

Classical equations for quantum systems

Murray Gell-Mann*

*Theoretical Astrophysics Group (T-6), Los Alamos National Laboratory, Los Alamos, New Mexico 87545
and Santa Fe Institute, 1660 Old Pecos Trail, Santa Fe, New Mexico 87501*

James B. Hartle

Department of Physics, University of California—Santa Barbara, Santa Barbara, California 93106

(Received 23 October 1992)

The origin of the phenomenological deterministic laws that approximately govern the quasiclassical domain of familiar experience is considered in the context of the quantum mechanics of closed systems such as the universe as a whole. A formulation of quantum mechanics is used that predicts probabilities for the individual members of a set of alternative coarse-grained histories that *decohere*, which means that there is negligible quantum interference between the individual histories in the set. We investigate the requirements for coarse grainings to yield decoherent sets of histories that are quasiclassical, i.e., such that the individual histories obey, with high probability, effective classical equations of motion interrupted continually by small fluctuations and occasionally by large ones. We discuss these requirements generally but study them specifically for coarse grainings of the type that follows a distinguished subset of a complete set of variables while ignoring the rest. More coarse graining is needed to achieve decoherence than would be suggested by naive arguments based on the uncertainty principle. Even coarser graining is required in the distinguished variables for them to have the necessary inertia to approach classical predictability in the presence of the noise consisting of the fluctuations that typical mechanisms of decoherence produce. We describe the derivation of phenomenological equations of motion explicitly for a particular class of models. Those models assume configuration space and a fundamental Lagrangian that is the difference between a kinetic energy quadratic in the velocities and a potential energy. The distinguished variables are taken to be a fixed subset of coordinates of configuration space. The initial density matrix of the closed system is assumed to factor into a product of a density matrix in the distinguished subset and another in the rest of the coordinates. With these restrictions, we improve the derivation from quantum mechanics of the phenomenological equations of motion governing a quasiclassical domain in the following respects: Probabilities of the correlations in time that define equations of motion are explicitly considered. Fully nonlinear cases are studied. Methods are exhibited for finding the form of the phenomenological equations of motion even when these are only distantly related to those of the fundamental action. The demonstration of the connection between quantum-mechanical causality and causality in classical phenomenological equations of motion is generalized. The connections among decoherence, noise, dissipation, and the amount of coarse graining necessary to achieve classical predictability are investigated quantitatively. Routes to removing the restrictions on the models in order to deal with more realistic coarse grainings are described.

PACS number(s): 03.65.Bz, 03.65.Ca, 03.65.Db, 98.80.Bp

I. INTRODUCTION

In a universe governed at a fundamental level by quantum-mechanical laws, characterized by indeterminacy and distributed probabilities, what is the origin of the phenomenological, deterministic laws that approximately govern the quasiclassical domain of everyday experience? What features and limitations of these classical laws can

be traced to their underlying quantum-mechanical origin? This paper addresses such questions in the context of the quantum mechanics of closed systems—most realistically and generally the universe as a whole.

It is a familiar notion in the quantum mechanics of simple “measured” systems that some coarseness in their description is needed if they are to approximate classical behavior. The Heisenberg uncertainty principle, for example, limits the accuracy with which position and momentum can be specified simultaneously. Successions of such suitably crude measurements of position and momentum can be correlated by the classical equations of motion following from the fundamental action of the sys-

*Permanent address: Lauritsen Laboratory, California Institute of Technology, Pasadena, CA 91125.

tem, provided there is a suitable initial state—typically a narrow wave packet.

In discussing the quasiclassical domain of familiar experience, however, we are dealing with a much more general situation than is envisioned by elementary analyses of the above type. We are concerned, first of all, with the classical features that may be exhibited by the behavior of the closed system irrespective of whether those features are receiving attention from “observers.” We should be able to deal, for example, with the classical behavior of the Moon whether or not any “observer” is looking at it. Second, we are necessarily concerned, not just with classical behavior exhibited by correlations among events at a few times of our choosing, but also with the classical behavior of whole orbits in as refined a description of the system as is possible. Third, we are concerned with phenomenological equations of motion, the form of which may be only very indirectly related to that of the fundamental action. The fundamental action, after all, may be that of heterotic superstring theory, while the equations of motion (such as the Navier-Stokes equation) governing the familiar quasiclassical domain involve such quantities as the averages of the densities of field energy and momenta over volumes very much larger than the Planck scale. Further, the applicability of effective classical equations of motion may be branch dependent, that is, contingent on events that have happened. The classical equations governing the motion of the Moon, for example, are contingent on its actually having formed in the early history of the solar system. In such general situations, simply identifying the form of the phenomenological classical equations of motion becomes an important problem.

It is a characteristic feature of the general situations described above that a much coarser graining is needed for quasiclassical behavior would be naively suggested by arguments based on the uncertainty principle. As we shall argue below, a large amount of coarse graining is needed to accomplish decoherence, which is an important ingredient of quasiclassical behavior as well as a sufficient (stronger than necessary) condition for assigning probabilities to the coarse-grained histories of the closed system. Further coarse graining is then necessary to achieve the “inertia” required for approximate predictability in the presence of the noise from the fluctuations that typical mechanisms of decoherence involve. All this coarse graining has important consequences for the form of the effective classical equations of motion. Their form may be as much influenced by the character of the coarse graining and the mechanisms of decoherence as by the fundamental equations of motion. The effective classical equations of motion necessarily include phenomena like dissipation arising from the mechanisms that produce decoherence. This paper is concerned with general methods of deriving the form of the phenomenological classical equations of motion and with the description of the noise that causes deviations from those equations and from classical predictability. As a result of the coarse graining, the noise includes the effects of classical (typically statistical-mechanical) fluctuations as well as quantum fluctuations, and these effects are mixed. The result-

ing indeterminacy, as indicated above, goes far beyond the elementary indeterminacy of the Heisenberg uncertainty principle. An accurate framework for prediction may be achieved by incorporating a generalized Langevin force that represents this noise into the classical equation of motion. Our paper is therefore concerned with the derivation of the general form and distribution of such forces as well as with the equation itself.

It is known, of course, that even in the classical deterministic limit one can encounter in nonlinear systems the phenomenon of chaos, in which the sensitivity of the outcome to the initial conditions is exponential in time. In the presence of chaos, even small fluctuations (including quantum fluctuations) can be amplified to produce large uncertainties in later behavior. A treatment of this combined effect of classical chaos and of indeterminacy arising from quantum mechanics, including the associated coarse graining, has often proved elusive in discussions of quantum chaos, but is amenable to analysis by the methods we shall describe [1].

II. DECOHERENCE, INERTIA, AND EQUATIONS OF MOTION

In this section we give a qualitative discussion of the role of decoherence in the derivation of phenomenological classical equations of motion. This will serve to review some aspects of the quantum mechanics of closed systems, motivate the subsequent mathematical derivations, and make connections with earlier work known to us. For the quantum mechanics of closed systems, we follow our discussion in Refs. [2–4], where references to the earlier literature may be found.

Most generally, quantum mechanics predicts the probabilities of the individual members of a set of alternative, coarse-grained, time histories of a closed system. By a coarse-grained history we mean, for example, one for which not every variable is specified and those that are specified are not fixed at every time or with arbitrary precision. It is evident why coarse-grained histories are of interest to us as observers of the universe. Our observations fix only a tiny fraction of the variables describing the universe and fix those only very imprecisely. As observers we therefore necessarily deal with a very coarse-grained description of the closed system in which we live. However, from the theoretical point of view, it is not necessary that the description be so very coarse-grained or that the coarse graining be so dependent on us. There is a more fundamental reason for interest in sets of coarse-grained alternative histories: In the quantum mechanics of closed systems, probabilities may be assigned only to those sets of histories for which there is negligible quantum-mechanical interference between the individual histories in the set (given the system’s Hamiltonian and initial quantum state) [5,6,2]. We shall define

various kinds of decoherence¹ [3], all of which imply the vanishing of this interference and for all of which coarse graining is necessary. An arbitrarily fine-grained description of the universe would reveal the phase correlations between histories, while in a coarse-grained description they may be absent. The probabilities of decoherent sets of coarse-grained alternative histories constitute the useful predictions in quantum mechanics.

Among the coarse-grained decohering sets of alternative histories of this universe must be the sets that describe the quasiclassical domain of familiar experience. These consist of histories that, for the most part, are defined by ranges of values of “quasiclassical operators” correlated in time with high probability by classical phenomenological laws. We have discussed elsewhere the problem of distinguishing such quasiclassical domains from all other decohering sets of histories and in particular the problem of deriving the form of the “quasiclassical operators” that characterize them [2,3]. Here, we want to focus on a more specific and less general question. That is the question of the derivation and form of the classical phenomenological equations of motion *assuming* a coarse graining is given that leads to decoherence.

A simple form of coarse graining consists of averaging over some variables and following the remaining ones. Let us call the two classes of variables “ignored” and “distinguished,” respectively. Widely occurring mechanisms of decoherence involve the rapid dispersal of

quantum-mechanical phase information among the ignored variables as they interact with the distinguished ones. Thus, for example, sets of histories that distinguish the center-of-mass positions of bodies as light as a grain of dust can be very efficiently “decohered” by the collisions of the bodies with the omnipresent 3 K cosmic background radiation [8]. Such interactions can be expected to produce deviations from the predictability that characterizes classical behavior. That is, they produce noise. For quasiclassical behavior, such that the distinguished variables mostly resist the noise and follow approximately classical predictable paths, a high inertia is required for the distinguished variables. In general, therefore, we expect that a *much* coarser graining is necessary to achieve long stretches of predictable behavior than is needed to achieve mere decoherence, and decoherence, as we remarked, requires much coarser graining than is needed for mere consistency with the uncertainty principle. Furthermore, mechanisms that produce decoherence naturally lead to processes, such as dissipation, that are necessarily included in the equations of motion that describe predictable behavior. There are thus connections among decoherence, noise, dissipation, and the amount of coarse graining necessary to achieve classical predictability. This paper explores those connections.

The habitually decohering quasiclassical operators that characterize our everyday quasiclassical domain include such “hydrodynamic” variables as the averages, over suitable volumes, of densities of energy, momentum, and other conserved or approximately conserved quantities. Such coarse grainings are not of the simple type described above in which the coordinates of a configuration space are separated once and for all into a set “distinguished” by the coarse graining and a set that is “ignored.” First, the coarse grainings corresponding to these averages are not defined by ranges of *coordinates*. In addition, realistic coarse grainings are, in general, branch dependent, meaning in this case, that the volumes over which the averages can be usefully taken are contingent on prior events in specific histories [2]. However, since the simple types of coarse graining are more easily analyzed than the realistic ones, we shall begin our discussion with a class of model problems that are based on distinguished and ignored coordinates and later return to how to generalize our results to more realistic situations.

Our central result is a derivation of the classical equation of motion, including effective forces and noise, for a specific type of coarse graining of the histories of a class of quantum systems. Each system is assumed to have a Lagrangian that is the difference between a kinetic energy quadratic in the velocities and a potential energy independent of velocities but allowed to be fully nonlinear. Coarse grainings are considered that distinguish a fixed subset of the coordinates of the system’s configuration space while ignoring the rest. The initial density matrix of the closed system is assumed to factor into a product of a density matrix in the distinguished variables and another density matrix in the ignored variables. We show that when such sets of coarse-grained histories decohere, the quantum-mechanical probabilities of the in-

¹The term “decoherence” is used in several different ways in the literature. We have used the term to refer to a property of a set of alternative histories of a closed system. Specifically, a set of coarse-grained histories decoheres when there is negligible interference between the individual histories in the set as measured by one of the several conditions discussed in this section. In the literature the term “decoherence” has also been used to refer to the decay in time of the off-diagonal elements of a reduced density matrix defined by a coarse graining of variables at a single moment of time, for example, the density matrix defined by Eq. (4.22). These two notions are not exactly the same. A reduced density matrix can be defined for those special coarse grainings that distinguish a fixed set of coordinates. However, the vanishing of the off-diagonal elements of this reduced density matrix at a succession of times is not identical with the decoherence of the corresponding histories, as will be discussed in Sec. IV. Yet the two notions are not unconnected. In the accessible, although unrealistic, model coarse grainings of the kind studied in this paper, where both notions are defined, typical *mechanisms* of decoherence ensure the validity of both. (See, e.g., Ref. [4], Sec. II.6.4). A general notion of mechanism of decoherence can be defined [7] that generalizes the reduced density-matrix definition of decoherence in the context of the decoherence of histories and characterizes more precisely how they are connected. It would be clearer to use the terms “decoherence of histories” when referring to one notion and “decoherence of density matrices” when referring to the other. In this paper, by “decoherence” we always mean the decoherence of sets of histories as defined precisely in this section.

dividual coarse-grained histories can be represented as the probabilities of the histories of a *classical* system evolving from distributed initial conditions under the action of a stochastic force. The initial conditions follow a Wigner distribution derivable from the initial density matrix. The distribution of the “total force” (including the inertial term “ $-ma$ ”) is a kind of generalization to histories of the Wigner distribution; it is derivable from the decoherence functional that measures quantum coherence. Neither of these distributions is, in general, positive. The effective classical equation of motion is the condition that the expected value of the total force vanish. When the noise arising from the stochastic force is almost negligible, so that the deviations from the effective classical equation of motion are small, then we achieve classical predictability. The noise, whether small or not, can be treated by incorporating a generalized Langevin force into the effective equation of motion. The distribution of the “total force” can also be regarded as a distribution of this Langevin force; it is, in general, non-Gaussian and, as mentioned above, not necessarily positive, as is to be expected from the quantum-mechanical nature of the problem.

Our characterization of the effective classical equation of motion and the Langevin noise distribution leads to a method of identifying the equation of motion and a systematic expansion procedure for calculating the noise distribution. These techniques are not restricted to linear problems or Gaussian noise. The key to the method is that the decoherence functional, which depends on a pair of coarse-grained histories, can be expressed as a path integral over a quantity $\exp(iA)$, where we can expand A in a power series in the difference between the distinguished coordinates in one history and those in the other. For the familiar mechanisms of decoherence, to which we alluded above, the higher-order terms in the expansion are expected to be small, and we can start by retaining only the linear and quadratic terms, which permit us to treat nonlinear equations of motion with Gaussian noise. The higher-order terms then give the non-Gaussian corrections to the distribution of the Langevin force.

As mentioned above, we would like to treat still more general and more realistic problems, in which we escape the limitation to coarse graining that begins with distinguished and ignored coordinate variables, as well as the restriction to factored initial density matrices and the prohibition of velocity-dependent potentials. We discuss at length some ideas of how to free ourselves from these limitations.

The details of the models are described in Sec. III. The various types of decoherence are discussed in Sec. IV. Section V introduces distribution functionals for the “total force” described above and the representation of quantum-mechanical probabilities in terms of them. In Sec. VI the equation of motion and distribution of noise are calculated for the well-known case of linear systems. The explicit generalization to nonlinear cases is given in Sec. VII. The comparison of these results with the corresponding classical analyses is discussed in Sec. VIII. Section IX describes routes to more general coarse grainings

and Sec. X contains some brief conclusions.

There have been, of course, a great many discussions of the derivation of classical behavior from quantum mechanics, and it is perhaps appropriate to offer a few comments on the similarities and differences between the present discussion and that great body of literature. As already mentioned, we aim beyond the elementary discussions of the classical behavior of *measured* systems based on Ehrenfest’s theorem, the WKB approximation, or the Wigner distribution. Such analyses do not usually treat noise, cover the *effective* classical equations of motion including such phenomena as dissipation, consider coarse grainings (other than very obvious ones), or deal seriously with the probabilities of time histories by which roughly predictable quasiclassical behavior is inevitably defined. (Analyses based on the steepest-descent approximation to Feynman’s path integral *do* consider histories but do not typically address the other issues.)

Derivations of the equations of hydrodynamics from statistical physics, as, for example, in Ref. [9,10,11,12], necessarily include phenomena like dissipation. However, those accounts known to us derive equations of motion for the *expectation values* of hydrodynamic variables. In quantum mechanics a system may be said to obey a classical equation of motion when the probability is high for the correlations *in time* that the equation of motion requires. For example, the center of mass of the Earth can be said to obey Newton’s law of motion when the probability is high that successive determinations of the position of the center of mass of the Earth will be correlated according to that law. A complete derivation of classical equations of motion for quantum systems therefore requires the consideration of the probabilities for time histories, not just the study of the evolution of expected values. That is important, because the requirements of decoherence include the restriction on which sets of alternative histories may be assigned probabilities, while there is no such restriction on which expected values may be studied. Furthermore, it is through the study of the probabilities for histories that the probabilities for the inevitable deviations from classical predictability are most directly assessed. That is, only through a study of the probabilities of time histories can we accurately characterize the mixed quantum and classical-statistical noise that characterizes those deviations.

The derivation of classical-statistical equations incorporating both classical determinism and stochastic noise has been extensively discussed for linear systems. Evolution equations for probability distributions on phase space were derived (in certain approximations) from the evolution equation for the Wigner distribution by Caldeira and Leggett [13] and more recently by Hu, Paz, and Zhang [14] and by Żurek [15]. Langevin equations have also been extensively discussed for linear systems. Caldeira and Leggett [13], for example, discuss such equations, using techniques developed by Feynman and Vernon [16], and review earlier efforts. The extensive investigation of linear systems in the quantum optics literature is reviewed in Ref. [17] and treatments from the point of view of statistical mechanics can be found in Refs. [18,12]. None of that work, however, explicitly considers

the probabilities of histories, by which classical behavior is necessarily defined.² Neither is there explicit consideration of the decoherence of histories, which is a prerequisite for the calculation of probabilities in the quantum mechanics of a closed system. By using decoherence and the probabilities of histories we not only recover the standard Langevin equations for linear systems but also we can generalize these results to nonlinear cases.

III. MODEL SYSTEMS AND MODEL COARSE GRAINING

This paper is concerned with classical predictability in a certain class of model quantum-mechanical systems described by particular classes of model coarse grainings. In this section we shall specify these models and coarse grainings and review the theoretical framework in which the decoherence and the probabilities of histories are defined. We shall give a simplified version of this theoretical framework and later discuss how it is related to more general ideas in Sec. IV. That way the reader only interested in the models and not in these general connections can proceed immediately from Sec. III to Sec. V. We shall be brief. For greater detail the reader may consult Refs. [2,3], and the references to the earlier literature found therein.

We consider the quantum mechanics of a closed system in the approximation that gross fluctuations in the geometry of space time may be neglected. A background space time is thus fixed and, in particular, there is a well-defined time t . The usual apparatus of Hilbert space, states, operators, etc. may then be applied in the quantum-mechanical process of prediction. The fundamental dynamics of the system is governed by an action S or its equivalent Hamiltonian H . The initial condition is specified by a density matrix ρ .

The most refined possible description of a closed system makes use of a set of fine-grained histories. The most familiar set of fine-grained histories are the possible paths, $q^\beta(t)$, in a configuration space of generalized coordinates that completely describe the system. These paths are the *single-valued* functions $q^\beta(t)$ on a fixed time interval, for example, $[0, T]$. For example, in a system of scalar fields each $q^\beta(t)$ might be the value of a field at a different spatial point considered as a function of time. (Our description is thus not restricted to nonrelativistic physics.) Configuration-space fine-grained histories are the starting point for sum-over-histories formulations of quantum mechanics and for the model coarse grainings we shall mostly consider in this paper. More general possibilities are discussed in Sec. IV.

A partition of an entire set of fine-grained histories into

exhaustive and exclusive classes defines a set of *coarse-grained histories*; each class is an individual coarse-grained history. The individual coarse-grained histories in an exhaustive set may be grouped into new exclusive sets. That is an operation of further *coarse graining*, yielding a coarser-grained set of alternative histories. The inverse operation of dividing a set of coarse-grained histories into smaller classes of the fine-grained histories is an operation called *fine graining*. Sets of coarse-grained histories are partially ordered with respect to the operations of fine and coarse graining. A rich variety of coarse grainings is possible. As we mentioned before, the histories of the quasiclassical domain of everyday experience, for example, are defined by coarse grainings utilizing ranges of values of averages over suitable spatial regions of such "hydrodynamic" variables as the densities of energy, momentum, charges, and currents. Such realistic coarse grainings are, in general, branch dependent, that is, contingent on which of many possible events have happened.

Because of their branch dependence, and because of their indirect relation to the fundamental fields, realistic coarse grainings are not as theoretically tractable as some model coarse grainings that can be studied. In this paper we shall study a familiar and instructive class of model coarse grainings in which the coordinates q^β of configuration space are divided into ones x^a that are distinguished by the coarse graining and the remaining ones Q^A , which are ignored. (For example, in a simplified model of a universe of particles, the x^a might label the center-of-mass positions and orientations of a group of massive bodies such as the planets and the Q^A would then be all the rest of the coordinates, including the internal coordinates of the bodies' constituents and the coordinates of gas molecules, etc. that interact with the planets.) Coarse-grained histories of this type are labeled by partitions of the paths $x^a(t)$. An individual coarse-grained history consists of a path $x^a(t)$ along with *all* possible paths $Q^A(t)$.

Further coarse graining of the classes of histories labeled by the paths $x^a(t)$ can be defined by sets of intervals exhausting the whole range of the x^a at a discrete sequence of times $t_1 < t_2 < \dots < t_n$. These correspond to intervals on the entire configuration space of q^β that are unrestricted in the Q^A but are subject to the stated restrictions on the x^a . We denote the exhaustive sets of such intervals at the successive times by $\{\Delta_{\alpha_1}^1\}$, $\{\Delta_{\alpha_2}^2\}, \dots, \{\Delta_{\alpha_n}^n\}$. The index k on $\{\Delta_{\alpha_k}^k\}$ labels the particular set, α_k labels the particular interval in the set, and t_k is the time. An individual coarse-grained history in such a set consists of the paths, $q^\beta(t)$, that thread a *particular* sequence of regions, e.g., $\Delta_{\alpha_1}^1$ at t_1 , $\Delta_{\alpha_2}^2$ at t_2 , etc., and the whole set of coarse-grained histories is exhausted as the different possible ways which paths pass through the regions are enumerated. A particular coarse-grained history thus corresponds to a sequence $(\alpha_1, \dots, \alpha_n)$, which we shall often abbreviate as just α .

The coherence between individual histories in a coarse-grained set is measured by the decoherence functional. This is a complex functional defined on all pairs

²Recently, Dowker and Halliwell [19] have studied the decoherence of histories in explicit linear models and also, in effect, derived, their classical equations of motion, although not including a description of dissipation and noise.

of coarse-grained histories in an exhaustive set of such histories. The sets of histories under discussion are coarse grainings of fine-grained histories that are paths $q^\beta(t)$ on the time interval $[0, T]$. The sum-over-histories formulation of quantum mechanics is, therefore, convenient for introducing the decoherence functional; more general formulations are discussed in the next section. For the coarse grainings under discussion, the decoherence functional is

$$D(\alpha', \alpha) = \int_{\alpha'} \delta q' \int_{\alpha} \delta q \delta(q'_f - q_f) \times \exp(i\{S[q'(\tau)] - S[q(\tau)]\} / \hbar) \rho(q'_0, q_0). \quad (3.1)$$

The path integral over $q^\beta(t)$ is over all paths that start at q_0^β at $t=0$, pass through the intervals $\Delta_{\alpha_1}^1, \Delta_{\alpha_2}^2, \dots, \Delta_{\alpha_n}^n$ of the x^a at times $t_1 < \dots < t_n$, and wind up at q_f^β at time T . The path integral over $q'^\beta(t)$ is similarly defined. The integrals include an integral over the initial q_0^β and final q_f^β . Here, $\rho(q'_0, q_0)$ is the initial density matrix in the q^β representation and $S[q(\tau)]$ is the fundamental action. The measure for the path integrals is the standard one in-

duced by the Liouville measure in phase space. It is described explicitly in Sec. A of the Appendix. Equation (3.1) has been compressed by omitting the indices on the $q^\beta(t)$ and by denoting the entire sequence $(\alpha_1, \dots, \alpha_n)$ at (t_1, \dots, t_n) by a single index α . We shall employ similar conventions in the rest of the paper.

The coarse grainings under consideration distinguish only the x^a in the division $q^\beta = (x^a, Q^A)$. Following an analysis of Feynman and Vernon [16], the integrals over the Q^A may therefore be carried out over their whole ranges unrestricted by the particular coarse-grained histories considered. Suppose that the action may be decomposed as

$$S[q(\tau)] = S_{\text{free}}[x(\tau)] + S_0[Q(\tau)] + S_{\text{int}}[x(\tau), Q(\tau)]. \quad (3.2)$$

(The use of the subscript “free” does not mean that there is no potential-energy term in S_{free} . There is, in general. Rather it means that the action of the x 's is free of any interaction with the Q 's.) The integral over the Q 's defines \mathcal{W} , a functional of the paths $x'(t)$ and $x(t)$ and a function of their initial end points x'_0 and x_0 , as

$$\begin{aligned} & \exp\{i\mathcal{W}[x'(\tau), x(\tau); x'_0, x_0] / \hbar\} \bar{\rho}(x'_0, x_0) \\ & \equiv \int \delta Q' \int \delta Q \delta(Q'_f - Q_f) \exp(i\{S_0[Q'(\tau)] + S_{\text{int}}[x'(\tau), Q'(\tau)] - S_0[Q(\tau)] - S_{\text{int}}[x(\tau), Q(\tau)]\} / \hbar) \rho(x'_0, Q'_0; x_0, Q_0). \end{aligned} \quad (3.3)$$

Here we have introduced the reduced density matrix $\bar{\rho} = \text{Sp } \rho$ associated with the coarse graining:

$$\bar{\rho}(x'_0, x_0) \equiv \int dQ_0 \rho(x'_0, Q_0; x_0, Q_0). \quad (3.4)$$

The functional \mathcal{W} is only a slight generalization of the Feynman-Vernon influence phase and we shall continue to call it that. It depends on the end points x'_0 and x_0 *implicitly* through the paths $x'(t)$ and $x(t)$ because the actions on the right-hand side of Eq. (3.3) are functionals of these paths. There is also an *explicit* dependence on x'_0 and x_0 arising from the dependence of ρ on these variables in Eqs. (3.3) and (3.4). We use the notation $\mathcal{W}[x'(\tau), x(\tau); x'_0, x_0]$ to indicate this dependence, the square bracket to indicate dependence on functions and the parentheses to indicate dependence on variables, and we maintain this notation for other cases. The quantity \mathcal{W} depends on the time interval T as well, but we have not indicated this explicitly.

The decoherence functional may then be expressed in terms of \mathcal{W} and $\bar{\rho}$ as

$$D(\alpha', \alpha) = \int_{\alpha'} \delta x' \int_{\alpha} \delta x \delta(x'_f - x_f) \exp(i\{S_{\text{free}}[x'(\tau)] - S_{\text{free}}[x(\tau)] + \mathcal{W}[x'(\tau), x(\tau); x'_0, x_0]\} / \hbar) \bar{\rho}(x'_0, x_0). \quad (3.5)$$

Thus all the contribution from the ignored variables is summarized by the functional $\mathcal{W}[x'(t), x(t); x'_0, x_0]$.

Restrictions on the form of the actions $S_{\text{free}}[x]$, $S_0[Q]$, and $S_{\text{int}}[x, Q]$, as well as on the form of the density matrix ρ , will be needed for the explicit derivation of the equation of motion, as described in the subsequent sections. We shall, for example, generally make the usual assumption that $S_{\text{free}}[x(\tau)]$ has a simple “kinetic minus potential” form

$$S_{\text{free}}[x(\tau)] = \int_0^T dt \left[\frac{1}{2} \dot{x}^\dagger(t) M \dot{x}(t) - V(x(t)) \right], \quad (3.6)$$

where we have used an obvious matrix notation $x^\dagger M x = \sum_{ab} x^a M_{ab} x^b$. Similar assumptions will be made for $S_0[Q(\tau)]$. We shall assume that $S_{\text{int}}[x(\tau), Q(\tau)]$ is local in time, that is, of the form

$$S_{\text{int}}[x(\tau), Q(\tau)] = \int_0^T dt L_{\text{int}}[x(t), Q(t)], \quad (3.7)$$

with L_{int} independent of velocities. These assumptions restrict us only to a widely applicable class of models and it is likely that similar results can be obtained from weaker assumptions.

A more restrictive assumption concerns the form of the initial ρ . We shall make the conventional assumption

[13,16] that ρ factors into a density matrix in x and another in Q ,

$$\rho(x'_0, Q'_0; x_0, Q_0) = \bar{\rho}(x'_0, x_0) \rho_B(Q'_0, Q_0), \quad (3.8)$$

so that the variables distinguished by the coarse graining are initially uncorrelated with those it ignores. Of course, we do not necessarily expect initial density matrix of the whole universe to factor as in Eq. (3.8), but, for widespread mechanisms of decoherence that operate essentially locally in space and time when compared with cosmological scales, Eq. (3.8) is an excellent approximation. For example, scattering by the cosmic background radiation can efficiently decohere alternative positions of the center of mass of a massive body coarse grained on centimeter scales [8]. The coordinates of the body and radiation may be correlated in the wave function of the universe, but on the local scales where the mechanism operates they are effectively uncorrelated, as described by Eq. (3.8).

Factorization has a number of helpful consequences for the form of W and $\bar{\rho}$ defined in Eqs. (3.3) and (3.4). First, $\bar{\rho}$ is given by

$$\bar{\rho}(x'_0, x_0) = \bar{\rho}(x'_0, x_0). \quad (3.9)$$

Second, and most importantly, the influence phase W contains no explicit dependence on x'_0 and x_0 and we may write $W[x'(\tau), x(\tau)]$. This will simplify the form of the equation of motion, which would otherwise contain terms arising from the explicit dependence on x'_0 and x_0 .

For most of this paper, therefore, we are considering a class of models defined by coarse grainings that distinguish a fixed subset of the coordinates of configuration space, by actions that have the simple forms (3.2), (3.6), and (3.7), and by an initial density matrix that factors as in Eq. (3.8). Especially simple examples of such models are the linear oscillator models studied by Feynman and Vernon [16] and by Caldeira and Leggett [13]. In these models a distinguished oscillator is coupled linearly to a large number of other oscillators constituting a thermal bath characterized by a temperature T_B . The density matrix is assumed to factor as in Eq. (3.8). Let $x(t)$ be the coordinate of the distinguished oscillator and ω_R its frequency renormalized by its interactions with the others. Then, in the case of a continuum of oscillators, cut off at frequency Ω , and in the Fokker-Planck limit of $kT_B \gg \hbar\Omega \gg \hbar\omega_R$, Caldeira and Leggett find, for the influence phase,

$$\begin{aligned} W[x'(\tau), x(\tau)] = & -M\gamma \int_0^\tau dt (x'\dot{x}' - x\dot{x} + x'\dot{x} - x\dot{x}') \\ & + i \frac{2M\gamma kT_B}{\hbar} \int_0^\tau dt [x'(t) - x(t)]^2, \end{aligned} \quad (3.10)$$

where γ is a coupling constant summarizing the interaction of the distinguished oscillator with the rest.

IV. DECOHERENCE

A. Decoherence in general

Quantum mechanics predicts the probabilities for the individual members of a set of alternative coarse-grained histories only when there is negligible quantum-mechanical interference between the individual members of the set [5,6,2]. Only then do the squares of amplitudes define probabilities that are consistent with the sum rules of probability theory. Sets of histories that exhibit negligible interference are said to *decohere weakly* [3].

However, in quantum mechanics we are not interested just in sets of histories that are consistent in the sense that they can be assigned probabilities satisfying probability sum rules. We are interested also in sets of histories that constitute the quasiclassical domain of everyday experience. It is this quasiclassical domain that lies at the root of the interpretation of quantum mechanics. It is through an understanding of this domain that quantum mechanics acquires utility for our experience; "measurement" situations arise precisely when variables become highly correlated with the quasiclassical domain. Stronger notions of decoherence are therefore useful to characterize the realistic mechanisms of decoherence that lead to a quasiclassical domain [3]. In this section we shall review several notions of decoherence that we have described in previous work, and we shall discuss their connections with each other and with some other notions of decoherence that have been introduced in the literature. We shall be brief. For greater detail the reader may consult Refs. [2,3].

We begin by recalling how a set of alternative, coarse-grained histories of a closed system is described generally in quantum mechanics. The simplest kinds of histories are specified by giving independent sets of alternatives at a sequence of times $t_1 < t_2 < \dots < t_n$. In the Heisenberg picture, alternatives at one moment of time t_k correspond to a set of projection operators $\{P_{\alpha_k}^k(t_k)\}$. The index k denotes the set of alternatives at time t_k , while the index α_k denotes the particular alternative within that set. These projections represent exclusive alternatives, so they are orthogonal for different alternatives, and they represent an exhaustive set, so they sum to unity over all alternatives. For example, for the coarse graining by ranges of a distinguished set of variables x^a described in Sec. III, the projection $P_{\alpha_k}^k(t_k)$ would just be the projection onto the range $\Delta_{\alpha_k}^k$ at time t_k . An individual history in a set defined by a sequence of such sets of alternatives corresponds to a sequence of particular alternatives $\alpha = (\alpha_1, \dots, \alpha_n)$. Each history is represented by the corresponding chain of projection operators:

$$C_\alpha = P_{\alpha_n}^n(t_n) \cdots P_{\alpha_1}^1(t_1). \quad (4.1)$$

A completely fine-grained set of histories would consist of one-dimensional projections onto complete sets of states at each and every time. Sets of histories defined by sets of projections that are not all one-dimensional or not at every moment of time are said to be coarse-grained.

In the Heisenberg picture, every exhaustive set of or-

thogonal projection operators in Hilbert space $\{P_{\alpha_k}^k\}$ represents, at *any* time, *some* set of alternatives for the system. Of course, the alternatives corresponding to a given set of projection operators will have different descriptions in terms of fundamental fields when different times are assigned to them. Similarly a sequence of sets of projection operators define different alternative histories when different times are assigned to the sets of projections. However, an assignment of time leads to meaningful alternative histories only if the ordering of the times corresponds to ordering of the projections as in Eq. (4.1). If this time ordering is not respected, two inconsistent sets of alternatives could be assigned the same time and the resulting alternatives would not be meaningful.

While in quantum mechanics we usually consider sets of histories consisting of independent sets of alternatives at sequences of times, a more realistic description is achieved by generalizing this notion in two related ways [2,4]. First, if $\alpha=(\alpha_1, \dots, \alpha_n)$ is a history and $\beta=(\beta_1, \dots, \beta_n)$ is a distinct history then we may consider the coarser-grained alternative that the system followed either history α or history β . The alternative α or β is represented by the *sum* of the chains for α and β ,

$$C_{\alpha \text{ or } \beta} = C_{\alpha} + C_{\beta}, \quad (4.2)$$

but is not itself necessarily a chain of projections of the form (4.1). More precisely, if $\{\alpha\}$ is a set of alternative histories for the closed system defined by sets of alternative projections at a sequence of times, then coarser-grained sets are defined by partitions of the $\{\alpha\}$ into exclusive classes $\{\bar{\alpha}\}$. The classes correspond to the individual histories in these coarser-grained sets and are represented by operators $C_{\bar{\alpha}}$ that are sums of the C_{α} :

$$C_{\bar{\alpha}} = \sum_{\alpha \in \bar{\alpha}} C_{\alpha} = \sum_{(\alpha_1, \dots, \alpha_n) \in \bar{\alpha}} P_{\alpha_n}^n(t_n) \cdots P_{\alpha_1}^1(t_1). \quad (4.3)$$

These generalized $C_{\bar{\alpha}}$ need not themselves be chains of projections and thus we sometimes extend the use of C_{α} to denote the operators representing individual histories in these more general coarse-grained sets.

The second important generalization is to allow the histories to be branch dependent, that is, for the set of alternatives at time t_k to depend on the values of earlier labels $\alpha_1, \dots, \alpha_{k-1}$ [2,6]. Branch dependence is important, because in a quasiclassical domain past events may determine what is a suitable quasiclassical variable. For example, if a quantum fluctuation gets amplified so that it leads to condensation of a galaxy in one branch and no such condensation in other branches, then the outcome clearly influences what are suitable quasiclassical variables in the region where the galaxy would form.

For the general case of branch dependence, a better notation than Eq. (4.1) for chains of projections would be

the following:

$$C_{\alpha} = P_{\alpha_n}^n(t_n; \alpha_{n-1}, \dots, \alpha_1) \\ \times P_{\alpha_{n-1}}^{n-1}(t_{n-1}; \alpha_{n-2}, \dots, \alpha_1) \cdots P_{\alpha_1}^1(t_1), \quad (4.4)$$

where the $P_{\alpha_k}^k(t_k; \alpha_{k-1}, \dots, \alpha_1)$ define an exhaustive set of orthogonal projection operators as α_k runs over all values for fixed $\alpha_1, \dots, \alpha_{k-1}$, corresponding to an exhaustive set of mutually exclusive alternatives for the closed system.³ In limiting ourselves to projections that depend only on previous alternatives rather than future ones we have incorporated a notion of causality consistent with the usual arrow of time in quantum mechanics. Further generalizations to formulations without an arrow of time are possible [20].

In the models that we treat in detail in this article, we do not make use of branch dependence, since we assume a fixed division of coordinates into those distinguished x^a and those ignored Q^A , and the Δ 's (intervals of x values) are taken to be branch independent. We can thus employ the simplified notation (4.1). It should be borne in mind, though, that realistic coarse-grained histories do involve branch dependence.

Having in hand this discussion of the possible sets of alternative coarse-grained histories of a closed system, we can now turn to the various notions of their decoherence. The central quantity, the decoherence functional, is defined generally for pairs of histories in a coarse-grained set by

$$D(\alpha', \alpha) = \text{Tr}(C_{\alpha'} \rho C_{\alpha}^{\dagger}) \quad (4.5)$$

for a density matrix ρ representing the initial condition and operators C_{α} representing the individual histories.

The necessary and sufficient condition for probability sum rules to be satisfied is

$$\text{Re } D(\alpha', \alpha) = 0, \quad \alpha' \neq \alpha. \quad (4.6)$$

In previous work [3] we have called this the *weak*

³In Ref. [2], Sec. X, we unnecessarily eliminated the possibility of branch-dependent chains of the form (4.4) and restricted attention to sets of α_k 's that were independent of one another. We did that in order to safeguard a special derivation of the weak decoherence condition for assigning probabilities to alternative coarse-grained histories. In fact, that special derivation is unnecessary, and as we shall see below weak decoherence can easily be seen to be the necessary and sufficient condition for the probability calculus to apply to histories that are chains or sums of chains, whether or not the choice of the set $P_{\alpha_k}^k$ is branch dependent.

decoherence condition⁴ to distinguish it from stronger decoherence conditions we shall discuss below. When a set of histories weakly decoheres, the probability of a history α is the corresponding “diagonal” element of the decoherence functional

$$p(\alpha) = D(\alpha, \alpha). \quad (4.7)$$

Equation (4.6) is the necessary and sufficient condition that these numbers obey the sum rules of probability theory. All that is needed to show this is to notice that the probability that either of two histories will happen is the sum of the probabilities of the two individual histories if and only if the sum of the interference terms represented by Eq. (4.6) vanishes. Weak decoherence is the criterion by which quantum mechanics discriminates between those sets of histories that can be assigned probabilities and those that cannot.

A stronger notion of decoherence is provided by the *medium decoherence* condition [3]

$$D(\alpha', \alpha) = 0, \quad \alpha' \neq \alpha. \quad (4.8)$$

Clearly medium decoherence implies weak decoherence, but not the other way around. Medium decoherence is a consequence of realistic mechanisms that are widespread in the universe. It is, therefore, a natural condition to impose in characterizing a quasiclassical domain. It is not, by itself, sufficient to single out a quasiclassical domain. The necessary further criteria are a large part of the sub-

ject of this paper.

When the initial state is pure, exact medium decoherence is equivalent to the existence of generalized records for each history in the decohering set [3]. To see this, notice that, for a pure initial condition $\rho = |\Psi\rangle\langle\Psi|$,

$$D(\alpha', \alpha) = (\langle\Psi|C_\alpha^\dagger)(C_{\alpha'}|\Psi\rangle). \quad (4.9)$$

To every set of alternative histories there corresponds a resolution of the pure initial state into branches:

$$|\Psi\rangle = \sum_\alpha (C_\alpha|\Psi\rangle) = \sum_{\alpha_1, \dots, \alpha_n} P_{\alpha_n}^n(t_n) \cdots P_{\alpha_1}^1(t_1)|\Psi\rangle. \quad (4.10)$$

For a set of histories obeying exact medium decoherence, the branches are orthogonal, as Eqs. (4.8) and (4.9) show. Therefore, there exists at least one set of orthogonal projection operators $\{R_\alpha\}$ that project onto these branches

$$R_\alpha|\Psi\rangle = C_\alpha|\Psi\rangle, \quad (4.11)$$

where

$$R_\alpha R_{\alpha'} = \delta_{\alpha\alpha'} R_\alpha \quad \text{and} \quad \sum_\alpha R_\alpha = I, \quad (4.12)$$

so that the projections $\{R_\alpha\}$ are exclusive and exhaustive.

Now we can also assign projection operators R_α to coarse grainings of the histories $\{\alpha\}$, that is, sums of chains of projections $\{C_\alpha\}$. To every such coarser-grained history, representing the union of a subset of the histories $\{\alpha\}$ or the sum of the corresponding $\{C_\alpha\}$, we assign the projection operator that is the sum of the corresponding $\{R_\alpha\}$. That is perfectly consistent with the extension of Eq. (4.8) to the coarser-grained histories. The resulting R 's have the property that progressive fine graining of the coarser-grained histories results in a sequence of *nested* R 's, projecting onto smaller and smaller subspaces of Hilbert space, where each such space is a subspace of the preceding one.

When the branches $C_\alpha|\Psi\rangle \neq 0$ do not form a *complete* set of orthogonal states for the Hilbert space, there can be many sets of projections $\{R_\alpha\}$ that obey the conditions (4.11) and (4.12). When the branches do form a complete set, then the $\{R_\alpha\}$ are unique; they are just the projections onto the single states $C_\alpha|\Psi\rangle$. The set of histories $\{\alpha\}$ is then said to be *full* [3]. The R 's for coarse grainings of the histories $\{\alpha\}$ are then also unique: for the further coarse-grained history that corresponds to the union of a given subset of the $\{\alpha\}$, the corresponding R is just the sum of the relevant R_α and projects onto the space spanned by the corresponding vectors $C_\alpha|\Psi\rangle = R_\alpha|\Psi\rangle$.

When assigned a time after t_n in the sequence $t_1 < t < \cdots < t_n$, the R_α 's may be thought of as representing generalized records of the histories. They may not represent records in the usual sense of being constructed from quasiclassical variables accessible to us, but the condition (4.11) means that at any time there is complete information somewhere in the universe about the histories $\{\alpha\}$.

⁴As pointed out to us by Bob Griffiths, we have incorrectly attributed the weak decoherence condition (4.6) to him (in Ref. [2] and elsewhere). In fact, Griffiths [5] and Omnès [6] employ a weaker condition than Eq. (4.6) as the necessary and sufficient condition for the “consistency of histories.” That is because they require, in the notation of Eq. (4.1), that two chains of projections C_α and $C_{\alpha'}$ must have the real part of their interference term vanish *only if* $C_\alpha + C_{\alpha'}$ is another chain of projections, whereas we require it in the case of *all* the chains C_α and $C_{\alpha'}$ (see below). Our weak decoherence applies to coarse grainings that are allowed to unite any two histories in the set being studied, while the “consistent histories” condition of Griffiths and Omnès applies only to some of those coarse grainings. For example, if α and α' differ in only one index, our approach and theirs give the same conditions, but if α and α' differ in more than one index, then Griffiths and Omnès do not always require that $\text{Re}[\text{Tr}(C_{\alpha'}\rho C_\alpha^\dagger)] = 0$ but rather a weaker condition. In essence Griffiths and Omnès restrict themselves to histories defined by independent alternatives at a sequence of times. Each history corresponds to a sequence of such alternatives and therefore to a chain of projections. They do not therefore incorporate branch dependence, at least in the sense of Eq. (4.4). In the models studied in this paper, the stronger conditions are, in fact, satisfied. The whole of any off-diagonal element of the decoherence functional approximately vanishes—not just the real part. Also, when more of the indices in α and α' differ, the decoherence condition (4.8) is satisfied more strongly, not more weakly. This gives us some confidence that our stronger conditions are physically realistic.

The above discussion shows that medium decoherence in a pure initial state implies the existence of generalized records. The converse is also true. The existence of orthogonal generalized record projections satisfying Eq. (4.11) ensures the medium decoherence of the corresponding set of histories through Eqs. (4.8) and (4.9). Exact medium decoherence can thus be characterized by records, and the physical formation of records is a way to understand mechanisms by which medium decoherence occurs. In the example implicit in the work of Joos and Zeh⁵ [8], histories describing successive alternative positions of a dust grain, initially in a superposition of positions about a millimeter apart, very accurately satisfy the condition of medium decoherence simply by virtue of the scattering by cosmic background radiation photons. The successive scatterings of these photons effectively create records of the histories of positions of the dust grain in the electromagnetic degrees of freedom. The commuting records of successive positions are stored independently in the vastness of cosmological space as the photons move off at the speed of light. They may not be accessible to us, but their existence is a way of understanding how this mechanism of medium decoherence works.

The permanence of the past is a feature of the quasi-classical domain that is naturally explained by medium decoherence when there is a pure initial condition. By permanence of the past we mean the feature of a quasi-classical domain that what has happened in the past is independent of any information expressed by a future projection. Neither the decoherence of past alternatives nor the selection of a particular past alternative is threatened by new information. “The Moving Finger writes; and having writ, Moves on: nor all thy Piety nor Wit Shall lure it back to cancel half a Line, Nor all thy Tears wash out a Word of it” [21].

In other words, we are discussing the property of the decohering coarse-grained histories $\{\alpha\}$ that, at any of the times t_k , there is, for each history up to that time, an effective density matrix [4]

$$\frac{P_{\alpha_k}^k(t_k) \cdots P_{\alpha_1}^1(t_1) \rho P_{\alpha_1}^1(t_1) \cdots P_{\alpha_k}^k(t_k)}{\text{Tr}[P_{\alpha_k}^k(t_k) \cdots P_{\alpha_1}^1(t_1) \rho P_{\alpha_1}^1(t_1) \cdots P_{\alpha_k}^k(t_k)]}, \quad (4.13)$$

which can be utilized for all future predictions without concern about the other values of $\alpha_1, \dots, \alpha_k$. For a pure state, this corresponds to the “reduction of the state vector”:

$$|\Psi\rangle \rightarrow N^{-1/2} P_{\alpha_k}^k(t_k) \cdots P_{\alpha_1}^1(t_1) |\Psi\rangle, \quad (4.14)$$

where N is the trace in Eq. (4.13).

When the past is permanent, we may still lose the ability to retrodict the probabilities of alternatives in the past through the impermanence or inaccuracy of present records *but not from the failure of those past alternatives to decohere in the face of the projections that describe information we acquire as we advance into the future*. Yet

we know that such continued decoherence of the past is not guaranteed *in general* by quantum mechanics.⁶ Adjoining future alternatives to a set of histories is a fine graining of that set and in general a fine graining of a decoherent set of histories may no longer decohere. Verifying the continued decoherence of all the past alternatives as we fine grain our set of histories to deal with the future would in general require significant computation. We would have to check that the branches corresponding to *every* alternative past that *might* have happened continue to be orthogonal in the presence of their newly adjoined sets of projections. Yet we adjoin sets of projections onto ranges of quasiclassical operators without making this calculation, secure in the faith that previous alternatives will continue to decohere despite this fine graining. It is this assumption of continued decoherence of the past that permits the focus for future predictions on the one branch corresponding to our particular history and the discarding of all others. In other words, we pointed out above, it is the permanence of the past that permits the “reduction of the state vector.”

If we consider, instead of the set of histories $\{\alpha\} = \{(\alpha_1, \dots, \alpha_n)\}$, the set of abbreviated histories $\{(\alpha_1, \dots, \alpha_k)\}$ with $k < n$, running up to time $t_k < t_n$, that is an example of a coarser graining of the set of coarse-grained histories $\{\alpha\}$. As these abbreviated histories get further α 's (α_{k+1} then α_{k+2} , etc.) adjoined to them, that represents a sequence of fine grainings of the coarser-grained histories $\{(\alpha_1, \dots, \alpha_k)\}$.

If ρ is pure, there are nested records R corresponding to these abbreviated histories. When further α 's are adjoined, as the histories unfold from $(\alpha_1, \dots, \alpha_k)$ to $(\alpha_1, \dots, \alpha_{k+1})$, etc., up to $(\alpha_1, \dots, \alpha_n)$, the nested record operators are projections onto subspaces of Hilbert space that progressively narrow. In such a situation, the past always continues to decohere as the histories advance into the future. In fact, the physical formation of nested, generalized record operators R guarantees, in general, the permanence of the past, including not only the permanence of its decoherence, but also the permanence of the selection of particular past alternatives as well.

If we relax the condition of exact decoherence and consider approximate medium decoherence, defined by the approximate validity of Eq. (4.8), then it is possible to understand more about the formation of records and the origins of medium decoherence for the kind of model coarse grainings studied in this paper, which distinguish particular coordinates and ignore a large number of others. The Hilbert space is a tensor product of a Hilbert space of functions of the distinguished coordinates, \mathcal{H}^x , and one for the ignored coordinates, \mathcal{H}^Q . The coarse-grained histories consist of chains of projections at times t_1, t_2, \dots that, in the Schrödinger picture, act only on \mathcal{H}^x and represent partitions of some complete set of states in \mathcal{H}^x at each time. In the models, these states are, in fact,

⁵For a more detailed analysis in terms of histories, see Ref. [4].

⁶For further and less informal discussion see Ref. [4], Sec. II.3.2.

just the localized states in x^a at each time, but for greater generality we shall consider different complete orthogonal sets at each time represented by wave functions $\{\phi_{r_1}^1(x)\}$ at time t_1 , $\{\phi_{r_2}^2(x)\}$ at time t_2 , etc. In the subsequent discussion we shall augment this notation to indicate the branch dependence of the possible orthogonal sets.

In order to make a connection with the thinking of some authors, we shall take a brief excursion into the Schrödinger picture, while assuming a pure initial state represented by a wave function $\Psi(x, Q, t_0)$. Let us follow for three steps the Schrödinger evolution of this initial Ψ . It may be evolved to the time t_1 of the first set of alternatives and expanded in the first complete set $\{\phi_{r_1}^1(x)\}$ as

$$\Psi(x, Q, t_0) \rightarrow \sum_{r_1} \phi_{r_1}^1(x) \chi_{r_1}(Q). \quad (4.15)$$

The right-hand side is $\Psi(x, Q, t_1)$ in the Schrödinger picture and $\langle x(t_1) = x, Q(t_1) = Q | \Psi \rangle$ in the Heisenberg picture. The coefficients χ_{r_1} are in general neither normalized nor orthogonal for different r_1 . A coarse graining divides the r_1 up into exclusive and exhaustive sets $\{\alpha_1\}$. The sum on the right of Eq. (4.15) may be similarly decomposed:

$$\Psi(x, Q, t_1) = \sum_{\alpha_1} \sum_{r_1 \in \alpha_1} \phi_{r_1}^1(x) \chi_{r_1}(Q). \quad (4.16)$$

The result of the evolution of each branch $\sum_{r_1 \in \alpha_1} \phi_{r_1}^1(x) \chi_{r_1}(Q)$ in Eq. (4.16) to the time t_2 of the next set of alternatives may again be expanded in a new complete set of functions in the x 's. In general, this set will be branch dependent, that is dependent on α_1 . We use the notation $\{\phi_{r_2}^{2\alpha_1}(x)\}$ to indicate this, and the expansion then has the form

$$\sum_{r_1 \in \alpha_1} \phi_{r_1}^1(x) \chi_{r_1}(Q) \rightarrow \sum_{\alpha_2} \sum_{r_2 \in \alpha_2} \phi_{r_2}^{2\alpha_1}(x) \chi_{r_2}^{\alpha_1}(Q). \quad (4.17)$$

Similarly, for the evolution from t_2 to t_3 ,

$$\sum_{r_2 \in \alpha_2} \phi_{r_2}^{2\alpha_1}(x) \chi_{r_2}^{\alpha_1}(Q) \rightarrow \sum_{\alpha_3} \sum_{r_3 \in \alpha_3} \phi_{r_3}^{3\alpha_2\alpha_1}(x) \chi_{r_3}^{\alpha_2\alpha_1}(Q) \quad (4.18)$$

and so forth. Thus,

$$\Psi(x, Q, t_3) = \sum_{\alpha_3\alpha_2\alpha_1} \left[\sum_{r_3 \in \alpha_3} \phi_{r_3}^{3\alpha_2\alpha_1}(x) \chi_{r_3}^{\alpha_2\alpha_1}(Q) \right]. \quad (4.19)$$

The term in the parentheses is the branch wave function corresponding to the history $\alpha = (\alpha_1, \alpha_2, \alpha_3)$. The overlap of branches gives the decoherence functional according to Eqs. (4.9) and (4.10). Thus

$$D(\alpha', \alpha) = \sum_{r'_3 \in \alpha'_3} \sum_{r_3 \in \alpha_3} (\phi_{r'_3}^{3\alpha'_2\alpha'_1}, \phi_{r_3}^{3\alpha_2\alpha_1}) (\chi_{r'_3}^{\alpha'_2\alpha'_1}, \chi_{r_3}^{\alpha_2\alpha_1}). \quad (4.20)$$

If the complete set at time t_3 is independent of previous alternatives, so that it is not branch dependent, then the decoherence of the last alternative is automatic. The sca-

lar product $(\phi_{r'_3}^{3\alpha'_2\alpha'_1}, \phi_{r_3}^{3\alpha_2\alpha_1}) = \delta_{r'_3 r_3}$ and so the right-hand side of Eq. (4.20) is diagonal in r_3 . Otherwise it is a nontrivial condition because $\{\phi_{r_3}^{3\alpha_2\alpha_1}\}$ and $\{\phi_{r'_3}^{3\alpha'_2\alpha'_1}\}$ are *different* orthogonal sets when α 's do not coincide. Suppose the Hilbert space \mathcal{H}^Q is effectively very large compared to \mathcal{H}^x , in the sense that at each time t_k a one-to-one correspondence is established between \mathcal{H}^x and a tiny portion of \mathcal{H}^Q . Then we may expect the scalar products between χ 's differing by any index to be typically very small,⁷ leading to small values of $D(\alpha, \alpha')$ for $\alpha' \neq \alpha$. The summations over unorganized phases between different values of r_3 may make the off-diagonal elements of $D(\alpha', \alpha)$ even smaller. That is approximate medium decoherence. In the approximation in which *all* the χ 's are really orthogonal to one another, the records R_α that accomplish projections onto branch wave functions such as those in Eq. (4.19) may then be taken to be projections onto the corresponding sets of χ 's: $\{\chi_{r_3}^{\alpha_2\alpha_1} | r_3 \in \alpha_3\}$.

The reduced density matrix in x may also be constructed. It is

$$\begin{aligned} \bar{\rho}(x', x) = & \sum_{\alpha'_3\alpha'_2\alpha'_1} \sum_{\alpha_3\alpha_2\alpha_1} \sum_{r'_3 \in \alpha'_3} \sum_{r_3 \in \alpha_3} \phi_{r'_3}^{3\alpha'_2\alpha'_1}(x') \phi_{r_3}^{3\alpha_2\alpha_1*}(x) \\ & \times (\chi_{r'_3}^{\alpha'_2\alpha'_1}, \chi_{r_3}^{\alpha_2\alpha_1}), \quad (4.21) \end{aligned}$$

involving a sum over all branches. If indeed all the χ 's are approximately orthogonal to one another because of the largeness of the Hilbert space \mathcal{H}^Q , then $\bar{\rho}(x', x)$ is approximately diagonal in the histories $\{\alpha\}$.

Many authors⁸ have considered, not the full density matrix (4.21), but the reduced effective density matrix for each branch [cf. (4.13)], relevant when the decoherence of earlier alternatives may be assumed. The applicable portion of $\bar{\rho}(x', x)$ [Eq. (4.21)] is

$$\begin{aligned} \rho^{\alpha_2\alpha_1}(x', x) = & \sum_{\alpha_3} \sum_{\alpha'_3} \sum_{r'_3 \in \alpha'_3} \sum_{r_3 \in \alpha_3} \phi_{r'_3}^{3\alpha_2\alpha_1}(x') \phi_{r_3}^{3\alpha_2\alpha_1*}(x) \\ & \times (\chi_{r'_3}^{\alpha_2\alpha_1}, \chi_{r_3}^{\alpha_2\alpha_1}). \quad (4.22) \end{aligned}$$

Approximate orthogonality of these χ 's for different α_3 leads to approximate diagonality of $\rho^{\alpha_2\alpha_1}$ in the α_3 's. However, a much stronger condition on the χ 's is needed to ensure the decoherence of whole histories as in Eq. (4.20) than is needed to ensure the diagonality of Eq. (4.22). We need the approximate orthogonality of the $\{\chi_{r_3}^{\alpha_2\alpha_1}\}$ not only when the r_3 belong to different sets α_3 ,

⁷Of course, as more and more times are added to the histories, and the set of α 's grows larger accordingly, we expect that eventually even the space \mathcal{H}^Q will be exhausted and χ 's no longer orthogonal.

⁸For example, Zeh [22] and Żurek [23] in their efforts to find a class of ϕ 's (e.g., Żurek's "pointer basis") that would, in part, characterize a quasiclassical domain.

but also between functions corresponding to different values of α_1 and α_2 . The strength of these conditions may be appreciated by noting two facts: First, since any coarse graining of a decoherent set is also decoherent, the χ 's must be approximately orthogonal in all previous steps. That is, the $\{\chi_{r_2}^{\alpha_1}\}$ in the example must be approximately orthogonal for different α_1 's and when the r_2 lie in different α_2 's. There is a similar condition for the $\{\chi_{r_1}\}$. Since x and Q interact between the times t_1, \dots, t_n none of these conditions is a simple consequence of the others. Second, nothing in our discussion has fixed the choice of the times t_1, \dots, t_n . Of course, they must be sufficiently separated for the interactions to disperse the phases, but once that is satisfied, we expect decoherence to hold for a range of times giving rise effectively to even more conditions. The decoherence of histories is a much stronger requirement than the diagonality of density matrices.

Some authors [24–26] have discussed how, in the case of a *pure* initial density matrix $\Psi(x', Q')\Psi^*(x, Q)$, one can use the Schmidt decomposition to achieve some of the objectives of decoherence. In the notation we have been using for the Schrödinger picture in Eqs. (4.15)–(4.19), we can choose, at each time t_k , the functions $\phi_{r_k}^{k\alpha_{k-1}\dots\alpha_1}(x)$ to be Schmidt functions, which means that their coefficients $\chi_{r_k}^{\alpha_{k-1}\dots\alpha_1}(Q)$ are *orthogonal* for different values of r_k (although still not normalized). In that way the reduced density matrix for each branch at time t_k becomes diagonal in r_k , as we can see for the case $k=3$ in Eq. (4.22). It is then, of course, also diagonal in α_k (α_3 for the case $k=3$). However, this Schmidt construction does not guarantee the decoherence of even the final alternative as defined by the decoherence functional (4.20) (unless we have $\alpha_1=\alpha'_1$ and $\alpha_2=\alpha'_2$). That is because the Schmidt functions are *necessarily branch-dependent* and the orthogonality of the χ 's for one branch does not guarantee the orthogonality between different branches that would lead to decoherence of the final alternative. In any event, the main thing for the decoherence of *histories* is to have $D(\alpha', \alpha)$ diagonal in α_2 and α_1 as well. That is not guaranteed at all by the Schmidt procedure, which does not imply the orthogonality of χ functions for different values of the α 's. In particular, the Schmidt procedure does not guarantee the permanence of the past discussed above. The relation between diagonalization of the reduced density matrix at successive times and the decoherence of histories will be further discussed in [7].

Now let us consider the situation with an impure initial state represented by a density matrix ρ . An impure initial state could be fundamental, representing the initial condition of the universe. However, even if the cosmological initial condition is pure, a reduced density matrix, in which some coordinates already have been traced over, can be a useful description of local physics. A familiar example is the cosmic background radiation. Imagine that the background photons have all been pair produced in a pure cosmological initial state. In that case, for every photon near us, there would be a correlated photon with equal but opposite momentum on the other side of

the universe. The local physics, however, would be accurately described by a nearly thermal reduced density matrix in which the distant photons has been traced over.

In Ref. [3] we discussed a “strong decoherence” condition that is the analogue of (4.11) for density matrices. A set of histories was said to decohere strongly when, for each history in the set, there exists record projections satisfying Eq. (4.12) such that

$$C_\alpha \rho = R_\alpha \rho . \quad (4.23)$$

However, we shall see that for highly impure states this is too strong a condition to usefully characterize a quasi-classical domain.

We can express an impure density matrix ρ in terms of its eigenstates and eigenvalues:

$$\rho = \sum_\mu \pi_\mu |\Psi^\mu\rangle \langle \Psi^\mu| , \quad (4.24a)$$

or, in the special case of x and Q variables,

$$\rho(x', Q'; x, Q) = \sum_\mu \pi_\mu \Psi^\mu(x', Q') \Psi^{*\mu}(x, Q) . \quad (4.24b)$$

Here the π_μ are the probabilities of the initial states $|\Psi^\mu\rangle$.

If ρ has nonzero probabilities for many states, then Eq. (4.23) is difficult to satisfy. If ρ has nonzero probabilities for an orthogonal set of states $|\Psi^\mu\rangle$, then Eq. (4.23) would imply

$$C_\alpha |\Psi^\mu\rangle = R_\alpha |\Psi^\mu\rangle \quad \text{for all } |\Psi^\mu\rangle . \quad (4.25)$$

If the set $|\Psi^\mu\rangle$ is *complete*, then Eq. (4.25) implies that $C_\alpha = R_\alpha$. That can be satisfied only in the trivial case in which all the P 's for all times commute with one another. We shall henceforth ignore “strong decoherence” for impure density matrices.

We could define, when ρ is impure, a kind of “medium strong decoherence”, in which we would have, for each $|\Psi^\mu\rangle$ with nonzero probability π_μ , a generalized record projection operator R_α^μ such that

$$C_\alpha |\Psi^\mu\rangle = R_\alpha^\mu |\Psi^\mu\rangle , \quad (4.26)$$

where for *each* μ the R_α^μ are exclusive and exhaustive projections. (When the eigenvalues π_μ are degenerate, so that the $|\Psi^\mu\rangle$ are not uniquely determined by ρ , we require (4.26) for at least one set of $|\Psi^\mu\rangle$ that diagonalizes ρ .) This requirement would mean medium decoherence separately for each $|\Psi^\mu\rangle$ (with $\pi_\mu \neq 0$) with respect to the same set of histories C_α . While not so difficult to satisfy as strong decoherence, it is still a very stiff requirement. For example, we shall see, in the linear oscillator models discussed below, that it is not very well satisfied there. However, if satisfied, medium strong decoherence would supply, in the case of an impure ρ , the same attractive features that medium decoherence yielded for the pure case, including the permanence of the past.

Approximate medium decoherence for density matrices continues to be defined by the approximate satisfaction of the condition (4.8) and can be discussed for the type of model coarse grainings considered in this paper, where a fixed set of coordinates is distinguished and the other coordinates included in the model are ignored. The mechanism of the formation of correlations between dis-

tistinguished coordinates and ignored ones can continue to operator for each state in the density matrix. Now, however, there is the possibility for improvement in the effectiveness of approximate decoherence from summations over the states in the impure density matrix.

To make this explicit, consider the Schrödinger evolution of an initial density matrix of the form (4.24b). The evolution described by Eqs. (4.15)–(4.19) is as before except that each χ acquires an index μ . In particular Eqs. (4.20) and (4.21) become

$$D(\alpha', \alpha) = \sum_{r_3 \in \alpha_3} \sum_{r'_3 \in \alpha_3} (\phi_{r_3}^{3\alpha_2\alpha_1}, \phi_{r'_3}^{3\alpha'_2\alpha'_1}) \sum_{\mu} \pi_{\mu}(\chi_{r_3}^{\mu\alpha_2\alpha_1}, \chi_{r'_3}^{\mu\alpha'_2\alpha'_1}), \quad (4.27)$$

$$\begin{aligned} \bar{\rho}(x', x) = & \sum_{\alpha'_1\alpha'_2\alpha'_3} \sum_{\alpha_1\alpha_2\alpha_3} \sum_{r'_3 \in \alpha'_3} \sum_{r_3 \in \alpha_3} \phi_{r'_3}^{3\alpha'_2\alpha'_1}(x') \phi_{r_3}^{3\alpha_2\alpha_1}(x) \\ & \times \sum_{\mu} \pi_{\mu}(\chi_{r_3}^{\mu\alpha_2\alpha_1}, \chi_{r'_3}^{\mu\alpha'_2\alpha'_1}). \end{aligned} \quad (4.28)$$

The additional sum over μ can lead to further phase cancellations and more effective decoherence. We will see an illustration of this in the discussion of the oscillator models discussed below.

B. Decoherence in the linear oscillator models

Medium decoherence can be treated explicitly in the kind of oscillator models worked out by Feynman and Vernon [16] and Caldeira and Leggett [13], which we described in Sec. III. Generally these models assume an impure initial density matrix factored as in Eq. (3.8), with the ignored coordinates representing a continuum of oscillators in a thermal state.

An explicit illustration of the medium decoherence of a pure state can be found in the zero-temperature limit of the linear oscillator model. In that limit all the ignored oscillators are in a pure state and the initial state of the distinguished oscillators may be taken to be pure. Denote the initial ground state of the bath by $\chi_0(Q_0)$ and the initial state of the distinguished oscillators by $\psi_0(x_0)$, so that the pure initial wave function of the whole system is $\psi_0(x_0)\chi_0(Q_0)$.

The Schrödinger evolution of this state can be represented in path integral form. At time t we have

$$\begin{aligned} \Psi(x, Q, t) = & \int \delta x \delta Q \exp\{iS[x(\tau), Q(\tau)]/\hbar\} \\ & \times \psi_0(x_0)\chi_0(Q_0). \end{aligned} \quad (4.29)$$

The integration is performed first over paths that start at x_0 and Q_0 at time t_0 and end at x and Q at time t , and then over x_0 and Q_0 . The integral over the distinguished coordinates x may be written as a sum over coarse grainings $\alpha_1, \dots, \alpha_k$, $t_k < t$, and an integral over the paths restricted by the coarse graining, in the form

$$\begin{aligned} \Psi(x, Q, t) = & \sum_{\alpha_1, \dots, \alpha_k} \int_{\alpha} \delta x \delta Q \exp\{iS[x(\tau), Q(\tau)]/\hbar\} \\ & \times \psi_0(x_0)\chi_0(Q_0). \end{aligned} \quad (4.30)$$

This can be rewritten in the form

$$\Psi(x, Q, t) = \sum_{\alpha_1, \dots, \alpha_k} \int_{-\infty}^{+\infty} dr \delta(x-r) \chi^{\alpha_k \dots \alpha_1}(r, Q), \quad (4.31)$$

where

$$\begin{aligned} \chi^{\alpha_k \dots \alpha_1}(r, Q) = & \int_{\alpha} \delta x \delta Q \exp\{iS[x(\tau), Q(\tau)]/\hbar\} \\ & \times \psi_0(x_0)\chi_0(Q_0) \end{aligned} \quad (4.32)$$

and the integral is over paths, consistent with the coarse graining, that start at (x_0, Q_0) and end at (r, Q) , including an integral over the values of x_0 and Q_0 . This is evidently the analogue of Eq. (4.19) with r being a continuous index and $\phi_r^k(x) = \delta(x-r)$, branch independent and the same at each time. The overlap that occurs in Eq. (4.20), of course, gives the decoherence functional.

In the zero-temperature limit, Caldeira and Leggett find, for the imaginary part of the influence phase [cf. Eq. (6.8)],

$$\text{Im } W[x'(\tau), x(\tau)] = \frac{1}{4} \int_0^T dt \int_0^T dt' \xi(t) k_I(t-t') \xi(t'). \quad (4.33)$$

Here, $\xi(t) = x'(t) - x(t)$ and

$$k_I(\tau) = \frac{4M\gamma}{\pi} \int_0^{\Omega} d\omega \omega \cos(\omega\tau), \quad (4.34)$$

where Ω is the cutoff of the oscillator spectrum. This imaginary part of the influence phase favors contributions to the functional integral defining the decoherence functional (3.5) from values of ξ near zero, or $x'(t)$ close to $x(t)$.

A crude estimate of the time intervals by which coarse-grained alternatives of position $\{\Delta_{\alpha_k}^k\}$ must be spaced in order to ensure decoherence may be obtained as follows: Approximate $\xi(t)$ by a constant value d that is characteristic of the sizes of the intervals $\{\Delta_{\alpha_k}^k\}$. Find the time interval $t_{\text{decoherence}}$ such that the integral in Eq. (4.33) evaluated over that time interval is of order of magnitude unity. One then has a rough estimate of a time interval long enough for $\text{Im } W/\hbar$ to be large enough for $\exp(-\text{Im } W/\hbar)$ to be small. The answer is

$$t_{\text{decoherence}} \sim \frac{1}{\Omega} \exp\left[\frac{\hbar}{M\gamma d^2}\right], \quad (4.35)$$

assuming $\Omega t_{\text{decoherence}} \gg 1$. Note that $t_{\text{decoherence}}$ decreases as the coupling γ is made stronger or the graining is made coarser or the number of oscillators becomes larger. In the present case of a pure initial state, the time scale for decoherence is essentially set by the cutoff, for fixed coupling and coarse graining.

The improvement in decoherence from the sum over

states in an impure density matrix is well illustrated in the linear oscillator models. In the high-temperature Fokker-Planck limit, where many states contribute with nearly equal probability to the density matrix ρ_B , the imaginary part of the influence phase is given by [cf. Eq. (3.10)]

$$\text{Im } W[x'(\tau), x(\tau)] = \frac{2M\gamma kT_B}{\hbar} \int_0^\tau dt [x'(t) - x(t)]^2. \quad (4.36)$$

The large value of $\text{Im } W$ suppresses contributions to the functional integral (3.5) defining the decoherence functional when $x'(t)$ is significantly different from $x(t)$, provided enough time has elapsed between successive alternatives so that significant values of $\text{Im } W/\hbar$ are built up,

$$\int_{\alpha'} \delta x' \int_{\alpha} \delta x M(x'_{k+1}, x_{k+1}) \exp(i\{S_{\text{free}}[x'(\tau)] - S_{\text{free}}[x(\tau)] + W[x'(\tau), x(\tau)]\}/\hbar) N(x'_k, x_k), \quad (4.37)$$

where the path integrals, as well as the integrals defining the actions and the influence phase, are over the interval of time from the time t_k of a set of alternatives $\{\alpha_k\}$ to the time of the next set t_{k+1} . The values x'_k, x_k are the end points at time t_k and x'_{k+1}, x_{k+1} are the end points at time t_{k+1} . We consider any functions M and N . The expression (4.37) for the imaginary part of the influence phase implies that, if the intervals $\{\Delta_{\alpha_k}^k\}$ by which the paths $x(t)$ are coarse grained have a characteristic size d and the interval between t_k and t_{k+1} is larger than the characteristic decoherence time scale [27],

$$t_{\text{decoherence}} \sim \frac{1}{\gamma} \left[\frac{\hbar}{\sqrt{2MkT_B}} \frac{1}{d} \right]^2, \quad (4.38)$$

then the ‘‘off-diagonal’’ terms in Eq. (4.8) will be very small and approximate decoherence will be achieved for the alternatives $\{\alpha_k\}$ for a large class of functions M and N .

Consider a set of coarse-grained histories defined by regions $\{\Delta_{\alpha_k}^k\}$ at times t_1, \dots, t_n separated by time intervals longer than the characteristic decoherence time (4.38). Fine-graining this set by adjoining *further* sets of intervals at similarly spaced times greater than t_n does not affect the decoherence of those already present because the mechanism of decoherence exhibited by (4.37) is operative over a time scale of the order of $t_{\text{decoherence}}$ about the time of each set of alternatives. Physically, that is reasonable. In this model, phases are carried away by interactions, local in time, of the distinguished variables x^a with the rest. Once dispersed among the continuum of oscillators described by the Q^A , they cannot be recovered by finer grainings beyond t_n that involve the x^a

yielding approximate medium decoherence. Further, in this limit of very large T_B such that $kT_B \gg \hbar\Omega$, there is more efficient medium decoherence than is provided in the pure ground-state example of (4.35).

Another feature that can be illustrated in the linear oscillator models is the permanence of the past. We have seen how the successive narrowing of the records implied by medium decoherence in the case of a pure state gives a natural explanation of the permanence of the past. For an impure state there is no satisfactory corresponding notion of generalized record, but in the oscillator models it is still possible to show how the past becomes permanent.

History is approximately permanent for a suitably restricted class of coarse grainings in the oscillator model in the Fokker-Planck approximation. The reason is that decoherence there is essentially local in time. More precisely, consider the integral

alone. To recover the phases, one would need a much finer graining that involved the whole set of variables.

V. DISTRIBUTIONS FOR DECOHERENCE FUNCTIONALS

As the review presented in the preceding two sections makes clear, quantum theory can be organized into two parts: First, there is the calculus of amplitudes for histories or the bilinear combinations that are the decoherence functionals. The rules of this calculus derive ultimately from the principle of superposition. Second, there are the rules for deriving probabilities from these amplitudes, most generally, the several notions of decoherence of sets of alternative histories of a closed system. In this section, we show how the first part, the calculus of decoherence functionals for histories, may be usefully reexpressed in terms of distribution functionals analogous to the Wigner distribution for alternatives at a single moment of time.

In familiar quantum mechanics the probability that a determination of coordinates x^a at one moment of time will yield a result in a volume V of the reduced configuration space spanned by these coordinates is

$$p(V) = \int_V dx \bar{\rho}(x, x), \quad (5.1)$$

where $\bar{\rho}(x', x)$ is the reduced density matrix on the reduced configuration space [cf. Eq. (3.4)] and dx is the reduced volume element. As is well known, the density matrix $\bar{\rho}(x', x)$ that gives such probabilities may be usefully expressed in terms of the Wigner distribution on phase space

$$\bar{\rho} \left[X + \frac{\xi}{2}, X - \frac{\xi}{2} \right] = \int dP w(X, P) e^{i\xi^\dagger P/\hbar}. \quad (5.2)$$

The probability $p(V)$ is then given by

$$p(V) = \int_V dX \int dP w(X, P). \quad (5.3)$$

It is also true that the probability density for the momentum P conjugate to x is given by $\int dX w(X, P)$. In these properties $w(X, P)$ is like a classical distribution on phase space. Other ways in which it is similar have been extensively discussed (see, e.g., Refs. [28,29]). It differs from a classical distribution in that it is not in general positive and it does not provide analogues for all probabilities that are defined on classical phase space. For example, it would be incorrect to think of $w(X, P)$ itself as a probability for a simultaneous determination of position *and* momentum. The calculus of amplitudes must therefore be supplemented by the rules that specify which classical quantities on phase space can be assigned probabilities. Ignoring those rules, and considering just the calculus of amplitudes for alternatives at a single time, we see from Eqs. (5.2) and (5.3) how the calculus may be reformulated in terms of a distribution function on phase space, although not, in general, in terms of a positive one. In the following we shall give an analogous formulation for the calculus of amplitudes for time histories.

The decoherence functional defined in Eq. (3.1) is the bilinear combination of amplitudes of which the diagonal elements give the probabilities of the individual histories in a decoherent set [cf. Eq. (4.7)]. In this sense, it plays the role for histories that the reduced density matrix does for alternatives at a single moment of time. We now construct a distribution functional for the decoherence function in much the same way in which the Wigner distribution was constructed from the reduced density matrix.

We consider the partially coarse-grained decoherence functional $D[x'(\tau), x(\tau)]$ for histories that are fine-grained in the variables x distinguished by the initial coarse graining discussed in the previous section. Introduce variables $X(t)$ and $\xi(t)$ that are the average and difference respectively of the arguments of the decoherence functional:

$$X(t) = \frac{1}{2}[x'(\tau) + x(\tau)], \quad (5.4a)$$

$$\xi(t) = x'(\tau) - x(\tau). \quad (5.4b)$$

In terms of these variables the fine-grained decoherence functional defined by Eq. (3.5) may be written

$$D[X(\tau), \xi(\tau); X_0, \xi_0] = \delta(\xi_f) \exp \left[\frac{i}{\hbar} A[X(\tau), \xi(\tau); X_0, \xi_0] \right] \times \bar{\rho} \left[X_0 + \frac{\xi_0}{2}, X_0 - \frac{\xi_0}{2} \right], \quad (5.5)$$

where

$$\begin{aligned} A[X(\tau), \xi(\tau); X_0, \xi_0] \\ = S_{\text{free}}[X(\tau) + \xi(\tau)/2] - S_{\text{free}}[X(\tau) - \xi(\tau)/2] \\ + W[X(\tau), \xi(\tau); X_0, \xi_0]. \end{aligned} \quad (5.6)$$

A distribution $G[R(\tau), X(\tau), X_0, \xi_0]$ may be introduced for the decoherence functional by taking its functional Fourier transform with respect to $\xi(t)$. We define G by the formula

$$D[X(\tau), \xi(\tau); X_0, \xi_0] = \int \delta R \exp \left[\frac{i}{\hbar} \int_0^T dt \xi^\dagger(t) R(t) \right] \times G[R(\tau), X(\tau); X_0, \xi_0]. \quad (5.7)$$

The functional G may be calculated from the decoherence functional by an inverse functional Fourier transform. Expressions such as Eq. (5.7) are to be interpreted as limits of multiple integrals over paths that are piecewise linear between a discrete set of time slices $\tau_0=0, \tau_1, \dots, \tau_N=T$ as the number of slices, N , tends to infinity. This is a standard way of defining path integrals; details for the particular integrals of interest are discussed more fully in Sec. C of the Appendix. As an aid to the present discussion, however, it is useful to note that we always represent the integral such as that in the exponent of Eq. (5.7) by discrete sums of the form

$$\int_0^T d\tau \xi^\dagger(\tau) R(\tau) = \sum_{k=1}^N \epsilon \xi_k^\dagger R_k, \quad (5.8)$$

where ϵ is the separation between time slices. The variable ξ_0 is thus not transformed and appears on both sides of the equation.

The transform variable $R(t)$ has the dimension of force. As we shall see later, $G[X(\tau), R(\tau); X_0, \xi_0]$ may be regarded as a classical distribution of $R(t)$, given the path $X(\tau)$ and X_0 and ξ_0 , although, in general, a nonpositive one. It is therefore instructive to consider its moments. We define

$$\begin{aligned} \langle R(t_1) \cdots R(t_n) \rangle_c \\ = \frac{\int \delta R R(t_1) \cdots R(t_n) G[R(\tau), X(\tau); X_0, \xi_0]}{\int \delta R G[R(\tau), X(\tau); X_0, \xi_0]}. \end{aligned} \quad (5.9)$$

Clearly, $\langle R(t_1) \cdots R(t_n) \rangle_c$ is a functional of the path $X(\tau)$ and a function of X_0 and ξ_0 .

The first moment we *define* to be the average "total force":

$$\mathcal{E}(t, X_0, \xi_0; X(\tau)) \equiv \langle R(t) \rangle_c. \quad (5.10)$$

By "total force" we mean the force minus the inertial term, so that $\mathcal{E}=0$ is the effective or phenomenological

classical equation of motion on the average. The deviations of the force $R(t)$ from its expected value define the Langevin force $\mathcal{L}[t, X_0, \xi_0; X(\tau)]$ for a given path $X(\tau)$:

$$\mathcal{L}(t, X_0, \xi_0; X(\tau)) \equiv R(t) - \mathcal{E}(t, X_0, \xi_0; X(\tau)). \quad (5.11)$$

The reason for these designations will become clear when we express the quantum-mechanical probability for an individual history α in a *decoherent* set of histories in terms of the distribution G .

The fully coarse-grained decoherence functional (3.5) is given by

$$D(\alpha', \alpha) = \int_{\alpha'} \delta x' \int_{\alpha} \delta x D[X(\tau), \xi(\tau); X_0, \xi_0], \quad (5.12)$$

where X and ξ are connected to x' and x by Eq. (5.4) and their range is restricted by the coarse graining through Eq. (5.4). The partially coarse-grained decoherence functional is given by Eqs. (5.5) and (5.6). If the further graining defined by the successive sets of regions is coarse enough so that sufficient positive imaginary part of W is built up between one set of intervals and the next, then there will be a significant contribution to the integral defining the decoherence functional only for values of $\xi(t)$ near zero and for $\alpha = \alpha'$. (See Fig. 1.) That is medium decoherence. Further, in the diagonal elements of the decoherence functional, which are the probabilities of the individual coarse-grained histories, the integral over $\xi(t)$ may be carried out, to an excellent approximation, as though unrestricted by the coarse graining, provided the intervals are sufficiently coarse. (See Fig. 1.) When reexpressed in terms of the distribution G , the probabilities for the individual histories $p(\alpha)$ are, in this approximation,

$$p(\alpha) = \int_{\alpha} \delta X \int \delta R \int \delta \xi G[R(\tau), X(\tau); X_0, \xi_0] \times \exp \left[\frac{i}{\hbar} \int_0^T dt \xi^\dagger(t) R(t) \right], \quad (5.13)$$

where the integrals over $\xi(t)$, including that over ξ_0 , are unconstrained.

The expression (5.13) for the probabilities of the individual histories in a coarse-grained set has an especially transparent interpretation in the special case where the initial density matrix factors as in Eq. (3.8) and a simple kinetic- minus potential energy form (3.6) is assumed for the action $S_{\text{free}}[x(\tau)]$. A simplification following from factorization is that the influence phase W has no direct dependence on x'_0 and x_0 , and hence on X_0 and ξ_0 , except through the paths $x'(\tau)$ and $x(\tau)$. Thus, we write simply $W[x'(\tau), x(\tau)]$. More accurately, in a time slicing implementation of the functional integral in which integrals over the paths are discretized as in Eq. (5.8), W is independent of X_0 and ξ_0 . Further discussion is given in Sec. B of the Appendix. When the action $S_{\text{free}}[x(\tau)]$ has a simple kinetic minus potential form, the dependence of

the rest of the exponent $A[X(\tau), \xi(\tau); X_0, \xi_0]$ may be easily isolated. An integration by parts in Eq. (5.6) yields

$$A[X(\tau), \xi(\tau); X_0, \xi_0] = -\xi_0^\dagger M \dot{X}_0 + \tilde{A}[X(\tau), \xi(\tau)], \quad (5.14)$$

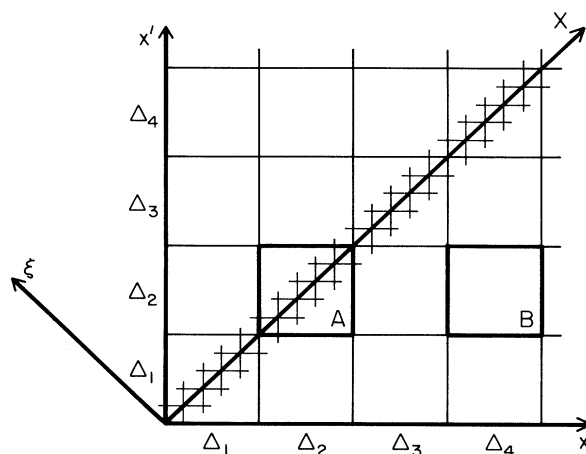


FIG. 1. The decoherence of histories coarse grained by intervals of a distinguished set of configuration-space coordinates. The decoherence functional for such sets of histories is defined by the double path integral of Eq. (3.5) over paths $x'(t)$ and $x(t)$ that are restricted by the coarse graining. These path integrals may be thought of [cf. Eq. (A14)] as the limits of multiple integrals over the values of x' and x on a series of discrete time slices of the interval $[0, T]$. A typical slice at a time when the range of integration is constrained by the coarse graining is illustrated. Of course, only one of the distinguished coordinates x^a and its corresponding x'^a can be shown and we have assumed for illustrative purposes that the regions defining the coarse graining correspond to a set of intervals Δ_α , $\alpha = 1, 2, 3, \dots$ of this coordinate. On each slice where there is a restriction from the coarse graining, the integration over x' and x will be restricted to a single box. For the "off-diagonal" elements of the decoherence functional corresponding to distinct histories, that box will be off the diagonal (e.g., B) for some slice. For the diagonal elements, corresponding to the same histories, the box will be on the diagonal (e.g., A) for all slices. If the imaginary part of the influence phase $W[x'(t), x(t)]$ grows as a functional of the difference $\xi(t) = x'(t) - x(t)$, as it does in the oscillator models [cf. Eq. (3.10)], then integrand of the decoherence functional will be negligible except when $x'(t)$ is close to $x(t)$ —a regime illustrated by the shaded band about the diagonal in the figure. When the characteristic sizes of the intervals Δ_α are large compared to the width of the band in which the integrand is nonzero, the off-diagonal elements of the decoherence functional will be negligible because integrals over those slices where the histories are distinct is negligible (e.g., over box B). That is decoherence of the coarse-grained set of histories. Further, the evaluation of the diagonal elements of the decoherence functional that give the probabilities of the individual histories in decoherent set can be simplified. If the integrations over x' and x are transformed to integrations over $\xi = x' - x$ and $X = (x' + x)/2$ the restrictions on the range of the ξ integration to one diagonal box may be neglected with negligible error to the probability.

where

$$\begin{aligned} \tilde{A}[X(\tau), \xi(\tau)] &= \int_0^T dt \left[-\xi^\dagger(t) M \ddot{X}(t) - V \left[X(t) + \frac{\xi(t)}{2} \right] \right. \\ &\quad \left. + V \left[X(t) - \frac{\xi(t)}{2} \right] \right] + W[X(\tau), \xi(\tau)]. \end{aligned} \quad (5.15)$$

Then, under the ground rules for discretizing functionals discussed above, the functional \tilde{A} is independent of both ξ_0 and X_0 . In particular, the only terms in the fine-grained decoherence functional (5.5) that depend explicitly on ξ_0 are the density matrix and the surface term in Eq. (5.14). The integral over ξ_0 may therefore be carried out in Eq. (5.13), giving

$$\begin{aligned} p(\alpha) &= \int_\alpha \delta X \int \delta R \int \delta \xi g[R(\tau), X(\tau)] w(X_0, M\dot{X}_0) \\ &\quad \times \exp \left[\frac{i}{\hbar} \int_0^T dt \xi^\dagger(t) R(t) \right], \end{aligned} \quad (5.16)$$

where w is the Wigner function defined by Eq. (5.2) and the distribution $g[R(\tau), X(\tau)]$ is defined by

$$\begin{aligned} \delta(\xi_f) \exp\{ (i/\hbar) \tilde{A}[X(\tau), \xi(\tau)] \} \\ = \int \delta R g[R(\tau), X(\tau)] \exp \left[\frac{i}{\hbar} \int_0^T dt \xi^\dagger(t) R(t) \right]. \end{aligned} \quad (5.17)$$

These expressions acquire precise meaning in the time slicing implementation of the path integrals discussed in Sec. B of the Appendix. The important point for the present discussion is that the assumptions that the initial ρ factorizes and that S_{free} has a simple kinetic- minus potential-energy form lead to a factorization of the general distribution G into a distribution w of initial values X_0 and $M\dot{X}_0$ and a distribution g involving forces $R(t)$ and paths $X(t)$. Further, the equation of motion $\mathcal{E}(t, X(\tau))$ and the Langevin force $\mathcal{L}(t, X(\tau))$ become independent of the initial conditions and calculable just from the distribution g of force $R(t)$ through expressions of the same form as Eqs. (5.9)–(5.11) with G replaced by g . These features allow a simple interpretation, which we shortly describe.

The Wigner distribution by which the initial conditions are distributed in the expression for the probabilities (5.16) is not generally positive and neither is the distribution g . However, the probabilities $p(\alpha)$ must be positive. It is not difficult to see how this comes about. Were the restriction on the range of integration arising from the coarse graining restored in expressions (5.13) or (5.16), the numbers $p(\alpha)$ they define would be manifestly positive. That is because they are expressions for a diagonal element of a decoherence functional which is always posi-

tive [cf. Eq. (4.5)]. This is not unlike other smearings of the Wigner distribution [30] which are known to give generally positive results. Of course, the approximation in which the restrictions of the coarse graining on the integration are ignored may result in small negative probabilities, but, to the extent the approximation is good, these are equivalent to zero for physical purposes.

The representation (5.16) allows the probabilities of decohering coarse-grained histories $p(\alpha)$ to be thought of as the probabilities of the histories of a system moving classically under the action of a stochastic force. The Wigner function gives the distribution of initial conditions. The distribution $g[R(\tau), X(\tau)]$ may be thought of as the distribution of total force $R(t)$ acting on a system that describes the path $X(t)$. Alternatively, if reexpressed in terms of $\mathcal{L}(t, X(\tau))$ and $\mathcal{E}(t, X(\tau))$ defined by Eqs. (5.11) and (5.10), the quantity

$$g[\mathcal{E}(t, X(\tau)) + \mathcal{L}(t, X(\tau)); X(\tau)]$$

may be thought of as the distribution of $\mathcal{L}(t)$ given the path $X(t)$. In the approximation we have discussed, the unconstrained integration over $\xi(t)$ in Eq. (5.16) leads to a functional δ function that enforces the condition, $R(t)=0$, that the “total force” on the system vanish. That is, it enforces the effective classical equation of motion, corrected by the Langevin force:

$$\mathcal{E}(t, X(\tau)) + \mathcal{L}(t, X(\tau)) = 0. \quad (5.18)$$

In this equation the average total force $\mathcal{E}(t, X(\tau))$ is the known functional of $X(\tau)$ defined by Eq. (5.10). The force $\mathcal{L}(t, X(\tau))$ is distributed according to the distribution g . It is for this reason that we have called \mathcal{L} the Langevin force; it can be thought of as noise.

The probabilities $p(\alpha)$ for decohering coarse-grained histories are thus obtained from a mathematical description of classical dynamical system characterized by an equation of motion $\mathcal{E}(t, X(\tau))$ together with distributed initial conditions and distributed noise. We should stress that this does not mean that quantum mechanics is equivalent to some kind of classical physics. For one thing, the distributions of noise and initial conditions are not generally positive and that is certainly a nonclassical feature. For another thing, there is the coarse graining needed for decoherence—an entirely nonclassical requirement that must be satisfied before the $p(\alpha)$ may be considered as the probabilities of histories.

The nonpositivity that distinguishes quantum-mechanical distributions from classical ones may be regarded as the reason that Bell’s classical inequalities [31] are violated in quantum mechanics, leading to important experimental tests of the theory. Of course, the Wigner distribution considered here is for continuous variables while Bell’s discussion of the EPRB (Einstein-Podolsky-Rosen-Bohm) problem was for a discrete spin system. However, Feynman [32] showed that there is an analogue of the Wigner distribution for spin- $\frac{1}{2}$ systems and that the departure of quantum mechanics from Bell’s inequality is traceable to the fact that the analogous distribution is not generally positive. At the level of the calculus of amplitudes, the difference between classical and quantum

mechanics is just the possibility of negative distributions. That possibility alone does not completely characterize the difference when we go beyond the calculus of amplitudes, because in quantum mechanics we also have the requirement of decoherence of histories.

When the initial ρ does not factor, or when S_{free} is not of simple kinetic minus potential-energy form, the interpretation of the general expression (5.13) for the probabilities is less direct. However, building on the analogy of the special case, we may still think of $G[R(\tau), X(\tau); X_0, \xi_0]$ as a combined distribution of total force and initial conditions given the path $X(\tau)$. Now, however, the distribution of the force $R(\tau)$ is not independent of the initial conditions but depends on them. Further, the distribution of initial momenta is not given directly but only implicitly through the integral over ξ_0 . Finally, the equation of motion and noise depend on the initial values of X_0 and ξ_0 . The integral over $\xi(t)$ in Eq. (5.13) continues to enforce the classical condition that the total force $R(t)=0$ vanish. These features complicate the interpretation of Eq. (5.13) but they do not vitiate its validity.

When the noise $\mathcal{L}(t, X(\tau))$ is small compared to the equation of motion term $\mathcal{E}(t, X(\tau))$ in Eq. (5.18), we expect approximate classical determinism. More precisely we expect significant probabilities for histories correlated in time by the classical equation $\mathcal{E}(t, X(\tau))=0$ with small deviations produced by the noise. In the following sections we shall analyze the circumstances where this is so.

VI. LINEAR SYSTEMS

A. Equations of motion

We begin our discussion of the derivation of classical equations of motion for quantum systems by considering

$$\begin{aligned} W[x'(t), x(t)] &= \frac{1}{2} \int_0^T dt \int_0^t dt' [x'(t) - x(t)]^\dagger [k(t, t')x'(t') + k^*(t, t')x(t')] \\ &= \frac{1}{2} \int_0^T dt \int_0^t dt' [x'(t) - x(t)]^\dagger \{k_R(t, t')[x'(t') + x(t')] + ik_I(t, t')[x'(t') - x(t')]\}, \end{aligned} \quad (6.2)$$

where $k(t, t')$ is a complex matrix kernel with real and imaginary parts $k_R(t, t')$ and $k_I(t, t')$, respectively. We shall explicitly assume that W depends on x_0 and x'_0 only implicitly through its dependence on the paths $x'(t)$ and $x(t)$ and not explicitly as in the general case.

Only under very restrictive conditions will the influence phase be exactly quadratic as in Eq. (6.2) with no explicit dependence on x'_0 and x_0 . Such an influence phase will certainly follow if (i) action of the ignored variables, $S_0[Q(t)]$, is quadratic in the Q 's, (ii) the interaction between the distinguished and ignored variables is exactly linear in each, giving

$$S_{\text{int}}[x(t), Q(t)] = - \int_0^T dt x^\dagger(t) f(Q(t)), \quad (6.3)$$

the simplest possible example—linear systems. This is the class of models studied by Feynman and Vernon [16], Caldeira and Leggett [13], and many others and for which there is a wealth of information available on the specific forms of the influence functional, its dependence on the initial condition, etc. Either implicitly or explicitly, equations of motion have been considered for these models by several authors. We are thus on familiar territory.

Linear systems may be characterized precisely, following Feynman and Vernon, by the following two requirements:

(1) A free action for the variables distinguished by the coarse graining with a kinetic energy that is quadratic in the coordinates and their velocities:

$$S_{\text{free}}[x(t)] = \int_0^T dt \left[\frac{1}{2} \dot{x}^\dagger(t) M \dot{x}(t) - \frac{1}{2} x^\dagger(t) K x(t) \right]. \quad (6.1)$$

We use here and throughout an obvious vector notation so that M and K are positive constant matrices and $\frac{1}{2}(\dot{x}^\dagger M \dot{x}) = \frac{1}{2} \sum_{ab} \dot{x}^a M_{ab} \dot{x}^b$, etc., in this case where the variables are real.

(2) An influence phase that is at most quadratic in the variables $x(t)$. Its most general form has been deduced by Feynman and Vernon [16] from general symmetries and quantum-mechanical causality. (See also the exposition in Ref. [33].) It consists of terms linear in the distinguished variables $x(t)$ and $x'(t)$ and terms quadratic in them. The linear terms may be eliminated by a time-dependent shift in $x(t)$ and for simplicity we imagine this has been done. The general form is then

where the force f is homogeneous and linear in the Q 's, and (iii) the initial density matrix factors as in Eq. (3.8) with a density matrix $\rho_B(Q'_0, Q_0)$ that is of the form

$$\rho_B(Q'_0, Q_0) = \exp[-B(Q'_0, Q_0)], \quad (6.4)$$

where $B(Q'_0, Q_0)$ is quadratic in its arguments. Under these conditions the integral in Eq. (3.3) is a Gaussian and a quadratic influence phase results. The models defined by Eqs. (3.6), (6.1), (6.2), and (6.4) are, in fact, just those considered by Feynman and Vernon [16] and Caldeira and Leggett [13].

A useful explicit example, studied in [13,16], is the case of a single coordinate x interacting with an assembly of "bath" oscillators, with an initial condition that factors

as in Eq. (3.8) and with a state of thermal equilibrium for the "bath" oscillators at temperature $T=(k\beta)^{-1}$. The oscillators are described by a free action

$$S_0[Q(t)] = \sum_A \int_0^T dt \left[\frac{1}{2} m [\dot{Q}^A(t)]^2 - \frac{1}{2} m \omega_A^2 [Q^A(t)]^2 \right]. \quad (6.5)$$

The force f entering Eq. (6.3) may be written explicitly as

$$f(Q(t)) = \sum_A C_A Q^A \quad (6.6)$$

with coupling constants C_A , and the function B is independent of x'_0 and x_0 and given, up to an additive constant (determinable through normalizing ρ_B), by the expression

$$B(Q'_0, Q_0) = \sum_A \frac{m \omega_A}{2\hbar \sinh(\hbar\beta\omega_A)} \times \{ [Q_0^A]^2 + (Q'_0)^2 \} \cosh(\hbar\beta\omega_A) - 2Q'_0 Q_0^A. \quad (6.7)$$

Then, as in Ref. [13] or [16],

$$k_R(t, t') = - \sum_A \frac{C_A^2}{m \omega_A} \sin \omega_A (t - t') \quad (6.8a)$$

and

$$k_I(t, t') = \sum_A \frac{C_A^2}{m \omega_A} \coth \left[\frac{1}{2} \hbar\beta\omega_A \right] \cos \omega_A (t - t'). \quad (6.8b)$$

The influence phase quoted in Eq. (3.10) for a distinguished oscillator interacting with a high-temperature thermal bath is a special case of these expressions in which a continuum limit of oscillators with the special couplings

$$\rho_D(\omega) C^2(\omega) = 4mM\gamma\omega^2/\pi \quad (6.9)$$

below the cutoff was assumed, where ρ_D is the density of states. With this coupling, we have the following, with the high-temperature limit needed only in Eq. (6.10b) and the high cutoff, which facilitates phase dispersal, used in both equations:

$$k_R(t, t') = -4M\gamma\delta'(t - t'), \quad (6.10a)$$

$$k_I(t, t') = \frac{8M\gamma kT_B}{\hbar} \delta(t - t'), \quad (6.10b)$$

where the δ' function occurring in retarded integrals is a distribution such that

$$\int_{-\infty}^t g(t') [\delta'(t - t')] dt' = \frac{1}{2} g'(t). \quad (6.10c)$$

[Equation (6.10a) gives the renormalized value of $k_R(t, t')$. Infinite terms in Eq. (6.2) proportional to $\delta(t - t')$ that arise from the continuum limit taken by Caldeira and Leggett have been absorbed in a renormal-

ization of the frequency of the distinguished oscillator in its free action.] Although the influence phase is exactly quadratic only under restrictive circumstances, the linear cases supply useful models for more general ones, as we shall see.

With these preliminaries, we can now give a derivation of the classical equation of motion for these linear systems. The equation we shall derive is, of course, the same as considered, for example, by Caldeira and Leggett. Even in this linear case, however, we believe that there are several important new features of this derivation. It is consistent with the general discussion of the average equation of motion and noise in Sec. V. The probabilities that coarse-grained histories of the distinguished particle are correlated in time by equations of motion are explicitly considered. The form of the phenomenological equation of motion is derived from a consideration of these probabilities. The amount of coarse graining necessary for the decoherence of histories and their classical behavior is discussed quantitatively, and the connection between decoherence and quantum noise is made explicit. Most importantly, the derivation suggests how the generalization to the nonlinear case is to be carried out.

The imaginary term in the influence phase (6.2) gives rise to decoherence between the trajectories of the distinguished variables, provided that the coarse graining is such that the integral of this term is sufficiently large for different coarse-grained histories so that the corresponding "off-diagonal" elements of the decoherence functional are exponentially small. To exhibit this decoherence explicitly it is useful to change variables, in the integral (3.5) defining the decoherence functional, from $x'(t)$ and $x(t)$ to the average and difference $X(t)$ and $\xi(t)$ defined by Eq. (5.4). The exponent in the decoherence functional (3.5) can be explicitly expressed in terms of $X(t)$ and $\xi(t)$ using Eq. (6.2). Denoting this exponent by $A[X(\tau), \xi(\tau)]$ as in Eq. (5.6), one finds after a few integrations by parts that

$$D(\alpha', \alpha) = \int_{(\alpha', \alpha)} \delta X \delta \xi \delta(\xi_f) \exp\{iA[X(\tau), \xi(\tau)]/\hbar\} \times \bar{\rho}(X_0 + \xi_0/2, X_0 - \xi_0/2), \quad (6.11)$$

where

$$A[X(\tau), \xi(\tau)] = -\xi_0^\dagger M \dot{X}_0 + \int_0^T dt \xi^\dagger(t) e(t, X(\tau)) + \frac{i}{4} \int_0^T dt \int_0^T dt' \xi^\dagger(t) k_I(t, t') \xi(t'). \quad (6.12a)$$

Here, $e(t, X(\tau))$ is the average "total force,"

$$e(t, X(\tau)) = -M\ddot{X}(t) - KX(t) + \int_0^t dt' k_R(t, t') X(t'), \quad (6.12b)$$

and k_I is defined to be symmetric in its argument, so that the limit of integration on t' can be extended to T . The quantities \dot{X}_0 , X_0 , etc. are understood in the usual path integral sense as finite difference expressions in a time-sliced implementation of the path integral. (See Sec. B of the Appendix.) The integrals over $X(t)$ and $\xi(t)$ are con-

strained by the coarse graining defining the histories α' and α .

The imaginary term in Eq. (6.12a) leads to decoherence. We shall give below, in the discussion of noise, an explicit construction of the kernel k_I that shows it to be a positive kernel. The imaginary part of W is proportional to k_I and occurs in the expression $\exp(iW/\hbar)$, giving a decreasing exponential. If the graining defined by the successive sets of regions $\{\Delta_{\alpha_k}^k\}$ is coarse enough so that sufficient positive imaginary part of W is built up between one set of intervals and the next, then there will be a significant contribution to the integral defining the decoherence functional only for values of $\xi(t)$ near zero and for $\alpha=\alpha'$. (See Fig. 1.) That is medium decoherence.

If only values of $\xi(t)$ near zero contribute significantly to the integral (6.11), then in the diagonal elements of the decoherence functional, which are the probabilities of the individual coarse-grained histories, the integral over $\xi(t)$ may be carried out, to an excellent approximation, as though unrestricted by the coarse graining when the intervals are sufficiently coarse. (See Fig. 1.) The integral over ξ_0 leads to the Wigner distribution (5.2) as in Eq. (5.16). The result of the unrestricted Gaussian integrals over the rest of $\xi(t)$ is again a Gaussian functional. We thus obtain for the probabilities of the individual histories in the coarse-grained set, the expression⁹

$$p(\alpha) \cong \int_{\alpha} \delta X [\det(k_I/4\pi)]^{-1/2} \times \exp \left[-\frac{1}{\hbar} \int_0^T dt \int_0^T dt' e^\dagger(t, X(\tau)) \times k_I^{\text{inv}}(t, t') e(t', X(\tau)) \right] w(X_0, M\dot{X}_0), \quad (6.13)$$

where k_I^{inv} is the inverse kernel to k_I . The integral is over all paths that proceed from $t=0$ to T and lie in the class corresponding to the coarse-grained history α . The integral includes an integration over the initial and final endpoints X_0 and X_f , respectively. Again, for further details and an explicit representation of Eq. (6.13), see Sec. B of the Appendix.

The Gaussian exponential in Eq. (6.13) means that, for given X_0 and $M\dot{X}_0$, the histories with the largest probabilities will be those with $e(t)=0$, that is, those for which the time evolution is correlated according to the effective average classical equation of motion

$$e(t, X(\tau)) = -M\ddot{X}(t) - KX(t) + \int_0^t dt' k_R(t, t') X(t') = 0 \quad (6.14)$$

This is, of course, not the equation of motion following

from the free action of the distinguished oscillator. It differs by the additional force, nonlocal in time, that arises from the interactions of the distinguished variables with the rest. The presence of such a force will, in general, mean that energy is not conserved, leading sometimes to dissipation. Although nonlocal in time, the additional force in Eq. (6.14) is retarded, expressing classical causality. The origin of this retardation can be traced to the retarded form of the general influence phase (6.2). That, in turn, follows from quantum-mechanical causality—the fact that the decoherence functional has a trace in the future and a density matrix differing from the unit matrix in the past. Causality in quantum mechanics thus implies the causality of classical physics.

A special case of a linear system is the Fokker-Planck limit of the oscillator model, for which the influence phase is exhibited in Eq. (3.10). With the corresponding k_R of Eq. (6.10a), the equation of motion away from $t=0$ becomes¹⁰

$$e(t, X(\tau)) = -M\ddot{X}(t) - KX(t) - 2M\gamma\dot{X}(t) = 0. \quad (6.15)$$

This is local in time, but that is a special property of the way in which the limit of a continuous spectrum of oscillators was taken in the Caldeira-Leggett model, not a general one. In that limit, Eq. (6.15) explicitly exhibits the familiar form of frictional dissipation, not necessarily a general characteristic of the additional force in Eq. (6.14).

The individual classical histories in Eq. (6.13) are distributed according to the probabilities of their initial conditions X_0 and $P_0 = M\dot{X}_0$ given by the Wigner function $w(X_0, P_0)$. Although the Wigner function is not generally positive, we know, as discussed in Sec. V, that apart perhaps from small errors introduced by the approximation in which the constraints of the coarse graining on the ξ integrations were neglected, the result of the integral (6.13) must be positive even though the Wigner function is not.¹¹

B. Noise and predictability

The distribution of probabilities for histories (6.13) predicts the largest probability for histories obeying the classical equations of motion but also predicts probabilities for deviations from classical predictability. Those give the noise, including quantum noise. The same interactions (of the distinguished variables with the others) that carry away phase information to produce decoherence also produce the quantum and classical-statistical buffeting of the trajectory of the distinguished variables that constitutes the noise.

In Sec. V we showed how the probabilities of coarse-grained histories $p(\alpha)$ could be thought of as the proba-

⁹Dowker and Halliwell [19] have obtained analogous expressions in linear models for the probabilities of histories defined by a finite number of “Gaussian slits.”

¹⁰We are thus for simplicity ignoring the terms proportional to $x(0)$ that arise when the integral in Eq. (6.12b) is carried out using Eq. (6.10). For further discussion see Ref. [34]. We thank J. P. Paz for a discussion of this point.

¹¹For explicit examples of this see Ref. [35].

bilities of a classical dynamical problem with (generally nonpositive) distributions of force and initial conditions. We now specialize that discussion to the linear systems of the present section.

Since we have assumed factorization of the initial ρ , we may compute separate distributions for the initial conditions and for the total force as in Eq. (5.16). The initial conditions are distributed according to the Wigner distribution as shown by that equation or directly from Eq. (6.13). The total force is then distributed according to the distribution function $g[R(\tau), X(\tau)]$ defined by Eq. (5.17). Using Eq. (6.12a) for A and Eq. (5.14) to define \tilde{A} , we may calculate $g[R(\tau), X(\tau)]$ directly. It is

$$g[R(\tau), X(\tau)] = [\det(k_I/4\pi)]^{-1/2} \times \exp \left[-\frac{1}{\hbar} \int_0^T dt \int_0^T dt' [R(t) - e(t, X(\tau))]^\dagger \times k_I^{\text{inv}}(t, t') [R(t') - e(t', X(\tau))] \right], \quad (6.16)$$

where $e(t, X(\tau))$ is given by Eq. (6.12b) and the precise meaning of the inverse kernel $k_I^{\text{inv}}(t, t')$ is discussed in the Appendix.

Evidently, we have, for this case,

$$\mathcal{E}(t, X(\tau)) \equiv \langle R(t) \rangle_c = e(t, X(\tau)). \quad (6.17)$$

The equation of motion *defined* as the expected value of R in Sec. VB, therefore, coincides with the equation of motion $e(t, X(\tau))$ whose correlations are favored by the probabilities (6.13). The Langevin force $\mathcal{L}(t, X(\tau))$ that governs the deviations from classical predictability is distributed according to

$$g[e(t) + \mathcal{L}(t, X(\tau)), X(\tau)].$$

As is easily seen from Eq. (6.16) for the linear models under consideration this noise is distributed with a *positive* Gaussian probability distribution that is independent of the path $X(\tau)$. To emphasize this we write

$$\mathcal{L}(t, X(\tau)) = l(t) \quad (6.18)$$

in the linear case, and $l(t)$ is then distributed according to

$$[\det(k_I/2\pi)]^{-1/2} \times \exp \left[-\frac{1}{\hbar} \int_0^T dt \int_0^T dt' l^\dagger(t) k_I^{\text{inv}}(t, t') l(t') \right]. \quad (6.19)$$

The spectrum of the Gaussian noise is summarized by the formula

$$\langle l(t)l(t') \rangle_c = (\hbar/2) k_I(t, t'). \quad (6.20)$$

In the Fokker-Planck limit of the oscillator model, we have

$$\langle l(t)l(t') \rangle_c = 4M\gamma kT_B \delta(t-t'), \quad (6.21)$$

giving rise to the model for Brownian motion discussed

by Caldeira and Leggett [13].

For linear systems instructive, explicit expressions for the functions $k_R(t, t')$, and $k_I(t, t')$, which describe the influence of the ignored variables on the distinguished ones, may be obtained in terms of quantum-mechanical expectation values of the force $f(Q)$ defined by Eq. (6.3). As these are straightforwardly derived as special cases of the similar formulas, applicable to nonlinear situations, to be discussed in Sec. VII, we shall just quote the results here.

Consider the Hilbert space of square-integrable functions in the ignored Q 's. Define an expected value $\langle \rangle_0$ of an operator $A(t)$ evolving by the Hamiltonian of S_0 by

$$\langle A(t) \rangle_0 = \text{Sp}[A(t)\rho_B], \quad (6.22)$$

where Sp denotes the trace over the ignored Q 's. The subscript zero means that the time dependence of the operators inside the expected value is calculated using the Hamiltonian H_0 of the Q 's alone neglecting interactions with the x 's.

As a consequence of our convention that the influence phase has no terms linear in $x(t)$, the expected value of $f(Q(t))$ vanishes:

$$\langle f(Q(t)) \rangle_0 = 0. \quad (6.23)$$

The real and imaginary parts of the kernel $k(t, t')$ may be expressed in terms of expected values of fluctuations in the force as

$$\hbar k_R(t, t') = i \langle [f(Q(t)), f(Q(t'))] \rangle_0, \quad (6.24)$$

$$\hbar k_I(t, t') = \langle \{f(Q(t)), f(Q(t'))\} \rangle_0, \quad (6.25)$$

where square brackets and curly brackets denote the commutator and anticommutator, respectively, and the matrix elements of k are understood to be the tensor product of the f 's. Expression (6.25) shows explicitly that the kernel $k_I(t, t')$ is positive in the sense that

$$\int_0^T dt \int_0^T dt' \xi^\dagger(t) k_I(t, t') \xi(t') \geq 0 \quad (6.26)$$

for any real vector $\xi(t)$. The same equation and Eq. (6.20) demonstrates that the spectrum of the random Gaussian force in the Langevin equation $e(t) + l(t) = 0$ is directly given by the quantum correlation function of the fluctuation in the force $f(Q)$, viz.,

$$\langle l(t)l(t') \rangle_c = \frac{1}{2} \langle \{f(Q(t)), f(Q(t'))\} \rangle_0. \quad (6.27)$$

Expressions (6.24) and (6.25) lead to the essential content of the fluctuation-dissipation theorem. If ρ_B is diagonal in the energy representation defined by the Hamiltonian of S_0 as it is for equilibrium distributions, then $k_R(t, t')$ and $k_I(t, t')$ are functions of the time difference $t-t'$, and their spectral weights are simply related. More specifically, it follows from the symmetries of commutator and anticommutator that we could write

$$k_R(t, t') = \int_0^\infty d\omega \tilde{k}_R(\omega) \sin \omega(t-t'), \quad (6.28a)$$

$$k_I(t, t') = \int_0^\infty d\omega \tilde{k}_I(\omega) \cos \omega(t-t'). \quad (6.28b)$$

If p_i is the probability of an energy eigenstate $|i\rangle$ with ei-

genvalue E_i in the density matrix ρ_B , we have

$$\hbar\tilde{k}_R(\omega) = \sum_{ij} i(p_i - p_j) |\langle i | f(Q(0)) | j \rangle|^2 \delta(E_j - E_i - \hbar\omega), \quad (6.29a)$$

$$\hbar\tilde{k}_I(\omega) = \sum_{ij} (p_i + p_j) |\langle i | f(Q(0)) | j \rangle|^2 \delta(E_j - E_i - \hbar\omega). \quad (6.29b)$$

Then, in a thermal bath where $p_i = \exp(-\beta E_i)/Z$, we recover the famous relation (see, e.g., Ref. [18])

$$\hbar\tilde{k}_I(\omega) = \hbar \coth(\beta\omega\hbar/2) \tilde{k}_R(\omega). \quad (6.30)$$

This connects the kernel $k_R(t, t')$ governing the effective force in Eq. (6.14) with the kernel $k_I(t, t')$ governing the fluctuations in Eq. (6.20). This connection is the fluctuation-dissipation theorem.

An important fact that emerges clearly from these linear models is that the same coarse graining and interactions that accomplish decoherence also lead to dissipation and noise. The fluctuation-dissipation theorem derived above is a well-known example of the connection between noise and dissipation. The connection between decoherence and these two phenomena appears to have been less widely stressed. We now consider it.

There is, in effect, a competition between decoherence and classical predictability. Consider, for example, a model of an oscillator interacting with a thermal bath. Increase the temperature of the thermal bath, T_B , and decoherence is more effective. The characteristic time $t_{\text{decoherence}}$, by which successive intervals of typical length d must be spaced to give decoherence, *decreases* with T_B according to Eq. (4.38). However, deviations from classical predictability expressed by Eq. (6.13) also increase. To ensure *both* classical predictability and decoherence we must consider a further limit, the limit of high inertia. In the present model, that is the limit of large M , and the exponent in Eq. (6.13) can be written

$$-\frac{M}{8\gamma k T_B} \int_0^T dt [\ddot{X} + M^{-1}V'(x) + 2\gamma\dot{X}]^2. \quad (6.31)$$

Assuming that V does not increase faster than M , we see that the probabilities for histories will become sharply peaked about the certainties implied by a classical equation of motion in the limit of large M/T_B even as T_B itself is becoming large to ensure efficient decoherence.

Of course, in realistic situations the parameters of a given system, such as the mass, are fixed. The same kind of limit of high inertia can be achieved, however, by considering coarser and coarser graining of a kind that keeps increasing the inertia of the variables distinguished by the coarse graining. Take the case of coarse grainings defined by hydrodynamic variables that are integrals, over suitable volumes, of densities of exactly or approximately conserved quantities such as mass, energy, or momentum. By making the size of these volumes larger, the resistance to noise can be increased. In the present model, decoherence and classical predictability can be achieved only by varying the parameters of the model. In realistic situations they are achieved by a suitable coarse

graining.

Classical behavior requires sufficient coarse graining and interactions for decoherence but sufficient "inertia" to resist the deviation from predictability that the coarse graining and interactions produce. Traditionally other descriptions have been given of the requirements for classical behavior of measured subsystems. Large action or high quantum numbers are often mentioned. While such criteria are not as precise or as complete as those deduced here, it can be seen from simple dimensional arguments that in typical situations an action *will* be large compared to \hbar when the two requirements of decoherence and sufficient inertia are satisfied. Let us consider a one-dimensional oscillator model of the kind just discussed in the high temperature limit. Let t_{dyn} be the shortest dynamical time scale of interest and assume that the coarse graining is characterized by sets of intervals of characteristic size Δ separated in time by t_{dyn} . From Eq. (3.10), it follows that decoherence requires

$$\frac{M\gamma k T_B}{\hbar^2} (t_{\text{dyn}} \Delta^2) \gg 1. \quad (6.32)$$

From Eq. (6.31), the requirement of sufficient inertia is

$$\frac{M}{\gamma k T_B} t_{\text{dyn}} \left[\frac{\Delta}{t_{\text{dyn}}} \right]^2 \gg 1, \quad (6.33)$$

if we assume $t_{\text{dyn}} \ll t_{\text{relaxation}} \equiv 1/\gamma$.

These two requirements may be reexpressed in terms of the characteristic scale of classical actions

$$S \sim t_{\text{dyn}} M \left[\frac{\Delta}{t_{\text{dyn}}} \right]^2 \quad (6.34)$$

and the thermal correlation time $t_{\text{thermal}} \equiv \hbar/kT_B$. One finds

$$(S/\hbar) \gg (1/\eta), \quad (S/\hbar) \gg \eta \quad (6.35)$$

for the requirements (6.32) and (6.33), respectively, where η is the ratio

$$\eta = (t_{\text{dyn}}/t_{\text{thermal}})(t_{\text{dyn}}/t_{\text{relaxation}}). \quad (6.36)$$

Whatever the size of η , the relations (6.35) imply that $S/\hbar \gg 1$ in the classical limit.

VII. EQUATIONS OF MOTION FOR NONLINEAR SYSTEMS

In this section we generalize the results of the preceding two sections to the more realistic situation where the action $S_{\text{free}}[x(\tau)]$, and the influence phase $W[x'(\tau), x(\tau)]$ are not necessarily quadratic functionals of their arguments. We begin in Sec. A by deriving some useful general properties of the influence phase. These are used in Sec. B to derive the form of the phenomenological equations of motion and analyze the restrictions on the coarse graining that permit the histories to stay close to solutions of those equations with high probability. In Sec. C we derive the classical causality of these equations of motion from quantum-mechanical causality. The linear

theory, discussed in the preceding two sections, is recovered in Sec. D.

A. General relations for the influence phase

The influence phase $W[x'(\tau), x(\tau)]$ is defined by the functional integral (3.3) over the ignored variables Q^A . We continue to assume a factored initial condition as in Eq. (3.8), so that W has no explicit dependence on x'_0 and x_0 . A useful operator expression for $W[x'(\tau), x(\tau)]$ may be derived by noting the following: The integrals over the $Q(t)$ in Eq. (3.3) are over paths between $t=0$ and T that are *unrestricted* except at their initial and final end points. They may, therefore, be thought of as defining the unitary evolution of a family of operators $\rho(x'_0, x_0)$ in the Hilbert space \mathcal{H}_Q of square-integrable functions of the Q 's. The dynamics of this evolution of the Q 's is specified by the action

$$S_Q[x(\tau), Q(\tau)] = S_0[Q(\tau)] + S_{\text{int}}[x(\tau), Q(\tau)], \quad (7.1)$$

which depends on the path $x(\tau)$ as a parameter. There is a corresponding Hamiltonian operator on \mathcal{H}_Q . If we assume that the interaction is local in time, specifically such that

$$S_{\text{int}}[x(\tau), Q(\tau)] = \int_0^T dt L_{\text{int}}(x(t), Q(t)), \quad (7.2)$$

then that Hamiltonian at time t depends only on the instantaneous value of $x(t)$: viz.,

$$H_Q(x(t)) = H_0 + H_{\text{int}}(x(t)). \quad (7.3)$$

Here H_0 is the Hamiltonian of the Q variables, omitting their interaction with the x variables, corresponding to the action $S_0[Q]$. The operator effecting the unitary evolution generated by this Hamiltonian between times t' and t'' is

$$U_{t'', t'}[x(\tau)] = \mathcal{T} \exp \left[-\frac{i}{\hbar} \int_{t'}^{t''} dt H_Q(x(t)) \right], \quad (7.4)$$

where \mathcal{T} denotes the time ordered product.

Write ρ_B for the density operation on \mathcal{H}_Q whose matrix elements are

$$\langle Q'_0 | \rho_B | Q_0 \rangle = \rho_B(Q'_0, Q_0), \quad (7.5)$$

where $\rho_B(Q'_0, Q_0)$ is the factor of the initial density matrix (3.8) referring to the Q 's. Utilizing the U 's defined by Eq. (7.4) and the ρ_B defined by Eq. (7.5) the path integral relation (3.3) defining the influence phase $W[x'(\tau)]$ may be reexpressed (see Sec. D of the Appendix) as

$$\begin{aligned} \exp\{iW[x'(\tau), x(\tau)]/\hbar\} \\ = \text{Sp}\{U_{T,0}[x'(\tau)]\rho_B U_{T,0}^\dagger[x(\tau)]\}, \quad (7.6) \end{aligned}$$

where Sp denotes the trace operation on the Hilbert space \mathcal{H}_Q . We shall now use this relation to derive some useful general properties of $W[x'(\tau), x(\tau)]$.

First, it is an immediate consequence of $\text{Sp}(A^\dagger) = (\text{Sp}(A))^*$ and the form of right-hand side of Eq. (7.6) that interchanging $x'(t)$ and $x(t)$ on the left-hand side is equivalent to complex conjugation. (This

elementary result also follows directly from Eq. (3.3), as noted in Refs. [16,33].) Thus,

$$\text{Re } W[x'(\tau), x(\tau)] = -\text{Re } W[x(\tau), x'(\tau)], \quad (7.7a)$$

$$\text{Im } W[x'(\tau), x(\tau)] = +\text{Im } W[x(\tau), x'(\tau)]. \quad (7.7b)$$

In particular, if the influence phase is written as a functional $W[X(\tau), \xi(\tau)]$ of the average of $x'(t)$ and $x(t)$ and the difference between them [Eq. (5.4)], then $\text{Re}(W)$ is an *odd* functional of $\xi(t)$ while $\text{Im}(W)$ is an *even* functional.

As shown by Brun [35] in the following paper, an elementary application of Schwarz's inequality shows that

$$\begin{aligned} \exp(-\text{Im } W[X(\tau), \xi(\tau)]) \\ \leq \sum_i p_i |\langle \psi_i | U_{T,0}^\dagger[x(\tau)] U_{T,0}[x'(\tau)] | \psi_i \rangle| \\ \leq \sum_i p_i = 1, \quad (7.8) \end{aligned}$$

where p_i are the eigenvalues and $|\psi_i\rangle$ the eigenvectors of ρ_B . Thus $\text{Im } W[X(\tau), \xi(\tau)]$ is positive, which is essential for the convergence of the functional integral defining the decoherence functional as well as decoherence itself.

The expression (7.6) may be used to find convenient operator expressions for the coefficients of the expansion of $W[X(\tau), \xi(\tau)]$ in powers of $\xi(t)$. Generally,

$$\begin{aligned} W[x(\tau), \xi(\tau)] = W[x(\tau), 0] + \int_0^T dt \xi^\dagger(t) \left[\frac{\delta W}{\delta \xi(t)} \right]_{\xi(t)=0} \\ + \frac{1}{2} \int_0^T dt \int_0^T dt' \xi^\dagger(t) \left[\frac{\delta^2 W}{\delta \xi(t) \delta \xi(t')} \right]_{\xi(t)=0} \xi(t') + \dots \quad (7.9) \end{aligned}$$

For the leading term in Eq. (7.9), we have, evaluating the right-hand side of Eq. (7.6) at $\xi(t) = x'(t) - x(t) = 0$,

$$\exp\{iW[X(\tau), 0]/\hbar\} = \text{Sp}\{U_{T,0}[X(\tau)]\rho_B U_{T,0}^\dagger[X(\tau)]\}. \quad (7.10a)$$

Using the cyclic property of Sp and the unitarity of $U_{T,0}[X(\tau)]$ it is easy to see that the right-hand side of Eq. (7.10a) is unity. Thus, the leading term in Eq. (7.9) vanishes:

$$W[X(\tau), 0] = 0. \quad (7.10b)$$

To evaluate the next term in the expansion (7.9), we must consider the derivatives

$$\delta U_{T,0}[X(\tau) \pm \xi(\tau)/2] / \delta \xi(t).$$

To do this we introduce the definition

$$F(x(t)) = -\frac{\partial H_Q(t)}{\partial x(t)} = \frac{\partial L_{\text{int}}(x(t), Q(t))}{\partial x(t)}. \quad (7.11)$$

The operator $F(x(t))$ is an operator in the Schrödinger picture. It is a function of x because L_{int} is a function of x and it becomes a function of t because x is a function of t . It is an operator representing the force on the distinguished coordinates $x(t)$ due to their interaction with

the rest of the system.

Carrying out the indicated differentiations of the U 's yields

$$\begin{aligned} & \{\delta U_{T,0}[X(\tau) \pm \xi(\tau)/2] / \delta \xi(t)\}_{\xi(t)=0} \\ &= \pm (i/2\hbar) U_{T,t}[X(\tau)] F(X(t)) U_{t,0}[X(\tau)] \\ &\equiv \pm (i/2\hbar) F(t, X(\tau)). \end{aligned} \quad (7.12)$$

The operator $F(t, X(\tau))$ is a representative of the Schrödinger picture operator (7.11) in a picture similar to the Heisenberg picture but determined by both $U_{T,t}$ and $U_{t,0}$. It is a function of t but also a functional of the path $X(\tau)$. We indicate this dependence by writing $F(t, X(\tau))$, using a parenthesis on the left to indicate that it is a function and the bracket on the right to indicate that it is also a functional.

With the result (7.12), it is only a short calculation to find the coefficient of the linear term in Eq. (7.9). It is

$$[\delta W / \delta \xi(t)]_{\xi(\tau)=0} = \text{Sp}\{F(t, X(\tau)) \rho_B\}. \quad (7.13)$$

If we define an expected value by

$$\langle A \rangle = \text{Sp}(A \rho_B), \quad (7.14)$$

Eq. (7.13) may be written in the compact form

$$[\delta W / \delta \xi(t)]_{\xi(\tau)=0} = \langle F(t, X(\tau)) \rangle. \quad (7.15)$$

As required by Eq. (7.7), this contribution to the part of W that is odd in $\xi(t)$ is purely real.

The coefficient

$$[\delta^2 W / \delta \xi(t) \delta \xi(t')]_{\xi(\tau)=0}$$

of the quadratic term in the expansion of W in powers of $\xi(t)$ is similarly evaluated. The expansion of the U 's on the right-hand side of Eq. (7.6) will result in products of first derivatives, such as those in Eq. (7.12), but also in second derivatives. Those second derivatives yield a term in the expansion proportional to

$$\delta(t-t') \text{Sp}\{[H''_{\text{int}}(t, X(\tau)) \rho_B]\}, \quad (7.16)$$

where $H''_{\text{int}}(t, X(\tau))$ is the Heisenberg picture representative of the operator $\partial^2 H_{\text{int}} / \partial x(t)^2$. A term like Eq. (7.16) vanishes, of course, for a linear interaction such as Eq.

(6.3). It formally vanishes in the nonlinear case too because it is the trace of a commutator. Of course, that is a delicate issue in the case of unbounded operators, as the nonvanishing value of $\text{Sp}([\Pi, Q])$ shows. However, in the present case, where ρ_B is bounded and H_{int} is a function of the Q 's and not of their conjugate momenta, we may reasonably assume that Eq. (7.16) vanishes and we shall do so in what follows.

The remaining contribution to the coefficient of the quadratic term in the expansion of $W[X(\tau), \xi(\tau)]$ comes from products of first derivatives of U 's such as those in Eq. (7.12) and is straightforwardly evaluated. One finds

$$\begin{aligned} & [\delta^2 W / \delta \xi(t) \delta \xi(t')]_{\xi(\tau)=0} \\ &= (i/2\hbar) \langle \{\Delta F(t, X(\tau)), \Delta F(t', X(\tau))\} \rangle, \end{aligned} \quad (7.17)$$

curly brackets denote the anticommutator and ΔF is the operator

$$\Delta F(t, X(t)) = F(t, X(\tau)) - \langle F(t, X(\tau)) \rangle \quad (7.18)$$

representing fluctuations in the force F about its mean. We note that, as required by Eq. (7.7), this contribution to the even part of $W[X(\tau), \xi(\tau)]$ in $\xi(t)$ is purely imaginary. When divided by i , it is also manifestly positive in the sense of Eq. (6.26).

With these preliminaries we may now derive the nonlinear equations of motion and discuss their form.

B. Nonlinear equations of motion

We consider a set of alternative coarse-grained histories specified at a sequence of times t_1, \dots, t_n by sets of exhaustive and exclusive regions of the x 's which we denote by $\{\Delta_{\alpha_1}^1\}, \{\Delta_{\alpha_2}^2\}, \dots, \{\Delta_{\alpha_n}^n\}$. The decoherence functional for such sets of histories is given by Eq. (3.5). We assume that the regions and times are chosen so that there is a negligible contribution to the path integrals in the decoherence functional except when $\xi(t) = x'(t) - x(t)$ is small. We expect to have such coarse grainings, for example, if the imaginary part of the influence phase $W[X(\tau), \xi(\tau)]$ has its minimum at $\xi(t) = 0$. Our assumption about the integrals in Eq. (3.5) implies the decoherence of such a set of alternative coarse-grained histories (cf. Fig. 1) and the following formula for their probabilities:

$$p(\alpha) \equiv \int_{\alpha} \delta X \delta \xi \delta(\xi_f) \exp \left\{ \frac{i}{\hbar} \left[S_{\text{free}} \left[X(\tau) + \frac{\xi(\tau)}{2} \right] - S_{\text{free}} \left[X(\tau) - \frac{\xi(\tau)}{2} \right] + W[X(\tau), \xi(\tau)] \right] \right\} \bar{\rho} \left[X_0 + \frac{\xi_0}{2}, X_0 - \frac{\xi_0}{2} \right]. \quad (7.19)$$

The functional integral is over paths in both X and ξ as restricted by the coarse-grained history α .

If only small values of $\xi(t)$ contribute to the integrals in Eq. (7.19), we may make a further approximation by expanding the exponent in powers of $\xi(t)$, utilizing the

expansion in Eq. (7.9), up to the quadratic terms. We then have an integral for the probability of a history that is of precisely the same form as the one occurring in the discussion of the linear theory, (6.13), with $e(t)$ replaced by

$$\mathcal{E}(t, X(\tau)) = \frac{\delta S_{\text{free}}[X(\tau)]}{\delta X(t)} + \langle F(t, X(\tau)) \rangle \quad (7.20)$$

and with $k_I(t, t')$ replaced by

$$K_I(t, t') = \hbar^{-1} \langle \{ \Delta F(t, X(\tau)), \Delta F(t', X(\tau)) \} \rangle . \quad (7.21)$$

In writing out these identifications we have made use of Eqs. (7.10b), (7.15), and (7.17) for the expansion coefficients of $W[X(\tau), \xi(\tau)]$. Equation (7.21) shows the kernel $K_I(t, t'; X(\tau))$ is manifestly positive—a necessary

condition for the mechanism of decoherence being discussed, not to mention the convergence of the integral (7.19) with the expanded exponent.

Under the assumption that only a narrow range of $\xi(t)$ near zero contributes to the integral (7.19), it is a good further approximation to neglect the constraints on the integration range of $\xi(t)$ arising from the coarse graining. The resulting Gaussian integrals can then be carried out, yielding an expression for $p(\alpha)$ that is the generalization of Eq. (6.13):

$$p(\alpha) \cong \int_{\alpha} \delta X [\det(K_I/4\pi)]^{-1/2} \exp \left[-\frac{1}{\hbar} \int_0^T dt \int_0^T dt' \mathcal{E}^\dagger(t, X(\tau)) K_I^{\text{inv}}(t, t'; X(\tau)) \mathcal{E}(t', X(\tau)) \right] w(X_0, P_0) , \quad (7.22)$$

where $K_I^{\text{inv}}(t, t'; X(\tau))$ is the inverse kernel to $K_I(t, t'; X(\tau))$ and P_0 is the momentum of the free action expressed in terms of X_0 and \dot{X}_0 . The measure δX is discussed in the Sec. B of the Appendix.

The derivation and analysis of the equations of motion now proceed as in the linear example, with important differences that we shall mention. The Gaussian form of the exponent in Eq. (7.22) means that for given X_0 and P_0 the most significant contribution comes from the histories with $\mathcal{E}(t, X(\tau)) \approx 0$; that is, those whose evolution in time nearly follows the effective classical equation of motion

$$\mathcal{E}(t, X(\tau)) = \frac{\delta S_{\text{free}}[X(\tau)]}{\delta X(t)} + \langle F(t, X(\tau)) \rangle = 0 . \quad (7.23)$$

The probabilities predicted by Eq. (7.22) are, therefore, those of an ensemble of classical histories individually correlated in time by the equation of motion (7.23) and with initial conditions distributed according to the Wigner function $w(X_0, P_0)$.

The first term in Eq. (7.23) is the equation of motion of the distinguished coordinates x^a in the absence of any interaction with the remaining coordinates Q^A . The second term is the expected value of the force arising from that interaction. This is a *functional* of the trajectory and will, in general, be nonlocal in time. As we shall show below, it is retarded as a consequence of quantum-mechanical causality. It typically leads to dissipation, although under some conditions the energy might actually increase. The phenomenological force is also generally dependent on the initial condition ρ_B through Eq. (7.13). The familiar, phenomenological equations describing, for example, *dissipative* friction, are characterized by a few parameters *independent* of initial conditions and are the result of further approximations to Eq. (7.23). These are typically good in situations where there is a significant contribution only from retarded times that are short compared to the relaxation times of that part of the bath that interacts significantly. If the distinguished system has energy large compared to kT_B , it will lose energy to the bath on the average. The result is a dissipative phenomenological equation, local in time, with parameters

independent of initial conditions, like Eq. (6.15).

Viewed as a generalization of the linear case, the important point about the equation of motion (7.23) is that both the free part of the equation of motion and the contribution from the interaction of the x 's with the Q 's are, in general, nonlinear in $X(t)$. For the special coarse grainings in which the variables are divided into a set distinguished by the coarse graining and a set ignored, we therefore have a general derivation of the form of the phenomenological equations of motion. We now discuss in more detail the implications of quantum-mechanical causality and quantum noise.

C. Quantum-mechanical causality implies classical causality

Feynman and Vernon [16] used path integral arguments to show that if $\xi(t)$ is set to zero for $t > t_*$, when $W[X(\tau), \xi(\tau)]$ is independent of $X(t)$ for $t > t_*$. This result could be used to show that the force $\langle F(t, X(\tau)) \rangle$ is retarded, that is independent of the path $X(\tau)$ for values of τ greater than t . However, the result also follows easily from the definitions (7.12), (7.14), and (7.15). Writing out the expected value of F in the Schrödinger picture, we have

$$\langle F(t, X(\tau)) \rangle = \text{Sp} \{ U_{T,t}[X(\tau)] F(X(t)) U_{t,0}[X(\tau)] \times \rho_B U_{T,0}^\dagger[X(\tau)] \} . \quad (7.24)$$

Using the cyclic property of the spur, the composition law and unitarity of the evolution operators defined by Eq. (7.4), we may write Eq. (7.24) in the form

$$\langle F(t, X(\tau)) \rangle = \text{Sp} \{ F(X(t)) U_{t,0}[X(\tau)] \rho_B U_{t,0}^\dagger[X(\tau)] \} . \quad (7.25)$$

Since $U_{t,0}[X(\tau)]$ depends on $X(\tau)$ only for $0 < \tau < t$ [cf. Eq. (7.4)], this shows that $\langle F(t, X(\tau)) \rangle$ is retarded.

The expression (3.1) for the decoherence functional incorporates a quantum-mechanical notion of causality. At one end of the histories, information about the specific

closed system in the form of a density matrix must be supplied. At the other end is a unit matrix in the form of a δ function representing a condition of indifference with respect to states at the end of the histories. This asymmetry between the two ends of the histories is the arrow of time in quantum mechanics. (See, e.g., Refs. [36,2].) It is by *convention* that we call the extremity of the histories next to the density matrix ρ *the past* and consider time as increasing from it.

To predict the future in quantum mechanics, we need the initial ρ and information about histories up to the present. We need no information about the future. That is quantum-mechanical causality. The retardation of $\langle F(t, X(\tau)) \rangle$ expressed by Eq. (7.25) shows that we need know the trajectory of the distinguished system only from the time of the initial ρ to the present to predict the system's future. That is classical causality. Equation (7.25) thus shows that quantum-mechanical causality implies classical causality.

The origin and implications of the arrow of time in quantum mechanics may be usefully discussed using a hypothetical generalization of the quantum mechanics of closed systems that employs both initial and final conditions [5,4,20]. In this generalization the decoherence functional would be given by

$$D(\alpha'\alpha) = N \int_{\alpha'} \delta q' \int_{\alpha} \delta q \rho_f(q_f, q'_f) \times \exp(i\{S[q'(\tau)] - S[q(\tau)]\} / \hbar) \rho_i(q'_0, q_0), \quad (7.26a)$$

where

$$N^{-1} = \int dq \int dq' \rho_f(q, q) \rho_i(q, q'). \quad (7.26b)$$

This generalization of quantum mechanics would permit the future and the past to be treated similarly. Arrows of time would arise in particular universes where ρ_f is different from the time-reversed version of ρ_i . In particular, usual quantum mechanics, represented by Eq. (3.1) and its associated causality and arrow of time arise for those universes, like ours, where $\rho_f \propto I$ is a good representation of the final condition. In more general situations, with ρ_f not proportional to I , there would be neither a notion of quantum-mechanical causality nor a notion of "state of the system at a moment of time." Nor would the argument described above succeed in deriving classical causality. There would be advanced as well as retarded effects.

D. Quantum noise

In Sec. V we showed how the probabilities of decoherent sets of coarse-grained histories of the type under discussion could be thought of as the probabilities of a classical system in which the distinguished coordinates x^a evolve from probabilistically distributed initial conditions according to an equation of motion in the presence of noise. In the case of a factored initial density matrix, initial conditions and noise are separately distributed and the system may be thought as obeying the Langevin equation (5.18),

$$R(t) = \mathcal{E}(t, X(\tau)) + \mathcal{L}(t, X(\tau)) = 0, \quad (7.27)$$

where the total force $R(t)$ is distributed according to the generally nonpositive distributing $g[R(\tau), X(\tau)]$ constructed from the decoherence functional according to Eq. (5.17). The analysis of this section provides an explicit form for $\mathcal{E}(t, X(\tau))$ in Eq. (7.23) and a systematic approximation scheme for the spectrum of the noise.

When decoherence is good enough that the restriction on the range of the ξ integration in Eq. (7.19) arising from the coarse graining can be neglected, a systematic perturbation scheme for the approximate probabilities $p(\alpha)$ can be obtained by expanding the exponent in powers of $\xi(t)$. In the leading approximation (7.22), an explicit expression for the distribution of the Langevin force is obtained, which is

$$\left[\det \left[\frac{K_I}{2\pi} \right] \right]^{-1/2} \exp \left[-\frac{1}{\hbar} \int_0^T dt \int_0^T dt' \mathcal{L}^\dagger(t, X(\tau)) \times K_I^{\text{inv}}(t, t'; X(\tau)) \mathcal{L}(t, X(\tau)) \right]. \quad (7.28)$$

In this leading approximation, the noise is distributed with a *positive* Gaussian distribution function whose spectrum is fixed by the correlation function

$$\begin{aligned} \langle \mathcal{L}(t, X(\tau)) \mathcal{L}(t', X(\tau)) \rangle_c \\ = (\hbar/2) K_I(t, t'; X(\tau)) \\ = \frac{1}{2} \langle \{ \Delta F(t, X(\tau)), \Delta F(t', X(\tau)) \} \rangle. \end{aligned} \quad (7.29)$$

For the linear problem, that approximation is exact [cf. Eq. (6.19)], with the further simplifying feature that the spectrum of Langevin force is independent of the path $X(t)$.

Higher-order terms in the expansion of the exponent of Eq. (7.19) may be regarded as providing corrections to the Gaussian noise. The general expression for the correlation functions is Eq. (5.10). Other corrections to the Gaussian noise arise from the coarse-graining restrictions on the range of ξ integration.

VIII. EXAMPLES AND COMPARISON WITH CLASSICAL CASES

In this section we specialize the general nonlinear theory of the preceding section to some particular cases considered by ourselves and other authors. We begin with the linear models described in Sec. VI.

A. Recovery of the linear theory

In Sec. VI we derived the equations of motion for linear systems defined by an influence phase (6.2) that was quadratic in the coordinates x distinguished by the coarse graining but containing no linear terms. An influence phase of this form will arise when the free action is quadratic in the coordinates distinguished by the coarse graining as in Eq. (6.1) when the interaction with the ignored coordinates Q is linear as in Eq. (6.3), and when the densi-

ty matrix has the special form (6.4). To recover the linear equation of motion (6.14) from the nonlinear (7.20), we need to evaluate $\langle F(t, X(\tau)) \rangle$ under these conditions. Equation (7.25) gives a general expression for $\langle F(t, X(\tau)) \rangle$. We know from the general arguments described in Sec. VI that, when the interaction between distinguished and ignored variables is linear as in Eq. (6.3), the influence phase is quadratic in the x 's and $\langle F(t, X(\tau)) \rangle$ is linear in $X(\tau)$. It therefore suffices to evaluate Eq. (7.25) to linear order in a perturbation expansion in $X(\tau)$; the higher orders must cancel. The result is

$$\langle F(t, X(\tau)) \rangle = \langle f(t) \rangle_0 + \frac{i}{\hbar} \int_0^t dt' \langle [f(t), f(t')] \rangle_0 X(t'). \quad (8.1)$$

Here, the forces $f(t)$ are the Heisenberg picture representatives of the homogeneous, linear functions of the ignored coordinates $f(Q(t))$ [cf. Eq. (7.11)], and the expected values $\langle \rangle_0$ are computed using the time dependence of operators provided by Hamiltonian H_0 of the ignored coordinates Q alone, neglecting their interaction with the x 's [cf. Eq. (6.5)]. If we assume, as we did in Sec. VI, that $\langle f(t) \rangle_0 = 0$ [this may always be achieved by a time-dependent shift in $x(t)$] then we recover both the form of the linear equation of motion (6.14) and the expression for the additional force anticipated in Eq. (6.24).

The derivation of the expression (6.25) for the spectrum of quantum noise in the linear problem is even more straightforward. From the form (7.21) we see that $K_I(t, t')$ is equal to $k_I(t, t')$ of the linear case. Since we assumed $\langle f(t) \rangle$ vanished for the linear problem [cf. Eq. (6.23)], Eq. (7.21) involves only the expected value of $\{f(t), f(t')\}$. However, general arguments for the linear problem show that when the interaction is of the linear form (6.3) this expected value, proportional to $k_I(t, t')$, is *independent* of the x 's. It may thus be evaluated in zeroth order in perturbation theory in $x(t)$; that is, the time dependence of the operators in Eq. (7.21) is provided by the Hamiltonian H_0 with no interactions between Q 's and x 's. This is the time dependence we denoted by a subscript zero in Sec. VI, and thus we recover Eq. (6.25).

B. Semilinear systems

Simple expressions for the spectrum of quantum noise and the kernel of the associated fluctuating force analogous to Eqs. (6.24) and (6.25) have been obtained in more general cases than strict linearity. A particularly simple case occurs when the dependence of the total action on the ignored variables (the Q 's) is restricted to be quadratic, but arbitrary dependence on the distinguished variables (the x 's) is allowed in the potential energy and interaction terms. This case was discussed¹² in purely clas-

sical situations by Zwanzig [38], and in particular field theory examples by Ryang and Saito [39], and more recently in a general survey of such problems by Brun [36], who applies our methods.

Let us consider a specific class of problems. For simplicity, we assume a single distinguished variable x to avoid matrix notation. Suppose that (i) the action of the ignored variables $S_0[Q(t)]$ is a quadratic functional of the Q 's, (ii) the interaction of the Q 's and x 's is of the form

$$S_{\text{int}}[x(\tau), Q(\tau)] = - \int_0^T dt a(x(t)) f(Q(t)), \quad (8.2)$$

where f is homogeneous and linear in the Q 's, but a is not necessarily a linear function of x , and (iii) the density matrix factors as in Eq. (3.8) and has the special form (6.4). Under these conditions the influence phase will be a functional that is at most quadratic in $a(x(t))$, and its form is given by Eq. (6.2) with $x(t)$ and $x'(t)$ replaced by $a(x(t))$ and $a(x'(t))$, respectively, assuming, as before, that linear terms vanish. The equation of motion will then be of the form (7.23), with the force arising from the interaction given by

$$\langle F(t, X(\tau)) \rangle = a'(X(t)) \int_0^t dt' k_R(t, t') a(X(t')), \quad (8.3)$$

where $a' = da/dX$. Here, k_R , as before in Eq. (6.24), is proportional to the expected value of the commutator of $f(Q(t))$ evolved in time without the interaction with the x 's. Linear terms in the influence phase would give rise to a nonzero value of $\langle f(t) \rangle_0$ and an additional term in Eq. (8.3) of the form $a'(X(t)) \langle f(t) \rangle_0$. With the addition of this term, the result (8.3) is equivalent to that of Brun [36], who does not eliminate linear terms in the influence phase. The Langevin force $\mathcal{L}(t)$ is given by the formula

$$\mathcal{L}(t, X(\tau)) = a'(X(t)) l(t), \quad (8.4)$$

where $l(t)$ is a Gaussian random force with a spectrum given by Eqs. (6.20) and (8.4). The Langevin equation describing both dissipation and fluctuations in this approximation is then

$$\mathcal{E}(t, X(\tau)) = \frac{\delta S_{\text{free}}[X(\tau)]}{\delta X(t)} + a'(X(t)) \left[\int_0^t dt' k_R(t, t') a(X(t')) + l(t) \right] = 0. \quad (8.5)$$

A fluctuation-dissipation theorem of the form given by Eq. (6.30) continues to hold.

C. Equations of motion for semilinear classical coarse-grained systems

The evolution of a classical Hamiltonian system is, of course, deterministic and essentially reversible when followed in all detail. In a coarse-grained description, however, such systems will, in general, approximately obey *effective* equations of motion including the effects of dissipative forces and with deviations from these equations produced by classical noise. The classical problem analogous to that considered in this paper would be to derive

¹²We thank T. Brun for these references.

the Langevin equation for the motion of some followed variables $x(t)$ that are interacting with some ignored variables $Q(t)$ whose initial conditions are probabilistically distributed according to some given rule. To our knowledge, this kind of problem has not been considered classically for the nonlinear situations discussed in Sec. VI. However, it has been worked out by Zwanzig [38] for the semilinear systems treated in the preceding subsection. Brun [36] has shown that the Langevin equation deduced from quantum mechanics coincides with Zwanzig's result in the limit $\hbar \rightarrow 0$ as it must. We briefly review Zwanzig's derivation and Brun's demonstration here.

Again, we assume for simplicity a single distinguished variable $x(t)$ interacting with an assembly of oscillators according to Eq. (8.2). The free action $S_0[Q(t)]$ and the function $f(Q(t))$ are given explicitly by Eqs. (6.5) and (6.6), respectively. The classical equation of motion for the Q 's following from Eqs. (6.5) and (8.2) may be solved explicitly with the result

$$Q^A(t) = l^A(t) - \frac{1}{m\omega_A} \int_0^t dt' \sin[\omega_A(t-t')] C_A a(X(t')). \quad (8.6)$$

Here, $l^A(t)$ is the solution of the free oscillator equations following from the action (6.5) with the same initial position and momentum as $Q^A(t)$. This result for $Q^A(t)$ may be substituted into the classical equation of motion for $X(t)$. The result is an equation of motion for X of the form

$$\frac{\delta S_{\text{free}}[X(\tau)]}{\delta X(t)} + a'(X(t)) \left[\int_0^t dt' k_R(t, t') a(X(t')) + l(t) \right] = 0, \quad (8.7)$$

where

$$l(t) = \sum_A C_A l^A(t) \quad (8.8)$$

and we have made use of the fact that the sum of retarded Green's functions from Eq. (8.6) enters Eq. (8.7) as the combination we called $k_R(t, t')$ in Eq. (6.8a).

If the initial conditions for the Q 's are probabilistically distributed, then the motion $X(t)$ will be probabilistically distributed as well. The initial conditions of the Q 's are the initial conditions of the free oscillator motions $l^A(t)$. In Eq. (8.7), the only way these initial conditions enter is through the function $l(t)$ defined in Eq. (8.8). Equation (8.7) may, therefore, be interpreted as a Langevin equation with a stochastic force $a'(X(t))l(t)$. The time dependence of this force is known because the $l^A(t)$ in Eq. (8.8) satisfy the harmonic oscillator equations of motion. When the values of the Q 's and their conjugate momenta are distributed thermally, the correlation functions of this noise are easily calculated. The classical phase-space distribution analogous to the thermal bath used in Refs. [13,16] is

$$\rho^{\text{cl}}(\Pi_0, Q_0) \propto \exp[-H_0(\Pi_0, Q_0)/kT_B], \quad (8.9)$$

where H_0 is the classical Hamiltonian corresponding to S_0 . (A slightly different distribution was assumed by

Zwanzig in Ref. [38] with slightly different results.) The result is Gaussian noise with $\langle l(t) \rangle_c = 0$ and

$$\begin{aligned} \langle l(t)l(t') \rangle_c &= kT \sum_A \frac{C_A^2}{m\omega_A^2} \cos[\omega_A(t-t')] \\ &\equiv \frac{1}{2} k_I^{\text{cl}}(t, t'). \end{aligned} \quad (8.10)$$

Equations (8.7) and (8.10) are essentially the results of Zwanzig [38]. We note that the spectral weights of the spectrum of the fluctuations (8.10) and of the kernel of the dissipative force term (6.8a) as defined by Eq. (6.29) are related by

$$\tilde{k}_I^{\text{cl}}(\omega) = (2kT_B/\omega) \tilde{k}_R(\omega). \quad (8.11)$$

This is the classical fluctuation-dissipation theorem.

As discussed by Brun, the Langevin equation of the quantum-mechanical problem (8.5) coincides in form with Eq. (8.7) derived classically by Zwanzig [38]. The only difference is that the noise spectrum $\hbar k_I(t, t')$ is given by Eq. (6.8b) in the quantum-mechanical case and by Eq. (8.10) in the classical one. Indeed, the two expressions coincide in the classical limit as they must. The classical noise is entirely thermal. When quantum mechanics is taken into account there is quantum noise as well.

IX. MORE GENERAL COARSE GRAININGS

The coarse grainings discussed in the previous sections are limited to those that distinguish a fixed subset of the coordinates q^β . Coarse grainings that realistically describe a quasiclassical domain are not of this simple type. As we have discussed elsewhere [2], it is likely that a quasiclassical domain will be described by, among other things, coarse graining with respect to ranges of values of the averages of densities of conserved or approximately conserved quantities over suitably small volumes. Examples are the densities of energy, momentum, charge, current, nuclear species, etc. Together with field averages, these *are* the "hydrodynamic" variables that enter into the differential equations of classical physics. Sufficiently large volumes would give these variables enough "inertia" to enable them to resist the deviations from predictability caused by the interactions that effect decoherence, as we have described in the earlier sections.

The coarse grainings discussed in this paper must be generalized in two ways to discuss such variables. They must be generalized to allow the original fine-grained description to involve momenta as well as coordinates. They must also be generalized to permit coarse grainings by ranges of values of averaged densities. These correspond to no particular fixed subset of coordinates. In this section we introduce the machinery necessary to consider such coarse grainings, although we do not carry out an analysis of the circumstances in which they decohere or behave quasiclassically.

A. Phase-space coarse grainings

A fairly general class of coarse grained histories may be obtained by considering partitions of the coordinates

q^β at some times and their conjugate momenta π_β at others. More specifically, we consider partitions defined by alternatives at a set of times $\{t_k\}$ by an exhaustive set of exclusive configuration space regions $\{\Delta_{\alpha_k}^k\}$ at some of the times t_k and by an exhaustive set of regions of momentum space $\{\tilde{\Delta}_{\alpha_k}^k\}$ at other of the times t_k . In this section, concerned only with this type of coarse graining, we will reserve the notation $\{P_{\alpha_k}^k(t_k)\}$ for projections of configuration-space alternatives and introduce the nota-

tion $\{\tilde{P}_{\alpha_k}^k(t_k)\}$ for a set of orthogonal projections onto a set of exclusive momentum space regions. The sets of histories we are considering consist of sequences of sets of either P 's or \tilde{P} 's at times t_1, \dots, t_n . The decoherence functional for such a set of histories is given generally by Eq. (4.5). We now show that it has a sum-over-histories representation by path integrals in *phase space*.

Utilizing complete sets of coordinates eigenstates, we may write the decoherence functional (4.5) in the form

$$D(\alpha', \alpha) = \int dq'_f \int dq_f \int dq'_0 \int dq_0 \delta(q'_f - q_f) \langle q'_f T | C_{\alpha'} | q'_0 0 \rangle \rho(q'_0, q_0) \langle q_0 0 | C_{\alpha}^\dagger | q_f T \rangle. \quad (9.1)$$

The matrix elements of the C_α may be written as the compositions of sequences of propagators between definite coordinate or momentum eigenstates. For example, if $\tilde{P}_{\alpha_1}^1(t_1)$ is a projection onto a momentum region $\tilde{\Delta}_{\alpha_1}^1(t_1)$ and $P_{\alpha_2}^2(t_2)$ onto a coordinate region $\Delta_{\alpha_2}^2(t_2)$ we can write

$$\langle q_f T | C_\alpha | q_0 0 \rangle = \langle q_f T | P_{\alpha_2}^2(t_2) \tilde{P}_{\alpha_1}^1(t_1) | q_0 0 \rangle \quad (9.2)$$

in the Schrödinger picture as

$$\int_{\Delta_{\alpha_2}^2} dq_2 \int_{\tilde{\Delta}_{\alpha_1}^1} d\pi_1 \langle q_f T | q_2 t_2 \rangle \langle q_2 t_2 | \pi_1 t_1 \rangle \langle \pi_1 t_1 | q_0 0 \rangle. \quad (9.3)$$

The propagators in Eq. (9.3) may be represented as phase-space path integrals if the Hamiltonian $H(\pi, q)$ associated with the action $S[q(t)]$ is of a suitably simple form [39,40]. In particular, they can be so represented if $H(\pi, q)$ is a sum of a function of the π 's and a function of the q 's. For example,

$$D(\alpha', \alpha) = \int_{\alpha'} \delta\mu' \int_{\alpha} \delta\mu \delta(q'_f - q_f) \exp(i\{S[\pi'(\tau), q'(\tau)] - S[\pi(\tau), q(\tau)]\} / \hbar) \rho(q'_0, q_0). \quad (9.7)$$

The integral is over phase-space paths restricted by the coarse graining. For example, the integral over $\pi(t)$ and $q(t)$ is over the phase-space paths that thread the intervals in either coordinate or momentum space corresponding to the history α . There is no integration over the initial momenta π'_0 and π_0 but unrestricted integrations over the final momenta π'_f and π_f . Again, the details of this and the measure are in Sec. A of the Appendix.

In usual cases where the Hamiltonian in Eq. (9.5) is quadratic in all momenta, the momentum dependence in the integrand in Eq. (9.7) is that of a Gaussian. For those coarse grainings that restrict only the coordinates q^β and ignore the momenta π_β , the integrals over momenta may be carried out explicitly. The result is the Lagrangian

$$\langle q'' t'' | q' t' \rangle = \int \delta\mu \exp\{iS[\pi(t), q(t)] / \hbar\}, \quad (9.4)$$

where $S[\pi, q]$ is the canonical form of the action:

$$S[\pi, q] = \int_{t'}^{t''} dt [\pi_\beta(t) \dot{q}^\beta(t) - H(\pi(t), q(t))] \quad (9.5)$$

and of course,

$$\langle \pi'' t'' | q' t' \rangle = \int \frac{dq''}{(2\pi\hbar)^{1/2}} e^{-i\pi'' q'' / \hbar} \langle q'' t'' | q' t' \rangle. \quad (9.6)$$

The integral is over phase-space paths between t' and t'' weighted by the invariant Liouville measure and restricted by the conditions that they intersect π' at t' and q'' at t'' . The details of these integrations are spelled out in Sec. A of the Appendix.

By inserting Eq. (9.5) into Eq. (9.4) and Eqs. (9.6) and (9.4) into Eq. (9.2), one arrives at a sum-over-histories form for the decoherence functional for phase-space coarse grainings:

path integral for the decoherence functional (3.1). Indeed, it is by this route that the measure in that path integral is usually derived from the canonical, Liouville, " $dp dq / (2\pi\hbar)$ " measure on paths in phase space.

There is no obstacle to letting the time of a coarse graining by momenta coincide with that of a coarse graining by coordinates. Even in the quantum mechanics of measured subsystems it is possible to consider a measurement of position followed after an arbitrarily short time interval by a measurement of momentum. Care must be taken, however, to specify the order of the coarse grainings when two such times coincide. Since the corresponding operators do not commute, a projection on a range of momentum at one time followed immediately by

a projection on a range of position defines a different history from one with the operators in the opposite time order.

With this machinery in hand, we may now consider phase-space coarse grainings, analogous to those of Sec. III, in which the phase-space coordinates (π_β, q^β) are divided into canonically conjugate pairs (p_a, x^a) that are distinguished by the coarse graining while the remaining pairs (Π_A, Q^A) are ignored. A simple, interesting, and important class of models is obtained by assuming that the action decomposes according to Eq. (3.2) with S_{free} and S_0 quadratic in time derivatives and the interaction independent of all time derivatives. (A slight generalization would then be needed to deal with a system of parti-

cles interacting electromagnetically.) The decomposition of the Hamiltonian corresponding to these assumptions is

$$H(\pi, q) = H_{\text{free}}(p, x) + H_0(\Pi, Q) + H_{\text{int}}(x, Q), \quad (9.8)$$

where H_{free} and H_0 have a quadratic momentum dependence. The simplifying consequence of these assumptions is that the Gaussian integrals over the momenta Π_A in Eq. (9.7) may all be carried out explicitly. The remaining integrals over the Q^A have the same form as they do in the Lagrangian path integral and may be summarized by a single influence phase $W[x'(t), x(t)]$ defined by Eq. (3.3). The decoherence functional for phase-space coarse grainings may then be written

$$D(\alpha', \alpha) = \int_{\alpha'} d\mu' \int_{\alpha} d\mu \delta(x_f - x'_f) \exp(i\{S_{\text{free}}[p'(\tau), x'(\tau)] - S_{\text{free}}[p(\tau), x(\tau)] + W[x'(\tau), x(\tau)]\} / \hbar) \bar{\rho}(x'_0, x_0), \quad (9.9)$$

with $\bar{\rho}$ defined by Eq. (3.4).

Equation (9.9) shows that coarse grainings in which a fixed set of coordinates or their conjugate momenta are followed at a sequence of times while all others are ignored may be studied by path integral techniques. The form of Eq. (9.5), however, allows an immediate and important distinction to be drawn between coordinate coarse grainings and momentum coarse grainings. Coordinate alternatives will decohere if W has a positive imaginary part that becomes large as $x'(t)$ and $x(t)$ are increasingly distinct. However, there is no corresponding mechanism leading to the decoherence of momentum alternatives for this class of models.

In the case of the linear models discussed in Sec. VI, this conclusion may be made more precise by following a few steps that led to the derivation of the equations of motion. Introduce variables for the momenta analogous to those for the coordinates in Eq. (5.4):

$$\varpi(t) = p'(t) - p(t), \quad (9.10a)$$

$$P(t) = \frac{1}{2}[p'(t) + p(t)]. \quad (9.10b)$$

The exponent in Eq. (9.7) may now be reexpressed in terms of the variables of Eqs. (5.4) and (9.7) and, after a few integrations by parts, put in the form

$$\begin{aligned} & S_{\text{free}}[p'(\tau), x'(\tau)] + S_{\text{free}}[p(\tau), x(\tau)] + W[x'(\tau), x(\tau)] \\ &= -\xi_0^\dagger P_0 + \int_0^T dt \left[\varpi^\dagger(t) \left[\dot{X}(t) - \frac{\partial H}{\partial P(t)} \right] + \xi^\dagger(t) \left[-\dot{P}(t) + \frac{\partial H}{\partial X(t)} + \int_0^T dt' k_R(t, t') X(t') \right] \right] \\ &+ \frac{i}{4} \int_0^T dt \int_0^T dt' \xi^\dagger(t) k_I(t, t') \xi(t'). \end{aligned} \quad (9.11)$$

The terms in square brackets on the right-hand side of Eq. (9.11) are Hamilton's equations of motion augmented by terms describing the additional forces arising from the interaction of the (p, x) subsystem with the rest. However, we cannot conclude that these equations of motion are valid for phase-space coarse grainings. The last term in Eq. (9.11) makes the integrand of Eq. (9.9) small when $\xi \neq 0$ and thereby enforces the decoherence of coordinate alternatives. However, the absolute value of that integrand is *uniformly* distributed in ϖ . Unless the integration over the other variables makes Eq. (9.11) small for $\varpi \neq 0$ there will be no decoherence of momentum alterna-

tives.¹³ In cases when momentum alternatives cannot even be assigned probabilities there is *a fortiori* no issue

¹³Such cancellation leading to the decoherence of momentum alternatives occurs, for example, in the case of a free, nonrelativistic particle. The conservation of momentum means that projections onto ranges of momentum at different times commute and a history composed exclusively of such projections will automatically decohere. We thank J. Halliwell for pointing this out to us.

of whether the probability is high for their correlation in time by an effective equation of motion. The origin of this distinction between coordinates and momenta in this model is, as has been remarked by many authors, that the interaction Hamiltonian is local in coordinates but not in momentum.

This analysis of both coordinate and momentum coarse grainings stresses an important if familiar point. Two coarse grainings that would be essentially indistinguishable classically may have very different properties quantum mechanically. The present example illustrates this with coarse grainings by momentum, $p(t)$, and by the difference in coordinates at two nearby times $M[x(t+\epsilon)-x(t)]/\epsilon$. These two types of coarse graining may have essentially different properties with respect to decoherence even though they would be indistinguishable on the basis of classical physics when the decoherence time-scale is much shorter than characteristic dynamical time scales of the classical motion. Evidently, considerable care is required in identifying the variables through which quasiclassical behavior is to be defined in quantum mechanics.

B. Densities

The most important generalization of the coarse grainings studied in this paper is to cases where the variables that are distinguished are not limited to a fixed set of fundamental coordinates. To coarse grain by the value of the baryon number in a small spatial volume, for example, is not the same as following some particular subset of the fields of a fundamental field theory. We shall discuss

how, in principle, more general and realistic coarse grainings can be treated by techniques analogous to those used to discuss the special cases of this paper. Our considerations are essentially formal and we have not pushed the analysis far enough to derive equations of motion in these realistic cases. Our discussion, however, indicates a route by which that might be accomplished.

A very general and useful class of coarse grainings is obtained by partitioning the fine-grained histories according to the values of functionals of them. To illustrate an interesting case in a manageable notation, we consider a field theory with a single charged scalar field $\phi(\mathbf{x}, t)$. A set of functionals leading to coarse grainings relevant for the present discussion consists of the values of the charge density at time t averaged over a small spatial volume V at different spatial points \mathbf{x} :

$$J^0(\mathbf{x}, t) = \frac{1}{V} \int_V d^3\xi j^0(\mathbf{x} + \xi, t) \quad (9.12)$$

where $j^0(\mathbf{x}, t)$ is the charge density expressed in terms of $\phi(\mathbf{x}, t)$. The fine-grained histories, $\phi(\mathbf{x}, t)$, may be partitioned into exhaustive and exclusive classes by the values of $J^0(\mathbf{x}, t)$. (Partition by *ranges* of values of these averaged densities would be a further coarse graining.) Consider a particular value $v(\mathbf{x}, t)$. (The conventional use of ρ for charge density is precluded by its use here for the density matrix.) The coarse-grained history corresponding to $v(\mathbf{x}, t)$ consists of all those $\phi(\mathbf{x}, t)$ for which the integral (9.12) has this value.

The decoherence functional for a pair of histories coarse-grained by particular values of the charge density is

$$D[v'(\mathbf{x}, t), v(\mathbf{x}, t)] = \int \delta\phi' \int \delta\phi \delta[\phi'(\mathbf{x}, T) - \phi(\mathbf{x}, T)] \delta[J^0(\mathbf{x}, t) - v'(\mathbf{x}, t)] \\ \times \exp((i/\hbar)\{S[\phi'(\mathbf{x}, t)] - S[\phi(\mathbf{x}, t)]\}) \delta[J^0(\mathbf{x}, t) - v(\mathbf{x}, t)] \rho[\phi'(\mathbf{x}, 0), \phi(\mathbf{x}, 0)]. \quad (9.13)$$

The first δ functional in Eq. (9.13) enforces the coincidence of the histories at the final time T as in Eq. (9.1). The other two δ functionals restrict the fields in the integral to the coarse-grained histories labeled by $v'(\mathbf{x}, t)$ and $v(\mathbf{x}, t)$. Of course, making precise sense of a formal expression like Eq. (9.13) raises many mathematical issues that we shall not pursue here.

The decoherence functional $D[v'(\mathbf{x}, t), v(\mathbf{x}, t)]$ is the generalization of $D[x'(t), x(t)]$ considered in the simple models of the earlier sections of this paper. In those cases a fixed set of variables describing the fine-grained histories of a single oscillator were distinguished by the coarse graining while the variables describing the other oscillators were ignored. In the case of a coarse graining by the average value of a density over a volume, one could loosely say that the "variables" describing the different field configurations internal to the volume and consistent with a given average value are ignored. Hence

there is no subset of the fundamental fields or linear transformation of them that describes these "internal variables." That is why $D[v'(\mathbf{x}, t), v(\mathbf{x}, t)]$ is a generalization of $D[x'(t), x(t)]$.

The form of $D[v'(\mathbf{x}, t), v(\mathbf{x}, t)]$ may be brought in closer analogy with Eq. (3.5) by representing some of the δ functionals in Eq. (9.13) as exponential integrals:

$$\delta[J^0(\mathbf{x}, t) - v(\mathbf{x}, t)] \\ = \int \delta y \exp \left[i \int d^4x y(\mathbf{x}, t) [J^0(\mathbf{x}, t) - v(\mathbf{x}, t)] \right]. \quad (9.14)$$

Then, Eq. (9.13) becomes

$$\begin{aligned}
D[v'(\mathbf{x}, t)v(\mathbf{x}, t)] = & \int \delta y' \int \delta y \int \delta \phi' \int \delta \phi \delta[\phi'(\mathbf{x}, T) - \phi(\mathbf{x}, T)] \\
& \times \exp \left[\frac{i}{\hbar} \left[S[\phi'(\mathbf{x}, t)] - S[\phi(\mathbf{x}, t)] + \int d^4x y'(\mathbf{x}, t)[J'^0(\mathbf{x}, t) - v'(\mathbf{x}, t)] \right. \right. \\
& \left. \left. - \int d^4x y(\mathbf{x}, t)[J^0(\mathbf{x}, t) - v(\mathbf{x}, t)] \right] \right] \rho[\phi'(\mathbf{x}, 0), \phi(\mathbf{x}, 0)] . \quad (9.15)
\end{aligned}$$

The integral in Eq. (9.15) is of a familiar form in a field theory with sources and should be accessible to standard approximation techniques. Of course, to study the putative “hydrodynamic” variables of a quasiclassical domain, coarse graining by the densities of other approximately conserved quantities such as energy, momentum, baryon number, as well as the averages of long-range fields, must be considered. Once the decoherence functional is calculated from expressions such as Eq. (9.14), the circumstances in which such coarse-grained sets of histories decohere can be investigated and their equations of motion derived by the methods of Sec. VII. An important question will be the closure of any such set of equations of motion. However, we have not yet progressed beyond this formal sketch of the route to a derivation of the classical equations of motion for realistic quantum systems.

X. CONCLUSIONS AND PROGRAM

Instead of merely summarizing the results obtained in this article and the future directions of research sketched in the last section, let us relate them to the program (of understanding quantum mechanics and its relation to quasiclassical experience) in which we have been engaged, and which involves many elements elucidated by other authors over the last 35 years. We shall briefly review that program and where the present work fits into it.

We start with the quantum mechanics of a closed system representing the universe (with a deliberately simplified treatment of the complications caused by quantum gravity). There is a dynamical theory of all the elementary particles and their interactions, which we take to be described by an action or a Lagrangian or a Hamiltonian H , and also a Heisenberg density matrix (or initial density matrix) ρ . The question is how the quasiclassical domain of familiar experience comes to be an emergent feature of the system characterized by H and ρ .

A “measurement situation” can often be characterized as one in which some variable comes into strong correlation with the quasiclassical domain. The quasiclassical domain also permits, for certain probability tracks, certain spatial regions, and certain epochs of time, enough classical predictability for the evolution of complex adaptive systems that learn, observe, and record, and also utilize, in some approximation, the probabilities assigned by quantum mechanics, on the basis of ρ and H , to different alternative coarse-grained histories. (Presumably, it is the quasiclassical correlations, representing near certainties, that are easiest to utilize.) Observations by such

complex adaptive systems, which we call IGUS’s (information gathering and utilizing systems) when they are functioning as observers, are then considered to be actual measurements.

A quasiclassical domain is defined by decoherence, a measure of classical predictability, and some sort of maximality (such as what we have called “fullness” [4]). The first requirement is decoherence, that is, enough coarse graining of the alternative histories of the universe so that there is, exactly or to a very good approximation, no interference between the alternative coarse-grained histories, as measured by the decoherence functional.

Decoherence requires coarse graining that goes far beyond the modest requirement imposed at each instant of time by the uncertainty principle. The mechanism of decoherence involves the loss of phase information as a result of the coarse graining and is associated with noise that inextricably combines quantum fluctuations and classical statistical fluctuations, both necessary for the decoherence.

All of that is most obvious, of course, in the instructive, although oversimplified models in which the coarse grainings include an average over some fixed set of variables (the ignored ones) while following others (the distinguished ones) more or less coarsely. The ignored variables, through their interactions with the distinguished ones, carry away the phases and are responsible for the fluctuations.

The fluctuations, of course, cause departures from any effective (or phenomenological) classical equations of motion for the distinguished variables. But a high degree of decoherence requires very large fluctuations, which threaten to produce very great departures from classical predictability. Thus, for the coarse graining to yield a quasiclassical domain, it is essential that the distinguished variables carry very high inertia so as to resist most of the large fluctuations and follow the effective equations of motion with only small deviations over long stretches of time and with only occasional large ones. That high inertia is achieved by even much coarser graining than was required for the decoherence itself.

The requirement of fullness, a kind of maximality, was discussed in Ref. [3]. For the coarse graining defining a quasiclassical domain to be an emergent feature of the universe characterized by H and ρ rather than an artifact chosen by some IGUS, it should be as refined a description of the universe as possible consistent with the requirements of decoherence and quasiclassicality. In Ref. [3] we proposed, for the case of perfect decoherence, the notion of a “full” set of alternative coarse-grained his-

tories to capture this idea of maximal refinement. Any “full” set belongs to an equivalence class defined by a basis in Hilbert space, provided the density matrix ρ corresponds to a pure state. (If it does not, and especially if it has a great many eigenstates with nonzero eigenvalues, then the condition of “strong decoherence” discussed in Ref. [3] is too strong and the associated discussion of maximality requires modification.)

We have posed the question as to whether there could be various kinds of essentially inequivalent quasiclassical domains or whether any quasiclassical domain is more or less equivalent to any other. The former case poses some challenging intellectual puzzles, especially if we imagine IGUS’s evolving in relation to each of the essentially inequivalent quasiclassical domains.

There is, of course, an indication from our everyday experience of some of the features of a particular quasiclassical domain. It involves distinguished variables that are more or less hydrodynamic in character; they are integrals or averages over small regions of space of conserved or nearly conserved densities and of fields coupled to those densities. The regions must be large enough to produce sufficient inertia to resist most of the fluctuations associated with the coarse graining and small enough to implement the requirements of “fullness.” It is clear, however, that the distinguished variables cannot be defined once and for all, but depend on history. For example, the suitable hydrodynamic variables under the conditions that prevailed before the condensation of the solar system involved much bigger volumes than those that were suitable inside the planets after they were formed. That is why the models in which the ignored variables are fixed for all time are only instructive examples and not general enough for the realistic case.

Our program thus aims at describing a quasiclassical domain with history-dependent distinguished variables resembling hydrodynamic ones and obeying effective classical equations of motion apart from small fluctuations and occasional large ones, some of which result in the need for redefinition (for later times) of the distinguished variables.

In this article we have, for the most part, confined ourselves to a model in which the distinguished variables are separated once and for all from the ignored ones and are also, unlike the hydrodynamic ones, coordinate variables of the fundamental theory, like modes of a scalar field, with kinetic energy bilinear in the time derivatives and the rest of the energy not involving the time derivatives. We have also restricted attention to initial density matrices that factor into the product of a density matrix for the distinguished variables and one for the ignored ones. With those simplifying assumptions, we identify the effective classical equations of motion and the Gaussian part of the noise that disturbs them.

We divided into two parts the quantum-mechanical process of prediction for further coarse grainings of the histories of the distinguished system. First, there is the calculus of amplitudes for histories fine-grained in the distinguished variables and of the bilinear combinations of amplitudes that define the decoherence functional. Second, there are the requirements of decoherence for

deriving consistent probabilities from these amplitudes for histories that are further coarse-grained.

We assumed that the requirements of decoherence were satisfied (although we discussed mechanisms by which this happens) and examined when the probabilities for histories favored classical predictability.

We started with a general formulation that applies to fully nonlinear systems. By performing a functional Fourier transform on the decoherence functional with respect to the difference in the distinguished coordinates on the left and right, we introduced a distribution function for the “total force” (including the inertial term) that acts on the distinguished system along its history. Like the analogous Wigner distribution at a moment of time, it is not generally positive.

Utilizing the distribution of the “total force,” we were able, even in nonlinear cases, to represent the probabilities of a decoherent set of coarse-grained histories as the probabilities of a classical dynamical system governed by a Langevin equation incorporating history dependent noise. The initial conditions of this classical system are distributed according to the Wigner distribution and the noise according to the distribution of “total force” mentioned above. Both distributions are generally nonpositive, which distinguishes them from classical ones, although the resulting probabilities for decohering sets of histories are, of course, positive. When the noise is small compared with the inertial term in the equation of motion, the coarse-grained histories are classically deterministic.

As suggested by the simple physical picture of phases being carried away by the interaction of the distinguished subsystem with the ignored variables, there is a connection between decoherence and noise. This was exhibited by writing the decoherence functional for the distinguished variables as an integral of an average density matrix times an exponential, and expanding the argument of the exponential in powers of the generalized vector $\xi(t)$ that measures the difference between the distinguished variables as functions of time on the left- and right-hand sides of the decoherence functional. After a partial integration, the linear term has the form $i/\hbar \int \xi^\dagger(t) \mathcal{E}(t) dt$ and the equation $\mathcal{E}(t) = 0$ is precisely the effective classical equation of motion for $X(t)$, which is the average of the distinguished coordinate variables on the left and on the right. The second term in the expansion of the argument of the exponential is

$$-(1/4\hbar) \int \int \xi^\dagger(t) K_I(t, t') \xi(t') dt dt' ,$$

where K_I is the positive kernel describing the Gaussian noise, and can be thought of as the self-correlation [depending in general on $X(t)$] of a Langevin force $\mathcal{L}(t)$ added to the equation of motion.

There are, in general, higher terms in the expansion, corresponding to the fact that coarse-grained quantum mechanics is not exactly equivalent to effective classical equations of motion accompanied by Gaussian noise. However, we are interested in the case of approximate decoherence of the coarse-grained histories, meaning that $\xi(t)$ is mostly confined to very small values. When K_I is

very large, this is achieved, provided that the higher terms are not somehow still more important. However, large K_I means large noise and that is the connection between decoherence and noise. We, therefore, also impose the requirement of a very large inertia matrix M , so that the huge noise is mostly resisted and the effective classical equations of motion are followed with high probability by the distinguished variables.

For pedagogical reasons, we started by reviewing the completely linear case, studied extensively by Feynman and Vernon [16], Caldeira and Leggett [13], Unruh and Żurek [40], etc. There the noise correlation $k_I(t, t')$ is just a numerical function and the effective equation of motion for $X(t)$ is linear, with a frictional force

$$\int k_R(t, t')X(t')dt'$$

that is, in general, nonlocal in time. When the effective density matrix is diagonal in the energy of the ignored variables, then the numerical functions k_I and k_R depend on $t-t'$ only and are related to each other by the fluctuation-dissipation theorem. In the Fokker-Planck limit, k_I is proportional to $\delta(t-t')$ and k_R to $\delta'(t-t')$ and the coefficients are related by a very simple fluctuation-dissipation relation.

We then generalized to the nonlinear case discussed above. *That is the principal content of this paper, and the work can be applied to many problems, as in Refs. [1] and [36].* We also treated more general coarse grainings, in which both generalized coordinates and generalized momenta are utilized and what had been a Lagrangian formulation turns into what is essentially a Hamiltonian formulation.

Finally, we pointed out the desirability of removing the remaining two unrealistic conditions, that the distinguished variables have their time derivatives occurring in the action only in a bilinear kinetic energy term, and that the distinguished variables be defined independent of history. We are studying the mathematical problems posed by these generalizations. If the first difficulty can be overcome, then one should be able to treat the effective classical equations of motion and the Langevin force for distinguished variables that include the hydrodynamic quantities of familiar quasiclassical experience.

If the second and greater difficulty can be overcome, and history dependence introduced into the coarse graining, then we may begin to tackle the deep problem of introducing individuality into quantum mechanics. Actual alternative histories deal, of course, in large part with individual objects such as our galaxy, the Sun, the Earth, biological organisms on the Earth, and so forth. Yet discussions of quantum mechanics up to now have typically treated such individual objects only as external systems, labeled as "observers" and "pieces of apparatus." If history dependence can be properly introduced into the explicit treatment of quantum mechanics, then we may be able to handle individuality with the care that it deserves.

ACKNOWLEDGMENTS

Part of this work was carried out while the authors were in residence at the Aspen Center for Physics. The

work of M.G.-M. was supported in part by DOE Contract No. DE-AC-03-81 ER40050 and by the Alfred P. Sloan Foundation. He also gratefully acknowledges support from the Air Force Office of Scientific Research under the University Resident Research Program for research performed at Phillips Laboratory (PL/OLAL). The work of J.B.H. was supported in part by NSF Grant No. PHY90-08502.

APPENDIX: PATH INTEGRALS

In the body of this paper we have formally manipulated both configuration space and phase-space path integrals. We define those integrals more carefully in this Appendix and use the definitions to show that the formal manipulations we have used are legitimate.

A. Phase-space integrals

We begin by deriving an explicit expression for the most general path integral occurring in the decoherence functional for sets of histories coarse grained both by regions of configuration space and regions of momentum space, as discussed in Sec. IX. To keep the notation manageable, we shall assume for the moment that we are considering a one-dimensional system with coordinate q , momentum π , and a Hamiltonian of the form

$$H(\pi, q) = \frac{\pi^2}{2M} + V(q, t). \quad (\text{A1})$$

The generalization to larger-dimensional configuration spaces is obvious. A set of histories consisting of chains of just two projections, one on momentum space region $\{\tilde{\Delta}_{\alpha_1}\}$ at time t_1 and the other on configuration space region $\{\Delta_{\alpha_2}\}$ at time t_2 , will suffice to illustrate the construction in more general situations. The important matrix elements for the construction of the decoherence functional are of the form (9.2). They may be expressed in the Schrödinger picture as

$$\begin{aligned} \langle q_f T | C_\alpha | q_0 0 \rangle &\equiv \langle q_f T | P_{\alpha_2}^q(t_2) P_{\alpha_1}^\pi(t_1) | q_0 0 \rangle \\ &= \langle q_f | e^{-iH(T-t_2)/\hbar} P_{\alpha_2}^q e^{-iH(t_2-t_1)/\hbar} \\ &\quad \times P_{\alpha_1}^\pi e^{-iHt_1/\hbar} | q_0 \rangle. \end{aligned} \quad (\text{A2})$$

Here we have written $P_{\alpha_2}^q$ and $P_{\alpha_1}^\pi$ to recall explicitly that we are dealing with projections onto a configuration space region Δ_{α_2} and on a momentum space region $\tilde{\Delta}_{\alpha_1}$, respectively.

We now cast the right-hand side of Eq. (A2) into a phase-space path integral form. We shall be brief because the construction is a standard one [41,42]. We divide the interval $[0, T]$ up into N equally spaced time slices $\tau_0=0, \tau_1, \tau_2, \dots, \tau_{N-1}, \tau_N=T$ with an interval $\epsilon=T/N$ between them. We assume that the times t_1 and t_2 coincide with two of these slices for a sequence of N 's tending to infinity. Let K_1 and K_2 be the labels of the slices corresponding to t_1 and t_2 , respectively, understanding that these are functions of N . Write the propagators in Eq.

(A2) as the product of an appropriate number of factors of $\exp(-iH\epsilon/\hbar)$. Between these factors, on each time slice except the first, insert a resolution of the identity of the form

$$\int d\pi_k \int dq_k |q_k\rangle \langle q_k | \pi_k \rangle \langle \pi_k | = I. \quad (\text{A3})$$

The result is the following expression for the matrix element in (A2):

$$\begin{aligned} \langle q_f T | C_\alpha | q_0 0 \rangle &= \int \prod_{k=1}^N d\pi_k dq_k \langle q_f | q_N \rangle \langle q_N | \pi_N \rangle \langle \pi_N | e^{-iH\epsilon/\hbar} | q_{N-1} \rangle \\ &\quad \times \langle q_{N-1} | \pi_{N-1} \rangle \langle \pi_{N-1} | e^{-iH\epsilon/\hbar} | q_{N-2} \rangle \cdots \langle q_{K_2} | P_{\alpha_2}^q | \pi_{K_2} \rangle \langle \pi_{K_2} | e^{-iH\epsilon/\hbar} | q_{K_2-1} \rangle \cdots \\ &\quad \times \langle q_{K_1} | P_{\alpha_1}^\pi | \pi_{K_1} \rangle \langle \pi_{K_1} | e^{-iH\epsilon/\hbar} | q_{K_1-1} \rangle \cdots \langle q_1 | \pi_1 \rangle \langle \pi_1 | e^{-iH\epsilon/\hbar} | q_0 \rangle. \end{aligned} \quad (\text{A4})$$

Now note the relations

$$\langle q | \pi \rangle = e^{i\pi q/\hbar} / (2\pi\hbar)^{1/2}, \quad (\text{A5})$$

$$\langle q | P_{\alpha_2}^q | \pi \rangle = e_{\alpha_2}(q) e^{i\pi q/\hbar} / (2\pi\hbar)^{1/2}, \quad (\text{A6})$$

$$\langle q | P_{\alpha_1}^\pi | \pi \rangle = e_{\alpha_1}(\pi) e^{i\pi q/\hbar} / (2\pi\hbar)^{1/2}, \quad (\text{A7})$$

where $e_\alpha(x)$ is unity for x in the interval α and zero outside it. (The symbol π stands for momentum unless it occurs in the combination $2\pi\hbar$.) Further, to first order in small ϵ ,

$$\langle \pi | e^{-iH\epsilon/\hbar} | q \rangle \approx e^{-iH(\pi,q)\epsilon/\hbar} e^{-i\pi q/\hbar} / (2\pi\hbar)^{1/2}, \quad (\text{A8})$$

where $H(\pi, q)$ is the function given by Eq. (A1). Inserting these relations in Eq. (A4), noting that there is a δ function that identifies q_N and q_f , we find the following expression, which is exact as $N \rightarrow \infty$:

$$\begin{aligned} \langle q_f T | C_\alpha | q_0 0 \rangle &= \lim_{N \rightarrow \infty} \int \frac{d\pi_N}{2\pi\hbar} \left[\sum_{k=1}^{N-1} \frac{d\pi_k dq_k}{2\pi\hbar} \right] e_{\alpha_2}(q_{K_2}) e_{\alpha_1}(\pi_{K_1}) \\ &\quad \times \exp \left\{ \frac{i}{\hbar} \sum_{j=1}^N \epsilon \left[\pi_j \left[\frac{q_j - q_{j-1}}{\epsilon} \right] - H(\pi_j, q_{j-1}, t_{j-1}) \right] \right\}. \end{aligned} \quad (\text{A9})$$

This is the definition of the phase-space path integral that we have written in Sec. IX as

$$\langle q_f T | C_\alpha | q_0 0 \rangle = \int_\alpha \delta\pi \delta q \exp \left[\frac{i}{\hbar} \int_0^T dt [\pi(t)\dot{q}(t) - H(\pi(t), q(t), t)] \right]. \quad (\text{A10})$$

It is an integral over phase-space paths in the class specified by the coarse graining, that is, over paths which pass through the momentum space region $\tilde{\Delta}_{\alpha_1}$ at time t_1 and configuration-space region Δ_{α_2} at time t_2 .

B. Path integrals for the decoherence functional

When the coarse graining is only by regions of configuration space, and there is no coarse graining by momentum space, then the Gaussian integrals over the π_k , $k=1, \dots, N$ may be carried out in expressions like Eq. (A9) leading to Lagrangian path integrals for the matrix elements $\langle q_f T | C_\alpha | q_0 0 \rangle$ corresponding to individual histories. For example, a history defined by a sequence of q intervals $\Delta_{\alpha_1}, \Delta_{\alpha_2}, \dots, \Delta_{\alpha_n}$ at times t_1, \dots, t_n would have

$$\begin{aligned} \langle q_f T | C_\alpha | q_0 0 \rangle &= \lim_{N \rightarrow \infty} \int \left[\prod_{k=1}^{N-1} \left[\frac{M}{2\pi i \epsilon \hbar} \right]^{1/2} dq_k \right] e_{\alpha_n}(q_{K_n}) \cdots e_{\alpha_1}(q_{K_1}) \\ &\quad \times \exp \left\{ \frac{i}{\hbar} \sum_{j=0}^{N-1} \epsilon \left[\frac{1}{2} M \left[\frac{q_{j+1} - q_j}{\epsilon} \right]^2 - V(q_j, t_j) \right] \right\}. \end{aligned} \quad (\text{A11})$$

Here, K_i is the label of the slice corresponding to t_i and, as before, $q_N = q_f$.

Equation (A11) is the sum-over-paths usually written

$$\langle q_f T | C_\alpha | q_0 0 \rangle = \int_{[q_0, \alpha, q_f]} \delta q \exp \{ iS[q(\tau)]/\hbar \}, \quad (\text{A12})$$

where $S[q(t)]$ is the Lagrangian form of the action corresponding to the Hamiltonian (A1):

$$S[q(\tau)] = \int_0^T dt \left[\frac{1}{2} M \dot{q}^2 - V(q, t) \right]. \quad (\text{A13})$$

In an even more compact notation, where, as in Eq. (3.1), there is an integration over q_0 and q_f , we have denoted the restrictions on the range of integration arising from the coarse graining by an unadorned subscript α on the integral sign.

The path integral for the decoherence functional consists of two multiple integrals like Eq. (A11) over two polygonal paths $\{q_k\}$ and $\{q'_k\}$ with additional integrals over their initial and final end points weighted by the initial density matrix and final δ function, respectively. When, as discussed in Sec. III, the action is of a suitable form and the coarse graining constrains only a fixed subset x^a of the variables of configuration space, the integrals over the remaining $\{Q_k^A\}$ may be carried out yielding a path integral involving an influence phase (3.5). To better understand how formal manipulations are carried out on that path integral we shall now write out the explicit time-slicing implementation of it following from Eq. (A11). To keep the notation manageable we shall

consider the case where the coarse graining refers to a single coordinate x and limit attention to the explicitly linear problem discussed in Sec. VI having a quadratic influence phase given by Eq. (6.2). The generalizations of this case should be obvious. The integral (3.5) for the decoherence functional is then explicitly

$$D(\alpha', \alpha) = \lim_{N \rightarrow \infty} \int \left[\left[\frac{M}{2\pi i \epsilon \hbar} \right]^{N-1} \prod_{k=0}^N dx'_k dx_k \right] \times \delta(x'_N - x_N) E_{\alpha'\alpha}(x'_j, x_j) \times \exp[iA(x'_j, x_j)/\hbar] \bar{\rho}(x'_0, x_0), \quad (\text{A14})$$

where the functions $E_{\alpha'\alpha}$ and A are defined as follows: The function $E_{\alpha'\alpha}$ is

$$E_{\alpha'\alpha}(x'_j, x_j) = \prod_{i=1}^n e_{\alpha'_i}(x'_{K_i}) e_{\alpha_i}(x_{K_i}) \quad (\text{A15})$$

and enforces the constraints of the coarse graining. The exponent A is the discrete form of Eqs. (6.1) and (6.2), viz.

$$A(x'_j, x_j) = \sum_{j=1}^N \epsilon \left[\frac{1}{2} M \left[\frac{x'_j - x'_{j-1}}{\epsilon} \right]^2 - \frac{1}{2} K x_j'^2 - \frac{1}{2} M \left[\frac{x_j - x_{j-1}}{\epsilon} \right]^2 + \frac{1}{2} K x_j^2 \right] + \frac{1}{2} \sum_{j=1}^N \sum_{l=1}^j \epsilon^2 (x'_j - x_j) [k_R(j, l)(x'_j + x_l) + ik_I(j, l)(x'_j - x_l)]. \quad (\text{A16})$$

Here, $k_R(j, l)$ and $k_I(j, l)$ are the real and imaginary parts of the function $k(t', t)$ evaluated on the discrete time slices. In passing from Eq. (A11) to (A16) we have made use of the fact that in the limit $N \rightarrow \infty$ it makes no difference in the Lagrangian path integral whether the integral $\int V(x(t)) dt$ between time slices is approximated using the value of V at the start or end of the interval. The form (A16) is slightly more convenient for what follows.

We now change variables in the multiple integral (A11) from $\{x'_k, x_k\}$ to the discrete versions of Eq. (6.11):

$$\xi_k = x'_k - x_k, \quad X_k = \frac{1}{2}(x'_k + x_k). \quad (\text{A17})$$

The Jacobian is unity on each slice so that Eq. (A14) becomes

$$D(\alpha', \alpha) = \lim_{N \rightarrow \infty} \int \left[\left[\frac{M}{2\pi i \epsilon \hbar} \right]^{N-1} \prod_{k=0}^N dX_k d\xi_k \right] \times \delta(\xi_N) E_{\alpha'\alpha} \left[X_j + \frac{\xi_0}{2}, X_j - \frac{\xi_0}{2} \right] \exp \left[\frac{iA(X_j, \xi_j)}{\hbar} \right] \bar{\rho} \left[X_0 + \frac{\xi_0}{2}, X_0 - \frac{\xi_0}{2} \right]. \quad (\text{A18})$$

After a little algebra the exponent may be written

$$A(X_j, \xi_j) = -\xi_0 M \left[\frac{X_1 - X_0}{\epsilon} \right] + \sum_{j=1}^N \epsilon \xi_j \left[- \left[\frac{X_{j+1} - 2X_j + X_{j-1}}{\epsilon^2} \right] - K X_j \right] + \frac{1}{4} \sum_{j=1}^N \sum_{l=1}^j \epsilon^2 \xi_j [k_R(j, l) X_l + ik_I(j, l) \xi_l]. \quad (\text{A19})$$

Equation (A19) is the discrete analogue of Eq. (6.12) and shows precisely how the second and other derivatives of the path in that expression are to be interpreted in a time-slicing representation.

We next assume decoherence and carry out the integration over the ξ_k 's neglecting the constraints of the coarse graining as discussed in Sec. III. Note that A in Eq. (A19) depends on ξ_0 only through the first term and that $\xi_N = 0$ because of the δ function in Eq. (A14). The result for the diagonal elements of the decoherence functional is explicitly

$$p(\alpha) = \lim_{N \rightarrow \infty} \int \left[\left(\frac{M}{2\pi i \epsilon \hbar} \right) \left(\frac{M}{2\pi i \epsilon^2} \right)^{N-2} \prod_{k=0}^N dX_k \right] \left[\det \left(\frac{k_I}{4\pi} \right) \right]^{-1/2} E_{\alpha'}(X_j, X_j) \\ \times \exp \left[-\frac{1}{\hbar} \sum_{j=1}^{N-1} \sum_{l=1}^{N-1} \epsilon^2 e_j(X_m) k_I^{\text{inv}}(j, l) e_l(X_m) \right] w \left[X_0, M \left(\frac{X_1 - X_0}{\epsilon} \right) \right]. \quad (\text{A20})$$

In this expression $\det k_I$ and k_I^{inv} are the determinant and inverse, respectively, of the $(N-1) \times (N-1)$ matrix $k_I(j, l)$, $j, l = 1, \dots, N-1$. The quantity e_k is the discrete version of the equation of motion (6.14), namely,

$$e_j(X_l) = -M \left[\frac{X_{j+1} - 2X_j + X_{j-1}}{\epsilon^2} \right] \\ - KX_j - \sum_{l=1}^N \epsilon k_R(j, l) X_l \quad (\text{A21})$$

and $w(X, P)$ is the Wigner function defined in Eq. (5.2). Equation (A20) contains the precise measure for the path integral (6.13) and the precise meaning of the integral over the Wigner function in it.

C. Functional Fourier transforms

In Sec. V we utilized a functional Fourier transform of the decoherence functional to define a distribution functional for the total force. Here, we offer a more explicit definition of what such transforms mean. We consider the case of one-dimensional paths for simplicity. Consider a functional $D[\xi(\tau)]$. On paths that are piecewise linear between time slices $\tau_0, \tau_1, \dots, \tau_N = T$ this defines a function, $D(\xi_0, \xi_1, \dots, \xi_N)$ of the values that $\xi(\tau)$ assumes on these slices. This function may be Fourier transformed in the following way:

$$G(R_1, \dots, R_N; \xi_0) = \mathcal{N} \int \left[\prod_{k=1}^N d\xi_k \right] \exp \left[-\frac{i}{\hbar} \sum_{k=1}^N \epsilon \xi_k R_k \right] \\ \times D(\xi_0, \xi_1, \dots, \xi_N) \quad (\text{A22})$$

where $\epsilon = T/N$ is the spacing between the time slices and \mathcal{N} is a normalizing factor. The inverse of this is

$$D(\xi_0, \xi_1, \dots, \xi_N) \\ = \mathcal{N}^{-1} \int \left[\prod_{k=1}^N \left(\frac{dR_k}{2\pi} \right) \right] \exp \left[\frac{i}{\hbar} \sum_{k=1}^N \epsilon \xi_k R_k \right] \\ \times G(R_1, \dots, R_N; \xi_0). \quad (\text{A23})$$

It is this latter expression, in the limit where the number of slices, N , goes to infinity that gives an explicit meaning to the functional Fourier transform that we wrote in Eq. (5.7) as

$$D[\xi(\tau)] = \int \delta R \exp \left[\frac{i}{\hbar} \int_0^T dt \xi(t) R(t) \right] G[R(\tau), \xi_0]. \quad (\text{A24})$$

In particular, Eq. (A23) defines the measure δR . Note that because of the way that we have differenced the exponent in Eq. (A24), ξ_0 remains untransformed and occurs on both sides of Eq. (A23).

The normalizing factor \mathcal{N} is typically arbitrary and will cancel from physical expressions such as the definition of the equation of motion in Eqs. (5.9) and (5.10). However, mathematically, it must be chosen carefully in order that expressions such as Eq. (A22) exist in the limit $N \rightarrow \infty$. For example, when the decoherence functional is given as in Eqs. (6.11) and (6.12) for linear problems the normalizing factor would depend on K_I and certainly on ϵ . We assume that such a normalizing factor always exists for interesting cases.

D. An operator expression for the influence functional

The influence functional $\exp\{iW[x'(\tau), x(\tau)]/\hbar\}$ was defined by the path integral (3.3) but can be represented as the operator expression (7.6) when the initial density matrix factors as in Eq. (3.6). We now spell out the details of this connection. By inserting complete sets of states of the Hilbert space \mathcal{H}_Q that are eigenfunctions of the coordinates Q^A , the right-hand side of Eq. (7.6) can be written

$$\text{Sp} [U_{T,0}[x'(\tau)] \rho_B U_{T,0}^\dagger[x(\tau)]] = \int dQ'_f \int dQ_f \int dQ'_0 \int dQ_0 \delta(Q'_f - Q_f) \\ \times \langle Q'_f | U_{T,0}[x'(\tau)] | Q'_0 \rangle \langle Q'_0 | \rho_B | Q_0 \rangle \langle Q_0 | U_{T,0}^\dagger[x(\tau)] | Q_f \rangle. \quad (\text{A25})$$

But $\langle Q_f | U_{T,0}[x(\tau)] | Q_0 \rangle$ is just the propagator in the Hilbert space \mathcal{H}_Q corresponding to unitary evolution by the time-dependent Hamiltonian $H_Q(x(t))$ [cf. Eq. (7.4)] over the time interval $[0, T]$. This propagator has an elementary path integral representation whose derivation we have reviewed in Sec. B of this Appendix [cf. Eq. (A12)]. It is

$$\langle Q_f | U_{T,0}[x(\tau)] | Q_0 \rangle = \int_{[Q_0, Q_f]} \delta Q \exp\{iS_Q[x(\tau), Q(\tau)]/\hbar\}, \quad (\text{A26})$$

where $S_Q[x(\tau), Q(\tau)]$ is the action (7.1) that corresponds to the Hamiltonian $H_Q(x(t))$. The path integral is over paths that start at time $t=0$ at Q_0 and proceed to Q_f at time T and are otherwise unrestricted. Inserting (A26) into the right-hand side of Eq. (A25), noting [Eq. (7.5)] that $\langle Q'_0 | \rho_B | Q_0 \rangle \equiv \rho_B(Q'_0, Q_0)$, and using Eq. (7.1) for $S_Q[x(\tau), Q(\tau)]$, we recover the expression (3.3) that defines the influence functional $\exp\{iW[x'(\tau), x(\tau)]/\hbar\}$. Thus the identity (7.6) is verified.

-
- [1] T. Brun (unpublished).
- [2] M. Gell-Mann and J. B. Hartle, in *Complexity, Entropy, and the Physics of Information*, edited by W. Żurek, SFI Studies in the Sciences of Complexity Vol. VIII (Addison Wesley, Reading, 1990); or in *Proceedings of the 3rd International Symposium on the Foundations of Quantum Mechanics in the Light of New Technology*, edited by S. Kobayashi, H. Ezawa, Y. Murayama, and S. Nomura (Physical Society of Japan, Tokyo, 1990).
- [3] 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 (South East Asia Theoretical Physics Association and Physical Society of Japan, World Scientific, Singapore, 1990).
- [4] J. B. Hartle, in *Quantum Cosmology and Baby Universes*, Proceedings of the 7th Jerusalem Winter School, Jerusalem, Israel, 1989, edited by S. Coleman, J. Hartle, T. Piran, and S. Weinberg (World Scientific, Singapore, 1991).
- [5] R. Griffiths, *J. Stat. Phys.* **36**, 219 (1984).
- [6] R. Omnès, *J. Stat. Phys.* **53**, 893 (1988); **53**, 933 (1988); **53**, 957 (1988); **57**, 359 (1989); *Rev. Mod. Phys.* **64**, 339 (1992).
- [7] M. Gell-Mann and J. B. Hartle (unpublished).
- [8] E. Joos and H. D. Zeh, *Z. Phys. B* **59**, 223 (1985).
- [9] L. P. Kadanoff and P. C. Martin, *Ann. Phys. (N.Y.)* **24**, 419 (1963).
- [10] D. Forster, *Hydrodynamic Fluctuations, Broken Symmetry, and Correlation Functions* (Addison-Wesley, Redwood City, CA 1975).
- [11] J. Lebovitz and H. Spohn, *J. Stat. Phys.* **28**, 539 (1982); **29**, 39 (1982).
- [12] R. Balian, Y. Alhassid, and H. Reinhardt, *Phys. Rep.* **131**, 1 (1986).
- [13] A. Caldeira and A. Leggett, *Physica* **121A**, 587 (1983).
- [14] B.-L. Hu, J. P. Paz, and Y.-H. Zhang, *Phys. Rev. D* **45**, 2843 (1992); **47**, 1576 (1993).
- [15] W. Żurek, *Phys. Today* **44**, 36 (1991).
- [16] R. P. Feynman and J. R. Vernon, *Ann. Phys. (N.Y.)* **24**, 118 (1963).
- [17] C. W. Gardiner, *Quantum Noise* (Springer-Verlag, Berlin, 1992).
- [18] R. Kubo, M. Toda, and N. Hashitsume, *Statistical Physics II* (Springer, Berlin, 1978), p. 167ff.
- [19] H. Dowker and J. Halliwell, *Phys. Rev. D* **46**, 1580 (1992).
- [20] M. Gell-Mann and J. B. Hartle, in *Proceedings of the 1st International A. D. Sakharov Conference on Physics*, Moscow, USSR, 1991 (Nova Science Publishers, New York, 1992); or in *Proceedings of the NATO Workshop on the Physical Origins of Time Asymmetry*, Mazagon, Spain, 1991, edited by J. Halliwell, J. Perez-Mercader, and W. Żurek (Cambridge University Press, Cambridge, MA, 1993).
- [21] E. Fitzgerald, *Rubáiyát of Omar Khayyám* (1859).
- [22] H. Zeh, *Found. Phys.* **1**, 69 (1971).
- [23] W. Żurek, *Phys. Rev. D* **24**, 1516 (1981); **26**, 1862 (1981); in *Proceedings of the NATO Workshop on the Physical Origins of Time Asymmetry* [20].
- [24] O. Kübler and H. D. Zeh, *Ann. Phys. (N.Y.)* **76**, 405 (1973).
- [25] S. Kochen, in *Symposium on the Foundations of Modern Physics*, edited by P. Lahti and P. Mittelstaedt (World Scientific, Singapore, 1985).
- [26] A. Albrecht, *Phys. Rev. D* **46**, 5504 (1992).
- [27] W. Żurek, in *Non-Equilibrium Quantum Statistical Physics*, edited by G. Moore and M. Scully (Plenum, New York, 1984).
- [28] S. R. deGroot and L. G. Suttrop, *Foundations of Electrodynamics* (North-Holland, Amsterdam, 1972), Chap. VI.
- [29] M. Hillery, R. F. O'Connell, M. O. Scully, and E. P. Wigner, *Phys. Rep.* **106**, 121 (1984).
- [30] K. Husimi, *Proc. Phys. Math. Soc. Jpn.* **22**, 264 (1940).
- [31] J. S. Bell, *Physics* **1**, 195 (1964).
- [32] R. P. Feynman, in *Quantum Implications: Essays in Honor of David Bohm*, edited by B. J. Hiley and F. D. Peat (Routledge and Kegan Paul, London, 1987).
- [33] R. P. Feynman and A. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).
- [34] C. Morais Smith and A. Caldeira, *Phys. Rev. A* **41**, 3103 (1990).
- [35] J. Halliwell, *Phys. Rev. D* **46**, 1610 (1992).
- [36] T. Brun, following paper, *Phys. Rev. D* **47**, 3383 (1993).
- [37] Y. Aharonov, P. Bergmann, and J. Lebovitz, *Phys. Rev.* **134**, B1410 (1964).
- [38] R. Zwanzig, *J. Stat. Phys.* **9**, 215 (1973).
- [39] S. Ryang and T. Saito, *Prog. Theor. Phys.* **71**, 1108 (1984).
- [40] W. Unruh and W. Żurek, *Phys. Rev. D* **40**, 1071 (1989).
- [41] R. P. Feynman, *Phys. Rev.* **84**, 108 (1951).
- [42] C. Garrod, *Rev. Mod. Phys.* **38**, 483 (1966).