Definition of decoherence

J. Finkelstein*
Department of Physics, San José State University, San José, California 95192
(Received 8 January 1993)

We examine the relationship between the decoherence of quantum-mechanical histories of a closed system (as discussed by Gell-Mann and Hartle) and environmentally induced diagonalization of the density operator for an open system. We study a definition of decoherence which incorporates both of these ideas, and show that it leads to a consistent probabilistic interpretation of the reduced density operator.

PACS number(s): 03.65.Bz, 98.80.Hw

I. INTRODUCTION

In classical physics, there is no apparent need to analyze the role of the observer. As long as one assumes that observations can be made with negligible effect on the properties observed, one can ascribe properties to objects without considering whether or not those properties have actually been observed. In quantum physics, at least in the Copenhagen interpretation, the situation is radically different: objects are assumed not to have any properties at all, unless and until those properties have been measured by an outside observer. Under such an assumption, an analysis of measurement becomes an indispensable part of the interpretation of the theory.

In quantum cosmology the "object" of study is the entire Universe. The quantum theory of the Universe must be interpreted without relying on the idea of measurement. This is not, as was the case in classical physics, because measurement is innocuous; rather, it is because it is impossible. There cannot be any outside observer of the entire universe.

It has been suggested [1, 2] by Gell-Mann and Hartle (GH) that the Copenhagen notion of measurement could be replaced by the concept of "decohering histories." This concept is a generalization of the idea of "consistent histories" advanced by Griffiths [3] and studied also by Omnès [4]; a recent paper by Dowker and Halliwell [5] analyzes several examples of decohering histories. GH suggest several different versions of the condition for decoherence (which they refer to as the weak, medium, medium-strong, and strong conditions), and it is an open question within the program of GH of what is the best definition of decoherence to impose.

Other authors [6] have investigated how, when a system interacts with its environment, the density operator for the system becomes diagonal (in a particular basis). This diagonalization can also be argued to provide an answer to the question of how, without relying on the idea of measurement, it is possible to say when a property is real or when an event has happened; in fact, this diagonalization has also been termed decoherence. In the following, we shall refer to the diagonalization of the density operator for a system, due to its interaction with its environment, as Z decoherence.

In this paper we study the properties of yet another definition of the decoherence of histories, which we will refer to as partial trace (PT) decoherence. We will see that PT decoherence implies the medium decoherence condition of GH, and that it also involves the diagonalization of the density operator; in a sense, PT decoherence can be taken to characterize those histories which decohere by the mechanism of Z decoherence. In the next section we will review the definitions of histories and decoherence as they were used in Refs. [1, 2, 5]; we will then define a generalization of the decoherence functional, and use it to state the condition that we call PT decoherence. In the final section we will establish and discuss some of the properties of histories which exhibit PT decoherence, and compare them with properties implied by other definitions.

II. FORMALISM

Since we are interested in those sets of histories for which we can discuss Z decoherence, we will limit our discussion to the case in which, from among all the variables which describe the world, we distinguish a certain fixed subset of them, and say that this subset describes the "system"; the remaining variables then describe the "environment." Formally, we write the state space $\mathcal{H}$ of the world as the tensor product of spaces $\mathcal{S}$ of the system and $\mathcal{E}$ of the environment: $\mathcal{H} = \mathcal{S} \otimes \mathcal{E}$. The discussion of histories will involve (Schrödinger picture) projection operators P, and we only consider those projection operators which act trivially on $\mathcal{E}$, and so can be written $P = P_{\mathcal{S}} \otimes I_{\mathcal{E}}$.

Following GH, we define a history by a sequence of projections made at a fixed sequence of times $t_1 \cdots t_n$ which satisfy $t_1 < \cdots < t_n$. For each given time t_k, the set of projection operators will be denoted by $\{P^k_{\alpha_k}\}$, where a particular value of the index α_k denotes a particular operator in the set. The superscript k on P is required

*Electronic address: JFINKEL@SJSUVM1.BITNET

since we are allowing different sets of projectors at different times. For each fixed value of k, the projections are supposed to represent an exhaustive set of exclusive alternatives, which implies that these operators satisfy $P_\alpha P_\beta = \delta_{\alpha\beta} P_\alpha$ and $\sum_\alpha P_\alpha = I$. A particular history is then labeled by a sequence $\alpha = [\alpha_1, \ldots, \alpha_n]$. For a particular sequence α, define

$$C_\alpha = P^n_{\alpha_n} e^{-iH(t_n - t_{n-1})} \times P^{n-1}_{\alpha_{n-1}} e^{-iH(t_{n-1}-t_{n-2})} \cdots P^1_{\alpha_1} e^{-iH(t_1 - t_0)}. \quad (1)$$

The condition of the world at the initial time $t_0 < t_1$ will be represented by the density operator $\rho(t_0)$. Then the decoherence functional is [1, 2]

$$D(\alpha', \alpha) = \mathrm{Tr}[C_{\alpha'} \rho(t_0) C^\dagger_\alpha]. \quad (2)$$

Although the definition of C_α given in Eq. (1) differs slightly from that used in [1,2], the decoherence functional defined in Eq. (2) agrees exactly with the one defined in [1,2].

The definition of decoherence that was emphasized in [2] was the "medium decoherence" condition; this is the requirement that, for two alternative histories labeled by α and α',

$$D(\alpha', \alpha) = 0. \quad (3)$$

We will say that a family of histories for which Eq. (3) holds satisfies "GH decoherence." Probabilities $p(\alpha)$ can be assigned to the members of a family of histories satisfying Eq. (3) by $p(\alpha) = D(\alpha, \alpha)$; Eq. (3) then guarantees that $p(\alpha \,\mathrm{or}\, \beta) = p(\alpha) + p(\beta)$, where $p(\alpha \,\mathrm{or}\, \beta)$ is calculated from Eq. (2) with $C_{\alpha \text{ or } \beta} = C(\alpha) + C(\beta)$.

One would certainly expect that, in cases where the GH decoherence condition [Eq. (3)] is (at least approximately) satisfied, this would come about because tracing over the environment in Eq. (2) would give (at least approximately) zero; this is well illustrated by examples considered in [2] and in [5]. In particular, the example that GH present in Sec. IV of [2] in order to discuss the relation with Z decoherence was constructed with this expectation. In order to formalize this expectation, we now define a generalized decoherence functional $\bar{D}$ by restricting the trace in Eq. (2) to be just over the environment; that is, we define

$$\bar{D}(\alpha', \alpha) = \mathrm{Tr}_{\mathcal{E}}[C_{\alpha'} \rho(t_0) C^\dagger_\alpha]. \quad (4)$$

For fixed α and α', $\bar{D}(\alpha', \alpha)$ is thus an operator in $\mathcal{S}$, which can easily be shown to satisfy

$$\mathrm{Tr}\bar{D}(\alpha', \alpha) = D(\alpha', \alpha), \quad (5)$$

$$\bar{D}(\alpha, \alpha') = \bar{D}^\dagger(\alpha', \alpha), \quad (6)$$

$$\sum_{\alpha', \alpha} \bar{D}(\alpha', \alpha) = \rho_{\mathcal{S}}(t_n), \quad (7)$$

where, in Eq. (7), $\rho_{\mathcal{S}}$ is the density operator for the system $\rho_{\mathcal{S}} = \mathrm{Tr}_{\mathcal{E}}[\rho]$, and we have used $\sum_\alpha C_\alpha = e^{-iH(t_n - t_0)}$. We can now state the condition that we call (since it is defined by a partial trace) PT decoherence: we will say a family of histories satisfies PT decoherence if, for any alternative histories α and α',

$$\bar{D}(\alpha', \alpha) = 0. \quad (8)$$

In a realistic situation, we would expect Eq. (8) to be, at best, approximately satisfied. The matrix elements of $\bar{D}$ can be shown to satisfy the following inequality:

$$|[\bar{D}(\alpha', \alpha)]_{ij}|^2 \le [\bar{D}(\alpha', \alpha')]_{ii} [\bar{D}(\alpha, \alpha)]_{jj}.$$

The generalization of the condition for approximate decoherence suggested in [5] would be to require the left-hand side of this inequality to be, if not zero, at least much less than the right-hand side, for all i and j. However, we shall not consider further the notion of approximate decoherence; in the following, decoherence will be taken to be exact.

In the next section we will establish and discuss the following properties of PT decoherence.

(1) PT decoherence implies the medium-decoherence condition of GH.

(2) For a family of histories satisfying PT decoherence, the projectors at the final time t_n must be, in the space $\mathcal{S}$, onto (perhaps a coarse graining of) a basis which diagonalizes $\rho_{\mathcal{S}}(t_n)$. In this sense a PT-decoherent family of histories satisfies Z decoherence.

(3) For each history α in a decoherent family of histories, one can define an "effective" density operator for the system $\rho^\alpha_{\mathcal{S}}$; PT decoherence then implies, and is essentially implied by, a consistency condition for these effective density operators:

$$p(\alpha \,\mathrm{or}\, \beta) \rho^{\alpha \text{ or } \beta}_{\mathcal{S}} = p(\alpha) \rho^\alpha_{\mathcal{S}} + p(\beta) \rho^\beta_{\mathcal{S}}.$$

(4) If ρ, the density operator for the world, represents a pure state, then GH decoherence implies [1, 2] the existence of "generalized records" of the histories. PT decoherence then implies further that these generalized records exist *in the environment*.

(5) To the extent to which it is possible to neglect the interaction between the system and the environment for times later than t_n, histories extended past t_n satisfy the following property: any two histories which are PT decoherent at t_n will continue to be PT decoherent at all later times, for any choice of projections at these later times. In the case where ρ represents a pure state, this would imply the *persistence* of records.

In this paper we are considering a world with a fixed split into "system" and "environment," and so we have required that all projection operators act trivially on $\mathcal{E}$. However, we will only need this requirement to establish the properties numbered (2) and (5) above, and then only for projection operators at time t_n or later. Thus, we could drop this requirement at times earlier than t_n without changing any of our results.

III. IMPLICATIONS OF THE FORMALISM

First, any two histories which are PT decoherent [Eq. (8)] are necessarily GH-decoherent [Eq. (3)]. This follows from Eq. (5).

Second, to discuss a relationship between PT decoherence and Z decoherence, suppose we sum a family of histories over all projectors at all times earlier than t_n; if α represents such a summed history, it follows from Eq. (1) that $C_\alpha = P^n_{\alpha_n} e^{-iH(t_n-t_0)}$. Since all of our projectors act trivially in $\mathcal{E}$, we can write $P^n_{\alpha_n} = P_{\mathcal{S}\alpha} \otimes I_{\mathcal{E}}$, where $P_{\mathcal{S}\alpha}$ is a projection operator acting on $\mathcal{S}$. So if α and α' are two such histories, we have

$$\begin{aligned}\bar{D}(\alpha',\alpha) &= \mathrm{Tr}_{\mathcal{E}}[(P_{\mathcal{S}\alpha'} \otimes I_{\mathcal{E}})e^{-iH(t_n-t_0)} \\ &\quad \times \rho(t_0)e^{iH(t_n-t_0)}(P_{\mathcal{S}\alpha} \otimes I_{\mathcal{E}})] \\ &= P_{\mathcal{S}\alpha'}\mathrm{Tr}_{\mathcal{E}}[\rho(t_n)]P_{\mathcal{S}\alpha} = P_{\mathcal{S}\alpha'}\rho_{\mathcal{S}}(t_n)P_{\mathcal{S}\alpha}. \qquad (9)\end{aligned}$$

Thus, since when $\alpha' \neq \alpha$, PT decoherence requires $\bar{D}(\alpha',\alpha) = 0$, Eq. (9) implies, in this case,

$$P_{\mathcal{S}\alpha'}\rho_{\mathcal{S}}(t_n)P_{\mathcal{S}\alpha} = 0. \qquad (10)$$

So $\rho_{\mathcal{S}}(t_n)$ has no matrix elements between the subspaces of $\mathcal{S}$ projected onto by $P_{\mathcal{S}\alpha'}$ and $P_{\mathcal{S}\alpha}$. If each member of the set $\{P_{\mathcal{S}\alpha}\}$ projects onto a one-dimensional subspace of $\mathcal{S}$, i.e., if $P_{\mathcal{S}\alpha} = |\alpha\rangle\langle\alpha|$, then Eq. (10) obviously says that $\rho_{\mathcal{S}}(t_n)$ is diagonal in the basis $\{|\alpha\rangle\}$. More generally, Eq. (10) says that $\rho_{\mathcal{S}}(t_n)$ is block diagonal in the subspaces projected onto by the $\{P_{\mathcal{S}\alpha}\}$; since it is always possible to diagonalize $\rho_{\mathcal{S}}(t_n)$ on each such subspace separately, this means that there exists a basis in which $\rho_{\mathcal{S}}(t_n)$ is diagonal and such that each $P_{\mathcal{S}\alpha}$ is a sum of projectors onto that basis.

This establishes that the set of projectors for the final time t_n of a PT-decoherent family necessarily project onto (possibly a coarse graining of) a basis which diagonalizes $\rho_{\mathcal{S}}(t_n)$. Therefore, at time t_n a PT-decoherent family also exhibits Z decoherence. However, at earlier times the situation is not so simple. If we consider, for example, two histories α and α' for which the only nontrivial projections are at $t = t_{n-1}$, we could write, in place of Eq. (9),

$$\begin{aligned}\bar{D}(\alpha',\alpha) &= \mathrm{Tr}_{\mathcal{E}}[\, e^{-iH(t_n-t_{n-1})}(P_{\mathcal{S}\alpha'} \otimes I_{\mathcal{E}})\rho(t_{n-1}) \\ &\quad \times (P_{\mathcal{S}\alpha} \otimes I_{\mathcal{E}})e^{iH(t_n-t_{n-1})}]. \qquad (11)\end{aligned}$$

The exponential factors in Eq. (11) do not necessarily cancel because $\mathrm{Tr}_{\mathcal{E}}$ is not cyclic in operators which act upon $\mathcal{S}$. Hence we *cannot* conclude that, in analogy with Eq. (10), $P_{\mathcal{S}\alpha'}\rho_{\mathcal{S}}(t_{n-1})P_{\mathcal{S}\alpha} = 0$, which would be the Z-decoherence condition at $t = t_{n-1}$.

Third, for each history α in a GH-decoherent family one can define an "effective" density operator ρ^α by

$$\rho^\alpha = \frac{C_\alpha\rho(t_0)C_\alpha^\dagger}{\mathrm{Tr}[C_\alpha\rho(t_0)C_\alpha^\dagger]}. \qquad (12)$$

Using this effective density operator corresponds, in the case of a pure state, to the "collapse of the state vector." We can also define the effective density operator for the system by

$$\rho_{\mathcal{S}}^\alpha \equiv \mathrm{Tr}_{\mathcal{E}}[\rho^\alpha] = \frac{\mathrm{Tr}_{\mathcal{E}}[C_\alpha\rho(t_0)C_\alpha^\dagger]}{\mathrm{Tr}[C_\alpha\rho(t_0)C_\alpha^\dagger]} \qquad (13)$$

which implies

$$\bar{D}(\alpha,\alpha) = p(\alpha)\rho_{\mathcal{S}}^\alpha. \qquad (14)$$

Now let α and β represent two alternative histories; from Eq. (4) with $C_{\alpha \,\mathrm{or}\, \beta} = C_\alpha + C_\beta$ we get

$$\bar{D}(\alpha\,\mathrm{or}\,\beta,\alpha\,\mathrm{or}\,\beta) = \bar{D}(\alpha,\alpha)+\bar{D}(\beta,\beta)+\bar{D}(\alpha,\beta)+\bar{D}(\beta,\alpha). \qquad (15)$$

If α and β represent alternative members of a PT-decoherent family, the last two terms in Eq. (15) vanish; then Eqs. (14) and (15) imply

$$\rho_{\mathcal{S}}^{\alpha \,\mathrm{or}\, \beta} = \frac{p(\alpha)}{p(\alpha)+p(\beta)}\rho_{\mathcal{S}}^\alpha + \frac{p(\beta)}{p(\alpha)+p(\beta)}\rho_{\mathcal{S}}^\beta, \qquad (16)$$

or if we sum over *all* alternative members of the family, we get

$$\rho_{\mathcal{S}} = \sum_\alpha p(\alpha)\rho_{\mathcal{S}}^\alpha. \qquad (17)$$

Equation (16) is the consistency condition for the effective density operator which is implied by PT decoherence. A necessary and sufficient condition for Eq. (16) is that the sum of the last two terms in Eq. (15) vanishes, for any alternative histories α and β; by using Eq. (6), we can write this condition as

$$\bar{D}(\alpha,\beta) + \bar{D}^\dagger(\alpha,\beta) = 0. \qquad (18)$$

This condition is somewhat weaker than is PT decoherence; it implies the weak, but not the medium, decoherence condition defined by GH.

Fourth, if the density operator ρ represents a pure state, we can write $\rho(t_0) = |\Psi\rangle\langle\Psi|$. The decoherence functional defined in Eq. (2) is then $D(\alpha',\alpha) = \langle\Psi|C_\alpha^\dagger C_{\alpha'}|\Psi\rangle$, and the GH decoherence condition is that, for α and α' representing alternative histories, $C_\alpha|\Psi\rangle$ and $C_{\alpha'}|\Psi\rangle$ are orthogonal. The fact that alternative histories lead, at time $t = t_n$, to orthogonal states implies what GH refer to as "generalized records" of the histories. The PT-decoherence condition becomes, in the case of a pure state,

$$\mathrm{Tr}_{\mathcal{E}}[C_{\alpha'}|\Psi\rangle\langle\Psi|C_\alpha^\dagger] = 0. \qquad (19)$$

It can be shown that Eq. (19) is equivalent to the condition that there exist subspaces $\mathcal{E}_\alpha$ and $\mathcal{E}_{\alpha'}$ of $\mathcal{E}$ such that $C_\alpha|\Psi\rangle \in S\otimes\mathcal{E}_\alpha$, $C_{\alpha'}|\Psi\rangle \in S\otimes\mathcal{E}_{\alpha'}$, and $\mathcal{E}_\alpha$ is orthogonal to $\mathcal{E}_{\alpha'}$. We can call this condition "orthogonality in the environment."[1]

In the case in which ρ represents a pure state, the PT-decoherence condition is therefore equivalent to the statement that $C_\alpha|\Psi\rangle$ and $C_{\alpha'}|\Psi\rangle$ are orthogonal in the environment. We can thus say that a PT-decoherent family of histories produces records in the environment.

Fifth, the histories we have discussed so far extend until a time t_n, which we can call the "present." Suppose we now consider further extending these histories to a

[1]If $\mathcal{E}_\alpha$ and/or $\mathcal{E}_{\alpha'}$ has dimension greater than one, it is possible to make a finer graining (perhaps in a branch-dependent way). For our discussion it is simpler not to assume this has been done.

time $t_m > t_n$, the "future." Let α represent, as before, a sequence of alternatives at times $t_1 \cdots t_n$, and let β represent a sequence at times $t_{n+1} \cdots t_m$. Then if we denote by (α and β) the sequence consisting of first α and then β, we can write, using an obvious extention of the notation introduced in Eq. (1), $C_{\alpha \text{ and } \beta} = C_\beta C_\alpha$. The generalized decoherence functional becomes

$$\bar{D}(\alpha' \text{ and } \beta', \alpha \text{ and } \beta) = \mathrm{Tr}_\mathcal{E}[C_{\beta'} C_{\alpha'} \rho(t_0) C_\alpha^\dagger C_\beta^\dagger]. \tag{20}$$

Now let us suppose that we could ignore the interaction between the system and the environment for all times in the future. Then, for future times, we could write the Hamiltonian H of the world as a sum: $H = H_\mathcal{S} \otimes I_\mathcal{E} + I_\mathcal{S} \otimes H_\mathcal{E}$. This means that C_β could be written

$$C_\beta = C_{\mathcal{S}\beta} \otimes e^{-iH_\mathcal{E}(t_m - t_n)}, \tag{21}$$

where $C_{\mathcal{S}\beta}$ is an operator on $\mathcal{S}$; there would be a similar expression for $C_{\beta'}$. Then Eq. (20) would imply

$$\begin{aligned} \bar{D}(\alpha' \text{ and } \beta', \alpha \text{ and } \beta) &= C_{\mathcal{S}\beta'} \mathrm{Tr}_\mathcal{E}[C_{\alpha'} \rho(t_0) C_\alpha^\dagger] C_{\mathcal{S}\beta}^\dagger \\ &= C_{\mathcal{S}\beta'} \bar{D}(\alpha', \alpha) C_{\mathcal{S}\beta}^\dagger. \end{aligned} \tag{22}$$

If the original histories α and α' are PT decoherent, then $\bar{D}(\alpha', \alpha)$ vanishes and so the extended histories are PT decoherent also. Then (under the assumption that we could ignore interaction between the system and the environment in the future), we see that two histories which are PT decoherent in the present will continue to be PT decoherent when extended into the future; this is true for any choice of projection in the future. In the case in which ρ represents a pure state, GH decoherence, and so *a fortiori* PT decoherence, implies the existence of records; then if PT decoherence persists, the records will persist also. In a sense, the PT-decoherence condition picks out those GH-decoherent histories which would be expected to continue to decohere in the future.

Of course it is completely unrealistic to expect that a system would not interact with its environment in the future. Rather, we would expect that interaction would continue, but that coherence, once lost to the environment, would never be recovered by the system. However, this last expectation is not guaranteed by the formalism we are using (although the authors cited in [6] might consider that it should be part of the definition of Z decoherence); we have not specified, for example, that the environment be large.

Finally, let us summarize the relationship between the three kinds of decoherence we have discussed, in the special case in which ρ represents a pure state. In this case, GH decoherence implies the existence of "generalized records," i.e., orthogonal states at the end points of each history. However, these records might not be in the environment, where they would be expected to persist. Perhaps a family of histories which satisfied GH decoherence, but not PT decoherence, should be called a family of "orthogonal histories." Z decoherence, on the other hand, does describe a mechanism by which records of the present state of the system will appear in the environment. However, histories are more general than is the present state of a system; for example, two clearly distinguishable histories, with records set in stone out there in the environment, might not be Z decoherent if they happened to lead to the same state of the system. PT decoherence distinguishes those sets of histories for which there are records in the environment.

Consider again the two states $C_\alpha|\Psi\rangle$ and $C_{\alpha'}|\Psi\rangle$. If they are orthogonal, they are GH decoherent; if they are orthogonal in the environment, they are PT decoherent; if they are orthogonal both in the system and in the environment, they are Z decoherent.

Note added. After this work was completed, I received a copy of a paper by Zurek [7] which also discussed the relationship between GH decoherence and Z decoherence.

ACKNOWLEDGMENTS

I have benefited from conversations with Henry Stapp about the program of Gell-Mann and Hartle. I appreciate also the hospitality of the Lawrence Berkeley Laboratory, where this work was done.

[1] M. Gell-Mann and J. B. Hartle, in *Proceedings of the 25th International Conference on High Energy Physics*, Singapore, 1990, edited by K. K. Phua and Y. Yamaguchi (World Scientific, Singapore, 1990).

[2] M. Gell-Mann and J. B. Hartle, Phys. Rev. D **47**, 3345 (1993).

[3] R. B. Griffiths, J. Stat. Phys. **36**, 219 (1984).

[4] R. Omnès, Rev. Mod. Phys. **64**, 339 (1992).

[5] H. F. Dowker and J. J. Halliwell, Phys. Rev. D **46**, 1580 (1992).

[6] H. D. Zeh, Found. Phys. **1**, 69 (1970); W. H. Zurek, Phys. Today **44** (10), 36 (1991), and references cited therein.

[7] W. H. Zurek, in *The Physics of Time Asymmetry*, edited by J. J. Halliwell, J. Perez-Mercader, and W. H. Zurek (Cambridge Univerity Press, Cambridge, England, 1993).