Generalised histories and decoherence: an outline for a large-scale quantum theory

Chris Clarke
School of Mathematics,
University of Southampton,
Southampton, United Kingdom,
SO17 1BJ
cclarke@scispirit.com

December 19, 2007

Abstract

Combining J. Hartles "generalised quantum theory" with proposals by D. Page, R. Penrose and M.-W. Ho produces a testable version of cosmological quantum theory which fully accounts for the emergence of the classical world with no additional implicit assumptions.

Keywords: quantum cosmology, decoherence, history interpretation PACS Nos: 03.65.Ta, 03.65.Yz, 98.80.Qc

1 Introduction

1.1 Background

For a large part of its history, applications of quantum theory have been carried out within the (usually implicit) von Neumanns conceptual framework, in which the existence of a classical world surrounding the phenomenon under investigation is assumed. Two factors have changed this: the development of quantum cosmology, which has created the demand for a more general approach, and the growth in understanding of the phenomenon of decoherence which has gone a long way to producing such an approach. As a result we now understand in detail why some subsystems of the universe behave classically and others do not, and we can thereby demonstrate that, in principle, a classical world can, with high probability, emerge from the sorts of initial conditions envisaged in quantum cosmology (Giulini et al., 1996). What we cannot do, however, is to rule out the emergence with equally high probability of any of a far larger range of totally non-classical "worlds that would be incompatible with any imaginable form of life, or the existence of semi-permanent records, because there is nothing in the bare formalism of quantum theory to prevent this.

The clearest illustration of the need for such an approach comes from current theories of the early universe, where it is supposed that the universe begins with a homogeneous state, from which structure arises through "quantum fluctuations" (Turner, 1999;

Linde, 2001). In these accounts the initial quantum state of the universe is usually homogeneous; but the problem then is that, since the Hamiltonian commutes with the spatial symmetry group, the quantum state will remain homogeneous under Schrödinger evolution, so that without further structure there is nothing to single out the inhomogeneous universe that we are in.

The commonest way of handling these issues is to adopt an approach similar to, though much weaker than, the application of an "anthropic principle in cosmology, in which it is stipulated that we are only interested in those universes that produce something that could be identified as a brain (Donald, 1990, 1995) or sensation Page, 2001. But this is not the end of the matter. It has, for example, been suggested (private communication) that, having decided on a criterion, all that is then required is simply to identify the systems satisfying the criterion as observers and then apply the usual Bohr interpretation. That interpretation, however, imposes a "simultaneous collapse throughout all space, and not only is there no immediate possibility of formulating any mechanism for this, but more seriously the notion of absolute simultaneity has no place in a general relativistic theory applicable to the entire universe with all its structure (as opposed to a simplified cosmological model). We need to realise that the only role of "collapse is to impose mutual consistency restrictions on observations that are causally related, resulting in the probabilities for the outcomes of two such observations being dependent. This is simply the definition of the histories approach in the version of Hartle (1991): it is the extraction of the essence of the collapse hypothesis in a way that is relativistically invariant and does not presume any unknown mechanism.

The aim of this paper is thus to show how a criterion for sensation can be wedded to Hartles histories approach in a way that is well defined, physically consistent, and scientifically fruitful. It turns out that this requires significant mathematical machinery. I shall therefore outline the essence of the approach first, filling in the details in subsequent sections in order to demonstrate that the approach is in fact workable.

The criterion for mind/sensation and the details of the histories formalism should at least satisfy the following minimal requirements. They should

1. be philosophically reasonable in the sense that it does correspond to the essence of mind/sensation;
2. not be so strong that it essentially puts in by hand physical properties that we can already explain physically through decoherence theory
3. be based on a well defined causal relation between observations

To be scientifically productive the criterion should then imply that

4. the contents of sensation are (nearly) classical, in the sense that histories of sensations (nearly) satisfy the decoherence criterion for histories (Dowker and Kent, 1996);
5. quantum systems should not be subject to Zeno-like effects that are not in fact found in laboratory experiments.

The points above strongly suggest the particular criterion and structure of histories used here. In relation to 1, I have argued elsewhere (Clarke, 2007) for (a) basing the criterion on consciousness rather than on the computational or memory properties of the brain, because consciousness is the essence of sensation; and (b) using a panpsychist, dual-aspect philosophy of consciousness in which *all* systems satisfying a general physical criterion have consciousness. This rules out the approach of (Donald, 1990, 1995) which singles out a very restrictive class of systems which does not relate to the capacity of the brain to produce consciousness. Points 2 and 4 suggest that the boundary between the sensate system and the surrounding environment should be set at a point where decoherence takes over, allowing the known mechanisms for decoherence to produce a classical world. This also agrees with the arguments of Hameroff and Penrose (1996) and myself (Clarke, 2001) that quantum theory, rather than classical theory, is the most natural framework within which to understand consciousness. Thus the criterion should single out maximal regions which are in some sense coherent (the "extensive coherence defined below in subsection 1.3). Note that this approach is the opposite of that of Bohr as far as the place of the observer is concerned: here the observer is identified with the quantum micro-system (e.g. a small subsystem of the brain) with the macro-system (e.g. the body as a whole) providing classicality, whereas in Bohr's quantum theory the observer is the record forming macro-system which observes a separate quantum micro-system.

Point 5 refers to the possibility that if the criterion were, for example, to allow an excited atom to have sensation, then its self-observation could inhibit its decay even though it is not in interaction with a record-forming external system, contradicting what is in fact the case. Points 4 and 5 are formally the same, except that the first is considered from the point of view of the experiencer and the second from the point of view of the onlooker. The application of 5 depends on the dynamics of the quantum theory being used rather than on its formal structure, and here it turns out that in order to avoid spurious Zeno-like effects the dynamics of quantum mechanics needs to be modified by the inclusion of an effect of quantum gravity for which Penrose (2004) has argued in relation to the role of consciousness. The qualification "nearly" is added to 4 because the criterion, if it is to match the experience of sensation, will be a matter of degree rather a strict dichotomy, and there will therefore be a borderline of nearly classical behaviour which will provide an arena for testing the theory. Finally, the point 3 is most simply implemented by placing restrictions on the temporal extent of observations and on their inclusion in histories.

1.2 quantum states

To formulate the structure proposed here we need to make some assumptions about what sort of quantum theory is to be used. We require a generally applicable theory of cosmological scope; but on the other hand it would take us too far afield to speculate on possible theories of quantum gravity. We are thus considering varieties of quantum field theory in curved space time. For the sake of definiteness I will assume that the theory has the features enumerated below, though it is likely that these do not hold in all cases. In the Appendix I derive them for the unambiguous case of static space-time. Here I assume that the background space-time M is globally hyperbolic and require the following:

1. For any achronal 3-surface S in M
   (a) there is a complex Hilbert space $\mathcal{H}(S)$ whose rays represent quantum states on S;

(b) for any partition of S into disjoint submanifolds S^0 and S^1 with $(\overline{S^0} \cup \overline{S^1}) \cap S = S$ there are Hilbert spaces $\mathcal{H}^0(S)$ and $\mathcal{H}^1(S)$ and an isometric injection $\iota_{S^0,S^1} : \mathcal{H}(S) \to \mathcal{H}(S^0) \otimes \mathcal{H}(S^1)$.

2. For any 1-parameter foliation $\{S_t \mid t \in I\}$ by Cauchy surfaces

   (a) the Hilbert spaces $\mathcal{H}_t := \mathcal{H}(S_t)$ are related by maps $\{U_{t_1t_2} \mid t_1, t_2 \in I\}$, representing a Schrödinger-picture evolution, satisfying

   $$U_{t_1t_2} \circ U_{t_0t_1} = U_{t_0t_2}, \quad U_{t_1t_1} = \text{identity on } \mathcal{H}_{t_1} \qquad \forall\, t_1, t_2 \in I.$$

   (b) for any submanifold $S^0_{t_1} \subset S_{t_1}$, if $S^0_{t_2} \subset S_{t_2}$ is defined by Lie-dragging orthogonally to the Cauchy surfaces then there exist maps

   $$\{T^0_{t_1t_2} \mid t_2 \in I\} \text{ where } T^0_{t_1t_2} : \mathcal{H}(S^0_{t_1}) \to \mathcal{H}(S^0_{t_2}),$$

   representing time translation, satisfying

   $$T^0_{t_1t_2} T^0_{t_0t_1} = T^0_{t_0t_2}, \quad T^0_{t_1t_1} = \text{identity on } \mathcal{H}(S^0 t_1) \qquad \forall\, t_1, t_2 \in I.$$

3. If S is a submanifold of S_{t_1} partitioned into $S^0_{t_1}$ and $S^1_{t_1}$ as in 1(b) above, then the following diagram commutes:

$$\begin{array}{ccc} H(S_1) & \xrightarrow{\iota_1} & H(S^0_1) \otimes H(S^1_1) \\ \downarrow{\scriptstyle T_{12}} & & \downarrow{\scriptstyle T^0_{12} \otimes T^1_{12}} \\ H(S_2) & \xrightarrow{\iota_2} & H(S^0_2) \otimes H(S^1_2) \end{array}$$

(where for legibility the ts are omitted, so that $S^0_1 := S^0_{t_1}$ and so on).

1.3 Extensive Coherence

Let W^1_Λ, W^2_Λ be the sets of projections of the form $P = P_1 \otimes I$, $Q = I \otimes Q_2$ with P_1, Q_2 projections in the corresponding Hilbert spaces $\mathcal{H}^1_\Lambda$ and $\mathcal{H}^2_\Lambda$ respectively. (The definitions will be symmetric as between these two sides of the surface.) Write the diagonalisation of ρ (assumed for the moment to be unique up to ordering) as $\rho = \sum_{k=1}^m a_k |\alpha_k\rangle \langle\alpha_k|$, where possibly $m = \infty$. Then define

$$\begin{aligned} \mathcal{C}(\rho) &:= \min_{\Lambda\in\mathcal{P}} \min_{P\in W^1_\Lambda, Q\in W^2_\Lambda} \sum_{k=1}^m a_k |d(P, Q, \alpha_k)| \\ \text{where} \quad d(P, Q, \alpha_k) &:= 4\left(\langle\alpha, PQ\alpha\rangle - \langle\alpha, P\alpha\rangle \langle\alpha, Q\alpha\rangle\right). \end{aligned} \tag{1}$$

If the diagonalisation of ρ is not unique because of degeneracy we take the supremum over diagonalisations.

The function $d(P, Q, \alpha)$ with P and Q as above satisfies

$$d(I - P, Q, \alpha) = -d(P, Q, \alpha), \quad d(Q, P, \alpha) = d(P, Q, \alpha).$$

In (Clarke, 2007) it is shown that d lies between -1 and 1, with 0 attained if and only if α is factorisable into a tensor product of states in $\mathcal{H}^1_\Lambda$ and $\mathcal{H}^2_\Lambda$ (fully unentangled), and that $|d|$ can always attain the value 1 (fully entangled) for some α. In probability theory terms, $d(P, Q, \alpha)$ is, apart from the normalisation factor 4, the covariance of P and Q.

Coherence as measured by $\mathcal{C}$ is a variable quantity: states are more coherent the higher the value of $\mathcal{C}$ between 0 and 1. At the present exploratory stage, we can suppose chosen an arbitrary threshold $\mathcal{C}_0$ for $\mathcal{C}$. We then define an *extensively coherent state* on a locus $(U, \mathcal{H}', \mathcal{H})$ as an effective state ρ such that

1. $\mathcal{C}(\rho) > \mathcal{C}_0$.
2. There is no locus $(U^*, \mathcal{H}'_{U^*}, \mathcal{H}_{U^*})$ with $U^* \supset U$ and satisfying $\mathfrak{P}1$ such that $\mathcal{C}(\rho^*) > \mathcal{C}_0$.

These two ingredients of extensive coherence and Penroses quantum gravity will be incorporated into a form of the histories interetation of quantum theory (widely regarded as most suitable for cosmological applications) as follows:

1. The probability of a given history is modified by a factors depending, for each pair of events in the history, on the ratio of their time separation to the timescale of Penrose's mechanism for the action of quantum gravity.
2. the criterion is imposed that all the quantum states corresponding to the projections that make up a history should be extensively coherent.

1.4 Outline of paper

The main body of this paper is taken up with establishing a description of a histories formalism that is as invariant (coordinate independent) as possible, within which to define extensive coherence rigorously. Full invariance is impossible, unfortunately, because we do not have a fully invariant formulation of a general quantum field theory in curved space-time.

Section 2 sets the scene for the construction, summarising in 2.1 decoherence theory and in 2.2 Penrose's recent theory. Section 2.3 specifies the limitations of scope of the theory attempted here, while 2.4 specifies the quantisation to be used, details being given later in an appendix. Section 3, the main part of the paper, then defines and justifies the new structures being proposed, building on previous work by Hartle (1991). Finally section 4 sketches the implications that remain to be explored in detail, both theoretically and experimentally, including possible departures from classical physics.

2 Overview of issues addressed

2.1 The role of decoherence

Decoherence theory is now an essential part of quantum theory, fully developed both theoretically and experimentally (see, for example, Giulini et al., 1996; Zurek, 2003 together with Sudbery, 2002, for a conceptual critique of the experimentation).

This theory addresses the same problem as I do here: as Zurek (2003) puts it, "The problem: Hilbert space is big." Recent refinements of the theory have certainly made much headway in whittling down the vast arena of possible universes. In particular, it resolves much of the problem by taking the view that any "observation" must involve the formation of a stable record in one subsystem of the universe (the apparatus, or the brain) that is correlated with a short-term state of another subsystem (the observed system). Stability in the face of even the smallest environmental perturbations then, on this view, "selects" a preferred basis of states of the apparatus (the process of "einselection") whose dynamics correspond to the classical world.

Three problems, however, prevent this consideration being the whole answer.

1. The processes of decoherence depend on the existence of sufficiently well-defined subsystems. But what is it that marks off a subsystem, and what causes this to happen? The condition of extensive coherence proposed here is designed to address this very point, which involves a shift of viewpoint to a particular sub-class of what are usually regarded as subsystems, to which subjectivity is ascribed, possibly of a primitive form. This issue of ascribing subjectivity to subsystems has been analysed in detail by Mathews (1991)—see also section 4.3.

2. The word "selects" has an ambiguity that can obscure the issues. If we overcome the previous problem and can designate subsystems, then the process of coupling with the environment establishes a clear mathematical line along which we (mathematical) human beings can "select" and divide up the state into components that we can observe. But this does not give selection in the sense that all other possibilities are objectively eliminated from the universe. For this we need additional structure.

3. The suggestion that the only phenomena of interest are those which involve the formation of stable records seems motivated almost entirely by the need to model the scientific process, rather than the desire to model the universe as it appears to us.

The thesis of this paper is that decoherence, while playing a vital role, does not remove the need for incorporating a further selection process in a quantum cosmological view of the universe. It should be noted (section 4.3) that this added process, although linked to subjective aspects of consciousness, will have observable physical effects, making the proposal potentially testable.

2.2 Penrose's theory

The approach of Roger Penrose (Penrose, 2004) is put forward by him as an *alternative* to decoherence theory for resolving some of the problems indicated in the introduction above. Like decoherence theory, it implicitly assumes a laboratory-like situation in which inhomogeneities and subsystems have already evolved, providing a preferred framework within which he formulates a gravitational criterion for the "collapse of the state". If one starts off with a homogeneous cosmology, there is no canonical way of splitting it into a particular superposition on which his collapse mechanism can operate. I therefore cannot regard it as the whole story, and I find the notion of "collapse", as opposed to less metaphysically demanding formulations, too dependent on a future theory of quantum gravity whose details are at present necessarily obscure. His *criterion* for collapse, however, does derive from a robust line of argument which has to be reflected in any large-scale theory: namely, the argument that it is basically inconsistent with what we already know about gravitation for certain sorts of superpositions to be manifested for time scales longer Penrose's gravitational decay time. Moreover, his criterion identifies a "natural break" between the microscopic and macroscopic worlds, even though we do not yet understand their underlying mechanism. I therefore incorporate Penrose's timescale in the presentation here. The advantage of this approach, which *combines* Penrose's criterion with decoherence theory, is that a tight "niche" for subjectivity is selected, in systems that are large enough for Penrose's criterion to be satisfied, but small enough and structured enough to be able to maintain some degree of decoherence.

2.3 Limitations of this paper

The proposals here are not presented as a definitive theory: my aim is to illustrate the principles involved, recognising that considerably more work, both theoretical and experimental, is required in order to define unambiguously the way ahead. As a result, some aspects of the approach have to be chosen arbitrarily, subject to later scrutiny. In addition to the lack of a full theory of quantum gravity needed to implement Penrose's hypothesis, noted above, there are three respects in which in is clear that the proposals are only an approximation to a subsequent more developed theory: in a general setting we will be dealing with a hierarchical situation rather than with standard quantum field theory; in the histories formalism a discrete approximation is made to what is presumably a continuous situation—an issue that has been extensively investigated by Isham et al. (1999); and in a histories approach where consistency (5) is not imposed *a priori* the decoherence functional does not yield true probabilities. I comment on these further below.

2.3.1 The hierarchy of physics

This concerns the implicit operational assumption in physics of a hierarchical structuring of theory according to length scale (into, for example, quantum field theory, nuclear and atomic physics, molecular and solid state physics, classical physics and so on). Theory at one level is a (quasi-) coarse-graining of the theory at the next finer level, in which the Hilbert space at the finer level is written 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. Strict coarse-graining, as applied in rigorous investigations of the validity of the idea, involves tracing out the fine structure, producing a mixed state. Actual scientific practice, however, focuses on the case where the quantum state has the components in the two spaces unentangled, i.e. where the state has the form $v \otimes w$. In this case the component v represents the large-scale structure as a pure state. It is usually assumed that this is approximately valid, to which level of approximation there exists a self-contained quantum theory at the level represented by $\mathcal{H}_1$.

Practical applications of quantum theory will necessarily depend on the hierarchical level involved. For simplicity, however, I shall present the argument here largely in terms of quantum field theory, while referring to the hierarchical context where necessary. The role of particular hierarchical level will be represented here by singling out a particular factor of the Hilbert space over each relevant region of space-time (section 3.1). Note that the hierarchical system just described depends on a property of non-entanglement (factorisation of the sate) between macro- and micro-states. This propoerty is precisely what is ensured by the definition of extensive coherence (section

1.3) below. Thus this theory provides the basis for a division into hieracrchical levels, as well as being applied at different hierarchical levels.

2.3.2 Discrete versus continuous descriptions

In the context of the generalised histories interpretation given below, a *history* corresponds to the subjectivity of a collection of organisms, in a very general and primitive sense of the term (section 4.3). Each organism has a continuous existence in space and time and so corresponds to a space-time "tube". Considerable mathematical simplification is achieved, however, by approximating this tube by a sequence of closely spaced globally hyperbolic space-time regions so that the history is then comprised discrete "moments" of awareness[1]. The space-time in each region can be treated as being approximately flat and the natural timescale defined by the light-crossing time of the organism is built into the time-extent of the regions. The artificiality of the approximation should, however, be borne in mind.

I note that Isham et al. (1999) have considered in detail the situation of continuous histories, but their treatment is limited to *paths* of zero spatial extent and uses quite different definitions and methodology from the treatment here.

2.3.3 The question of probability

We recall that the histories formalism for quantum theory (see, e.g. Griffiths, 1984) is based on sequences of quantum propositions (projections) $P_1, P_2, \ldots, P_n$ on an underlying Hilbert space of pure states $\mathcal{H}$, supposed measured at times $t_1, t_2, \ldots, t_n$ on an initial mixed state ρ_0 which we can suppose to be at time 0. In a Schrödinger representation the probability of all these propositions being measured as true is then

$$p = \overline{p}(P_1, P_2, \ldots, P_n; P_1, P_2, \ldots, P_n; t_1, t_2, \ldots, t_n; \rho_0) \tag{2}$$

where $\overline{p}$ is the *decoherence functional* given by

$$\overline{p}(P_1, P_2, \ldots, P_n; Q_1, Q_2, \ldots, Q_n; t_1, t_2, \ldots, t_n; \rho_0) = \mathrm{Tr}(P'_n \ldots P'_2 P'_1 \rho_0 Q'_1, Q'_2, \ldots, Q'_n) \tag{3}$$

$$\text{where} \qquad P'_i = e^{iHt_i/\hbar} P_i e^{-iHt_i/\hbar} \tag{4}$$

and similarly for Q_i.

It is essential that the numbers claimed in this way to be probabilities do in fact satisfy the axioms for probabilities (crucially, additivity). This requirement was originally enforced by imposing the the "consistency" condition

$$\overline{p}(P_1, P_2, \ldots, P_n; P_1, P_2, \ldots, P_{i-1}, Q_i, P_{i+1} \ldots P_n; t_1, t_2, \ldots, t_n; \rho_0) = 0 \tag{5}$$

if $P_i Q_i = Q_i P_i = 0$. In later work this condition was deduced from decoherence effects arising in experimental settings, rather than being imposed *ad hoc*. In the more general context being considered here it cannot be guaranteed that the quantities defined by (2) are additive; the observational consequences of this are briefly discussed in section 4.3. In view of this issue, I will refer to the quantities in the present theory analogous to those of (2) as *pre-probabilities*.

[1]The representation in terms of separated regions should not be confused with Whitehead's "pulsational" view of awareness (or "prehension" in his terminology), which I interpret as referring to the existence of a "specious present", not a discontinuity in experience. The specious present also has a much longer timescale than that of the "moments" here.

2.4 Quantisation

While using quantum field theory, rather than a hierarchical notion (section 2.3.1), gives a conceptual simplicity, it exacts the penalty of our having to decide on a formalism to adopt for quantum field theory in curved space-time. Because of the stress on local properties here, I adopt a quantisation in terms of a configuration space of classical fields on space-like hypersurfaces, describing the details of this in the Appendix. There I make the (inessential) simplification to a Klein-Gordon field and the (more essential) simplification that the space-time is static, since in the general case the nature of quantisation becomes ambiguous (Wald, 1994), probably indicating the need for a more comprehensive setting in which to formulate the full effect of the creation of particles by the gravitational field which occurs in this case. At least in this static case (and I assume that eventually it will be possible to extend this unambiguously to all globally hyperbolic space-times) one obtains a quantisation of the field where the Hilbert space is $\mathcal{H}(S) = L^2(\Phi'(S), \mu)$. Here $\Phi'(S)$ is a space of scalar distributions (dual to the space $\Phi(S)$ of smooth rapidly decreasing functions) on a hypersurface S orthogonal to the timelike Killing vector, and μ is a measure on $\Phi'(S)$, normalised to unity, derived from the Klein-Gordon 'potential' (the elliptic part of the wave operator). Note that this is not an alternative quantum field theory, but just a less usual representation of the customary Fock space version. The translation between the two versions is given by Simon (1974).

The virtue of this construction is that it is not only mathematically rigorous (at least in the restricted situation of linear fields in static space-time and hopefully more generally) but one also obtains a simple localisation construction, described in the Appendix, such that the space $\mathcal{H}(S)$ of configuration-space states over a hypersurface $S = S_0 \cup S_1$, with S_i being submanifolds with boundary, can immediately be written as a subspace of the tensor product $\mathcal{H}(S_0) \otimes \mathcal{H}(S_1)$ of states over the two submanifolds.

3 Generalised Histories

3.1 Basic ideas

I adopt a histories formalism for quantum theory on the grounds that it includes mathematically other formalisms such as Copenhagen, "collapse of the state", "many worlds", "many minds" etc. but without their metaphysical assumptions or contextual limitations. The original version of this formalism is given by (2), where in the particular case of conventional quantum theory, P_n is the laboratory observation and the satisfaction of $P_1, P_2, \ldots, P_{n-1}$ constitutes the preparation of the state presented to this observation. In a larger context, or a cosmological one, it is however unreasonable to single out a single instant of time in a particular time-coordinate to associate with an element of a history: a better, but still tractable, approximation to the continuous space-time tube of the organism (section 2.3.2 above) would be a sequence of *space-time regions* $U_1, U_2, \ldots, U_n$.

Here I develop the work of Hartle (1991) on this same idea (see also Isham, 1994), but adding to the space-time region U the specification of a particular factor space of the quantum Hilbert space defined by U, representing a particular level of coarse-graining and a particular subsystem within that. Specifically, we define a *locus* as a triple $(U, \mathcal{H}'_U, \mathcal{H}_U)$ as follows. Suppose (as can be assumed without loss of generality) that there is a Cauchy surface S_1 for U contained in a surface of simultaneity (a global Cauchy surface orthogonal to the static Killing vector) S, where $S = S_0 \cup S_1$ with $S_0 \cap S_1 = \emptyset$. Then the definition of a locus on U is that $\mathcal{H}(S_1)$ can be represented as $\mathcal{H}(S_1) = \mathcal{H}'_U \otimes \mathcal{H}_U$. It follows (using the notation of section 2.4) that

$$\mathcal{H}(S) \subset \mathcal{H}(S_{U,0}) \otimes \mathcal{H}'_U \otimes \mathcal{H}_U. \tag{6}$$

I will call U the *support* of the locus.

For defining "extensive coherence" later we will also need the idea of splitting a locus into two parts. Suppose that $\mathfrak{L}_A$ $(A = 0, 1, 2)$ are loci with $\mathfrak{L}_A = (U_A, \mathcal{H}'_A, \mathcal{H}_A)$ and that S is a surface of simultaneity such that the surfaces $S_A = S \cap U_A$ are Cauchy surfaces of U_A, with $S_0 = S_1 \cup S_2$, $\text{int}S_1 \cap \text{int}S_2 = \emptyset$. Suppose in addition that

$$\mathcal{H}'_0 = \mathcal{H}'_1 \otimes \mathcal{H}'_2, \qquad \mathcal{H}_0 = \mathcal{H}_1 \otimes \mathcal{H}_2. \tag{7}$$

Then we will express this situation by writing

$$\mathfrak{L}_0 = \mathfrak{L}_1 \otimes \mathfrak{L}_2.$$

We can now define a *generalised history* to be a sequence $(\mathfrak{P}_1, \mathfrak{P}_2, \dots \mathfrak{P}_n)$ of quadruples $\mathfrak{P}_i = (U_i, \mathcal{H}'_i, \mathcal{H}_i, P_i,)$, where $(U_i, \mathcal{H}'_i, \mathcal{H}_i)$ is a locus and P_i is a projection on the total quantum Hilbert space $\mathcal{H}$, satisfying the conditions $\mathfrak{P}1$–$\mathfrak{P}4$ below (see Hawking and Ellis, 1973, for detailed definitions of the notation for space-time properties). I will call[2] the $\mathfrak{P}$s "moments (of awareness)", and U_i the *support* of the moment P_i.

$\mathfrak{P}1$ Each connected component of U_i has a geodesically convex Cauchy surface.

$\mathfrak{P}2$ $D(U_i) = U_i$ where D denotes the domain of dependence

$\mathfrak{P}3$ Given a pair of Us, either one is entirely to the future of the other (denoted by the relational sign $\prec$), or they are entirely space-like related (denoted by the sign $\succ\!\prec$).[3]. More precisely: for any $i \neq j$,

either	(i)	$\forall(x \in U_i, y \in U_j)(x \in J^+(y))$	$(U_j \prec U_i)$
or	(ii)	$\forall(x \in U_i, y \in U_j)(x \in J^-(y))$	$(U_j \succ U_i)$
or	(iii)	$\forall(x \in U_i, y \in U_j)(x \notin J(y)\ \&\ y \notin J(x))$	$(U_j \succ\!\prec U_i)$

Moreover, if $i < j$ then either (ii) or (iii) holds.

$\mathfrak{P}4$ For every i, P_i can be written

$$P_i = I_{\mathcal{H}(S_{U,0}) \otimes \mathcal{H}'_i} \otimes \tilde{P}_i$$

with respect to the decomposition (6), for a projection $\tilde{P}_i : \mathcal{H}_i \to \mathcal{H}_i$.

[2]The terminology and the precise definition here differ from those of (Clarke, 2007) but are broadly in agreement with those of (Clarke, 2001).

[3]While this is not a major restriction in the case of a single organism, when there are many organisms close together it starts to become more difficult to fulfil—a further indication that a continuous tube would be a better model

Condition $\mathfrak{P}1$ covers situations such as the model by Penrose and Hameroff (see Hameroff and Penrose, 1996) in which consciousness arises from a network of filamentary microtubules in neurons, where the condition here ensures that the timescale of this structure as a locus is set by the overall extent of the network rather than the diameter of the microtubule. Condition $\mathfrak{P}2$ is a natural maximality condition on the temporal extent of a support (the spatial extent will be constrained in section 3.2). This will enable later conditions to be defined more simply. Condition $\mathfrak{P}3$ results in the set of supports being a partially ordered set under the chronology relation $\prec$. Condition $\mathfrak{P}4$ localises the projection to its support and allows the algebra of projections P to be further localised to a subsystem over the support as is required in a hierarchical context (section 2.3.1).

The sensed universe—the object of the subjectivity of all subsystems satisfying a coherence condition—is supposed to be represented by a generalised history.

It must be stressed that a locus U included in the history does not represent a measurement (measurements being described by particular sorts of physical interaction within the Hamiltonian dynamics), but rather the selection by subjectivity from the potentialities of the current state—which may include the outcomes of a preceding measurement (see section 4.3).

3.2 The effective state

In the original formulation of the histories interpretation the probability of a given history is given by (2) and (3), which involves the Heisenberg operators P' defined globally at a single instant of time. This needs to be modified to handle operators defined with a region restricted in space but extended in time. The process of restriction in space is based on the configuration quantisation described in the Appendix (see section 2.4). In the case of extension in time, a problem emerges in that a given projection P could be defined on any of the Cauchy surfaces in U, and it could be that the results would depend on this choice (since the Schrödinger equation is not hyperbolic). We would like to be able to restrict consideration, in a natural way, to quantum states where this dependence on the Cauchy surface was not the case, and we accordingly make use here of the suggestion (Clarke, 2002) that quantum states that are subject to decoherence are essentially also those are *stable* in time. To achieve this stability we will use states averaged in time, in a sense now to be defined.

We note first that the averaging in time cannot be done using Cauchy surfaces of the region U since some of these will be strongly curved, leading to anomalies in the representations of quantum states. Instead, continuing to treat the situation of a static space-time, we use the global time parameter t and global Cauchy surfaces S_t. Recalling that the supports U_i are only a representation of the space-time tube of an organism, we can suppose each U_i chosen so that, for some (unique) global time t_c, the surface $s_c := U \cap S_{t_c}$ is a Cauchy surface of U. We set $v_c := S_{t_c} \setminus s_c$ and denote the past and future t-extents of U by t_1 and t_2 respectively.

The timelike Killing vector then generates a family of diffeomorphisms $\theta_t : S_{t_c} \to S_t$ which can be used to identify states on S_t with states on S_{t_c}. Explicitly, for each $t \in [t_1, t_2]$ we can define the sate α_t on s_c by $\alpha_t(x) = \alpha(t)(\theta_t(x))$, where $\alpha(t)$ is the state on S_t and $x \in S_{t_c}$. From section 2.4 and the Appendix, we can decompose the state α_t as an element of $\mathcal{H}(v_c) \otimes \mathcal{H}(s_c)$ which from (6) can be further decomposed as $(\mathcal{H}(v_c) \otimes \mathcal{H}'_i) \otimes \mathcal{H}_i$. The probability p of satisfaction in α_t of a proposition (projection) of the form $P = 1 \otimes \tilde{P}$ in this decomposition is

$$p = \text{Tr}(\tilde{P} R_t)$$

where

$$R_t = \text{Tr}_{\mathcal{H}(v_c) \otimes \mathcal{H}'_i} |\alpha_t\rangle\langle\alpha_t|$$

is the *reduced state* defined by restriction to s_c and $\mathcal{H}_i$. This state can then be time-averaged. To avoid artifacts arising from a sharp cut-off at the past and future boundaries of the support the average must use a weighting factor tending to zero at its past and future limits. Recalling that this is only a representation of a possible continuous tune version where no such weighting is required, the form of the function is arbitrary; but a convenient choice is weighting by the volume of the support at a given time. Thus we can define the *effective state* at the locus $(\mathcal{H}'_i, \mathcal{H}_i, U)$ as

$$\begin{aligned} \text{Eff}_{U_i, \mathcal{H}_i}(\alpha) \equiv \text{Av}_t(R_t) &:= \frac{1}{W} \int_{t_1}^{t_2} w_t R_t \, dt \\ \text{where} \quad w_t &:= \int_{S_t \cap U_i} \sqrt{g(x)} \, d^3x, \quad W = \int_{t_1}^{t_2} w_t \, dt. \end{aligned} \tag{8}$$

Note that $\text{Eff}_{U_i, \mathcal{H}_i}(\alpha)$, while closely associated to U_i, necessarily involves states that extend outside U.

Both restriction and time averaging have a "decohering" effect on the state. A state—such as is produced in a measurement-like interaction—of the form

$$\alpha = a_i \sum_i \lambda_i \otimes \mu_i$$

with respect to the decomposition $\mathcal{H}(s_c) \otimes \mathcal{H}(v_c)$ (where λ_i and μ_i belong to orthonormal bases of their respective Hilbert spaces) restricts to the purely mixed state $\sum_i |a_i|^2 |\lambda_i\rangle\langle\lambda_i|$. A state such as a sum of eigenstates of different energy, of the form

$$R_t = |\lambda\rangle\langle\lambda| \quad \text{with} \quad |\lambda\rangle = a_1 e^{i\omega_1 t} |\lambda_1\rangle + a_1 e^{i\omega_2 t} |\lambda_2\rangle$$

averages to

$$\text{Av}_t(R_t) = |a_1|^2 |\lambda_1\rangle\langle\lambda_1| + |a_2|^2 |\lambda_2\rangle\langle\lambda_2| + O(1/|\omega_1 - \omega_2|)$$

since w_t is a uniformly Lipshitz function tending to zero at t_1 and t_2. The effect of restriction is the usual decoherence phenomenon whereby measurement, by entangling with an external environmental state, produces a purely mixed state; while time averaging produces a decoherence within U for states that are superpositions of states with different energies.

3.3 Extensively Coherent Subsystems

In my preceding paper (Clarke, 2007) I argued that the most likely criterion for the subsystems which had the attribute of subjectivity (i.e. awareness), and so entered into the composition of generalised paths, were those whose quantum state had a high value of a quantity $\mathcal{C}$ which measured the extent of the internal entanglement of the state. The "state" here is the effective state defined by the members of the generalised history chronologically preceding the locus in question, using the definition of $\mathrm{Eff}_{U_i,\mathcal{H}_i}(\alpha)$ in the preceding section. The definition which follows is a slight modification of that used in (Clarke, 2007) (which I termed "coherent entanglement"), in the light of further work on applications. Adaptation has also been made because of the field theoretic framework chosen here.

Let ρ be the effective state on $\mathcal{H}(s_c)$ when s_c is a compact submanifold with boundary and let $\mathcal{P}$ be the set of all minimal surfaces[4] in s_c with connected boundary in ∂s_c. Assume for simplicity that s_c is topologically simple so that any such surface Λ divides s_c into two disjoint parts S^1_Λ and S^2_Λ, with corresponding loci $\mathfrak{L}_1$ and $\mathfrak{L}_2$ so that the locus $\mathfrak{L}_0$ of s_c is $\mathfrak{L}_1 \otimes \mathfrak{L}_2$ (see (7) and (3.1)). Let W^1_Λ, W^2_Λ be the sets of projections of the form $P = P_1 \otimes I$, $Q = I \otimes Q_2$ with P_1, Q_2 projections in the corresponding Hilbert spaces $\mathcal{H}^1_\Lambda$ and $\mathcal{H}^2_\Lambda$ respectively. (The definitions will be symmetric as between these two sides of the surface.) Write the diagonalisation of ρ (assumed for the moment to be unique up to ordering) as $\rho = \sum_{k=1}^{m} a_k |\alpha_k\rangle\, \langle\alpha_k|$, where possibly $m = \infty$. Then define

$$\begin{aligned} \mathcal{C}(\rho) &:= \min_{\Lambda\in\mathcal{P}} \min_{P\in W^1_\Lambda, Q\in W^2_\Lambda} \sum_{k=1}^{m} a_k |d(P,Q,\alpha_k)| \\ \text{where} \quad d(P,Q,\alpha_k) &:= 4\left(\langle\alpha, PQ\alpha\rangle - \langle\alpha, P\alpha\rangle\, \langle\alpha, Q\alpha\rangle\right). \end{aligned} \tag{9}$$

If the diagonalisation of ρ is not unique because of degeneracy we take the supremum over diagonalisations.

The function $d(P,Q,\alpha)$ with P and Q as above satisfies

$$d(I-P,Q,\alpha) = -d(P,Q,\alpha), \quad d(Q,P,\alpha) = d(P,Q,\alpha).$$

In (Clarke, 2007) it is shown that d lies between -1 and 1, with 0 attained if and only if α is factorisable into a tensor product of states in $\mathcal{H}^1_\Lambda$ and $\mathcal{H}^2_\Lambda$ (fully unentangled), and that $|d|$ can always attain the value 1 (fully entangled) for some α. In probability theory terms, $d(P,Q,\alpha)$ is, apart from the normalisation factor 4, the covariance of P and Q.

Coherence as measured by $\mathcal{C}$ is a variable quantity: states are more coherent the higher the value of $\mathcal{C}$ between 0 and 1. At the present exploratory stage, we can suppose chosen an arbitrary threshold $\mathcal{C}_0$ for $\mathcal{C}$. We then define an *extensively coherent state* on a locus $(U, \mathcal{H}', \mathcal{H})$ as an effective state ρ such that

1 $\mathcal{C}(\rho) > \mathcal{C}_0$.

2 There is no locus $(U^*, \mathcal{H}'_{U^*}, \mathcal{H}_{U^*})$ with $U^* \supset U$ and satisfying $\mathfrak{P}1$ such that $\mathcal{C}(\rho^*) > \mathcal{C}_0$.

The use of the decomposition into $\mathcal{H}$ and $\mathcal{H}'$ in (7), in conjunction with the use of a minimum over projections in (1), is to be noted. As a result of this, the projections which enter the generalised history are only those which involve extensively coherent states, excluding those which involve subsystems that happen to occupy the same support or which are at a different level of coarse-graining, but which are not extensively coherent. In addition, condition 2 above ensures that U captures the whole of the decoherent system, rather that an arbitrary subsystem of it. The definition thus singles out for consideration a narrow class of subsystems.

[4]The use of minimal surfaces (the curved space generalisation of a plane) rather than some other form of division is, at present, an arbitrary choice for computational convenience.

The essential point of this definition is that it places context-dependent limits on the size of supports that can enter into a history. The quantity $d(P, Q, \alpha)$ will be reduced well below the threshold for extensive coherence by the mechanisms described above (subsection 3.2) as a result of perturbations in the Hamiltonian. This decoherence will have a timescale of $\tau_{\text{deco}} = \hbar/E$, in which E, the typical size of the perturbations, might vary very roughly as KL^3 where K is a measure of the ambient energy density and L a length-scale for the support. For an extensively coherent state τ_{deco} must be greater than than the temporal extent of the support, which is of order L/c, and so (on this very crude accounting) $L < (\hbar c/K)^{1/4}$. On the other hand condition 2 ensures that the size of the support must still be of the same order of magnitude as this limit; the locus must in a sense be "marginally" coherent. We can make an estimate of this within a laboratory context. If, for example, we use for K the density of ambient thermal electromagnetic radiation (as an order of magnitude for all perturbations, not just radiative fluctuations), this length is around 10^{-5}m; if we modify the argument to filamentary structures with length-scale L and a diameter of a few tens of Ångströms, we get an order of magnitude for L of 4cm. As is well known (e.g. Hagan et al. 2002), decoherence lengths have a very wide range variation depending on the details of the mechanism and the detailed geometry, so these figures serve only as a check that the proposal is not absurdly out of line.

3.4 The Specification of Pre-Probabilities

It remains to specify the analogue of the decoherence functional (3). We will restrict to the analogue of the case where the histories $P_1, P_2, \ldots, P_n$ and $Q_1, Q_2, \ldots, Q_n$ are identical, so that we are looking for a function $\overline{p}(\mathfrak{P}_1, \mathfrak{P}_2, \ldots, \mathfrak{P}_n, \rho_0)$ with each $\mathfrak{P}_i$ satisfying $\mathfrak{P}$1–$\mathfrak{P}$4 (see (3) and following) that gives the pre-probability of the propositions $P_1, P_2, \ldots, P_n$ all being realised at their loci given an initial state ρ_0: that is, a number between 0 and 1 that will be a probability in circumstances where the history belongs to a well defined Boolean algebra. For simplicity I explicitly only give the case where $\rho_0 = \alpha_0$, a pure state.

Denote by S^i the global Cauchy surface whose restriction to U_i is s_c^i, a Cauchy surface for U_i, with time coordinate value t_i, and let α_0 be defined at time t_0. Define recursively $\alpha_i := e^{-iH(t_i - t_{i-1})/\hbar} P_{i-1}\alpha_{i-1}$ for $i = 1, 2, \ldots$ (as in the usual case—compare (4)). P_0 is defined to be the identity. Note that the commutativity of P_i, P_j when these are not $\prec$-related ensures that this procedure is not sensitive to different orderings of the projections, provided they are consistent with the partial ordering $\prec$. The pre-probability of realising P_j given the realisation of a preceding sequence $P_1, \ldots, P_{i-1}$ is then formed of two factors. The first arises from the application of P_i to the effective state and is given by

$$p_i^{(1)} := \|P_i \operatorname{Eff}_{U_i}(\alpha_i)\|^2 \tag{10}$$

using the effective state defined by (8).

The second factor implements Penrose's criterion by multiplying the probability of realisation of P_j as

$$p_i^{(2)} := (1 - e^{-s_i/\tau_i^P}), \tag{11}$$

with the following definitions.

τ_i^P is Penrose's gravitational decay time (Penrose, 2004), given by $\tau_i^P = \hbar/|E_i^*|$ where E^* is the Newtonian gravitational energy of the difference between the expectation values of the energy-densities ϵ_1 and ϵ_2 in the states $P_i\alpha_i$ and $(1 - P_i)\alpha_i$, respectively. Explicitly:

$$E^* := \int_{S^i \times S^i} \gamma(x, y)(\epsilon_1(x) - \epsilon_2(x))(\epsilon_1(y) - \epsilon_2(y)) \sqrt{h(x)}\sqrt{h(y)}\, d^3x\, d^3y \quad (12)$$

where γ is the Green function for the Laplacian on S^i and h is the determinant of the 3-metric. Note that the hypothetical mechanism that presumably underlies this is a global influence of quantum gravity which affects the whole of S_i, not just the region within U.

s_i in (11) is given by

$$s_i := \min_{j:U_j \prec U_i} (t_i - t_j)$$

with $t_j = 0$ if there are no instances in the $\min$.

The introduction of this factor eliminates the influence of a locus U_j whose temporal distance from a predecessor U_i is small in relation to a timescale defined by the gravitational difference between the complementary possible outcomes of P_j. Loosely speaking, the smaller the energetic difference between P_j and its complement, the longer it takes the organism to become "aware" of it. Without this factor it would be possible for a sequence of closely spaced U_is to produce a Zeno-like effect that could, for example, prevent the decay of all coherent quantum states, contrary to what is observed. Note also that the introduction of this factor takes the value p further from being a true probability (over the usually defined event space for histories) in that the "probabilities" of P_j being realised and of $1 - P_j$ being realised do not sum to 1: there is now the third possibility of neither of these being realised.

To summarise:

$$\overline{p}(\mathfrak{P}_1, \mathfrak{P}_2, \ldots, \mathfrak{P}_n, \rho_0) = \prod_1^n p_i^{(1)} p_i^{(2)}$$

with the ps defined by (10) and (11).

4 Consequences and future research

The essential test of any scheme of this nature must be its reduction to ordinary physics in the laboratory situation, which will be achieved if it reduces to the conventional histories interpretation in that case. Particular interest attaches here to the notion of classicality and the "emergence of a classical world". Classicality is commonly defined in terms of formal properties: perhaps something like a situation where the variance of physical variables due to quantum uncertainty is small and the equations of motion can be well approximated by classical ones. This usage is clear, for example, in the account of the early universe by Linde (2001) when he writes, "The first quantum fluctuations to freeze [during inflation] are those with large wavelengths ... At that stage one cannot call these waves quantum fluctuations anymore ... What we obtain is an inhomogeneous distribution of the classical scalar field". In the present context, however, classicality emerges in two stages. The first is constituted by the inclusion of a locus in a generalised history, which only becomes possible when the universe is sufficiently evolved (see 4.1 below). Without this inclusion there can be no awareness and so no phenomenal universe. The second is the property of a family of generalised histories to generate genuine (i.e. additive, normalised) probabilities—the property corresponding to the "consistency" (5) of that family. Neither of these coincides exactly with the more conventional concept of classicality cited above. The interest of the present theory lies in this distinction of the notion of classicality, explored in the points below.

4.1 The emergence of the universe

We can now start to indicate how the structure just presented could elucidate the problem, noted in the introduction, of the emergence of a classical universe. This can only be a preliminary suggestion because of the limitations of the treatment (section 2.3), not least of which is the need to extend the quantisation approach to a wider range of space-times, essential for cosmology. We focus on the first stage of classicality, the condition for inclusion in a generalised history.

Suppose we look in the early universe for the occurrence of a region within the horizon size, and a subsystem of the fields therein, which is extensively coherent (that is, its size is such that it is on the margin of becoming non-coherent) and is at a time after the primordial quantum state greater than the Penrose time τ_P of the previous section. For extensive coherence to hold the region must be sufficiently small; for the Penrose time to be less than the age of the universe the region must be sufficiently large. Thus there may be a critical time in the development of the universe before which these conditions cannot both be fulfilled. This marks the earliest phenomenal manifestation of the universe. If one makes an order of magnitude estimate of this time, it turns out that it must be in the early inflationary period or even earlier.

To sketch the calculation: suppose this time were in the post-inflationary universe, and consider a region U of diameter $L = \alpha L_h$, where $L_h = 1/H$ is the horizon size. The most likely candidate for $\mathcal{H}_U$ is the restriction to U of the Hilbert space for the oscillatory modes of the model Klein-Gordon field of our quantisation having wavelengths around L (say $L/2 < \lambda < 2L$). Let us suppose that the universe at that stage still has a pure, homogeneous quantum state with a superposition of states corresponding to a spectral index of 1 (Barger et al., 2003) for perturbations of the energy-density, which implies fluctuations δM in the mass of the region proportional (at a given time) to L (see Kolb and Turner, 1994, p. 332). We can normalise the amplitude of these by setting the mass perturbation at decoupling on the scale of the horizon size then to be around 3×10^{45} kg. Taking these fluctuations to be constant at earlier times in any comoving region gives a mass fluctuation in U of $K_1 \alpha t^{1/2}$, K_1 a constant, and hence (from section 3.2) a decoherence time within U of $\hbar/(K_1 \alpha t^{1/2} c^2)$. For U to be maximal subject to being extensively coherent (i.e. only marginally coherent) this must be of the order of the light-crossing time of U, that is $\alpha L/c$, and hence we find $\alpha \simeq K_2 t^{-3/4}$. We might consider imposing the condition that U cannot exceed the horizon in extent, so that $\alpha < 1$, but in fact K_2 is so small that this is never reached.

For the Penrose time we need to calculate the gravitational energy E^*_L (see (12)) of the mass fluctuations. In doing this we should take into account the entanglement of the states in $\mathcal{H}_i$ with the state over the whole of space, presumably cut off at the horizon; but it turns out that even if we only use the contribution to the gravitational energy from U itself, of order $G(\delta M)^2/L$, we still obtain a Penrose time that is many tens of magnitudes smaller than t. The conclusion is that the first stage of the quantum-classical transition is likely to occur well into the inflationary or even the quantum epoch of the universe, but without classicality either in the sense of the consistency of histories, or as regards the form of the physical equations of motion. Once its details have been determined in the more certain area of laboratory physics, this theory could thus have a considerable relevance to the physics of the very early universe.

4.2 The quantum gap

Once a moment $\mathfrak{P}$ is part of a generalised history, there remains the question of whether this results in true probabilities obeying a classical logic. Following the usual approach to the histories formalism, this is expressed in terms of a *history framework*: a pair $(\Sigma, \mathcal{L})$, where the elements of $\mathcal{L} = (\mathfrak{L}_1, \mathfrak{L}_2, \ldots, \mathfrak{L}_n)$ are loci satisfying conditions $\mathfrak{P}1$ to $\mathfrak{P}3$, and the elements of $\Sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ are such that, for each i, σ_i is a Boolean algebra of projections on $\mathcal{H}_i$ with any P_i in σ_i satisfying $\mathfrak{P}4$. For a fully classical situation we require the set of all histories belonging to the history framework to be itself a Boolean algebra over which the decoherence functional defines an additive probability. This will be achieved if, for a given collection of loci $\mathfrak{L}_i$ there is a unique algebra at each $\mathfrak{L}_i$ satisfying the conditions we have imposed. In other words, the combination of the restriction of the states at each locus through decoherence acting on the effective state, and the Penrose condition preventing loci from lying so close together that the histories become inconsistent, results in a unique classical arena. The design of laboratory procedures ensures that this is the case for the results of observations, by forming a macroscopic (i.e. highly decohered) permanent record of events by automatic means which tightly couple the recording and the observing systems. The basic hypothesis of the theory proposed here is, however, that awareness—including our own awareness—can encompass generalised histories that do not belong to classical history frameworks. In such a case the Boolean algebra that is manifested over a given locus is not determined by the physics of the situation, and the space of possible contents of awareness does not form a classical logic.

In such situations the outcome of quantum events is not merely indeterminate in the usual sense, where indeterminacy is limited to which outcome occurs out of a uniquely given Boolean algebra; but it is also undetermined which Boolean algebra might arise. This new indeterminacy constitutes an explanatory *gap* within quantum theory. Stapp (2005) has argued that it is in this gap within quantum theory that the "freedom" of living organisms emerges. In the case of human experience, our real freedom is to be found in our capacity to shape, from moment to moment, that range of options with this gap. The indeterminism within a given algebra, on the other hand, conveys the impression of freedom but not the reality (Clarke, 2001).

4.3 Organisms

Extensively coherent subsystems provide a natural implementation of the concept of an organism or "self" introduced by Mathews (1991): a subsystem of the world whose dynamics is geared to maintaining its own identity using its interactions with its surroundings. In Mathews (2003) she further argues that this gives rise to a panpsychist universe. These properties of selves leave two questions open. First, what is the physical basis for self-identity, which singles out and defines the boundaries of organisms (given that in dynamical systems theory the specification of the boundary of a subsystem is always arbitrary and pragmatic); and secondly, what in broader psychological terms, is the "psyche" element of selves? These issues are resolved in the present system, in which extensively coherent systems naturally define the extent of organisms (Ho, 1998). Mathews' perspective then suggests the identification made earlier of awareness as the characteristic of such systems.

This proposal then raises the possibility that deviations from classicality might play a significant biological role through the following mechanisms.

- Non-classical macroscopic states might be used to enhance problem-solving capacity through forms of quantum computation. This would give rise to evolutionary pressure to develop extensively coherent states and their associated awareness.
- Closely spaced moments of awareness could produce an effect formally similar to the "Zeno effect" (Stapp, 2005), which could then be used to actively maintain extensive coherence on longer timescales than the normal decoherence time (Clarke, 2007).
- Awareness may have an active role in determining the Boolean algebra at a given locus from which a proposition becomes realised. This would enable awareness to play an active role in dynamics that was complementary to, and independent of, both the Hamiltonian dynamics and the stochastic rules for the appearance of a proposition within the specified algebra.

The presence of such processes in living organisms should be experimentally detectable and will be the subject of a subsequent paper.

Appendix: Configuration space quantisation

Definition of the Hilbert space

The aim of this approach is to define the Hilbert space for quantum field theory as an L^2-space of complex functions defined on a configuration space V that is a collection of classical fields. It is essentially an expansion of the quantisation used in Simon (1974). The key feature of this approach is a particular choice of V as a dual of an auxiliary nuclear space Φ (defined below), which ensures that constructions intended to define measures (and hence L^2-spaces) really do give measures in a rigorous sense. The formalism for all this is taken from Gel'fand and Vilenkin (1964), hereafter referred to as "G&V". For simplicity we consider scalar Klein-Gordon fields.

The auxiliary space Φ is a Schwartz space of rapidly decreasing infinitely differentiable real functions on a space-like hypersurface S, with induced Riemannian metric h. A global "radial" coordinate (not necessarily with only one point where the derivative is zero) is required in order to define "rapidly decreasing", and from this a topology on Φ can be defined via a sequence of inner products (see, for example, Becnel and Sengupta (2004), extended to a Riemannian manifold), and a corresponding sequence of Hilbert spaces $\{\Phi_k\}_{k=0}^{\infty}$ of functions of increasing differentiability. Alternative choices of the radial coordinate yield equivalent topologies and the particular choice does not enter into the formalism. The intersection $\Phi := \bigcap_{k=0}^{\infty} \Phi_k$ with the topology defined by the open balls of all the norms is a *countably Hilbert space* (G&V pp.57, 59) and is, moreover, *nuclear* in the defining sense that, for any m there is an n such that the natural inclusion map $\iota : \Phi_n \to \Phi_m$ has the form

$$\iota\phi = \sum_{i=1}^{\infty}(\phi, \phi_i)_n \lambda_i \psi_i \qquad\qquad \phi \in \Phi_n$$

for some orthonormal systems $\{\phi_i\}$ and $\{\psi_i\}$, with the series $\sum_{i=0}^{\infty} \lambda_i$ converging. We then take for the configuration space $V = \Phi'$, the dual space to Φ with respect to the above topology. Note that this means that V is a space of *distributional* fields.

In order to define a configuration space representation of quantum theory in which the wave function is a complex function on V we need a measure on V with which to defined the Hilbert inner product on such functions. To this we use the construction of G&V, pp 335-345, and in particular the result (p. 340) that on the dual of a nuclear space any cylinder set measure is a measure.

We begin with an inner product B on Φ to be defined later in terms of the dynamics of the Klein-Gordon equation. Then take Ψ to be a finite dimensional subspace of Φ and let Ψ^0 be the subspace of elements of Φ' that are zero on Ψ. There is a natural isomorphism $A_\Psi : \Psi \to \Phi'/\Psi^0 : \pi \mapsto \chi$ where χ is defined by

$$(\forall \rho \in \Phi)(\ (\chi, \rho) = B(\rho, \pi)\ )$$

(parentheses denote the natural pairing of the reflexive dual spaces Φ' and Φ). For any set $U \subset \Phi'/\Psi^0$ the corresponding set $\pi_\psi^{(-1)}(U)$, where $\pi_\psi : \Phi' \to \Phi'/\Psi^0$ is the natural projection, is termed a *cylinder set*. A non-negative function on all the cylinder sets of Φ' satisfying the properties of a measure including finite additivity but not necessarily countable additivity is called a *cylinder set measure*. It is a *measure* if it is countably additive. In the present case, a cylinder set measure can be defined by taking τ to be the normalised Gaussian measure on Ψ using the inner product B (in coordinates using a basis orthonormal with respect to B, $\tau(S) = (2\pi)^{-n/2} \int_s \exp(-|x|^2/2) d^n x$), and then defining a corresponding measure τ_Ψ on cylinder sets by $\tau_\Psi(\pi_\Psi^{-1}(U)) := \tau(A_\Psi^{-1}(U))$. It can be shown that this is consistent between different Ψs and hence defines a measure on Φ'. The Hilbert space for quantisation is then $\mathcal{H} = L^2(\Phi', \mathbb{C}, \mu)$.

Dynamical structure

In order to define the dynamical structure of the Klein-Gordon field we use the standard procedure (Abraham and Marsden, 1978) of passing to the cotangent bundle W of the configuration space V and defining there a canonical symplectic form Ω. Since V is a vector space we can identify W with $V \times V'$, where V' is the dual of V in the topology defined above. Since a countably Hilbert space is reflexive (G&S p.61), $V' = \Phi'' = \Phi$ and so $W = \Phi' \times \Phi$. We denote a generic element of W by $w = (\chi, \pi)$, $\chi \in \Phi'$, $\pi \in \Phi$, as a reminder that the cotangent component π has the character of a momentum. The standard expression for the symplectic form on the cotangent bundle

$$\Omega(w_1, w_2) = \Omega((\chi_1, \pi_1), (\chi_2, \pi_2)) = (\pi_1, \chi_2) - (\pi_2, \chi_1)$$

is thus well defined.

To handle the bilinear form Ω I will introduce the following notation that will be used several times below. For any non-degenerate bilinear form Θ on a vector space

W we can define maps $\overrightarrow{\Theta} : W \to W'$ and $\overleftarrow{\Theta} : W' \to W$ by requiring

$$\begin{aligned} & (\overrightarrow{\Theta}(\theta_1), \theta_2) = \Theta(\theta_1, \theta_2) \qquad \text{for all } \theta_1, \theta_2 \\ \text{and we set} \qquad & \overleftarrow{\Theta} := \overrightarrow{\Theta}^{-1}. \end{aligned} \tag{13}$$

Where Θ is also positive definite we can extend these maps to the Hilbert space closure W^Θ of W, canonically isomorphic to its dual W'_Θ, when we denote them by $\overrightarrow{\boldsymbol{\Theta}}$ and $\overleftarrow{\boldsymbol{\Theta}}$.

In the case of the classical dynamics on V, the map $\overrightarrow{\Omega} : V \to V'$ becomes $\overrightarrow{\Omega}(\chi, \pi) = (\pi, -\chi)$ and the Poisson bracket of functions f, g on W is defined by $\{f, g\} := -(df, \overleftarrow{\Omega}(dg))$ which in the present case becomes $\{f, g\}_w = (D^1_w f, D^2_w g) - (D^1_w g, D^2_w f)$, where D^1 and D^2 indicate the components of the derivative on $W = \Phi' \times \Phi$ corresponding to Φ' and Φ respectively.[5] The canonical Poisson brackets that are normative for quantisation are obtained in the special case where f and g are taken to be coordinates on Φ' and Φ, respectively, of the form $\Lambda^1_\xi(\chi, \pi) = (\chi, \xi)$ and $\Lambda^2_\psi(\chi, \pi) = (\psi, \pi)$ for fixed $\xi \in \Phi$ and $\psi \in \Phi'$. This gives

$$\{\Lambda^1_\xi, \Lambda^2_\psi\} = (\psi, \xi); \qquad \{\Lambda^1_\xi, \Lambda^1_\zeta\} = 0; \qquad \{\Lambda^2_\phi, \Lambda^2_\psi\} = 0. \tag{14}$$

Quantisation

The standard (Dirac) quantisation procedure replaces functions f on phase space by operators $\hat{f}$ on $\mathcal{H}$ so that $[\hat{f}, \hat{g}] = i\hbar\widehat{\{f, g\}}$. From now on we set $\hbar = 1$. In the case of the canonical Poisson brackets (14), $\widehat{\Lambda^1}$ is as usual taken to be the multiplication operator

$$(\widehat{\Lambda^1_\xi}\alpha)(\chi) = (\xi, \chi)\alpha(\chi) \qquad (\text{where } \alpha : V \to \mathbb{C}).$$

To define the unbounded operator $\widehat{\Lambda^2}$ we use the map $\overleftarrow{\mathbf{B}}$ (see (13) et seq.) and set

$$(\widehat{\Lambda^2_\psi}\alpha)(\chi) = -iD^1_\chi\alpha(\psi) + \frac{i}{2}B^*_\psi(\chi)\alpha(\chi)$$

$$\text{where} \quad B^*_\psi(\chi) = (\overleftarrow{\mathbf{B}}(\psi), \chi)$$

the domain of $\widehat{\Lambda^2_\psi}$ being those function α on $\mathcal{H}$ for which the above expression defines a unique element of $\mathcal{H}$. The second term is required to give formal self-adjointness with respect to the inner product defined by the Gaussian measure.

The standard treatment of Quantum Field Theory can now be carried out using the annihilation and creation operators

$$a_\psi = \widehat{\Lambda^2_\psi} - \frac{1}{2}i\widehat{\Lambda^1_{\overleftarrow{\mathbf{B}}(\psi)}} = \widehat{\Lambda^2_\psi} - \frac{1}{2}iB^*_\psi \tag{15}$$

$$a^\dagger_\psi = \widehat{\Lambda^2_\psi} + \frac{1}{2}i\widehat{\Lambda^1_{\overleftarrow{\mathbf{B}}(\psi)}} = \widehat{\Lambda^2_\psi} + \frac{1}{2}iB^*_\psi \tag{16}$$

with a vacuum state of 1. Note that $a_\psi\alpha = -iD^1\alpha(\psi)$ so that (putting $\alpha = 1$) a_ψ annihilates the vacuum state.

[5] Derivation can be defined as a Gâteaux derivative, which then can be shown to be a Fréchet derivative with respect to all of the norms on the spaces.

From now on I will restrict to the case of static space-time, considering a hypersurface S orthogonal to the timelike Killing vector τ. As noted above (section 2.4), this is necessary for definiteness because of the ambiguities that otherwise arise. A formal expression for the classical Klein-Gordon Hamiltonian is

$$\begin{aligned} H(\chi,\pi) &= \frac{1}{2}\left(T(\pi,\pi)+Q(\chi,\chi)\right) \\ \text{where} \qquad T(\pi_1,\pi_2) &= \int_S \pi_1(x)\pi_2(x)\sqrt{h}d^3x \\ \text{and} \qquad Q(\chi_1,\chi_2) &= \int_S \left(\chi_{,i}\chi_{,}{}^{i}+m^2\chi^2\right)\sqrt{h}d^3x \end{aligned} \tag{17}$$

The problem arises even in the non-quantum case that the quadratic form Q is only defined on the subset of Φ' consisting of functions with square integrable derivatives. We therefore proceed indirectly, introducing Q via the space Φ and using annihilation and creation operators to reach the Hamiltonian. We begin by embedding Φ in $L^2(S)$. On the latter space the map

$$\gamma(\pi)(x) := \int_S G(x,y)\pi(y)\sqrt{h}d^3y,$$

where G is the kernel of the elliptic operator $K(\chi) = -\chi_{,i;}{}^{i} + m^2\chi$, is a positive bounded self adjoint operator, so that its square root $\gamma^{1/2}$ is well defined. We then define B by

$$B(\pi_1,\pi_2) = T(\pi_1,\gamma^{1/2}(\pi_2)), \tag{18}$$

which can be verified to be symmetric and non-degenerate, followed by restriction to Φ, thus building the Klein-Gordon operator into the construction of the measure already described. This is motivated by the fact that, in the case of a finite-dimensional system of oscillators, this choice reproduces the usual Schrödinger quantisation.

The quantum Hamiltonian can then be defined by an elaboration of the formula $H = \sum_{\mathbf{k}} a^\dagger_{\mathbf{k}} a_{\mathbf{k}}$ for field theory in a finite domain admitting a discrete modes labelled by $\mathbf{k}$. Define the linear map $e_\alpha : \Phi' \supset Z \to \Phi$ by

$$(e_\alpha(\psi_1),\psi_2) := (\alpha, a^\dagger_{\psi_1} a_{\psi_2}\alpha)$$

for $\alpha \in \mathcal{H}$ and Z the domain on which this is well defined. We then further restrict the domain of ψ_1 to the (canonical dual of) the L^2 Hilbert space defined by using the space-integral inner product T on Φ (see 17): set

$$E_\alpha := \overrightarrow{T} \circ e_\alpha : \Phi' \to \Phi'.$$

For α such that this operator is of trace class the hamiltonian is then defined as the quadratic form

$$H(\alpha) = \mathrm{Tr}(E_\alpha)$$

Localisation

One obtains an immediate localisation construction. Let $S_1 \subset S$ be a submanifold with boundary, and possibly with corners, with S_2 its complement. If we define Φ_1

and Φ_2 on S_1 and S_2 as for Φ, but requiring the existence of all one-sided derivatives of functions on the boundaries, then Φ decomposes according to the exact sequence

$$0 \;\rightarrow\; \Phi \;\stackrel{\kappa}{\rightarrow}\; \Phi_1 \oplus \Phi_2 \;\stackrel{\lambda}{\rightarrow}\; \Phi_\partial \;\rightarrow\; 0 \tag{19}$$

where

$$\kappa(\phi) := (\phi|_{S_1}, \phi|_{S_2}) \tag{20}$$

$$\lambda(\phi_1, \phi_2) := \phi_1|_{\partial S_1} - \phi_2|_{\partial S_1} \tag{21}$$

and Φ_∂ is the space of sections of germs of C^∞ functions on S over the piecewise linear manifold ∂S_1. Dually,

$$0 \;\leftarrow\; \Phi' \;\stackrel{\kappa'}{\leftarrow}\; \Phi'_1 \oplus \Phi'_2 \;\stackrel{\lambda'}{\leftarrow}\; \Phi'_\partial \;\leftarrow\; 0 \tag{22}$$

For studying dynamics we can define inner products B_i on Φ_i ($i = 1, 2$) as in equation (18) above, noting that we use the Green function G for the whole of S, not the Green function on each S_i with conditions on their joint boundary. This enables us to quantise fields over each of the S_i by repeating the previous construction. Consistency with quantisation on S as a whole is obtained by defining a similar form jointly linear in Φ_1 and Φ_2, viz.

$$B_{12}(\phi_1, \phi_2) = T(\phi_1, \gamma^{1/2}(\phi_2))$$

so that a bilinear form can be defined on $\Phi_1 \oplus \Phi_2$ by

$$B_+((\phi_1, \phi_2), (\rho_1, \rho_2)) = B_1(\phi_1, \rho_1) + B_2(\phi_2, \rho_2) + B_{12}(\phi_1, \rho_2) + B_{12}(\rho_1, \phi_2).$$

The gaussian measure defined on $\Phi'_1 \oplus \Phi'_2$ by this form then projects under the map κ' of (22) to the measure μ obtained by considering S as a whole.

Given a state $\alpha : \Phi' \rightarrow C$ we can localise this to S_1 by noting that the projection $p_1 : \Phi_1 \oplus \Phi_2$ gives a map $\iota_1 := \kappa \circ p_1 : \Phi \rightarrow \Phi_1$, and hence by duality a map $\iota'_1 = p'_1 \circ \kappa' : \Phi'_1 \rightarrow \Phi'$ consistent with the measures on these spaces. Thus $\alpha \circ \iota'_1$ gives a state on S_1, the restriction of the state α.

One minor complication stems from the fact that, while states on S restrict to states on submanifolds, the space of states on S is not quite the same as the tensor product of states on S_1 and S_2, reflecting the sequence (22), which arises from the fact that distributions whose support is entirely in the common boundary of S_1 and S_2 are included in the field spaces of both submanifolds.

The space Φ' is thus $(\Phi'_1 \oplus \Phi'_2)/D$, where $D = \kappa'(\Phi'_\partial)$ consists of distributions that measure the discontinuity across the boundary of S_1; and hence quantum states do not constitute the whole of the tensor product space $\mathcal{H}_1 \otimes \mathcal{H}_2$ but only the functions in this space that are constant on the equivalence classes $D + \psi$ for $\psi \in \Phi'_1 \oplus \Phi'_2$.

This is a feature of focusing on quantisation in terms of field theory: on a hierarchical approach, if we were considering, for example, the quantum theory of a dynamical system supported by an organ in the brain, it would be clearer on physical grounds what was inside and what outside the system.

References

Ralph Abraham and Jerrold E. Marsden. *Foundations of mechanics*. Benjamin/Cummings Pub. Co, Reading, Mass., 2nd edition, 1978.

V. Barger, H.-S. Lee, and D. Marfatia. WMAP and inflation. *Physics Letters B*, 565: 33–41, 2003.

Jeremy J. Becnel and Ambar N. Sengupta. The Schwartz space: A background to white noise analysis. *Louisiana State University preprint series http://www.math.lsu.edu/ preprint/*, 2004.

C. J. S. Clarke. Consciousness and non-hierarchical physics. In Philip Van Loocke, editor, *The physical nature of consciousness*. John Benjamins, Amsterdam/Philadelphia, 2001.

Chris J. S. Clarke. Entanglement and statistical independence for mixed quantum states. *Found. Phys. Lett.*, 15:495–500, 2002.

Chris J. S. Clarke. The role of quantum physics in the theory of subjective consciousness. *Mind and Body*, 5:45–81, 2007.

Matthew J. Donald. Quantum theory and the brain. *Proceedings of the Royal Society (London) Series A*, 427:43–93, 1990.

Matthew J. Donald. A mathematical characterization of the physical structure of observers. *Foundations of Physics*, 25:529–571, 1995.

Fay Dowker and Andrew Kent. On the consistent histories approach to quantum mechanics. *J. Stat. Phys.*, 82:1575, 1996.

I. M. Gel'fand and N. Ya. Vilenkin. *Generalised functions*, volume 4: Applications of Harmonic Analysis. Academic Press, New York and London, 1964.

D. Giulini, E. Joos, C. Kiefer, J. Kupsch, I.-O. Stamatescu, and H. D. Zeh. *Decoherence and the appearance of a classical world*. Springer-Verlag, Berlin and Heidelberg, 1996.

R. Griffiths. Consistent histories and the interpretation of quantum mechanics. *J. Stat. Phys.*, 36:219–272, 1984.

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

Stuart Hameroff and Roger Penrose. Orchestrated reduction of quantum coherence in brain microtubules: A model for consciousness? In S.R. Hameroff, A.W. Kaszniak, and A.C. Scott, editors, *Toward a Science of Consciousness – The First Tucson Discussions and Debates*, pages 507–540. MIT Press, 1996.

James Hartle. The quantum mechanics of cosmology. In S. Coleman, P. Hartle, T. Piran, and S. Weinberg, editors, *Quantum cosmology and baby universes*. World Scientific, 1991.

Stephen W. Hawking and George F. R. Ellis. *The Large Scale Structure of Space-Time*. Cambridge University Press, 1973.

Mae-Wan Ho. *The Rainbow and the Worm*. World Scientific, Singapore, 1998.

Christopher J Isham. Quantum logic and the histories approach to quantum theory. *J. Math. Phys*, 35:2157–85, 1994.

Christopher J. Isham, N. Linden, K. Savvida, and S. Schreckenberg. Continuous time and consistent histories. *J. Math. Phys.*, 39:1818–34, 1999.

Edward W Kolb and Michael Turner. *The early universe*. Perseus, 1994.

Andrei Linde. Inflation. In P. Murdin, editor, *Encyclopedia of Astronomy & Astrophysics*. IOP Publishing Ltd, 2001. Article 2135.

Freya Mathews. *For Love of Matter: a contemporary panpsychism*. State University of New York Press, 2003.

Freya Mathews. *The Ecological Self*. Barnes & Noble Books-Imports, 1991.

D.N. Page. Mindless sensationalism: A quantum framework for consciousness. In Q Smith and A. Jokic, editors, *Consciousness: New Philosophical Essays*. Oxford University Press, Oxford, 2001.

Roger Penrose. *The Road to Reality: a Complete Guide to the Laws of the Universe*. Jonathan Cape, 2004.

Barry Simon. *The $P(\phi)_2$ Euclidean (quantum) field theory*. Princeton Series in Physics. Princeton Univ. Press, 1974.

Henry O. Stapp. Quantum interactive dualism: An alternative to materialism. *J. Consc. Studies*, 12:43–58, 2005.

Tony Sudbery. Diese verdammte Quantenspringerei. *Stud. Hist. Phil. Mod. Phys.*, 33: 387–411, 2002.

M. S. Turner. Large-scale structure from quantum fluctuations in the early universe. *Phil. Trans. R. Soc. A*, 357:7–20, 1999.

Robert M. Wald. *Quantum Field Theory in Curved Spacetime and Black Hole Thermodynamics*. Chicago Lectures in Physics. University of Chicago Press, Chicago, 1994.

Wojciech Hubert Zurek. Decoherence, einselection, and the quantum origins of the classical. *Rev. Mod. Phys.*, 75:715, 2003.