 Mind & Matter Vol. 5(1), pp. 45–81

The Role of Quantum Physics in the Theory of Subjective Consciousness

Chris J.S. Clarke
School of Mathematics
University of Southampton, United Kingdom

Abstract

I argue that a dual-aspect theory of consciousness, associated with a particular class of quantum states, can provide a consistent account of consciousness. I illustrate this with the use of *coherent* states as this class. The proposal meets Chalmers' requirements of allowing a structural correspondence between consciousness and its physical correlate. It provides a means for consciousness to have an effect on the world (it is not an epiphenomenon, and can thus be selected by evolution) in a way that supplements and completes conventional physics, rather than interfering with it. I draw on the work of Hameroff and Penrose to explain the consistency of this proposal with decoherence, while adding details to this work. The proposal is open to extensive further research at both theoretical and experimental levels.

1. Preliminary Ideas

1.1 Introduction

The idea of linking quantum theory and consciousness has several attractions, but is also beset by deep-rooted difficulties. The aim of this paper is twofold. First, it is to describe and illustrate some of the basic principles involved in analyzing the possibility of such a link, concentrating on the perspective of theoretical physics. Second, it is to introduce a new quantum approach to consciousness, based on the consistent histories interpretation of quantum theory, which overcomes difficulties that have beset previous quantum approaches, illustrating this with a possible implementation of the idea of coherence.

I start with a brief review of both the attractions and the difficulties of quantum approaches, before outlining the basis of the alternative to be used here.

1.1.1 The Attractions

1. Consciousness, if used to fill a gap in the overall scheme of quantum theory, can be given a role where it actually does something (that is, is not merely epiphenomenal), without thereby coming into conflict with physical laws (Stapp 2005).

2. Many core aspects of experience seem to match a quantum universe better than a classical one (Clarke 2004, Stapp 2005).
3. The connectivity of widely separated systems in quantum theory seems to offer the possibility of extending into the scientific domain "anomalous" phenomena previously regarded as inherently unscientific (Clarke 2001).
4. In particular, the idea of a coherent state can unify consciousness theory and biology (Ho 1998).

1.1.2 The Difficulties

1. The traditional view of quantum theory as presented in older textbooks[1] separates the observer and the quantum system. When applied to the conscious brain this leads to the postulating of an autonomous consciousness observing the physical brain (a view implicit in the proposals of Stapp (2005), whose ideas have influenced the approach here). The nature of this consciousness, and of its interaction with the brain, remain largely unspecified, producing an account which is a "theory of consciousness" only in a very weak sense.
2. All the proposals for quantum consciousness probably require the maintenance in the brain of states that are coherent over distances of the order of centimeters. This seems impossible to achieve in a system such as the brain because of the overwhelming efficiency of the mechanism of decoherence, now widely accepted as a central aspect of quantum theory.

1.1.3 Overcoming the Difficulties

The following are the main contributions of this paper to overcoming the above difficulties.

1. I replace the traditional interpretation of quantum theory by the "histories interpretation" (Sect. 3.1), which has been developed in order to apply quantum theory to the universe as a whole, and is thus well suited to including the observer within a single physical account. While not entirely eliminating the need for consciousness as a distinct property in its own right, it places tight constraints on how consciousness operates and interacts with physical properties, thus representing a significant advance in relation to Sect. 1.1.2 (1) above.
2. Linking Penrose's approach (Sect. 2.1 (a) and App. *A*3) with some of the proposals of Stapp (2005) then gives a mechanism for maintaining coherence, which could (subject to more detailed calculations) overcome objection (2) in Sect. 1.1.2 (see Sects. 2.2, 2.3),

[1] For a more nuanced analysis of early work see Stapp (1993).

1.1.4 The Basic Standpoint Adopted Here

Here it is necessary to make choices as to what to retain and what to reject, in a situation where experimental evidence is at present too sparse to determine a unique approach, so that some of these choices are necessarily somewhat arbitrary.

1. I take as a philosophical starting point Chalmers' (1995) distinction between function and quality in defining consciousness, and also a pan-experientialist approach akin to Whitehead's (1929), which separates consciousness *per se* from the content and structure that consciousness takes in the case of humans (Sects. 1.5, 1.7).
   I must stress that I am here considering "consciousness" in the sense of subjective awareness, "what it is like to be ...", which Chalmers has distinguished from, say, a particular sort of integrating sequential information processing. I call the first *subjective consciousness* and the second *functional consciousness.* I start from accepting Chalmers' conclusion that subjective consciousness cannot be derived from functional consciousness since any particular function can be conceived as being performed "in the dark" (Velmans 2000), without there being anything that it is "like to be" that function. Here, moreover, "function" refers to functions that can be defined within a discourse relating to the objective world; subjective consciousness can still have a "function" in a different sense within the organization of the subjective world, and that function can then make its influence felt within the objective world via the complementarity of causation that I discuss in Sects. 1.4 and 3.2.
2. Second, the considerations of Sects. 1.1.2 (1) and 1.1.3 (1) indicate that it is necessary to construct a quantum theory and a theory of consciousness at the same time. It is not possible to take an "off the shelf" quantum theory and apply it to consciousness. Consequently, though this paper focuses on consciousness theory, it should be borne in mind that the argument needs to be judged for its success or failure in making a start on these two coupled tasks.
3. To clarify my approach, I should also make it clear that I am choosing to concentrate on *individual* consciousness. I take seriously the empirical evidence from unitive mystical experience suggesting that individual consciousness is an aspect of a universal consciousness, but I take the view that universal consciousness (if the term consciousness is not to become so generalized as to become a misnomer) is still *embodied* consciousness (Washburn 2003). Consequently the task I address here of understanding the relation between "my" consciousness and *this* body may still shed light on the more general relation between universal consciousness and the universe.

1.2 The Structure of This Paper

I will be developing an overall picture with several closely linked components – a minimal concept of philosophical dualism, a particular implementation of the histories interpretation of quantum theory, and a specific illustrative proposal for a state in the brain that embodies consciousness. In order to keep the overall picture in view while at the same time explaining its components, Sects. 1 to 3 will each touch on all these components at increasing levels of technical development. Section 1 surveys previous work and explains my stance on a variety of controversial issues. Section 2 focuses on the illustrative proposal of coherent states. Section 3 describes the overall picture with a focus on quantum theory. Cross-references will, I trust, enable the reader to track the different components of the theory through these sections. Finally, Sect. 4 develops a variety of implications of the ideas for future work.

1.3 The Nature of Quantum Physics

The terms "quantum mechanics" and "quantum theory" are used interchangeably in the literature, but here I will for convenience distinguish them to describe the two main levels of generality in quantum physics. By *quantum mechanics*, I mean the detailed level of the study of particular sorts of dynamical laws within a given quantum formalism (atomic theory, quantum electrodynamics, supersymmetric field theories, etc.). By *quantum theory*, on the other hand, I will mean a more general level, including the study of the various frameworks and formalisms that have been considered as foundations for quantum physics. This level thus comprises such things as the different formalisms in use (Schrödinger, path integral, Bohm, etc.), different interpretations of the quantum formalisms (Bohr, many worlds, histories, etc.), the general appraisal of theories that embody essentially different physics from the conventional formalisms (Ghirardi *et al.* 1986, Penrose 1994), and the study of the place of different theories of quantum mechanics *as a whole* within this general understanding. The focus of this paper will be on quantum *theory*, which I believe holds the key to the issues, and I will be largely taking for granted the conventional consensus for the main physical theories within quantum *mechanics*.

Approaches to quantum theory vary according to the way in which the observer is treated. While in the work of Bohr and most expositions before about 1960 the observer was, at least implicitly, a human being, in succeeding works the "observer" can be a measuring apparatus, or anything that makes a permanent record, as well as a human being. Until the advent of decoherence theory, however, only a few papers gave a clear definition of these terms or demonstrated that the definition did indeed result in observers having the properties that were required of them. Very often the terminology used ("record", "observer", "information", etc.) tended implicitly to assume a laboratory context, thereby begging the question of how widespread in the universe such things as observers are, and how they might emerge. Since a human being will certainly be an "observer", we are only interested in versions of quantum theory that address these issues. The area where we find accounts of quantum theory most suitable to our purpose is (perhaps surprisingly) quantum cosmology – because in the early universe there are no observers, record-keepers or pieces of apparatus, and so these things are not presupposed in that theory. The aim of quantum cosmology (and related systems) is to provide a consistently quantum-theoretical description of the entire universe, including "observers" within it, rather than assuming and drawing on the presence of an observer external to the (laboratory) system under investigation. This is usually achieved by the histories interpretation, see Sect. 3.1.

Quantum theory in its original form relied on an external observer to perform a "collapse of the quantum state", in which an open-ended superposition of possibilities changed to a definite actuality. Removing this external observer and incorporating it within an enlarged quantum system thus leaves a gap in the theory: a mechanism now must be produced within quantum theory to explain how the transition from possibility to actuality takes place. The main contender for that mechanism is *decoherence*: the process (whose existence is not contested) whereby the interaction of any system with its environment rapidly obliterates, as far as observable consequences are concerned, any distinctively quantum aspects, leaving a system that behaves classically (Giulini *et al.* 1996).

There are two lines of argument suggesting that this mechanism, while clearly of central importance, does not fully close the gap left by the incorporation of the observer. One is that of Penrose (2004), who points out the loophole that after interaction with the environment the system is only classical "for all practical purposes"; in fact, it remains a quantum system. The other is the argument of Dowker and Kent (1996), who show that, while decoherence makes the existence of the classical world *consistent* with quantum theory, it is not the only world that is consistent with it. The problem remains of why it is that in fact we live in a largely classical world (with the possible exception of some aspects of consciousness) rather than in a different, equally consistent, non-classical world. The solution to this proposed by Penrose is that the classical world is formed by the action of quantum gravity, in a way that is linked with consciousness – a proposal to which we return in Sect. 3.3.

1.4 Duality and Consciousness: the Notion of "Carrier"

I have already espoused part of Chalmers' approach (Chalmers 1995) which regards the two categories of subjectively experienced consciousness and the physical structure that is correlated with it as non-identical (a view also strongly argued for by Velmans 2000). Chalmers proposes that they are two aspects of a single entity – namely information – and regards the proposal as a dual-aspect theory, information being for him a neutral category possessing both physical and mental aspects. My concern here will likewise be with identifying the sort of physical structure that is particularly correlated with consciousness, and with the precise nature of this correlation; but I will not be concerned with whether there is a neutral entity that has these two aspects, or whether they are actually independent substances. Since my approach is based on physics, it is the physical aspect that will be in focus, and so my language will be that of property dualism in the sense of, say, Bennett (1984), although in other contexts I would argue for a more idealist stance. In particular, I will use the phrase "the support of consciousness" principally to designate the physical side of this duality. As such, it would be applicable to substance-dualist theories, although there the more traditional "seat of consciousness" would seem more natural.

The key question to be answered is: What is the relation between mental properties and the corresponding physical properties ? We can distinguish a bare non-causal correspondence (à la Leibniz) from an epiphenomenal approach due to which the brain influences (and may determine) the mind, but not *vice versa*, and this in turn from a (two-way) interactionist approach in which each influences the other. I will be arguing for a form of the last of these, in which there is a complementarity of causation: within each property, there are sequences of states which "just happen" with no discernible causation in terms of that property, but which are correlated with states in the other propert which are causally connected. Both sides are required for the fullest possible causal account, but the interaction is indirect, in that it is not phrased in terms of a causal interaction of one side to the other. It is remarkable that such an account is possible (there is *a priori* no reason to suppose that the causal gaps should be mutually complementary in this way). The fact that this is so in the present approach (pointed out by Stapp 2005) is one of the major positive features of the quantum approach.

1.5 The Carrier of Subjective Consciousness in Chalmers' Approach

Chalmers proposes that the carrier of consciousness could be *information*. This has the attraction that "information" seems capable of being expounded either mentally (as significance) or physically (as negative entropy). In this section I will argue that this is inadequate, thus opening up the introduction, in the next section, of a related, but more physical concept based on the quantum state.

Quantitative information theory (Eco 1976, Shannon and Weaver 1949) deals with signals, conveyed along a given channel, which are interpreted and acted on by a receiving station. So, prior to measuring information quantitatively, a definition is required of *what is counted as information*, and this is always relative to a particular intended interpretation of a physical state as "signal" and of "signal" as a source of action. For example, the information in a hand of playing cards is presumed to be relative to the standard classification of their faces; but a conjurer might be interest in their backs, and an art historian in the design of the symbols. These different contexts give rise to different definitions of information based on the same set of physical entities.

Let us apply this sort of consideration to a particular state of part of the brain. What here is to count as the "information" which carries consciousness? Presumably, one might answer, that aspect of the state which is in closest correspondence with the contents of consciousness. But then this opens the danger of a circular argument, in which consciousness is explained as an aspect of a special class of states which can only be designated by referring back to the content of consciousness. Such circularity is often covertly introduced: we become so accustomed to identifying and referring to brain states in the light of their discovered conscious correlates that we fail to recognize the circularity of singling out the corresponding sort of information as the carrier of consciousness. If we are to implement Chalmers' proposal that consciousness is an aspect of information, then for a non-circular theory the relevant sense of information must be identifiable purely physically, without prior implicit reference to the content of consciousness.

The most promising approach to making such an identification of information would be in the spirit of the "enactive" approach to consciousness of Varela *et al.* (1993). Here the information conveyed by a train of nerve impulses proceeding from a sense organ, for example, is relative to the sensibilities of the brain into which they are flowing, where by "sensibilities" is meant the repertoire of highly non-linear responses that can be exhibited. For example, the sound of the words "They're coming: run!" may carry barely milliwatts of power but can produce several hundred watts of activity, in a way dependent on the subtleties of the structure of the sound. In other words, information subsists in the relationship between the condition of one physical system (here, sound vibration) and the capacity of another system (the body) to respond to it in a very non-linear and structure-dependent manner. Information, in other words, might be the extent to which the state of a system is liable to engender a highly non-linear and structure-dependent response in another system.

I would argue that such a concept of information – and, to generalize, any concept of information that is physically definable without (self-defeating) reference to consciousness – is inadequate for Chalmers' dual-aspect approach. It is far too general, involving a totally unspecific notion of "system" which could be absolutely any part, aspect, or projection of the universe, however large or small. Moreover, it is almost certainly not what Chalmers was thinking of. He restricted the concept to those properties of our brain states that correspond to the input of which we are aware when we make some decision; which then involves us in the circularity, already noted, of identifying the carrier of consciousness to be those properties of which we are conscious. Invoking information does not help. We still have to decide what is the signaling system that carries the information, and this problem is essentially just a reformulation of the original question as to what is the carrier of consciousness.

1.6 Why the Carrier of Consciousness is Quantum Mechanical

What other candidates might be available? Chalmers' guiding principle here is that the physical properties of the carrier should serve as a partial *explanation* of the subjective properties, and therefore that there must be a structural similarity between the two. That is, we require that the physical structure of the carrier, by reflecting particular perceptual constructs, verbal forms and so on, should (partially) explain why these form the contents of my consciousness (though it is not required to explain what it is like to have this contents, this being the other, mental, aspect of the carrier).

Here "structural similarity" must necessarily be a weak concept because of the difference in kind between the two aspects: it is essentially a stable correlation between structures on the two aspects that is consistent with observations of brain states and reports of subjective states, together with a homomorphism of these structures that respects this correlation and preserves logical relations such as subsets and implications. This concept is weak because any such structural similarity will in practice be somewhat fuzzy and fluid, because of the plasticity of the brain and the distributed nature of its functions. It is at a much shallower level than the vexed issue of "intentionality" in consciousness research, but it is nonetheless sufficient to rule out many seemingly arbitrary identifications of the carrier that have appeared.

One natural requirement at this structural level is that the carrier of *my* subjective consciousness should be part of my physical body. This is a good start, though we shall see later (Sect. 4.2) that the notion of "physical body", and what is or is not a part of it, and similarly the notion of "me" and what is or is not a subjective part of it, may not be clear cut.

Marshall (1989), in his proposal for taking excitonic condensations as the carrier, helpfully stressed one further aspect of the carrier and one that elaborates structural similarity. First, the carrier must be identifiable, at the physical level, as a single entity (though no doubt made up of parts). This reflects the most important structural aspect of subjective consciousness, its (comparative) unity. It is a qualified unity since, as Blackmore (2003), Dennett (1991), Douglas-Klotz (2001) and others have stressed, the subjective experience of unity is on inspection rather more like a shifting focus among a collection of voices. So within the human organism we may be looking for a few, perhaps fuzzily defined entities, or a continuum that manifests a natural connectivity. We are seeking a possibly qualified physical unity to correspond to a possibly qualified mental unity.

Marshall's second criterion is that this unitary physical system should be modulated in some physically identifiable way by brain processes that are to correspond structurally to the contents of subjective consciousness – as implied by the principle of structural similarity. A structure that is unchanging is no use, and neither is a structure whose changes cannot be linked physically, in principle, with the formation of percepts. Of course, neither I nor Marshall are demanding a complete neurological account of consciousness before we can accept any proposed structure as the carrier, but we do require an indication of the sort of things that might be going on in the brain so as to qualify the structure as such.

It is these two criteria that in my view point towards replacing the ontology[2] of classical physics by one of the ontologies arising from quantum theory. Classical physics is founded on an ontology that is spatial and realist. The physical world is a "given" absolute reality, laid out in an absolute space (or, for relativity, *mutatis mutandis* in an absolute space-time). Its constituents are *things* and *fields*. Things are hierarchically included in each other within a set of levels (astronomy, macro-physics, molecular physics, ...) but at each level are atomistic. At any level of consideration this reality is expressed in terms of separate physical entities at definite places in space, each one of which may possibly be resolved into more fundamental entities at a lower level, equally distinct and separately located. My use of the word *things* is shorthand for these properties. In the case of the brain, we can physically identify such things as "neurons", "hippocampus", and so on. The other classical entities are *fields*, which pervade all space and admit no natural division into separate regions. The only unitary structures in classical physics are things and fields.

I claim that neither things nor fields adequately satisfy the criteria just introduced. For this, and for the next subsection, I need to elaborate on this concept of "things". On a formalist view of classical physics, it might be argued that there are no "things". Matter, of various compositions, is the only existent, and the division of regions of matter into tables, chairs, rivers, and so on depends on social conventions rather than on physics. We could, it is true, try to divide the world along surfaces of

[2]I use the notion of ontology not in the sense of the general study of being, but with respect to a system of given existents that is presupposed within some particular theory or form of discourse.

discontinuity in the density, but because few discontinuities are absolute it then becomes very fuzzy as to what counts and what does not: Is a cloud a "thing" ? Or the Gulf Stream ? In practice division into "things" arises from a mixture of variable aspects of physical structures and a system of conventions within the relevant social (including scientific) context. Consequently, "things" seem too convention-laden to be likely candidates for such a fundamental role as the carriers of consciousness. Moreover, the particular "things" (structures in the brain) identified within neuroscience seem unable to deliver the goods because the carrier of subjective consciousness, as well as the carriers of various particular functions of consciousness, appear to be highly distributed. Fields, on the other hand, pervade the whole of space and are not individuated into the consciousnesses of separate individual beings. They would, thus, seem a non-starter for the carriers of individual consciousness.

Quantum theory, by contrast, has a much richer basic ontology, which combines locality and non-locality. This will be explored in detail below, and will be shown to offer a fruitful domain in which to search for structures that are wider than neurons and organs, but are still localized in, or at least hooked onto, the brain. Quantum states, in particular, can be seen as such structures. In Sect. 3 this conclusion will be reinforced by the idea that the *dynamics* of a quantum approach (over and above its ontology) allows us to look for a form of consciousness that does something, and is not merely epiphenomenal.

In basing my proposal on a sort of quantum state I will to some extent retain a concept close to that of information, since a (mixed) quantum state (see App. *A*1) is defined as the specification of the probabilities for the results of measurements on the system, a concept which has similarities with the concept of information. I will avoid, however, the difficulty just described of specifying what counts as information both by restricting the sort of state under consideration (in Sect. 2) and by generalizing the notion of consciousness (in Sect. 1.7).

1.7 Pan-Experientialism

I have argued (Sect. 1.5) that, if a dual-aspect theory is to have any explanatory force, and is not to be circular, then carriers of consciousness need to be *physically* identifiable in terms of some basic inherent property. To expand on this: suppose, for a purely hypothetical example, that on the basis of detailed neuroscience it was decided that consciousness resided only in the hippocampus. While this might, let us suppose, be correct in terms of exploring which areas were most closely correlated with the reported contents of consciousness, it would leave untouched the question of what it is about the hippocampus, as opposed, say, to the amygdala, that might make the first conscious but not the second. The physical side of the dual-aspect approach would still not have been grasped. If we take into account that we are more likely to be looking for a distributed entity than for an anatomically defined location, then whatever physical property it is that is identified will almost certainly be applicable to a great many other entities in the universe than human brains. We will thus be led to some form of pan-psychism or pan-experientialism[3] in the sense that being a carrier of consciousness becomes not a particular evolutionary achievement of our own species, but an attribute that is widespread, even beyond what we normally regard as living beings.

Some would regard this as a fatal objection, holding that consciousness (in the sense of subjective consciousness as I have defined it) resides only in humans and our nearest cousins. This, however, confuses consciousness *in itself* with the content of consciousness or with its function. Probably, but trivially, the only organisms that think like people are organisms that are anatomically like people. This statement, however, concerns what subjective awareness in such organisms is an awareness *of*, rather than whether or not there is anything that could be postulated (though not experimentally verified) as being an experiential awareness in itself. The fact that my position leads to a form of pan-experientialism should thus not be regarded as an objection.

It will transpire that quantum theory, unlike (a formalist view of) classical theory, contains at least one division of the universe into "things" which does not depend explicitly on social convention (though it may depend on it implicitly insofar as quantum theory itself is a social construct). Indeed it seems likely that the division of the world into carriers of consciousness is also the best candidate for its division into "things", as opposed to what Whitehead called "aggregates" (of many things) – whereas entities that have an internal unity he calls "organisms". In that case we would have pan-experientialism in the sense that every "thing" is conscious.

2. Quantum States as Carriers of Consciousness

2.1 Overview of Coherence

In Sect. 1.6 I described how, on a classical approach, matter is either a scattered collection of elements ("things" – atoms, molecules, neurons, etc. – often conventionally defined), or else globally extended fields. In neither case was it possible to discern physical wholes to carry consciousness. At first glance, quantum theory seems to suffer from exactly the

[3]The term pan-experientialism is often applied to Whitehead's philosophy of process (Whitehead 1929, see also references in Deckers 2004), in which the basic entities are elementary durations of organisms which "prehend" (Whitehead's term) the preceding actions in the world. Other aspects of Whitehead's vision will appear later on in this paper. See also the defenses of pan-psychism by Skrbina (2005) and de Quincey (2002).

same problem as do fields, in that everything is so interconnected that we cannot discern any boundaries to demarcate one entity from another. The universe threatens to disappear into a single blob. This arises from three factors:

(i) The indefiniteness of quantum phenomena, which are usually described by quantum states represented as wave functions which diffuse over the whole of space;
(ii) The development of more and more unified quantum mechanical theories: individual particles and distinctions between particle types are assimilated into a single quantum field; and finally, space-time itself is assimilated into an abstract structure which no longer has any basis at all for separation into distinct entities;
(iii) The phenomenon of quantum entanglement in which, whenever two systems interact, they subsequently share a quantum state, so that states become more and more global in character.

This analysis, however, looks only at the part of quantum theory (including quantum mechanics) that is characteristically different from classical theory. Any comprehensive theory must also contain within it a description of the classical world: The quantum and the classical need to be two different regimes within the same theory.. The classical world separates, while the quantum world makes wholes. So it seems that the place in which to look for entities that are internal unities while still retaining distinctness from each other is on a length scale that defines the boundary between the quantum and the classical regimes. This boundary will give a *natural* definition of "things", overcoming the problem of conventionality described in the preceding subsection. In order to locate this boundary we need to understand how it is that quantum theory seems to operate on a small scale, while classical theory operates on a large scale. Of course, this demarcation between classical and quantum is the central problem of quantum theory as described in Sect. 1.3. Recent progress, however, has now shown how, for each of the above unity-generating factors, there are mechanisms causing them to break down at particular length scales, thus pointing to the boundary between quantum and classical. The following are the two most important of these.

(a) Penrose has argued in a number of works (Penrose 1989, 1994, 2004) that the standard evolution equation for the quantum state (Schrödinger's equation) breaks down at the point where the state becomes a superposition of two constituent states representing a massive object at two different positions in space. This breakdown both prevents the indefinite diffusion of the quantum state (point (i) above) and enables the structure of space-time itself to be handled in a way different from other physical fields (point (ii) above). The core of his argument (Penrose 2004) is the point that, when the gravitational field of the object itself is taken into account, the space-times determined by each of the two positions of the object are different, and so the idea of the superposition of different positions of an object in the *same* space-time becomes meaningless. The sort of new physics that is required to resolve this situation remains speculative, but the fact that this stage marks an incompatibility between the quantum and the classical is based on established physics. The time scale on which this effect operates depends on the gravitational energy of the difference in mass-density between the states involved. Strictly speaking, Penrose's argument indicates that a transition to a classical framework has to take place *at the latest* by the stage defined by this time scale. The overall picture presented by his work strongly supports, however, his claim that this stage does mark the emergence of a specifically gravitational mechanism that causes a transition to a classical form. (I continue this argument in Sect. 3.3 (ii).)

(b) While quantum entanglement for pure states is the general situation, for mixed states it is effective only so long as the phase relations between the components are preserved. These are extremely sensitive to perturbations by environmental influences, which for states that are separated in space act differently on the different states that are entangled. This destroys the phase relations between such states and hence destroys the entanglement – an effect called *decoherence*. Because it is caused by spatial variations in the environment, its strength increases with increasing length scale. Point (iii) above thus ceases to be effective at larger length scales. (I continue this argument in Sect. 4.1.)
The time scale on which this effect operates depends on the differences in the energy conveyed to the various entangled states by fluctuations in the environment.

In this section I shall argue that the mechanisms (a) and (b) above allow the existence of quantum states that have both unity and individuality, qualifying them as carriers of consciousness. This importantly links with the work of Ho (1998) who has presented a number of lines of work indicating the role of *coherence* in living organisms, an idea which we will in turn link with Whitehead's terminology of "organism" for entities which have an internal unity (superseding our previous unsatisfactory terminology of "things").

We can recall here that the idea of coherence was originally a classical concept derived from phenomena such as Young's slits: the occurrence of interference bands in light which had passed through two closely separated parallel slits. This happens when the light has first passed through a single slit, which results in a beam such that, for any two points in space, the phases of the vibrations of the electric field at the two points differ by an angle that is fixed in time. Light with this property is called coherent. When, later, all matter was shown to have a wave aspect in quantum theory, the concept was extended to quantum coherence as a generally applicable condition. Decoherence as described in (b) above acts by introducing random fluctuations in the relative phases of the quantum states of systems separated in space, analogous to a lack of the coherence in light as just described. As will be described below, the situation is complicated by a range of possible definitions of this intuitive concept. The proposal of this paper can, however, be expressed roughly as follows: *Coherent states are the essence of organisms; all organisms have consciousness which is carried by such states.*

The various different sorts of coherence described by Ho (1998) can be seen as derivative from this quantum focus. The alignment of molecules which registers in micrographs using polarized light (the core experimental technique of her research) can be seen as the structure which supports, and is in the first instance controlled by, the coherent quantum state, and the coordinated thermodynamic processes that she also cites become other aspects of this global network of organization of which the coherent state is the coordinator.

2.2 Coherence and Entanglement

The concept of coherence has been most studied in the theory of lasers (quantum optics). The sort of system relevant to the brain, however, is more likely to be analogous to Marshall's model (Marshall 1989), based on the ideas of Fröhlich (1968) in which the structure consists of an array of molecules (either lipids in cell membranes, or electrically polarized proteins embedded in those membranes) regarded as charged mechanical oscillators each with a single relevant vibrational mode, coupled with each other through electrostatic forces and supporting a field of phonons. This is a particular case of the very general model of *coherent structures* (Dixon *et al.* 1997). Locally this model applies also to the microtubule model of Hameroff and Penrose (1996). Thus typically we will be considering a coherent structure that is embedded in a background non-coherent matrix.

Coherence in optics is about consistency of the relative quantum phases of parts of the system that are separated in space and time. The trouble is that this statement does not make sense, because "quantum phase" has no invariant meaning (it depends on the choice of how we represent the quantized system). One possible starting point would be the concept of first-order quantum coherence that is well established in optics (Loudon 1983). Since, however, we will most often be concerned with molecular coherent structures rather than purely electromagnetic ones, I will approach the concept via the more general idea of entanglement considered in Sect. 2.1, (iii) and (b).

I should stress here that, while I consider that strong arguments can be made for the general principles described so far, the particular definition of coherence I am about to introduce will certainly be subject to revision in its details. I believe, however, that the essential issues arising – such as the stability of the particular sort of quantum state in the presence of environmental fluctuations – will not depend greatly on the particular definitions adopted; whereas by considering a particular definition the arguments can be brought out in a more concrete form.

In App. *A*2 I describe a definition of coherence based on entanglement and use a characterization of entanglement in terms of the existence of correlations between the responses to observation of spatially separated parts of the state. (See also the brief summary of entanglement in Sect. 4.1.) This seems to capture Ho's conception of the coherence of an organism as a condition where there is non-local coordination between its active responses. The optics definition, by contrast, does not necessarily imply any sort of non-local connection, as can be seen by considering the laboratory case of interference arising from two separate but precisely tuned lasers (a case that is admittedly theoretical, because of technical difficulty). I define in App. *A*2 a quantity C, the degree of coherent entanglement, depending on the state being considered. C is zero if the state is made up of separate parts having no entanglement between them, and it is one if all parts of the system are maximally entangled with each other. This separates organisms ($C \approx 1$) from aggregates ($C \approx 0$). All approaches of this kind necessarily produce a gradation between the two categories rather than an absolute demarcation, so that consciousness is not suddenly and completely switched on. Since, however, the emergence of organismic properties seems likely to lead to processes that increase the internal coherence, it is unlikely that there will be many cases of systems lying half-way between the two limiting cases.

We can note here that the need for the state to be modulated by its containing matrix so as to reflect the content of consciousness is easily accommodated in this picture. If the containing matrix is slowly modified, that will not in itself affect the entanglement of the parts of the state that it supports.

Any state of a system will define a state (in general mixed) on any of its subsystems. It could therefore transpire that a state of a subsystem was coherent by virtue of its being part of a larger coherent state. Since I am identifying qualified unity and individuality as hallmarks of consciousness, we would not wish all substates of a state to be consciousnesses in their own right, and we therefore exclude this by demanding that the coherent state is *maximal* in the sense of not being a restriction of a larger coherent state. To summarize, my hypothesis is that *carriers of consciousness are* *maximal coherently entangled states, and all such states are carriers of consciousness.*

As with all quantum effects, the crucial question becomes the analysis of how decoherence removes coherence in this sense (noting that coherence and decoherence are separate phenomena and not opposites). I will return to this question and survey mechanisms for maintaining coherence in Sect. 3.3.

2.3 State and Context

To clarify the proposals just made, I conclude this section with a few remarks about the different levels at which a state may be defined. By "state" I mean the quantum state – that is, the totality of information necessary to supply probabilities for all possible quantum measurements in a given situation. By "levels" I refer to the hierarchical structure of physics, within a hierarchical structure of the sciences as a whole, whereby quantum fields give rise to atomic physics, atomic physics to solid-state (or liquid or gaseous) physics, and to chemistry, these in turn to microbiology and biochemistry, and so on. If we just look at the quantum-physics sections of this hierarchy, it is clear that a different assignment of what is meant by a quantum state applies at each level, with states at one level being defined in terms of the states at a lower level by a process analogous to coarse-graining, provided there are no correlations between the large and small scales.[4] If, then, we have a coherent state at, say, the level of phonons in a molecular array, the state at the lower level of individual atoms may not be coherent, because of phase factors at this level which are eliminated at the higher level. Coherence, and hence consciousness, is not a property just of the matter in a given region, but also of the level at which this matter is regarded. This is implicit whenever we talk of a physical "system": this is not the total matter, but the matter viewed at a particular level. So it is not the brain, or a part of the brain, as a whole that may be conscious, but a subsystem of the brain considered at a particular level of description.

We are, however, seeking an objective definition of the carrier of consciousness, and therefore we need an objective definition of the level at which matter is viewed. If this definition were left to the decision of the

[4]Coarse-graining of a state is its replacement by another state in which finer structure is ignored. In the context considered here the Hilbert space can be represented as a tensor product $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ where $\mathcal{H}_1$ and $\mathcal{H}_2$ represent, respectively, the large-scale structure and the small-scale refinements to this. Coarse-graining, in the strict sense, of a state in $\mathcal{H}$ is carried out by "tracing out" the component in $\mathcal{H}_2$, resulting in a *mixed* state. If, however, the two components are unentangled (i.e. uncorrelated, see App. *A*2), then the state has the form $v \otimes w$, where the component v represents the large-scale structure. If the large-scale and small-scale structures are correlated at the quantum level then – as would be expected – the large-scale structure cannot be represented as a pure state.

human experimenter, then we would be guilty of the sort of circularity of argument already criticized in Sect. 1.5. The resolution of this depends on the structure of quantum theory as a whole, as it will be developed in the next section. In this there is an interplay of the "top-down" influence of consciousness on the physical structure (i.e. causal interactions or correlations between the larger scale and the smaller scale) and the "bottom-up" influence of the small-scale structure on the large-scale structure, only the latter being recognized in classical science. The top-down influence restricts the algebra of propositions (cf. Sect. 3.2) and so provides a *context* that defines the level at which a physical system is to be defined. The dynamics of this system then determines the subsequent development of the carrier of consciousness at the higher level. Because the influences in the two directions are structurally different and non-interfering (the first is a selection of logic and the second is a free causal dynamics within that context) this can be a consistent scheme – which would not be the case if the two influences were causal in the classical mechanical sense.

3. A Quantum Setting for the Interaction of Consciousness with Matter

3.1 The Histories Interpretation of Quantum Theory

In order to show that this proposal is viable I need now to couple it with a version of quantum theory that does not require an appeal to an external observer, for the reasons discussed in Sects. 1.3 and 1.4. There I noted that the most suitable theories for an account of subjective consciousness are those which derive from a cosmological perspective. Out of these I will use a particular form of Hartle's (1991) "generalized quantum theory" on the grounds that it is as free as possible from implicit metaphysical assumptions. I will call my variant of generalized quantum theory the "generalized history formulation". The ideas I will present seem most natural in this version, but could probably equally well be expressed in other forms of cosmological quantum theory.

The (standard) histories interpretation can be regarded as a means of generalizing the language of laboratory physics to any wider context, up to and including the entire universe. Considering first the laboratory case, we note that an experiment involves the *preparation* of a particular system and observing apparatus, so as to control the relevant variables as far as possible, followed by the *observation* of a phenomenon that follows the preparation. The aim of the physical theory involved (i.e. quantum mechanics as distinct from quantum theory) is either to predict, unconditionally, the observed outcome, or in the case where there is a random or indeterminate element, to assign probabilities to the various possible outcomes. But there is no need to lay stress on the distinction between preparation and observation; what the theory is doing is just assigning probabilities to particular successions of physical events, whether designated "preparations" or "outcomes". We can also take into account that, whereas the observation is a physical event that usually occupies a brief interval of time, the preparation could involve a longer process and, thus, be more relevantly represented as a sequence of events. Thus, an experiment assigns a conditional probability to an event (the observation), given the occurrence of a sequence of prior events (the preparation). From here it is a short step to the conclusion that what a physical theory is, essentially, required to do is to assign probabilities to sequences of events. This is the basic philosophy of the histories interpretation, in which a sequence of events is called a *history.* Expressed in that generality, the theory will be applicable to contexts wider than that of the laboratory.

We can now indicate how to formalize this (though with no attempt at either physical or philosophical precision at this stage). In the simplest form of this approach, a *history* is a time-sequence of assertions about the world. A laboratory example might consider the sequence:

L1. An electron is emitted from a particular production device.
L2. It then travels through a magnetic field of a given intensity.
L3. It strikes a particle counter at a given position.

A cosmological example might be:

C1. There is a cloud of gas of some particular composition.
C2. It forms galaxies.
C3. These produce stars.
C4. At least one of the stars produces planets.

We would typically be interested in the probability of the sequence {L1, L2, L3 } *given* the sequence {L1, L2}, or of {C1, ..., C4} *given* {C1, ..., C3}. If this probability is 1/2 for the first example, we expect a 50:50 chance that electrons produced in this way will be deflected so as to miss the detector. In the second example we would then think that there was a 50:50 chance of the universe containing planets, and so the appearance of the planet earth was not that remarkable. "In principle" (words that often cover a multitude of fundamental problems), if we could work out absolute probabilities for all the sequences involved, then the conditional probabilities would be simply obtained by division. In practice the process would be the reverse: We would in each case use a particular mathematical model to calculate the conditional probability directly, and from this infer the existence of absolute probabilities.

The theoretical framework of the histories interpretation idealizes the physical theory as a means of specifying the probability of any history, given only an initial state (i.e. a quantum state; cf. Sect. 2.3) for the system in question (universe or laboratory) at some basic "zero-time" (e.g., the time of the separation of the fundamental forces in cosmology). The specification is given explicitly in Sect. 3.2 and in App. $A3$. The initial state could be of a very general nature, chosen deliberately so as to give minimal initial information. Thereafter, there is, strictly speaking, no need to refer to states at all except as a convenience for organizing calculations. However, because of the role being assigned to states here we shall freely refer to them, while bearing in mind that at each stage in a history the state is completely specified by the initial state and the past history up to that stage.

Hartle's "generalized quantum theory" is a generalization of this in which we also allow assertions about simultaneous events at different places. Thus a *generalized* history might include sequences like "... an atomic bomb is detonated on earth, and meanwhile in the Andromeda galaxy ...".

In order to combine this scheme with the concepts so far introduced, we will apply it by choosing as histories sequences of experiences of organisms (as defined in Sect. 2) – for example, sequences of classes of coherently entangled states.[5] If we idealize what is really a continuum of experiences, flowing through time, as a discrete sequence, then this sequence can itself be seen as a history, provided that we replace the notion of an "assertion about the universe" by simply "an experience". Then, if we aggregate the histories of all the organisms in the universe, we obtain a generalized history. The dynamics of the model (quantum mechanics) are required to specify probabilities for experiences.

Hartle's approach gives rise to a natural way of breaking the continuum of experiences up into discrete elements. Each "unit" of experience occupies a region of spacetime which I call a *locus*. I describe their structure in App. $A4$ including a temporal restriction on the *loci* ensuring that they are, in a sense, either unambiguously simultaneous or unambiguously time-ordered. The spatial extent of a *locus* is the extent of the support of the coherent state, and its temporal extent is essentially given by the time taken for light to cross this support (about 1 nanosecond for the brain).

I must emphasize here three points:

(i) The discretization of *loci* is performed as a matter of mathematical convenience. While there are theories of continuous measurement that could in principle be used instead (Clarke 1973, Sudbery 1984) they would at the moment introduce considerable complications without any benefit.

[5] A history, on this approach, consists of classes of these states rather than individual states, because the history consists of a sequence of verified propositions, and in general the propositions may be tested for a whole range of possible coherent states rather than a single precise one. The "fineness" of the propositions will be restricted by Penrose's gravitational criterion as described in Sect. 3.3, (ii).

(ii) Because the discrete *loci* are only a convenience, they should not be identified with the prehending events that in Whitehead's (1929) theory provide a real pulsational structure to time – though the present theory has many other points in common with Whitehead's theory (Clarke 1993).

(iii) The time scale determined by the *loci* is expected also to be relevant to a more exact continuous version, being linked to the quantum mechanical and relativistic processes that underlie experience. It is, however, *not* linked to the natural time scale of subjective experience, the "specious present", which is far longer.

In the next subsection I will show how, by linking this particular formulation of quantum theory to consciousness, a central gap in quantum theory can be closed. At the same time, a specific quantitative mechanism is introduced whereby consciousness "does something" and is not merely epiphenomenal. It is this simultaneous resolution of problems in consciousness studies and quantum theory that I consider the characteristic strength of this approach.

3.2 Quantum Logic and Conscious Interaction

In generalized quantum theory each *locus* is associated with a possible event that makes up the history, and an *event* is defined as the "satisfaction" of a proposition, with values TRUE or FALSE. In the basic formalism of quantum theory, a proposition is represented as a projection in the Hilbert space of states associated with the *locus*, regarded as being an observable with the possible values 1 (TRUE) or 0 (FALSE).

I can now describe the formalism for deriving the probabilities for different generalized histories. Explicitly, given a finite collection of propositions P with corresponding *loci* U as described above, then we order them as a sequence $P_1, P_2, P_3, ...$ with *loci* $U_1, U_2, U_3, ...$ and so on, satisfying the conditions of App. *A*4. The probability of this generalized history is then

$$p\,(P_1, P_2, P_3, ...; U_1, U_2, U_3, ...) = \mathrm{Tr}\,(...P_3' P_2' P_1' \;\rho\; P_1' P_2' P_3'...) \ ,$$

where ρ is the (mixed) initial state, Tr denotes the trace of the following operator and P_i' denotes the result of propagating P_i back to the initial time (see App. *A*3 for details). The stipulations imposed ensure that p is then independent of the particular ordering chosen, provided it satisfies the time ordering requirement just stated.

We now come to a core problem of quantum theory, which expresses itself differently in different formalisms. In the present formalism it amounts to the fact that the theory is incomplete in the sense of requiring a further restriction. This is because the numbers given by the values of p, which I have referred to as probabilities, are in fact no such thing: If we allow choices of all possible propositions subject to restrictions already imposed, then they do not satisfy the axioms of probability theory. To take the simplest case, if we consider histories consisting of a single proposition at a fixed *locus* U, and if a collection of propositions $P_1, P_2, P_3, ...$ are pairwise disjoint ($P_i P_j = 0$ whenever i and j are distinct) then in general, setting $P = P_1 \cup P_2 \cup P_3 \cup ...$ for the projection on the union of the image subspaces, it turns out that

$$p\,(P; U_1) < p\,(P_1; U) + p\,(P_2; U) + p\,(P_3; U) + ...$$

whereas for genuine probabilities the two sides of this relation should be equal. This is essentially because propositions satisfy a quantum logic, not a classical logic (Beltrametti 1981) – an aspect that Isham (1994) extends to general histories. In order for the numbers p to be interpretable as probabilities we need to place restrictions on the possible propositions so that (i) the collection of propositions at each *locus* constitutes a classical logic and (ii) the collection of all the resulting histories constitutes a classical logic. Collections satisfying a slightly stronger form of (ii) are called "consistent" in the literature.

The major achievement of work in this area over the last 20 years or so has been an understanding of how this issue within logic is related to the difference between the classical world and the quantum world. If we demand that our propositions are "ordinary" macroscopic propositions, and *not* propositions concerning such things as superpositions of live and dead cats, then the conditions (i) and (ii) above are satisfied because of the phenomenon of decoherence. Crucially, however, the converse is not true: there are non-ordinary "worlds" (collections of propositions) that produce genuine probabilities. The theory is *consistent* with the occurrence of a classical world, but it does not *explain* it (Dowker and Kent 1996).

There is, thus, a *lacuna* in quantum theory, which is present in different forms in all its interpretations, though it is most clearly seen in the histories approach. A criterion is needed for a process that selects, at each *locus* in the generalized histories approach, one particular logic of propositions.[6] Such a selection is independent of, and unconstrained by, both the dynamics of the quantum mechanical theory being used and the subsequent process (often referred to as "collapse") of the manifestation of one particular history within the probabilistic theory set by the basic histories interpretation. I will refer to a whole sequence of such logics, with their *loci*, as a *framework* for the corresponding histories.

A core proposal of this paper is that *the selection of a logic at each locus is the result of the conscious (mental) aspect of the state.*[7] This

[6] In what follows I use the term "logic of propositions", or just "logic", to refer to a Boolean sub-algebra of the full algebra of projections with support over a given *locus.*

[7] For future reference, note that the selection of a logic is, in conventional quantum

proposal arises naturally from consideration elsewhere (e.g. Clarke 2001, 2002b) of the nature of free will, where I have argued that the subjective experience so described is distinct from action caused by an acquired disposition, from spontaneous uncaused or arbitrary action, and from any mixture of the two. When introspectively examined, it appears as an action of reframing the overall meaning of a situation, which would seem to be the mental correlate of the use of one out of many alternative logics of propositions at a *locus*. This is the only action of consciousness with implications in the physical world. The impression we sometimes have that spontaneous choices from pre-defined alternatives represent an influence of consciousness on the world is not supported either by careful introspection or by neurological work such as that of Libet (2004).

It is appropriate to discuss further how this proposal fits with the idea of complementary causation raised in Sect. 1.1.4 (1) and in Sect. 1.4, and with general issues of discriminating different types of dualism, which on the whole I am avoiding here. If there is a specific process of the "reframing of meaning" whose outcome correlates with the selection of a logic of propositions, where is this process to be situated ? Do we situate it on the mental side, in which case on the physical side the selection just happens, without physical causation; or do we situate it on the physical side, as a physical cause of the selection, in which case our conscious perception of the process is epiphenomenal ? The fact that, at present, we only have a mental-language account of the process inclines us to the former, but this could change with further research. I would suggest that a more stable basis for deciding might be given by the arguments of Penrose (1989, 1984) suggesting that consciousness is responsible for non-algorithmic actions, though I do not consider these arguments to be at present conclusive. If we take the view that on the physical side any causal account must be algorithmic, in Penrose's sense, and that at least some creative mental processes are not, then this would give grounds for placing such processes purely on the mental side. This is the view to which I think the current evidence strongly points, and which I therefore adopt here. I should point out, however, that the view presented here differs from that of Penrose in that he places the action of consciousness at the point of the emergence of a definite outcome from a set of quantum possibilities already established, whereas I place it earlier, in the determination of this range of possibilities (the logic) in the first place (see Sect. 3.3, (ii)).

To summarize and conclude this section, therefore, I note that, as already described in Sect. 1.4, my thesis is a histories version of the thesis that has been developed by Stapp (2005). In common with that of Stapp, my thesis

---

theory, equivalent to the specification of the measurement that is about to be performed in a given experimental context – but not to the outcome of that measurement. To that extent, consciousness has connections with the idea of measurement.

(i) is consistent with physical causality as we currently understand it;
(ii) includes consciousness in our world view in a way that is interactional and not merely epiphenomenal.

It goes further in that

(iii) it gives a physical specification of the frequency of observation (App. *A*3), shown by Stapp to have experimental psychological significance;
(iv) it fills a gap in existing interpretations of quantum theory by specifying which *loci* enter into the histories interpretation (namely, the *loci* of organisms);
(v) it provides a proposal for consciousness studies specifying which entities are conscious.

3.3 Mechanisms for Establishing and Maintaining Coherence

Any proposal for linking quantum theory and consciousness must address the role of decoherence in annulling quantum effects in the brain. The background to this is the work of Hameroff and collaborators (Hagan *et al.* 2002) who calculate time scales of between 10^{-5} and 10^{-4} seconds for the removal of quantum effects in the microtubule system and elsewhere. The proposal of this paper, that the carrier of subjective consciousness is a quantum state, places very strong demands in this respect. For consciousness to include the breadth of content that it does this state must extend over macroscopic regions of the brain (centimeters), and to match the subjective introspective data it must be maintained continuously in time. For the particular quantum solution proposed, this means addressing the effect of the environment in removing entanglement (Sect. 2.2) from such macroscopically extended states. I will outline here the three main mechanisms which may be able to do this, each of which will require considerable research in order to flesh out the details. An essential part of the argument is that the maintenance of coherent entanglement will require the combined effects of these mechanisms.

(i) *Pumping.* This refers to a process in which metabolic energy is continuously supplied to the physical system in question, which is set up in such a way that the energy flows selectively into coherent states, and is then dissipated into a surrounding thermal sink. An example of this is the Fröhlich mechanism for maintaining a Bose condensation in the ground state of a system of oscillators (Fröhlich 1968). For current purposes, however, it has the weakness (Clarke 1994) that, in the systems of interest, probably only a quasi-condensation will be produced so that, on the face of it, there seems no reason to suppose that this will make the system robust against the removal of internal entanglement by decoherence. It remains to be seen whether the generalization of this to coherent structures (Dixon *et al.* 1997) might offer ways of extending the decoherence times of (Hagan *et al.* 2002) to larger systems than microtubules.

(ii) *Gravitationally induced state-collapse* (see Sect. 2.1(a)). Penrose phrases this in terms of a different interpretation of quantum theory in which the quantum state is given a real status and suffers an objective collapse as a result of gravitational effects, on a time scale of $t_G = h/E_G$ where E_G is the gravitational energy of the difference in mass distribution between states making up the superposition. Hameroff and Penrose (1989) link this mechanism with consciousness, holding that a conscious decision occurs at the point of gravitational collapse (and the decision determines the outcome), with previous states being pre-conscious. The problem with this is that our normal process of self-awareness clearly includes superpositions of thoughts prior to their resolution by a decision, so that the time scale t_G appears to be much shorter than the time scale of conscious decision with which it is identified. On the other hand, for the general relativistic reasons indicated by Penrose (2004), some mechanism of the form that he describes seems indispensable. It thus seems necessary to add Penrose's mechanism to the scheme already outlined, while restricting it to organisms, since on the current approach it is only these that exhibit this mechanism. To implement this within the histories interpretation I propose the concept of selection of a logic already discussed to provide a framework for decomposing states into superpositions (a gap in Penrose's current formulation of the idea) and then introduce into the theory an additional dynamic in which the fineness of the decomposition into different propositions is determined by the time elapsed since preceding measurements in relation to the time scale t_G (see App. *A*3). The non-algorithmic element in Penrose's scheme suggests that it could go further than this and actively bias the probabilities in favor of particular outcomes.

Adopting this implementation of the idea allows us to overcome the above objection based on the time scale t_G since, in the system outlined in App. *A*3, consciousness is exercized at all time scales in the organism, and the Penrose mechanism is no longer linked specifically to human decision-making.

(iii) *Zeno effect.* This is the phenomenon that a frequently repeated measurement can increase the persistence of a state that would otherwise dissipate. It has been invoked by Stapp (2005) in order to effect a link between conscious attention and the maintenance of quantum states. In Stapp's work the Zeno effect arises from the operation of consciousness with its own independent dynamics and rate of observation (the Zeno time scale). When we use the implementation of the Penrose effect suggested here through generalized histories, however, then the "observation" of Stapp's form of dualism becomes replaced by the inclusion of a *locus* (see Sect. 1.5) in the structure of histories. As shown in App. $A3$, the relevant time scale here is set by the temporal closeness of *loci*, and so the time scale of the Zeno effect becomes that of the Penrose time.[8]

Combining the above mechanisms can be very effective. For instance, if we consider a Zeno effect with a time scale of t_Z accompanied by decoherence on a longer time scale of t_D then it is easily seen that the combination extends the effective time scale of decoherence to $t'_D = t^2_D/t_Z$. When the Zeno time scale is set by the Penrose effect, the balance between these factors becomes favorable at larger length scales, because the Penrose effect increases in strength with the square of the mass scale (i.e. the sixth power of the length scale) whereas the decoherence effect increases with the square root of the mass scale (for, e.g., perturbations by low-energy photons) or at worst linearly with the mass scale for coupling with other internal modes not participating in the coherent state. As Hameroff and Penrose (1996) have discussed, the critical point is to find a sub-unit of the coherent state in which the decoherence time can be raised, without any Zeno effect, to the Penrose time. Thereafter, coupling these units together and applying the Zeno effect can produce large-scale coherent states.

4. Applications and Developments

4.1 Quantum Entanglement and the Extension of Consciousness

The ideas presented so far can, I believe, open up a rich area of research into the neurophysiology of consciousness. But it is also possible to develop them so as to engage with a most fundamental aspect of consciousness, namely our awareness of external events. In the past, much emphasis has been placed on the idea that quantum entanglement could play a role in this, and in my own work (Clarke 2002b) I have speculated that this could provide a mechanism for the association of qualia with percepts. I can now develop this using the postulated role of entanglement in coherence. We need to bear in mind, however, that entanglement is being used to define the extent of consciousness and *not* to communicate information. Indeed there are radical limits on the accessibility of

[8]This is essential: were it not for this link with the Penrose effect, the natural time scale of the Zeno effect would be the light-crossing time of the maximal coherent state, which would result in the possibility of the Zeno effect disrupting the behavior of microscopic systems in contradiction to observed physics

information, which essentially mean that entanglement could not be used to carry information instantaneously (Holevo 1973).

One of the central mysteries of consciousness is the fact that, while its mechanism appears to be in the brain, the perceptions of objects are located in a subjective space extending outside the (subjective) body (Velmans 2000). An attractive solution to the problem of how percepts are, as it were, "projected" into an external space might be available if our consciousness itself extended beyond our body (Sheldrake 2003) in the sense that the support of the carrier of consciousness could be so extended, at least while perception is taking place.

For this to be possible, the organism must first bring part of the remote system into coherent entanglement with its conscious state, and then maintain this coherence. Maintenance is the less problematic of these two: once consciousness is extended, then the Zeno effect could in principle maintain coherence, though the time scale will be extended to microseconds or milliseconds. For the initial extension of the carrier, it becomes necessary to speculate on the way in which the content of consciousness might affect the development of the logic of propositions used in the histories interpretation. The flexible nature of the brain makes it essential that it should be possible to extend the carrier when a possible larger coherent system becomes available, and for the application in mind here we might suppose that the occurrence of a physical perception might, through a selection of logic, bias such an extension in the direction of including an external system. If the probability of propositions including this system were non-zero (as could be the case in the presence of a physical interaction through the senses) then the required extension could take place.

We should note that the increased time scale of interactions with an external system impose a limitation on distance for this process. Looking at a starry sky, for example, immediately raises a problem that is insurmountable at the level being considered here. If we add to this the evidence from the many types of illusions in perception, it seems likely that external subjective space and the most elementary qualia of perception are established in the earliest stages of childhood by a direct mechanism of consciousness; but that subsequently (perhaps at the time of the "closing of the ground" discussed by Washburn 2003) these become subordinated to more mechanical types of neural processing. In the case of sensing another organism, however, which already had within itself a conscious coherent state, the joint effect of the two systems could extend the distance over which direct perception is possible.

4.2 Future Prospects

I have described a dual-aspect approach to consciousness which I hope will be found philosophically and physically persuasive by some others. Of necessity, given the underdeveloped nature of this area, I have made many arbitrary assumptions in order to draw out the sort of consequences that might be expected of such an approach. These all must, and can, now be tested, and where possible be modified or replaced, by further theoretical and experimental investigation. To this end, I will here outline some of the many questions that might now be addressed.

(i) What particular biological models are available for coherent states of sufficiently large extent, using the mechanisms outlined here? As I have described above, the most likely candidates seem to be generalizations of the Fröhlich model and the Hameroff-Penrose model with a specified physical (e.g. electromagnetic) mechanism for the orchestration of coherence, and hence of apparent collapse.

(ii) Are there other classes of states that might be as convincing as coherent states as candidates for the carrier of subjective consciousness? Is the particular definition of coherence suggested here the most appropriate one?

(iii) Are there suitable *in vitro* systems that might serve as test systems for the basic physics of these ideas?

(iv) Are there areas of anomalies research which can be given an improved focus and methodology by linking with the proposed theoretical framework?

(v) Can the concept of coherence explored here link together the different aspects of coherence in organisms (thermodynamic, quantum and molecular alignment) involved in Ho's (1998) work?

(vi) Can research be established to link the systematic subjective investigation of states of consciousness (plurality of strands, coexisting contradictory percepts, ground/content distinction) with features of a quantum model?

Having set out the principles and arguments in some detail, I hope that more work can follow to develop or to disprove the approach presented here.

Acknowledgment

I am most grateful to the two referees for *Mind and Matter* for many improvements to this paper.

Appendices

A1. Pure and Mixed States

A basic distinction in quantum theory is between pure and mixed states. It rests on a rather fuzzy distinction between classical uncertainty and quantum uncertainty. The first is due to human ignorance, it is governed by the rules of conventional logic and probability theory, and it is expressed through a real number between 0 and 1. The second is an expression (according to most commentators) of an uncertainty intrinsic to nature. It is governed by a distinctively different logic, and it requires more data to specify it, namely a complex number. A *pure state* is then one which contains only quantum uncertainty (and is expressed through a vector in Hilbert space), while a *mixed state* combines quantum and classical uncertainty and is expressed through a unit-trace Hermitian operator ρ in Hilbert space (often called a "density matrix"). For any observable A the expectation value of A in the state ρ is given by $\mathrm{Tr}(A\rho)$ where "Tr" denotes the trace. Thus, when A is a projection corresponding to a proposition P, this formula gives the probability for this proposition being true.

Any mixed state can be written in diagonal form as $\rho = \sum_{i=1}^{\infty} p_i |i\rangle\langle i|$, in which the numbers p_i can be interpreted as classical probabilities for the corresponding (pure) quantum states $|i\rangle$. This mathematical fact can, however, be misleading, since the expansion becomes non-unique when some of the p's are equal and the states involved may not be practically observable. If we instead consider the matrix of ρ in a practically convenient basis of measurable states $|k\rangle$, then (using the trace properties above) the diagonal elements $\langle k|\rho|k\rangle$ represent the probabilities for obtaining the different possible values for a measurement having this basis as its eigenvectors, while the off-diagonal elements $\langle k|\rho|k'\rangle$ are complex numbers representing the phase interference effects that distinguish this state from a classical probability ensemble (in which the off-diagonal elements are zero).

The two sorts of uncertainty, classical and quantum, are often linked with ontological distinctions. In the case of classical uncertainty, it is supposed that there is an underlying state of affairs, of which we happen to be partially ignorant at a given time. In the case of quantum uncertainty it is known (as a result of the Kochen-Specker theorem; Kochen and Specker 1967) that it is impossible within conventional logic for there to be an underlying state of affairs.

A2. Entanglement

A2.1 The Concept of Entanglement

There are a number of distinct properties loosely associated with quantum entanglement, which are equivalent for pure states but inequivalent for mixed states. These are differently related to the ideas of coherence and decoherence, and so it will be helpful to consider them separately (see Clarke 2002a for details and proofs). To avoid lots of negations, the following conditions for a mixed state ρ on two separated systems A and B, with associated Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, are phrased in terms relating to *not* being entangled. P, P' denote propositions on system A, Q, Q' denote propositions on system B, and for any observable X we write $\langle X \rangle$ for the expectation value $\text{Tr}(\rho X)$. When acting on the joint system defined by A and B, P will be identified with $P \otimes I_B$ and Q will be identified with $I_A \otimes Q$. The conditions below are connected by $1 \Leftrightarrow 2 \Rightarrow 3 \Rightarrow 4$ where the implications "$\Rightarrow$" are strict (not reversible).

1. Statistical independence: For all P, Q we have $\langle PQ \rangle = \langle P \rangle \langle Q \rangle$. For mixed states (which incorporate classical probabilities) the violation of this need not indicate entanglement because the events at A and B may be the result of a common preceding cause which results in their being correlated.
2. Factorizability: $\rho = \alpha \otimes \beta$ where α and β are states on A and B separately.
3. Separability: $\rho = a_1 \alpha_1 \otimes \beta_1 + a_2 \alpha_2 \otimes \beta_2 + ... + a_n \alpha_n \otimes \beta_n$ for positive numbers $a_1, ..., a_n$.
4. Bell's inequality: For all propositions P, P', Q, Q' as above we have $|\langle RS \rangle + \langle RS' \rangle + \langle R'S \rangle - \langle R'S' \rangle| \leq 2$, where $R = 2P - I$, $S = 2Q - I$, $R' = 2P' - I$, $S' = 2Q' - I$. Bell's inequality is generally regarded as the most solid test of locality, because it can be shown to hold for any sort of locality (not just quantum effects) under very general conditions and is accessible to experimental test. I will not consider it directly here, working with the previous more tractable conditions which imply it.

A2.2 Decoherence and Entanglement

I will next give a simple model to illustrate how entanglement can be very easily removed by decoherence. This is important in its own right, in order to understand the main obstacle to quantum states playing a role in the brain, and also in order to understand why a consideration of mixed states is essential.

Consider first a pure entangled state $|S\rangle = (1/\sqrt{2})\ (|a1\rangle|b2\rangle + |a2\rangle|b1\rangle)$, corresponding to two particles labeled by a and b at different positions, each capable of being in either of two states labeled by 1 and 2. To discuss decoherence we suppose these states are coupled to a state $|E\rangle$ describing an aspect of the environment, such as the electromagnetic field, so that the joint state is $\rho = |S\rangle|E\rangle$. We then suppose that an interaction takes place as a result of which particle a acquires a phase factor of either $e^{i\omega_1 T}$ or $e^{i\omega_2 T}$, where T is a time scale characteristic of the interaction and the ω's are (large) frequencies depending on the energies exchanged, which in turn depend on which state the particle is in. (The environmental state will also suffer a small change depending on the state, but to a good approximation this can be ignored here.) Then ρ changes to:

$$\begin{aligned}\rho' &= \frac{1}{2}\left(|a1\rangle\langle a1||b2\rangle\langle b2| + |a1\rangle\langle a2||b2\rangle\langle b1|e^{i(\omega_1-\omega_2)T}\right)|E\rangle\langle E| \\ &\quad +\frac{1}{2}\left(|a2\rangle\langle a1||b1\rangle\langle b2|e^{i(\omega_2-\omega_1)T} + |a2\rangle\langle a2||b1\rangle\langle b1|\right)|E\rangle\langle E|\end{aligned}$$

This is still a pure state, but in general a laboratory observer would be ignorant of the nature of the environmental state $|E\rangle$. From their perspective of ignorance, the probabilities of various results of observations would actually have to be calculated by averaging over the possible values of the unknown environmental variable, thus producing a mixed state. Such an averaging removes the oscillating terms with ω. Moreover, because ω is large, this is almost independent of any additional *a priori* knowledge of the likely probability distribution of the environmental variables. With the removal of these terms, the state becomes *separable* (as defined above), which is one definition of being unentangled. If we switch terminology from “observer” to “organism” and follow through this linguistic change we reach the same conclusion: unless the organism can compensate for an environmental influence about which it will have hardly any information stored, the effect of the state within the organism will be represented with high accuracy by a separable state, even though it is part of a larger state which is still entangled.

A2.3 Coherent Entanglement

In order to define the carrier of consciousness we are looking for a condition on a spatially distributed state which expresses the fact that its separated parts are entangled, and hence, with the interpretations proposed here, that it is an organism and a carrier of consciousness. Entanglement is a matter of degree, so that the concepts of organism and consciousness will be necessarily also be a matter of degree.

I shall first consider pure states, for which the situation is comparatively clear, and then extend this to mixed states. With pure states the conditions (1)–(4) in *A*2.1 are equivalent, and it is convenient to work with (1). The following elementary result establishes a quantitative measure of entanglement.

Proposition. Let $|S\rangle$ denote a state in a tensor product of Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$. Define

$$D(S, \mathcal{H}_A, \mathcal{H}_B) = \sup_{P,Q}\ 4\Big(\langle S|P|S\rangle\langle S|Q|S\rangle - \langle S|PQ|S\rangle\Big)\ ,$$

where P, Q range over projections on $\mathcal{H}_A$ and $\mathcal{H}_B$ and the identifications defined above are made when P, Q act on $\mathcal{H}_A \otimes \mathcal{H}_B$. Then $0 \leq D \leq 1$, and $D = 0$ if and only if $|S\rangle$ is unentangled (factorizable) and for given $\mathcal{H}_A$ and $\mathcal{H}_B$ of dimensions greater than 2 there exists a state $|S\rangle$ with $D = 1$.

Proof. Write

$$|S\rangle = \sum a^{ij}\nu_i^A \otimes \nu_j^B + \sum b^{ij}\nu_i^A \otimes \rho_j^B + \sum c^{ij}\rho_i^A \otimes \nu_j^B + \sum d^{ij}\rho_i^A \otimes \rho_j^B \ ,$$

where ρ and ν denote elements of bases of the range and null space, respectively, of P or Q in their respective spaces, and set

$$\alpha = \sum |a^{ij}|^2, \quad \beta = \sum |b^{ij}|^2, \quad \gamma = \sum |c^{ij}|^2, \quad \delta = \sum |d^{ij}|^2 \ .$$

Then $\langle S|P|S\rangle = \gamma + \delta$, $\langle S|Q|\rangle = \beta + \delta$, $\langle S|PQ|S\rangle = \delta$, and for the inequality $D \leq 1$ it is sufficient to maximize the expression in parentheses in the definition of D with respect to $\alpha, \beta, \gamma, \delta$ subject to $\alpha+\beta+\gamma+\delta = 1$. The condition for $D = 0$ is (1) above, while a state achieving $D = 1$ is given by

$$|S\rangle = \frac{1}{\sqrt{2}}\left(u^A \otimes v^B \pm v^A \otimes u^B\right)$$

for bases of $\mathcal{H}_A$ and $\mathcal{H}_B$ containing the vectors u, v.

We now want to extend this to a concept of spatial coherence. At this stage a number of minor variations are possible, which would need to be examined in the context of specific models. I will describe what seems to be the most straightforward extension.

Suppose we have a many-body state $|V\rangle$ in a space $\mathcal{H} = \otimes_{i=1}^n \mathcal{H}_i$, where the $\mathcal{H}_i$ are (e.g. molecular) state spaces at distinct spatial points x_i. Consider the set $\mathcal{P}$ of all partitions Λ of $\{1, ..., n\}$ into disjoint non-empty sets L_Λ, R_Λ such that the corresponding sets of the x_i are divided by a plane. Denote the corresponding Hilbert spaces by $\mathcal{H}_\Lambda^L = \otimes_{i\in L_\Lambda}\mathcal{H}_i$ and $\mathcal{H}_\Lambda^R = \otimes_{i\in R_\Lambda}\mathcal{H}_i$. Then we can define the *degree of coherent entanglement* of $|V\rangle$ as:

$$C(|V\rangle) = \min\nolimits_{\Lambda\in\mathcal{P}} \ D(V, \mathcal{H}_\Lambda^L, \mathcal{H}_\Lambda^R).$$

The proposition above, together with consideration of the maximally entangled state

$$|E\rangle = \frac{1}{\sqrt{2}} \sum_{i<j\leqslant n} \left(u_i^j \otimes v_j^i \pm v_i^j \otimes u_j^i\right)$$

(where, for each i, the set $\{u_1, ..., u_{i-1}, u_{i+1}, ..., u_n, v_1, ..., v_{i-1}, v_{i+1}, ..., v_n\}$ is a set of orthonormal vectors in $\mathcal{H}_i$), shows that $0 \leq C(|V\rangle) \leq 1$, $C(|V\rangle) = 0$ if and only if $|V\rangle$ is a product of states in $\mathcal{H}_\Lambda^L, \mathcal{H}_\Lambda^R$ for some partition Λ in $\mathcal{P}$ and there exists a state $|E\rangle$ with $C(|E\rangle) = 1$.

When we proceed to *mixed states* ρ, there are two possible approaches. The first involves imitating the definition of separability for mixed states, where we define a degree of coherent entanglement to be

$$C(\rho) = \min \max_i C(|V_i\rangle) ,$$

where the minimum is over all representations of ρ as $\rho = a_1|V_1\rangle\langle V_1| + ... + a_k|V_k\rangle\langle V_k|$ for positive numbers a_i. This extends the results established to mixed states and relates it to separable mixed states.

An alternative approach would be to borrow part of the definition of coherence for electromagnetic fields (Loudon 1983) by restricting the term $\langle S|PQ|S\rangle$ (now replaced by a trace over the mixed state) in the definition of D to the case where P has positive frequency and Q has negative frequency, including in the trace expression a time average over the *locus* of the observation (see App. $A4$) and modifying the whole expression for D so as to make all the terms separately real. While this would give a tidy continuity with the electromagnetic definition of coherence, it seems to be imposing a condition of equality of frequency over all parts of the state which could be too strong, and not applicable to coherent structures that are far from oscillators.

A3. Penrose-Modified Histories

Penrose's criterion for the rate of gravitational collapse is phrased in terms of an objective wave function, but without any specific mechanism or form. Here I will recast it in terms of the generalized histories interpretation. A great deal more work is, of course, needed in order to determine the true form of this, and the present note is only a start towards a feasibility study.

Let us consider a history framework corresponding to an organism, of the form $\Sigma = (\sigma_1, \sigma_2, ..., \sigma_n)$, where each σ_i is a Boolean algebra associated with a support U_i. Our aim is to specify the "probability" of any particular history $H = (P_1, P_2, ..., P_n)$, with each P_i in σ_i, for a given initial state. (The list of Us will not be given explicitly in the specification of a history, but is always to be understood.)

Before proceeding, however, I need to discuss a complication arising from the time separation of the different supports. By definition, different supports lie on different hypersurfaces, and so we need to relate the resulting different quantizations. Consider the general case of two different Cauchy surfaces V_1 and V_2 (surfaces that are each intersected once by any timelike curve). The propagation equations for quantum fields allow us to connect the quantum fields and momenta on V_1 to those on V_2 and, hence, to relate the algebras that they generate. From this we can seek to connect the Hilbert spaces of their respective quantizations and thence to establish a unitary map between them.

When the underlying space-time is static this is unproblematic: We can choose the Cauchy surfaces to be orthogonal to the timelike Killing vector and find a natural correspondence. If this is not the case, however, the situation is more complex (e.g. Wald 1994, Fulling 1989). The field propagation mixes the creation and annihilation operators with each other, so that the vacua of the two quantizations no longer correspond and particles are in general created from the vacuum between V_1 and V_2. Different choices of the Cauchy surfaces give rise to different particle production rates. Even if the two representations of the canonical algebra are equivalent, questions arise as to which is the "right" vacuum. Similar problems can arise when the quantum system is defined by a time-varying context of an organism.

I make no attempt to resolve these problems here; their solution will require a detailed examination of each individual case. Instead I assume, first, that we can unambiguously associate the Hilbert spaces on all the Cauchy surfaces of a given *locus*, so that we can simply talk of "the" Hilbert space of the *locus*. Second, I assume that between any two chronologically preceding *loci* U_i and U_j we can, by extending Cauchy surfaces through them and then restricting to the *loci*, establish an unambiguous unitary map F_{ij} between the first and second Hilbert space.

We can now proceed to associate numbers between 0 and 1 – which I shall call *pre-probabilities* – with histories, specifying more precisely the formula in Sect. 3.2 as

$$p(P_1, P_2, P_3, ...) = \mathrm{Tr}\ (...P_3 F_{32} P_2 F_{21} P_1 F_{10}\ \rho\ F_{01} P_1 F_{12} P_2 F_{23} P_3 ...)$$

(with the U's understood throughout). Note that we are *not* requiring these to be additive in any sense over the histories in a given framework, because such a requirement would force the theory to have the structure of the classical world (whereas our interest here is in the physical origins of such a structure, and possible departures from it) and would also be incompatible with incorporating Penrose's gravitational hypothesis. We note also that the state at the ith *locus* is ρ_i, defined by

$$\rho_i := Q_i^T \rho\, Q_i, \quad Q_i := \prod_{k=1}^{i} F_{k-1k} P_k\ ,$$

where ^{T} denotes the transpose and $\rho_0 = \rho$.

The following would then seem to be a reasonable modification of the histories formula to incorporate the gravitational hypothesis. We enumerate the members of σ_i as $P^1, ..., P^n$ and write T^k for the expectation value of the energy-momentum tensor in the state ρ_i defined above. Then let E^{kl} be the gravitational self-energy of the difference $T^k - T^l$, and define a characteristic gravitational collapse time for proposition P^k at the *locus* U_i by setting

$$t^{kG} = \max_{l:P^l \cap P^k = 0} \left(\frac{\hbar}{E^{kl}} \right) ,$$

interpreted as zero if $P^k = I$ (so that there are no terms in the max).

In order of magnitude, $t^{kG} \approx \hbar r / Gm^2$, where r and m are characteristic length and mass scales pertaining to the various $T^k - T^l$, so that the quantity t^{kG} is proportional to the ratio of the two very small constants G and $\hbar$. The quantities $T^k - T^l$ give a measure of the fineness of subdivision expressed in the algebra σ_i; the finer it is, the less distinguishable are the various propositions, and the longer it takes for the states to collapse (in Penrose's terminology) to a particular one.

The probability of a particular projection P^k being realized will then depend on both the quantum amplitude for this measurement result and the chance that a definite outcome will be realized at all. This last will be of the order of $p^{kG} := \exp\left(-t^{kG}/t_i^c\right)$, where $t_i^c = (d_{i,i-1} + d_{i-1,i})/2$ and $d_{i,j} = \min_{x \in U_i} \max_{y \in U_j} t(x,y)$ with $t(x,y)$ being the proper-time separation of x and y. In the trivial (but important) case where $P^k = I$, we have $p^{kG} = 1$.

Combining this with the quantum probability gives an overall pre-probability of

$$p(H) = \mathrm{Tr}(\rho_n) \prod_{i=1}^{n} p^{kG} .$$

For the usual case where the light-crossing time $d_{i,i-1}$ is much smaller than the gravitational collapse time, significant pre-probabilities will only be obtained if most of the propositions in the history are the identity (and thus have no effect). If the time-spacing of the non-identical propositions is large compared with the gravitational collapse time the expression will reduce to the standard one.

A4. Space-Time *Loci*

In this Appendix I will adopt a relativistic time ordering of events applicable to curved space-time. The main structure that is needed here is the existence of a relation of "before" in a chronological sense, which I assume can be defined globally.

In the idealization adopted here, experience is divided into discrete events, each occupying a region in space-time – which Isham (1994) called the support of the event, but which I will call the *locus* (to agree with Clarke (2001) and to avoid confusion with the use of support in a more technical sense). The spatial extent of a *locus* is the geodesic convex hull h of the support of the coherent state (i.e. the support "filled in"). For the temporal extent there is one very natural choice, namely the maximum extent for which the *locus* is causally self-contained, in the sense that if a and b are points in the *locus* such that a is before b, then every point x between a and b (i.e. such that a is before x and x is before b) is also in the *locus*. This is essentially the definition of a region's being *globally hyperbolic*. I call a globally hyperbolic region U *maximal* when there is no other globally hyperbolic region V containing U such that for any point x in V either every inextendible timelike curve to the future of x cuts U or every inextendible timelike curve to the past of x cuts U. If this is the case then all the points to the future of a that can influence b are themselves in the *locus* (hence the idea of being causally self-contained).

For a generalized history we require also that, for any two *loci* U and V, either (i) $U < V$ in the sense that every point in U is before every point in V, or (ii) $V < U$ in the sense that every point in V is before every point in U, or (iii) $U \sim V$ in the sense that every point in U is simultaneous with every point in V. This and the preceding stipulation on the duration of *loci* has the advantage that the formula used for probabilities in the ordinary histories interpretation now applies unambiguously and without change to the generalized interpretation.

References

Beltrametti E.G. and Cassinelli G. (1981): *The Logic of Quantum Mechanics*, Addison-Wesley, Reading MA.

Bennett J. (1984): *A Study of Spinoza's Ethics*, Cambridge University Press, Cambridge.

Blackmore S. (2003): *Consciousness, an Introduction*, Hodder & Stoughton, London.

Chalmers D.J. (1995): Facing up to the problem of consciousness. *Journal of Consciousness Studies* **2**, 200–219.

Clarke C.J.S. (1973): Smoothing the quantum collapse. *International Journal of Theoretical Physics* **8**, 231–235.

Clarke C.J.S. (1993): Process as a primitive physical category. In *Time and Process*, ed. by J. T. Fraser and L. Rowell, International Universities Press, Madison CT, pp. 53–59.

Clarke C.J.S. (1994): Distributed pumped dipole systems do not admit true Bose condensations. *Journal of Physics A* **27**, 5495–5500.

Clarke C.J.S. (2001): Consciousness and non-hierarchical physics. In *The Physical Nature of Consciousness*, ed. by P. van Loocke, Benjamins, Amsterdam, pp. 191–217.

Clarke C.J.S. (2002a): Entanglement and statistical independence for mixed quantum states. *Foundations of Physics Letters* **15**(5), 495–500.

Clarke C.J.S. (2002b): *Living in Connection*, Creation Spirituality Books, Warminster.

Clarke C.J.S. (2004): Quantum mechanics, consciousness and the self. In *Science, Consciousness and Ultimate Reality*, ed. by D. Lorimer, Imprint Academic, Exeter, pp. 65–92.

de Quincey C. (2002): *Radical Nature: Rediscovering the Soul of Matter*, Invisible Cities Press, Montpelier VT.

Deckers J. (2004): Christianity and ecological ethics: The significance of process thought and a panexperientialist critique on strong anthropocentrism. *Ecotheology* **9**(3), 359–387.

Dennett D.C. (1991): *Consciousness Explained*, Penguin Press, London.

Dixon J.M., Tuszynski J.A., and Clarkson P.A. (1997): *From Nonlinearity to Coherence: Universal Features of Nonlinear Behavior in Many-Body Physics*, Clarendon Press, Oxford.

Douglas-Klotz N. (2001): *The Hidden Gospel: Decoding the Spiritual Message of the Aramaic Jesus*, Quest Books, Wheaton IL.

Dowker F. and Kent A. (1996): On the consistent histories approach to quantum mechanics. In *Journal of Statistical Physics* **82**, 1575–1646.

Eco U. (1976): *A Theory of Semiotics*, Indiana University Press, Bloomington IN.

Fröhlich H. (1968): Long-range coherence and energy storage in biological systems. *International Journal of Quantum Chemistry* **2**, 641–649.

Fulling S.A. (1989): *Aspects of Quantum Field Theory in Curved Space-Time*, Cambridge University Press, Cambridge.

Ghirardi G.C., Rimini A., and Weber T. (1986): Unified dynamics for microscopic and macroscopic systems. *Physical Review D* **34**, 470–491.

Giulini D., Joos E., Kiefer C., Kupsch J., Stamatescu I.-O., and Zeh H.D., eds. (1996): *Decoherence and the Appearance of a Classical World in Quantum Theory*, Springer, Berlin.

Hagan S., Hameroff S., and Tuszynski J. (2002): Quantum computation in brain microtubules? Decoherence and biological feasibility. *Physical Review E* **65**, 061901.

Hameroff S.R. and Penrose R. (1996): Orchestrated reduction of quantum coherence in brain microtubules: A model for consciousness? In *Toward a Science of Consciousness – The First Tucson Discussions and Debates*, ed. by S.R. Hameroff, A.W. Kaszniak, and A.C. Scott, MIT Press, Cambridge MA, pp. 507–540.

Hartle J. (1991): The quantum mechanics of cosmology. In *Quantum Cosmology and Baby Universes,* ed. by S. Coleman, P. Hartle, T. Piran, and S. Weinberg, World Scientific, Singapore.

Ho M.-W. (1998): *The Rainbow and the Worm*, World Scientific, Singapore.

Holevo A.S. (1973): Statistical problems in quantum physics. In *Proceedings of the Second Japan-USSR Symposium on Probability Theory*, ed. by G. Murayama and J.V. Prokhorov, Springer, Berlin, pp. 104–109.

Isham C.J. (1994): Quantum logic and the histories approach to quantum theory. *Journal of Mathematical Physics* **35**, 2157–2185.

Kochen S. and Specker E. (1967): The problem of hidden variables in quantum mechanics. *Journal of Mathematics and Mechanics* **17**, 59–87.

Libet B. (2004): *Mind Time*, Harvard University Press, Cambridge MA.

Loudon R. (1983): *The Quantum Theory of Light*, Oxford University Press, Oxford.

Marshall I. (1989): Consciousness and Bose-Einstein condensates. *New Ideas in Psychology* **7**, 73–83.

Penrose R. (1989): *The Emperor's New Mind*, Oxford University Press, Oxford.

Penrose R. (1994): *Shadows of the Mind*, Oxford University Press, Oxford.

Penrose R. (2004): *The Road to Reality: A Complete Guide to the Laws of the Universe*, Jonathan Cape, London.

Shannon C.E. and Weaver W. (1949): *The Mathematical Theory of Communication*, University of Illinois Press, Urbana IL.

Sheldrake R. (2003): *The Sense of Being Stared at, and Other Aspects of the Extended Mind*, Hutchinson, London.

Skrbina D. (2005): *Panpsychism in the West*, Bradford Books, Cambridge MA.

Stapp H. (1993): *Mind, Matter and Quantum Mechanics*, Springer, Berlin.

Stapp H. (2005): Quantum interactive dualism: An alternative to materialism. *Journal of Consciousness Studies* **12**(11), 43–58.

Sudbery A.S. (1984): The observation of decay. *Annals of Physics* **157**, 512–536; erratum in *Annals of Physics* **176**, 181.

Varela F.J., Thompson E., and Rosch E. (1993): *The Embodied Mind*, MIT Press, Cambridge MA.

Velmans M. (2000): *Understanding Consciousness*, Routledge, London.

Wald R.M. (1994): *Quantum Field Theory in Curved Spacetime and Black Hole Thermodynamics*, University of Chicago Press, Chicago.

Washburn M. (2003): *Embodied Spirituality in a Sacred World*, SUNY Press, Albany NY.

Whitehead A.N. (1929): *Process and Reality*, Cambridge University Press, Cambridge.

Received: 27 February 2006
Revised: 15 October 2006
Accepted: 23 October 2006

Reviewed by William Seager and Jack Tuszynski