Reflections on the quantum Zeno paradox

1986 J. Phys. A: Math. Gen. 19 1847

(http://iopscience.iop.org/0305-4470/19/10/025)

J. Phys. A: Math. Gen. 19 (1986) 1847-1854. Printed in Great Britain

Reflections on the quantum Zeno paradox

D Home† and M A B Whitaker

Department of Physics, University of Ulster, Coleraine, Northern Ireland BT52 1SA, UK

Received 11 October 1985

Abstract. We clarify some of the controversial aspects involved in the treatment of the quantum Zeno paradox. An argument for the paradox is given on the basis of the uncertainty principle, and the conditions for the paradox are analysed. Fleming's rule is proved in a simple way. Connection with the standard time-dependent perturbation theory is discussed. We compare the Zeno paradox with the other well known paradoxes of quantum mechanics and point out that the quantum Zeno paradox is immune to the use of the ensemble interpretation which has some success in tackling the other paradoxes. Remarks are made concerning the broader significance of the quantum Zeno paradox.

1. Introduction

The quantum Zeno paradox was first formally stated in the literature by Misra and Sudarshan (1977) followed by several other discussions (Chiu *et al* 1977 (to be referred to as CSM), Peres 1980a, Singh and Whitaker 1982, Peres 1984). The paradox stems from the observation that the survival probability of a given quantum system (i.e. the probability for finding the system in the initial state after being left to itself for a certain period of time) tends to unity in the limit of a continuous series of observations to find out whether the system is in the original state or not. This result hinges on two factors: the presence of an initial range of time for which the survival probability falls off as t^2, and the notion of wavefunction collapse during the observation process.

Even though there have been extensive discussions about departure from the exponential nature of the survival probability for short times (Khalfin 1958, Winter 1961, Fleming 1973, Peres 1980b and references therein), there remain areas of controversy highlighted by the recent papers due to Fleming (1978, 1983) and Chiu *et al* (1982).

In the present work we wish to clarify these controversial aspects as well as the essence of the quantum Zeno paradox by much simpler arguments showing how deeply rooted it is in the quantum formalism. In particular, we show that it persists within the ensemble interpretation of quantum mechanics. We also discuss various facets of the possible significance of the quantum Zeno paradox.

2. Argument from the generalised uncertainty principle

Our argument starts from the well known inequality (Gillespie 1970, p 68)

$$\Delta A \Delta B \geq \tfrac{1}{2}|\langle [A, B]\rangle|. \qquad (1)$$

† Permanent address: Theoretical Nuclear Physics Division, Saha Institute of Nuclear Physics, Calcutta 700009, India.

Here

$$\Delta A = \{\langle A^2\rangle - \langle A\rangle^2\}^{1/2} \tag{2}$$

is the uncertainty of the corresponding observable A for the state corresponding to the wavefunction of the system and ΔB is defined in an analogous fashion. (It is assumed that neither A nor B depends explicitly on time.) We may take for B the total Hamiltonian of the system, $\mathscr{H}$. Then, using

$$\langle [A, \mathscr{H}]\rangle = \mathrm{i}\hbar\, \mathrm{d}\langle A\rangle/\mathrm{d}t \tag{3}$$

we obtain

$$\Delta A \geqslant (\hbar/2\Delta E)|\mathrm{d}\langle A\rangle/\mathrm{d}t|. \tag{4}$$

We now choose, for A, the projection operator

$$A = |\psi\rangle\langle\psi| \tag{5}$$

where ψ is the initial decaying wavefunction.

Then we have

$$\langle A\rangle = \langle A^2\rangle = P(t) = |\langle\psi|\phi\rangle|^2 \tag{6}$$

where $\phi(t)$ is the wavefunction at time t:

$$\phi(t) = \exp(-\mathrm{i}\mathscr{H}t)\psi. \tag{7}$$

Then

$$(\Delta A)^2 = \langle A^2\rangle - \langle A\rangle^2 = P(1-P) \tag{8}$$

and, using (4), we finally arrive at

$$[P(1-P)]^{1/2} \geqslant (\hbar/2\Delta E)|\mathrm{d}P/\mathrm{d}t|. \tag{9}$$

However at $t=0$, P must be equal to unity. This gives directly that $(\mathrm{d}P/\mathrm{d}t)$ must be equal to zero at $t=0$, and hence that there is no linear term in the expansion for $P(t)$. This corroborates in a general way the presence of an initial range of time for which the survival probability falls off as t^2, which leads to the quantum Zeno paradox using the postulate of wavefunction collapse during the act of measurement.

3. Conditions for the presence of a Zeno region

From (9), the condition that the quantum Zeno effect may be observed is that ΔE is finite. This is actually a *sufficient* condition, but has not been shown to be a *necessary* one; we have obtained no information on the case where ΔE is infinite.

This appears to be in conflict with the result given by CSM. They state explicitly that it is *essential* for the existence of the Zeno paradox that $\mathscr{H}$ is semi-bounded, and an additional condition that $\langle\mathscr{H}\rangle$ is finite is *sufficient* to give the proof. (Our condition, in comparison, is that both $\langle\mathscr{H}\rangle$ and $\langle\mathscr{H}^2\rangle$ must be finite or, in different terms, that the first and second moments should be finite. Of course the origin of $\mathscr{H}$ may be chosen freely, but moving the origin by a finite amount cannot affect whether or not $\langle\mathscr{H}\rangle$ and $\langle\mathscr{H}^2\rangle$ are finite.)

We first point out that the question of semi-boundedness raised by CSM does not appear to be relevant. Their argument consists only of the example that the non-bounded Lorentzian distribution does give rise to a linear term in survival probability; this certainly does not show that semi-boundedness is essential. Our condition also gives correct results for the Lorentzian since, for this case, $\langle \mathscr{H}^2 \rangle$ is infinite.

The question of which moments need to be finite for a Zeno paradox to occur has recently been one of controversy. Chiu *et al* (1982) claimed that the condition is that *all* moments should be finite, but Fleming (1983) pointed out that he had already claimed to have shown (Fleming 1978) that only the first two moments need be finite (see also Danos and Johnson 1984). Our result clearly supports Fleming and is obtained in a particularly simple way.

Of the discussion given by Chiu *et al* (1982), one may suggest that, while the ideas expressed in the first five equations of their § 2 are well known in many branches of physics, and straightforward in use when all moments are finite, the implications when one or more moments are infinite may be more subtle. Let us consider, for example, the Lorentzian distribution, $\psi(\lambda)$, discussed in (6) of CSM. The fact that the second moment of $\psi(\lambda)$ is infinite tells us that the second derivative of $a(t)$ is infinite at the origin, and hence there is a discontinuity of slope. Since $\psi(\lambda)$ is even, the only possibility is that $a(t)$ has a non-zero slope at the origin and there is no t^2 region. Let us now consider, though, a function $\psi(\lambda)$ with second moment finite but fourth moment infinite. By a similar argument, there will be a discontinuity in third derivative of $a(t)$ at the origin, but this does not rule out a t^2 region and hence a quantum Zeno paradox.

4. Proof of Fleming's rule

From (9) we may easily obtain Fleming's rule (Fleming 1973, Peres 1980b). We perform the integral in the inequality

$$\left| \int_1^P \mathrm{d}P'/[P'(1-P')]^{1/2} \right| \leqslant 2(\Delta E)t/\hbar \tag{10}$$

to obtain

$$\sin^{-1}(2P-1) \geqslant \pi/2 - 2(\Delta E)t/\hbar \tag{11}$$

and, after a little manipulation,

$$P \geqslant \cos^2[(\Delta E)t/\hbar] \tag{12}$$

which is Fleming's rule.

The only limitation on this result arises from taking sines under an inequality, which leads to the condition $(\Delta E)t/\hbar \leqslant \pi/2$. This though merely corresponds to the obvious demand that $0 \leqslant P \leqslant 1$. Fleming's rule is significant for low values of t; for these values, a t^2 decay is slower than any decay proportional to t, but for larger values it is faster, and one may estimate a time of transition from the t^2 region to the t region (Peres 1984).

5. Use of time-dependent perturbation theory

It is interesting to note that the usual textbook time-dependent perturbation theory approach to such problems (e.g. Dicke and Wittke 1960, Schiff 1955), which is usually stated to give rise to a t dependence (and hence to the usual exponential type of decay), does, in fact, give a t^2 dependence in the appropriate region.

Rather than using the full Hamiltonian of the system, $\mathscr{H}$, as we have done so far, the perturbation theory approach breaks the Hamiltonian into two parts, $\mathscr{H} = \mathscr{H}_0 + V$, where the decaying state is an eigenstate of $\mathscr{H}_0$, and V is to be regarded as a perturbation. It is easy to establish a connection between the two approaches as

$$\begin{aligned} \langle m|\mathscr{H}|j\rangle &= \langle m|V|j\rangle \qquad m \neq j \\ &= \langle m|\mathscr{H}_0|j\rangle \qquad m = j \end{aligned} \tag{13}$$

where $|m\rangle$ is the decaying state and $|j\rangle$ is any eigenstate of $\mathscr{H}_0$.

The usual formula for the probability of decay is

$$Q(t) = 1 - P(t) = \frac{1}{\hbar^2}\int_{-\infty}^{\infty} |\langle m|V|k\rangle|^2 \rho_k \sin^2\left(\frac{(E_k - E_m)t}{2\hbar}\right)\left(\frac{E_k - E_m}{2\hbar}\right)^{-2} \mathrm{d}E_k \tag{14}$$

where ρ_k is the density of states $|k\rangle$. It is then usually stated that the integral is small except for values of E_k such that $|E_k - E_m|$ is small, and therefore it is permissible to assume that $|\langle m|V|k\rangle|^2$ and ρ_k are constant. With z equal to $(E_k - E_m)t/2\hbar$, and using the fact that $\int_{-\infty}^{\infty} \mathrm{d}z \sin^2 z/z^2$ is equal to π, one easily obtains

$$Q(t) = 2\pi\rho_k t|\langle m|V|k\rangle|^2/\hbar \tag{15}$$

the decay, as required, being proportional to t.

However our initial assumption that the integral is small unless $|E_k - E_m|$ is small is essentially a long-time condition. For short times one obtains trivially

$$Q(t) = \frac{1}{\hbar^2}\left(\int_{-\infty}^{\infty} |\langle m|V|k\rangle|^2 \rho_k \, \mathrm{d}E_k\right)t^2 \tag{16}$$

clearly a t^2 dependence. Of course, the prediction is not exact, because we are working to first order in perturbation theory, but, assuming there are no divergences, the t^2 dependence should be exact for infinitesimally small time. Thus, contrary to the view usually expressed in textbooks, one will expect a t^2 region whether the final states are discrete or form a continuum.

Computed calculations of (14) for particular forms of $|\langle m|V|k\rangle|^2\rho_k$ show a gradual transition from t^2 dependence to t dependence. An example is given in figure 1, where $|\langle m|V|k\rangle|^2\rho_k$ has been taken as a finite set of equally spaced delta functions centred on $|m\rangle$. (Such simple systems demonstrate clearly the required trend, but also introduce extraneous, presumably unphysical, features, due to the discrete rather than continuous function used, and the periodicity also involved. Thus the results of figure 1 show oscillatory behaviour.)

The framework described obviously breaks down if the integral which is the coefficient of t^2 in (16) becomes infinite. Using (13) we note that

$$\langle \mathscr{H}^2 \rangle = \int_{-\infty}^{\infty} \mathrm{d}E_k \, \rho_k |\langle m|V|k\rangle|^2 + \langle m|\mathscr{H}|m\rangle^2 \tag{17}$$

and hence the integral becomes equal to $\langle \mathscr{H}^2\rangle - \langle \mathscr{H}\rangle^2$. We therefore expect a quantum Zeno paradox only if $\langle \mathscr{H}^2\rangle - \langle \mathscr{H}\rangle^2$ is finite, in conformity with the treatment of § 3.

Since completion of the present work we have become aware of an interesting paper by Joos (1984), who considers in some detail a problem similar to the one

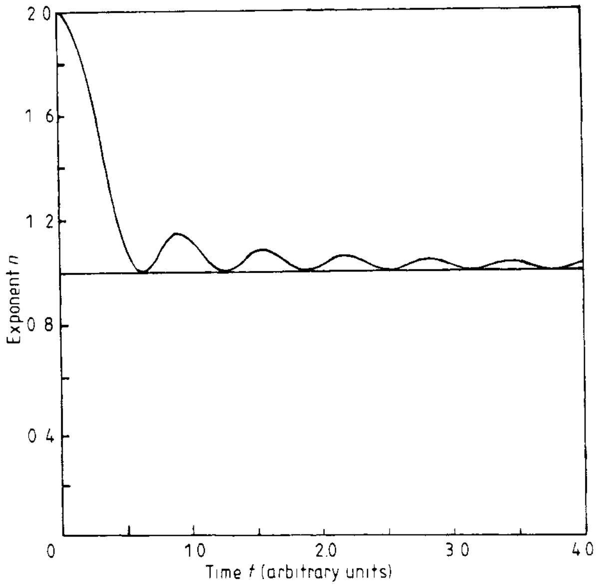

Figure 1. Form of dependence on t for the function given in (14). We take the particular case where s, given by $|\langle m|V|k\rangle|^2\rho_k$, is a series of 20 000 delta functions equally spaced and centred on E_m. The required exponent, n, may then be obtained from the ratio of the increments in the logarithms of s and t between successive closely spaced values of t. The results show clearly that n tends to 2 for infinitesimally small t, and to 1 for large values of t. (The oscillations are a result of the rather unphysical periodicity in s and are irrelevant.)

discussed here, but from a different point of view. He discusses the measurement process both using wavepacket reduction, and taking into account the measuring apparatus. He attempts to distinguish between circumstances which will cause a quantum Zeno effect (which he terms the 'watchdog effect') and those for which the behaviour may be represented by a master equation with suppression of interference terms and constant transition rates.

6. Zeno and other paradoxes in various interpretations of quantum mechanics

The Zeno paradox stands apart in a number of ways from the other well known paradoxes of quantum theory and interpretation such as Einstein–Podolsky–Rosen (EPR) and Schrödinger's cat. (Here, as elsewhere in the paper, we use the word 'paradox' non-controversially to describe an experiment with results or interpretation considered to be surprising.)

In the first place, the EPR and Schrödinger paradoxes are paradoxes of interpretation. The Zeno paradox, on the other hand, is a paradox of prediction; it is the predicted behaviour that is difficult to accept. In principle the Zeno paradox may be tested; either the decay is inhibited or not. Equally, if different interpretations of quantum mechanics gave different predictions as to the occurrence or otherwise of the Zeno paradox, the choice of interpretations would be open to experimental test. (In contrast, when one applies different interpretations to, for example, the EPR paradox, one is only asking whether they can explain the accepted facts without unduly stretching one's credulity. Of course there are extensions of the EPR example of the Bell type, where the results may genuinely be open to question and experimental test.)

All the paradoxes result from collapse of wavefunction at an observation but the Zeno paradox is connected with a different aspect of the problem from the other paradoxes. To clarify, we set out the basic facts of wavefunction collapse. Suppose a measurement is made of an observable O with associated operator $\hat{O}$, and which has eigenvalues and eigenfunctions O_m and α_m. If the wavefunction before the measurement is given by $\Sigma_n c_n \alpha_n$, the probability of obtaining the value O_m is $c_m^* c_m$; if the value O_m is obtained, the von Neumann theory of measurement says that the wavefunction immediately after the observation is α_m.

Paradoxes such as EPR and Schrödinger's cat exploit the collapse from a linear combination of eigenfunctions to a single one in a variety of ways. In the quantum Zeno paradox, the collapse of wavefunction has two distinct roles to play. First, rather trivially, the collapsed wavefunction tells us whether the system has decayed or not. The more important role is that the part of the wavefunction corresponding to the survival of the system is separated from the part corresponding to the decayed system. As Peres (1980a) clearly shows, it is the presence of the latter that gives rise to the further decay of the system. The crucial point responsible for the Zeno paradox, then, is not the change from a linear combination of eigenfunctions to a single one, but the change from pure state to mixed state due to the observation.

In an ensemble interpretation, the fact that a *particular* result is obtained in a measurement (survival or decay) is of considerably reduced importance in the development of the wavefunction. However, the fact that *any* result is obtained is sufficient that the state of the system becomes mixed. After each measurement in a Zeno sequence, one must work with a new sub-ensemble corresponding to non-decayed systems only. Thus the part of the wavefunction for this sub-ensemble is 'cut off from' that for the sub-ensemble of decayed systems. (The linear combination of the two parts is broken.) Again, this change reduces, or in the extreme case, eliminates, the possibility of decay and this is the quantum Zeno paradox.

Let us discuss this in density-matrix language, the natural language to use for the ensemble interpretation. We may use 2×2 matrices, the first state being for surviving systems, the second for decayed systems. (In other contexts we would need to regard the latter as a continuum of separate states, but here that is not necessary.) The density matrix may be written as

$$\begin{pmatrix} \rho_{ss} & \rho_{sd} \\ \rho_{ds} & \rho_{dd} \end{pmatrix}.$$

At $t=0$, naturally ρ_{ss} is unity, and other elements zero. At the time of the first measurement, ρ_{ss} has decreased from unity, ρ_{dd} has increased to $1-\rho_{ss}$ and the off-diagonal elements have become non-zero, so that the matrix is idempotent. During the measurement, the off-diagonal elements become zero, the diagonal elements do not change and the state becomes mixed as the density matrix is non-idempotent. Subsequently we need to consider only the matrix corresponding to the sub-ensemble for the surviving systems, so its form reverts to that at $t=0$, but corresponds, of course, to a smaller number of systems, the number surviving the first decay period. The procedure may be taken through each decay period and measurement. After each measurement the ensemble refers to a smaller number of systems. The slowing down of the decay is caused because, at each measurement, the off-diagonal terms are stripped off. Since the rate of growth of ρ_{dd} is proportional to the magnitude of these off-diagonal terms, and the rate of growth of these elements is proportional in magnitude to ρ_{ss}, the t^2 dependence of ρ_{dd} after each measurement is clear. It is this that leads to the paradox.

(The fact that the quantum Zeno paradox persists within an ensemble interpretation appears to make it more acute a problem than EPR and Schrödinger paradoxes, which are usually claimed (e.g. Ballentine 1970) not to exist within the ensemble interpretation. This is because only the diagonal elements are important in the study of such paradoxes, and the washing-out of the off-diagonal elements due to measurement is not significant.)

7. Significance of the quantum Zeno paradox

Should the quantum Zeno paradox be regarded as a genuine problem, or merely an abstract metaphysical curiosity? For many important types of decay, and with a certain outlook to what constitutes a measurement, it has been claimed to be unimportant. This approach assumes the measurement to be a physical process lasting a certain period of time, which may be, in general, the apparatus response time. This time is usually considered to be longer than the Zeno time (defined as the time for which the t^2 dependence of the survival probability is predicted) for most atomic and nuclear decays (Peres 1984).

Even with this approach, one still needs to pay particular attention to the possibility that the Zeno paradox may affect the proton decay conjectured by grand unified theories. The proton lifetime estimate based on the grand unified theories is about 10^{31} yr (Georgi and Glashow 1974, Georgi *et al* 1974, Langacker 1981), but the experimental lower limit is larger by about two orders of magnitude (Bionta *et al* 1983). Two different explanations have been suggested, both using the concept of Zeno time. The one by Khalfin (1982) considers the Zeno time to be greater than the age of the universe, taken to be about 10^{10} yr. The other one, suggested by Fleming (1983), argues that the Zeno interval for proton decay is longer than the experimental time resolution for observing the proton decay. However, the latter contention contradicts the argument by Chiu *et al* (1982) who claim to demonstrate the implausibility of the Zeno time for proton decay being as large as the pertinent time resolution for the experiment of about 10^{-12} s. Since the theoretical basis for estimating the Zeno time is rather dubious, it is difficult to assess properly this range of ideas.

The importance of the quantum Zeno paradox is accentuated if one adheres to the concept of measurement closer to that of Dirac or von Neumann, the idea that an observation is accompanied by an instantaneous collapse of wavefunction, irreducible to the evolution of wavefunction via the Schrödinger equation. If the measurement process is instantaneous, one is led inevitably to the quantum Zeno paradox whatever the Zeno time.

We believe that the quantum Zeno paradox is an instructive problem to probe the foundations of quantum mechanics, particularly since it appears relatively immune to the type of interpretation of quantum mechanics one adopts.

Empirical realisability of the Zeno paradox calls for comprehensive investigation. Perhaps at this stage one needs to envisage a concrete form of gedanken experiment, similar to the type formulated by Bohm for the EPR paradox.

Finally we would like to emphasise an unexplored question: does the quantum Zeno paradox persist in the macroscopic limit? It is transparent that the quantum Zeno paradox is a direct consequence of the quantum formalism using the definition of survival probability based on the collapse of wavefunction. However it is not clear how the quantum Zeno paradox can disappear, as one would expect, in the macroscopic limit, where the concept of continuous measurement involving instantaneous observations is surely an operationally viable notion. The 'disturbance' idea of the Copenhagen interpretation associated with the collapse of wavefunction seems to be rather vague for clarifying this issue. It would be interesting to explore whether approaches which claim to furnish 'realist' descriptions of quantum mechanics can provide insight into this problem (Bohm and Hiley 1985, Conrad *et al* 1985).

Acknowledgments

One of the authors (DH) is indebted to the British Council Division (British Deputy High Commission, Calcutta) for financial support, and to Professors S Sengupta, B Datta Roy, David Bohm and Brian Josephson for stimulating discussions.

References

Ballentine L E 1970 *Rev. Mod. Phys.* **42** 358
Bionta R M *et al* 1983 *Phys. Rev. Lett.* **51** 27
Bohm D and Hiley B J 1985 *Paper presented at Int. Conf. on Microphysical Reality and Quantum Formalism, Urbino, Italy*
Chiu C B, Sudarshan E C G and Misra B 1977 *Phys. Rev.* D **16** 520
—— 1982 *Phys. Lett.* **117B** 34
Conrad M, Home D and Josephson B D 1985 *Paper presented at the Int. Conf. on Microphysical Reality and Quantum Formalism, Urbino, Italy*
Danos M and Johnson A B 1984 *Phys. Rev.* D **30** 2692
Dicke R H and Wittke J P 1960 *Introduction to Quantum Mechanics* (Reading, MA: Addison-Wesley)
Fleming G N 1973 *Nuovo Cimento* A **16** 232
—— 1978 *Nuovo Cimento* A **46** 579
—— 1983 *Phys. Lett.* **125B** 287
Georgi H and Glasgow S L 1974 *Phys. Rev. Lett.* **32** 438
Georgi H, Quinn H R and Weinberg S 1974 *Phys. Rev. Lett.* **33** 451
Gillespie D T 1970 *A Quantum Mechanics Primer* (Leighton Buzzard, UK: Intertext)
Joos E 1984 *Phys. Rev.* D **29** 1626
Khalfin L A 1958 *Sov. Phys.-JETP* **6** 1053
—— 1982 *Phys. Lett.* **112B** 223
Langacker P 1981 *Phys. Rep.* **72** 185
Misra B and Sudarshan E C G 1977 *J. Math. Phys.* **18** 765
Peres A 1980a *Am. J. Phys.* **48** 931
—— 1980b *Ann. Phys., NY* **129** 33
—— 1984 *Found. Phys.* **14** 1131
Schiff L I 1955 *Quantum Mechanics* (New York: McGraw-Hill)
Singh I and Whitaker M A B 1982 *Am. J. Phys.* **50** 882
Winter R G 1961 *Phys. Rev.* **123** 1503