Quantum Mechanics versus Macroscopic Realism: Is the Flux There when Nobody Looks?

Leggett and Garg

March 4, 1985

Abstract

It is shown that, in the context of an idealized "macroscopic quantum coherence" experiment, the predictions of quantum mechanics are incompatible with the conjunction of two general assumptions which are designated "macroscopic realism" and "noninvasive measurability at the macroscopic level". The conditions under which quantum mechanics can be tested against these assumptions in a realistic experiment are discussed.

Despite sixty years of schooling in quantum mechanics, most[1] physicists have a very non-quantum-mechanical notion of reality at the macroscopic level, which implicitly makes two assumptions. (A1)Macroscopic realism: A macroscopic system with two or more macroscopically distinct[2,3] states available to it will at all times *be* in one or the other of these states. (A2) Noninvasive measurability at the macroscopic level: It is possible, in principle, to determine the state of the system with arbitrarily small perturbation on its subsequent dynamics. A direct extrapolation of quantum mechanics to the macroscopic level denies this. The aim of this Letter is (1) to point out that under certain conditions the experimental predictions of the conjunction of (A1) and (A2) are incompatible with those of quantum mechanics extrapolated to the macroscopic level, and (2) to investigate how far these conditions may be met in a realistic experiment.

To this end, let us consider the (as yet unobserved) phenomenon of "macroscopic quantum coherence" (MQC) in an rf SQUID.[4] We take the potential $V(q)$ for the trapped magnetic flux q to be reflection symmetric (see Fig. 1) with minima at $\pm q_0$ far enough apart that states in which q is close to $+q_0$ and $-q_0$ can be regarded as macroscopically distinct. For an isolated SQUID, quantum mechanics predicts that if the flux is initially in one well, it will oscillate back and forth with some frequency Δ_0. A more realistic quantum mechanical calculation[5] which includes the irremovable environmental effects shows that for low enough temperature and weak enough coupling to the environment, the oscillations are not entirely destroyed, but merely underdamped. Since it is under these conditions that our argument is most pertinent, we shall assume that the experimental constraints on achieving them, while stringent[6], can, in fact, be met.

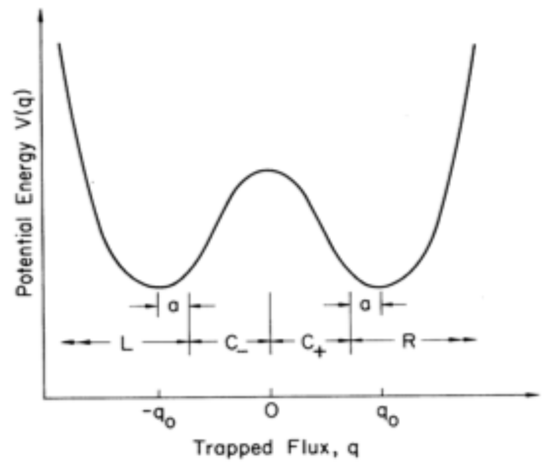

Figure 1: The potential $V(q)$ for the trapped flux q. The various notations are explained in the text.

Let us divide the possible values of q into four regions L, C_-, C_+, and R as shown in Fig. 1, where $x_0 << a << q_0$, x_0 being the zero-point width that a wave packet would have in either well if the other were absent. We define a quantity Q, which equals $+1(-1)$ if the system is observed to be in region $R(L)$. If we temporarily ignore the minuscule probability of finding the system in $C_\pm$, quantum mechanics predicts (and we assume that experiment will find) that any observation of Q will find only the values ± 1.

It immediately follows from (A1) that for an ensemble of systems prepared in some way at time t_0,[7] we can define (i) joint probability densities $\rho(Q_1, Q_2)$, $\rho(Q_1, Q_2, Q_3)$, etc. for Q to have the values Q_l at times t_i (we take $t_0 < t_1 < t_2 ....$), (ii) correlation functions $K_{ij} \equiv \langle Q_i Q_j \rangle$. The probability densities must be consistent with one another, which implies, e.g.,

$$\sum_{Q_2=\pm 1} \rho(Q_1, Q_2, Q_3) = \rho(Q_1, Q_3) \tag{1}$$

From this, we can derive inequalities similar to those of Bell[8] or of Clauser et al.[9] for the Einstein-Podolsky-Rosen (EPR) experiment[10] with the times t_i playing the role of the polarizer settings. For example, we have

$$1 + K_{12} + K_{23} + K_{13} \geq 0 \tag{2a}$$

$$|K_{12} + K_{23} + K_{14} - K_{24}| \leq 2 \tag{2b}$$

If we assume that (A2) can be realized in an actual experiment (we shall discuss this below), then these correlations and probabilities can be measured, and we can test whether (1) and (2) hold.

We can also test (1) and (2) against the predictions of quantum mechanics. For definiteness, we consider the case of "Ohmic" dissipation, which has been studied in detail by Chakravarty and Leggett.[5] The behavior of the system can be parametrized by Δ_r, a renormalized tunneling frequency, by ω_c the highest frequency scale at which the environment can respond, and by α, a dimensionless dissipation coefficient. Typically $\Delta_r << \Delta_0 << \omega_c$. If, as in Chakravarty and Leggett, we ignore the so-called "interblip effects" (which is a good approximation both for very low T and α, when the flux executes underdamped oscillations, and for high values of T and α, when we have overdamped relaxation), then although we cannot rigorously prove, we can very plausibly argue that for $t_i, |t_j - t_i| >> \omega_c^{-1}$, K_{ij} is essentially independent of the choice of the initial ensemble and equals $P(t_j - t_i)$ as defined there.[5] One can further argue that

$$\rho(Q_1, Q_2, Q_3) \simeq \rho(Q_1, Q_2)\rho(Q_2, Q_3) \tag{3}$$

It is now clear from experience with Bell-type inequalities that if $P(t)$ is not too heavily damped, then quantum mechanics will violate conditions (1) and (2). Consider, for example, the expression (24) of Chakravarty and Leggett for $P(t)$ at $T = 0$, and set $\Delta P(t) = 0$.[5] Since the "incoherent" part, $P_{\text{inc}}(t)$, is always negative, we will overestimate the left-hand side of (2a) if we neglect this altogether. Any value of α for which a violation of (2a) is thus obtained will be less than the critical value of α beyond which (2a) is always satisfied. The reader can verify that for

$$t_2 - t_1 = t_3 - t_2 = 2.3\Delta_{\text{eff}}^{-1}(\approx \tfrac{2}{3}\pi\Delta_{\text{eff}}^{-1})$$

and $\alpha \leq 0.11$, Eq. (2a) is indeed violated. A similar underestimate of the critical α value can be obtained from (2b) but with $P_{\text{inc}}(t)$ replaced with its asymptotic long-time form (which overestimates its magnitude).[11] Doing this we find that Eq. (2b) is violated for

$$t_2 - t_1 = t_3 - t_2 = t_4 - t_3 = 0.84\Delta_{\text{eff}}^{-1}(\approx \tfrac{1}{4}\pi\Delta_{\text{eff}}^{-1})$$

and $\alpha \leq 0.08$.[12] Note, however, that quantum mechanics and rnacroscopic realism continue to differ even in the over- damped regime. Using the methods of Ref. 5, we can show that for $Q_1 = Q_3 = 1$, quantum mechanics would have the left-hand side of Eq. (1) exceed the right-hand side by

$$\left[\hbar \frac{\cot(\pi\alpha)}{4\tau k_B T} e^{[-(t_3-t_2)/\tau]}\right]$$

a quantity that can assume negative values.[13]

There is a slight difficulty in this argument arising from the nonzero (but exponentially small) probability of finding the system in regions $C_\pm$ (see Fig. 1), which is that once the system can have nearby q values, the concept of "macroscopically distinct states" becomes somewhat blurred. The easiest solution to this problem is to modify the macroscopic realism postulate (A1) to allow the system to be in a superposition of only two neighboring states (R and C_+, C_+ and C_-, etc.). We now assign to Q the value $+1(-1)$ if the system is in $R(L)$ alone, in $C_+(C_-)$ alone, or in a superposition of R and C_+ (L and C_-). The only combination which can affect Eqs. (1) and (2) is C_+ and C_-. Its contribution, however, cannot be more than a few times the total probability for find- ing the system in either C_+ or C_-, which is vanishingly small, and the incompatibility of quantum mechanics and macroscopic realism is not affected.

We now turn to the vexing question of whether the assumption (A2) of noninvasive measurability is likely to hold in practice. Indeed, ever since Heisenberg's "invention" of the "γ-ray microscope", we have all learned *not* to make such assumptions when dealing with microsystems, and at first sight there is no reason to treat macrosysterns differently. We can, nevertheless, make (A2) seem extremely natural and plausible by introducing the idea of an *ideal negative result experiment*. This is defined to be an experiment in which the measuring apparatus interacts with the system (and then very strongly) only if the latter has one value of $Q(t)$ (say $+1$), and does not interact at all otherwise. We can then confidently infer that $Q(t)$ has the value -1, if at time t the system does not elicit a response from the apparatus. Conjoined with the assumption of (A1) and (A2), this would, of course, not be formally macroscopic reality, this strongly suggests that the system also had $Q(t') = -1$ for t' immediately prior to the measurement at time t, and therefore that (at least in the limit of an arbitrarily short measurement) the apparatus could not have affected the dynamics of the system, i.e., that (A2) holds. Unlike the two-slit experiment where such a measurement can be made by shining light on one slit only, it is highly doubtful whether the analogous measurement could be made for an rf SQUID, but the difficulty seems to be technical and not conceptual. Under the assumption that an ideal negative-result experiment can be conducted, it is plain that all the quantities in Eqs. (1) and (2) can be measured. Suppose, for example, we wish to measure $\rho(Q_1 = 1, Q_3 = 1)$. Since the dynamics after t_3 are not of interest, we can use an ordinary measurement at t_3, and an ideal negative-result setup at t_1, which responds only if $Q(t_1) = -1$. We then simply discard those members of our ensemble which produce a response at t_1. Of the remainder, we count the number which have $Q(t_3) = 1$ and divide by the total number of members of the ensemble to obtain $\rho(Q_1 = 1, Q_3 = 1)$. By using a different ideal negative-result setup on another large and identical ensemble, we can obtain $\rho(Q_1 = -1, Q_3 = 1)$. We can thus calculate a value of K_{13}, and assumption (A2) al- lows us to assert that this is the K_{13} characteristic of the original ensemble.

An alternative to making ideal negative-result measurements is to couple the system to a microscopic probe. For example, in principle one could fire a neutron through the SQUID ring with its spin transverse to the magnetic field with a velocity such that it would precess precisely through an angle $\pm\pi/2$ if $q - \pm q_0$, and with a Larmor frequency much larger than Δ_{eff} but much less than the small oscillation frequency in either well. Let us consider how this method could be used to measure $\rho(Q_1, Q_2, Q_3)$, for example. For simplicity, let us prepare the system in a definite state (say $Q_1 = +1($ at time t_1 itself. We then fire our neutron to pass through the ring at t_2, and measure the flux at t_3 directly. Since the SQUID-neutron interaction is effectively instantaneous on the scale of Δ_{eff}^{-1}, we can infer the value of Q at time t_2 by measuring the neutron spin at any time after t_2, or even t_3!' A little thought shows that the quantum mechanical prediction (3) still holds with extra (small) corrections due to the finite duration of the measurement at t_2. Similar small corrections enter into the macroscopic-realisitic predictions (1) and (2), so that once again, the conflict between quantum mechanics and assumptions (A1) and A(2) is not affected.

In conclusion it should be emphasized that, should the quantum mechanically predicted results be obtained in a situation where they conflict with postulates (A1) and (A2), this would, of course, not be formally in conflict with the arguments so often given in discussions of the quantum theory of measurement to the effect that once a microsystem has interacted with a realistic measuring device, the device (and, if necessary, the microsystem) behave *as if* it were in a definite (and noninvasively measurable) macroscopic state: The macroscopic systems suitable for a macroscopic quantum coherence experiment are certainly *not* suitable to be measuring devices, at least under the conditions specified. But such a result might cause us to think a great deal harder about the significance of the "as if"!

This research was supported through the MacArthur Professorship endowed by the John D. and Catherine T. MacArthur Foundation. One of us (A.J.L.) is grateful to the Department of Physics at Harvard University for hospitality during the period where some of this work was performed.

[1] One must, of course, exclude here the genuine adherents of the relative-state ("many worlds") and mentalistic ("reduction-by-consciousness") interpretations of quantum mechanics. We strongly suspect that the number of physicists who in fact genuinely adhere to either of these interpretations (in the sense that it really makes a difference to the way they think about the macroscopic world) is considerably less than the number who claim to!

[2] 0ne can, of course, argue *ad nauseam* about the precise meaning of the phrase "macroscopically distinct". One specific objection which is sometimes raised with respect to a hypothetical experiment on a SQUID ring is that the difference in flux values between the two potential minima can be at most a fraction of the flux quantum ϕ+0 ≡ $\pi\hbar/e$ [A. J. Leggett, in Proceedings ofÃA TO Advanced Study Institute on Percolation, Localization, and Superconductivity, edited by A. Goldman and S. Wolf (Plenum, New York, 1984)]; it is therefore (it is argued) "only of order h" and therefore still in the quantum domain. We would regard this particular objection as merely verbal, since it is a historical accident that we treat the constants e and $\hbar$ as independent "fundamental constants" rather than say $\hbar$ and ϕ_0. A more sweeping objection is that any phenomenon which involves quantum interference effects can by definition not occur "at the macroscopic level". One can no more argue with this view than with the claim that the mere fact that a certain kind of behavior can be programmed into a computer *ipso facto* disqualifies it from being "intelligent". For our present purpose it is adequate that the "disconnectivity" (as defined as Ref. 3) of the superposition of "left"

and "right" states is of the order of the total number of electrons in the device ($\sim 10^{15} - 10^{23}$).

[3] A. J. Leggett, Prog. Theor. Phys. 69, 80 (1980).

[4] Leggett, Ref. 2.

[5] S. Chakravarty and A. J. Leggett, Phys. Rev. Lett. 52, 5 (1984). A much more detailed treatment of the argument leading to the results quoted in this reference, and of the corrections $\Delta P(t)$ in Eq. (24) due to"interblip" effects, is contained in A. J. Leggett et al. , to be published. In particular, it is shown that for any finite t a rigorous upper bound, which tends to zero as α^2 for small α, can be placed on the magnitude of the deviation of $P(t)$ from the expression given by the first two terms of (24) with $A(\alpha) = 1, q(\alpha) = 0$. Further, inspection of the formalism used by Chakravarty and Leggett makes it extremely plausible (though we have not yet succeeded in giving a rigorous proof) that the difference between the K_{ij} defined below and $P(t_j - t_i)$ is itself at most of the order of this deviation. If this is so, the effect of all these corrections would be at most a (probably very small) correction to the "critical" values of α, estimated in the text.

The work of Chakravarty and Leggett applies to the case of "Ohmic" dissipation (the case almost certainly realized in a SQUID), for which the spectral function $J(\omega)$ defined there is proportional to ω for small ω. One can show (see Leggett et al.) that for environments with $J(\omega) \sim \omega^p, p > 1$, quantum effects are less severely suppressed than in the Ohmic case. While environments with spectra corresponding to $0 < p < 1$ do not appear to be excluded by any *a priori* consideration, no mechanism which would produce such a state of affairs is known for SQUIDS, and it would presumably have dramatic (and, so far at least, unobserved) effects on the dynamics in the classically accessible regime. However, it is not known at present whether "$1/f$ noise" can be treated within the general framework of these references, and so our results may not be applicable to systems where such noise is appreciable.

[6] See, e.g. , R. de Bruyn Ouboter, in Proceedings of the Inter national Symposium on the Foundation of Quantum Mechanics in the Light of

New Technology, edited by S. Kamefuchi (Physical Society of Japan, Tokyo, 1984).

[7] In practice this is more likely to be a time ensemble.

[8] J.S. Bell, Physics (N,Y.) 1, 195 (1964). The inequality (15) of this paper is essentially the same as (2a) of our text.

[9] J.F. Clauser, M.A. Horne, A. Shimony, and R.A. Holt, Phys. Rev. Lett. 23, 880 (1969) The inequality (1a) of this papaser is (2b) of our text. See J.F. Clauser and A. Shimony, Rep. Prog. Phys. 41, 1881 (1978), for the various interpretations of this inequality, and also for discussions pertaining to imperfect/ inefficient detectors, many of which are applicable to the problem at hand.

[10] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. 47, 777 (1935)

[11] This argument exploits the additional fact that the error made in replacing $P_{\mathrm{inc}}(t)$ by its asymptotic form decreases with increasing t. The replacement, therefore, adds three negative quantities and one positive quantity (whose magnitude is less than that of any of the negative quantities) to the left-hand side of (2a). The net effect is to underestimate the left-hand side.

[12] It is easy to show that irrespective of the form of $P(t)$, (2a) and (2b) are maximally violated (if at all) for a given value of α for equally spaced times t_i.

[13] If α is too close to an integer or half-integer, the discrepancy is not accurately given by the expression in the text.

[14] We note in passing that since the neutron can be quite far from the SQUID at t_3, the situation has many of the seemingly paradoxical aspects of the EPR experiment. For example, suppose that the neutron spin was measured before the flux was measured at t_3, and that the two measurements were separated by a timelike interval. A local realist could argue that a measurement on the microsystem (neutron) was affecting the macrosystem (SQUID)!