# Orchestrated Objective Reduction: Quantum Physics and its Implications in Human Consciousness

A thesis submitted in partial fulfillment of the requirements for the degree of Bachelor of Science degree in Physics from the College of William and Mary

by


Amanda L. Collins


Advisor: Dr. David Armstrong

Senior Research Coordinator: Dr. Gina Hoatson

Date: April 17, 2015

# 1 ABSTRACT

Consciousness is a complex human phenomenon that has only been studied in the most abstract sense. Sir Roger Penrose and his colleague, Stuart Hameroff, propose a theory of consciousness, that has its roots in quantum mechanics, called Orchestrated Objective Reduction. From Gödel's Theorem to Diósi-Penrose Gravity, in this paper I will deconstruct the key components of this theory and discuss its implications for physics as a whole. Several studies have been cited as evidence for different aspects of Orchestrated Objective Reduction, though in most cases this is the result of speculation and inference rather than actual data. In truth, a significant portion of this theory cannot be directly studied with current technology. However, given time, this theory could prove to be a ground-breaking foray into the world of quantum consciousness.

# TABLE OF CONTENTS



# 2 INTRODUCTION

Artificial intelligence (AI) has been attempting to duplicate the human mind for decades. AI was originally built on the premise that each input to a brain cell has one output. This assumption, known as the Hodgkin-Huxley model, would mean that human brain activity is digitally computable and can be replicated by a digital computer once the correct code has been discovered. The Turing Test is plausibly the ultimate judgment on whether an AI has successfully achieved this goal. In this test, a human and a computer each generate answers to questions, and another human must decide which is the bona fide human and which is artifice. To this day, no AI has been able to pass this test.

In The Emperor's New Mind [1], Roger Penrose takes a new approach to the problem of artificial intelligence. He asserts that if consciousness were truly a one-to-one computable phenomenon, an AI sophisticated enough to pass the Turing Test surely would have been developed by now. For some reason, this has not happened yet. In order to come up with a solution, Penrose turned to Gödel's Theorem. In short, Gödel's Theorem states that in any mathematical model, there will always be a statement that is true but not provable within the framework of this model – that is, no model can be complete. This true but unprovable statement is known as a non-computable factor.

Penrose took Gödel's Theorem and concluded that human consciousness is a non-computable factor, meaning it cannot be explained by classical mechanics. So he turned to quantum mechanics, which certainly contains non-computable elements – most notably, the underlying cause of a wave function collapse. So Penrose formulated Orchestrated objective reduction (Orch OR), a theory that attempts to explain how quantum mechanical processes in the brain result in consciousness in the human brain.

A central idea in quantum mechanics is that any particle can exist in more than one state at once. It has a probability of being found in one state or another at any given time, and so the wave function is actually a superposition of states. When a measurement is taken, the wave function collapses to one and only one state, and then it becomes a matter of classical mechanics. Objective reduction is the idea that an objective threshold of some physical quantity exists, and that exceeding this threshold is what causes the wave function to collapse. As of yet, researchers have not been able to determine what this threshold might be. Penrose realized that this is one of the unanswered questions in quantum mechanics, and so he looked to another unknown in quantum theory for a solution.

General relativity is one of the most celebrated theories of the twentieth century. However its major pitfall is that it is only compatible with classical mechanics, and there is currently no widely-accepted theory that allows general relativity to be compatible with quantum mechanics. Penrose attempted to unify the two in Diosi-Penrose gravity, which will be discussed further in section 6.3. This new interpretation of quantum gravity would become the proposed threshold for Orch OR.

At this point, Penrose had proposed an intriguing, albeit incomplete and speculative, theory of how quantum mechanics gives rise to consciousness. However, it was not until a later collaboration with Stuart Hameroff that Penrose proposed a biological mechanism that could support his theory. This mechanism is proposed to be housed in microtubules, which will be discussed in section 7.

This paper will provide a review of Penrose's proposal as well as bring the perspective of a student of both physics and psychology to provide a rounded interpretation of this controversial topic. In order to talk about consciousness and its underlying physical foundation, it is necessary to first define consciousness. In neuroscience, consciousness is made up of sensation and perception. Sensation is the processing of physical stimuli by the sensory systems, whereas perception is the organization and interpretation of that information that allows an individual to gain an understanding of the world. The difference between sensation and perception is best illustrated by the auditory system. Sound waves are actually just pressure waves in the air, and these waves have a frequency, amplitude and wavelength. These pressure waves are captured by the ear and become tiny vibrations in the inner ear. This is sensation – the auditory system is detecting and processing the physical stimulus of sound. However, perception does not occur until this information reaches the brain and is interpreted as music or a voice.

Using this definition, consciousness is the ability to interpret the world, as well as to respond to it. True consciousness is often described as being self-aware - as being capable of self-reflection, memory, and planning for the future. In the eastern tradition of Buddhist philosophy, consciousness is described as a series of moments, not a continuous absolute reality. Buddhist monks often describe "flickers" in their conscious experience when they practice meditation, as though consciousness is a flip-book of moments [2].

It is also important to understand what is meant by reality. Following in the Buddhist tradition, reality is simply what is absolutely true in the world, not an individual's interpretation of it. For example, the visual system sees and interprets the color blue, but all that is absolutely present in the world is a series of photons with a certain wavelength. This definition of reality can be compared to the Western "objective reality," in which the experience varies from one person to the next. In the Western world, reality is taken as an absolute truth. However, there are many instances where this is not quite the case. A schizophrenic person experiences a different reality from mentally healthy individuals, but it is a reality nonetheless; dreams are also a form of reality. In both cases there is perception without sensation, and the objective reality is based on something that does not exist in the external world.

If consciousness and reality rely on sensation and perception, but are also possible in the absence of sensation, then perhaps the basis of consciousness really does lie not in the external world, but in the brain itself. If consciousness is divided into a series of discrete moments, is each moment not similar to the instant during which a quantum event takes place?

# 3 Basics of Neuroscience

Here I will provide a brief overview of the most relevant topics in neuroscience. For more information and details, see [3,4].

The brain is an intricate network of two main types of cells: neurons and glia. There are two main types of neurons, motor and sensory. Motor neurons are located in the central nervous system and are responsible for sending signals from the brain to the muscles to control movement. Consciousness, in the most widely-accepted meaning of the word, requires some input of sensory information. This information is collected by sensory neurons, which are spread out through the peripheral nervous system. These neurons respond specifically to a certain kind of stimulus, and are unique in that they are able to regenerate when damaged or destroyed.

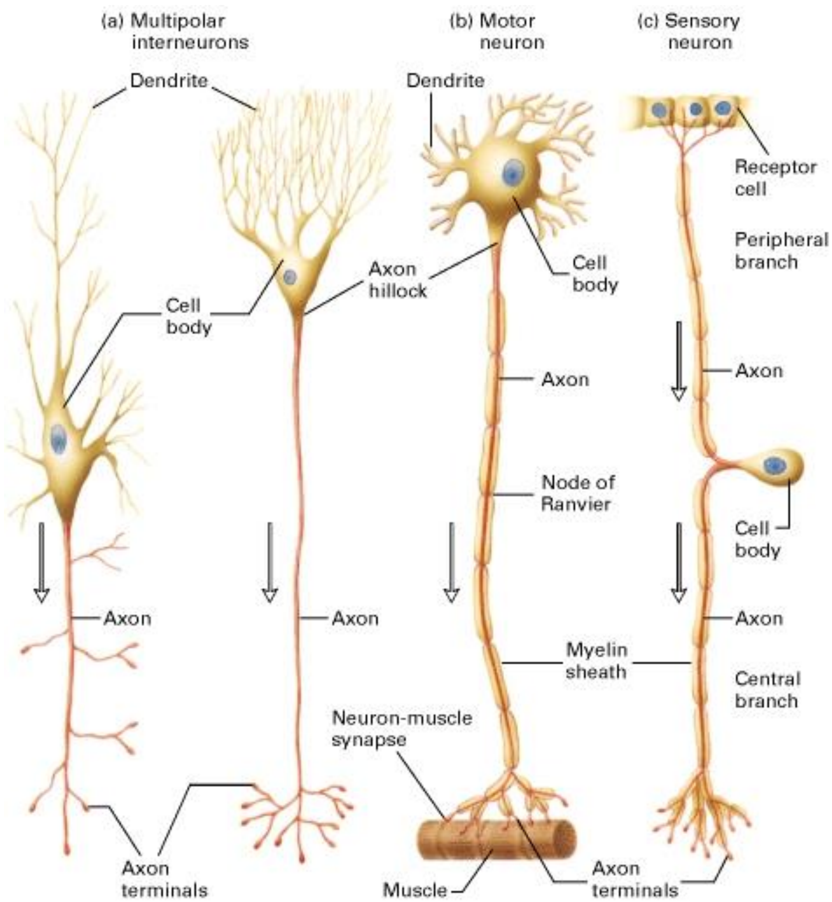


Figure 1: There are three types of neurons, but the most relevant to the topic of consciousness are sensory neurons, which convey information about the external world to the brain. A sensory neuron consists of three main components – the dendrites, the cell body, and the axon. Reprinted from Lodish H, et al., Molecular Cell Biology, 4th ed., W.H. Freeman, New York, 2000.

A sensory neuron (Figure 1) has 3 main parts: the dendrites; the soma, or cell body; and the axon. The dendrites are branch-like structures that receive information from other neurons or sensory receptors located in the skin, eyes, etc.. The soma contains the nucleus, and is responsible for maintaining the cell. The axon contains synaptic terminals that send information to other neurons. A synapse is a small gap between the synaptic terminals of one cell and the dendritic spines of another; the synaptic cleft is the space where chemicals called neurotransmitters pass along information. Some cells use electrical conduction, and in this case the space between cells is called the gap junction.

Glia cells do not pass along information like neurons, but they have a myriad of other functions. There are many different types of glia cells, but the most relevant in this case is the astrocyte (Figure 2). These cells are named thus because they are star shaped, and they were originally thought to simply remove waste material from the brain. However, recent research has shown that these cells actually synchronize the activity of neighboring axons, allowing the information to be sent in electromagnetic waves. As will be discussed later, these waves can actually represent a superposition of states, which is an essential aspect of Penrose’s theory.

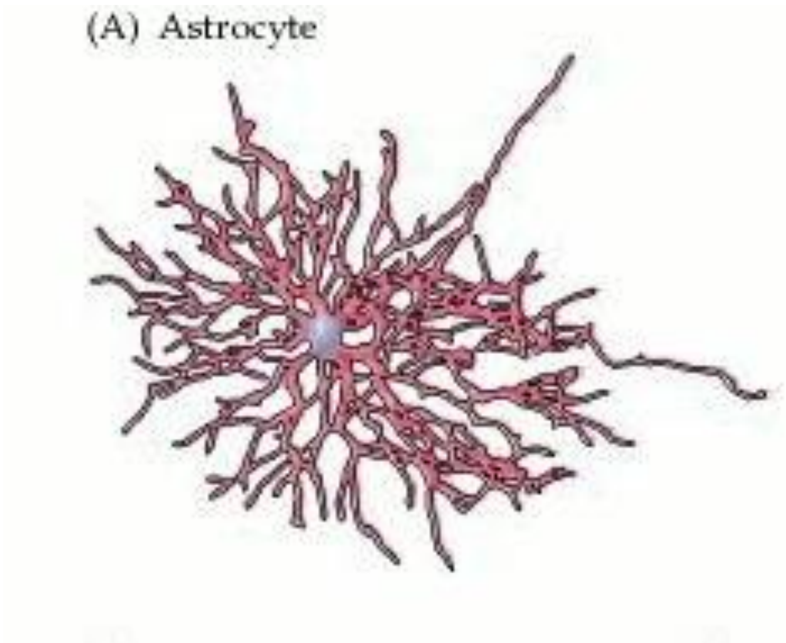


Figure 2: An astrocyte glial cell is shaped something like a star. These cells are believed to synchronize the activity of neighboring axons. Reprinted from Purves D, et al., Neuroscience, 2nd ed., Sinauer, Sunderland (MA), 2001.

## 3.1 Mechanism of Signal Transmission by Neurons

When a neuron's electric potential relative to its surroundings surpasses a certain threshold, the neuron releases chemical signals called neurotransmitters. Each cell is coated in a fatty membrane that is selectively permeable; this membrane maintains a slightly negative potential inside the cell with respect to its surroundings, as the anions contained inside the neuron are too large to pass through the membrane. This is called its resting potential, and is characterized by a higher concentration of potassium ions inside the cell and sodium ions outside of the cell. Both potassium and sodium ions are positively charged, and so they will be attracted to the anions inside the membrane. If these ions pass through the membrane, the resting potential will be disrupted. Neurons have a sodium-potassium pump mechanism that lets two sodium ions out for every two potassium ions it lets in, which maintains the resting potential.

This pump works because the membrane is dotted with voltage-gated channels that open and close to allow potassium and sodium ions to pass through. When the potential difference across the membrane no longer matches the resting potential, these gates will open to restore the balance. However, a neuron that is always at resting potential will never send a signal and release neurotransmitters. This is dictated by the action potential.

The action potential is the objective threshold involved in neural signaling. This is an all-or-nothing mechanism; the neuron's potential must surpass this threshold for a signal to be released; there is no gradient of progressively stronger signals. The voltage-gated channels that maintain the resting potential can also let ions flow in or out of the cell to cause either depolarization or hyperpolarization of the cell. Depolarization results in a positive potential inside the cell, which is the action potential, and the release of neurotransmitters. Hyperpolarization actually makes the cell potential more negative, and so inhibits the release of neurotransmitters (this hyperpolarization stops the spontaneous fire and release of neurotransmitters present in all cells).

Axons can be up to two meters long – if an electrical signal was pushed through this “wire,” it would dampen by the time it reached the end. This does not happen in reality, and so to explain this observation researchers turned to chemical conduction of the signal down the axon. In this model, the axon is set up like a line of dominos. When the soma transmits a signal, it creates a positive charge in one point on the axon. The neighboring points are still negatively charged, and so the positive ions flow into these areas. This creates depolarization in the new point and begins the whole process again. Each point on the axon is actually emitting a new signal, not just passing along the old one. When you push one domino over, it pushes the next over, and so on down the line – which is much more effective than trying to push one domino in order to have an effect on another domino a foot away.

Neurons can either send signals to only their closest neighbors, or can pass information to large areas of the brain. This is where the astrocyte glia cells come in – they make sure the neurons are all giving one unified response. If different neurons were telling different stories, then it would be impossible to have one cohesive conscious experience.

# 4 QUANTUM PHYSICS IN CONSCIOUSNESS

This section assumes basic knowledge of quantum mechanics, classical mechanics, and modern physics. For more information, please see [27-30].

Consciousness has been defined in a number of ways. As mentioned previously, Eastern philosophies tend to view it as a series of discrete moments. To some it is simply awareness of content, a definition that alludes to the difference between external realities and mental images. An external, physical fire is hot to the touch, but the mental image of a fire is not. Further, thoughts about water are not themselves wet [5]. Consciousness is also described as an experience that allows us to differentiate between perception and memory [6]. Perhaps the most comprehensive explanation of consciousness is that it is the “internal, first-person, subjective experience” and “all that one is consciously aware of or consciously experiencing at once” [7]. But how is it that consciousness is a subjective experience? Consciousness being subjective implies that there are many different interpretations of the external world, and that the human mind actually selects which interpretation is experienced. This is where quantum physics comes into play.

The Copenhagen Interpretation has long been the orthodoxy of quantum physics. It includes familiar principles such as the Heisenberg uncertainty principle and de Broglie’s wave-particle duality. However, this interpretation, it has been suggested, also implies that observation is what spurs a particle to come into existence or for an action to take place – in short, that consciousness creates reality [8]. Eugene Wigner [43] echoed the Buddhist philosophical tradition that the physical and chemical properties of an object (i.e. a person or soul) is the foundation for consciousness [2]. That is, a person or animal can only experience consciousness if they possess the biological mechanism that makes consciousness possible. According to Wigner, the body gives rise to mind, but the mind cannot have any effect on the body - consciousness does not create reality, nor does it have any influence over the external world. Classic experiments such as the Double-Slit experiment have demonstrated that an electron behaves as a wave until it is directly observed, at which point it behaves as a particle. Experiments of this nature could provide supportive evidence for consciousness actually creating reality.

For example, the Mach-Zender interferometer is a device that splits and then reunites an incident beam of electrons. After splitting, the two beams interfere with each other until a measurement is taken. A Michelson interferometer is a type of Mach-Zender interferometer that uses a mirror to reflect the beam so the same splitter can be used to split and reunite the beam; this apparatus is more commonly used in lab settings than its more elaborate counterpart. Inanimate measurement devices are a different case entirely from another conscious being [43], and as such it is necessary to test situations where a second consciousness is making these measurements. At least one controversial experiment has been purported to show that the human mind, when in a focused conscious state, actually reduced the interference in the beam even when no measurement apparatus was present [9]. The results of this study suggest that consciousness creates reality, but Vannini and Di Corpo [8] had another theory altogether.

## 4.1 Advanced Waves and Consciousness

The arrow of time is another famous concept in physics. Directly related to the law of entropy, time can only move forward as the universe tends toward chaos. Waves that emit and diverge in accordance with the law of entropy are called retarded waves, but there is another kind of wave that absorbs and converges – in fact, they move backwards in time. Vannini and Di Corpo [8] propose that advanced waves follow the law of syntropy (which is proposed to be the opposite complement of the law of entropy) , which originates from the negative solution of the Klein-Gordon Equation (Eq. 1).

$$E\Psi = \sqrt{p^2 + m^2\Psi} \qquad (1)$$

The Klein-Gordon equation, in which *p* represents momentum and *m* represents mass, is a relativistic version of the Schrödinger equation for particles with no spin. This equation is commonly used in particle physics, and in fact gave rise to the Dirac equation. The Dirac equation is also a relativistic equation, though for spin-½ particles. Solving this equation results in a quadratic equation, which of course has two solutions. Dirac found that one solution corresponded to a positive particle moving forward in time, and that the other solution corresponded to a particle of opposite charge moving backward in time[1]. This second solution was called an antiparticle, and with this result Dirac predicted the existence of the positron, or the electron's antiparticle.

Diverging from the traditional interpretation of the Klein-Gordon equation, Vannini and Di Corpo [8] asserted that "models that originate from the Klein-Gordon equation are not pure quantum physics and unite Schrodinger's wave equation with special relativity to make a biologically compatible model." This was an important statement, because the brain does not seem to be a quantum-friendly environment – for one thing, it is very warm when most quantum interactions require temperatures close to absolute zero. The positive solution of this equation results in the commonly accepted positive solution with time moving forward. The negative solution, according to Vannini and Di Corpo, results in advanced waves moving from the future to the past. This seems to violate the law of causality, which dictates that the

---

[1] Many combinations are possible. The first solution could also be a positive particle moving backward in time, a negative particle moving forward in time, etc.

cause should come before the consequence. However, time is simply a conceptualization of reality, and there are no hard and fast rules about how it should be described. It is of course much easier to describe time as always moving linearly forward, but even special relativity shows that that is an oversimplification of a complex universe. Furthermore, the Copenhagen Interpretation is almost an apotheosis of the human mind, and advanced waves may be more likely than a reality created by consciousness.

The law of syntropy is the exact opposite of the law of entropy in that it states that the universe tends away from chaos and toward order. Advanced waves follow this law, and they absorb and converge to concentrate matter and energy in smaller spaces. This would only be possible at the microscopic level, as time only moves forward at the macroscopic level. Consciousness is proposed to be a product of these advanced waves, and while the idea is not expanded upon, it is implied that some form of quantum information is carried by these advanced waves to create a conscious experience.

The existence of advanced waves and their role in consciousness would lead to an anticipatory response before a stimulus is applied, and there is a surprising amount of supporting research for this. In fact the results of dozens of research projects point to physiological anticipation of a stimulus that has not yet been applied or even hinted at [10, 11]. Physiological measures include things like electroencephalogram (EEG) readings, blood-oxygen levels in the brain, and heart rate, all of which are variables that cannot be easily controlled by the subject. This is could be viewed as direct evidence for the existence of advanced waves, and could explain commonplace phenomena such as déjà vu. However, more research will be needed before the existence of advanced waves can be conclusively confirmed.

## 4.2 QUANTUM THEORY OF CONSCIOUSNESS

Erwin Schrödinger is most widely known for the paradox of Schrödinger’s cat, which questioned the Copenhagen Interpretation. He believed that consciousness did not create reality, but that reality already exists in superpositions of states with infinite degrees of freedom until the conscious mind causes the entangled wave function to collapse [12]. It can be argued that the Schrödinger Equation (Eq. 2) forms the backbone of modern physics.

$$i\hbar \frac{\partial}{\partial t}\Psi(r,t) = -\frac{\hbar^2}{2m}\nabla^2\Psi(r,t) + V(r,t)\Psi(r,t) \qquad (2)$$

Normalization of the Schrödinger Equation results in the probabilities associated with each eigenstate, or possible state. This is not the probability that the state exists before the measurement, as all states exist to varying degrees. This is simply the probability that a given state will be represented as reality. “Mindless Sensationalism” [7] proposes that each conscious perception has a measure that is given by the expectation value of a corresponding quantum “awareness operator” in a fixed quantum state of the universe. Each perception exists to varying degrees of realness, and the measure can be interpreted as the probability that a given perception becomes part of the conscious experience. It is important to note that this is just a framework, and that a complete theory will not be possible until more is known about consciousness.

Wave functions can be entangled, meaning that a system of more than one particle is treated as a single holistic unit. This is similar to what we experience in our day-to-day lives – when you look at a chair, you do not see all the individual particles forming a rigid lattice. In entanglement, it is not so much the positions of each particle that is important, but rather the configuration of the particles. That chair would look quite different if three of the legs stuck straight out the back, and it might not even be considered a chair anymore. Once relativity is taken into consideration, each particle would have its own time as well. This is a messy statement – if all these particles have different times, how could they come together to form one cohesive experience? Remember that time is not concrete, it is a complex abstraction that involves other phenomena such as relativity. Advanced waves may come into play here, or there may be another explanation that has not yet been explored.

Penrose postulated his own form of entanglement, termed quanglement. This is supposed to represent quantum information, but Penrose is careful to say that these qubits (packets of quantum information necessary for quantum computing) are not transmittable information. According to him, this would require that time move backward, which is simply not possible [13]. Clearly Penrose is against the advanced wave theory, but he does not offer up another explanation in its place.

# 5 Gödel's Theorem and the Non-Computability Problem

To understand Penrose's proposal regarding consciousness, it is necessary to address the non-computability problem in both quantum mechanics and consciousness.

A digital computer is an electronic circuit that responds to binary code. A value of zero produces no response, while a value of one produces a full response. There is no gradient in binary, and so there are no partial responses in digital computers. This is profoundly similar to the way the human brain operates. Recall that an action potential, which produces the neurotransmitters that result in brain activity, is an all-or-nothing response. Just as in binary code, there is no gradient for neurotransmitter release. This has spurred many physicists, psychologists, and neuroscientists to treat the mind like a computer and to attempt to build an AI that can perfectly emulate human thought.

According to Gödel's Theorem [13], no algorithm can prove all true statements within the framework of its own rules. That is, if a formal system F is assumed to give only mathematically correct solutions, then some statement G(F) must be accepted as correct even though G(F) is not provable through the methods of F alone. Furthermore, no algorithm can prove whether the system is faulty (meaning the code does not terminate every time it is run). This does not mean that these statements (G(F)) are unprovable, they are simply unprovable within that particular set of formal laws.

One example of Gödel's Theorem in action is a program that is coded to find the smallest natural number that is not the sum of n squares. Lagrange's Four-Square Theorem states that all natural numbers are the sum of four integer squares. Therefore, if n is set to four, the code will never terminate. It does not inform the user that the problem cannot be solved, that all numbers are the sum of four squares, because it is not within its design parameters to do so. This unprovable statement is called a non-computable factor, because it cannot be computed (i.e. proven) within this frame of reference.

Returning to the topic of AI, the Turing Test was designed to determine whether a computer had truly achieved artificial intelligence or not. The Test is simple – the computer simply has to fool a human being into thinking that it too is a sentient being. Though recent programs have come close to passing this test, no AI in history has been able to fully pass for human. There may be a non-computable part of the human mind that cannot be characterized algorithmically [14], and this non-computable part might just have something to do with consciousness [1]. If there are non-computable laws of nature [15] and non-computable parts of consciousness, perhaps there is a connection between the two. Penrose goes on to suggest that attempting to adhere to classical mechanics will be AI's downfall, and that the answer lies in quantum mechanics. This seemingly innocuous statement will later blossom into a full-blown theory uniting consciousness and quantum mechanics in Orchestrated objective reduction.

# 6 Orchestrated Objective Reduction

What follows is a brief overview of a complex theory that has developed over decades known as Orchestrated Objective Reduction, or Orch OR. For a more in-depth recount, please see [1, 13, 16-25].

Penrose believes that objective reduction is the non-computable aspect of the human mind, as well as the physical source of consciousness. Objective reduction occurs when eigenstates of a quantum superposition differ from one another by a miniscule space-time factor where the time measure is particularly small compared to relative large space-difference measurements [14]. A more concise way of putting it is that some objective threshold, probably related to space-time configurations, causes quantum state reduction. Orchestration occurs as a result of a beat phenomenon where axonal firings are integrated and coherent, but are preceded and caused by dendritic-somatic integrations [14]. That is, axonal firings are caused by chemical and electrical interactions, and so are not necessarily in synch with one another. Orchestration is some process that ensures these firings are integrated properly while maintaining their coherence so as to produce a conscious moment.

## 6.1 Decoherence

A substantial obstacle for this theory is the phenomenon of quantum decoherence. This is the process where the phases of quantum wavefunctions are randomized to reduce quantum phenomena to classical processes [25]. As a result of quantum decoherence, the superposition of states would collapse too quickly into a classical process for any effect to be had on the system, which would circle right back to the non-computability problem. Quantum coherence is typically thought to require temperatures near absolute zero, but Hameroff and Penrose cite research that shows that thermal energy actually enhances quantum coherence, and that these processes sometimes occur in hydrophobic pockets [21]. This will be explored further in section 8. Furthermore, electrical signals in the brain are carefully calibrated and correlated, which would eliminate any concern of "noise" disrupting coherence. However, decoherence would still be a significant problem for the Orch OR theory if the brain had no way of preventing it, or at least compensating for it.

Microtubules (discussed in section 7) produce mechanical vibrations not at the original frequency of the signal being transmitted, but faster vibrations that produce beat frequencies. These beat frequencies create a wave-packet around the original sine wave, and so there is a slightly longer stretch of time between nodes. These nodes correlate to the point at which decoherence reduces the superposition to classical processes, and so beat frequencies give the superposition more time to collapse of its own accord (due to the proposed objective threshold) and give rise to a qubit. Ultimately, operating at beat frequencies would increase the likelihood that microtubules are somehow involved in the production of a discrete conscious moment by making it more plausible for a qubit to be produced by these wavefunctions. Furthermore, quantum entanglement (discussed previously in section 4.2) allows for the interconnection of qubits from various microtubules in a lattice. The quantum states of each microtubule therefore affects those in nearby microtubules, which spreads the conscious content throughout the brain. It would not be biologically favorable for one part of the brain to have one experience, and for another part to have a completely different perspective. Quanglement is a domino effect; one qubit does not have to directly send information to each microtubule itself, but instead is able to indirectly have an effect on them. This is similar to the way a signal travels down an axon, in which it is regenerated at each point rather than having to face damping and decoherence as it attempts to reach the end. Because the qubit does not have to travel from one end of the brain to the other, it does not have to stay coherent for quite so long.

Finally, it is possible that qubits actually tunnel through the gap junctions. If this is the case, qubits will travel much faster, drastically reducing the amount of time it must avoid decoherence. Orchestration is absolutely essential at this point – if qubits both affect all microtubules at once at the same time as they are tunneling through gap junctions, it would be very easy for signals to get crossed. A well-designed system would certainly have to be in place to ensure a unified conscious experience.

## 6.2 Reduction Time

We have discussed decoherence and that the superposition must collapse within a specific time to avoid losing its quantum nature. This time is called the reduction time, and in Penrose's theory it is related to the gravitational self-energy of the difference in the mass distribution between states. Consider the case of two states, $\phi_1$ and $\phi_2$ (Eq. 3) where G is the gravitational constant.

$$\tau = \frac{\hbar}{\Delta E_G} \text{ where } \Delta E_G = \frac{1}{G}\int (\nabla\varphi_2 - \nabla\varphi_1)^2 dx^3 \qquad (3)$$

The reduction time is related to the gravitational self-energy $E_G$, which is in turn dependent on the gradient of the eigenstate. The reduction time must be shorter than the decoherence threshold time in order for the proposed quantum computing process to run its course, thus resulting in a moment of consciousness. The gravitational self-energy is necessarily smaller than anything in the biological sphere, or there would be destructive interference resulting in noise. However, there has to be a large enough displacement in mass for this variable to be large enough – otherwise the reduction time would be too long. The literature is not clear on what exactly constitutes this mass. It is possible that the mass is provided by the microtubules, which are able to shrink and expand and are connected to each other. This could create mechanical oscillations and therefore displacements in mass, and as microtubules are proposed to be the biological mechanism for consciousness it is not too grand an assumption. However, mass could also refer to the energy divots in space-time, which can easily be converted from energy to mass using Einstein's equivalency formula.

## 6.3 Diósi-Penrose Gravity

As Penrose's theory began to take form, it became more and more apparent that the actual physical criterion for objective reduction had not yet been specified. Because the topic of human consciousness is not well understood, Penrose turned to another topic that is well understood but has not yet been connected to quantum mechanics – general relativity and gravity.

Einstein's theory of general relativity describes the universe as made up of space-time, which is laid out flat like a blanket[2]. Massive objects create proportionate divots in this fabric, and thereby generate the gravitational effect. Using Einstein's mass-energy equivalence formula (Eq. 4), this effect can be described as energy.

$$E = mc^2 \qquad (4)$$

General relativity is currently accepted as the best current theory of gravity, and there has been research confirming several predictions of the theory. Three observational effects have been observed, the first being the anomalous perihelion advance of Mercury. This represents a slight deviation from the predictions made by Newton's classical gravity. The second is the tiny bending of distant starlight by the sun, called gravitational lensing. Finally, the third is the slowing of clock rates in the gravitational potential. This last effect has been observed, though not conclusively, as it could be a result of general energy and basic quantum considerations rather than general relativity [13]. Additionally, general relativity predicts the existence of gravitational waves, or ripples in space-time caused by the interactions between two extremely massive objects. For example, two black holes orbiting around each other, or two stars merging and becoming one. The existence of these waves seem to be confirmed by Hulse-Taylor observations [13], which provides additional evidence for the theory of general relativity. However, this theory is infamous for its incompatibility with our understanding of quantum mechanics.

If an object exists in two divots at once, it can be described as a superposition of energy eigenstates. So quantum superpositions are like having the same coordinates in two different space-times – however this interpretation would require that two Killing vectors occupy a single space, which violates a fundamental principle of general relativity (two objects that occupy the same space cannot occupy the same time, and vice versa). Killing vectors are timelike, and represent d/dt in the Schrodinger equation [13].

According to Penrose, these superpositions collapse to form qubits. But what causes this collapse? Superposed quantum states each have slight differences in their space-time geometry, and the quantum state uses a gravitational criterion to choose between one of these two sections of space-time, which results in a moment of proto-consciousness [21]. Penrose partnered with Lagos Diósi to produce a new theory of gravity that would unite general relativity and quantum mechanics once and for all. The core axiom of Diósi-Penrose is the gravity actually only has an effect on the uncertainty of an energy, rather than the energy itself. The Killing vector is implicated as the source of uncertainty, and so the error is a measure of uncertainty in the d/dt operator needed for the Schrödinger equation. From this, the absolute uncertainty in the energy of the superposed states can be obtained. Finally, this uncertainty is converted into the gravitational self-energy, $E_G$. The gravitational self-energy in a mass distribution is the energy that is gained in assembling that mass distribution out of point masses completely dispersed at infinity [13]. Of course, this proposal seems to assume that these are stationary states, of which the position is known with little to no uncertainty. According to Heisenberg's uncertainty principle, this would cause a huge uncertainty in the momentum, and so the states would actually immediately disperse. That is, they would not be coherent long enough to give rise to a conscious moment.

---

[2] This is an over-simplification of the topic. Research has found that the universal space-time may not be flat, but instead may have a hyperbolic shape. Other research has suggested that while at a smaller scale space-time may be hyperbolic, it is still generally flat.

According to this proposal, a superposed state would spontaneously reduce into one of its eigenstates in an average time of $\hbar/E_G$. Microtubules (discussed in section 7) actually allow weak gravity-related influences to affect conscious behavior by keeping $E_G$ small and the mass high enough to stay within the necessary reduction time [13]. Of course, $E_G$ must be much smaller than any biological energies, or there would be conflict between them. In fact, a criticism of Diósi-Penrose gravity is that this theory is not consistent with the discreteness of space-time, which would actually shorten the reduction time, resulting in the qubit losing coherence far too early [22]. Another criticism is that the effect of gravity would be negligible compared to the effect of electrostatic, chemical, and Van der Waal's forces, not to mention damping of the quantum effects by thermal noise [24]. Penrose of course realized these objections and answered each critique, but quite frankly does not provide sufficient evidence for how gravity, which is one of the weakest forces, could overcome these stronger forces that are more relevant to processes in the brain.

# 7 MICROTUBULES

Originally Penrose and Hameroff thought it was possible that dendritic-somatic membranes between neurons were the source of consciousness, because they generate local field potentials that give rise to electroencephalogram (EEG) spikes. These dendrites would connect to axon terminals, which would fire in response. These axonal firings were thought to convey outputs of conscious processes to control behavior. However, the pair later realized that it is impossible for this to be the source of consciousness. Firstly, this would leave little room for free will, as a person's conscious experience would be produced by chemical reactions as opposed to interpretations of the environment. Furthermore, integrating all the components would provide the necessary computing abilities, but would greatly affect the membrane potential, which would be observable. [21]. There is no such effect, and so they decided to look inside the neuron itself.

Single cell organisms are able to perform all the necessary functions for survival, and yet they have no gap functions or synapses. They do, however, have cytoskeletons, which help to maintain the structure of the cell. Inside the cytoskeletons are microtubule (MT) lattices, with microtubule associated proteins (MAPs) dispersed throughout.

MTs consist of tubulin proteins (Figure 3), which contain a dipole that gives the MTs ferroelectric properties. Neuronal MTs are incredibly stable – the MAP caps prevent depolymerization for long-term encoding of information. Additionally, the tubulins in MTs can exist in many states – or as Penrose and Hameroff will propose, a superposition of states. MTs are widely believed to be highly influential in the brain – they can influence axonal firings and synaptic transmissions, regulate synaptic plasticity through placement of MAPs at particular locations on the MT lattice, and possibly pass on qubits to larger neuronal networks.

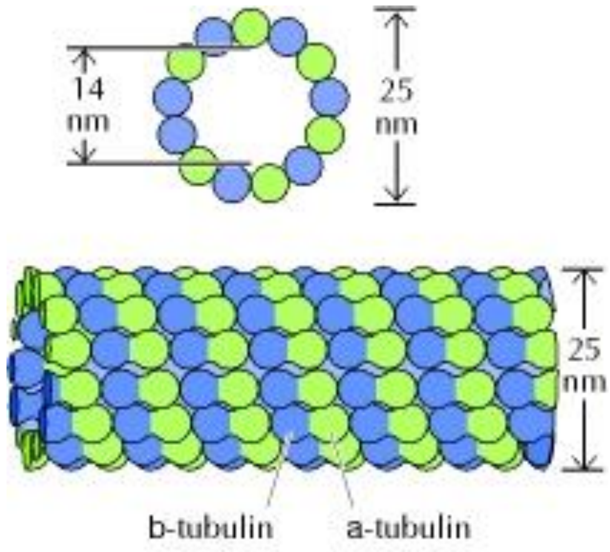


Figure 3: Basic structure of a microtubule. Tubulins combine to form a protofilament, and 13 protofilaments are bound together to form a microtubule. Reprinted from Cooper GM, The Cell: A Molecular Approach, 2nd ed., Sinauer, Sunderland (MA), 2000.

The origin of the dipoles in the tubulin proteins has been debated throughout the development of the theory. Early on, Penrose and Hameroff believed that hydrophobic pockets were created during protein folding, which allowed highly polarizable π orbital electron cloud dipoles to couple by Van der Waals London forces [16]. It was hypothesized that these electric dipoles somehow led to consciousness by binding anesthetic gas molecules in the hydrophobic regions. In one critique of the theory, Reimers et al. [22] pointed out that London force dipoles are typically viewed as fluctuations rather than states, and so could not give rise to qubits. However, because the dipoles couple and oscillate between two orientations, they form a superposition of the two states, which allows them to be a qubit.

However, Penrose and Hameroff now propose that the dipoles are actually magnetic in nature, and related to electron spin [21]. Polarization and depolarization result in current changes in the MTs, which could influence spin-flips. Ouyang & Awshalom [26] found that certain quantum flips are enhanced in a warm environment, such as the brain. It was also proposed that Frohlich condensation could produce a superconductive effect that would help propel the qubits forward through the MT lattice, allowing moments of consciousness to coalesce.

Little is understood about consciousness and how the brain produces this experience, and so each of these proposals only forms a foundation upon which a later complete theory may be built. Regardless of the physics behind this phenomenon, Hameroff and Penrose seem convinced that microtubules are the integral biological structures for this process.

# 8 Testable Aspects of Orch OR

A criticism of Orch OR is that the brain is too warm for quantum processes to take place. Ouyang and Awschalom [26] used Faraday rotation spectroscopy to study the instantaneous transfer of spin coherence. Not only did this study reveal that quantum coherence is possible at room temperature, but that coherence is actually enhanced at warmer temperatures. In fact, temperatures of up to 80 K had a spin-transfer percentage twice as high as those at low temperatures. While this clearly provides evidence that these processes are in fact possible in the human brain, more importantly it underlines Vigner's assertion that it is "premature to believe that the present philosophy of quantum mechanics will remain a permanent feature of future physical theories," [43] as the theory of quantum mechanics is constantly evolving and long-held beliefs may be proven incorrect in the future.

Psychoactive drugs often produce hallucinations and delusions, which are twisted perceptions of reality. If consciousness is to be defined as being the individual perception of reality, then these drugs alter an individual's state of consciousness. Therefore, if microtubules are the biological mechanism of consciousness, the actions of psychoactive drugs should involve neuronal microtubules. Fluoxetine is an antidepressant – specifically, it is a selected serotonin reuptake inhibitory (SSRI) that has a direct effect on brain chemistry. Antidepressants loosely affect the perception of reality by adjusting an individual's mood and decreasing the effect of negative experiences. One particular study [33] injected fluoxetine into the brains of rats *in vito* and observed that the drug did in fact work through the microtubules. The study observed measurable chemical changes in the makeup of the cytoskeleton, as well as decreased microtubule stabilization and a change in the way MAPs were expressed. Additionally, chronic exposure to fluoxetine actually decreased microtubule assembly rates. However, this study did not address whether the drug acted exclusively through the microtubules, so it cannot be regarded as absolute evidence that microtubules are the biological basis for consciousness. Furthermore, a study involving schizophrenic patients when they are taking anti-psychotic drugs might be more relevant to the study of consciousness.

Continuing with potential evidence for the role of microtubules in consciousness, Hameroff and Penrose hypothesized that dynamic microtubule vibrations would correlate with cellular activity. That is, if microtubules are instrumental in producing consciousness, their activity should coincide with that of neural cells. Hameroff conducted a study involving transcranial ultrasound (TUS) which has shown to have an effect on brain function [38]. In Hameroff's study, TUS was applied to chronic pain patients and when compared to a placebo produced measurable differences in the patients' moods and perception of pain. In his discussion of the results, Hameroff proposes that TUS has an effect on cellular activity, which correlates with microtubule vibrations, thus producing a slight alteration in an individual's mental state. While it is certainly interesting that TUS can improve mood, this is not evidence for Orch OR. It was presented as such by Penrose [14], but this is merely speculation. In fact, current fMRI is incapable of directly observing correlations between microtubule variations and cell activity, due to fMRI spatial and temporal resolution being too low. However, this could easily be the subject of future study, and if this hypothesis were to be supported by evidence it would be a compelling case for microtubules as the source of consciousness.

A core component of Orch OR is that groups of neurons must be in synch in order to produce a coherent conscious experience. Typically, this synchronization is described as being chemical transmission from one neuron to the next. However, it is proposed that electrotonic[3] gap junctions mediate coupling between astrocytic glial cells [35]. This would allow for communication between

[3] Electrotonic refers to the direct spread of current through electrical conduction without the production of new action potentials.

diverse cell populations, rather than being limited to neurons with the right chemical receptors. In fact, the observed frequency of interneural gap junctions is in accordance with this proposal [35]. One study observed that low-pass filtering of extracellular field potentials in the rat hippocampus revealed repetitive short bursts of high-frequency discharge [36]. Also observed were independent and coherent oscillations at two or more recording sites that persist when measures are taken to inhibit chemical transmission at gap junctions. Not only is it unlikely that these observations are the product of a single cell's activity, but it is almost certainly through electrotonic transmission. Furthermore, blocking the gap junctions suppressed these high-frequency oscillations [36]. This is evidence that gap junctions do in fact use electrotonic transmission to synchronize neurons. However, more study is needed before this can be considered as evidence for Orch OR. What effect do these high-frequency oscillations have on microtubules? Does this produce any changes in an individual's mental state and therefore consciousness? The answers to these questions may be far off, but are essential to proving or disproving Orch OR.

The reduction time is instrumental to this theory as well – if quantum superpositions do not undergo OR by $t = \hbar/E_G$, then the entire gravitational model no longer applies. Bouwmeester et al. studied quantum teleportation, which involves transferring the state of one system to another system without copying the actual object [34]. This study used two-photon interferometry, and found that any two quantum systems can be entangled through quantum teleportation. In order for this study to work, it was necessary to either detect or to generate interfering photons within a time interval shorter than the coherence time. That is, within the reduction time proposed by Penrose. Bouwmeester et al. did not actually measure the reduction time of naturally occurring superpositions, but rather forced the superpositions to collapse within this time. While this shows that superpositions *can* undergo OR (though not necessarily *orchestrated* OR) within the reduction time, it does not necessarily prove that they *do*. In order to prove this phenomenon, a study would have to be designed that would directly study the reduction time, rather than using it as a qualification or intermediate step.

If microtubules are indeed the biological mechanism for consciousness, then they should emit coherent photons, show phases of quantum coherence, and contain coherent gigahertz excitations. An optical study used combined excitation emission spectroscopy (CEES) to compare tubulin proteins with microtubules [41], and the results were compelling. CEES works by detecting the excitation and emission of photons; this implies that microtubules do in fact emit coherent photons, although it was not directly studied. The study found that tubulins interact but do not change the overall energy level of the microtubule polymer, and that energy emission is the same regardless of the amount of energy absorbed. Furthermore, properties of individual tubulin actually become properties of the microtubule itself. These findings could account for noise alleviation. The study also observed AC resonance peaks in microtubules that could be the method of communication between cells. Another study on multi-level memory switching found that the conductivity of tubulins does not vary between open-air and vacuum conditions [42]. This could be evidence that microtubules themselves are capable of quantum coherence, and that there are no outside influences promoting coherence. Neither of these studies directly address coherent photons, phases of quantum coherence, or coherent gigahertz excitations, but the findings imply that all three occur in microtubules. Further study is necessary to definitively prove that microtubules possess these qualities.

Several hypotheses that could provide significant support for Orch OR have not yet been tested, in most cases because current technology is incapable of making these observations. Vision plays a large part in perception, and so naturally the mechanisms of vision should have some involvement in the production of consciousness. Microtubule-based cilia in retinal rod and cone cells should detect quantum information, rather than simply taking in classical information about the wavelength of a photon (for more information about the visual system, please see 3 and 4]. That is, quantum information should not just be limited to neural cells in the brain, but should be detected by sensory neurons, as well. Additionally, quantum tunneling should occur across gap junctions. A few studies [35-37, 42] have hinted at observational evidence for this, but as yet none have been able to replicate those findings, nor has it been studied directly. Furthermore, it is necessary to combine the Hameroff study [38] with the studies involving the synchronization of neurons [35, 36]. That is, not only should neurons be synchronized to produce the same conscious experience, but this synchronization should act through microtubules. These quantum correlations between the microtubules of different neurons should occur via gap junctions. Each of these untested hypotheses are critical to providing support for Orch OR, but may be years away.

# 9 CONCLUSION

Consciousness is often described as the individual, subjective experience of reality. However in order for consciousness to be subjective, there must be something in the brain that results in differing perceptions. Gödel's Theorem, while written explicitly for mathematics, has been applied to consciousness. This allows for the interpretation that there is something that is not classically computable about the human mind, and so it must involve quantum mechanics instead. Advanced waves attempted to explain this non-computable factor, but fell short due to its reliance on time moving backwards. Penrose postulates that reality exists in a superposition of states, and that this quantum information is absorbed by sensory neurons and then collapses to one state or another to result in a conscious moment. This objective reduction is proposed to have a gravitational threshold, and is orchestrated – that is, synchronized among the many neurons in one individual – to provide a unified conscious experience. There is some evidence to support aspects of this theory, but most of this is simply inferred or speculated from existing observations, rather than being based on direct observation. Current technology, such as fMRI, lacks the spatial and temporal resolution to observe these phenomena directly, and so it may be several decades before this theory is fully testable. Furthermore, Orch OR seems to challenge quantum theory, to the point that proving Orch OR may actually result in the development of new physical laws. While some may dismiss the radical models proposed by Penrose and his colleagues, physics was never meant to be a stagnant field. New observations frequently disprove long-held conventions, and so there is no reason to dismiss this theory unless and until direct observation proves it to be incorrect.

# 10 REFERENCES


[1] Penrose R, *The Emperor's New Mind* (Oxford University Press, New York, 1989).

[2] *Buddhist Philosophy: Essential Readings*. Edited by Edelglass W and Garfield J. Oxford University Press: 2009.

[3] Wolfe JM, Kluender KR, and Levi DM, *Sensation & Perception*, 3rd Ed. (Sinauer Associates, Inc., 2011).

[4] Kalat JW, *Biological Psychology*, 11th Ed. (Cengage Learning, 2013).

[5] James W, "Does 'Consciousness' Exist?,' in Consciousness and the Universe: Quantum Physics, Evolution, Brain & Mind, edited by Penrose R, Hameroff S, and Kak S (Cosmology Science Publishers, Cambridge, 2011).

[6] Shevrin H, "What Does Consciousness Do?,' in Consciousness and the Universe, op cit.

[7] Page DN, "Consciousness and the Quantum,' in Consciousness and the Universe, op cit.

[8] Vannini A and Di Corpo U, "Quantum Physics, Advanced Waves and Consciousness,' in Consciousness and the Universe, op cit.

[9] Dobyns YH, "Consciousness and Quantum Measurement: New Empirical Data,' in Consciousness and the Universe, op cit.

[10] Mossbridge JA, P Tressoldi, J Utts, JA Ives, D Radin, and WB Jonas, Predicting the unpredictable: critical analysis and practical implications of predictive anticipatory activity. Frontiers in Human Neuroscience **8** (146), (2014).

[11] Mossbridge J, Tressoldi P, and Utts J, Predictive physiological anticipation preceding seemingly unpredictable stimuli: a meta-analysis. Frontiers in Psychology **3** (390), (2012).

[12] Nauenberg M, "Does Quantum Mechanics Require a Conscious Observer?,' in Consciousness and the Universe, op cit.

[13] Penrose R, *The Road to Reality* (Alfred A. Knopf, New York, 2005).

[14] Lucas J, The face of freedom. Comment on "Consciousness in the universe. A review of the 'Orch OR' theory" by Hameroff and Penrose. Phys Life Rev **11** (1), 87-88 (2014).

[15] Tandy C, Are you (almost) a zombie? Conscious thoughts about "Consciousness in the universe" by Hameroff and Penrose. Phys Life Rev **11** (1), 89-90 (2014).

[16] Hameroff S 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, edited by Hameroff S, Kaszniak AW, and Scott AC (MIT Press, Cambridge, 1996).

[17] Hameroff S, Penrose R, Reply to criticism of the 'Orch OR qubit' – 'Orchestrated objective reduction' is scientifically justified. Phys Life Rev **11** (1), 104-112 (2014).

[18] Hameroff S, Penrose R, Reply to seven commentaries on "Consciousness in the universe: Review of the 'Orch OR' theory." Phys Life Rev **11** (1), 94-100 (2014).

[19] Globus G, "Consciousness and Quantum Physics: A Deconstruction of the Topic,' in Consciousness and the Universe, op cit.

[20] Ghosh S, Sahu S, Bandyopadhyay A, Evidence of massive global synchronization and the consciousness. Comment on "Consciousness in the universe. A review of the 'Orch OR' theory" by Hameroff and Penrose. Phys Life Rev **11** (1), 83-84 (2014).

[21] Hameroff S, Penrose R. Consciousness in the universe. A review of the ‘Orch OR’ theory. Phys Life Rev **11** (1), 39-78 (2014).

[22] Reimers JR, McKemmish LK, McKenzie RH, Mark AE, Hush NS, The revised Penrose-Hameroff Orchestrated objective-reduction proposal for human consciousness is not scientifically justified. Comment on “Consciousness in the universe. A review of the ‘Orch OR’ theory” by Hameroff and Penrose. Phys Life Rev **11** (1), 101-103 (2014).

[23] Pino S, Di Mauro E, How to conciliate Popper with Cartesius. Comment on “Consciousness in the universe. A review of the ‘Orch OR’ theory” by Hameroff and Penrose. Phys Life Rev **11** (1), 91-93 (2014).

[24] Tuszynski JA, The need for a physical basis of cognitive process. Comment on “Consciousness in the universe. A review of the ‘Orch OR’ theory” by Hameroff and Penrose. Phys Life Rev **11** (1), 79-80 (2014).

[25] Jumper CC, Scholes GD, Life – Warm, wet and noisy? Comment on “Consciousness in the universe. A review of the ‘Orch OR’ theory” by Hameroff and Penrose. Phys Life Rev **11** (1), 85-86 (2014).

[26] Ouyang M, Awschalom DD. Coherent spin transfer between molecularly bridged quantum dots. Science 2003; 301:1074-8.

[27] Taylor JR, *Classical Mechanics*, (University Science Books, 2005).

[28] Serway RA, Moses CJ, and Moyer CA, *Modern Physics*, 3rd Ed. (Cengage Learning, 2004).

[29] Mandl F, *Statistical Physics*, 2nd Ed. (Wiley, 1988).

[30] Griffiths DJ, *Introduction to Quantum Mechanics*, 2nd Ed. (Pearson Prentice Hall, 2005).

[31] Deutsch D, *The Fabric of Reality* (Penguin Books, New York, 1997).

[32] Bennett MVL and Zukin S, Electrical coupling and neuronal synchronization in the mammalian brain. Neuron **41**, 495-511 (2004).

[33] Bianchi M, Shah AJ, Fone KCF, Atkins AR, Dawson LA, Heidbreder CA, Hows ME, Hagan JJ, and Marsden CA, Fluoxetine administration modulates the cytoskeletal microtubular system in the rat hippocampus. Synapse **63**, 359-364 (2009).

[34] Bouwmeester D, Pan J, Mattle K, Eibl M, Weinfurter H, and Zeilinger A, Experimental quantum teleportation. Phil Trans R Soc Lond A **356**, 1733-1737 (1998).

[35] Dermietzel R, Gap junction wiring: a ‘new’ principle in cell-to-cell communication in the nervous system? Brain Research Reviews **26**, 176-183 (1998).

[36] Draguhn A, Traub RD, Schmitz D, and Jeffrys JGR, Electrical coupling underlies high-frequency oscillations in the hippocampus *in vitro*. Nature **394**, 189-192 (1998).

[37] Galarreta M and Hestrin S, Electrical synapses between GABA-releasing interneurons. Nature Reviews Neuroscience **2**, 425-433 (2001).

[38] Hameroff S, Trakas M, Duffield C, Annabi E, Bagambhrini Gerace M, Boyle P, Lucas A, Amos Q, Buadu A, Badal JJ, Transcranial ultrasound (TUS) effects on mental states: a pilot study. Brain Stimulation **6**, 409-415 (2013).

[39] Mandelkow E, Song Y, Mandelkow A, The microtubule lattice – dynamic instability of concepts. Trends in Cell Biology **5**, 262-266 (1995).

[40] Moroz IM, Penrose R, and Tod P, Spherically-symmetric solutions of the Schrodinger-Newton equations. Classical and Quantum Gravity 15 (9), 2733-2742 (1998).

[41] Sahu S, Ghosh S, Ghosh B, Aswani K, Hirata K, Fujita D, and Bandyopadhyay A, Atomic water channel controlling remarkable properties of a single brain microtubule: Correlating single protein to its supramolecular assembly. Biosensors and Bioelectronics **47**, 141-148 (2013).

[42] Sahu S, Ghosh S, Hirata K, Fujita D, and Bandyopadhyay A, Multi-level memory-switching properties of a single brain microtubule. Applied Physics Letters **102**, 123701 (2013).

[43] Wigner EP, “Remarks on the Mind-Body Question,” in The Scientist Speculates, edited by Good IJ (Heinemann, London, 1961; Basic Books, New York, 1962).