"Funda-Mentality" Is the Conscious Mind Subtly Linked to a Basic Level of the Universe?

Summary

Stuart Hameroff

Age-old battle lines over the puzzling nature of mental experience are shaping a modern resurgence in the study of consciousness. On one side are the long-dominant "physicalists" (reductionists, materialists, functionalists, computationalists. .) who see consciousness as an emergent property of the brain's neural networks ("brain = mind = computer"). On the alternative, rebellious side are those who see a necessary added ingredient: proto-conscious experience intrinsic to reality, perhaps understandable through modern physics (panpsychists, pan-experientialists, "funda-mentalists"). It is argued here that the physicalist premise alone is unable to solve completely the difficult issues of consciousness (e.g. experience, binding, pre-conscious conscious transition, non-computability and free will) and that to do so will require supplemental panpsychist/pan-experiential philosophy expressed in modern physics. In one such scheme proto-conscious experience is a basic property of physical reality accessible to a quantum process associated with brain activity. The proposed process is Roger Penrose's *objective reduction* (OR), a self-organizing "collapse" of the quantum wave function related to instability at the most basic level of spacetime geometry. In the Penrose-Hameroff model of "orchestrated objective reduction" ("Orch OR"), OR quantum computation occurs in cytoskeletal microtubules within the brain's neurons and links cognition with proto-conscious experience and Platonic values embedded in spacetime geometry. The basic idea is that consciousness involves brain activities coupled to self-organizing ripples in fundamental reality.

Introduction - A Burning Issue

Can conscious experience-feelings, qualia, our "inner life"-be accommodated within present-day science? Those who believe it can (e.g. physicalists, reductionists, materialists, functionalists, computationalists) see conscious experience as an emergent property of complex neural network computation. Others see conscious experience either outside science (dualists), or believe science must expand to include experience (idealists, panpsychists, pan-experientialists, "funda-mentalists"). These philosophical battle lines were originally drawn in ancient Greece between Socrates, who believed the cerebrum created consciousness, and Aristotle, Democritus, Thales and others who argued that mental qualities belonged to fundamental reality. Perhaps both sides were correct.

Brain = Mind = Computer?

The basic physicalist idea is that the mind is a computer functioning in the brain's neural networks. The current leading candidate for a computer-like "neural correlate" of consciousness involves synchronously oscillating feedback loops of thalamo-cortical neurons. Higher frequencies (collectively known as "coherent 40 Hz") have been suggested to mediate temporal binding of conscious experience (e.g. Singer, Gray, Crick and Koch, etc.). The proposals vary, for example as to whether coherence originates in thalamus or resonates in cortical networks, but "thalamo-cortical 40 Hz" stands as a prevalent view of the substrate for consciousness.

But how do synchronized neural firings and synaptic transmissions produce experiential qualia, emotions or free will? Physicalists believe this to be relatively straightforward (brain = mind = computer) however others find the question intractable, or as vexing as trying to coax a reluctant genie from a magic lamp. I see three problems with the brain = mind = computer analogy:

1. *Is consciousness classical computation?* In a controversial stance Roger Penrose [1-3] has asserted that essential aspects of consciousness are non-computable. But regardless, classical computers appear to be evolving toward quantum computers. Beginning in the early 1980's Benioff, Feynman and others proposed that states in a system - bits in a computer - could interact while in quantum superposition of all possible states, effecting near-infinite parallel computation. Rather than classical Boolean bit states 1 or 0, quantum computers would utilize interactive "qubits" of 1 *and* 0. If quantum computers can ever be constructed they will have huge advantages in important applications. As the brain/mind has always been cast as current information technology, consciousness may inevitably be seen as some form of quantum computation.

2. *Are neural firings the "fine grain" of consciousness?* Cells and synapses are far more complex than simple onoff switches. Consider the paramecium, a single cell organism which gracefully swims, avoids predators, learns to escape from capillary tubes, and finds food and mates. Observing intelligent behavior in unicellular creatures C.S. Sherrington said in 1951: "Of nerve there is no trace. But the cell framework, the cyto-skeleton, might serve." Lacking synapses, paramecium utilizes its cytoskeleton for communication and organization. Neurons have a rich and dynamic set of cytoskeletal microtubules which regulates synapses, and tremendously increases potential computational capacity (e.g. 10^{16} bit states/second/neuron) [4]. More importantly, neurons are alive and we don't yet know what that implies for consciousness.

3. *Details which don't fit the brain = mind = computer scheme are overlooked.*

For example:

    a. Neurotransmitter vesicle release and cognitive reaction times are "noisy", and exhibit apparent probabilistic randomness (?noncomputable quantum indeterminacy [5]).
    b. Axonal firing patterns (rather than average frequency) and dendriticdendritic processing may be relevant to consciousness [6].
    c. Apart from chemical synapses, primitive electrotonic gap junctions couple neurons and glia synchronously and may play an important role in consciousness.
   d. Glial cells (80% of the brain) are ignored in the brainascomputer view.

Quibbling aside, the physicalist view fails to address difficult issues. For example the problem of 'binding' in vision and self is often attributed to temporal correlation (e.g. coherent 40 Hz), but it is unclear why temporal correlation *per se* should bind experience without an explanation of experience. Regarding transition from preconscious or implicit processing to consciousness itself, the physicalist view sees emergence at a critical level of neural-level computational complexity. But no conscious threshold is apparent, nor is there a reasonable suggestion why such an emergent property should have conscious experience. As physicalism is based on deterministic computation, it is also unable to account for free will or Penrose's proposed non-

computability. But the major problem remains experience, for which physicalism offers no testable predictions. Something is missing.

Panpsychism Meets Modern Physics

Perhaps panpsychists are in some way correct and components of mental processes are fundamental, like mass, spin or charge. Following the ancient Greek panpsychists, Spinoza (1677) saw some form of consciousness in all matter. Leibniz (1766) portrayed the universe as an infinite number of fundamental units ("monads") each having a primitive psychological being. Whitehead (e.g. 1929) was a process philosopher who viewed reality as a collection of events occurring in a basic field of protoconscious experience ("occasions of experience"). Abner Shimony[7] observed that Whitehead's occasions were comparable to quantum state reductions-actual events in physical reality (see below). But what of Whitehead's "basic field" of proto-conscious experience? How could experience (qualia) simply exist in empty space?

What *is* empty space? This question also stems from ancient Greece. Democritus argued that empty space was a true void whereas Aristotle contended that it was in fact a plenum (background filled with substance)-a medium in which heat and light traveled. Siding with Aristotle, Maxwell's 19th century theory of the luminiferous ether described a plenum that carried electromagnetic waves. However attempts to detect the ether failed and Einstein's special relativity in 1905 reverted to Democritus in that empty space was an absolute void. However ten years later Einstein's general relativity with its curved space and distorted geometry reversed his stand to opt for a richly-endowed plenum termed the spacetime metric.
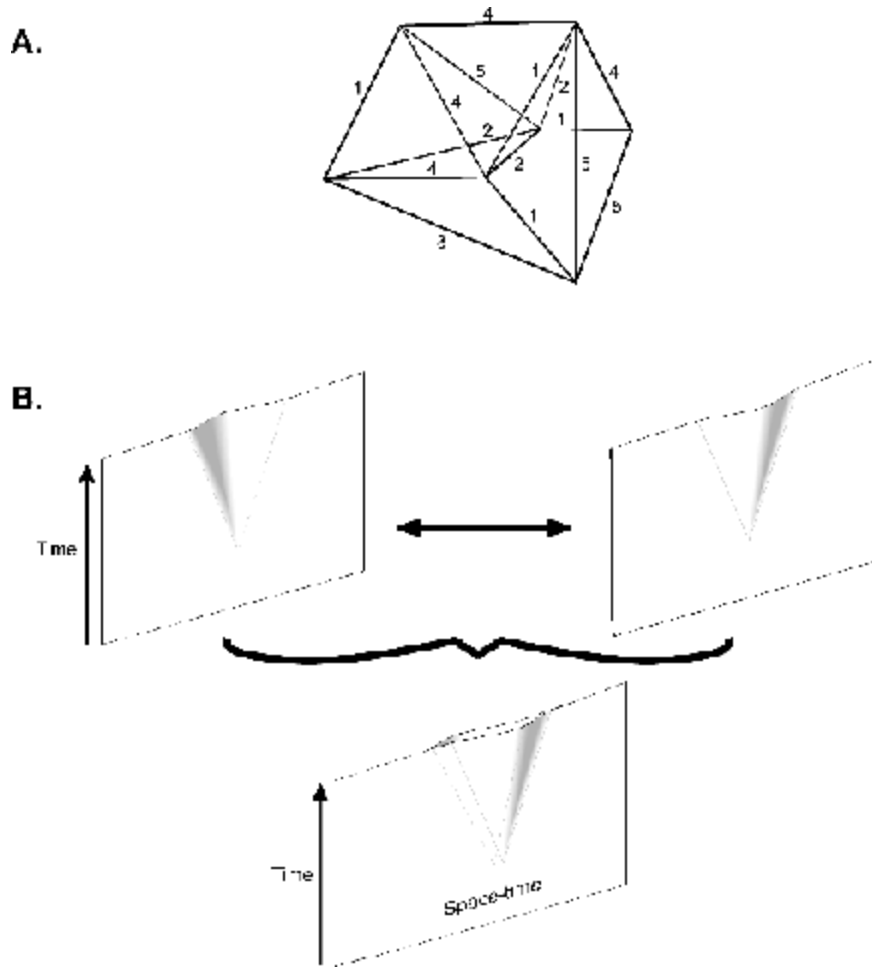
Figure 1. *Two descriptions of fundamental spacetime geometry. a) A quantum spin network. Introduced by Roger Penrose[8] as a quantum mechanical description of the geometry of space, spin networks describe a spectra of discrete Planck scale volumes and configurations[9,10]. Average length of each link is the Planck length (10^{33} cm). b) Four dimensional spacetime may be schematically represented by one dimension of space and one dimension of time: a two dimensional "spacetime sheet." Mass is curvature in spacetime, and the two spacetime curvatures in the top of Figure 1b represent mass (e.g. a tubulin protein) in two different locations, or conformations respectively. In quantum superposition mass separated from itself is simultaneous spacetime curvature in opposite directions, a separation or "bubble" of spacetime. At a critical degree of separation, the system becomes unstable and must select either one state or the other[2].*

We now know that at very small scales space and time are not smooth, but quantized. This granularity occurs at the incredibly small dimensions of the "Planck scale" at 10^{-33} centimeters and 10^{-43} seconds. Roger Penrose portrays this basic makeup of the universe as a dynamical spider-web of quantum spins[8]. These "spin networks" create an evolving array of Planck scale geometric volumes defining four dimensional spacetime (Figure 1a). Penrose applies Einstein's general relativity (in which mass equates to curvature, or perturbation of spacetime) all the way down to this near-infinitesimal geometry (Figure 1b). Thus everything is *in reality* particular arrangements of spacetime geometry. Building on these ideas, Lee Smolin[9,10] likens spin network volumes to Leibniz monads and suggests that self-organizing processes at this level constitute a flow of time, raising the issue of whether the universe is in some sense alive. Could infinitesimally small, weak and fast processes be coupled to biology? A reasonable possibility

for such a link is Penrose's objective reduction-a particular type of quantum state reduction in which new macroscopic information emerges.

At the Edge of Reality: Quantum State Reduction and Consciousness

Quantum theory describes the bizarre wave/particle duality of energy and matter at very small scales. The behavior is so strange that the American physicist Richard Feynman once commented "anyone who claims to understand quantum theory is either lying or crazy."

Strange as it is, quantum theory offers features which may be relevant to consciousness. One is that large collections of quantum particle/waves can merge into unitary coherent states of macroscopic size and influence. Superconductors, BoseEinstein condensates and lasers are unitary states in which component atoms or molecules give up individual identity and behavior. Such coherent quantum states have been suggested to occur among brain proteins to provide unitary "binding" in vision and sense of self[11,12].

Another feature involves "quantum superposition." Components of isolated small scale systems can exist in different states or locations simultaneously. This is contrary to our perceived macroscopic world in which objects have well defined positions and are decidedly concrete. The problem is the transition-why and how do microscopic quantum superposed states become classical and definite in the macro-world? This problem is called quantum state reduction, or collapse of the wave function, and it may be the key to both consciousness and reality.

Experimental evidence in the early part of this century led great theorists Bohr, Heisenberg and Wigner to conclude (the "Copenhagen interpretation") that objects remain in wave-like quantum superposition until observed by a conscious human being-consciousness causes collapse of the wave function! To illustrate the apparent absurdity of this conclusion, in the 1930's Schrodinger devised his famous thought experiment Schrodinger's cat. A living cat is placed in a box into which poison can be released by a quantum event, e.g., sending a photon through a half-silvered mirror. So after the photon has been sent there are equal possibilities that the cat is either dead or alive. But according to the Copenhagen interpretation until a conscious being opens the box to observe, the cat is both dead *and* alive. Schrodinger's point was that the conscious observer interpretation was incorrect.

Many physicists now believe that intermediate between tiny quantum-scale systems and "large" cat-size systems some objective factor disturbs the superposition to cause collapse, or "objective reduction (OR)." According to Roger Penrose[2,13] this objective factor is an intrinsic feature of spacetime itself (quantum gravity). As mass is equivalent to spacetime curvature, Penrose begins with the notion that quantum superpositionactual separation (displacement) of mass from itself is equivalent to simultaneous spacetime curvatures in opposite directions, causing "bubbles," or separations in fundamental reality (Figure 1b). Penrose reasons that these bubble-like separations are unstable and reduce to specific states and locations after a critical degree of separation. If proto-conscious experience is rooted in the Planck scale, then objective reductions (Whitehead's occasions of experience) may ripple through an experiential medium.

Could OR events occur in the brain? The critical spacetime separation precipitating Penrose's OR is given by the uncertainty principle $E=h/T$. E is the energy of the superposed mass, h is Planck's

constant over 2pi, and T is the coherence time until reduction. The size (and energy) of a superposed system (degree of spacetime separation) is inversely related to the time T until self-collapse. If isolated, a large system (e.g. Schrodinger's one kilogram cat) will undergo OR very quickly, e.g. in only 10^{-37} seconds. A small system such as a single isolated superposed atom would undergo OR only after 10^7 years. OR brain events would be linked to neural processes with T in the range of tens to hundreds of milliseconds (e.g. 25 msec intervals in coherent 40 Hz). For T=25 msec (40 Hz) OR events, E corresponds to roughly 3 nanograms (3×10^{-9} gram) of superposed brain mass.

Nanograms of what? Which biological structures could function as both classical and quantum computers, avoid environmental decoherence and couple to neural-level activities? Microtubules are the logical candidates.

Are Microtubules Quantum Computers? The PenroseHameroff "Orch OR" Model

Interiors of neurons and glia are functionally organized by webs of protein polymersthe cytoskeleton[14]. Its major components are microtubules, actin and intermediate filaments. Microtubules are selfassembling hollow cylinders whose walls are crystalline lattices of subunit proteins known as tubulin. Evidence links the neuronal cytoskeleton to cognitive functions, and theoretical models suggest interactive microtubule subunits function as molecular automata capable of nanosecond-scale computation (Figure 2a)[4].
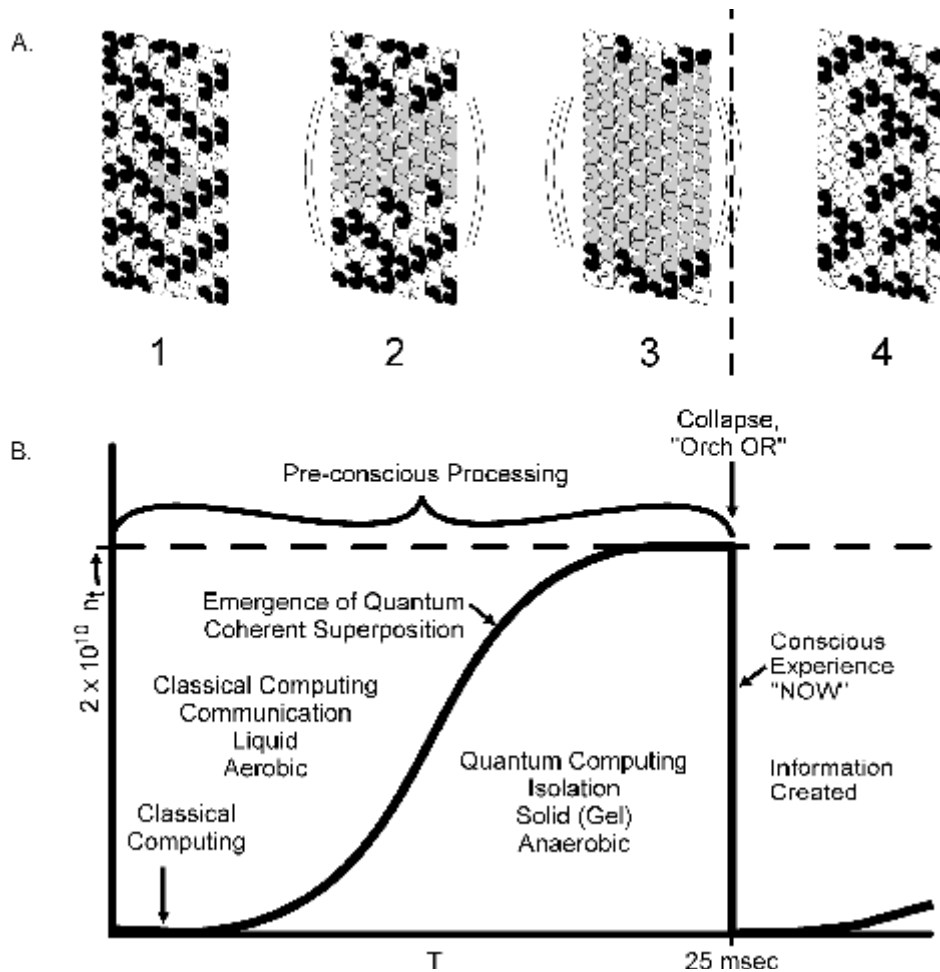
Figure 2. *Schematic of neural synapse showing cytoskeletal structures within two neurons. Left: Pre-synaptic axon terminals releases neurotransmitter vesicles (black spheres) into synaptic cleft. Thick, black rod-like structures at top indicate microtubules; thinner filaments (e.g. synapsin) facilitate vesicle release. Right: Dendrite on post-synaptic neuron with two dendritic spines. Microtubules in main dendrite are interconnected by microtubule-associated proteins. Other cytoskeletal structures (fodrin, actin filaments, etc.) connect membrane receptors to microtubules. Based on Hirokawa[14].*

Roger Penrose and I have developed a model in which quantum superposition, objective reduction and quantum computation occur in microtubule automata within brain neurons and glia. Microtubuleassociatedproteins (MAPs) provide feedback and "tune" the quantum oscillations; the proposed OR is thus selforganized ("orchestrated" objective reduction"Orch OR")[15-21]. In the Orch OR model microtubule quantum computation is isolated from decoherence (Box 1) and continues until threshold is met ($E = h/T$) and an OR event occurs (Figure 2b). For example an OR event coinciding with one 40 Hz cycle (T = 25 msec) would require $E = 2 \times 10^{10}$ superposed tubulins (roughly 20,000 neurons).

Quantum computation in the Orch OR scheme differs from technological proposals in that superpositions in the latter will reduce to output states by environmental decoherencecomputation is terminated by intervention and choice of states has an element of randomness. On the other hand, in the Orch OR scheme isolated superpositions self-reduce due

to instability in spacetime separation. The choice of outcome states, according to Penrose, is therefore neither completely deterministic nor random, but has an element of non-computabilityinfluenced by Platonic logic embedded in spacetime. Penrose has also suggested that aspects of human understanding and consciousness involve non-computability, a controversial and widely assailed claim. Although outnumbered by his critics, Penrose has thoroughly and systematically answered them[3]. Non-computability is a clue, a delicate thread with which to unravel the mystery of consciousness.

Orch OR and Cognition

Each proposed Orch OR event consists of 1) an isolated quantum computing phase identified with preconscious, implicit processing which culminates in 2) instantaneous reduction corresponding with a discrete conscious "now" event-a Whitehead "occasion of experience."

Each event selects (noncomputably) particular configurations of Planck-scale experiential geometry and corresponding classical states of microtubule automata which regulate synaptic/neural functions (Figure 2; Figure 3a,b). Sequences of discrete conscious events (e.g. at 40 Hz) can provide a "stream" of consciousness.
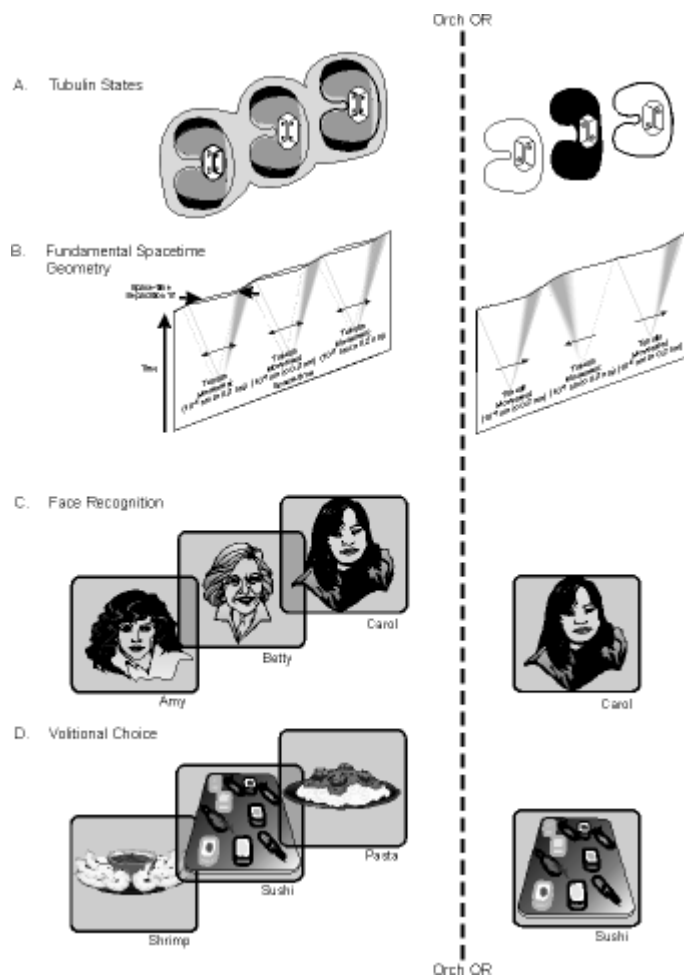
***Figure 3.** Orch OR events in conscious experience. a) (left) Three tubulins in quantum superposition prior to 25 msec Orch OR After reduction (right), particular classical states are selected. b) Fundamental spacetime geometry view. Prior to Orch OR (left), spacetime corresponding with three superposed tubulins is separated as Planck scale bubbles: curvatures in opposite directions. The Planckscale spacetime separations S are very tiny in ordinary terms, but relatively large mass movements (e.g., hundreds of tubulin conformations, each moving from 10^6 to 0.2 nm) indeed have precisely such very tiny effects on the spacetime curvature. A critical degree of separation causes Orch OR and an abrupt selection of single curvatures (and a particular geometry of experience). c) Cognitive facial recognition. A familiar face induces superposition (left) of three possible solutions (Amy, Betty, Carol) which "collapse" to the correct answer Carol (right). d) Cognitive volition. Three possible dinner selections (shrimp, sushi, pasta) are considered in superposition (left), and collapse via Orch OR to choice of sushi (right).*

Consider Orch OR in the context of two cognitive tasks: facial recognition and ordering dinner (Figure 3c,d). Each may occur in a series of steps yielding intermediate solutions, however for the purpose of illustration consider how single Orch OR conscious events could accomplish these tasks. (Although classical neurallevel parallel computation can partially explain these functions, the Orch OR scheme provides far greater information capacity, conscious experience, binding, and noncomputability consistent with free will.)

Imagine you briefly see a familiar woman's face (Figure 3c). Is she Amy, Betty, or Carol? All possibilities may superpose in quantum computation. For example during 25 milliseconds of pre-conscious processing quantum computation occurs with information (Amy, Betty, Carol) in the form of "qubits", superposed states of microtubule automata. As threshold for objective reduction is reached, superposed tubulin qubits reduce (collapse) to definite states, becoming bits. Now, you recognize Carol as a particular experiential geometry is selected! (Many more than three possibilities, in fact an astronomically high number of possibilities could be superposed in microtubule quantum computing.)

In a volitional act possible choices may be superposed. Suppose you are selecting dinner from a menu. During preconscious processing, shrimp, sushi and pasta are superposed. As threshold for objective reduction is reached, the quantum state reduces to a single classical state whose selection results from deterministic quantum computation influenced at the moment of reduction by Platonic logic embedded in the Planck scale. A choice is made. You'll have sushi!

Conclusion

The Orch OR model is consistent with known neurophysiological processes, generates numerous testable predictions[18,19] and is the type of multi-level, trans-disciplinary theory required to address the mind's enigmatic features. Consciousness may involve subtle links between the brain and fundamental spacetime geometry.

References

1. Penrose, R. (1989) The Emperor's New Mind, Oxford Press, Oxford, U.K.
2. Penrose, R. (1994) Shadows of the Mind, Oxford Press, Oxford, U.K.
3. Penrose, R. (1997) On understanding understanding. International Studies in the Philosophy of Science 11(1):7-20, 1997.
4. Rasmussen, S., Karampurwala, H., Vaidyanath, R., Jensen, K.S., and Hameroff, S. (1990) Computational connectionism within neurons: A model of cytoskeletal automata subserving neural networks. Physica D 42:428449.
5. Beck, F. and Eccles, J.C. (1992) Quantum aspects of brain activity and the role of consciousness. Proc. Natl. Acad. Sci. USA 89(23):1135711361.
6. Pribram, K.H. (1991) Brain and Perception (Lawrence Erlbaum, New Jersey)
7. Shimony, A., (1993) Search for a Naturalistic World View Volume II. Natural Science and Metaphysics. Cambridge University Press, Cambridge, U.K.
8. Penrose, R. (1971) in Quantum Theory and Beyond. ed E.A. Bastin, Cambridge University Press, Cambridge, U.K.
9. Smolin, L. (1997) Life of the Cosmos, Oxford Press, New York.
10. Rovelli C, Smolin L (1995b) Spin networks in quantum gravity. Physical Review D 52(10)57435759
11. Penrose, R. (1987) Newton, quantum theory and reality. In 300 Years of Gravity, Eds. S.W. Hawking and W. Israel, Cambridge, U.K., Cambridge Press
12. 12. Marshall, I.N. (1989) Consciousness and BoseEinstein condensates. New Ideas in Psychology 7, 3 83.
13. Penrose, R. (1996) On gravity's role in quantum state reduction. General Relativity and Gravitation. 28(5): 581600
14. Hirokawa, N. (1991) Molecular architecture and dynamics of the neuronal cytoskeleton. In The neuronal cytoskeleton, Ed. R. Burgoyne. New York, Wiley- Liss, pp. 5-74
15. Penrose, R., Hameroff, S.R. (1995) What gaps? Reply to Grush and Churchland. Journal of Consciousness Studies 2(2):99112.
16. Hameroff, S.R., and Penrose, R., (1996a) 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. Also published in Mathematics and Computers in Simulation 40:453480, c.f. http://www.u.arizona.edu/~hameroff/penrose1
17. Hameroff, S.R., and Penrose, R. (1996b) Conscious events as orchestrated spacetime selections. Journal of Consciousness Studies 3(1):3653 http://www.u.arizona.edu/~hameroff/penrose2
18. Hameroff S.R. (1997a) Quantum computing in microtubules-An Intra-neural correlate of consciousness? Japanese Bulletin of Cognitive Science 4 (3): 67-92
19. Hameroff, S.R. (1998) Fundamental geometry: The PenroseHameroff Orch OR model of consciousness. In: Geometrical Issues in the Foundations of Science, S. Huggett, L.J. Mason, K.P. Tod, S.T. Tsou, N.M.J. Woodhouse (eds.) Oxford University Press, Oxford, U.K. pp 103-127
20. Hameroff, S (1998) Did consciousness cause the Cambrian evolutionary explosion? in: Toward a Science of Consciousness II The Second Tucson Discussions and Debates. Eds S Hameroff, A Kaszniak, A Scott. MIT Press pp 421-437

21. Hameroff, S.R. (1998) Quantum computation occurring in brain microtubules? The Penrose-Hameroff "Orch OR" model of consciousness (1998) Philosophical Transaction of the Royal Society (in press)


Box 1

Isolated Macroscopic Quantum States in the Brain?

At first glance the brain is a noisy, thermal environment, hardly hospitable to delicate quantum effects which require (in the technological realm) extreme cold to prevent thermal excitations and environmental decoherence. However nature may have solved the problem of quantum state isolation. For example Dan Sackett[a] of NIH has recently shown that microtubules may be insulated from thermal noise by a surrounding sleeve of plasma-like charge condensation.

But how could isolated cytoplasmic quantum states traverse membranes and synapses to occur macroscopically among microtubules in (e.g. 20,000) neurons throughout the brain? One possibility involves quantum tunneling through gap junctions, primitive electrotonic windows between neurons and glia. Neurons interconnected by gap junctions form networks which fire synchronously, "behaving like one giant neuron"[f] (possibly accounting for synchronized 40 Hz neural activity[g]). Unlike chemical synapses which separate neural processes by 3050 nanometers, gap junction separations are 3.5 nanometers, within range for quantum tunneling. Gap junctions are widespread but unevenly distributed. Immunolabeling of gap junction protein (connexin) precursor demonstrates high levels in thalamic subnuclei, layers 2 and 3 of cortex, and midbrain[h]. Thalamocortical networks of gap junctionconnected neurons with sol-gel phases coupled to synchronized 40 Hz activity could isolate microtubules across large brain volumes and provide cycles of isolated macroscopic quantum coherence.
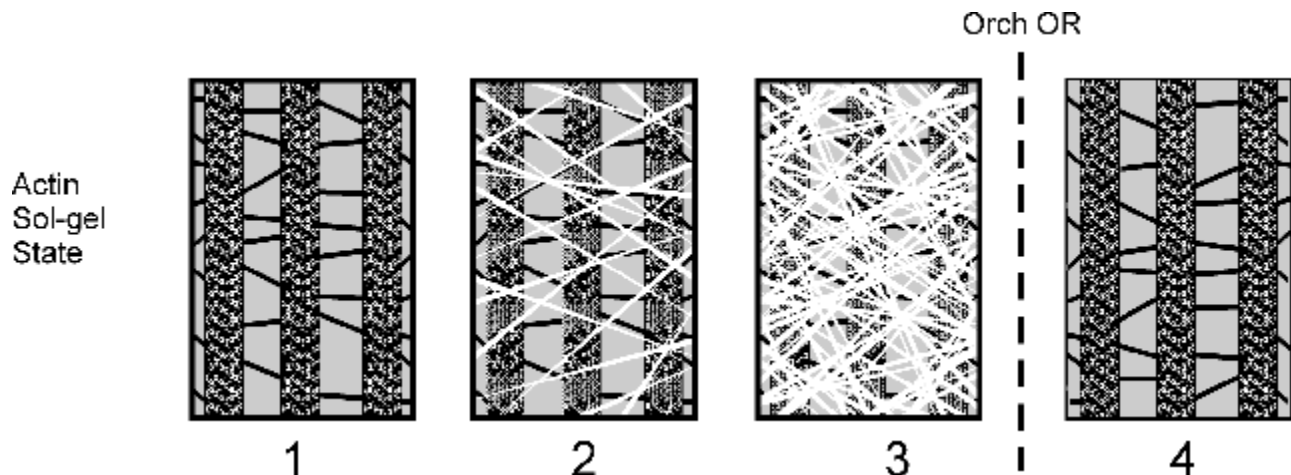
Figure *Schematic sequence of phase of actin gelation/quantum isolation (13) and solution/environmental communication (4) around MT. Cycles may occur rapidly, e.g., 25 msec intervals (40Hz).*

References

a. Sackett, D. (1997) personal communication; paper presented at "Molecular Biophysics of the Cytoskeleton" sponsored by the University of Alberta, Banff, Alberta, Canada, August 18-22, 1997

b. Muallem, S., Kwiakowska, K., Xu, X. and Yin, H.L. (1995) Actin filament disassembly is a sufficient final trigger for exocytosis in nonexcitable cells. Journal of Cell Biology 128, 589-598.

c. Miyamoto, S. (1995) Changes in mobility in synaptic vesicles with assembly and disassembly of actin network. Biochimica et Biophysica Acta 1244, 85-91.

d. Wachsstock, D.H., Schwartz, W.H. and Pollard, T.D. (1994) Cross-linker dynamics determine the mechanical properties of actin gels. Biophysical Journal 66, 801-809.

e. Walleczek, J. (1995) Magnokinetic effects on radical pairs: a possible paradigm for understanding sub kT magnetic field interactions with biological systems. In: Electromagnetic Fields: Biological Interactions and Mechanisms. M. Blank (ed) Advances in Chemistry No. 250 American Chemical Society Books: Washington, D.C.

f. Kandel, E.R., Siegelbaum, S.A., Schwartz, J.H. (1991) Synaptic transmission. In: Principles of Neural Science (3rd Edition). E.R. Kandel, J.H Schwartz, T.M. Jessell (eds). Elsevier: New York, pp. 121-134.

g. Jibu, M. (1990) On a heuristic model of the coherent mechanism of the global reaction process of a group of cells. Bussei Kenkyuu (Material Physics Research) 53, 431-436.

h. Micevych, P.E. and Abelson, L. (1991) Distribution of mRNAs coding for liver and heart gap junction protein in the rat central nervous system. Journal of Comparative Neurology 305, 96-118.