CONSCIOUSNESS AND QUANTUM MEASURMENT

Chris Allen Broka
(chris.broka@gmail.com)

Abstract

A variant of the von Neumann-Wigner Interpretation is proposed. It does not make use of the familiar language of wave functions and observers. Instead it pictures the state of the physical world as a vector in a Fock space and, therefore not, literally, a function of any spacetime coordinates. And, rather than segregating consciousness into individual points of view (each carrying with it a sense of its proper time), this model proposes only unitary states of consciousness, Q(t), where t represents a fiducial time with respect to which both the state of the physical world and the state of consciousness evolve. This would seem to impose a preferred Lorentz frame on the world. But it will be argued that no physical violations of relativity are engendered thereby.

Introduction.

The relationship between consciousness and quantum mechanics has long been discussed in a serious context. Schrödinger (1), Wigner (2), and von Neumann (3) were, early on, most associated with these lines of investigation. And Everett's Relative State Interpretation is very much about our states of consciousness. Still, it seems that little effort has gone into developing mathematical formalisms appropriate to the description of this relationship. Perhaps this is because mental states seem mathematically indefinable. In this paper I will try to incorporate consciousness into a straightforward generalization of ordinary quantum physics. It is both possible and illustrative to do this.

States of the World.

Designate the state of the physical world as $|\Psi(t)>$. Such states are to be understood as vectors in a Fock space. Vectors in this space are constructed from the vacuum state $|0>$ by the repeated application of creation operators appropriate to the various kinds of particles that inhabit our universe. The parameter t recognizes that this state evolves with respect to a fiducial time that we can identify with a particular Lorentz frame. It is essential to recognize that $|\Psi(t)>$ is *not* a function of the spatial coordinates (x, y, z). We will work in the Dirac Interaction Picture. Here we regard our Fock space as built using the creation operators appropriate to free, non-interacting particles and write $i\,\partial_t\ |\Psi(t)> \ =\ H'\ |\Psi(t)>$ where H' is the Hamiltonian describing the interactions amongst these particles. The complete Hamiltonian for the system is written $H = H_0 + H'$. Every effort will be made *not* to express anything in terms of 'wave functions.' It has been recognized that this concept is problematic since the inception of quantum field theory (4).

There are, of course, circumstances where the wave function idea works adequately and provides useful answers to practical questions. Mostly these are circumstances under which an experiment can be conducted where only one or a very few particles do something interesting and are observed. The observer can, in these fortunate circumstances, be absorbed into the background without much mathematical consequence. (Experimental physicists work diligently to construct such situations.) But our tendency to mistake these artificial situations for reality has introduced a host of seeming paradoxes.

States of Consciousness.

Suppose that, at any time t, the conscious state of the universe can be designated as $Q(t)$. $Q(t)$ describes the *qualia* - the totality of sensations experienced by any and all consciousness anywhere at that time; that could include physicists recording a quantum measurement or worms tasting sugar in a pond on a distant planet. We will try not to make particular distinctions between various "observers." Proceeding in analogy to quantum mechanics we suppose that there are 'states of consciousness' and that these states can be represented as vectors in our Fock space. There acts upon this space a linear 'consciousness operator'– $\mathbb{C}$– which has the property that, for some of these vectors,

1) $\mathbb{C}\ |C_i\rangle = Q_i\ |C_i\rangle$.

The $|C_i\rangle$ constitute eigenvectors of $\mathbb{C}$– they correspond to what we will call *definite states of consciousness*. By this we mean that Q_i specifies a unique and unambiguous state of awareness possessed by the totality of its sentient observers. Now it is immediately clear what a strange sort of "operator" $\mathbb{C}$ is. We are accustomed to seeing c-numbers as eigenvalues, maybe a few other things, but sensations? We assign $\mathbb{C}$ no explicit time dependence. Any state that is not an eigenstate of $\mathbb{C}$ will be called a mixed state.

There should be no segregation of consciousness into any set of individual observers. We will just designate it *Q(t)*. This will prevent us from trying to describe the world in terms of separable and independent "wave functions," one for your brain, my brain, or other things; it would be difficult having to make sense of trillions of independent consciousness operators. While it is true that we seem to experience our individual conscious states independently, we have to argue that this apparent separateness is merely an illusion born of facts like "I" can remember my memories but not "yours." And "I" can experience my sensation of blue but not "yours." These facts are indisputable, and surely interesting. But they only obscure matters if we wish to study the problem at hand. And we do not want to be misinterpreted as proposing anything mystical here. The important point is mathematical– we will regard consciousness (in its totality) as something that can be indexed by a *single* parameter t.

We want these states of consciousness to stand in some relation to those of the material world. We will say that every $|C_i\rangle$ is identical with a particular physical state vector and that the $|C_i\rangle$ constitute a *complete set of orthonormal basis states* in terms of which any $|\Psi(t)\rangle$ may be decomposed. Regarding the nature of the $|C_i\rangle$ one thing is obvious; they are highly degenerate with respect to the Q_i. It would, for instance, make no difference to the overall state of consciousness whether an electron had been originally created in a region with no observers. And we must recognize a sort of null state of consciousness—a state where there just aren't any sentient observers at all. (The state vector a few seconds after the big bang would correspond to such a state. So would many others.) So we must picture the states $|C_i\rangle$ as existing in a space that is broken up into many separate subspaces each with a particular Q_i that designates the unique conscious experience corresponding to it. We will denote these subspaces C_i. Recalling that every state $|\Psi\rangle$ can be written as a superposition of the states $|C_i\rangle$, and knowing that every region having Q_i in common must be spanned by a number (n_i) of the $|C_i\rangle$, we can write:

2) $|\Psi\rangle = \sum_{i=0}^{\infty} \left(\sum_{j=1}^{n_i} \alpha_j^i\ \Phi_i^j \right)$.

Φ_i^j represents the j-th state in our Fock space that corresponds to one of the n_i eigenvectors $|C_i>$ that span the subspace C_i (5). We sum over these subspaces. Since $\mathfrak{C}$ does not depend on t, neither do the Φ_i^j. It should be remembered that two states that are eigenstates of $\mathfrak{C}$ having different qualia-eigenvalues must necessarily be orthogonal to one another.

The time variable that appears in $|\Psi(t)>$ and $Q(t)$ requires a comment. As it pertains to the former case it causes no problems with relativity since the equations that determine the evolution of $|\Psi(t)>$ are themselves relativistically invariant—t only represents an arbitrary choice of Lorentz frame. $Q(t)$ does cause a problem, however. By choosing not to regard consciousness as broken up into separate observers (each of which needing to be assigned its own proper time) we have more-or-less forced ourselves to select a particular set of space-like hypersurfaces to designate the various 't's. Now perhaps because consciousness is a non-material sort of thing such a violation of relativity is permissible—we can't be sure how physics treats non-material things. But it is essential that we construe $Q(t)$ in such a way as to end up with no *physical* violations of relativity.

The Evolution of these States with Time.

Taking no account of consciousness we could picture $|\Psi(t)>$ evolving according to $i\,\partial_t\,|\Psi(t)> = H'(t)\,|\Psi(t)>$ where H' designates the interaction operator for the material world. ($H'(t) = \int \mathcal{H}'(x,t)\,d^3\boldsymbol{x}$ where $\mathcal{H}(x,t)$ is the corresponding Hamiltonian density operator). We assume normal ordering. All operators and state vectors are being represented in the Interaction Picture. There would, in consequence, exist a unitary operator, $U(t_2, t_1)$, having the property that $|\Psi(t_2)> = U(t_2, t_1)\,|\Psi(t_1)>$. Let us imagine the world at time t_1 being in a definite state of consciousness. That is to say $|\Psi(t_1)>$ can be written as $\sum_{j=1}^{n_i} \alpha^i_j(t_1)\,\Phi_i^j$ for some *specific* i. Q_i, to which the Φ_i^j all correspond, is understood to designate a definite state of consciousness. Now the Φ_i^j are not in any necessary way eigenstates of H_0 or H'. $|\Psi(t_1)>$ will, therefore, evolve into a state vector having no particular Q_i specified. Under this scenario we can write $|\Psi(t_2)>$ as $\sum_{i=0}^{\infty}(\sum_{j=1}^{n_i} \alpha^i_j(t_2)\,\Phi_i^j)$ where we have some probability of finding the conscious state of the world in any of quite a number of configurations. But this is never what we seem to experience; our common awareness appears unconfused and composed of a well-defined set of qualia. So let us suppose that reality will only tolerate definite states of consciousness–that is to say $|\Psi(t)>$ must *always* lie entirely within one of the eigenspaces of $\mathfrak{C}$.

We can arrange for this to happen by amending the previous equation for the time-evolution of $|\Psi(t)>$ to also require $\mathfrak{S}\,|\Psi(t)> = |\Psi(t)>$ where $\mathfrak{S}$ is a (non-linear) operator having the property:

3) $\mathfrak{S}\,(\sum_{i=0}^{\infty}(\sum_{j=1}^{n_i} \alpha^i_j\,\Phi_i^j)) \to \frac{1}{N}\sum_{j=1}^{n_i} \alpha^i_j\,\Phi_i^j$ at random with the probability $\sum_{j=1}^{n_i} |\alpha^i_j|^2$ for all specific i (N is only for normalization).

$\mathfrak{S}$ functions as a projection operator taking mixed states (with respect to $\mathfrak{C}$) into definite states of consciousness. We give up the idea of a unitary time-evolution operator. Such an operator has an inverse. We cannot go backwards in time according to Eqn 3) since the decision how to go forward is made at random. This imparts a natural directionality to time. $\mathfrak{S}^2 = \mathfrak{S}$ and $\mathfrak{S}$ has no explicit time dependence. Since $|\Psi(t)>$ is always an eigenstate of $\mathfrak{C}$ we may write $\mathfrak{C}\,|\Psi(t)> = Q(t)\,|\Psi(t)>$. The qualia-state is assumed independent of phase so, if $|\Psi(t)>$

corresponds to a particular Q_i, $e^{i\theta}$ |Ψ*(t)*> will correspond to it also. We suppose that ℭ |0> = ϕ |0> where ϕ designates the null state of consciousness.

The Anatomy of a Measurement.

Consider a very simple experiment in which an electron is sent through a Stern-Gerlach apparatus. It can be prepared as either spin-up or spin-down or in any superposition of these states. If it comes in spin-up it always veers up and strikes a detector that causes a light to shine green. If it's down it goes the other way and a red light is triggered. This device, the electron whose spin it measures, and something like an observer, constitute a physical universe described by |Ψ*(t)*>. The conscious states of this universe, we will imagine, belong to this single observer whose only possible states of awareness are 1) seeing a green color, 2) seeing a red color, or 3) seeing nothing (the null-state of consciousness). So the space in which the conscious state of the universe is a vector contains three subspaces– one corresponding to each of the above possibilities. These are the eigen-spaces defined by ℭ. Each is spanned by a number of eigenvectors of ℭ, each of these a state |Ψ> in the Fock space characterizing the physical world. Since this world is simple we think that we can get away with describ-ing it in a simple manner. Let us describe its initial state as |Ψ(0)> = |+, I> where + says that our electron is spin-up. 'I' simply says that the rest of the measurement system ("observer" and all) are in their initial state. ℭ|Ψ(0)> = ϕ |Ψ(0)> since this initial state is an eigenstates of ℭ corresponding to the null-state of consciousness. When the spin-up electron is detected at t_d |Ψ(0)> evolves into |Ψ(t_d)> which we can write as |+, G>. ℭ|Ψ(t_d)> = G |Ψ(t_d)> meaning that this new state is an eigenvector corresponding to the qualia 'seeing a green color.' (If the election had been spin-down, relative to our arbitrary coordinate system, we'd have ended up with a red qualia and a state |-, R>.) If things happen to start out as (|+, I> + |-, I>)/$\sqrt{2}$ our system will, obviously, evolve into a superposition of states which is no longer an eigenvector of ℭ. Since, according to the above-mentioned princi-ple, reality cannot tolerate any state that is not an eigenstate of ℭ it is necessary that 𝔖 project |Ψ(t_d)> into either |+, G> or |-, R> with equal probability. Let us make it clear that no wave function collapses. Instead, a state vector |Ψ(t)>– which is *not* a function of the spatial coordinates (x, y, z)– tries to evolve into a state (in Fock space) where it no longer resides entirely within a particular C_i but rather exists as a superposition of 'red' and 'green' qualia states. 𝔖 immediately corrects this by projecting |Ψ*(t)*> back into only one of the *definite* states of consciousness available to it according to Eqn 3).

There is something undeniably awkward about such a phenomenon. And it is not obvious that adjoin-ing consciousness to the problem by way of 𝔖 does much to improve things. Everett elects to throw out 𝔖 and freely allow non-definite states of consciousness. These are, presumably, able to sort themselves out into separate, conscious worlds. Clever as the Relative-State Interpretation is, it suffers from a serious problem. Suppose that the electron is sent out in such a state that the green light should illuminate 99% of the time and the red one only 1%. I know perfectly well that, in situations like this, I will see the green light almost all the time. But one cannot be "just a little bit conscious." One either is or one isn't. If there are two conscious "observers"– one seeing green and one seeing red– there ought, really, to be a 50/50 chance of "my" being either. In fact, there does not seem to be a satisfactory solution to this inconsistency (6). For this reason we will want to reject the Everett Interpretation and not burden ourselves with the uneconomical existence of realities we can have no contact with or knowledge of.

Since 𝔖 seems to measure the state vector constantly (7) a concern may arise regarding the quantum Zeno effect; if the state is always being observed can it really ever go from null to green or red? The answer would seem to be 'yes' as a deceptively simple example will demonstrate: Consider a universe with only one spatial dimension. In it there exists a special particle (there is only one particle of this type in the entire universe and it can never decay away) whose state, with respect to H_0 can be decomposed into momentum eigenstates $|\mathbf{k}>$. There is also an observer coupled to a momentum-measuring device. (We assume that its measurements disrupt the special particle's momentum in only an insignificant way.) Let us suppose that the consciousness eigenstates of interest here are two in number, red and green. If $N \leq \mathbf{k} < N + 1$ ($N \in \mathbb{Z}$) for N even the observer sees green. For N odd he sees red. Let us now subject the special particle to a force which causes it to accelerate so that, if it is in a state $|\mathbf{k}>$ at $t = 0$, a later time t will see it in $|\mathbf{k} + t>$. If the particle starts out in state $|1/2>$ it will, at $t = 1/2$, cross the boundary from a green qualia-state into a red one. If the particle started out, say, in the state $(|1/3> + |2/3>)/\sqrt{2}$ initially all is well because we are in a green eigenstate. But at $t = 1/3$ part of the vector crosses into the red eigenspace and a mixed state tries to occur. $\mathfrak{S}$ immediately projects this into one of the available possibilities with equal likelihood. After this the chosen state continues to evolve normally. While seeming a bit unphysical, we do think this example captures what takes place in reality. It illustrates the "hard-edged" nature of the boundaries of the qualia eigenspaces C_i. In order for the consciously evolving world we live in to be possible (according to this interpretation) a state vector must be able cross from one C_i to a different one, or into a superposed state residing primarily in a different one, instantaneously. And note that eigenstates in two different C_i must be orthogonal to one another.

We will define continuous time-evolution in our Fock space as follows:

4) $$\lim_{\Delta t \to 0} < (\Psi(t + \Delta t) - \Psi(t)) \mid (\Psi(t + \Delta t) - \Psi(t)) > = 0 \text{ for all } t.$$

Clearly this equation cannot hold if physics is to admit of consciousness under our assumptions. The action of $\mathfrak{S}$ upon a mixed state will violate this. But let us consider the most flagrant violation of continuous evolution– that being the instantaneous transit of $\Psi(t)$ into an orthogonal state. Here the left-hand side of Eqn 4) goes to 2. If we write $\Psi(t + \Delta t)$ as $e^{-i H' \Delta t}\, \Psi(t) \approx (1 - i H' \Delta t)\, \Psi(t)$. The left hand side of Eqn 4) becomes $\lim_{\Delta t \to 0} <\Psi(t) \mid H'^2 \mid \Psi(t)> \Delta t^2$. Certainly this will go to zero with Δt *unless* $<\Psi(t) \mid H'^2 \mid \Psi(t)>$ is infinite. (If it is finite $<\Psi(t) \mid H'^2 \mid \Psi(t)> \approx \tau_z^{-2}$, in many situations. τ_z represents the Zeno time– a period during which the probability of a state's remaining the same decreases quadratically with time. It is during this period that frequent observations will inhibit a state's evolution.) We can see the same thing through a different argument. Consider $\lim_{\Delta t \to 0} <\Psi(t) \mid \Psi(t+\Delta t)>$. For the above-described instantaneous transit to occur this must be 0. But write $|\Psi(t+\Delta t)>$ as $e^{-i H' \Delta t}\, |\Psi(t)>$ and expand. If all the $<\Psi(t) \mid H'^n \mid \Psi(t)>$ terms are finite the limit will be 1 as Δt goes to 0.

But it occurs to us that $<\Psi(t) \mid H'^2 \mid \Psi(t)>$ *is*, in fact infinite. $\Psi(t)$ represents the entire universe (and cosmology tells us this is very likely infinite). $<\Psi(t) \mid H'^2 \mid \Psi(t)>$ therefore includes the (squared) interaction energies between infinitely many particles. For the universe, considered as a whole, $\tau_z = 0$. For *it* there can be no quadratic regime or quantum Zeno effect.

Now the quantum Zeno effect has been experimentally demonstrated. It is a real thing. How can this be possible? When physicists measure this effect they isolate a small "part" of the larger state vector in such a way as to segregate it from the wider world. (Perhaps they put the whole measuring system, observer, and decaying atom inside a finite-dimensional lead box.) And they arrange things so that only one particle is doing anything interesting. Under these circumstances terms like $<\Psi(t) \mid H'^2 \mid \Psi(t)>$ are easily taken to be finite and the effect can be observed (8).

The EPR Paradox.

There must be no physical contradiction between $\mathfrak{S}$, *Q(t)* and special relativity. If such a contradiction is to be found, the most obvious place to look for it is in the EPR paradox. Suppose that our universe is enlarged to consist of two detectors, two "observers," and a source that can fire a pair of particles whose spins are anticorrelated. They go off in opposite directions toward the detectors. These detectors are imagined to be aligned. Let the initial state, |Ψ(0)> be described as |(+, -), (I, I)>. Let the observer-detector system on the left experience a green sensation at t_{d1}. |Ψ(t_{d1})> is now ||(+, -), (G, I)>. At a later time, t_{d2}, the other observer sees red and we have |Ψ(t_{d2})> = |(+, -), (G, R)>. So far the worst problem seems to be that a signal is sent from the first observer to the second (perhaps faster than light) without any obvious mechanism. The thing to recognize is that |Ψ*(t)*> is not a function of the (x, y, z) spatial coordinates whatsoever. It is a vector in a Fock space. Such measurements as may pertain to its evolution do not take place in physical space at all. That they do take place with respect to the fiducial parameter *t* is a source of concern since this seems to give preference to a specific Lorentz frame. But let us also take note of the fact that, in all this measurement, no transfer of actual information has taken place (the No-Communication Theorem). Indeed, no *physical* thing has occurred that violates relativity. And it would be unthinkable that our second "observer" experienced green since that outcome would have to have |Ψ*(t)*> evolving into two, mutually incompatible futures for the particles' spins. (If the entire picture seems to speak of a certain non-locality in physics, so be it. There is ample and well-known experimental evidence to prove that such correlations do exist.)

A question will arise as to which of the red or green qualia *actually* occurred first. Wanting to ascribe some reality to *Q(t)* we believe this question to have an answer. But we have no way of determining it. If we are observers looking at the above experiment from a rocket ship we might, depending on our state of motion, see either the red or green color first. If we happen to be stationary in the preferred frame we will see the green color first and conclude that the state vector projected and was in its new condition throughout the hypersurfaces of greater *t.* Someone else, moving otherwise, may imagine that the state vector projected with the perception of the red color and that the (necessary) green color was only seen later. (Both of these observers just contribute to the larger *Q(t)*.) Now the second one is "wrong" but there can be no way of demonstrating this. We can propose all sorts of additional measurements that might have been made in one region of spacetime or another so as to verify "when" the state vector really projected. But these measurements were not made and quantum mechanics does not recognize contrafactual definiteness. If they had been made we'd have had a different problem and it would have to be analyzed differently. Since we assume that our underlying theory obeys microcausality, no physical measurement event can have physical consequences outside its future light-cone. And the light-cone structure of spacetime is invariant under Lorentz transformations. Therefore no physical experiment can be devised that would identify the preferred frame.

Conclusion.

By replacing wave functions with states in Fock space, |Ψ*(t)*>, we have created an interpretive picture that is in better agreement with the view adopted by physics since the inception of modern field theory in the 1950s. Unfortunately, much of the philosophical discussion related to quantum physics originated in the 1920s and 30s, before relativistically correct models had been developed.

A price to be paid for this "better agreement" is the acceptance of a unitary view of consciousness in which the idea of individual observers is ignored. Peculiar as this may seem, it does not bring with it any observable consequences. But it allows us to refer to an instantaneous consciousness-state (qualia-state) of the universe as *Q(t)*. We have to do this if we want to put such a state into relation with |Ψ*(t)*>.

The fiducial time parameter *t* is taken directly from the Interaction Picture and is, as discussed above, arbitrary. Although it appears to violate relativity by singling out a preferred Lorentz frame, this results in no observable, physical, events that contradict relativity. Instead, it frees us from trying to make sense of a multitude of separated consciousnesses each observing the world through the lens of its own proper time.

Acknowledgement.

The author is grateful to Professors David J. Chalmers, Lawrence S. Schulman, and Saverio Pascazio for useful and interesting discussions regarding this work.

References and Footnotes.

1) Schrödinger, E. (1935). *The Present Situation in Quantum Mechanics. Naturwissenschaften* **23** (49); 807.

2) Wigner, E. P. (1961), *Remarks on the Mind-Body Question.* In: I. J. Good,"The Scientist Speculates," London, Heinemann.

3) von Neumann, J. (1932). *The Mathematical Foundations of Quantum Mechanics.*

4) In free field theory it easy to express a universe consisting of only one particle with a definite momentum k as $a_{\boldsymbol{k}}^{\dagger}$|0> = |**k**>. If we attempt to restrict this "particle" to a single point, let us say x = 0 at t = 0, by representing |Ψ(0)> as $\sum_{\boldsymbol{k}} e^{i\,\boldsymbol{k}\,x}$ |**k**>, a so-called Newton-Wigner state, we have a chance of finding the particle infinitely far away at the slightest future time– seeming to violate relativity. If we try to formulate things in a relativististically invariant way by representing |Ψ(0)> as $\sum_{\boldsymbol{k}} \frac{1}{\sqrt{2\,\omega_{\boldsymbol{k}}}}\, e^{i\,\boldsymbol{k}\,x}$ |**k**> then we end up with a situation where the states of two particles, localized at different places are no longer orthogonal (see Teller, P. (1995), *An Interpretive Introduction to Quantum Field Theory*, Princeton University Press).

5) It may be objected that we should treat the various quantum states as continua rather than as discreet states. This is probably true. We have chosen the above method of presentation simply because it makes the underlying ideas a bit more simple and transparent in description. The reader wanting a more realistic representation can replace the summations by integrals. There is a caveat to be mentioned concerning Haag's theorem (Haag, R. *Mathematisk-fysiske Meddelelser*, **29**, 12 (1955)) which states that, in the continuum picture, we cannot unambiguously specify a vacuum state for our Fock space. This is an unresolved (and, largely, ignored) defect of quantum field theory which lies outside the scope of this article.

6) Chalmers (Chalmers, D. J. (1996), *The Conscious Mind*. Oxford University Press) defends the Everett idea by means of arguments that try to show that superpositions of states will automatically organize themselves into separate (we would say definite) conscious experiences. The present writer does not see how these arguments resolve the "relative probabilities" problem. This matter has been analyzed in a recent paper (Byrne, A and Hall, N., *Chalmers on Consciousness and Quantum Mechanics*, http://web.mit.edu/abyme/www/Conc&-QM.html). More recently Chalmers has introduced the notion of M-properties, which he associates clearly with consciousness. He also suggests that the wave function must always remain an eigenstate of M– an idea very

much like our requirement that the state vector be an eigenstate of C. See Chalmers, D. J., *Consciousness and its Place in Nature*, Sec. 9, in *Philosophy of Mind: Classical and Contemporary Readings* (Oxford, 2002). See also https://www.tubule.coma/watch?v=DIBT6E2GtjA.

7) While the simplest way to think of continuous measurement is just to imagine an infinite number of von Neumann measurements being performed infinitely quickly there are other, and more subtle, ways of looking at the problem. See Schulman, L. S., Physical Review A, 1509 (1998).

8) For an excellent recent review of the quantum Zeno effect see Pascazio, S. (2013). arXiv 1311.6645v1.pdf.