

William G. Unruh

Statement

and

Readings

Decoherence and Entanglement

William G. Unruh

Decoherence is the obverse side of entanglement, the peculiarly quantum nature of correlations between systems. By Bell's theorem, we know that entanglement has a number of non-intuitive properties, implying that quantum correlations can in some cases be stronger than classical, and in some cases violate transitivity ($A \Rightarrow B$, $B \Rightarrow C$, $C \Rightarrow D$ but *not* $A \Rightarrow D$, where \Rightarrow is implication). If we disregard these correlations, looking at only one of the systems on its own, the statistical properties of that system suffer decoherence. Interference terms which would in general be present for a quantum system with a variety of possible values for some attribute, are not present.

This decoherence has been argued to solve a variety of problems including the measurement problem in quantum mechanics.

However it is also true that the presence or absence of decoherence is far more subtle than usually described. A system, quantum correlated with an "environment" (another quantum system), can for certain measurements appear to be highly decohered, while still exhibiting interference between the apparently decohered values with suitable, long time, experiments.

False loss of coherence

W. G. Unruh

CIAR Cosmology Program, Dept. of Physics

University of B.C.

Vancouver, Canada V6T 1Z1

email:unruh@physics.ubc.ca

Abstract

The loss of coherence of a quantum system coupled to a heat bath as expressed by the reduced density matrix is shown to lead to the miss-characterization of some systems as being incoherent when they are not. The spin boson problem and the harmonic oscillator with massive scalar field heat baths are given as examples of reduced incoherent density matrices which nevertheless still represent perfectly coherent systems.

I. MASSIVE FIELD HEAT BATH AND A TWO LEVEL SYSTEM

How does an environment affect the quantum nature of a system? The standard technique is to look at the reduced density matrix, in which one has traced out the environment variables. If this changes from a pure state to a mixed state (entropy $Tr\rho \ln\rho$ not equal to zero) one argues that the system has lost quantum coherence, and quantum interference effects are suppressed . However this criterion is too strong. There are couplings to the environment which are such that this reduced density matrix has a high entropy, while the system alone retains virtually all of its original quantum coherence certain experiments.

The key idea is that the external environment can be different for different states of the system. There is a strong correlation between the system and the environment. As usual, such correlations lead to decoherence in the reduced density matrix. However, the environment in these cases is actually tied to the system, and is adiabatically dragged along by the system. Thus although the state of the environment is different for the two states, one can manipulate the system alone so as to cause these apparently incoherent states to interfere with each other. One simply causes a sufficiently slow change in the system so as to drag the environment variables into common states so the quantum interference of the system can again manifest itself.

An example is if one looks at an electron with its attached electromagnetic field. Consider the electron at two different positions. The static coulomb field of the two charges differ, and thus the states of the electromagnetic field differ with the electron in the two positions. These differences can be sufficient to cause the reduced electron wave function loose coherence for a state which is a coherent sum of states located at these two positions. However, if one causes the system to evolve so as to cause the electron in those two positions to come together (

eg, by having a force field such that the electron in both positions to be brought together at some central point for example), those two apparently incoherent states will interfere, demonstrating that the loss of coherence was not real.

Another example is light propagating through a slab of glass. If one simply looks at the electromagnetic field, and traces out over the states of the atoms in the glass, the light beams travelling through two separate regions of the glass will clearly decohere– the reduced density matrix for the electromagnetic field will lose coherence in position space– but those two beams of light will also clearly interfere when they exit the glass or even when they are within the glass.

The above is not to be taken as proof, but as a motivation for the further investigation of the problem. The primary example I will take will be of a spin $\frac{1}{2}$ particle (or other two level system). I will also examine a harmonic oscillator as the system of interest. In both cases, the heat bath will be a massive one dimensional scalar field. This heat bath is of the general Caldera Leggett type [1] (and in fact is entirely equivalent to that model in general). The mass of the scalar field will be taken to be larger than the inverse time scale of the dynamical behaviour of the system. This is not to be taken as an attempt to model some real heat bath, but to display the phenomenon in its clearest form. Realistic heat baths will in general also have low frequency excitations which will introduce other phenomena like damping and genuine loss of coherence into the problem.

II. SPIN- $\frac{1}{2}$ SYSTEM

Let us take as our first example that of a spin- $\frac{1}{2}$ system coupled to an external environment. We will take this external environment to be a one dimensional massive scalar field. The coupling to the spin system will be via purely the 3 component of the spin. I will use the velocity coupling which I have used elsewhere as a simple example of an environment (which for a massless field is completely equivalent to the Caldera Leggett model). The Lagrangian is

$$L = \int \frac{1}{2} ((\dot{\phi}(x))^2 - (\phi(x)')^2 + m^2 \phi(x)^2 + 2\epsilon \dot{\phi}(x) h(x) \sigma_3) dx \quad (1)$$

which gives the Hamiltonian

$$H = \int \frac{1}{2} ((\pi(x) - \epsilon h(x) \sigma_3)^2 + (\phi(x)')^2 + m^2 \phi(x)^2) dx \quad (2)$$

$h(x)$ is the interaction range function, and its Fourier transform is related to the spectral response function of Leggett and Caldera.

This system is easily solvable. I will look at this system in the following way. Start initially with the field in its free ($\epsilon = 0$) vacuum state, and the system is in the +1 eigenstate of σ_1 . I will start with the coupling ϵ initially zero and gradually increase it to some large value. I will look at the reduced density matrix for the system, and show that it reduces one which is almost the identity matrix (the maximally incoherent density matrix) for strong coupling. Now I let ϵ slowly drop to zero again. At the end of the procedure, the state of the system will again be found to be in the original eigenstate of σ_1 . The intermediate maximally incoherent density matrix would seem to imply that the system no longer has any quantum

coherence. However this lack of coherence is illusionary. Slowly decoupling the system from the environment should in the usual course simply maintain the incoherence of the system. Yet here, as if by magic, an almost completely incoherent density matrix magically becomes coherent when the system is decoupled from the environment.

In analyzing the system, I will look at the states of the field corresponding to the two possible σ_3 eigenstates of the system. These two states of the field are almost orthogonal for strong coupling. However they correspond to fields tightly bound to the spin system. As the coupling is reduced, the two states of the field adiabatically come closer and closer together until finally they coincide when ϵ is again zero. The two states of the environment are now the same, there is no correlation between the environment and the system, and the system regains its coherence.

The density matrix for the spin system can always be written as

$$\rho(t) = \frac{1}{2}(1 + \vec{\rho}(t) \cdot \vec{\sigma}) \quad (3)$$

where

$$\vec{\rho}(t) = Tr(\vec{\sigma}\rho(t)) \quad (4)$$

We have

$$\vec{\rho}(t) = Tr\left(\vec{\sigma}\mathcal{T}[e^{-i\int_0^t H dt}]\frac{1}{2}(1 + \vec{\rho}(0) \cdot \vec{\sigma})R_0\mathcal{T}[e^{-i\int H dt}]^\dagger\right) \quad (5)$$

where R_0 is the initial density matrix for the field (assumed to be the vacuum), and $\mathcal{T}[]$ is the time ordering operator. (Because ϵ and thus H is time dependent, the H at different times do not commute. this leads to requirement for the time ordering in the expression. As usual, the time ordered integral is the way of writing the time ordered product $\prod_n e^{-iH(t_n)dt} = e^{-iH(t)dt}e^{-iH(t-dt)dt} \dots e^{-iH(0)dt}$.)

Let us first calculate $\rho_3(t)$. We have

$$\rho_3(t) = Tr\left(\sigma_3\mathcal{T}[e^{-i\int_0^t H dt}]\frac{1}{2}(1 + \vec{\rho}(0) \cdot \vec{\sigma})R_0\mathcal{T}[e^{-i\int H dt}]^\dagger\right) \quad (6)$$

$$= Tr\left(\mathcal{T}[e^{-i\int_0^t H dt}]\sigma_3\frac{1}{2}(1 + \vec{\rho}(0) \cdot \vec{\sigma})R_0\mathcal{T}[e^{-i\int H dt}]^\dagger\right) \quad (7)$$

$$= Tr\left(\sigma_3\frac{1}{2}(1 + \vec{\rho}(0) \cdot \vec{\sigma})R_0\right) \quad (8)$$

$$= \rho_3(0) \quad (9)$$

because σ_3 commutes with $H(t)$ for all t . We now define

$$\sigma_+ = \frac{1}{2}(\sigma_1 + i\sigma_2) = |+\rangle\langle -|; \quad \sigma_- = \sigma_+^\dagger \quad (10)$$

Using $\sigma_+\sigma_3 = -\sigma_+$ and $\sigma_3\sigma_+ = \sigma_+$ we have

$$\begin{aligned} Tr & \left(\sigma_+\mathcal{T}[e^{-i\int_0^t H dt}]\frac{1}{2}(1 + \vec{\rho}(0) \cdot \vec{\sigma})R_0\mathcal{T}[e^{-i\int H dt}]^\dagger \right) \\ = & Tr_\phi\left(\mathcal{T}[e^{-i\int(H_0-\epsilon(t))\int\pi(x)h(x)dx}dt]^\dagger \right. \end{aligned} \quad (11)$$

$$\left. \mathcal{T}[e^{-i\int(H_0+\epsilon(t))\int\pi(x)h(x)dx}dt]\right) < -|\frac{1}{2}(1 + \vec{\rho}(0) \cdot \vec{\sigma})|+\rangle$$

$$= (\rho_1(0) + i\rho_2(0))J(t)$$

where H_0 is the Hamiltonian with $\epsilon = 0$, i.e., the free Hamiltonian for the massless scalar field and

$$J(t) = Tr_\phi \left(\mathcal{T} [e^{-i \int (H_0 - \epsilon(t)) \int \pi(x) h(x) dx dt}]^\dagger \mathcal{T} [e^{-i \int (H_0 + \epsilon(t)) \int \pi(x) h(x) dx dt}] R_0 \right) \quad (12)$$

Breaking up the time ordered product in the standard way into a large number of small time steps, using the fact that $e^{-i\epsilon(t) \int h(x) \phi(x) dx}$ is the displacement operator for the field momentum through a distance of $\epsilon(t)h(x)$, and commuting the free field Hamiltonian terms through, this can be written as

$$J(t) = Tr_\phi \left(e^{-i\epsilon(0)\Phi(0)} \prod_{n=1}^{t/dt} \left[e^{-i(\epsilon(t_n) - \epsilon(t_{n-1}))\Phi(t_n)} \right. \right. \\ \left. \left. e^{i\epsilon(t)\Phi(t)} e^{i\epsilon(t)\Phi(t)} \prod_{n=t/dt}^1 \left[e^{i\epsilon(t_n - \epsilon(t_{n-1}))\Phi(t_n)} \right] e^{i\epsilon(0)\Phi(0)} R_0 \right) \right) \quad (13)$$

where $t_n = ndt$ and dt is a very small value, $\Phi(t) = \int h(x) \phi(t, x) dx$ and $\phi_0(t, x)$ is the free field Heisenberg field operator. Using the Campbell-Baker-Hausdorff formula, realizing that the commutators of the Φ s are c-numbers, and noticing that these c-numbers cancel between the two products, we finally get

$$J(t) = Tr_\phi \left(e^{2i(\epsilon(t)\Phi(t) - \epsilon(0)\Phi(0) + \int_0^t \dot{\epsilon}(t')\Phi(t') dt')} R_0 \right) \quad (14)$$

from which we get

$$\ln(J(t)) = -2Tr_\phi \left(R_0 \left(\epsilon(t)\Phi(t) - \epsilon(0)\Phi(0) + \int_0^t \dot{\epsilon}(t')\Phi(t') dt' \right)^2 \right) \quad (15)$$

I will assume that $\epsilon(0) = 0$, and that $\dot{\epsilon}(t)$ is very small, and that it can be neglected. (The neglected terms are of the form

$$\int \int \dot{\epsilon}^2 \langle \Phi(t')\Phi(t'') \rangle dt' dt'' \approx \dot{\epsilon}^2 t \tau \langle \Phi(0)^2 \rangle$$

which for a massive scalar field has τ , the coherence time scale, $\approx 1/m$. Thus, as we let $\dot{\epsilon}$ go to zero these terms go to zero.)

We finally have

$$\ln(J(t)) = -2\epsilon(t)^2 \langle \Phi(t)^2 \rangle \\ = -2\epsilon(t)^2 \int |\hat{h}(k)|^2 \frac{1}{\sqrt{(k^2 + m^2)}} dk \quad (16)$$

Choosing $\hat{h}(k) = e^{-\Gamma|k|/2}$, we finally get

$$\ln(J(t)) = -4 \int_0^\infty \epsilon(t)^2 \frac{e^{-\Gamma|k|} dk}{\text{sqr}(k^2 + m^2)} \quad (17)$$

This goes roughly as $\ln(\Gamma m)$ for small Γm , (which I will assume is true). For Γ sufficiently small, this makes J very small, and the density matrix reduces to essentially diagonal form ($\rho_z(t) \approx \rho_y(t) \approx 0$, $\rho_z(t) = \rho_z(0)$.)

However it is clear that if $\epsilon(t)$ is now lowered slowly to zero, the decoherence factor J goes back to unity, since it depends only on $\epsilon(t)$. The density matrix now has exactly its initial form again. The loss of coherence at the intermediate times was illusionary. By decoupling the system from the environment after the coherence had been lost, the coherence is restore. this is in contrast with the naive expectation in which the loss of coherence comes about because of the correlations between the system and the environment. Decoupling the system from the environment should not in itself destroy that correlation, and should not reestablish the coherence.

The above approach, while giving the correct results, is not very transparent in explaining what is happening. Let us therefor take a different approach. Let us solve the Heisenberg equations of motion for the field $\phi(t, x)$. The equations are (after eliminating π)

$$\partial_t^2 \phi(t, x) - \partial_x^2 \phi(t, x) + m^2 \phi(t, x) = -\dot{\epsilon}(t) \sigma_3 h(x) \quad (18)$$

$$\pi(t, x) = \dot{\phi}(t, x) + \epsilon(t) h(x) \sigma_3 \quad (19)$$

If ϵ is slowly varying in time, we can solve this approximately by

$$\phi(t, x) = \phi_0(t, x) + \dot{\epsilon}(t) \int \frac{1}{2m} e^{-m|x-x'|} h(x') dx' \sigma_3 + \psi(t, x) \epsilon(0) \sigma_3 \quad (20)$$

$$\pi(t, x) = \dot{\phi}_0(t, x) + \epsilon(t) h(x) \sigma_3 + \dot{\psi}(t, x) \epsilon(0) \sigma_3 \quad (21)$$

where $\phi_0(t, x)$ and $\pi_0(t, x)$ are free field solution to the equations of motion in absence of the coupling, with the same initial conditions

$$\dot{\phi}_0(0, x) = \pi(0, x) \quad (22)$$

$$\phi_0(0, x) = \phi(0, x) \quad (23)$$

, while ψ is also a solution of the free field equations but with initial conditions

$$\psi(0, x) = 0 \quad (24)$$

$$\dot{\psi}(0, x) = -h(x). \quad (25)$$

If we examine this for the two possible eigenstates of σ_3 , we find the two solutions

$$\phi_{\pm}(t, x) \approx \phi_0(t, x) \pm (\dot{\epsilon}(t) \int \frac{1}{2m} e^{-m|x-x'|} h(x') dx' + \psi(t, x)) \quad (26)$$

$$\pi_{\pm}(t, x) \approx \dot{\phi}_0(t, x) + O(\dot{\epsilon}) \pm (\epsilon(t) h(x) + \epsilon(0) \dot{\psi}(t, x)) \quad (27)$$

These solutions neglect terms of higher derivatives in ϵ . The state of the field is the vacuum state of ϕ_0, π_0 . ϕ_{\pm} and π_{\pm} are equal to this initial field plus c number fields. Thus in terms of the ϕ_{\pm} and π_{\pm} , the state is a coherent state with non-trivial displacement from the vacuum. Writing the fields in terms of their creation and annihilation operators,

$$\phi_{\pm}(t, x) = \int A_{k\pm}(t) e^{ikx} + A_{k\pm}^{\dagger} e^{-ikx} \frac{dk}{\sqrt{2\pi\omega_k}} \quad (28)$$

$$\pi_{\pm}(t, x) = i \int A_{k\pm}(t) e^{ikx} - A_{k\pm}^{\dagger} e^{-ikx} \sqrt{\frac{k^2 + m^2}{2\pi}} dk \quad (29)$$

we find that we can write $A_{k\pm}$ in terms of the initial operators A_{k0} as

$$A_{k\pm}(t) \approx A_{k0}e^{-i\omega_k t} \pm \frac{1}{2}i(\epsilon(t) - \epsilon(0)e^{-i\omega_k t})(h(k)/\sqrt{\omega_k} + O(\dot{\epsilon}(t))) \quad (30)$$

where $\omega_k = \sqrt{k^2 + m^2}$. Again I will neglect the terms of order $\dot{\epsilon}$ in comparison with the ϵ terms. Since the state is the vacuum state with respect to the initial operators A_{k0} , it will be a coherent state with respect to the operators $A_{k\pm}$, the annihilation operators for the field at time t . We thus have two possible coherent states for the field, depending on whether the spin is in the upper or lower eigenstate of σ_3 . But these two coherent states will have a small overlap. If $A|\alpha\rangle = \alpha|\alpha\rangle$ then we have

$$|\alpha\rangle = e^{\alpha A^\dagger - |\alpha|^2/2}|0\rangle \quad (31)$$

Furthermore, if we have two coherent states $|\alpha\rangle$ and $|\alpha'\rangle$, then the overlap is given by

$$\langle \alpha|\alpha'\rangle = \langle 0|e^{\alpha^* A - |\alpha|^2/2}e^{\beta A^\dagger - |\beta|^2/2}|0\rangle = e^{\alpha^* \beta - (|\alpha|^2 + |\beta|^2)/2} \quad (32)$$

In our case, taking the two states $|\pm_\phi\rangle$, these correspond to coherent states with

$$\alpha = -\alpha' = \frac{1}{2}i(\epsilon(t) - \epsilon(0)e^{-i\omega_k t}) = \frac{1}{2}i\epsilon(t)h(k)/\sqrt{\omega_k} \quad (33)$$

Thus we have

$$\langle +_\phi, t | -_\phi, t \rangle = \prod_k e^{-\epsilon(t)^2 |h(k)|^2 / (k^2 + m^2)} = e^{-\epsilon(t)^2 \int \frac{|h(k)|^2}{\omega_k} dk} = J(t). \quad (34)$$

Let us assume that we began with the state of the spin as $\frac{1}{\sqrt{2}}(|+\rangle + |-\rangle)$. The state of the system at time t in the Schrodinger representation is $\frac{1}{\sqrt{2}}(|+\rangle + |+\phi(t)\rangle + |-\rangle + |-\phi(t)\rangle)$ and the reduced density matrix is

$$\rho = \frac{1}{2}(|+\rangle\langle +| + |-\rangle\langle -| + J^*(t)|+\rangle\langle -| + J(t)|-\rangle\langle +|). \quad (35)$$

The off diagonal terms of the density matrix are suppressed by the function $J(t)$. $J(t)$ however depends only on $\epsilon(t)$ and thus, as long as we keep $\dot{\epsilon}$ small, the loss of coherence represented by J can be reversed simply by decoupling the system from the environment slowly.

The apparent decoherence comes about precisely because the system in either the two eigenstates of σ_3 drives the field into two different coherent states. For large ϵ , these two states have small overlap. However, this distortion of the state of the field is tied to the system. π changes only locally, and the changes in the field caused by the system do not radiate away. As ϵ slowly changes, this bound state of the field also slowly changes in concert. However if one examines only the system, one sees a loss of coherence because the field states have only a small overlap with each other.

The behaviour is very different if the system or the interaction changes rapidly. In that case the decoherence can become real. As an example, consider the above case in which $\epsilon(t)$ suddenly is reduced to zero. In that case, the field is left as a free field, but a free field whose state (the coherent state) depends on the state of the system. In this case the field radiates away as real (not bound) excitations of the scalar field. The correlations with the system are carried away, and even if the coupling were again turned on, the loss of coherence would be permanent.

III. OSCILLATOR

For the harmonic oscillator coupled to a heat bath, the Hamiltonian can be taken as

$$H = \frac{1}{2} \int (\pi(x) - \epsilon(t)q(t)\tilde{h}(x))^2 + (\partial_x \phi(x))^2 + m^2 \phi(t, x)^2 dx + \frac{1}{2} (p^2 + \Omega^2 q^2) \quad (36)$$

Let us assume that m is much larger than Ω or that the inverse time rate of change of ϵ . The solution for the field is given by

$$\phi(t, x) \approx \phi_0(t, x) + \psi(t, x)\epsilon(0)q(0) - \frac{\dot{\epsilon}(t)q(t)}{\epsilon(t)q(t)} \int \frac{e^{-m|x-x'|}}{2m} h(x') dx' \quad (37)$$

$$\pi(t, x) \approx \dot{\phi}(t, x) + \dot{\psi}(t, x)\epsilon(0)q(0) - \frac{\ddot{\epsilon}(t)q(t)}{\epsilon(t)q(t)} \int \frac{e^{-m|x-x'|}}{2m} h(x') dx' + \epsilon(t)q(t)h(x) \quad (38)$$

where again ϕ_0 is the free field operator, ψ is a free field solution with $\psi(0) = 0$, $\dot{\psi}(0) = -h(x)$. Retaining terms only of the lowest order in ϵ

$$\psi(t, x) \approx \phi_0(t, x) \quad (39)$$

$$\pi(t, x) \approx \dot{\phi}(t, x) + \epsilon(t)q(t)h(x) \quad (40)$$

The equation of motion for q is

$$\dot{q}(t) = p(t) \quad (41)$$

$$\dot{p}(t) = -\Omega^2 q + \epsilon(t)\dot{\Phi}(t) \quad (42)$$

where $\Phi(t) = \int h(x)\phi(t, x)dx$. Substitution in the expression for ϕ , we get

$$\ddot{q}(t) + \Omega^2 q(t) \approx \epsilon(\dot{\Phi}_0(t)) + \epsilon(t)\frac{\ddot{\epsilon}(t)q(t)}{\epsilon(t)q(t)} \int \int h(x)h(x') \frac{e^{-m|x-x'|}}{2m} dx dx' \quad (43)$$

Neglecting the derivatives of ϵ (i.e., assuming that ϵ changes slowly even on the time scale of $1/\Omega$), this becomes

$$\left(1 + \epsilon(t)^2 \int \int h(x)h(x') \frac{e^{-m|x-x'|}}{2m} dx dx' \right) \ddot{q} + \Omega^2 q = \partial_t(\epsilon(t)\Phi(t)) \quad (44)$$

The interaction with the field thus renormalizes the mass of the oscillator to

$$M = \left(1 + \epsilon(t)^2 \int \int h(x)h(x') \right)$$

The solution for q is thus

$$q(t) \approx q(0) \cos\left(\int_0^t \tilde{\Omega}(t) dt\right) + \frac{1}{\Omega} \sin\left(\int_0^t \tilde{\Omega}(t) dt\right) p(0) + \frac{1}{\Omega} \int_0^t \sin\left(\int_{t'}^t \tilde{\Omega}(t) dt\right) \partial_t(\epsilon(t')\overline{\epsilon(t)\dot{\Phi}_0(t')}) dt' \quad (45)$$

where $\tilde{\Omega}(t) \approx \Omega/\sqrt{M(t)}$.

The important point is that the forcing term dependent on Φ_0 is a rapidly oscillating term of frequency at least m . Thus if we look for example at $\langle q^2 \rangle$, the deviation from the free evolution of the oscillator (with the renormalized mass) is of the order of $\int \sin(\tilde{\Omega}t - t') \sin(\omega(t - t'')) \langle \dot{\Phi}_0(t') \dot{\Phi}_0(t'') \rangle dt' dt''$. But $\langle \dot{\Phi}_0(t') \dot{\Phi}_0(t'') \rangle$ is a rapidly oscillating function of frequency at least m , while the rest of the integrand is a slowly varying function with frequency much less than m , Thus this integral will be very small (at least $\tilde{\Omega}/m$ but typically much smaller than this depending on the time dependence of ϵ). Thus the deviation of $q(t)$ from the free motion will in general be very very small, and I will neglect it.

Let us now look at the field. The field is put into a coherent state which depends on the value of q , because $\pi(t, x) \approx \dot{\phi}_0(t, x) + \epsilon(t)q(t)h(x)$ Thus

$$A_k(t) \approx a_{0k} e^{-i\omega_k t} + i \frac{1}{2} \hat{h}(k) \epsilon(t) q(t) / \omega_k \quad (46)$$

The overlap integral for these coherent states with various values of q is

$$\prod_k \langle i \frac{1}{2} \hat{h}(k) \epsilon(t) q / \omega_k | i \frac{1}{2} \hat{h}(k) \epsilon(t) q' / \omega_k \rangle = e^{-\frac{1}{8} \int |\hat{h}(k)|^2 dk (q - q')^2} \quad (47)$$

The density matrix for the Harmonic oscillator is thus

$$\rho(q, q') = \rho_0(t, q, q') e^{-\frac{1}{8} \int |\hat{h}(k)|^2 dk (q - q')^2} \quad (48)$$

where ρ_0 is the density matrix for a free harmonic oscillator (with the renormalized mass).

Ie, we see a strong loss of coherence of the off diagonal terms of the density matrix. However this loss of coherence is false. If we take the initial state for example with two packets widely separated in space, these two packets will loose their coherence. However, as time proceeds, the natural evolution of the Harmonic oscillator will bring those two packets together ($q - q'$ small across the wave packet). For the free evolution they would then interfere. They still do. The loss of coherence which was apparent when the two packets were widely separated disappears, and the two packets interfere just as if there were no coupling to the environment. The effect of the particular environment used is thus to renormalise the mass, and to make the density matrix appear to loose coherence.

IV. SPIN BOSON PROBLEM

Let us now complicate the spin problem in the first section by introducing into the system a free Hamiltonian for the spin as well as the coupling to the environment. Following the example of the spin boson problem, let me introduce a free Hamiltonian for the spin of the form $\frac{1}{2} \Omega \sigma_1$, whose effect is to rotate the σ_3 states (or to rotate the vector $\vec{\rho}$ in the 2 - 3 plane with frequency Ω).

The Hamiltonian now is

$$H = \frac{1}{2} \left(\int (\pi(t, x) - \epsilon(t)h(x)\sigma_3)^2 + (\partial_x \phi(x))^2 + m^2 \phi(t, x)^2 dx + \Omega \sigma_1 \right) \quad (49)$$

where again $\epsilon(t)$ is a slowly varying function of time. We will solve this in the manner of the second part the first section.

If we let Ω be zero, then the eigenstates of σ_z are eigenstates of the Hamiltonian. The field Hamiltonian (for constant ϵ) is given by

$$H_{\pm} = \frac{1}{2} \int (\pi - (\pm\epsilon(t)h(x)))^2 + (\partial_x\phi)^2 dx. \quad (50)$$

Defining $\tilde{\pi} = \pi - (\pm h(x))$, $\tilde{\pi}$ has the same commutation relations with π and ϕ as does π . Thus in terms of $\tilde{\pi}$ we just have the Hamiltonian for the free scalar field. The instantaneous minimum energy state is therefor the ground state energy for the free scalar field for both H_{\pm} . Thus the two states are degenerate in energy. In terms of the operators π and ϕ , these ground states are coherent states with respect to the vacuum state of the original uncoupled ($\epsilon = 0$) free field, with the displacement of each mode given by

$$a_k|\pm \rangle = \pm i\epsilon(t) \frac{h(k)}{\sqrt{\omega_k}}|\pm \rangle \quad (51)$$

or

$$|\pm \rangle = \prod_k |\pm \alpha_k \rangle |\pm \rangle_{\sigma_3} \quad (52)$$

where the $|\alpha_k \rangle$ are coherent states for the k^{th} modes with coherence parameter $\alpha_k = i\epsilon(t) \frac{h(k)}{\sqrt{\omega_k}}$, and the states $|\pm \rangle_{\sigma_3}$ are the two eigenstates of σ_3 . (In the following I will eliminate the \prod_k symbol.) The energy to the next excited state in each case is just m , the mass of the free field.

We now introduce the $\Omega\sigma_x$ as a perturbation parameter. The two lowest states (and in fact the excited states) are two fold degenerate. Using degenerate perturbation theory to find the new lowest energy eigenstates, we must calculate the overlap integral of the perturbation between the original degenerate states and must then diagonalise the resultant matrix to lowest order in Ω . The perturbation is $\frac{1}{2}\Omega\sigma_1$. All terms between the same states are zero, because of the $\langle \pm |_{\sigma_3} \sigma_1 | \pm \rangle_{\sigma_3} = 0$. Thus the only terms that survive for determining the lowest order correction to the lowest energy eigenvalues are

$$\frac{1}{2} \langle + | \Omega \sigma_1 | - \rangle = \frac{1}{2} \langle - | \Omega \sigma_1 | + \rangle^* \quad (53)$$

$$= \frac{1}{2} \Omega \prod_k \langle \alpha_k | - \alpha_k \rangle = \frac{1}{2} \Omega \prod_k e^{-2|\alpha_k|^2} \quad (54)$$

$$= \frac{1}{2} \Omega e^{-2 \int \epsilon(t)^2 |h(k)|^2 / \omega_k dk} = \frac{1}{2} \Omega J(t) \quad (55)$$

The eigenstates of energy thus have energy of $E(t)_{\pm} = E_0 \pm \frac{1}{2}\Omega J(t)$, and the eigenstates are $\sqrt{\frac{1}{2}}(|+ \rangle \pm |- \rangle)$ If epsilon varies slowly enough, the instantaneous energy eigenstates will be the actual adiabatic eigenstates at all times, and the time evolution of the system will just be in terms of these instantaneous energy eigenstates. Thus the system will evolve as

$$|\psi(t) \rangle = \sqrt{\frac{1}{2}} e^{-iE_0 t} \left((c_+ + c_-) e^{-i \int \frac{1}{2}\Omega J(t) dt} (|+ \rangle + |- \rangle) \right. \quad (56)$$

$$\left. + (c_- - c_+) e^{+i \int \frac{1}{2}\Omega J(t) dt} (|+ \rangle - |- \rangle) \right) \quad (57)$$

where the c_+ and c_- are the initial amplitudes for the $|+\rangle_{\sigma_3}$ and $|-\rangle_{\sigma_3}$ states. The reduced density matrix for the spin system in the σ_3 basis can now be written as

$$\vec{\rho}(t) = (J(t)\rho_{01}(t), J(t)\rho_{02}(t), \rho_{03}(t)) \quad (58)$$

where $\vec{\rho}_0(t)$ is the density matrix that one would obtain for a free spin half particle moving under the Hamiltonian $J(t)\Omega\sigma_1$.

$$\begin{aligned} \rho_{01}(t) &= \rho_1(0) \\ \rho_{02}(t) &= \rho_2(0) \cos(\Omega \int J(t')dt') + \rho_3(0) \sin(\Omega \int J(t')dt') \\ \rho_{03}(t) &= \rho_3(0) \cos(\Omega \int J(t')dt') - \rho_2(0) \sin(\Omega \int J(t')dt') \end{aligned} \quad (59)$$

Thus if $J(t)$ is very small (i.e., ϵ large), we have a renormalized frequency for the spin system, and the off diagonal terms (in the σ_3 representation) of the density matrix are strongly suppressed by a factor of $J(t)$. Thus if we begin in an eigenstate of σ_3 the density matrix will begin with the vector $\vec{\rho}$ as a unit vector pointing in the 3 direction. As time goes on the 3 component gradually decreases to zero, but the 2 component increases only to the small value of $J(t)$. The system looks almost like a completely incoherent state, with almost the maximal entropy that the spin system could have. However as we wait longer, the 3 component of the density vector reappears and grows back to its full unit value in the opposite direction, and the entropy drop to zero again. This cycle repeats itself endlessly with the entropy oscillating between its minimum and maximum value forever.

The decoherence of the density matrix (the small off diagonal terms) obviously represent a false loss of coherence. It represents a strong correlation between the system and the environment. However the environment is bound to the system, and essentially forms a part of the system itself, at least as long as the system moves slowly. However the reduced density matrix makes no distinction between whether or not the correlations between the system and the environment are in some sense bound to the system, or are correlations between the system and a freely propagating modes of the medium in which case the correlations can be extremely difficult to recover, and certainly cannot be recovered purely by manipulations of the system alone.

V. INSTANTANEOUS CHANGE

In the above I have assumed throughout that the system moves slowly with respect to the excitations of the heat bath. Let us now look at what happens in the spin system if we rapidly change the spin of the system. In particular I will assume that the system is as in section 1, a spin coupled only to the massive heat bath via the component σ_3 of the spin. Then at a time t_0 , I instantly rotate the spin through some angle θ about the 1 axis. In this case we will find that the environment cannot adjust rapidly enough, and at least a part of the loss of coherence becomes real, becomes unrecoverable purely through manipulations of the spin alone.

The Hamiltonian is

$$H = \frac{1}{2} \int \left((\pi(t, x) - \epsilon(t)h(x)\sigma_3)^2 + (\partial_x \phi(t, x))^2 + m^2 \phi(t, x) \right) dx + \theta/2\delta(t - t_0)\sigma_1 \quad (60)$$

Until the time t_0 σ_3 is a constant of the motion, and similarly afterward. Before the time t_0 , the energy eigenstates state of the system are as in the last section given by

$$|\pm, t \rangle = \{|+\rangle_{\sigma_3} |\alpha_k(t)\rangle \text{ or } \{|-\rangle_{\sigma_3} |-\alpha_k(t)\rangle\} \quad (61)$$

An arbitrary state for the spin–environment system is given by

$$|\psi\rangle = c_+|+\rangle + c_-|-\rangle \quad (62)$$

Now, at time t_0 , the rotation carries this to

$$\begin{aligned} |\phi(t_0)\rangle &= c_+(\cos(\theta/2)|+\rangle_{\sigma_3} + i\sin(\theta/2)|-\rangle_{\sigma_3}) |\alpha_k(t)\rangle \\ &\quad + c_-(\cos(\theta/2)|-\rangle_{\sigma_3} + i\sin(\theta/2)|+\rangle_{\sigma_3}) |-\alpha_k(t)\rangle \\ &= \cos(\theta/2)(c_+|+\rangle + c_-|-\rangle) \\ &\quad + i\sin(\theta/2)(c_+|-\rangle_{\sigma_3} |\alpha_k(t)\rangle - c_-|+\rangle_{\sigma_3} |-\alpha_k(t)\rangle) \end{aligned} \quad (63)$$

The first term is still a simple sum of eigenvectors of the Hamiltonian after the interaction. The second term, however, is not. We thus need to follow the evolution of the two states $|-\rangle_{\sigma_3} |\alpha_k(t_0)\rangle$ and $|+\rangle_{\sigma_3} |-\alpha_k(t_0)\rangle$. Since σ_3 is a constant of the motion after the interaction again, the evolution takes place completely in the field sector. Let us look at the first state first. (The evolution of the second can be derived easily from that for the first because of the symmetry of the problem.)

I will again work in the Heisenberg representation. The field obeys

$$\dot{\phi}_-(t, x) = \pi_-(t, x) + \epsilon(t)h(x) \quad (64)$$

$$\dot{\pi}_-(t, x) = \partial_x^2 \phi_-(t, x) - m^2 \phi_-(t, x) \quad (65)$$

with solution At the time t_0 the field is in the coherent state $|\alpha_k\rangle$. This can be represented by taking the field operator to be of the form

$$\phi_-(t_0, x) = \phi_0(t_0, x) \quad (66)$$

$$\pi_-(t_0, x) = \dot{\phi}_0(t_0, x) + \epsilon(t_0)h(x) \quad (67)$$

where the state $|\alpha_k\rangle$ is the vacuum state for the free field ϕ_0 . We can now solve the equations of motion for ϕ_- and obtain (again assuming that $\epsilon(t)$ is slowly varying)

$$\phi_-(t, x) = \phi_0(t, x) + 2\psi(t, x)\epsilon(t_0) \quad (68)$$

$$\pi_-(t, x) = \dot{\phi}_0(t, x) + 2\psi(t, x)\epsilon(t_0) - \epsilon(t)h(x) \quad (69)$$

where $\psi(t_0, x) = 0$ and $\dot{\psi}(t_0, x) = h(x)$. Thus again, the field is in a coherent state set by both $2\epsilon(t_0)\psi$ and $\epsilon(t)h(x)$. The field ψ propagates away from the interaction region determined by $h(x)$, and I will assume that I am interested in times t a long time after the time t_0 . At these times I will assume that $\int h(x)\psi(t, x)dx = 0$. (This overlap dies out as $1/\sqrt{mt}$. The calculations can be carried out for times nearer t_0 as well— the expressions are just messier and not particularly informative.)

Let me define the new coherent state as $|-\alpha_k(t) + \beta_k(t)\rangle$, where α_k is as before and

$$\beta_k(t) = 2\epsilon(t_0)\omega_k\tilde{\psi}(t, k) = 2i\epsilon(t_0)e^{i\omega_k t}\tilde{h}(k)/\omega_k \quad (70)$$

(The assumption regarding the overlap of $h(x)$ and $\psi(t)$ corresponds to the assumption that $\int \alpha_k^*(t)\beta_k(t)dk = 0$). Thus the state $|-\rangle_{\sigma_3} |\alpha_k\rangle$ evolves to the state $|-\rangle_{\sigma_3} |-\alpha_k + \beta_k(t)\rangle$. Similarly, the state $|+\rangle_{\sigma_3} |-\alpha_k\rangle$ evolves to $|+\rangle_{\sigma_3} |\alpha_k - \beta_k(t)\rangle$.

We now calculate the overlaps of the various states of interest.

$$\langle \alpha_k | \alpha_k \pm \beta_k \rangle = \langle -\alpha_k | -\alpha_k \pm \beta_k \rangle = e^{-\int |\beta_k|^2 dk} = J(t_0) \quad (71)$$

$$\langle -\alpha_k | \alpha_k \pm \beta_k \rangle = \langle \alpha_k | -\alpha_k \pm \beta_k \rangle = J(t)J(t_0) \quad (72)$$

$$\langle -\alpha_k + \beta_k | \alpha_k - \beta_k \rangle = \langle -\alpha_k - \beta_k | \alpha_k + \beta_k \rangle = J(t)J(t_0)^4 \quad (73)$$

The density matrix becomes

$$\rho_3 = \cos(\theta)\rho_{03} + \sin(\theta)J(t_0)\rho_{02} \quad (74)$$

$$\rho_1 = J(t) \left(\cos(\theta) + J^4(t_0) \sin(\theta) \right) \rho_{01} \quad (75)$$

$$\rho_2(t) = J(t) \left(-\sin(\theta)\rho_{03} + (\cos(\theta/2) - J^4(t_0)\sin(\theta))\rho_{02} \right) \quad (76)$$

where

$$\rho_{03} = \frac{1}{2}(|c_+|^2 - |c_-|^2) \quad (77)$$

$$\rho_{01} = \text{Re}(c_+c_-^*) \quad (78)$$

$$\rho_{02} = \text{Im}(c_+c_-^*) \quad (79)$$

If we now let $\epsilon(t)$ go slowly to zero again (to find the ‘real’ loss of coherence), we find that unless $\rho_{01} = \rho_{02} = 0$ the system has really lost coherence during the sudden transition. The maximum real loss of coherence occurs if the rotation is a spin flip ($\theta = \pi$) and ρ_{03} was zero. In that case the density vector dropped to $J(t_0)^4$ of its original value. If the density matrix was in an eigenstate of σ_3 on the other hand, the density matrix remained a coherent density matrix, but the environment was still excited by the spin.

We can use the models of a fast or a slow spin flip interaction to discuss the problem of the tunneling time. As Leggett et al argue [3], the spin system is a good model for the consideration of the behaviour of a particle in two wells, with a tunneling barrier between the two wells. One view of the transition from one well to the other is that the particle sits in one well for a long time. Then at some random time it suddenly jumps through the barrier to the other side. An alternative view would be to see the particle as if it were a fluid, with a narrow pipe connecting it to the other well- the fluid slowly sloshing between the two wells. The former is supported by the fact that if one periodically observes which of the two wells the particle is in, one sees it staying in one well for a long time, and then between two observations, suddenly finding it in the other well. This would, if one regarded it as a classical particle imply that the whole tunneling must have occurred between the two observations. It is as if the system were in an eigenstate and at some random time an interaction flipped the particle from one well to the other. However, this is not a good picture. The environment is continually observing the system. If it really moved rapidly from one to the other, the environment would see the rapid change, and would radiate. Instead, left on its own, the environment in this problem (with a mass much greater than the frequency of transition of the system) simply adjust continually to the changes in the system. The tunneling thus seems to take place continually and slowly.

VI. DISCUSSION

The high frequency modes of the environment lead to a loss of coherence (decay of the off-diagonal terms in the density matrix) of the system, but as long as the changes in the system are slow enough this decoherence is false— it does not prevent the quantum interference of the system. The reason is that the changes in the environment caused by these modes are essentially tied to the system, they are adiabatic changes to the environment which can easily be adiabatically reversed. Loosely one can say that coherence is lost by the transfer of information (coherence) from the system to the environment. However in order for this information to be truly lost, it must be carried away by the environment, separated from the system by some mechanism or another so that it cannot come back into the system. In the environment above, this occurs when the information travels off to infinity. Thus the loss of coherence as represented by the reduced density matrix is in some sense the maximum loss of coherence of the system. Rapid changes to the system, or rapid decoupling of the system from the environment, will make this a true decoherence. However, gradual changes in the system or in the coupling to the external world can cause the environment to adiabatically track the system and restore the coherence apparently lost.

This is of special importance to understanding the effects of the environmental cutoff in many environments [3]. For “ohmic” or “superohmic” environments (where \hbar does not fall off for large arguments), one has to introduce a cutoff into the calculation for the reduced density matrix. This cutoff has always been a bit mysterious, especially as the loss of coherence depends sensitively on the value of this cutoff. If one imagines the environment to include say the electromagnetic field, what is the right value for this cutoff? Choosing the Plank scale seems silly, but what is proper value? The arguments of this paper suggest that in fact the cutoff is unnecessary except in renormalising the dynamics of the system. The behaviour of the environment at frequencies much higher than the inverse time scale of the system leads to a false loss of coherence, a loss of coherence which does not affect the actual coherence (ability to interfere with itself) of the system. Thus the true coherence is independent of cutoff.

As far as the system itself is concerned, one should regard it as “dressed” with a polarization of the high frequency components of the environment. One should regard not the system itself as important for the quantum coherence, but a combination of variables of the system plus the environment. What is difficult is the dependence of which the degrees of freedom of the environment are simply dressing and which are degrees of freedom which can lead to loss of coherence depends crucially on the motion and the interactions of the system itself. They are history dependent, not simply state dependent. This make it very difficult to simply find some transformation which will express the system plus environment in terms of variables which are genuinely independent, in the sense that if the new variable loose coherence, then that loss is real.

These observations emphasis the importance of not making too rapid conclusions from the decoherence of the system. This is especially true in cosmology, where high frequency modes of the cosmological system are used to decohere low frequency quantum modes of the universe. Those high frequency modes are likely to behave adiabatically with respect to the low frequency behaviour of the universe. Thus although they will lead to a reduced density matrix for the low frequency modes which is apparently incoherent, that incoherence

is likely to be a false loss of coherence.

ACKNOWLEDGMENTS

I would like to thank the Canadian Institute for Advanced Research for their support of this research. This research was carried out under an NSERC grant 580441.

REFERENCES

- [1] A. O. Caldeira, A. J. Leggett *Physica* **121A** 587(1983), *Phys Rev* **A31** 1057 (1985) . See also the paper by W. Unruh, W. Zurek, *Phys Rev* **D40**1071(1989) where a field model for coherence instead of the oscillator model for calculating the density matrix of an oscillator coupled to a heat bath.
- [2] Many of the points made here have also been made by A. Leggett. See for example A. J. Leggett in *Applications of Statistical and Field Theory Methods to Condensed Matter*(Proc. 1989 Nato Summer School, Evora, Portugal), ed D. Baeriswyl, A.R. Bishop, and J. Carmelo. Plenum Press (1990) and **Macroscopic Realism: What is it, and What do we know about it from Experiment** in *Quantum Measurement: Beyond Paradox* ed R. A. Healey, and G. Hellman, U. Minnesota Press (Minneapolis, 1998)
- [3] See for example the detailed analysis of the density matrix of a spin 1/2 system in an oscillator heat bath, where the so called superohmic coupling to the heat bath leads to a rapid loss of coherence due to frequencies in the bath much higher than the frequency of the system under study. A.J. Leggett et al *Rev. Mod.Phys* **59** 1 (1987)
- [4]

Thoughts on Non-Locality and Quantum Mechanics

W. Unruh

Program in Cosmology and Gravity of CIAR

Dept Physics and Astronomy

University of B.C.

Vancouver, Canada V6T 1Z1

Abstract

The debate about the non-locality of quantum mechanics is old, but still lively. Numerous people use non-locality as a (bad) shorthand for quantum entanglement. But some have a long standing commitment to the validity of this characterisation. This paper examines two separate streams in this debate. The first is the arguments of Stapp, and especially his recent paper where he simplifies his contractually argument in the Hardy situation to argue for the non-locality of quantum mechanics. He has maintained his contention that an analysis of a Hardy type correlation between two spatially separated observers proves that quantum mechanics itself is non-local, without any additional assumption of realism or hidden variables.

In the second section I try to carefully examine the Bell argument, in the CHSH variant to see where the difference between the quantum and classical situations differ.

Asher Peres was one of the great physicists on the late 20th century, especially in his intense concern with the fundamental nature of quantum mechanics. His courage in devoting his life to an area many considered “philosophical” (ie non-physical) paved the way for the rest of us to reveal our interests and confusions about this area. I am not sure that he would agree with everything in this paper, but I offer it as a tribute to him.

I. STAPP

Stapp[1] has long maintained the position that quantum mechanics must be considered to be a non-local theory in its own right. He believes that the the assumption of ”hidden variables” or local realism in Bell’s argument is unnecessary, and that no local theory or any form could mimic quantum mechanics. It is not that any hidden variable theory, or locally realistic theory must be non-local in order to mimic quantum mechanics, as Bell showed. It is that quantum mechanics itself is non-local.

In much of the popular vocabulary of physicists, his war has been won. Many physicists, including many of those with an interest in the foundational issues of quantum theory, refer to quantum mechanics as non-local– using Bell’s arguments as a justification. By this they usually mean that quantum mechanical entanglement has non-classical features and when pushed, they will back off and agree that that non-locality is not really what Bell’s arguments mean. However, they stubbornly insist on using the terminology. (Names or references are purposely omitted to protect the guilty).

Stapp would however like to put this popular misnaming onto a firm footing. Despite a large amount of criticism, he still insists that his analysis of a Hardy type experiment shows that quantum mechanics itself is non-local. Unfortunately, in the face of this criticism, his claims have become more and more diluted.

He has recently published another paper in the American Journal of Physics [1] with new arguments on the non-locality of quantum mechanics. The end of the paper states ”This conclusion represents *some sort of failure* of the notion that no influence *of any kind* can act over a space-like interval”. “Some sort of failure” is so vague that almost anything can be subsumed under its mantle. Meanwhile “no influence of any kind” is so strong that many innocuous aspects of both classical and quantum physics can fall under this rubric.

Of course neither quantum mechanics nor classical mechanics has never argued that

no influence of any kind can act over a space-like interval. The existence of correlations between widely separated bodies could be taken to imply some sort of action over space-like intervals. A measurement operation, in which the measuring apparatus is only read when widely separated from the object could be taken to act over space-like intervals, since the value of the variable measured on the system in question has changed from unknown to known when the measuring apparatus is read. This is especially true in quantum mechanics where one cannot regard the system in question as having a value for the quantity of interest even in the absence of measurement. Ie, this sentence makes it unclear as to what Stapp is claiming. With a suitably diluted notion of non-locality, any theory could be said to be non-local.

The above paragraph may be taken as unfairly using his infelicitous language to erect and demolish a straw man. Let us therefor look a bit more closely at his argument.

He uses a Hardy-type experiment in his argument. The quantum Hardy-type experiment has been extensively described and generalised. It is a thought experiment in which two (spatially separated) physical systems are described by some state which is weakly entangled between the two systems. The weaker the entanglement, the more striking is the violation of the classical expectations, although the more rare the conditions under which it applies. We can consider the two systems to each be a two level system, and the state to be any state which is not a product state. For any such state, one can find a set of two dynamical variables for each sub-system, call them **L1** and **L2** for the one sub-system, and **R1** and **R2** for the other, with each variable having a pair of eigenvalues, denoted by + and -. These attributes have the following four properties in the given state.

In all experiments with the system in that given initial state and in which L1 and R1 are measured, and L1 is found to have value +, then R1 always has value +.

If R1 and L2 are measured, and R1 has value +, then L2 always has value +.

If L2 and R2 are measured, and L2 has value +, then R2 always has value +.

If L1 and R2 are measured, and L1 has value +, then R2 has value - with a probability which approaches unity as the state approaches a product state. This is clearly in conflict with the logical chain

$$L1 = + \Rightarrow R1 = + \Rightarrow L2 = + \Rightarrow R2 = + \tag{1}$$

which one would naively deduce from the chain of bipartite measurements.

To make the above more definite, consider the two systems to be two two level systems, with the usual Pauli matrices . Assume that in the σ_z basis for each the state of the system is

$$|\Psi \rangle = \sin(\phi)|++\rangle + \cos(\phi)|--\rangle \quad (2)$$

Ie, if ϕ is small, this state is almost a product state. Take L1 to be $\cos(2\mu)\sigma_{Lz} + \sin(2\mu)\sigma_{Lx}$ where $\tan(\mu) = \tan(\phi)^2$, and take R1 to be $\cos(2\phi)\sigma_{Rz} + \sin(2\phi)\sigma_{Rx}$, Choosing L2 to be $\frac{1}{\sqrt{2}}(\sigma_{Lz} + \sigma_{Lx})$ and R2 to be $\sin(2\phi)\sigma_{Rz} + \cos(2\phi)\sigma_{Rx}$, where These operators obey the above conditions. This choice comes very close to maximizing the probability, $\cos(2\phi)^2$, that if L1 and R2 are measured and L1 is +, then R2 is -. For $\phi \ll 1$, this probability becomes very close to unity. Note that attribute R2 is almost exactly the negative of R1, and its + eigenvector is almost exactly the - eigenvector of R1. Ie, L2 having value + implies R1 has value + while L1 having value + implies that R2 has value - with high probability. However, for any value of ϕ except 0 (no entanglement but the probability of L1=+ is zero) or $\pi/2$ (maximum entanglement) these operators obey the conditions of this generalised Hardy system.

For any classical system, the first three properties would imply that if L1 has value + then R2 must have value +. The fourth property contradicts this. Stapp's argument is that this chain of reasoning also applies in quantum mechanics. The argument is subtle and uses the language of counterfactuals.

Counterfactual arguments are tricky (see for example Shimony's criticism of this paper by Stapp which is similar to my criticism)[2], and are invariably heavily theory laden. They are not statements about the world, but rather about one's theory of the world. This is especially clear in the example which could be called the argument of Peres's mother [3].

When young his mother asked herself the counterfactual question of whose child, her mother's or her father's, she would have been if her mother and father had each married different people. While she ultimately decided the question was meaningless, it is clear that it would not have been meaningless, and would furthermore have had a definite answer, had her theory of human essence rested upon matrilinear reincarnation. Furthermore, had she asked instead whether her father's or mother's child would have had her blue eyes, we would have had no difficulty giving an answer based on our theories of genetic inheritance. Ie, the meaningfulness and answer to a counterfactual question depends crucially on the theoretical

context in which it is embedded.

The central point of Stapp's argument rests on a proposition which he calls SR. This proposition is (translated to the notation I am using)

If $R1$ is measured and gives outcome $+$, then if, instead, $R2$ had been measured, the outcome would have been $+$.

By "instead" he does not mean in some other experiment, but means a counterfactual replacement of the measurement of $R1$ by a measurement of $R2$ in the same experiment in which $R1$ was measured. As a counterfactual statement, it can of course never actually be tested by experiment in the real world. As with all counterfactual statements, it is a statement made within the context of a theoretical framework. As such one must be careful to ensure that the replacement makes sense within the context of the theory. Within quantum theory this becomes especially ticklish, since the attribute $R2$ does not commute with $R1$, and quantum mechanics thus rules out any interpretation of "instead of" which makes it synonymous with "as well as". I.e., the measurement of $R1$ inherently destroys the probability structure of the outcomes for $R2$ and interferes with any measurement of $R2$.

The first question to ask is whether, within the theoretical context of quantum mechanics, the statement makes any sense. The statement assumes a number of other postulates—namely that the state of the system before any measurement is the Hardy state, a state which explicitly refers to both L and R . One can certainly argue that in fact, as in the case of Peres' mother, this statement does not make any sense within the context of quantum mechanics. Because attributes do not have values in the absence of measurement, because the values found in a measurement occur without sufficient cause, are generated out of thin air by the measurement itself, the question of what quantum mechanics would have to say about the counterfactual replacement of $R1$ with $R2$ is "nothing" in the absence of any other conditions. But let us push the analysis a little bit further.

Within quantum mechanics, the validity of this counterfactual replacement hinges on whether or not $L2$ was actually measured. If it was measured, then, because of the prior condition that the state is the Hardy state and the assumption about the measured value of $R1$, it is a fact that both $R1$ and $L2$ have values $+$. The validity of the counterfactual replacement of $R1$ with $R2$ giving the value $+$ then rests on the reality of the measurement of $L2$. If, on the other hand, $L2$ was not actually measured, then the validity of the argument rests on a double counterfactual—namely that if instead of not being measured at all, $L2$

had been measured, it would have had value +, and thus on the second counterfactual substitution of R2 for R1, R2 would have value +. There is no reason to believe that quantum physics makes any sense at all out of such a double counterfactual substitution.

Stapp argues that if, one had, in the past (but space-like separated from) of the measurement of either R1 or R2 referred to in this statement, L2 had been measured, then this statement is true. The measurement result of + for the measurement of R1 would ensure through the correlations inherent in the state that the outcome of the L2 measurement in the past must have been + as well. But, since that result is surely independent of whether or not R1 or R2 were measured in the future, it would still have had outcome + if the experimenter in R had decided to measure R2 instead, and thus, because of the correlations in the state, R2 would then have had value + as well. Thus, given only the knowledge that L2 was measured, the statement SR is true. Of course it is true only because of the existence of the measurement of L2. Without the existence of that measurement, the statement SR is nonsense (ie, untrue).

However Stapp here uses the free will of the experimentalist and his notion of locality to argue that, as a statement about region R, SR must surely be independent of what experiment was carried out in region L, since, it being space-like separated from R, one can consider the measurement in region L to occur after that in R. Thus SR should continue to be true if L2 were replaced with L1, in which case however, the inference of SR does not follow (and is in fact negated with high probability if the outcome of the L1 measurement is +).

However this notion of locality is strange. SR, is not a statement about region R, rather it is a statement about two different counterfactual worlds, the one in which R1 was measured and the other where R2 was measured. There seems to me to be no argument from locality or anything else which could demand that such a counterfactual relationship should be independent of the actions in region L. The existence of the measurement of L2 plays a crucial role in the establishment of the truth of SR, and there is no reason why that truth should be independent of that measurement. IF SR referred to some actual state of affairs in a single world (established even by counterfactual reasoning) then such a locality requirement might be reasonable. But as I have stated, the assumption that SR says something about the single real world is a form of realism.

II. VON NEUMAN MEASUREMENT

This Hardy type system can also be used to point out some features and limitations of the von Neuman description of measurement. In establishing the logical consistency of quantum mechanics and in particular of the measurement hypothesis, von Neuman introduces a measurement hypothesis. A measurement on a system could be regarded as a primate operation on that system. Alternatively, one could introduce a measuring apparatus which was itself a quantum system, and whose interactions with the system were fully governed by the laws of quantum mechanics. The measurement process on the original system was now regarded as the establishment of correlations between some dynamic attribute of the apparatus with the “measured” attribute of the system. The measurement, in the primate sense, on this pointer attribute of the apparatus, could be used to infer, by means of the correlations between the two systems, a value for the attribute of the system. He argued that regarding a measurement on the system either as a primitive, or as being inferred from a measurement on an apparatus, are consistent, and equivalent.

However, this model demonstrates limitations of this equivalence in some situations. Because of the correlations inherent in this Hardy state, one can regard the either the system on the left or on the right as the system of interest and the other to be a measuring apparatus. The correlations created by the interaction which placed the system into the partially entangled state are of the kind discussed by von Neuman. In particular, a measurement, in the primitive sense, of R1 giving value + is perfectly correlated with L2 having value +. Ie, a primitive measurement of R1 giving value + is a measurement in the von-Neuman sense of L2 giving value +. (The primitive measurement of R1 giving any value is not equivalent to a generic von Neuman measurement of L2, since the correlation is not valid for R1 having value -.) Now, the primitive measurement of L2, giving value + can also be regarded as a measurement in the von Neuman sense of R2 giving value +. But von Neuman also insisted that there is no difference between a von Neuman and a primitive measurement as far as the system is concerned. Thus, we can take the primate measurement of R1 with value + to be equivalent to the measurement of L2 referred to the above, which was also a measurement of R2 with value +. Ie, by the double application of von Neuman’s argument the (primitive) measurement of R1 giving value + can apparently also be regarded as a von Neuman measurement of R2 giving value +.

One might make two objections. The first is that the measurement of R1 destroys the probability distribution of R2, leaving R2 with an entirely different probability distribution. What is R1 then measuring? However, physicists have long engaged in measurements which destroy the system being measured even more completely. When a photon impinges on a photographic plate or a CCD, the fact that the photon is completely destroyed in the process does not change physicists' notion that the photographic plate has measured the position of the photon. Yes, it is destroyed, but just before the destruction the photon had that position.

The second possible objection is that the measurement is very indirect. After all we are operating through the intermediary of L2. Without L2, the measurement of R1 would not allow anything to be inferred about the value of R2. But again, this possibility was already envisioned by von Neuman, who discussed a whole chain of measuring apparatuses. One could "measure" the pointer of the apparatus, either as a primitive operation, or by coupling it again to another super-apparatus, whose pointer we correlated with the pointer of the first apparatus. This chain could be as long as one wished, as long as one had established the chain of correlations between the various pointers and the original attribute in the system to be measured. Ie, there is nothing in the von Neuman equivalence which limits our right to regard R1 having value + as being a measurement of R2.

Note of course that this is a system to which we cannot apply the arguments of "Wigner's Friend". Ie, a separate attempt to measure R2 either by coupling it to some other apparatus, or via a primitive measurement will not give the same result as the result inferred from the measurement of R1. But nowhere in the naive von Neuman analysis is there any requirement that the "Wigner's Friend" argument apply.

But, of course, if one does allow the measurement of R1 with value + to be a valid measurement of R2, the plot grows even more convoluted. One could regard the measurement of L1, giving value +, to be a measurement of R1 (with value +) which is a measurement of L2 (with value +) which is a measurement of R2 (with value +). Again the fact that L2 is destroyed in the primitive measurement of L1 would seem to be irrelevant.

But this leads to a contradiction. For exactly the same correlated state between the measuring apparatus L1 and the system attribute R2 allows one to assume that if L1 has value +, R2 almost certainly has value -. Ie, the equivalence between primitive measurements and von Neuman causal chain measurements fails spectacularly. At the same time it is not

clear exactly where it fails.

Ie, it would seem that one needs to restrict the von Neuman measurement chain such that at each step one can apply a "Wigner's Friend" argument to obtain the same outcome for the measurement as the one inferred from the von Neuman chain. Or equivalently one must restrict the measurement chain so that at no point can a measuring apparatus be regarded as measuring itself.

Bell's theorem and Quantum Systems

Ultimately all arguments for the non-locality of quantum mechanics can be traced back to Bell's arguments [4] in establishing his theorem for "Locally realistic" systems. It seems to be because of the powerful fascination of realism that the violation Bell's inequality for quantum mechanics and for the real physical world is interpreted as a violation of locality. It is worth looking in more detail at Bell's argument and at the differences between quantum and classical systems for each step in the argument. In the following I will use the name Bell to refer to the Clauser, Horn, Shimony and Holt [5] version of the argument.(See also Jarret[6] for a discussion of the experiment).

The setup is that we have two attributes L1 and L2 on the left and R1 and R2 on the right. (these are not the same as the attributes above in the Stapp argument.) Each takes values of ± 1 . In the quantum system we will take L1 and L2 to be maximally non-commuting attributes, and can take them as σ_1 and σ_2 , the two Pauli spin matrices, and R1 and R2 are also the two sigma matrices for another two level system. The system is set up in a correlated state, and a sequence of measurements are made on the L and R systems. In particular L1 or L2 is measured on the left and R1 or R2 on the right. In each measurement only one of the pair are measured. After the measurements have all been made, a set of correlation functions is measured. Namely

$$[L1 R1] = \sum_{11} L1_i R1_i / \sum_{11} 1 \tag{3}$$

$$[L1 R2] = \sum_{12} L1_i R2_i / \sum_{12} 1 \tag{4}$$

$$[L2 R1] = \sum_{21} L2_i R1_i / \sum_{21} 1 \tag{5}$$

$$[L2 R2] = \sum_{22} L2_i R2_i / \sum_{22} 1 \tag{6}$$

where in each case terms like $L1_i$ refer to the value obtained for L1 in the i^{th} trial and the sum over i is over all instances in which the corresponding attributes were measured. (ie,

\sum_{12} is the sum over all instances in which $L1$ and $R2$ were measured.)

Now, of course each of these correlation function is taken over disjoint sets. It is never the case that both $R1$ and $R2$ were measured in the same instance, and similarly no case where $L1$ and $L2$ were measured in the same instance.

The critical procedure in Bell's proof is to argue using local realism, that even though they were not measured in any instance, all of the operators $L1$, $L2$, $R1$, $R2$ actually have values in each of the instances of measurement. Furthermore, he uses locality to argue that if this is true, then the measured correlation function $\langle La Rb \rangle$, with a, b both taking values 1 and 2 is a good estimator of the (counterfactual) correlator

$$\langle La Rb \rangle \approx \sum_j La_j Rb_j / \sum_j 1 \quad (7)$$

where this time the sum is taken over all instances in which any measurement was taken. If we assume that the sets are or roughly equal size, in 1/4 of the values of j , these correspond to real values for La and Rb and in 3/4 of the cases at least one of them is the value assumed to exist by counterfactual realism.

Furthermore, locality is used to argue that we can write

$$[L1 R1] + [L1 R2] + [L2 R1] - [L2 R2] = \langle L1(R1 + R2) \rangle + \langle L2(R1 - R2) \rangle = \langle L1(R1 + R2) + L2(R1 - R2) \rangle$$

This is the critical relation. Ie, the whole use of locality and local realism is to argue that the sum of the correlators is equal to the correlation of the sum of the operators.

What is of course interesting about quantum mechanics is this property comes free. If we define $\mathbf{L}a$ and $\mathbf{R}a$ as the quantum operators and the expectation values as the quantum expectation values, then quantum mechanics gives us, for free, that

$$\begin{aligned} [L1 R1] + [L1 R2] + [L2 R1] - [L2 R2] &= \langle \psi | \mathbf{L}1(\mathbf{R}1 + \mathbf{R}2) | \psi \rangle + \langle \psi | \mathbf{L}2(\mathbf{R}1 - \mathbf{R}2) | \psi \rangle \\ &= \langle \psi | \mathbf{L}1(\mathbf{R}1 + \mathbf{R}2) + \mathbf{L}2(\mathbf{R}1 - \mathbf{R}2) | \psi \rangle \end{aligned} \quad (10)$$

Since the use of locality in the classical case is solely to demonstrate the truth of something which quantum mechanics apparently gives us for free, the question now arises as to where the difference between the quantum and classical resides.

The first instance is when we examine the meaning of these expectation values. In the classical case, for example $\langle L1(R1 + R2) \rangle$ is taken to mean something different from $\langle \psi | \mathbf{L}1(\mathbf{R}1 + \mathbf{R}2) | \psi \rangle$. In the classical case, Bell took $R1 + R2$ in each instance to be the

sum of the values of $R1$ and $R2$ for each particular instance. Since by assumption Ra took values of ± 1 , $R1+R2$ has values of ± 2 or 0 . However, a critical feature of quantum mechanics is that $\mathbf{R1}+\mathbf{R2}$ is an operator, and attribute in its own right, and will take values of $\pm\sqrt{2}$. Furthermore, in all situations in which the operators $R1$ and $R2$ are measured separately, or their sum is measured,

$$|\psi\rangle\mathbf{M}\mathbf{R1}\langle\psi| + |\psi\rangle\mathbf{M}\mathbf{R2}\langle\psi| = |\psi\rangle\mathbf{M}(\mathbf{R1} + \mathbf{R2})\langle\psi| \quad (11)$$

where M is any operator which commutes with \mathbf{Ra} . Ie, measured separately or measured as a sum, these two correlators are identical.

If the classical system is to mimic the quantum system, this must also be true of the classical system. In general since $R1 + R2$ has different values than $\mathbf{R1} + \mathbf{R2}$ (namely $\pm 2, 0$ instead of $\pm\sqrt{2}$) this mimicking is difficult for the classical system to maintain.

Secondly, Bell makes use of another feature. Both of the attributes $R1 + R2$ and $R1 - R2$ are assumed to have possible values of $\pm 2, 0$. Furthermore they are perfectly anti-correlated in that one and only of the two ever has the value 0 in any one instance of the experiment. Thus in each element of the sum, either $R1 + R2$ or $R1 - R2$ is zero. Since $L1$ and $L2$ have values of ± 1 we immediately get Bell's theorem, namely that

$$-2 \leq [L1 R1] + [L1 R2] + [L2 R1] - [L2 R2] \leq 2 \quad (12)$$

The quantum violation comes about by noting that we can find a state, $|\psi\rangle$ such that $L1$ and $(R1+R2)$ are maximally correlated– ie every-time $L1$ has value $+1$, $R1+R2$ has value $+\sqrt{2}$ and every time $L1$ has value -1 , $R1+R2$ has value $+\sqrt{2}$. That same state $|\psi\rangle$ can be chosen so that $L2$ and $R1-R2$ are also maximally correlated. This immediately leads to the quantum correlation

$$[\mathbf{L1}\mathbf{R1}] + [\mathbf{L1}\mathbf{R2}] + [\mathbf{L2}\mathbf{R1}] - [\mathbf{L2}\mathbf{R2}] = 2\sqrt{2} \quad (13)$$

Where can one locate the difference between the quantum and classical case. A key location is the assumption that the values of $R1 + R2$ take values of $\pm 2, 0$ rather than the $\pm\sqrt{2}$ of the quantum system. Ie, in quantum mechanics the sum of the values is not the same as the values of the sum. This is clearly crucial in Bell's argument.

WE can express this in a slightly different way. If we look at the correlation $\langle (R1 + R2)(R1 - R2) \rangle$ for the classical system, it is crucial to Bell's argument that this is zero.

Quantum mechanically of course, this expression would not necessarily be zero and in fact in the quantum state under consideration it is non-zero.

The second point, related to the first, is that $R1+R2$ is anti-correlated with $R1-R2$ in that a non-zero value of one is perfectly correlated with a zero value of the other. Clearly if the values are not 0 and ± 2 this correlation between the two makes little sense. Finally, the perfect correlations between $\mathbf{R1+R2}$ and $\mathbf{L1}$ at the same time as a perfect correlation between $\mathbf{R1-R2}$ and $\mathbf{L2}$ obtains in the quantum system is also critical to the possibility of its violating the classical limits. Can a classical system be set up so as to have this same correlation? The answer is of course yes. We take $R1 \pm R2$ to have values $\pm\sqrt{2}$ as for the quantum system. Set up the four states $\{+1, +1, +\sqrt{2}, +\sqrt{2}\}$, $\{+1, -1, +\sqrt{2}, -\sqrt{2}\}$, $\{-1, +1, -\sqrt{2}, +\sqrt{2}\}$, and $\{-1, -1, -\sqrt{2}, -\sqrt{2}\}$ where these four values are the classical values of $L1, L2, (R1+R2)$, and $(R1-R2)$ respectively. The classical state is now defined by taking each of these states with probability of $1/4$. Thus we see that the critical difference between quantum and classical system is in the fact that the sum of values is not the same as the values of the sum. Classically, the values of $R1+R2$ are just the values of $R1$ added to those of $R2$, namely $\pm 2, 0$ while quantum mechanically they are just $\pm\sqrt{2}$.

We note that the locality has played a weak role. It has acted to allow us to argue that for the classical system, the correlations behave in just the way we would expect the quantum system to behave—namely that the sum of the correlators is just the theoretical correlation of the sum.

Acknowledgments

I would like to thank A. Shimony for showing me a copy of his paper before publication, and drawing my attention to Stapp's latest paper. I would also thank the Canadian Institute for Advanced Research, and the NSERC for their support during this research.

-
- [1] H.P. Stapp, Am. J. Phys **71**, 30 (2004).
[2] A. Shimony, Found. Phys (to Appear as a special edition for A. Peres on his 70th birthday) (2005) See also the extensive bibliography in this paper.
[3] W. Unruh, Phys Rev **A 59**, 126 (1999).

- [4] See the reprints of the original papers in J.S.Bell **Speakable and Unspeakable in Quantum Mechanics** Cambridge U Press (Cambridge, 1987)
- [5] J. F. Clauser, M.A. Horne, A. Shimony and R. A. Holt, Phys. Rev. Lett. 23, 880-884 (1969),
- [6] Jon Jarrett “Bell’s Theorem: A Guide to the Implications” in J. Cushing, E. McMullin, eds **Philosophical Consequences of Quantum Theory** University of Notre Dame Press (1989)

