful in computations. However, I do not expect it to be directly relevant to the set \mathcal{N}_T of postulate four:

example 9.8 Mathematically, it is simplest to define closeness of states σ and ρ on a set \mathcal{B} by the closeness of the values $\sigma(B)$ and $\rho(B)$ for $B \in \mathcal{B}$. However, physically, it is necessary also to consider unbounded operators. This arises implicitly, for example, in the differentiability requirement of Hypothesis V(2) of Donald (1990). The present example shows that app can yield a satisfactory theory dealing with such operators, and suggests that, despite lemma 9.7, it can be important not to take closures of sets of states.

Let $(\psi_n)_{n \geq 1}$ be an orthonormal basis for \mathcal{H} and let $\rho = \sum_{n=1}^{\infty} 2^{-n} |\psi_n\rangle\langle\psi_n|$. Let $H = \sum_{n=1}^{\infty} n |\psi_n\rangle\langle\psi_n|$ and consider H as an energy operator. A typical set of states, which we might wish to consider collapse to, is $K = \{\sigma : 3 \leq \sigma(H) < \infty\}$. $\sup\{\text{app}_{\mathcal{B}(\mathcal{H})}(\sigma | \rho) : \sigma \in K\} = 27/32$ and this supremum is uniquely attained at $\tilde{\sigma} = \frac{1}{2} \sum_{n=1}^{\infty} (2/3)^n |\psi_n\rangle\langle\psi_n|$. However, the closure of K in the w^* -topology (resp. the norm topology) is the set of all states (resp. all normal states) on $\mathcal{B}(\mathcal{H})$. On these closures, the supremum would be 1 and would be uniquely attained at ρ . ■

lemma 9.9 *Suppose that $\mathcal{N}(1)^M \subset \mathcal{N}(2)^M$ and that, for some $m \leq M$, $\text{app}(\mathcal{N}(1), m) \neq \text{app}(\mathcal{N}(2), m)$. Then there exists $m' \leq m$ such that $\text{app}(\mathcal{N}(1), k) \leq \text{app}(\mathcal{N}(2), k)$ for $k = 1, \dots, m' - 1$ and $\text{app}(\mathcal{N}(1), m') < \text{app}(\mathcal{N}(2), m')$.*

proof Suppose not. Then there would exist $m' \leq m$ such that $\text{app}(\mathcal{N}(1), k) = \text{app}(\mathcal{N}(2), k)$ for $k = 1, \dots, m' - 1$ and $\text{app}(\mathcal{N}(1), m') > \text{app}(\mathcal{N}(2), m')$.

Let $((\sigma_i^n)_{i=1}^M)_{n \geq 1} \subset \mathcal{N}(1)^M$ be a sequence given by lemma 9.6 such that $\text{app}((\sigma_i^n)_{i=1}^k | \omega) \rightarrow \text{app}(\mathcal{N}(1), k)$ for $k = 1, \dots, M$.

Then $\text{app}((\sigma_i^n)_{i=1}^k | \omega) \rightarrow \text{app}(\mathcal{N}(2), k)$ for $k = 1, \dots, m' - 1$. The definition of $\text{app}(\mathcal{N}(2), m')$ requires that

$\text{app}(\mathcal{N}(2), m') \geq \limsup_{n \rightarrow \infty} \text{app}((\sigma_i^n)_{i=1}^{m'} | \omega) = \text{app}(\mathcal{N}(1), m')$. This contradiction proves the result. \blacksquare

This lemma emphasizes the absolute priority given by the definitions to the maximization of a priori probability at an early stage over its maximization at a later stage. Similar results can be proved for decreasing sets \mathcal{B} and, as exemplified by 9.5, if $\text{app}(m) \neq \text{app}^0(m)$.

lemma 9.10 For $\delta > 0$, let $(\sigma_i^\delta)_{i=1}^m \in \tilde{\mathcal{N}}_\delta^m$.

Then, as $\delta \rightarrow 0$, $\text{app}((\sigma_i^\delta)_{i=1}^k | \omega) \rightarrow \text{app}(k)$ for $k = 1, \dots, m$.

proof By definition $\limsup_{\delta \rightarrow 0} \text{app}((\sigma_i^\delta)_{i=1}^k | \omega) \geq \text{app}(k)$ for $k = 1, \dots, m$. If the result does not hold then there exists $m' \leq m$ such that $\text{app}((\sigma_i^\delta)_{i=1}^k | \omega) \rightarrow \text{app}(k)$ for $k = 1, \dots, m' - 1$, while $\limsup_{\delta \rightarrow 0} \text{app}((\sigma_i^\delta)_{i=1}^{m'} | \omega) > \text{app}(m')$. This contradicts the definition of $\text{app}(m')$. \blacksquare

Proposition 9.11 (6.2) and (6.4) hold under the conditions of postulate nine.

proof Use the method of lemma 9.6 to find a sequence $((\sigma_m^n)_{m=1}^{M+1})_{n \geq 1}$ with $((\sigma_m^n)_{m=1}^M)_{n \geq 1} \subset \mathcal{N}^M$ and $(\sigma_{M+1}^n)_{n \geq 1} \subset \Sigma$ such that

$\text{app}_{\mathcal{B}_{SM}}((\sigma_i^n)_{i=1}^k | \omega) \rightarrow \text{app}(\mathcal{N}_T, \mathcal{B}_{SM}, k, \omega)$ for $k = 1, \dots, M$, and

$\text{app}_{\mathcal{B}_{SM}}((\sigma_m^n)_{m=1}^{M+1} | \omega) \rightarrow \text{app}(O, T, \mathcal{B}_S, \Sigma | \omega)$. Then the sequence $((\rho_m^n)_{m=1}^{M+1})_{n \geq 1}$ defined by $\rho_{M+1}^n = \sigma_{M+1}^n$ for $n \geq 1$ and $\rho_m^n = \sigma_m^n$ for $n \geq 1$ and $m = 1, \dots, M$ has the property that

$$\begin{aligned} \text{app}(O, T, \mathcal{B}_S, \Sigma | \omega) &\geq \limsup_{n \rightarrow \infty} \text{app}_{\mathcal{B}_{SM}}((\rho_m^n)_{m=1}^{M+1} | \omega) = \limsup_{n \rightarrow \infty} \text{app}_{\mathcal{B}_{SM}}((\sigma_m^n)_{m=1}^M | \omega) \\ &= \text{app}(\mathcal{N}_T, \mathcal{B}_{SM}, M, \omega) \geq \limsup_{n \rightarrow \infty} \text{app}_{\mathcal{B}_{SM}}((\sigma_m^n)_{m=1}^{M+1} | \omega) = \text{app}(O, T, \mathcal{B}_S, \Sigma | \omega). \end{aligned}$$

This shows that $\text{app}(O, T, \mathcal{B}_S, \Sigma | \omega) = \text{app}(\mathcal{N}_T, \mathcal{B}_{SM}, M, \omega)$. (9.12)

Let $(\sigma_m)_{m=1}^{M+1} \in \tilde{\mathcal{N}}_\delta^{M+1}(\mathcal{B}_{SM}, C_a)$. Note that $(\sigma_m)_{m=1}^M \in \tilde{\mathcal{N}}_\delta^M(\mathcal{B}_{SM})$. By (7.6), (7.7), and A of postulate nine, for δ sufficiently small,

$$\begin{aligned} \text{app}_{\mathcal{B}_{SM}}(\sigma_{M+1} | \sigma_M) &\leq \exp\{-\sigma_{M+1}(Q_a) \log(\sigma_{M+1}(Q_a)/\sigma_M(Q_a)) \\ &\quad - \sigma_{M+1}(1 - Q_a) \log(\sigma_{M+1}(1 - Q_a)/\sigma_M(1 - Q_a))\} \sim p_a. \end{aligned}$$

Applying the method of lemma 9.6 again, this yields

$$\text{app}(O, T, \mathcal{B}_S, C_a | \omega) \leq (\sim) p_a \text{app}(\mathcal{N}_T, \mathcal{B}_{SM}, M, \omega) = p_a \text{app}(O, T, \mathcal{B}_S, \Sigma | \omega).$$

Let $(\sigma_m^\delta)_{m=1}^M \in \tilde{\mathcal{N}}_\delta^M(\mathcal{B}_{SM})$ be given by B of postulate nine and set $\sigma_{M+1}^\delta = \rho_a^\delta$. By (5.1) and (6.1), $\text{app}_{\mathcal{B}_{SM}}(\rho_a^\delta | \sigma_M^\delta) \sim p_a$ and $\text{app}_{\mathcal{B}_{SM}}(\rho_a^\delta |_{\mathcal{B}_S} | \sigma_M^\delta |_{\mathcal{B}_S}) \sim p_a$, so that (6.4) holds.

By lemma 9.10, as $\delta \rightarrow 0$, $\text{app}_{\mathcal{B}_{SM}}((\sigma_i^\delta)_{i=1}^k | \omega) \rightarrow \text{app}(\mathcal{N}_T, \mathcal{B}_{SM}, k, \omega)$ for $k = 1, \dots, M$ and by (4.7) and (9.12),

$$\begin{aligned} \text{app}(O, T, \mathcal{B}_S, C_a | \omega) &\geq \limsup_{\delta \rightarrow 0} \text{app}_{\mathcal{B}_{SM}}((\sigma_m^\delta)_{m=1}^{M+1} | \omega) \\ &\sim p_a \text{app}(\mathcal{N}_T, \mathcal{B}_{SM}, M, \omega) = p_a \text{app}(O, T, \mathcal{B}_S, \Sigma | \omega). \end{aligned} \quad \blacksquare$$

10. Elementary Models.

In this section, three mathematically elementary models of the set \mathcal{B}_M of postulate four are discussed in order to explain the development of some of the abstract structures of earlier sections. The results used will be quite simple properties of the function defined by (8.2), so proofs have been omitted. We begin with the assumption that an observer comes into contact with a succession of independent physical subsystems. These are modelled as independent Hilbert spaces $(\mathcal{H}_m)_{m=1}^M$ forming a tensor product subspace $\otimes_{m=1}^M \mathcal{H}_m$ of the total universal Hilbert space \mathcal{H} . The set of operators that define states individually on these subsystems is $\mathcal{B}_M^0 = c(\cup_{m=1}^M \mathcal{B}(\mathcal{H}_m))$, where, as usual, $\mathcal{B}(\mathcal{H}_m)$ is identified as a subalgebra of $\mathcal{B}(\mathcal{H})$ and c is defined in 3.13.

The independence assumption made here is a simplification of the more realistic situation considered earlier in which the $\mathcal{B}(\mathcal{H}_m)$ correspond to distinct, but not necessarily commuting, local algebras. It is not unreasonable however to claim on physical grounds that local algebras at distinct times are “nearly” independent, because of the variety of physically independent states that may be imposed on such regions. A version of this argument was used in section 6 in imagining the duplicitous graduate student.

\mathcal{B}_M^0 contains no operators measuring correlations between the M subsystems. Thus $\text{app}_{\mathcal{B}_M^0}$ is an inappropriate function for determining physical a priori probabilities. For example, take $M = 2$ and consider the problem of finding the state $\tilde{\sigma}_2$ on $\mathcal{B}(\mathcal{H}_2)$ maximizing $\text{app}_{\mathcal{B}_2^0}((\sigma_i)_{i=1}^2 | \omega)$ for given ω and $\sigma_1|_{\mathcal{B}(\mathcal{H}_1)}$. This problem has the solution that $\tilde{\sigma}_2|_{\mathcal{B}(\mathcal{H}_2)} = \omega|_{\mathcal{B}(\mathcal{H}_2)}$. This is useless for present purposes as it takes no account of the information gained by the observer from his observations on $\mathcal{B}(\mathcal{H}_1)$.

The most obvious way to include correlations is to replace \mathcal{B}_M^0 by the von Neumann algebra $\mathcal{B}_M^1 = \mathcal{B}(\otimes_{m=1}^M \mathcal{H}_m) = \otimes_{m=1}^M \mathcal{B}(\mathcal{H}_m)$. For particular choices of ω , \mathcal{B}_M^1 would be satisfactory. For example, once again choose $M = 2$. Let $(\psi_i)_{i=1}^\infty$ be an orthonormal basis of \mathcal{H}_1 and $(\varphi_j)_{j=1}^\infty$ be an orthonormal basis of \mathcal{H}_2 . Suppose that ω is diagonal in the product basis $(\psi_i \otimes \varphi_j)_{i=1, j=1}^\infty$, so that

$$\omega = \sum_{i,j=1}^{\infty} r_{ij} |\psi_i \varphi_j\rangle \langle \psi_i \varphi_j|. \quad (10.1)$$

Then consider the problem of finding the state $\tilde{\sigma}_2$ on $\mathcal{B}(\mathcal{H}_2)$ maximizing $\text{app}_{\mathcal{B}_2^1}((\sigma_i)_{i=1}^2 | \omega)$ given that $\sigma_1|_{\mathcal{B}(\mathcal{H}_1)} = |\psi_1\rangle \langle \psi_1|$. This problem has unique solution $\tilde{\sigma}_2|_{\mathcal{B}(\mathcal{H}_2)} = \sum_{j=1}^{\infty} r_{1j} |\varphi_j\rangle \langle \varphi_j| / \sum_{j=1}^{\infty} r_{1j}$ as long as $\sum_{j=1}^{\infty} r_{1j} > 0$.

As a model, this is fairly satisfactory. Indeed, the whole of Everett’s analysis of quantum mechanics is based on this sort of idea. The problem lies in the justification of (10.1) and in its generalization. The function app has been developed to generalize the idea of the r_{ij} in (10.1) being probabilities. This generalization, however, has the property demonstrated by example 7.11. Thus if ω happens to be a pure state on \mathcal{B}_M^1 then the problem considered has no solution except in the trivial case that $\sigma_1|_{\mathcal{B}(\mathcal{H}_1)} = \omega|_{\mathcal{B}(\mathcal{H}_1)}$. To resolve this difficulty, one either has to change the function app , and I know of no plausible alternatives; or argue that an assumption analogous to (10.1) is physically natural, which, in view of the problem raised by 3.9, I cannot see how to do; or look for some appropriate set \mathcal{B}_M^2 between \mathcal{B}_M^0 and \mathcal{B}_M^1 .

Suppose then that \mathcal{Z}_M is an Abelian algebra generated by suitable projections from $\cup_{m=1}^M \mathcal{B}(\mathcal{H}_m)$, and let $\mathcal{B}_M^2 = c\{BC : B \in \cup_{m=1}^M \mathcal{B}(\mathcal{H}_m), C \in \mathcal{Z}_M\}$. \mathcal{Z}_M is a version of the algebra \mathcal{C}_M of postulate [four](#), so that this model envisages that the projections defined in [\(4.1\)](#) are mutually commuting. A further simplification, reducing the mathematics of this paper to classical probability theory, would result if it could be claimed that

$$\text{app}_{\mathcal{B}_M^2}((\sigma_m)_{m=1}^M | \omega) = \text{app}_{\mathcal{Z}_M}((\sigma_m)_{m=1}^M | \omega). \quad (10.2)$$

The theory presented has indeed been developed under the assumptions that \mathcal{C}_M behaves as if it were Abelian, and that [\(10.2\)](#) is a good analogy for the states of highest a priori probability of postulate [eight](#) with \mathcal{B}_M^2 replaced by \mathcal{B}_M and \mathcal{Z}_M by \mathcal{C}_M . While the final result is a mathematical theory independent of the correctness of these assumptions, they remain plausible because the world appears to behave so classically.

In the sequel to this paper, I intend to consider variations in geometrical structure for an observer. It follows from [\(7.6\)](#) that maximizing a priori probability over such variations should make most likely those structures which satisfy (or effectively satisfy) both of the assumptions just mentioned. The commutativity of \mathcal{C}_M is made likely because the algebra generated by the projections defined in [\(4.1\)](#) will be smallest precisely when those projections commute. Because, by [\(7.6\)](#), we always have $\text{app}_{\mathcal{B}_M}((\sigma_m)_{m=1}^M | \omega) \leq \text{app}_{\mathcal{C}_M}((\sigma_m)_{m=1}^M | \omega)$, the second assumption can be interpreted as the claim that the states in the neighbourhood \mathcal{N}_T of postulate [four](#) are determined by their values on \mathcal{C}_M and can be varied freely off \mathcal{C}_M until maximum a priori probability is achieved.

A determining algebra of this sort would be a set of “definitive observables” in the sense discussed in section [2](#). The idea then arises that one should throw away all of \mathcal{B}_M except for \mathcal{C}_M . This, however, would be inappropriate, both because of the approximate nature of the second assumption and because, as can be seen from hypothesis V of [Donald \(1990\)](#), the abstract theory which allows \mathcal{C}_M to be defined depends on the behaviour of the states in \mathcal{N}_T on the whole of \mathcal{B}_M .

11. Consistency.

Three different consistency issues arise in this paper. These are; overall consistency of the underlying interpretation of quantum theory; consistency between the observations of distinct observers; and consistency between the various types of probability mentioned. Little can be said about the first issue until the full details of the interpretation have been presented. It is intended in a sequel to this paper to propose a complete characterization for the physical structure of observers. This will involve, in particular, developing a formalism to permit the claim that a given observer has only a finite number of possible distinct futures within a given bounded complexity.

Consistency between the observations of distinct observers depends, in the first place, on each individual observer seeing other observers as observers rather than as superpositions or mixtures. This, in other words, is part of the general problem of what it is that an individual observer can observe. In [Donald \(1990\)](#), it was proposed that the brain acts as an observer by processing definite information and that its

quantum state, when so acting, is a state characterized by certain neural proteins having definite status. When these proteins do have definite status, the brain is processing information in a way which, in theory, can be interpreted by an external neurophysiologically-expert observer. The neighbourhoods \mathcal{N}_T of postulate four are to be taken as consisting of sequences of brain states of this type. When this is done, the only sets of states C in (4.7) of high a priori probability will be states compatible with the prior information processing by the brain. In particular, when a human observer interacts with a colleague who occupies a macroscopically mixed quantum state, the brain of the original observer moves into a mixed quantum state. That mixture then must be disentangled by the set of possible neighbourhoods \mathcal{N}_T at the new time T . The consistency between the observations of observer A and the observations that observer A observes observer B to be making then stems from the correlations within the states of each \mathcal{N}_T . There is a correlation between the number I write down in my notebook as we look at the bottom line of our computer printout and the number I see you writing down, because otherwise, assuming neither of us is making mistakes, there must be additional “collapses” costing (logarithmically) large amounts of a priori probability in the quantum state of my brain between when I write down my number and when I look over your shoulder. This large logarithmic cost is argued for by the justification of (6.3). The underlying fact is that single states in quantum theory provide good descriptions of most observed causal processes – like light reflecting off the printout and carrying the same message to two different observers. Only intermittent collapse is required.

The central focus of this paper has been on probability. A mathematical tool has been provided and its relationship to other types of probability has been discussed. Whether that tool has any use independent of the framework of Donald (1990) is up to the reader.

References.

- Araki, H. 1963 “A generalization of Borchers’ theorem.” *Helv. Phys. Acta.* **36**, 132–139.
- Araki, H. 1976 “Relative entropy of states of von Neumann algebras.” *Publ. Res. Inst. Math. Sci. (Kyoto)* **11**, 809–833.
- Araki, H. 1977 “Relative entropy for states of von Neumann algebras II.” *Publ. Res. Inst. Math. Sci. (Kyoto)* **13**, 173–192.
- Araki, H. 1980 “A remark on Machida-Namiki theory of measurement.” *Prog. Theor. Phys.* **64**, 719–730.
- Araki, H. 1986 “A continuous superselection rule as a model of classical measuring apparatus in quantum mechanics.” In *Fundamental Aspects of Quantum Theory*, Gorini, V. and Frigerio, A. (eds) pp 23–33, (Plenum).
- Borchers, H.J. 1961 “Über die Vollständigkeit lorentzinvarianter Felder in einer zeitartigen Röhre.” *Il Nuovo Cimento* **19**, 787–793.
- Bratteli, O. and Robinson, D.W. 1979 *Operator Algebras and Quantum Statistical Mechanics*. (Springer-Verlag) Vol. I.
- Bratteli, O. and Robinson, D.W. 1981 *Operator Algebras and Quantum Statistical Mechanics*. (Springer-Verlag) Vol. II.

- Buchholz, D. and Wichmann, E.H. 1986 “Causal independence and the energy-level density of states in local quantum field theory.” *Commun. Math. Phys.* **106**, 321–344.
- Buchholz, D., D’Antoni, C., and Fredenhagen, K. 1987 “The universal structure of local algebras.” *Commun. Math. Phys.* **111**, 123–135.
- Cohen, L.J. 1989 *An Introduction to the Philosophy of Induction and Probability*. (Oxford).
- Daneri, A., Loinger, A., and Prosperi, G.M. 1962 “Quantum theory of measurement and ergodicity conditions.” *Nucl. Phys.* **33**, 297–319. Reprinted in [Wheeler and Zurek \(1983\)](#).
- De Facio, B. and Taylor, D.C. 1973 “Commutativity and causal independence.” *Phys. Rev. D* **8**, 2729–2731.
- DeWitt, B.S. and Graham, N. 1973 *The Many-Worlds Interpretation of Quantum Mechanics*. (Princeton).
- Donald, M.J. 1986 “On the relative entropy.” *Commun. Math. Phys.* **105**, 13–34.
- Donald, M.J. 1987a “Further results on the relative entropy.” *Math. Proc. Camb. Phil. Soc.* **101**, 363–373.
- Donald, M.J. 1987b “Free energy and the relative entropy.” *J. Stat. Phys.* **49**, 81–87.
- Donald, M.J. 1990 “Quantum theory and the brain.” *Proc. R. Soc. Lond. A* **427**, 43–93.
- Driessler, W., Summers, S.J., and Wichmann, E.H. 1986 “On the connection between quantum fields and von Neumann algebras of local operators.” *Commun. Math. Phys.* **105**, 49–84.
- Ekstein, H. 1969 “Presymmetry II.” *Phys. Rev.* **184**, 1315–1337 with correction in *Phys. Rev. D* **1**, 1851(E) (1970).
- Everett, H., III 1957 “The theory of the universal wave function.” In [DeWitt and Graham \(1973\)](#).
- Exner, P. 1985 *Open Quantum Systems and Feynman Integrals*. (Reidel).
- Garber, W.-D. 1975 “The connexion of duality and causal properties for generalized free fields.” *Commun. Math. Phys.* **42**, 195–208.
- Glimm, J. and Jaffe, A. 1971 “Field theory models.” In *Statistical Mechanics and Quantum Field Theory*, DeWitt, C. and Stora, R. (eds) pp 1–108, (Gordon and Breach). Reprinted in Glimm and Jaffe (1985).
- Glimm, J. and Jaffe, A. 1972 “Boson quantum field models.” In *Mathematics of Contemporary Physics*, Streater, R.F. (ed.) pp 77–143 (Academic). Reprinted in Glimm and Jaffe (1985).
- Glimm, J. and Jaffe, A. 1979 “The resummation of one particle lines.” *Commun. Math. Phys.* **67**, 267–293.
- Glimm, J. and Jaffe, A. 1985 *Quantum Field Theory and Statistical Mechanics – Expositions*. (Birkhäuser).
- Haag, R. and Kastler, D. 1964 “An algebraic approach to quantum field theory.” *J. Math. Phys.* **5**, 848–861.
- Haag, R. and Schroer, B. 1962 “Postulates of quantum field theory.” *J. Math. Phys.* **3**, 248–256.

- Hepp, K. 1972 “Quantum theory of measurement and macroscopic observables.” *Helv. Phys. Acta* **45**, 237–248.
- Hiai, F. and Petz, D. 1991 “The proper formula for relative entropy and its asymptotics in quantum probability.” *Commun. Math. Phys.* **143**, 99–114.
- Lindblad, G. 1974 “Expectations and entropy inequalities for finite quantum systems.” *Commun. Math. Phys.* **39**, 111–119.
- Machida, S. and Namiki, M. 1980 “Theory of measurement in quantum mechanics.” I – *Prog. Theor. Phys.* **63**, 1457–1473, II – *Prog. Theor. Phys.* **63**, 1833–1847.
- Petz, D. preprint “Characterization of the relative entropy of states of matrix algebras.”
- Reed, M. and Simon, B. 1972 *Methods of Modern Mathematical Physics I: Functional Analysis*. (Academic).
- Roos, H. 1970 “Independence of local algebras in quantum field theory.” *Commun. Math. Phys.* **16**, 238–246.
- Sanov, I.N. 1957 “On the probability of large deviations of random variables.” *Mat. Sbornik* **42**, 11–44. Translation for the Institute of Mathematical Statistics published by the American Mathematical Society in *Selected Translations in Mathematical Statistics and Probability* **1**, 213–244 (1961).
- Seiler, E. 1982 *Gauge Theories as a Problem of Constructive Quantum Field Theory and Statistical Mechanics*. (Springer-Verlag).
- Shore, J.E. and Johnson, R.W. 1980 “Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy.” *IEEE Trans. Inform. Theory* **26**, 26–37.
- Shore, J.E. and Johnson, R.W. 1981 “Principles of cross-entropy minimization.” *IEEE Trans. Inform. Theory* **27**, 472–482.
- Strătilă, S. and Zsidó, L. 1979 *Lectures on von Neumann Algebras*. (Abacus).
- Streater, R.F. and Wightman, A.S. 1964 *PCT, Spin and Statistics, and All That*. (Benjamin).
- Takesaki, M. 1979 *Theory of Operator Algebras I*. (Springer-Verlag).
- Umegaki, H. 1962 “Conditional expectation in an operator algebra, IV (entropy and information).” *Kodai Math. Sem. Rep.* **14**, 59–85.
- Wheeler, J.A. and Zurek, W.H. 1983 *Quantum Theory and Measurement*. (Princeton).
- Whitten-Wolfe, B. and Emch, G.G. 1976 “A mechanical quantum measuring process.” *Helv. Phys. Acta* **49**, 45–55.
- Zurek, W.H. 1981 “Pointer basis of quantum apparatus: Into what mixture does the wave packet collapse?” *Phys. Rev. D* **24**, 1516–1525.
- Zurek, W.H. 1982 “Environment-induced superselection rules.” *Phys. Rev. D* **26**, 1862–1880.