

Conscious Events as Orchestrated Space-Time Selections

Stuart Hameroff and Roger Penrose¹

Abstract

What is consciousness? Some philosophers have contended that "qualia," or an experiential medium from which consciousness is derived, exists as a fundamental component of reality. Whitehead, for example, described the universe as being comprised of "occasions of experience." To examine this possibility scientifically, the very nature of physical reality must be re-examined. We must come to terms with the physics of space-time--as is described by Einstein's general theory of relativity--and its relation to the fundamental theory of matter--as described by quantum theory. This leads us to employ a new physics of *objective reduction*: "OR" which appeals to a form of quantum gravity to provide a useful description of fundamental processes at the quantum/classical borderline. Within the OR scheme, we consider that consciousness occurs if an appropriately organized system is able to develop and maintain quantum coherent superposition until a specific "objective" criterion (a threshold related to quantum gravity) is reached; the coherent system then self-reduces (objective reduction: OR). We contend that this type of objective self-collapse introduces non-computability, an essential feature of consciousness. OR is taken as an instantaneous event--the climax of a self-organizing process in fundamental space-time--and a candidate for a conscious Whitehead "occasion" of experience. How could an OR process occur in the brain, be coupled to neural activities, and account for other features of consciousness? We nominate an OR process with the requisite characteristics to be occurring in cytoskeletal microtubules within the brain's neurons. In this model, quantum-superposed states develop in microtubule subunit proteins ("tubulins"), remain coherent and recruit more superposed tubulins until a mass-time-energy threshold (related to quantum gravity) is reached. At that point, self-collapse, or objective reduction (OR) abruptly occurs. We equate the pre-reduction, coherent superposition ("quantum computing") phase with pre-conscious processes, and each instantaneous (and non-computable) OR, or self-collapse, with a discrete conscious event. Sequences of OR events give rise to a "stream" of consciousness. Microtubule-associated-proteins can "tune" the quantum oscillations of the coherent superposed states; the OR is thus self-organized, or "orchestrated" ("Orch OR"). Each Orch OR event selects (non-computably) microtubule subunit states which regulate synaptic/neural functions using classical signaling. The quantum gravity threshold for self-collapse is relevant to consciousness, according to our arguments, because macroscopic superposed quantum states each have their own space-time geometries. These geometries are also superposed, and in some way "separated," but when sufficiently separated, the superposition of space-time geometries becomes significantly unstable, and reduce to a single universe state. Quantum gravity determines the limits of the instability; we contend that the actual choice of state made by Nature is non-computable. Thus each Orch OR event is a self-selection of space-time geometry, coupled to the brain through microtubules and other biomolecules.

If conscious experience is intimately connected with the very physics underlying space-time structure, then Orch OR in microtubules indeed provides us with a completely new and uniquely promising perspective on the hard problem of consciousness.

Key Words: consciousness, quantum theory, quantum gravity, space-time, collapse, objective reduction, microtubules

NeuroQuantology 2003;1: 10-35

¹ Corresponding author: Stuart Hameroff, Department of Anesthesiology, 1501 North Campbell Avenue. Tucson, Arizona 85724 USA. E-mail: hameroff@u.arizona.edu

Introduction: Self-Selection in an Experiential Medium?

The "hard problem" of incorporating the phenomenon of consciousness into a scientific world-view involves finding scientific explanations of qualia, or the subjective experience of mental states (Chalmers, 1995; 1996). On this, reductionist science is still at sea. Why do we have an inner life, and what exactly is it?

One set of philosophical positions, addressing the hard problem, views consciousness as a fundamental component of physical reality. For example an extreme view - "panpsychism" - is that consciousness is a quality of all matter: atoms and their subatomic components having elements of consciousness (e.g. Spinoza, 1677; Rensch, 1960). "Mentalists" such as Leibniz and Whitehead (e.g. 1929) contended that systems ordinarily considered to be physical are constructed in some sense from mental entities. Bertrand Russell (1954) described "neutral monism" in which a common underlying entity, neither physical nor mental, gave rise to both. Recently Stubenberg (1996) has claimed that qualia are that common entity. In monistic idealism, matter and mind arise from consciousness - the fundamental constituent of reality (e.g. Goswami, 1993). Wheeler (1990) has suggested that information is fundamental to the physics of the universe. From this, Chalmers (1995; 1996) proposes a double-aspect theory in which information has both physical and experiential aspects.

Among these positions, the philosophy of Alfred North Whitehead (1929; 1933) may be most directly applicable. Whitehead describes the ultimate concrete entities in the cosmos as being actual "occasions of experience," each bearing a quality akin to "feeling." Whitehead construes "experience" broadly - in a manner consistent with panpsychism - so that even "temporal events in the career of an electron have a kind of 'protomentality'." Whitehead's view may be considered to differ from panpsychism, however, in that his discrete 'occasions of experience' can be taken to be related to "quantum events" (Shimony, 1993). In the standard descriptions of quantum mechanics, randomness occurs in the events described as quantum state reductions--these being events which appear to take place when a quantum-level process gets magnified to a macroscopic scale.

Quantum state reduction (here denoted by the letter R; cf. Penrose, 1989, 1994) is the random procedure that is adopted by physicists in their descriptions of the quantum measurement process. It is still a highly controversial matter whether R is to be taken as a "real" physical process, or whether it is some kind of illusion and not to be regarded as a fundamental ingredient of the behavior of Nature. Our position is to take R to be indeed real--or, rather to regard it as a close approximation to an objectively real process OR (objective reduction), which is to be a non-computable process instead of merely a random one (see Penrose 1989; 1994). In almost all physical situations, OR would come about in situations in which the random effects of the environment dominate, so OR would be virtually indistinguishable from the random R procedure that is normally adopted by quantum theorists. However, when the quantum system under consideration remains coherent and well isolated from its environment, then it becomes possible for its state to collapse spontaneously, in accordance with the OR scheme we adopt, and to behave in non-computable rather

than random ways. Moreover, this OR scheme intimately involves the geometry of the physical universe at its deepest levels.

Our viewpoint is to regard experiential phenomena as also inseparable from the physical universe, and in fact to be deeply connected with the very laws which govern the physical universe. The connection is so deep, however, that we perceive only glimmerings of it in our present day physics. One of these glimmerings, we contend, is a necessary non-computability in conscious thought processes; and we argue that this non-computability must also be inherent in the phenomenon of quantum state *self-reduction*--the "objective reduction" (OR) referred to above. This is the main thread of argument in *Shadows of the Mind* (Penrose, 1994). The argument that conscious thought, whatever other attributes it may also have, is non-computable (as follows most powerfully from certain deductions from Gödel's incompleteness theorem) grabs hold of one tiny but extremely valuable point. This means that at least some conscious states cannot be derived from previous states by an algorithmic process - a property which distinguishes human and other animal minds from computers. Non-computability *per se* does not directly address the 'hard problem' of the nature of experience, but it is a clue to the kind of physical activity that lies behind it. This points to OR, an underlying physical action of a completely different character from that which seems to underlie non-conscious activity. Following this clue with sensitivity and patience should ultimately lead to real progress towards understanding mental phenomena in their inward manifestations as well as outward.

In the OR description, consciousness occurs if an organized quantum system is able to isolate and sustain coherent superposition until its quantum gravity threshold for space-time separation is met; it then *self-reduces* (non-computably). For consciousness to occur, *self-reduction* is essential, as opposed to reduction being triggered by the system's random environment. (In the latter case, the reduction would itself be effectively random and would lack useful non-computability, being unsuitable for direct involvement in consciousness.) We take the *self-reduction* to be an instantaneous event -- the climax of a *self-organizing* process fundamental to the structure of space-time - and apparently consistent with a Whitehead "occasion of experience."

As OR could, in principle, occur ubiquitously within many types of inanimate media, it may seem to imply a form of 'panpsychism' (in which individual electrons, for example, possess an experiential quality). However according to the principles of OR (as expounded in Penrose 1994; 1996), a single superposed electron would spontaneously reduce its state (assuming it could maintain isolation) only once in a period much longer than the present age of the universe. Only large collections of particles acting coherently in a single macroscopic quantum state could possibly sustain isolation and support coherent superposition in a time frame brief enough to be relevant to our consciousness. Thus only very special circumstances could support consciousness:

1. High degree of coherence of a quantum state - a collective mass of particles in superposition for a time period long enough to reach threshold, and brief enough to be useful in thought processes.

2. Ability for the OR process to be at least transiently isolated from a 'noisy'

environment until the spontaneous state reduction takes place. This isolation is required so that reduction is not simply random. Mass movement in the environment which entangles with the quantum state would effect a random (not non-computable) reduction.

3. *Cascades of ORs to give a "stream" of consciousness, and huge numbers of OR events taking place during the course of a lifetime.*

By reaching quantum gravity threshold, each OR event has a fundamental bearing on space-time geometry. One could say that a cascade of OR events charts an actual course of physical space-time geometry selections.

It may seem surprising that quantum gravity effects could plausibly have relevance at the physical scales relevant to brain processes. For quantum gravity is normally viewed as having only absurdly tiny influences at ordinary dimensions. However, we shall show later that this is not the case, and the scales determined by basic quantum gravity principles are indeed those that are relevant for conscious brain processes.

We must ask how such an OR process could actually occur in the brain. How could it be coupled to neural activities at a high rate of information exchange; how could it account for preconscious to conscious transitions, have spatial and temporal binding, and both simultaneity and time flow?

We here nominate an OR process with the requisite characteristics occurring in cytoskeletal microtubules within the brain's neurons. In our model, microtubule-associated proteins "tune" the quantum oscillations leading to OR; we thus term the process "orchestrated objective reduction" (Orch OR).

Space-Time: Quantum Theory and Einstein's Gravity

Quantum theory describes the extraordinary behavior of the matter and energy which comprise our universe at a fundamental level. At the root of quantum theory is the wave/particle duality of atoms, molecules and their constituent particles. A quantum system such as an atom or sub-atomic particle which remains isolated from its environment behaves as a "wave of possibilities" and exists in a coherent complex-number valued "superposition" of many possible states. The behavior of such wave-like, quantum-level objects can be satisfactorily described in terms of a state vector which evolves deterministically according to the Schrödinger equation (unitary evolution), denoted by U .

Somehow, quantum microlevel superpositions lead to unsuperposed stable structures in our macro-world. In a transition known as wave function collapse, or reduction (R), the quantum wave to alternative possibilities reduces to a single macroscopic reality, an "eigenstate" of some appropriate operator. (This would be just one out of many possible alternative eigenstates relevant to the quantum operator.) This process is invoked in the description of a macroscopic measurement, when effects are magnified from the small, quantum scale to the large, classical scale.

According to conventional quantum theory (as part of the standard "Copenhagen interpretation"), each choice of eigenstate is entirely random, weighted according to a probability value that can be calculated from the previous state according to the precise procedures of quantum formalism. This probabilistic

ingredient was a feature with which Einstein, among others, expressed displeasure: "You believe in a God who plays dice and I in complete law and order"(from a letter to Max Born). Penrose (1989; 1994) has contended that, at a deeper level of description, the choices may more accurately arise as a result of some presently unknown "non-computational" mathematical/physical (i.e., "Platonic realm") theory, that is they cannot be deduced algorithmically. Penrose argues that such non-computability is essential to consciousness, because (at least some) conscious mental activity is unattainable by computers.

It can be argued that present-day physics has no clear explanation for the cause and occurrence of wave function collapse R. Experimental and theoretical evidence through the 1930's led quantum physicists (such as Schrödinger, Heisenberg, Dirac, von Neumann and others) to postulate that quantum-coherent superpositions persist indefinitely in time, and would, in principle be maintained from the micro to macro levels. Or perhaps they would persist until conscious observation collapses, or reduces, the wave function (subjective reduction, or "SR"). Accordingly, even macroscopic objects, if unobserved, could remain superposed. To illustrate the apparent absurdity of this notion, Erwin Schrödinger (e.g. 1935) described his now-famous "cat in a box" being simultaneously both dead and alive until the box was opened and the cat observed.

As a counter to this unsettling prospect, various new physical schemes for collapse according to objective criteria (objective reduction - "OR") have recently been proposed. According to such a scheme, the growth and persistence of superposed states could reach a critical threshold, at which collapse, or OR rapidly occurs (e.g. Pearle, 1989; Ghirardi et al, 1986). Some such schemes are based specifically on gravitational effects mediating OR (e.g. Károlyházy, 1986; Diósi, 1989; Ghirardi et al., 1990; Penrose, 1989; 1994; Pearle and Squires, 1994; Percival, 1995).

Table 1. categorizes types of reduction.

Context	Cause of Collapse (Reduction)	Description	Acronym
Quantum coherent superposition	No collapse	Evolution of the wave function (Schrödinger equation)	U
Conventional quantum theory (Copenhagen interpretation)	Environmental entanglement, Measurement, Conscious observation	Reduction; Subjective reduction	R SR
New physics (Penrose, 1994)	Self-collapse -quantum gravity induced (Penrose, Diósi, etc)	Objective reduction	OR
Consciousness (present paper)	Self-collapse, quantum gravity threshold in microtubules orchestrated by MAPs etc	Orchestrated objective reduction	Orch OR

The physical phenomenon of gravity, described to a high degree of accuracy by Isaac Newton's mathematics in 1687, has played a key role in scientific understanding. However, in 1915, Einstein created a major revolution in our scientific world-view. According to Einstein's theory, gravity plays a unique role in physics for several reasons (cf. Penrose, 1994). Most particularly, these are:

1. Gravity is the only physical quality which influences causal relationships between space-time events.

2. Gravitational force has no local reality, as it can be eliminated by a change in space-time coordinates; instead, gravitational tidal effects provide a curvature for the very space-time in which all other particles and forces are contained.

It follows from this that gravity cannot be regarded as some kind of "emergent phenomenon," secondary to other physical effects, but is a "fundamental component" of physical reality.

There are strong arguments (e.g. Penrose, 1987; 1995) to suggest that the appropriate union of general relativity (Einstein's theory of gravity) with quantum mechanics - a union often referred to as "quantum gravity" - will lead to a significant change in *both* quantum theory and general relativity, and, when the correct theory is found, will yield a profoundly *new* understanding of physical reality. And although gravitational *forces* between objects are exceedingly weak (feebler than, for example, electrical forces by some 40 orders of magnitude), there are significant reasons for believing that gravity has a fundamental influence on the behavior of quantum systems as they evolve from the micro to the macro levels. The appropriate union of quantum gravity with biology, or at least with advanced biological nervous systems, may yield a profoundly new understanding of consciousness.

CURVED SPACE-TIME SUPERPOSITIONS AND OBJECTIVE REDUCTION ("OR")

According to modern accepted physical pictures, reality is rooted in 3-dimensional space and a 1-dimensional time, combined together into a 4-dimensional space-time. This space-time is slightly curved, in accordance with Einstein's general theory of relativity, in a way which encodes the gravitational fields of all distributions of mass density. Each mass density effects a space-time curvature, albeit tiny.

This is the standard picture according to *classical* physics. On the other hand, when *quantum* systems have been considered by physicists, this mass-induced tiny curvature in the structure of space-time has been almost invariably ignored, gravitational effects having been assumed to be totally insignificant for normal problems in which quantum theory is important. Surprising as it may seem, however, such tiny differences in space-time structure *can* have large effects, for they entail subtle but fundamental influences on the very rules of quantum mechanics.

Superposed quantum states for which the respective mass distributions differ significantly from one another will have space-time geometries which correspondingly differ. Thus, according to standard quantum theory, the superposed state would have to involve a quantum superposition of these differing space-times. In the absence of a coherent theory of quantum gravity there is no accepted way of handling such a superposition. Indeed the basic principles of Einstein's general relativity begin to come into profound conflict with those of quantum mechanics (cf. Penrose, 1996).

Nevertheless, various tentative procedures have been put forward in attempts to describe such a superposition. Of particular relevance to our present proposals are the suggestions of certain authors (i.e., Karolyhazy, 1996; 1974; Karolyhazy et al., 1986; Kibble, 1991, Diósi, 1989; Ghirardi et al, 1990; Pearle and Squires, 1995; Percival, 1995; Penrose, 1993; 1994; 1996) that it is at this point that an objective quantum state reduction (**OR**) ought to occur, and the rate or timescale of this process can be calculated from basic quantum gravity considerations. These particular proposals differ in certain detailed respects, and for definiteness we shall follow the specific suggestions made in Penrose (1994; 1996). Accordingly, the quantum superposition of significantly differing space-times is unstable, with a lifetime given by that timescale. Such a superposed state will decay - or "reduce" - into a single universe state, which is one or the other of the space-time geometries involved in that superposition.

Whereas such an **OR** action is not a generally recognized part of the normal quantum-mechanical procedures, there is no plausible or clear-cut alternative that standard quantum theory has to offer. This **OR** procedure avoids the need for "multiple universes" (cf. Everett, 1957; Wheeler, 1957, for example). There is no agreement, among quantum gravity experts, about how else to address this problem. For the purposes of the present article, it will be assumed that a gravitationally induced **OR** action is indeed the correct resolution of this fundamental conundrum.

Figure 1 (adapted from Penrose, 1994, p. 338) schematically illustrates the way in which space-time structure can be affected when two macroscopically different mass distributions take part in a quantum superposition. Each mass distribution gives rise to a separate space-time, the two differing slightly in their curvatures. So long as the two distributions remain in quantum superposition, we must consider that the two space-times remain in superposition. Since, according to the principles of general relativity, there is no natural way to identify the points of one space-time with corresponding points of the other, we have to consider the two as separated from one another in some sense, resulting in a kind of "blister" where the space-time bifurcates.

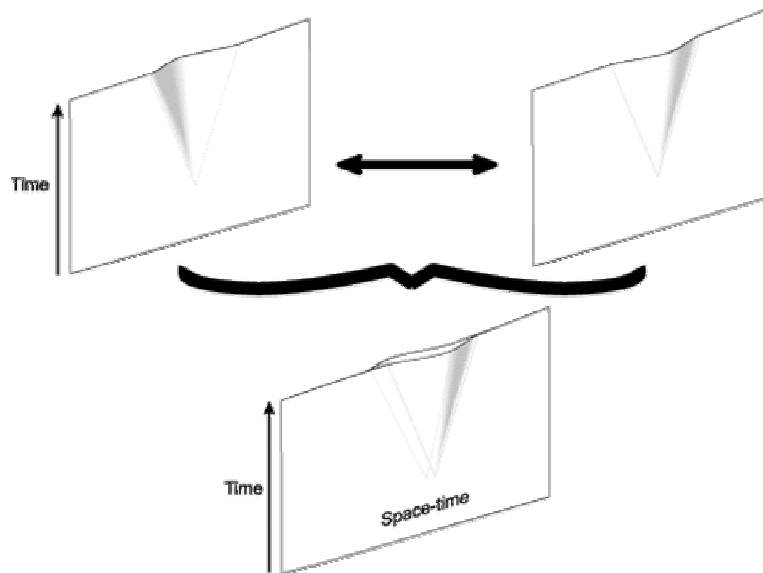


Figure 1. Quantum coherent superposition represented as a separation of space-time. In the lowest of the three diagrams, a bifurcating space-time is depicted as the union ("glued together version") of the two alternative space-time histories that are depicted at the top of the Figure. The bifurcating space-time diagram illustrates two alternative mass distributions actually in quantum superposition, whereas the top two diagrams illustrate the two individual alternatives which take part in the superposition (adapted from Penrose, 1994 - p. 338).

A bifurcating space-time is depicted in the lowest of the three diagrams, this being the union ("glued together version") of the two alternative space-time histories that are depicted at the top of Figure 1. The initial part of each space-time is at the lower end of each individual space-time diagram. The bottom space-time diagram (the bifurcating one) illustrates two alternative mass distributions actually in quantum superposition, whereas the top two illustrate the two individual alternatives which take part in the superposition. The combined space-time describes a superposition in which the alternative locations of a mass move gradually away from each other as we proceed in the upward direction in the diagram. Quantum-mechanically (so long as **OR** has not taken place), we must think of the "physical reality" of this situation as being illustrated as an actual superposition of these two slightly differing space-time manifolds, as indicated in the bottom diagram. As soon as **OR** has occurred, one of the two individual space-times takes over, as depicted as one of the two sheets of the bifurcation. For clarity only, the bifurcating parts of these two sheets are illustrated as being one convex and the other concave. Of course there is additional artistic license involved in drawing the space-time sheets as 2-dimensional, whereas the actual space-time constituents are 4-dimensional. Moreover, there is no significance to be attached to the imagined "3-dimensional space" within which the space-time sheets

seem to be residing. There is no "actual" higher dimensional space there, the "intrinsic geometry" of the bifurcating space-time being all that has physical significance. When the "separation" of the two space-time sheets reaches a critical amount, one of the two sheets "dies" - in accordance with the **OR** criterion - the other being the one that persists in physical reality. The quantum state thus reduces (**OR**), by choosing between either the "concave" or "convex" space-time of Figure 1.

It should be made clear that this measure of separation is only very schematically illustrated as the "distance" between the two sheets in the lower diagram in Figure 1. As remarked above, there is no physically existing "ambient higher dimensional space" inside which the two sheets reside. The degree of separation between the space-time sheets is a more abstract mathematical thing; it would be more appropriately described in terms of a *symplectic measure* on the space of 4-dimensional metrics (cf. Penrose, 1993) - but the details (and difficulties) of this will not be important for us here. It may be noted, however, that this separation is a space-time separation, not just a spatial one. Thus the *time* of separation contributes as well as the spatial displacement. Roughly speaking, it is the product of the temporal separation T with the spatial separation S that measures the overall degree of separation, and **OR** takes place when this overall separation reaches the critical amount. [This critical amount would be of the order of unity, in absolute units, for which the Planck-Dirac constant h (actually " \hbar ": Planck's constant over 2π), the gravitational constant G , and the velocity of light c , all take the value unity, cf. Penrose, 1994 - pp. 337-339.] Thus for small S , the lifetime T of the superposed state will be large; on the other hand, if S is large, then T will be small. To calculate S , we compute (in the Newtonian limit of weak gravitational fields) the gravitational self-energy E of the difference between the mass distributions of the two superposed states. (That is, one mass distribution counts positively and the other, negatively; see Penrose, 1994; 1995.) The quantity S is then given by:

$$S = E$$

Thus

$$E = h / T$$

Schematically, since S represents three dimensions of displacement rather than the one dimension involved in T , we can imagine that this displacement is shared equally between each of these three dimensions of space - and this is what has been depicted in Figure 3 (below). However, it should be emphasized that this is for pictorial purposes only, the appropriate rule being the one given above. These 2 equations relate the mass distribution, time of coherence, and space-time separation for a given **OR** event. If, as some philosophers contend, experience is contained in space-time, **OR** events are *self-organizing* processes in that experiential medium, and a candidate for consciousness.

But where in the brain, and how, could coherent superposition and **OR** occur? A number of sites and various types of quantum interactions have been proposed. We strongly favor microtubules as an important ingredient, however various organelles and biomolecular structures including clathrins, myelin (glial cells), pre-synaptic vesicular grids (Beck and Eccles, 1992) and neural membrane proteins (Marshall, 1989) might also participate.

Microtubules

Properties of brain structures suitable for quantum coherent superposition, **OR** and relevant to consciousness might include: 1) high prevalence; 2) functional importance (for example regulating neural connectivity and synaptic function); 3) periodic, crystal-like lattice dipole structure with long range order; 4) ability to be transiently isolated from external interaction/observation; 5) functionally coupled to quantum-level events; 6) hollow, cylindrical (possible wave guide); and 7) suitable for information processing. Membranes, membrane proteins, synapses, DNA and other types of structures have some, but not all, of these characteristics. Cytoskeletal microtubules appear to qualify in all respect.

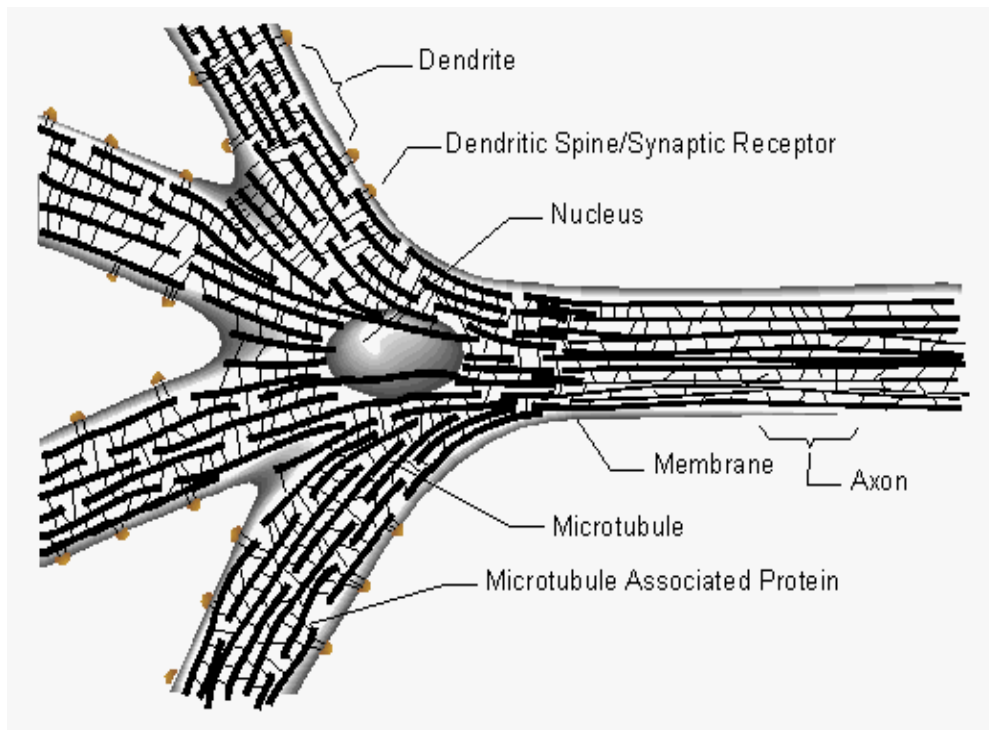


Figure 2. Schematic of central region of neuron (distal axon and dendrites not shown), showing parallel arrayed microtubules interconnected by MAPs. Microtubules in axons are lengthy and continuous, whereas in dendrites they are interrupted and of mixed polarity. Linking proteins connect microtubules to membrane proteins including receptors on dendritic spines.

Interiors of living cells, including the brain's neurons, are spatially and dynamically organized by *self-assembling* protein networks: the cytoskeleton. Within neurons, the cytoskeleton establishes neuronal form, and maintains and regulates synaptic connections. Its major components are microtubules, hollow cylindrical polymers of individual proteins known as tubulin. Microtubules ("MTs") are interconnected by linking proteins (microtubule-associated proteins: "MAPs") to other microtubules and cell structures to form cytoskeletal lattice networks (Figure 2).

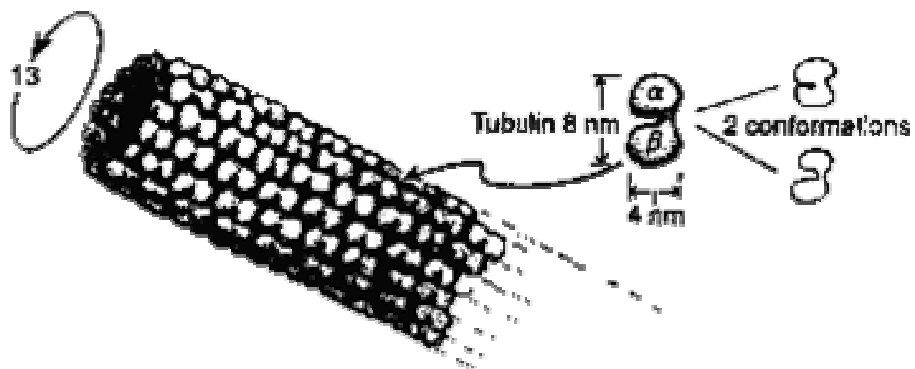


Figure 3. Microtubule structure: a hollow tube of 25 nanometers diameter, consisting of 13 columns of tubulin dimers. Each tubulin molecule is capable of (at least) two conformations. (Reprinted with permission from Penrose, 1994, p. 359.)

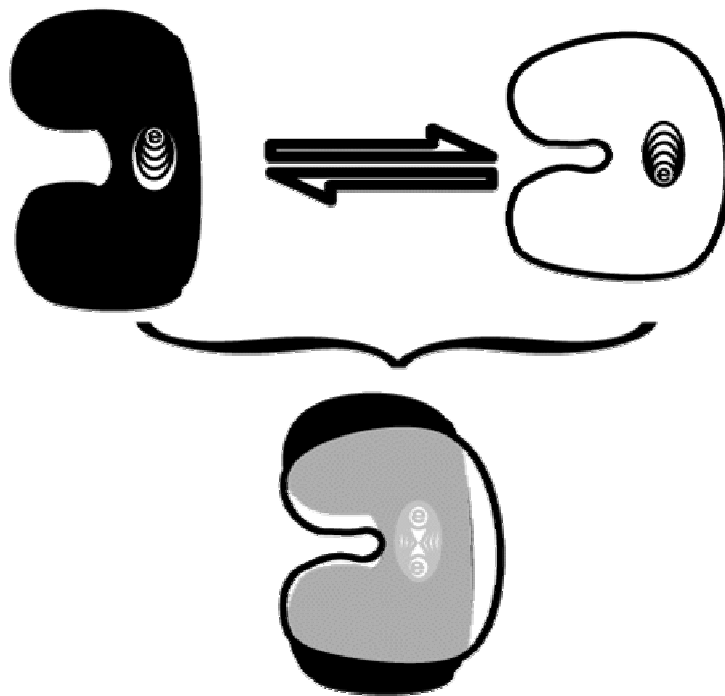


Figure 4. Top: Two states of tubulin in which a single quantum event (electron localization) within a central hydrophobic pocket is coupled to a global protein conformation. Switching between the two states can occur on the order of nanoseconds to picoseconds. Bottom: Tubulin in quantum coherent superposition of both states.

MTs are hollow cylinders 25 nanometers (nm) in diameter whose lengths vary and may be quite long within some nerve axons. MT cylinder walls are comprised of 13 longitudinal protofilaments which are each a series of subunit proteins known as tubulin (Figure 3). Each tubulin subunit is a polar, 8 nm dimer which consists of two slightly different 4 nm monomers (alpha and beta tubulin - Figure 4). Tubulin dimers are dipoles, with surplus negative charges localized toward monomers (DeBrabander, 1982), and within MTs are arranged in a hexagonal lattice which is slightly twisted, resulting in helical pathways which repeat every 3, 5, 8 and other numbers of rows. Traditionally viewed as the cell's "bone-like" scaffolding, microtubules and other cytoskeletal structures also appear to fill communicative and information processing roles. Numerous types of studies link the cytoskeleton to cognitive processes (for review, cf. Hameroff and Penrose, 1996). Theoretical models and simulations suggest how conformational states of tubulins within microtubule lattices can interact with neighboring tubulins to represent, propagate and process information as in molecular-level "cellular automata," or "spin glass" type computing systems (Figure 5; e.g. Hameroff and Watt, 1982; Rasmussen et al, 1990; Tuszynski et al, 1995).

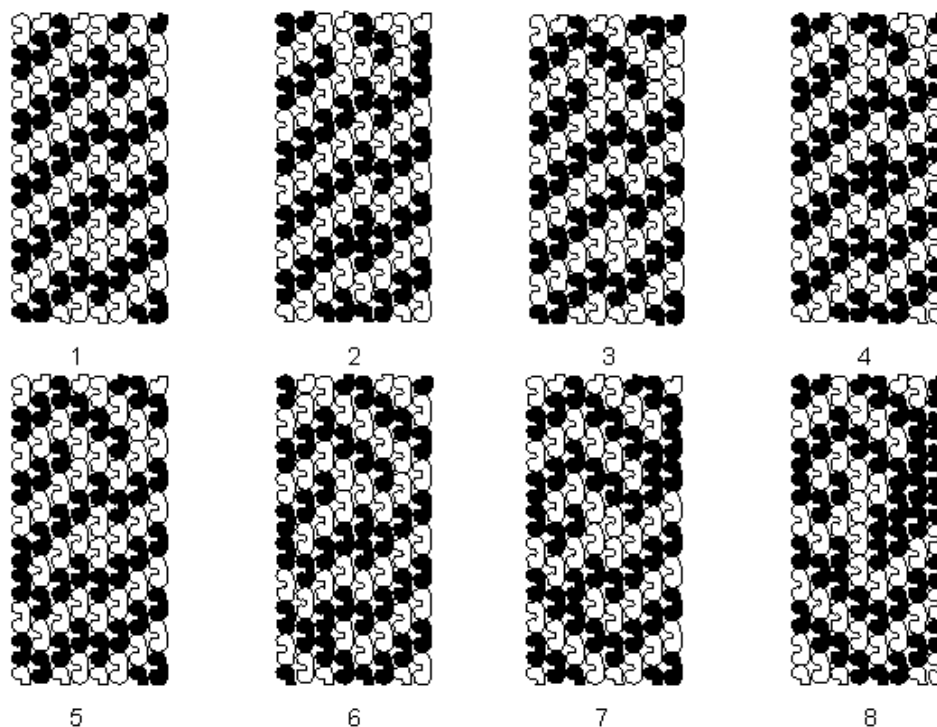


Figure 5. Microtubule automaton simulation (from Rasmussen et al, 1990). Black and white tubulins correspond to states shown in Figure 2. Eight nanosecond time steps of a segment of one microtubule are shown in "classical computing" mode in which patterns move, evolve, interact and lead to emergence of new patterns.

In Hameroff and Penrose (1996; and in summary form, Penrose and Hameroff, 1995), we present a model linking microtubules to consciousness, using quantum theory as viewed in the particular "realistic" way that is described in *Shadows of the Mind* (Penrose, 1994).

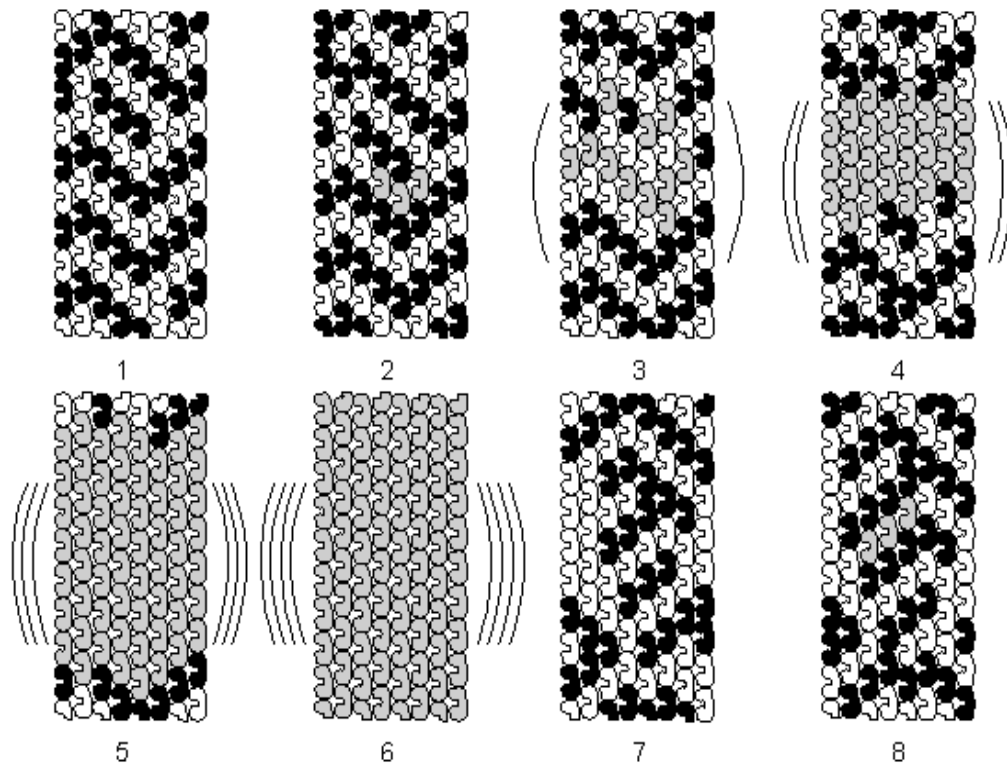


Figure 6. Microtubule automaton sequence simulation in which classical computing (step 1) leads to emergence of quantum coherent superposition (steps 2-6) in certain (gray) tubulins due to pattern resonance. Step 6 (in coherence with other microtubule tubulins) meets critical threshold related to quantum gravity for self-collapse (Orch OR). Consciousness (Orch OR) occurs in the step 6 to 7 transition. Step 7 represents the eigenstate of mass distribution of the collapse which evolves by classical computing automata to regulate neural function. Quantum coherence begins to re-emerge in step 8.

In our model, quantum coherence emerges, and is isolated, in brain microtubules until the differences in mass-energy distribution among superposed tubulin states reach the threshold of instability described above, related to quantum gravity (Figure 6). The resultant *self-collapse* (OR), considered to be a time-irreversible process, creates an instantaneous "now" event. Sequences of such events create a flow of time, and consciousness (Figures 7 and 8).

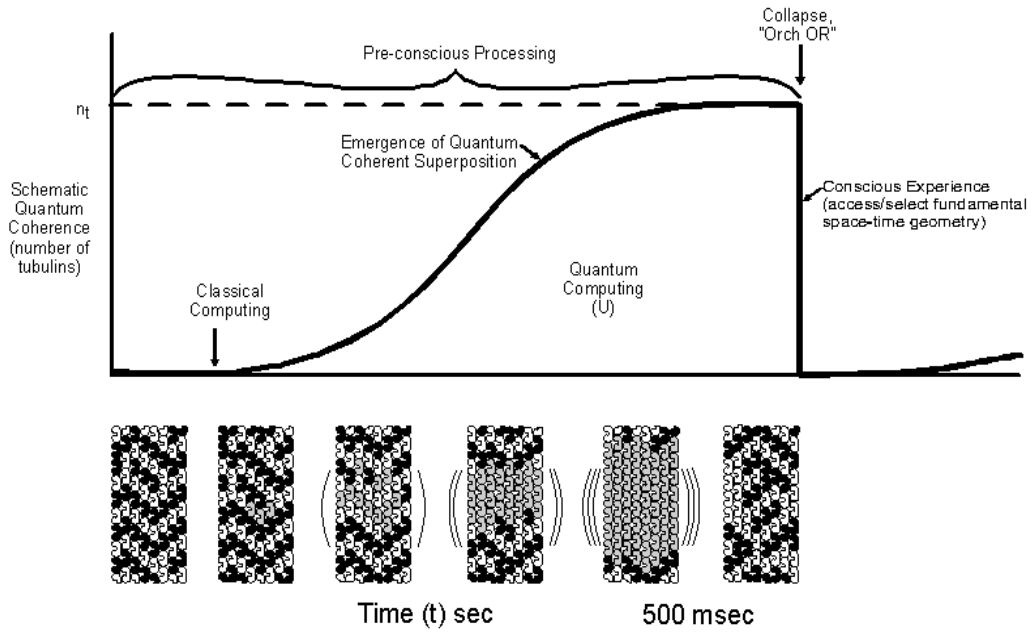


Figure 7. Schematic graph of proposed quantum coherence (number of tubulins) emerging vs time in microtubules. 500 milliseconds is time for pre-conscious processing (e.g. Libet, 1979). Area under curve connects mass-energy differences with collapse time in accordance with gravitational OR. This degree of coherent superposition of differing space-time geometries leads to abrupt quantum classical reduction ("self-collapse" or "orchestrated objective reduction: Orch OR").

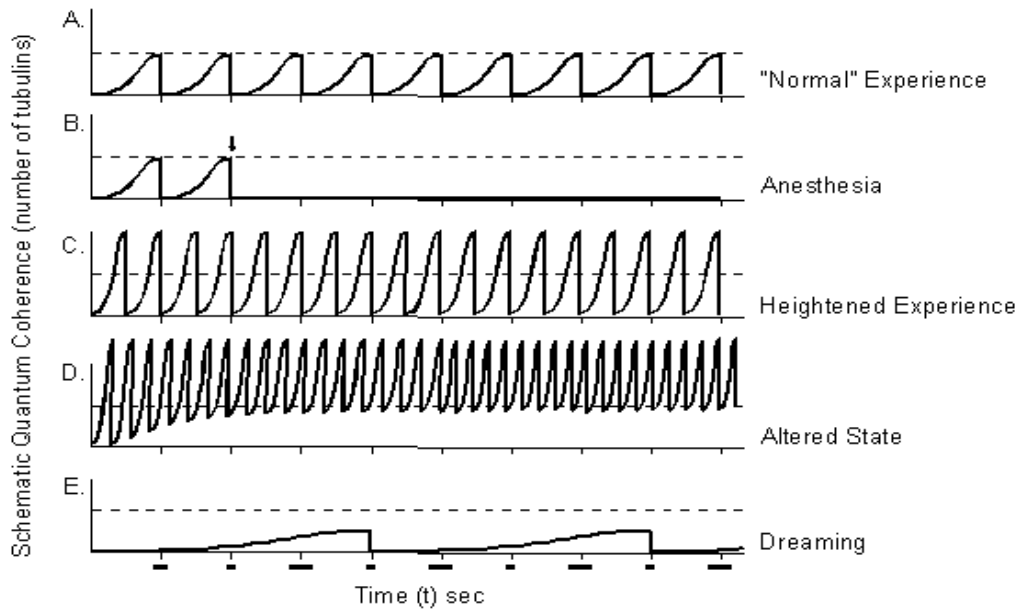


Figure 8. Quantum coherence in microtubules schematically graphed on longer time scale for 5 different states related to consciousness. Area under each curve equivalent in all cases. A. Normal experience: as

in Figure 8. B. Anesthesia: anesthetics bind in hydrophobic pockets and prevent quantum delocalizability and coherent superposition (e.g. Louri and Hameroff, 1996). C. Heightened Experience: increased sensory experience input (for example) increases rate of emergence of quantum coherent superposition. Orch Or threshold is reached faster (e.g. 250 msec) and Orch Or frequency is doubled. D. Altered State: even greater rate of emergence of quantum coherence due to sensory input and other factors promoting quantum state (e.g. meditation, psychedelic drug etc.). Predisposition to quantum state results in baseline shift and only partial collapse so that conscious experience merges with normally sub-conscious quantum computing mode. E. Dreaming: prolonged quantum coherence time.

We envisage that attachments of MAPs on microtubules "tune" quantum oscillations, and "orchestrate" possible collapse outcomes (Figure 9). Thus we term the particular *self-organizing OR* occurring in MAP-connected microtubules, and relevant to consciousness, orchestrated objective reduction ("**Orch OR**"). **Orch OR** events are thus *self-selecting* processes in fundamental space-time geometry. If experience is truly a component of fundamental space-time, **Orch OR** may begin to explain the "hard problem" of consciousness.

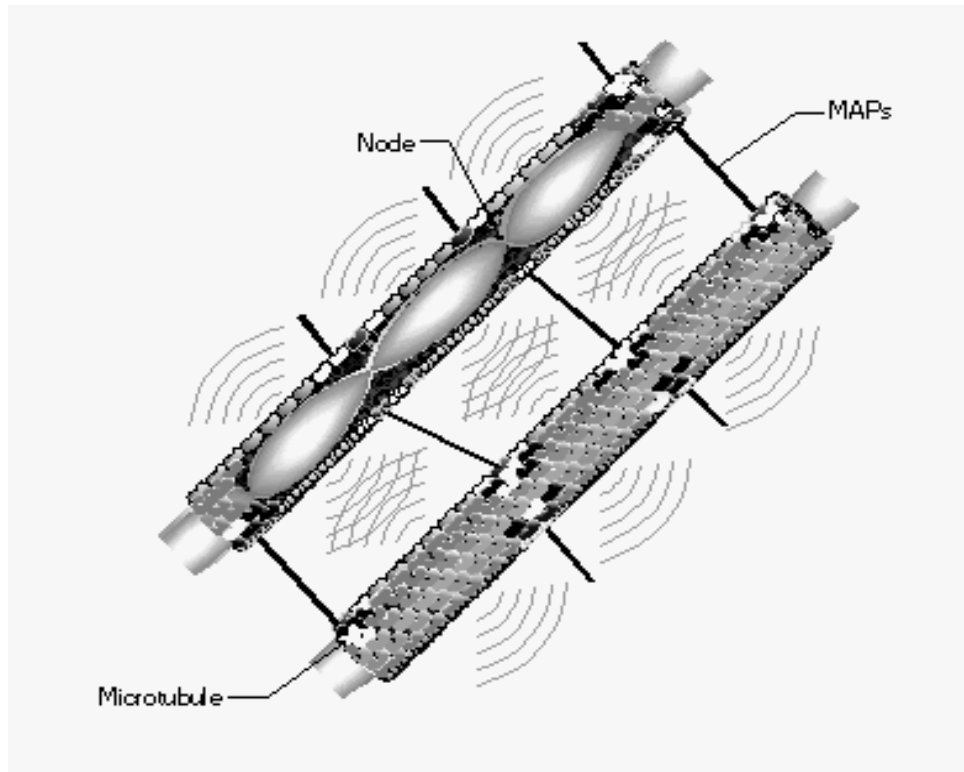


Figure 9. Quantum coherence in microtubules. Having emerged from resonance in classical automaton patterns, quantum coherence non-locally links superpositioned tubulins (gray) within and among microtubules. Upper microtubule: cutaway view shows coherent photons generated by quantum ordering of water on tubulin surfaces, propagating in microtubule waveguide. MAP (microtubule-associated-protein) attachments breach isolation and prevent quantum coherence; MAP attachment sites thus act as "nodes" which tune and orchestrate quantum oscillations and set possibilities and probabilities for collapse outcomes ("**Orch OR**").

Summary of the Orch OR Model for Consciousness

The full details of this model are given in Hameroff and Penrose (1996). The picture we are putting forth involves the following ingredients:

1. Aspects of quantum theory (e.g. quantum coherence) and of the suggested physical phenomenon of quantum wave function "*self-collapse*" (objective reduction: OR - Penrose, 1994; 1996) are essential for consciousness, and occur in cytoskeletal microtubules (MTs) and other structures within each of the brain's neurons.

2. Conformational states of MT subunits (tubulins) are coupled to internal quantum events, and cooperatively interact with other tubulins in both classical and quantum computation (Hameroff et al, 1992; Rasmussen et al, 1990 - Figures 4, 5 and 6).

3. Quantum coherence occurs among tubulins in MTs, pumped by thermal and biochemical energies (perhaps in the manner proposed by Frohlich, 1968; 1970; 1975). Evidence for coherent excitations in proteins has recently been reported by Vos et al (1993).

It is also considered that water at MT surfaces is "ordered," dynamically coupled to the protein surface. Water ordering within the hollow MT core (acting like a quantum wave guide) may result in quantum coherent photons (as suggested by the phenomena of "super-radiance" and "*self-induced transparency*" - Jibu et al, 1994; 1995). We require that coherence be sustained (protected from environmental interaction) for up to hundreds of milliseconds by isolation a) within hollow MT cores; b) within tubulin hydrophobic pockets; c) by coherently ordered water; d) sol-gel layering (Hameroff and Penrose, 1996). Feasibility of quantum coherence in the seemingly noisy, chaotic cell environment is supported by the observation that quantum spins from biochemical radical pairs which become separated retain their correlation in cytoplasm (Walleczek, 1995).

4. During pre-conscious processing, quantum coherent superposition/computation occurs in MT tubulins and continues until the mass-distribution difference among the separated states of tubulins reaches a threshold related to quantum gravity. *Self-collapse* (OR) then occurs (Figures 6 & 7).

5. The OR *self-collapse* process results in classical "outcome states" of MT tubulins which then implement neurophysiological functions. According to certain ideas for OR (Penrose, 1994), the outcome states are "non-computable"; that is, they cannot be determined algorithmically from the tubulin states at the beginning of the quantum computation.

6. Possibilities and probabilities for post-OR tubulin states are influenced by factors including initial tubulin states, and attachments of microtubule-associated proteins (MAPs) acting as "nodes" which tune and "orchestrate" the quantum oscillations (Figure 9). We thus term the *self-tuning* OR process in microtubules "orchestrated objective reduction - Orch OR").

7. According to the arguments for OR put forth in Penrose (1994), superposed states each have their own space-time geometries. When the degree of coherent mass-energy difference leads to sufficient separation of space-time geometry, the system must choose and decay (reduce, collapse) to a single universe state. Thus Orch OR involves *self-selections* in fundamental space-time geometry (Figures 10 & 11).

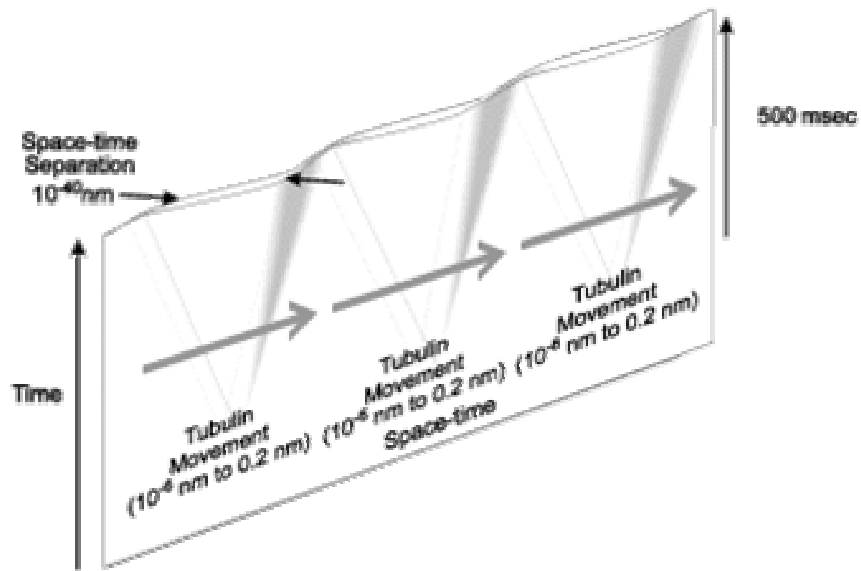


Figure 10. Schematic space-time separation illustration of three superposed tubulins. The space-time differences are very tiny in ordinary terms (10^{-40}nm), but relatively large mass movements (e.g. hundreds of tubulin conformations, each moving from 10^{-6} nm to 0.2 nm) indeed have precisely such very tiny effects on the space-time curvature.

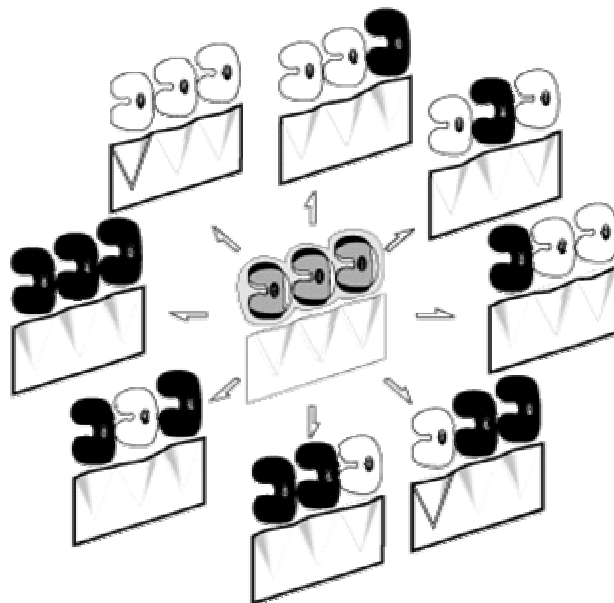


Figure 11. Center: Three superposed tubulins (e.g. Figure 4) with corresponding schematic space-time separation illustrations (Figures 1 and 10). Surrounding the superposed tubulins are the eight possible post-reduction "eigenstates" for tubulin conformation, and corresponding space-time geometry.

8. To quantify the Orch OR process, in the case of a pair of roughly equally superposed states, each of which has a reasonably well-defined mass distribution, we calculate the gravitational *self-energy* E of the difference between these two mass distributions, and then obtain the approximate lifetime T for the superposition to decay into one state or the other by the formula $T = \hbar/E$. Here \hbar is Planck's constant over 2π . We call T the coherence time for the superposition (how long coherence is sustained). If we assume a coherence time $T = 500$ msec (shown by Libet, 1979, and others to be a relevant time for pre-conscious processing), we calculate E , and determine the number of MT tubulins whose coherent superposition for 500 msec will elicit Orch OR. This turns out to be about 10^9 tubulins.

9. A typical brain neuron has roughly 10^7 tubulins (Yu and Baas, 1994). If, say, 10 percent of tubulins within each neuron are involved in the quantum coherent state, then roughly 10^3 (one thousand) neurons would be required to sustain coherence for 500 msec, at which time the quantum gravity threshold is reached and occurs.

10. Orch OR then 10. We consider each *self-organized* Orch OR as a single conscious event; cascades of such events would constitute a "stream" of consciousness. If we assume some form of excitatory input (e.g. you are threatened, or enchanted) in which quantum coherence emerges faster, then, for example, 10^{10} coherent tubulins could Orch OR after 50 msec (e.g. Figure 8c). Turning to see a bengal tiger in your face might perhaps elicit 10^{11} in 5 msec, or more tubulins, faster. A slow emergence of coherence (your forgotten phone bill) may require longer times. A single electron would require more than the age of the universe.

11. Quantum states are non-local (because of quantum entanglement--or "Einstein-Podolsky-Rosen" (EPR) effects), so that the entire non-localized state reduces all at once. This can happen if the mass movement that induces collapse takes place in a small region encompassed by the state, or if it takes place uniformly over a large region. Thus, each instantaneous Orch OR could "bind" various superpositions which may have evolved in separated spatial distributions and over different time scales, but whose net displacement *self-energy* reaches threshold at a particular moment. Information is bound into an instantaneous event (a "conscious now"). Cascades of Orch ORs could then represent our familiar "stream of consciousness," and create a "forward" flow of time (Aharonov and Vaidman, 1990; Elitzur, 1996; Tollaksen, 1996).

It may be interesting to compare our considerations with subjective viewpoints that have been expressed with regard to the nature of the progression of conscious experience. For example, support for consciousness consisting of sequences of individual, discrete events is found in Buddhism; trained meditators describe distinct "flickerings" in their experience of reality (Tart, 1995). Buddhist texts portray consciousness as "momentary collections of mental phenomena", and as "distinct, unconnected and impermanent moments which perish as soon as they arise." Each

conscious moment successively becomes, exists, and disappears - its existence is instantaneous, with no duration in time, as a point has no length. Our normal perceptions, of course, are seemingly continuous, presumably as we perceive "movies" as continuous despite their actual makeup being a series of frames. Some Buddhist writings even quantify the frequency of conscious moments. For example the Sarvaastivaadins (von Rospatt, 1995) described 6,480,000 "moments" in 24 hours (an average of one "moment" per 13.3 msec), and some Chinese Buddhism as one "thought" per 20 msec. These accounts, including variations in frequency, are consistent with our proposed Orch OR events. For example a 13.3 msec pre-conscious interval would correspond with an Orch OR involving 4×10^{10} coherent tubulins, a 0.13 msec interval would correspond with 4×10^{12} coherent tubulins, and a 20 msec interval with 2.5×10^{10} coherent tubulins. Thus Buddhist "moments of experience," Whitehead "occasions of experience," and our proposed Orch OR events seem to correspond tolerably well with one another.

The Orch OR model thus appears to accommodate some important features of consciousness:

1. *control/regulation of neural action*
2. *pre-conscious to conscious transition*
3. *non-computability*
4. *causality*
5. *binding of various (time scale and spatial) superpositions into instantaneous "now"*
6. *a "flow" of time*
7. *a connection to fundamental space-time geometry in which experience may be based.*

CONCLUSION

WHAT IS IT LIKE TO BE A WORM?

The Orch OR model has the implication that an organism able to sustain quantum coherence among, for example, 10^9 tubulins for 500 msec might be capable of having a conscious experience. More tubulins coherent for a briefer period, or fewer for a longer period ($E = h/T$) will also have conscious events. Human brains appear capable of, for example, 10^{11} tubulin, 5 msec "bengal tiger experiences," but what about simpler organisms?

From an evolutionary standpoint, introduction of a dynamically functional cytoskeleton (perhaps symbiotically from spirochetes, e.g. Margulis, 1975) greatly enhanced eukaryotic cells by providing cell movement, internal organization, separation of chromosomes and numerous other functions. As cells became more specialized with extensions like axopods and eventually neural processes, increasingly larger cytoskeletal arrays providing transport and motility may have developed quantum coherence via the Fröhlich mechanism as a by-product of their functional coordination.

Another possible scenario for emergence of quantum coherence leading to Orch OR and conscious events is "cellular vision." Albrecht-Buehler (1992) has observed that single cells utilize their cytoskeletons in "cellular vision" - detection, orientation and directional response to beams of red/infra-red light. Jibu et al (1995) argue that this process requires quantum coherence in microtubules and ordered water, and Hagan (1995) suggests the quantum effects/cellular vision provided an evolutionary advantage for cytoskeletal arrays capable of quantum coherence. For whatever reason quantum coherence emerged, one could then suppose that, one day, an organism achieved sufficient microtubule quantum coherence to elicit Orch OR, and had a "conscious" experience.

At what level of evolutionary development might this primitive consciousness have emerged? A single cell organism like *Paramecium* is extremely clever, and utilizes its cytoskeleton extensively. Could a paramecium be conscious? Assuming a single paramecium contains, like each neuronal cell, 10^7 tubulins, then for a paramecium to elicit Orch OR, 100% of its tubulins would need to remain in quantum coherent superposition for nearly a minute. This seems unlikely.

Consider the nematode worm *C elegans*. It's 302 neuron nervous system is completely mapped. Could *C elegans* support Orch OR? With 3×10^9 tubulins, *C elegans* would require one third of its tubulins to sustain quantum coherent superposition for 500 msec. This seems unlikely, but not altogether impossible. If not *C elegans*, then perhaps *Aplysia* with a thousand neurons, or some higher organism. Orch OR provides a theoretical framework to entertain such possibilities.

Would a primitive Orch OR experience be anything like ours? If *C elegans* were able to *self-collapse*, what would it be like to be a worm? (Nagel, 1974) A single, 10^9 tubulin, 500 msec Orch OR in *C elegans* should be equal in gravitational *self-energy* (and thus perhaps, experiential intensity) to one of our "everyday experiences." A major difference is that we would have many Orch OR events sequentially (up to, say, 10^9 per second) whereas *C elegans* could generate, at most, 2 per second. *C elegans* would also presumably lack extensive memory and associations, and have poor

sensory data, but nonetheless, by our criteria a 10^9 tubulin, 500 msec Orch OR in *C elegans* would be a conscious experience: a mere smudge of known reality, the next space-time move.

Consciousness has an important place in the universe. Orch OR in microtubules is a model depicting consciousness as sequences of non-computable *self*-selections in fundamental space-time geometry. If experience is a quality of space-time, then Orch OR indeed begins to address the "hard problem" of consciousness in a serious way.

References

- Aharonov Y and Vaidman L. Properties of a quantum system during the time interval between two measurements. *Phys Rev A* 1990;41:11.
- Albrecht-Buehler G. Rudimentary form of "cellular vision" *Cell Biol* 1992;89: 8288-8292
- Beck F and Eccles JC. Quantum aspects of brain activity and the role of consciousness. *Proc. Natl. Acad. Sci. USA* 1992;89(23):11357-11361.
- Chalmers D. Facing up to the problem of consciousness. In: *Toward a Science of Consciousness - The First Tucson Discussions and Debates*, S.R. Hameroff, A. Kaszniak and A.C. Scott (eds.), MIT Press, Cambridge, MA. 1996.
- Chalmers D. *Toward a Theory of Consciousness*. Springer-Verlag, Berlin. 1996.
- Conze E. *Buddhist Thought in India*, Louis de La Vallee Poussin (trans.), *Abhidharmako"sabhaa.syam*: English translation by Leo M. Pruden, 4 vols (Berkeley) 1988;85-90.
- Diği L. Models for universal reduction of macroscopic quantum fluctuations. *Phys Rev A* 1989;40:1165-1174.
- Everett H. Relative state formulation of quantum mechanics. In *Quantum Theory and Measurement*, J.A. Wheeler and W.H. Zurek (eds.) Princeton University Press, 1983; originally in *Rev Mod Physics* 1957;29:454-462.
- Frohlich H. Long range coherence and the actions of enzymes. *Nature* 1970;228:1093.
- Frohlich H. Long-range coherence and energy storage in biological systems. *Int J Quantum Chem* 1968;2:641-9.
- Frohlich H. The extraordinary dielectric properties of biological materials and the action of enzymes. *Proc Natl Acad Sci* 1975;72:4211-4215.

Ghirardi GC, Grassi R and Rimini A. Continuous-spontaneous reduction model involving gravity. Phys

Rev A 1990;42:1057-1064.

Ghirardi GC, Rimini A and Weber T. Unified dynamics for microscopic and macroscopic systems. Phys

Rev D 1986;34:470.

Goswami A The Self-Aware Universe: How Consciousness Creates the Material World.

Tarcher/Putnam, New York. 1993.

Hameroff SR and Penrose R. Orchestrated reduction of quantum coherence in brain microtubules: A

model for consciousness. Neural Network World 1995;5(5): 793-804.

Hameroff SR and Penrose R. Orchestrated reduction of quantum coherence in brain microtubules: A

model for consciousness. In: Toward a Science of Consciousness - The First Tucson Discussions and Debates, S.R. Hameroff, A. Kaszniak and A.C. Scott (eds.), MIT Press, Cambridge, MA. 1996.

Hameroff SR, Dayhoff JE, Lahoz-Beltra R, Samsonovich A and Rasmussen S. Conformational automata in

the cytoskeleton: models for molecular computation. IEEE Computer (October Special Issue on Molecular Computing) 1992;30-39.

Jibu M, Hagan S, Hameroff SR, Pribram KH and Yasue K. Quantum optical coherence in cytoskeletal

microtubules: implications for brain function. BioSystems 1994;32:195-209.

Kaszniak A and Scott AC(eds.), Time and consciousness: The uneasy bearing of relativity theory on the

mind-body problem. In: Toward a Science of Consciousness - The First Tucson Discussions and Debates, Hameroff SR., MIT Press, Cambridge, MA. 1996

Krolhzy F, Frenkel A and Lukacs B. On the possible role of gravity on the reduction of the wave function.

In Quantum Concepts in Space and Time, R. Penrose and C.J. Isham (eds.), Oxford University Press. 1986.

Libet B, Wright EW Jr, Feinstein B and Pearl DK. Subjective referral of the timing for a conscious sensory experience. *Brain* 1979;102:193-224.

Louria D and Hameroff S. Computer simulation of anesthetic binding in protein hydrophobic pockets. In: *Toward a Science of Consciousness - The First Tucson Discussions and Debates*, S.R. Hameroff, A. Kaszniak and A.C. Scott (eds.), MIT Press, Cambridge, MA. 1996.

Margulis L. *Origin of Eukaryotic Cells*. Yale University Press, New Haven. 1975.

Marshall IN. Consciousness and Bose-Einstein condensates. *New Ideas in Psychology* 1989;7:73-83.

Nagel T. What is it like to be a bat? in *The Mind's I. Fantasies and Reflections on Self and Soul*. (eds) D.R. Hofstadter and D.C. Dennett, Basic Books, N.Y. 1981;391-403, (Originally published in *The Philosophical Review*, October, 1974)

Pearle P and Squires E. Bound-state excitation, nucleon decay experiments and models of wave-function collapse. *Phys Rev Letts* 1994;73(1):1-5.

Pearle P. Combining stochastic dynamical state vector reduction with spontaneous localization. *Phys Rev D* 1989;13:857-868.

Penrose R and Hameroff SR. What gaps? Reply to Grush and Churchland. *Journal of Consciousness Studies* 2(2):99-112.

Penrose R. Gravity and quantum mechanics. In *General Relativity and Gravitation. Proceedings of the Thirteenth International Conference on General Relativity and Gravitation held at Cordoba, Argentina 28 June-4 July 1992. Part 1: Plenary Lectures*. (eds. R.J. Gleiser, C.N. Kozameh and O.M. Moreschi) Institute of Physics Publications, Bristol. 1993.

Penrose R. Newton, quantum theory and reality. In *300 Years of Gravity* S.W. Hawking and W. Israel (eds.) Cambridge University Press. 1987.

Penrose R. *Shadows of the Mind*, Oxford Press, Oxford, U.K. 1994.

Penrose R. *The Emperor's New Mind*, Oxford Press, Oxford, U.K. 1989.

Rasmussen S, Karampurwala H, Vaidyanath R, Jensen KS and Hameroff S. (Computational connectionism within neurons: A model of cytoskeletal automata subserving neural networks. *Physica D* 1990;42:428-449.

Rensch B. *Evolution Above the Species Level*. Columbia University Press, New York. 1960.

Russell B. *The Analysis of Matter*. Dover, New York. 1954.

Schrödinger E. Die gegenwärtigen situation in der quantenmechanik. *Naturwissenschaften*, 1935;23:807-812, 823-828, 844-849. (Translation by J. T. Trimmer (1980) in *Proc. Amer. Phil. Soc.*, 124:323-338.) In *Quantum Theory and Measurement* (ed. J.A. Wheeler and W.H. Zurek). Princeton University Press, 1983.

Shimony A. *Search for a Naturalistic World View - Volume II. Natural Science and Metaphysics*. Cambridge University Press, Cambridge, U.K. 1993.

Spinoza B. *Ethica in Opera quotque reperta sunt*. 3rd edition, eds. J. van Vloten and J.P.N. Land (Netherlands: Den Haag). 1677.

Stubenberg L. The place of qualia in the world of science. In: *Toward a Science of Consciousness - The First Tucson Discussions and Debates*, S.R. Hameroff, A. Kaszniak and A.C. Scott (eds.), MIT Press, Cambridge, MA. 1996.

Tart CT. personal communication and information gathered from "Buddha-I newnet" 1995.

Tollaksen J. New insights from quantum theory on time, consciousness, and reality. In: *Toward a Science of Consciousness - The First Tucson Discussions and Debates*, S.R. Hameroff, A. Kaszniak and A.C. Scott (eds.), MIT Press, Cambridge, MA. 1996.

von Rospatt A. The Buddhist Doctrine of Momentariness: A survey of the origins and early phase of this doctrine up to Vasubandhu (Stuttgart: Franz Steiner Verlag). 1995.

Vos MH, Rappaport J, Lambry J Ch, Breton J, Martin JL. Visualization of coherent nuclear motion in a membrane protein by femtosecond laser spectroscopy. Nature 1986;363:320-325.

Walleczek J. Magnetokinetic effects on radical pairs: a possible paradigm for understanding sub-kT magnetic field interactions with biological systems. in Biological Effects of Environmental Electromagnetic Fields. M Blank (ed) Advances in Chemistry No 250 American Chemical society Books Washington DC, 1995.

Wheeler JA. Assessment of Everett's "relative state" formulation of quantum theory. Revs Mod Phys 1957;29:463-465.

Wheeler JA. Information, physics, quantum: The search for links. In (W. Zurek, ed.) Complexity, Entropy, and the Physics of Information. Addison-Wesley. 1990.

Whitehead AN. Process and Reality. Macmillan, N.Y. 1929.

Whitehead AN. Science and the Modern World. Macmillan, N.Y. 1929.

Yu W and Baas PW. Changes in microtubule number and length during axon differentiation. J Neuroscience 1994;14(5):2818-2829.