Decoherence properties of finite quantum systems

Göran Lindblad *

Theoretical Physics
Royal Institute of Technology
S-100 44 Stockholm, Sweden

(December 1993)

There has recently been a revival of interest in Griffiths' consistent histories approach to quantum theory. In this scheme a set of correlation functions are used to define probabilities of sequences of events (histories) without necessarily postulating an outside observer. Instead, the vanishing of certain quantum coherence terms is claimed to be sufficient for a consistent classical interpretation to exist. In spite of the large volume of work in this field there seems to be a lack of general mathematical results on the consequences of such decoherence conditions. In this paper some of the properties of the correlation functions for finite quantum systems (with unitary dynamics and discrete spectra) are investigated using methods from operator algebras and quantum statistical mechanics. Their information content and decoherence properties are measured by entropy functionals in a formalism which is capable of dealing with approximate measurements as well as the ideal ones defined by orthogonal projectors. Necessary and sufficient decoherence conditions are found for the observables to form a classical commutative system. These conditions will involve the correlation functions of all orders, and the systems satisfying them exactly are exceptional. On the other hand it is shown that for large but finite quantum systems most observables will approximately satisfy a chaotic form of decoherence condition for correlation functions of not too high order. It is also possible to choose observables which define nearly deterministic (and hence decoherent) histories over a finite time.

03.65.Bz, 05.30Ch

I. INTRODUCTION

Recently Omnès [1] Gell-Mann and Hartle [2–4], Dowker and Halliwell [5] and many others have taken up the ideas of Griffiths [6,7] on the foundations of a quantum-mechanical theory of closed systems. The goal of his approach is to do away with the necessity of introducing measuring instruments or classical observers looking at the system from the outside, dealing instead with a single closed quantum system with a reversible, Hamiltonian evolution. The motivation comes from the desire to have an interpretation of quantum theory adapted to the need of eventually constructing a quantum theory of gravity, as this may involve a description of a closed universe without outside observers. The technical arguments of these works are based on the so called decoherence functional (or quantum correlation kernel, which will be defined again below). It has been argued that if this quantity, with a suitable choice of arguments, has certain consistency properties, called "decoherence" as they imply the vanishing of quantum interference effects, then some probability distributions can be defined as objective quantities without the introduction of external observers. This is then taken to mean that the closed system has certain intrinsically classical properties.

A superficially different but closely related set of ideas concern the emergence of classical properties in a finite quantum system as a result of a weak interaction of the observed system with the environment [8–13]. The dynamics of the small system is that of an open quantum system, which means that the Hamiltonian evolution is modified by dissipative terms which will not preserve all the information content of the quantum state. This environment-induced decoherence can thus destroy the quantum phases which hinder us from assigning definite properties to macroscopic systems (Schrödinger's cats and all). Some authors expect this kind of formalism to explain why there is an objectively existing macroscopic world despite the quantum nature of the basic equations of motion.

---

*E-mail: gli@theophys.kth.se

Preprint TRITA-TFY-93-9

This paper will not deal directly with these ambitious projects. Instead the goal is to state and prove some mathematical properties of the quantum correlations which are likely to prove useful in this and other contexts. For our purposes it is inessential which point of view is taken on the philosophical issues. However, in order to have a well-defined mathematical setup we choose the dynamics to be of the reversible, conservative type represented by unitary maps, the spectra of which are assumed to be discrete. Furthermore, we choose the time parameter to be discrete and the Hilbert spaces to be of finite dimension in most of the paper. The technical tool used to develop the argument are entropy inequalities which relate the quantum coherence properties to the information entropy of certain density operators defined by the quantum correlations.

The quantum systems considered here have a fixed unitary (Hamiltonian) dynamics. There is also a fixed set of operations representing observations of the system in an open system approach, or defining the histories in the closed system picture. The operations are given by a set of orthogonal projectors or by a more general partition of unity describing an approximate measurement, perhaps a measurement of a continuous variable. The given set of observations is repeated an arbitrary number of times, allowing us to follow the evolution of the system starting from an initial state which is assumed stationary, representing a microcanonical state for the system. This paper will not go into the problem of actually choosing these ingredients of the model from some physical principles, we start from a suitable level of generality chosen to give some results which are mathematically accessible and potentially interesting for applications.

The paper starts in section II with a definition of the timeordered (causal) quantum correlation kernels (also called decoherence functionals). Their positivity and normalization properties mean that they define density operators in suitably chosen Hilbert spaces. In section III entropy functionals for these density operators are used to measure the amount of coherence in the correlation kernels. There is also defined a measure of the information obtained in a sequence of observations which applies also to nonideal (approximate) measurements of continuous observables. Some useful inequalities for the entropy functionals are given in section III, based on a number of mathematical results which are listed in the Appendix together with some of the proofs.

In sections IV and V some necessary and sufficient conditions are given on the correlations in order that the full set of time translated observations shall correspond to a commutative algebra of observables. This is done first for the case where the observations are given by projectors (section IV), and then analogous results are proved for the more general and complex case of approximate, nonrepeatable observations (section V).

In section VI a simple example shows that it is possible to have complete decoherence for a finite number of observations, while all following ones show a maximal amount of quantum coherence. Hence there is a problem of finding the decoherence properties from a finite sequence of observations. In this context there are two extreme types of correlation functions showing decoherence. There is a chaotic type where the decoherence comes from the statistical independence of the successive events in the history, and a deterministic type where the decoherence is due to the predictability of the later events from the earlier ones. Only in the latter case can we conclude anything about the overall decoherence from a limited set of information.

In section VII it is shown that for a system where a large number of energy levels are involved most choices of observables will give histories which show approximate decoherence of the chaotic type for low order correlations. This demonstration is based on the introduction of certain ensembles of systems or choices of observables, and estimating the probability measure of the set of elements in the ensemble which depart from this decoherence property.

In section VIII it is shown under similar conditions that it is always possible to introduce projectors which approximate a commutative, classical situation well under a large but finite number of iterations.

In section IX some tentative conclusions are drawn from the mathematical results. It is argued that decoherence is insufficient for the purpose of defining a classical domain, still less capable of making it unique.

II. CORRELATION KERNELS

The fundamental quantities in this paper are the quantum correlation kernels (QCKs for short, or multitime correlation functions), in recent publications sometimes called decoherence functionals. Their definition and most basic properties will be set down here, and a few remarks on their importance and interpretation ends this section.

Introduce a Hilbert space $\mathcal{K}$, which we can take to be of large but finite dimension N, a set of non-negative (and consequently self-adjoint) operators $\{P(\alpha); \alpha \in \mathcal{I}\}$ in $\mathcal{K}$,

$$\begin{aligned} P(\alpha) &\geq 0 \\ \sum_\alpha P(\alpha)^2 &= \mathbb{1} \end{aligned} \tag{2.1}$$

There is an associated completely positive (CP) map on the operators $X \in B(\mathcal{K})$

$$T : X \mapsto T[X] = \sum_\alpha P(\alpha) X P(\alpha)$$

See e.g. [14,15] for a physical and a mathematical background on CP maps. There is a dual map on the density operators ρ which is actually the same map, but this self-duality does not hold for general CP maps. As ρ defines a state, denoted by the same symbol, through $\rho(X) = \mathrm{Tr}[\rho X]$, the dual map can symbolically be written as $\rho \mapsto \rho \circ T$. Each term in the sum above defines the probability of an outcome α

$$p(\alpha) = \rho(P(\alpha)^2)$$

and the state $\rho(\alpha)$ after the observation of this outcome by

$$p(\alpha)\rho(\alpha) = P(\alpha)\rho P(\alpha)$$

This construction has been used under various names, like CP instruments, operation-valued measures and so on [16–18]. In most of the papers in this field the operators are restricted to be orthogonal projectors, which means that

$$P(\alpha)P(\beta) = \delta_{\alpha\beta} P(\alpha) \tag{2.2}$$

but it is convenient to deal also with the more general case. This allows the same formalism to work for measurements of observables with continuous spectra and measurements continuous in time, though these subjects will not appear explicitly in this paper. Typically the rank of the operators (projectors) will be large, of the same order of magnitude as the total dimension of $\mathcal{K}$, thus representing an incomplete (coarsegrained) observation of the system. We represent repeated observations of the system by combining the operators (2.1) with a unitary map U representing the dynamics (for simplicity we let the time parameter be discrete with a fixed time unit). The timeordered (causal) operator products $V_n(\alpha)$ are defined iteratively through $V_1(\alpha) = P(\alpha)$ and

$$V_n(\alpha_1, \alpha_2, \ldots, \alpha_n) \equiv V_n(\alpha) = P(\alpha_n) U V_{n-1}(\alpha_1, \alpha_2, \ldots, \alpha_{n-1}) \tag{2.3}$$

Here the index 1 refers to the first observations in the sequence which takes place at time $t = 0$. The notation α will be used also for the causal sequences (*histories*) in (2.3) and $V(\alpha)$ for the operators (2.3) of arbitrary order n. In order to have a complete notation we should also introduce an extra symbol, say ι referring to a trivial event with $P(\iota) = \mathbb{1}$, and consider sequences including arbitrary numbers of this symbol. In order not to complicate the notation this is subsumed under the notation in (2.3). The normalization condition follows from (2.1), for each n

$$\sum_{\alpha \in \mathcal{I}^n} V_n(\alpha)^\dagger V_n(\alpha) = \sum_{\alpha \in \mathcal{I}^n} V_n(\alpha) V_n(\alpha)^\dagger = \mathbb{1} \tag{2.4}$$

In order to simplify the typography we also introduce $M(\alpha) = V(\alpha)^\dagger V(\alpha)$ which is a positive-operator-valued measure (POVM) [19,20]. This means that it satisfies

$$\begin{aligned} M(\alpha) &\geq 0 \\ \sum_\alpha M(\alpha) &= \mathbb{1} \end{aligned} \tag{2.5}$$

Given an initial state ρ for the whole system, the timeordered (causal) quantum correlation kernel of order n is given by

$$\mathcal{D}_n(\alpha|\beta) = \rho(V_n(\alpha)^\dagger V_n(\beta)) \tag{2.6}$$

For the formal development it is convenient to assume ρ to be stationary. Let the Hilbert space be spanned by a finite number of energy eigenstates corresponding to a microcanonical ensemble for a finite system, and choose ρ to represent that ensemble. This state is tracial, which means that $\rho(XY) = \rho(YX)$ for all observables X, Y. In any realistic description of a large system there will inevitably exist many subensembles with different classical interpretations in such a large ensemble, but this should then be reflected in the statistics of the full set of histories. The diagonal elements

$$p(\alpha_1, \alpha_2, \ldots, \alpha_n) \equiv p_n(\alpha) = \mathcal{D}_n(\alpha|\alpha) \tag{2.7}$$

give the probability distributions $\{p_n(\alpha)\}$, each associated with sequential n-fold measurements of the (coarse-grained) observable, with the obvious normalization when summing over all indices

$$\sum_\alpha p_n(\alpha) = 1.$$

Let us set down a couple of immediate properties of $\mathcal{D}_n(\alpha|\beta)$.

(1) *Positivity.* $\mathcal{D}_n$ is a positive semidefinite matrix for every value of n. This means: for all $\lambda(\alpha) \in \mathbb{C}, \alpha \in \mathcal{I}^n$ it holds that

$$\sum_{\alpha,\beta} \lambda(\alpha)^* \lambda(\beta) \mathcal{D}_n(\alpha|\beta) \geq 0 \tag{2.8}$$

This implies (but is stronger than) the Schwarz inequality

$$|\mathcal{D}_n(\alpha|\beta)|^2 \leq \mathcal{D}_n(\alpha|\alpha)\mathcal{D}_n(\beta|\beta)$$

(2) *Compatibility.* Summing over the *last* index set in $\mathcal{D}_n$ gives $\mathcal{D}_{n-1}$

$$\sum_{\alpha_n} \mathcal{D}(\alpha_1, \alpha_2, \ldots, \alpha_n|\beta_1, \beta_2, \ldots, \alpha_n) = \mathcal{D}(\alpha_1, \alpha_2, \ldots, \alpha_{n-1}|\beta_1, \beta_2, \ldots, \beta_{n-1}). \tag{2.9}$$

If the initial state is stationary and if the invariance relation

$$\sum_\alpha \rho(P(\alpha) X P(\alpha)) = \rho(X) \tag{2.10}$$

holds for all operators X in the Hilbert space, then there is a corresponding result when summing over the first index. The characteristic feature of quantum coherence is that summing over one of the other outcomes $\{\alpha_2, \ldots, \alpha_{n-1}\}$ will in general not give the same result as leaving out the observation at that instant. Griffiths called a set of finite histories $\{\alpha_1, \ldots, \alpha_n\}$ *consistent* if this quantum coherence effect is absent in the probability distributions (2.7) [6]. In our notation the consistency condition reads: for any $k = 1, \ldots, n$

$$\sum_{\alpha_k} p_n(\alpha) = p_n(\alpha') \tag{2.11}$$

where in α' the symbol α_k is replaced by ι. We can use a slightly stronger consistency condition by imposing the analog of (2.9) also for the intermediate events $\{\alpha_2, \ldots, \alpha_{n-1}\}$, which means that (2.11) is extended to the offdiagonal elements. An even stronger condition is the operatorvalued counterpart

$$\sum_\alpha P(\alpha) V_m(\beta)^\dagger V_m(\gamma) P(\alpha) = V_m(\beta)^\dagger V_m(\gamma) \tag{2.12}$$

for all $\alpha, \beta, \gamma, m \leq n$. This is clearly satisfied if the operators

$$P_m(\gamma) = U^{m\dagger} P(\gamma) U^m \tag{2.13}$$

commute for $m = 0, 1, \ldots, n-1$. In section V results will be given on the equivalence of some of these decoherence conditions when they are satisfied for all orders of the correlation functions.

When (2.2) holds the consistency condition (2.11) is implied by the vanishing of the offdiagonal elements of the correlations up to order n:

$$\mathcal{D}_k(\alpha|\beta) = \delta_{\alpha\beta}\mathcal{D}_k(\alpha|\alpha) \tag{2.14}$$

for $k \leq n$. This property is often referred to as *decoherence* in the literature. It is easier to apply than the conditions (2.11).

The correlation kernel (2.6) is of a form ubiquitous in quantum physics. Recall the Wightman functions of relativistic QFT [21,22] where the stationary state is the physical vacuum. Here the observables are field operators in a spacetime point rather than projectors, and the analyticity properties of the Wightman functions makes the time order less relevant. Similar constructs are used in quantum optics [23,24] where the observables are replaced by creation and annihilation operators in order to model the action of photon counters, and the probability distributions give the photocount statistics.

In a general scheme where we observe the evolution of the system from the outside the multitime correlations have the central role of defining the intrinsic dynamics in analogy with the theory of stochastic processes (see [17,25] and many references given there). The quantum case is distinguished by the importance we have to attach to the causal time order in (2.6). The fact that this time order is the relevant one is plain already from an application of timedependent perturbation theory to an interaction of the system with an outside measuring apparatus. From a consideration of the most general types of measurements it is also appears that the offdiagonal elements in (2.6) must be considered to be observable, not just the diagonal ones defining the probabilities (2.7). In a formalism without external observers this may be a debatable point.

In the open system approach the $P(\alpha)$ are operators belonging to the small observed system, and the dynamics of this system interacting with an (infinite) reservoir can be irreversible. The simplest examples of such models are quantum dynamical semigroups, which define a quantum version of Markov processes [17,26]. They represent an idealized limit where the reservoir has an internal relaxation which is much faster than the evolution of the open system. In the open system picture it is natural to use a more general setup where instead of a fixed set $\{P(\alpha)\}$ we can use a different set of observables (or more general operations) for the small system at different instants [17]. When we use a full quantum description for this system, this possibility of having noncommuting, complementary observables is an essential ingredient in the theory. In the closed system picture, however, a complete arbitrariness in the choice of observables at different instants would make the intrinsic dynamics of the system quite irrelevant. An example of this situation is given by the "Schmidt" histories discussed in recent papers [11]. The point of view taken here is that it is the combination of a restricted set of observations with a fixed intrinsic dynamics of the closed system which can given an interesting structure to this problem. In the following we will use the closed system picture, and it is then convenient to restrict ourselves to the consideration of the same set $\{P(\alpha)\}$ for all instants.

It is common to apply the label "classical" to states of the quantum system rather than the full set of correlation functions. As an example consider the correspondence limit (large quantum numbers) of the motion of a free or bound particle. Using coherent states it is possible to construct initial states which evolve in an approximately classical manner over a certain time scale. This type of construction can be included in the present scheme if we allow instruments (2.1) which create coherent states (or approximations of them) out of the initial unstructured microcanonical state. The stationarity of the initial state also makes the mathematical formalism a good deal simpler. For this reason we do not consider, in this work, the different possible choices for the initial state.

III. ENTROPY FUNCTIONALS AND COHERENCE PROPERTIES

The positivity and normalization properties of the correlation functions defined in the previous section allow us to define entropy measures of their information content and the coherence properties. Basic to this development is a set of entropy inequalities, where some of the proofs appear in the Appendix.

The relations (2.8) and (2.9) mean that for each value of $n = 1, 2, \ldots$ we can consider $\mathcal{D}_n$ as a density operator in a Hilbert space

$$\begin{aligned} h_{n,\mathcal{I}} &= \bigotimes^{n} h_{\mathcal{I}} \\ h_{\mathcal{I}} &= \{\phi(\alpha) \in \mathbb{C}; \sum_{\alpha\in\mathcal{I}} |\phi(\alpha)|^2 \leq \infty\} \end{aligned} \tag{3.1}$$

and $\mathcal{D}_{n-1}$ is obtained from $\mathcal{D}_n$ by taking a partial trace. We can define a (dimensionless) entropy for each n

$$S_n\{\alpha\} = S(\mathcal{D}_n) = -\mathrm{Tr}[\mathcal{D}_n \ln \mathcal{D}_n]. \tag{3.2}$$

If we leave out the trace implicit in (2.6) there is a density operator

$$\mathcal{R}_n(\alpha|\beta) = V_n(\beta)\rho V_n(\alpha)^\dagger \tag{3.3}$$

which is again a positive semidefinite operator of trace 1, now acting in the Hilbert space $h_{n,\mathcal{I}} \otimes \mathcal{K}$. Introduce the diagonal elements of $\mathcal{R}_n$ as density operators $\rho_n(\alpha)$ through

$$p_n(\alpha)\rho_n(\alpha) = \mathcal{R}_n(\alpha|\alpha). \tag{3.4}$$

and the related density operators $\sigma_n(\alpha)$

$$p_n(\alpha)\sigma_n(\alpha) = \sqrt{\rho}\, V_n(\alpha)^\dagger V_n(\alpha)\sqrt{\rho}. \tag{3.5}$$

The latter set satisfy

$$\sum_\alpha p_n(\alpha)\sigma_n(\alpha) = \rho \tag{3.6}$$

but the corresponding relation does not hold for the $\rho_n(\alpha)$ in general. We will use the following information measure

$$\begin{aligned} I_n\{\alpha\} &= S(\rho) - \sum_\alpha p_n(\alpha) S(\sigma_n(\alpha)) \\ &= \sum_\alpha p_n(\alpha) S(\sigma_n(\alpha)|\rho) \geq 0. \end{aligned} \tag{3.7}$$

The relative entropy functional

$$S(\rho|\mu) = \mathrm{Tr}[\rho \ln \rho - \rho \ln \mu] \geq 0 \tag{3.8}$$

is zero if and only if $\rho = \mu$ [27–29]. Note that in [28] and many other places the arguments have the opposite order. For commuting observables I_n gives the information gain according to conventional wisdom while Ozawa [18] treats the quantum case. Finally there is a classical (Shannon) entropy associated with the probability distribution $\{p_n\}$

$$H_n\{\alpha\} \equiv H\{p_n(\alpha)\} = -\sum_\alpha p_n(\alpha) \ln p_n(\alpha). \tag{3.9}$$

and a corresponding relative entropy for two distributions

$$H_n\{p|q\} \equiv H\{p_n(\alpha)|q_n(\alpha)\} = \sum_\alpha p_n(\alpha)(\ln p_n(\alpha) - q_n(\alpha)) \tag{3.10}$$

We now come to the interrelations between these different entropy functions. The relation

$$I_n\{\alpha\} \leq S_n\{\alpha\} \leq H_n\{\alpha\}$$

holds, this is Proposition 4 in the Appendix. The second equality holds precisely when the decoherence condition (2.14) holds, hence in particular when the operators (2.13) are commuting-projectors, in which case the observations have a perfect classical structure. However, even in the commutative case there will be offdiagonal elements in (2.6) if we use operators (2.1) which are not projectors. In this case it turns out that $I_n(\alpha)$ is the most useful of the entropy functions (see Theorem 4).

The stationarity of the state and some well known relations for the functional (3.9) implies that [30]

$$\begin{aligned} H_n\{\alpha\} &\leq H_{n+1}\{\alpha\} \\ H_{n+1}\{\alpha\} - H_n\{\alpha\} &\leq H_n\{\alpha\} - H_{n-1}\{\alpha\} \end{aligned} \tag{3.11}$$

and from this follows that there is an asymptotic form as $n \longrightarrow \infty$

$$H_n\{\alpha\} \sim n \cdot h + k$$

where $h, k \geq 0$. Typically

$$H_\infty\{\alpha\} \equiv \lim_{n\to\infty} H_n\{\alpha\} = \infty$$

even though the Hilbert space is of finite dimension. A simple example in the Hilbert space of dimension 2 is given by the projectors corresponding to a measurement of the spin in the z-direction and a unitary U which represents a rotation of the z-axis into the x-axis. Here $H_n\{\alpha\} = n \ln 2$. In fact, we will see below that $h = 0, k < \infty$ only in the case of commuting projections. Hence we can use a value $h > 0$ as a measure of coherence in this particular case.

The entropy (3.2) satisfies, by Propositions 5, 6

$$S_{n-1}\{\alpha\} \leq S_n\{\alpha\} \leq S(\rho) + S(\rho_n)$$
$$\rho_n = \sum_\alpha \mathcal{R}_n(\alpha|\alpha)$$

It is clear that when $\mathcal{K}$ is of finite dimension N it holds that $S(\rho), S(\rho_n)$ are no larger than $\ln N$, hence there must then be a monotone convergence to a finite value

$$S_n\{\alpha\} \longrightarrow S_\infty\{\alpha\} \leq 2 \ln N \tag{3.12}$$

When the operators (2.5) are projectors for all n it holds that

$$S_\infty\{\alpha\} \leq \ln N \tag{3.13}$$

as the number of $M(\alpha)$ is bounded by N. Using the same type of arguments we find for (3.7)

$$I_{n-1}\{\alpha\} \leq I_n\{\alpha\} \leq S(\rho)$$

and a monotone convergence

$$I_n\{\alpha\} \longrightarrow I_\infty\{\alpha\} \leq S(\rho)$$

For the density operator (3.3) it holds that $S(\mathcal{R}_n) = S(\rho)$ for all n, using the proof of Proposition 6. If we project out the diagonal of $\mathcal{R}_n$ we obtain (3.4), and it is shown in Proposition 7 that

$$\sum_\alpha p_n(\alpha) S(\rho_n(\alpha)) \leq S(\rho) \leq \sum_\alpha p_n(\alpha) S(\rho_n(\alpha)) + H_n\{\alpha\}$$

where the second equality holds precisely when the offdiagonal elements of $\mathcal{R}_n$ are zero

$$\mathcal{R}_n(\alpha|\beta) = 0 \; \forall \alpha \neq \beta \Longleftrightarrow S(\rho) = \sum_\alpha p_n(\alpha) S(\rho_n(\alpha)) + H_n\{\alpha\}$$

This then means that

$$\mathcal{R}_n(\alpha|\beta) = 0 \; \forall \alpha \neq \beta \Longrightarrow H_n\{\alpha\} \leq S(\rho)$$

a stronger inequality than that coming from the vanishing of the offdiagonal elements of $\mathcal{D}_n$. It is shown in the Proposition 7 that $S(\sigma_n(\alpha)) = S(\rho_n(\alpha)) \; \forall n, \alpha$. Together with (3.6) this shows that $I_n\{\alpha\} = S_n\{\alpha\}$ holds precisely when the offdiagonal elements of $\mathcal{R}_n$ vanish. Of course, this holds when the operators (2.13) are commuting projectors and commute with the density operator ρ. Again the equality fails in the commutative but nonprojective case.

IV. THE COMMUTATIVE CASE I

A couple of results will be proved for the case (2.2) when the observations are orthogonal projections and we can use (2.14) as a definition of decoherence.

Theorem 1 *Let the system of operators $V(\alpha)$ be based on an orthogonal set of projectors (2.1), (2.2) and assume that the Hilbert space $\mathcal{K}$ is of finite dimension N. Then the equality $H_n\{\alpha\} = S_n\{\alpha\}$ for all n implies that there is a commutative system defining the correlation functions and consequently $S_\infty \leq \ln N$.*

Proof. The QCK (2.6) can be represented in the following way. Use the representation (A4) of a vector $\Omega \in \mathcal{K} \otimes \mathcal{K}$. There is then a subspace $\mathcal{K}_1 \subseteq \mathcal{K} \otimes \mathcal{K}$ spanned by the vectors

$$\psi(\alpha) = V(\alpha)\Omega$$

It follows from Proposition 4 that $H_n = S_n$ for all n implies that (2.14) must hold for all orders, and hence

$$(\psi(\alpha)|\psi(\beta)) = \delta_{\alpha\beta}\mathcal{D}(\alpha|\alpha) \tag{4.1}$$

Note that this quantity is defined for arbitrary α, β by the introduction of suitable numbers of trivial events ι. We will be able to choose a finite, complete orthogonal basis set where each vector is of this form

$$\phi(\alpha) = [\mathcal{D}(\alpha|\alpha)]^{-1/2}\psi(\alpha)$$

for a finite subset of α such that all $\mathcal{D}(\alpha|\alpha)$ are nonzero. It is then clear that any bounded operator X in $\mathcal{K}_1$ which satisfies $(\psi(\alpha)|X|\psi(\beta)) = 0 \ \forall \alpha \neq \beta$ will be diagonal in such an orthonormal basis, i.e. $X\phi(\alpha) = \xi(\alpha)\phi(\alpha)$, and is thus an element in a commutative subalgebra of $B(\mathcal{K}_1)$. From (4.1) follows that this conclusion must hold for each element $P(\gamma)$ and each time translate (2.13) for $m > 0$. As we have assumed the spectrum to be discrete the evolution is quasiperiodic and the same will hold for all integer m. The unitary time translation U is again a unitary operator in the subspace $\mathcal{K}_1$ but not diagonal except in trivial cases. The time translates (2.13) are then orthogonal projectors in $\mathcal{K}_1$, all diagonal in the same basis, hence commuting, and this implies that all the operators $M(\alpha) =$ (2.5) are projectors. From this and (3.13) follows that $S_\infty \leq \ln N$. □

This result can be reformulated in the following way. The argument of the theorem above serves to show that if (2.14) holds to all orders, then, from the completeness of the basis in $\mathcal{K}_1$, corresponding relations hold for the operators, not just the expectation values, for instance

$$P(\beta)M(\alpha)P(\gamma) = \delta_{\beta\gamma}P(\beta)M(\alpha)P(\beta)$$

In fact, the decoherence condition (2.12) for all orders m is sufficient, multiplication left and right with the projectors then leads to the equation above.

When $H_n > S_n$ from some n on, then there will be a constant $h > 0$ such that asymptotically $H_n \sim n \cdot h$, i.e. it is excluded that $H_n\{\alpha\}$ converges to a finite limit $H_\infty\{\alpha\} > S_\infty\{\alpha\}$.

Theorem 2 *If the increasing sequence $\{H_n(\alpha)\}$ converges to a finite value, then the time-translated projectors commute, hence $H_n\{\alpha\} = S_n\{\alpha\}$ and the limiting value is no larger than $\ln N$.*

Proof. For simplicity admit that the limit is achieved for a finite value of n (the general case can be approximated with arbitrary accuracy by increasing n). If the probability distribution $p(\alpha, \beta) = \rho(V_n(\alpha)^\dagger P(\beta)V_n(\alpha))$ has the same entropy as $p(\alpha) = \rho(V_n(\alpha)^\dagger V_n(\alpha))$ then it follows that for any fixed β either

$$\rho(V_n(\alpha)^\dagger P(\beta)V_n(\alpha)) = \rho(V_n(\alpha)^\dagger V_n(\alpha))$$

or it is zero. This means that $P(\beta)V_n(\alpha)|\Omega\rangle = V_n(\alpha)|\Omega\rangle$ or 0. We can iterate this argument to find that the same holds when the vectors $V_n(\alpha)|\Omega\rangle$ are acted on by operators $A = U^\dagger P(\gamma)UP(\beta)$ and so on, for higher order timeordered products. The vectors are all right eigenvectors of eigenvalue 0 or 1. All the operators $A^\dagger A$ have a norm not larger than one. On the other hand, if the eigenspaces of A are nonorthogonal then it is easy to construct vectors ϕ with $||A\phi|| > ||\phi||$. This shows that all the operators A have orthogonal eigenspaces and hence they are projectors. This means that all the time translates of the projectors $P(\beta)$ commute, and we obtain the desired result. □

V. THE COMMUTATIVE CASE II

In dealing with the case where the $P(\alpha)$ are not projectors but just non-negative operators we can no longer look to the vanishing of the offdiagonal terms in the QCK as a decoherence condition. Instead we must use (2.11) or the stronger relation which demands the corresponding condition for the offdiagonal elements of (2.6). Using the latter condition for all orders of the QCK, then an argument very similar to that of Theorem 1 above leads to a compatibility condition which is just (2.12) for all orders m. If we assume that the non-negative operators $P(\alpha)$ all commute, then they have a common spectral resolution Π_k

$$P(\alpha) = \sum_k p_k(\alpha)\Pi_k$$

where $\sum_\alpha p_k(\alpha)^2 = 1$. The compatibility condition (2.12) then reads: for all (k,l)

$$\sum_\beta p_k(\beta)p_l(\beta)\Pi_k U^{n\dagger}M(\alpha)U^n\Pi_l = \Pi_k U^{n\dagger}M(\alpha)U^n\Pi_l$$

This can hold if and only if $p_k(\beta) = p_l(\beta)$ for all β, and if the spectral resolution is defined to be nondegenerate this means that $k = l$, and all $M(\alpha)$ and their time translates are diagonal:

$$\Pi_k M(\alpha)\Pi_l = \delta_{kl}\Pi_k M(\alpha)$$

On the other hand the time translates of the $P(\alpha)$ cannot generate an algebra larger than that defined by the Π_k, consequently there is a decomposition of element of the POVM into a convex combination of projectors

$$M(\alpha) = \sum_k m_k(\alpha)\Pi_k \tag{5.1}$$

It seems desirable to have an expression for this kind of commutativity in terms of entropy functions just as in the projective case, and as an additional boon there will be, for the general case where the time translates of the observations do not commute, a quantitative measure of the lack of commutativity expressed in the information content. We will start with the following theorem.

Theorem 3 *Let $M(\alpha)$ satisfy (2.5) and let ρ,μ be two density operators with associated probability distributions $p(\alpha) = \rho(M(\alpha))$, $q(\alpha) = \mu(M(\alpha))$. With the notations of section II it then holds that*

$$H\{p(\alpha)|q(\alpha)\} \leq S(\rho|\mu)$$

with equality if and only if the density operators ρ and μ commute, the $M(\alpha)$ are projections commuting with both, and in addition the operators $\rho M(\alpha)$ and $\mu M(\alpha)$ are proportional for all α.

Proof. Start from the fact that a POVM (2.5) has a projective dilation [31,20]. This means that there is another Hilbert space $\mathcal{H} = \mathcal{K}\otimes\mathcal{K}_0$, a set of orthogonal projectors $Q(\alpha)$ in this space and a density operator ω in $\mathcal{K}_0$ such that the partial trace over $\mathcal{K}_0$ gives back $M(\alpha)$:

$$M(\alpha) = \mathrm{Tr}_0[\omega Q(\alpha)]$$

It is no essential restriction to take ω to be a tracial state. The second basic property is that

$$S(\rho\otimes\omega|\mu\otimes\omega) = S(\rho|\mu)$$

The third one is that the inequality above holds due to Proposition 1 in the Appendix. The problem is to find the precise conditions for equality. In general the projectors $Q(\alpha)$ have dimensions larger than one. We can then make the decomposition

$$S(\rho'|\mu') = H\{p(\alpha)|q(\alpha)\} + \sum_\alpha q(\alpha)S(\rho_{Q(\alpha)}|\mu_{Q(\alpha)})$$

where

$$p(\alpha)\rho_{Q(\alpha)} = Q(\alpha)(\rho\otimes\omega)Q(\alpha)$$

and

$$\rho' = \sum_\alpha p(\alpha)\rho_{Q(\alpha)} = (\rho \otimes \omega) \circ E$$

where E denotes the conditional expectation

$$E[X] = Q(\alpha)XQ(\alpha)$$

Consequently we find that

$$S(\rho'|\mu') = H\{p(\alpha)|q(\alpha)\}$$

if and only if $\rho_{Q(\alpha)} = \mu_{Q(\alpha)}$ for all α such that $q(\alpha) > 0$. This condition holds if the $Q(\alpha)$ are one-dimensional projectors, and conversely if these equalities hold, we will not get more information if the $Q(\alpha)$ are further subdivided into one-dimensional projectors. From Proposition 1 we know that

$$S(\rho'|\mu') \leq S(\rho \otimes \omega|\mu \otimes \omega)$$

so in order to get equality in the statement of the theorem we need the equality here. Now we can apply Proposition 2 in the Appendix. In the present case it says that equality holds in the inequality above if and only if the density operators ρ and μ commute, and in addition, for all real t it holds that $\rho^{it}\mu^{-it} \otimes \mathbb{1}$ is in the algebra generated by the $Q(\alpha)$, which means that there are non-negative λ_α such that

$$\rho^{it}\mu^{-it} \otimes \mathbb{1} = \sum_\alpha \lambda_\alpha^{it} Q(\alpha)$$

Then multiply with the density operator ω and take the partial trace over $\mathcal{K}_0$ to obtain

$$\rho^{it}\mu^{-it} = \sum_\alpha \lambda_\alpha^{it} M(\alpha)$$

It is then seen that each $M(\alpha)$ is itself a sum of projectors generating the common spectral resolution of ρ and μ, and satisfying $\rho M(\alpha) = \lambda_\alpha \mu M(\alpha)$. □

This result can now be applied to the case where μ is the tracial state. In the formulas above put

$$\rho = \rho(\beta) \propto V(\beta)\mu V(\beta)^\dagger$$
$$p(\alpha|\beta) = \mathrm{Tr}[\rho(\beta)M(\alpha)]$$

and $q(\alpha)$ defined as before. The α again represent n-sequences of observations and $M(\alpha) = V(\alpha)^\dagger V(\alpha)$ where the index n is left out for simplicity. The time parameters in the observations β are all chosen to be earlier than those in the α. We can now state the following result.

Theorem 4 *With the notations introduced above it holds that*

$$\lim_{n\to\infty} H_n\{p(\alpha|\beta)|q(\alpha)\} = S(\rho(\beta)|\rho)$$

if and only if the measurements at different times are compatible, in the sense of forming a commutative algebra of observables.

Proof. First the necessity. Due to the assumed finite dimension of the Hilbert space we can deal with the problem as if the limit was achieved for a finite value of n. Equality then holds if and only if the $M(\alpha)$ are orthogonal projectors commuting with $N(\beta) = V(\beta)V(\beta)^\dagger$ and such that

$$M(\alpha)N(\beta)M(\alpha) = \lambda(\alpha,\beta)M(\alpha)$$

for some non-negative constants $\lambda(\alpha,\beta)$. This means that for every β

$$N(\beta) = \sum_\alpha \lambda(\alpha,\beta)M(\alpha)$$

For an observation relating to one instant, this means that with $P(\beta)^2 = N(\beta)$ each time translate $U^{n\dagger}P(\beta)U^n$ must be in the algebra generated by the $M(\alpha)$, first for all n earlier than the time parameters in α, and then, by the quasiperiodic property of the dynamics, for all integer n.

Now for the sufficiency: if the observations all commute, then the equality will be achieved. We obtain the situation described in (5.1), and have to prove that for sufficiently large n each $M(\alpha)$ will be proportional to one of the projectors Π_k. The sesquilinear forms $V(\alpha)^\dagger V(\beta)$ create a convex set of operators with a natural partial order, and the $M(\alpha)$ form the positive cone. Because of the commutativity we have for all α, β

$$P(\beta)M(\alpha)P(\beta) = P(\beta)^2 M(\alpha) \leq M(\alpha)$$

which means that

$$M(\alpha) = \sum_\beta P(\beta)M(\alpha)P(\beta)$$

defines a convex decomposition of the POVM. Because of the finite dimension of the Hilbert space the process of decomposition must terminate in an extreme decomposition which can no more be decomposed in a nontrivial way. This means that for all α, β and n large enough

$$P(\beta)^2 M(\alpha) = \lambda_\beta(\alpha)M(\alpha)$$

so we can conclude that for an extreme decomposition we must have, for each α that $M(\alpha) \propto \Pi_k$ for some k. When μ is the tracial state and $\rho \propto P(\beta)\mu$ we find the conditions for equality in Theorem 3 are satisfied. The final conclusion is then that in the commutative case the equality in the present theorem is achieved in the limit $n \to \infty$. □

From these results we also obtain measure of the coherence associated with the set of histories defined by the the tracial state μ, the dynamics U and the observation $\{P(\alpha)\}$

$$\Delta(\rho) = \sum_\beta p_1(\beta)\,[S(\rho(\beta)|\rho) - H_\infty\{p(\alpha)|q(\alpha)\}] \geq 0 \tag{5.2}$$

By Theorem 4 it is zero precisely when the system is commutative, and a positive value is a sign of the existence of quantum coherence in the histories.

VI. CHAOTIC AND DETERMINISTIC HISTORIES

In both the previous sections the arguments depend on having an infinite sequence of observations, but the results do not tell us how many observations we really need to decide if the set of histories is nearly decoherent or not. Here there are quantitatively different types of behavior which can be associated with the concepts of regularity and chaos. In the theory of classical dynamical systems these two notions have a well defined qualitative meaning based on the asymptotic properties of the evolution. For finite quantum systems this is no longer so, there are characteristic time scales for the evolution, containing Planck's constant [32].

First it will be shown from a simple example that the vanishing of the coherence terms for correlation functions a finite order will not tell us much about the overall decoherence properties. We will give for each integer $M > 1$ a simple model with the following properties. There is a measurement of the projective type, a stationary dynamics and a stationary reference state. The coherence terms in the correlation vanishes for all correlation functions of order $n \leq 2M$, while the higher order correlations display coherence terms which are maximal in a well-defined sense.

We introduce a two-dimensional Hilbert space $\mathcal{H}$ in M copies and the tensor product

$$\mathcal{K} = \bigotimes_1^M \mathcal{H}_k.$$

A stationary discrete time dynamics is introduced by picking a unitary map V acting in $\mathcal{H}$ and defining the unitary U acting in $\mathcal{K}$ through

$$U(\phi_1 \otimes \phi_2 \otimes \cdots \phi_M) = V\phi_2 \otimes \cdots V\phi_M \otimes V\phi_1$$

for any set of vectors $\{\phi_k \in \mathcal{H}\}$. Iteration gives

$$U^M = V^M \otimes V^M \otimes \ldots \otimes V^M.$$

The stationary state ρ is picked to be the tracial state on $\mathcal{K}$ and the measurement that of the spin projection in a fixed direction. The entropy quantities defined in section III can be calculated, and we find

$$S_n\{\alpha\} = n \ln 2 \qquad \text{for } n \leq M$$

$$I_n\{\alpha\} = \begin{cases} n \ln 2 & \text{for } n \leq M \\ M \ln 2 & \text{for } n > M \end{cases}$$

The offdiagonal elements of (2.6) vanishes for $n \leq 2M$, hence

$$H_n\{\alpha\} = S_n\{\alpha\} \text{ for } n \leq 2M \tag{6.1}$$

but the higher orders depends on the properties of V. If V^M is a rotation which maps the direction vector of the measurement into an orthogonal direction, then we obtain

$$H_n\{\alpha\} = n \ln 2 \qquad \text{for all } n$$

$$S_n\{\alpha\} = \begin{cases} n \ln 2 & \text{for } n \leq 2M \\ 2M \ln 2 & \text{for } n > 2M \end{cases}$$

Consequently, in this case the coherence terms $H_n - S_n$ are of the maximal magnitude $(n - 2M) \ln 2$ for $n > 2M$. For other choices of V there is a slower increase of H_n and S_n with n, of course, down to the limiting case where $H_n = S_n = N \ln 2$ for all $n \geq M$.

The QCK for the model has the following simple structure for $n \leq M$ (and it can hold up to $n \leq 2M$ depending on the choice of V)

$$\mathcal{D}_n(\alpha|\beta) = \prod_{k=1}^{n} \mathcal{D}_1(\alpha_k|\beta_k) \tag{6.2}$$

This decomposition of the QCK into a product is well known from classical ergodic theory as being characteristic of Bernoulli processes [30]. For a quantum system it cannot hold for all n unless the system lives in an infinite dimensional Hilbert space. In addition the spectrum of the dynamics is then absolutely continuous with infinite multiplicity (as for the commutative case), so this type can be appropriate in the quantum context only for the description of the relaxation of a finite system in contact with an infinite heat bath (to be more precise, a heat bath at infinite temperature [25]). We have the following characterization of the factorization in terms of the entropy functions.

Theorem 5 *When the stationary state ρ and $\{P(\alpha)\}$ satisfy the invariance condition (2.10) then the factorization property (6.2) is satisfied for a given n (hence for all lower orders) precisely when it holds that*

$$S_n(\alpha) = n S_1(\alpha) \tag{6.3}$$

The proof is a direct consequence of Proposition 3 in the Appendix if we note that the compatibility conditions (2.9) and (2.10) can be used to define from $\mathcal{D}_n$, for any $m < n$, two complementary partial states which are $\mathcal{D}_m$ and $\mathcal{D}_{n-m}$. It is clear from (3.12) that (6.3), and hence (6.2), can hold at most for

$$n \leq 2 \ln N / S_1(\alpha)$$

($N = 2^M$ in the example above). This bound corresponds to the time scale (4.1) in [32].

It is a matter of interpretation if we want to accord any classical properties to the set of histories when (6.2) holds. This strong form of decoherence holds for the thermal fluctuations in equilibrium for time scales longer than the relaxation time, and this independently of the classical and quantum character of the system. It is just a reflection of the fact that the outcomes of successive observations are statistically independent because of an effective diffusion of the information throughout the system. From the example follows that the factorization for all orders up to a certain n does not tell us anything at all about the higher orders, and it does not help us in distinguishing between classical and quantum dynamical systems. In the following section we will see that most observables will decompose approximately according to (6.2) for the lowest order correlations, so this condition will introduce practically no restriction on the model unless there other restrictions which we have not treated in this context.

There can be decoherence from a completely different cause, namely that the outcomes of successive observations can be accurately predicted from a sufficient information of the past history of the system. The determinism holds in an asymptotic sense, as the initial state as chosen here will in general impart a certain lack of determinism for the lowest order correlations.

From (3.11) above it follows that

$$H_{m+1} - H_m \leq h \Longrightarrow H_{m+n} - H_m \leq mh \tag{6.4}$$

so if h is "small" in some sense, then H_n will continue to grow slowly for $n > m$. When the observations are given by projectors and $S_m \approx H_m$ this is enough to conclude that the quantum coherence will show up slowly in the probability distributions of the histories. This property will be applied to the example in section VIII. In the general, nonprojective case it is necessary to use the measure (5.2) for the coherence.

We will see in section VII that decoherence from statistical independence is holds true for almost any large system you can think of provided all observables are equally relevant. The deterministic property is a much more stringent decoherence condition. All the same it will be seen in section VIII that it is far from enough to single out a unique classical or macroscopic description.

VII. EXISTENCE OF APPROXIMATELY CHAOTIC HISTORIES

In this section the observations are defined by orthogonal projectors for simplicity, and we again take the stationary state to be tracial. It will be shown, by an argument that is not completely rigorous, that when the dimension of the Hilbert space is large enough and the discrete spectrum of the dynamics is nondegenerate and without higher order regularities, then most choices of the projectors will give correlation functions which are approximately decoherent for orders $n \ll \ln N$, where $N = \dim \mathcal{K}$. The decoherence will be of the chaotic type, i.e. the structure (6.2), (6.3) holds approximately for this range of n.

Start from the lowest order correlation function.

$$\begin{aligned} \mathcal{D}_2 &= \rho(PU^\dagger QUP) \\ U = U_t &= \sum_\omega |\omega)\,(\omega|\exp[-i\omega t] \end{aligned} \tag{7.1}$$

There are two essential ingredients in understanding the behavior of this quantity. The first is that the time parameter t represents a "macroscopic" time scale, which means that the function $\exp[-i\omega t]$ of the eigenvalues ω of the Hamiltonian (in the range which is relevant for a microcanonical ensemble) covers the unit circle a very large number of times. This assumption implies that the points are distributed essentially randomly on the unit circle. To be more precise, in the limit $N \to \infty$ they will look like a Poisson process on the circumference. On the other hand the time scale should be much smaller than that needed to observe the discreteness of the spectrum, but this restriction is not pertinent in this particular context. For the higher order correlations we used a discretized (unit) time step in section II. In the present context this should be interpreted as equality on a macroscopic time scale, but it would be meaningless to assert the equality of the time steps on a microscopic scale resolving all the relaxation processes in a macroscopic system.

The second important point is that matrix elements $(\omega|P|\omega')$ which will occur in the explicit expression for the correlation function can be dealt with as random variables if we introduce a suitable ensemble of systems. In order to get an overview of this problem we first consider the case where $U \in U(N)$, the group of unitary operators in the N dimensional Hilbert space. There is an invariant measure on the group which defines $U(N)$ as an ensemble of matrices. We denote by $\langle\ \rangle$ the average over this ensemble. It is not difficult to show, using the invariance properties of the measure, that it holds as an operator relation that

$$\langle U^\dagger QU \rangle = \rho(Q)\mathbb{1}$$

and hence when ρ is the tracial state

$$\langle \mathcal{D}_2 \rangle = \langle \rho(PU^\dagger QUP) \rangle = \rho(P)\rho(Q)$$

This result is only interesting in the present context if the fluctuations around the ensemble average is small under some useful conditions. In a fixed basis for the Hilbert space we write an operator relation

$$U^{\dagger}QU = \rho(Q)\mathbb{1} + \xi(Q)$$

The matrix elements of $\xi(Q)$ are random variables (functions on the ensemble $U(N)$) satisfying

$$\begin{aligned} \langle \xi(Q)_{kl} \rangle &= 0 \\ \langle |\xi(Q)_{kl}|^2 \rangle &= N^{-1}\rho(Q)[1-\rho(Q)] \\ \langle \xi(Q)_{kk}\xi(Q)_{ll} \rangle &= -N^{-2}\rho(Q)[1-\rho(Q)] \quad (k \neq l) \end{aligned} \tag{7.2}$$

For the QCK we find

$$\begin{aligned} \mathcal{D}_2 &= \rho(P)\rho(Q) + R_2 \\ R_2 &= \rho(P\xi(Q)P) \end{aligned} \tag{7.3}$$

and consequently $\langle R_2 \rangle = 0$. Using the ensemble averages displayed above we find the variance in leading order in N^{-1}

$$\langle |R_2|^2 \rangle = N^{-2}\rho(P)\rho(Q)(1-\rho(P))(1-\rho(Q)) \tag{7.4}$$

The $O(N^{-1})$ asymptotic behavior of the remainder term means that most elements in $U(N)$ give a value for $\mathcal{D}_2$ close to the ensemble average without averaging. There is thus a Chebychev inequality for the probability of a relative error ϵ

$$\text{Prob}\left\{ \frac{|\mathcal{D}_2 - \langle \mathcal{D}_2 \rangle|}{\langle \mathcal{D}_2 \rangle} \geq \epsilon \right\} \leq \frac{\langle |R_2|^2 \rangle}{\langle \mathcal{D}_2 \rangle^2 \epsilon^2} \leq \frac{1}{\rho(P)\rho(Q)\epsilon^2 N^2} \tag{7.5}$$

Now let U be fixed and consider the subset of $U(N)$ of elements with the spectrum of U, that is the subensemble

$$\{V^{\dagger}UV; V \in U(N)\} \tag{7.6}$$

The eigenvalue statistics of the ensemble $U(N)$ for large N is well known and locally similar to that of $GUE(N)$ [33]. The eigenvalues are distributed over the unit circle in a highly uniform way. If the spectrum of the chosen operator U conforms to this picture (making a statistics over the spectrum of this single operator) then the average over the ensemble (7.6) will be essentially the same as that over $U(N)$. Actually it is not important that the spectrum has the full rigidity of $U(N)$ with a strong level repulsion. A tedious calculation using the results of Mello [34] on averages over $U(N)$ shows that an ensemble average of $\mathcal{D}_2$ over (7.6) differs from $\langle \mathcal{D}_2 \rangle$ by a term of relative magnitude no larger than $|\rho(U)|^2$ which is $O(N^{-2})$ if the spectrum of U is uniform enough and consequently negligible compared to R_2.

Instead of the ensemble of evolution operators we can equivalently take U to be fixed and consider the ensemble of observables

$$\{V^{\dagger}P(\alpha)V; V \in U(N)\} \tag{7.7}$$

and obtain statements valid for a generic choice of observables. The previous argument is now applied to a fixed choice of evolution operator (7.1) with a macroscopic time t which gives a spectrum for U without extra structure or regularities. The results for ensemble averages and variances are then same as when the evolution operator is drawn at random from (7.6). Again the approximate factorization (7.3) holds with the order of magnitude estimate (7.4) of the rest term as $N \to \infty$. As already indicated above, in this particular case it is possible top make a more explicit calculation using averages over the ensemble $U(N)$ [34]. However, in dealing with higher order correlations such an approach would become very cumbersome as the technique for performing averages over the unitary group is complex and only give manageable exact results for the lowest order products. In a nonrigorous way the general structure can be found by the argument sketched below, where the essential point is to find an estimate of where the factorization must begin to break down.

In order to deal with the higher order correlations start by considering a general correlation function of the form

$$\mathcal{D} = \rho(V^{\dagger}U^{\dagger}MUV)$$

where $V = V(\alpha)$, $M = M(\beta)$ are operators of the type (2.3) and (2.5) respectively. Using the $U(N)$ ensemble for U we find again the ensemble average and variance

$$\mathcal{D} = \rho(V^\dagger V)\rho(M) + R$$
$$\langle R\rangle = 0$$
$$\langle R^2\rangle = N^{-2}\sigma_1^2\sigma_2^2$$

where $\sigma_1^2 = \rho(V^\dagger V V^\dagger V) - \rho(V^\dagger V)^2, \sigma_2^2 = \rho(M^2) - \rho(M)^2$, and for the relative magnitude of the deviation from the product form

$$\frac{\sqrt{\langle R^2\rangle}}{\langle\mathcal{D}\rangle} = \frac{\sigma_1\sigma_2}{\rho(V^\dagger V)\rho(M)N} \tag{7.8}$$

The factorization property clearly breaks down when this quantity is of order 1. The worst case is where the operators $V^\dagger V$ and M are projectors, say of dimension N_1 and N_2, respectively. We then obtain the order of magnitude estimate

$$\frac{\sqrt{\langle R^2\rangle}}{\langle\mathcal{D}\rangle} = \frac{1}{\sqrt{N_1 N_2}}$$

and a corresponding estimate of the probability in (7.5) which is $(N_1 N_2\epsilon^2)$. For simplicity let the partition (2.1) of orthogonal projectors satisfy $p(\alpha) = \rho(P(\alpha)) = \exp(-h)$ for all α and some $h > 0$. This means that $H_1(\alpha) = h$. If $V = V_m(\alpha)$, $M = V_n(\alpha)^\dagger V_n(\alpha)$ then an explicit bound on the deviation from product form is obtained by putting

$$N_1 = \exp(-mh)N, \qquad N_2 = \exp(-nh)N$$

and the condition for the smallness of the error term reads $(m+n)h \ll \ln N$.

We can now repeat the same type of estimate for each evolution operator in a multi-time correlation function provided that all the operators can be chosen independently. For the correlation function $\mathcal{D}_n$ there is then an $(n-1)$-tuple of elements in $U(N)$, and $\langle\mathcal{D}_n\rangle$ is of the form (6.2) if we average over the corresponding $n-1$ ensembles. Furthermore, there is a generalization of (7.5), where the RHS can be estimated by adding the terms coming from the $n-1$ unitary operators.

Finally this is applied to a sequence of unitary operators $U(t_1), U(t_2), \ldots, U(t_{n-1})$ which are generally widely different as operators but where the t_k are nevertheless equal on a macroscopic time scale. The final result is that given the ensemble of observables (7.7), the probability that $\mathcal{D}_n$ can be approximated by a product form (6.2) is close to 1 as long as $nh \ll \ln N$.

VIII. EXISTENCE OF APPROXIMATELY DETERMINISTIC HISTORIES

Again consider a set of observations defined by orthogonal projectors and choose the tracial state as stationary state. Introduce the Hilbert spaces indexed by sequences of symbols in $\mathcal{I}$

$$\mathcal{K}_n(\alpha) = [V_n(\alpha)\mathcal{K}]^-$$

We will show that $\mathcal{K}_n(\alpha) \perp \mathcal{K}_n(\beta)$ for $\alpha \neq \beta$ if and only if the operators (2.13) commute for $m = 0, \ldots, n-1$. The sufficiency is clear. For necessity we find from the orthogonality of the Hilbert subspaces that

$$P(\gamma)P_m(\beta)P(\alpha) = \delta_{\alpha\gamma}P_m(\beta)P(\alpha)$$

which implies the commutativity $P(\gamma)P_m(\beta) = P_m(\beta)P(\gamma)$. It is immediately clear that

$$\{\mathcal{K}_n(\alpha); \alpha \in \mathcal{I}^n\}_{n=1}^{\infty}$$

form a lattice of subspaces. In fact, in the commutative algebra generated by the projectors (2.13) there is a basis of minimal, but not necessarily 1-dimensional projectors. The dynamics maps this basis set into itself, hence it is a permutation of the minimal projectors (and of the corresponding subspaces). Only projectors of the same rank can belong to the same orbit under the action of the dynamics. The permutation can be decomposed into cycles, and it is no real restriction to consider only cyclic permutations. Such a setup clearly defines a deterministic process in an asymptotic sense.

We now turn to a more general situation where there is no exact commutativity for observations at different instants and no exact periodicity of the kind described above. It will be shown that given a discrete time dynamics with a sufficiently dense point spectrum on the unit circle represented by independent random variables, we can construct orthogonal projectors and cyclic permutations of the corresponding subspaces which approximate the dynamics during a finite number of iterations. The choice of projectors can be made in many, mutually noncommuting ways.

Let the Hilbert space $\mathcal{K}$ have dimension $M \cdot N$. This restriction is not essential, it is just there to make the computations easier. The unitary map U representing the unit step dynamics then has $M \cdot N$ eigenvalues on the unit circle. For the reasons given in the previous section they can be assumed to be independent and uniformly distributed random variables. They are renumbered and written in the form $\exp(2\pi i X_k)$, where the X_k are ordered

$$0 \leq X_1 \leq X_2 \ldots \leq X_{MN} \leq 1$$

Introduce the notation ϕ_k for the eigenvectors:

$$U\phi_k = \exp(2\pi i X_k)\phi_k$$

Define another basis set in the Hilbert space as follows:

$$\psi_{m,\alpha} = \frac{1}{\sqrt{N}} \sum_{p=0}^{N-1} \exp(2\pi i\alpha p/N)\phi_{m+pM}$$

for $m = 1, \ldots, M$, $\alpha = 1, \ldots, N$. These vectors form an orthonormal set

$$(\psi_{m,\alpha}|\psi_{m',\alpha'}) = \delta_{mm'} \frac{1}{N} \sum_p \exp(2\pi i(\alpha - \alpha')p/N) = \delta_{mm'}\delta_{\alpha\alpha'}$$

and from the inverse relation

$$\phi_{m+pM} = \frac{1}{\sqrt{N}} \sum_{\alpha=1}^{N} \exp(-2\pi i\alpha p/N)\psi_{m,\alpha}$$

follows that the set is complete. A set of orthogonal projectors are introduced

$$P(\alpha) = \sum_{m=1}^{M} |\psi_{m,\alpha})(\psi_{m,\alpha}| \tag{8.1}$$

and corresponding subspaces $\mathcal{K}(\alpha) = P(\alpha)\mathcal{K}$. In order to find the action of U on $\mathcal{K}(\alpha)$ consider the expression

$$(\psi_{m,\alpha+1}|U|\psi_{m,\alpha}) = \frac{1}{N}\exp(2\pi i X_m)\sum_p \exp\{2\pi i(X_{m+pM} - X_m - p/N)\}$$

Averaging over the tracial state we obtain

$$\begin{aligned}\rho(P(\alpha)U^\dagger P(\alpha+1)UP(\alpha)) &= (MN)^{-1} \sum_{p,q,\alpha} |(\psi_{m,\alpha+1}|U|\psi_{m,\alpha})|^2 \\ &= N^{-3}\sum_{p,q} \exp\{2\pi i(X_{(p-q)M} - (p-q)/N)\}\end{aligned} \tag{8.2}$$

The random variable X_{pM} has a beta distribution with first and second moment ([35], sections I.7 and III.3)

$$\begin{aligned}\langle X_{pM}\rangle &= \frac{pM}{MN+1} \\ \langle X_{pM}^2\rangle - \langle X_{pM}\rangle^2 &= \frac{pM[M(N-p)+1]}{(MN+1)^2(MN+2)}\end{aligned}$$

We can then calculate the expectation over the ensemble defined by the random variables X_{pM}

$$\langle \rho(P(\alpha)U^\dagger P(\alpha+1)UP(\alpha))\rangle = N^{-3}\sum_{p,q}\Phi((p-q)M, MN+1; 2\pi i)\exp(-2\pi i(p-q)/N)$$

where $\Phi(a,b;x)$ is the confluent hypergeometric (Kummer's) function. Standard asymptotic formulas (6.13.2(17) in [36]) show that

$$\lim_{M\to\infty}\frac{1}{N^3}\sum_{p}\Phi((p-q)M, MN+1; 2\pi i)\exp(-2\pi i(p-q)/N) = \frac{1}{N} = \rho(P(\alpha)) = p_1(\alpha)$$

The leading correction to the asymptotic behavior can be bounded by a more direct estimate as follows. Clearly

$$\rho(P(\alpha)U^\dagger P(\alpha+1)UP(\alpha)) \leq \rho(P(\alpha))$$

and from (8.2) follows for the ensemble average of the difference

$$\langle \rho(P(\alpha)) - \rho(P(\alpha)U^\dagger P(\alpha+1)UP(\alpha))\rangle \leq \frac{1}{N^3}\sum_{p,q}(2\pi)^2\langle |X_{(p-q)M} - (p-q)/N|^2\rangle$$

The right hand side should be much smaller than $p_1(\alpha)$, so using the leading term for large $M, N \gg 1$ the following condition is obtained

$$\Delta \equiv (2\pi)^2\frac{1}{N^2}\sum_{p,q}\frac{|p-q|(N-|p-q|)}{MN^3} \approx \frac{2\pi^2}{MN} \ll 1 \tag{8.3}$$

Using the ensemble averaged probability distributions (2.7) we then find that

$$\sum_{\beta\neq\alpha+1} p_2(\alpha,\beta) \leq p_1(\alpha)\Delta$$

The largest value for $H\{p_2\}$ is obtained when there is a uniform distribution over the $N-1$ values $\beta \neq \alpha+1$. Consequently it holds that

$$\begin{aligned} H_2\{\alpha\} &= H_1\{\alpha\} + \delta H = \ln N + \delta H \\ \delta H &\leq -\Delta\ln(\Delta/N) - (1-\Delta)\ln(1-\Delta) \end{aligned}$$

By the inequalities (3.11) we then know that

$$H_n\{\alpha\} \leq H_1\{\alpha\} + (n-1)\delta H$$

and this means that the increase in the coherence terms $H_n(\alpha) - S_n(\alpha)$ with n, for arbitrarily large n, is slow when $\Delta \ll 1$.

There are many ways of choosing the set of orthogonal projectors in a different way, such that they do not commute with the choice made above, but still have the properties under the evolution shown above to a good approximation. Instead of choosing eigenstates of U as in (8.1), first make a unitary transformation mixing the eigenstates corresponding to a small interval δ of the phase. As long as δ^2 is much smaller than the variance of the random variables X_k the properties of the evolution will not change significantly in the aspects which we consider here. Thus we can choose $\delta \approx (MN)^{-1/2}$, and as this interval contains a number $(MN)^{1/2}$ of levels, there is a very large number of choices when M and N are large.

IX. DISCUSSION

The conclusions to be drawn from the abstract results above are rather negative. One of them is certainly that decoherence in itself is quite insufficient to select a set of observables as a "classical" domain. Decoherence is generic (section VII) and is clearly present in equilibrium thermal fluctuations, also for observables with no obvious classical property. It seems to be more promising to look for significant classical aspects of the macroworld in the existence of structures which are departures from complete equilibrium but are metastable over a long time scale. This metastability can be interpreted as a particular case of determinism. Thus we should consider the deterministic property of section VI to be more relevant than decoherence in defining a macroscopic, classical domain.

In order to gain some physical insight and moderate the formal arguments, it is useful to recall some earlier work and ideas on the subject. One of the basic examples of the emergence of classical properties in relatively small quantum systems is that of molecular shape [37,8,38–40]. The shape is not a property of eigenstates of the Hamiltonian which on the contrary belong to irreducible representations of the symmetry group of the system. There is no real mystery here, when the density of levels is high enough, then the initial conditions and the method of observing the system decides what you will see just as much as the intrinsic properties of the Hamiltonian. There is one point of view that the interaction with the environment and the resulting decoherence defines the quantum states with a definite shape as a basis preferred to those of a definite symmetry. An alternative, but not necessarily contradictory, point of view is that the remarkable fact to be explained is the long term stability of certain nonstationary states, like those of a definite shape, in environments satisfying certain conditions (like a bound on the temperature) and a large class of interactions (where some form of locality will be essential).

Note that in equilibrium quantum statistical mechanics of infinite systems it is possible to introduce in a rigorous way some classical properties. In the quasilocal approach there is defined an algebra of observables at infinity which is commutative [41]. These observables can be said to define superselection rules, as the relative phases between different eigenspaces are unobservable. They describe global properties of an infinite system which can be observed by local observables but also outside any finite part of the system, e.g. the total charge. The drawback of this formalism is that these observables always are constants of the motion. There can be no dynamics for the global observables when the dynamics is generated in a quasilocal way. In the context of explaining the emergence of classical properties we should be interested in models which are not strictly infinite and which have the potential for nontrivial dynamics for variables or states which are "nearly" classical. Two different types of situations can be distinguished here.

There is the classical dynamics of the "correspondence limit" (limit of large quantum numbers) of a simple quantum system, like a quantum harmonic oscillator or quantum rotator. For a large mass (or moment of inertia) we have a classical situation of a kind we understand rather well. The nearly classical observables are very coarse quantum measurements simultaneously of coordinate and momentum. Here it is convenient to use observations which are not described by projectors but the more general form (2.1), eventually with a continuous rather than a discrete variable α. The relevant information measure is (3.7), and we can use a relatively small value of (5.2) as a sign of decoherence for the chosen set of observations. Introducing a heat bath here to destroy some quantum phases seems superfluous. As shown in section VIII the problem here is that there will be many nonequivalent choices of observations.

There is also the slow evolution of metastable observables like the geometric shape of molecules, where there is no reason to expect any obvious similarity with classical Hamiltonian dynamics. If there is a well-defined classical limit it will be given, e.g. by equations of the Fokker-Planck type. The essential classicality of these observables lies in an information-theoretical aspect rather than a dynamical one. The information stored is stable and can be observed and copied without destroying it. This property means that there is a possibility of several observers making observations and comparing the results, implying that this piece of information has an objective reality. One desirable result would be a proof, under some conditions yet to be specified, that all the information contained in the quantum state of a sufficiently complex quantum system which is stable over a very long time scale has this "classical" property to a good approximation. It should be clear from the kind of calculation done in section VIII that this cannot be true without additional assumptions. A likely candidate for such an extra input is a finite-system counterpart of the local structure used in the algebraic theory of infinite systems.

APPENDIX: ENTROPY INEQUALITIES

Let T be a positive linear map of $B(\mathcal{K})$ into itself which maps the unit operator on itself. There is then an associated map on the density operators: $\rho \mapsto \rho \circ T$. For a subclass of such maps the relative entropy has the following property, which can be interpreted as a general H-theorem.

Proposition 1 *Let T satisfy the Schwarz inequality*

$$T[X]^\dagger T[X] \leq T[X^\dagger X] \tag{A1}$$

and $T[\mathbb{1}] = \mathbb{1}$. Then, for any two density operators ρ, μ such that their relative entropy is defined, it holds that

$$S(\rho \circ T|\mu \circ T) \leq S(\rho|\mu) \tag{A2}$$

The proof follows most easily from the variational formula for the relative entropy proved by Kosaki [42]. For completely positive maps T the result is closely related to the strong subadditivity of the quantum entropy [43,44,28].

We are also interested in a particular situation when the equality holds in (A2). Let there be a set of orthogonal projectors Q_k, $\sum_k Q_k = \mathbb{1}$ and denote by τ_k the tracial state in the subspace $Q_k\mathcal{K}$. The map $E : X \mapsto \sum_k \tau_k(X)Q_k$ satisfies (A1), in fact it is a completely positive idempotent map into the subalgebra $\mathcal{M}$ generated by the projectors Q_k. Petz [45,46] proved a theorem which in the present context has the following form.

Proposition 2 *For two density operators ρ, μ the equality*

$$S(\rho \circ E|\mu \circ E) = S(\rho|\mu)$$

holds if and only if the density operators ρ, μ commute and the unitary operators $\rho^{it}\mu^{-it}$ are in $\mathcal{M}$ for all t.

The statement by Petz assumes that the density operators are nondegenerate (the states are faithful), but the result is extended in a straightforward way to states which are not necessarily faithful. Note that the RHS in the equality is finite only if the support projection of μ is contained in that of ρ, so these operators certainly commute, they must be in $\mathcal{M}$ and the part of ρ outside the support of μ does not contribute to the entropy.

Proposition 3 *Let ρ be a density operator in the Hilbert space $\mathcal{K}_1$. There is then a pure (vector) state ω in the tensor product space $\mathcal{K}_1 \otimes \mathcal{K}_2$, where $\mathcal{K}_1 \simeq \mathcal{K}_2$, such that $\omega_1 \equiv \mathrm{Tr}_2\, \omega = \rho$. Furthermore, for any pure state on such a tensor product, the two partial states are isometric, i.e. they have the same spectrum of nonzero eigenvalues with multiplicity, and consequently the same entropy: $\omega_1 \simeq \omega_2$, $S(\omega_1) = S(\omega_2)$. For a general mixed state μ on the tensor product the following triangle inequality holds (and all similar relations obtained by permutation of the three density operators):*

$$|S(\mu_1) - S(\mu_2)| \leq S(\mu) \leq S(\mu_1) + S(\mu_2) \tag{A3}$$

Equality holds in the second relation if and only if $\mu = \mu_1 \otimes \mu_2$.

The proof of this result is given in [28], section II.F. One simple representation of ω is as follows: if $\rho = \sum_k p_k |k)(k|$ then we can choose $\omega = |\Omega)(\Omega|$ where

$$|\Omega) = \sum_k \sqrt{p_k}\, |k) \otimes |k) \tag{A4}$$

Proposition 4 *With the notation of section II it holds that*

$$I_n\{\alpha\} \leq S_n\{\alpha\} \leq H_n\{\alpha\}$$

and the second equality holds if and only if the QCK (2.6) is diagonal.

Proof. From ρ we construct ω in $\mathcal{K} \otimes \mathcal{K}$ space according to (A4). From this we obtain a pure state $\omega_n^\sharp = V_n(\beta)\omega V_n(\alpha)^\dagger$ in a still larger space $\mathcal{K} \otimes \mathcal{K} \otimes h_{n,\mathcal{I}}$ where $h_{n,\mathcal{I}} = (3.1)$. From $\omega^\sharp$ we define two complementary partial traces, the first by tracing over $\mathcal{K} \otimes \mathcal{K}$, the other by summing over $\mathcal{I}$. The two partial states are $\mathcal{D}_n$ and

$$\sum_\alpha V_n(\alpha)\omega V_n(\alpha)^\dagger \equiv \sum_\alpha p_n(\alpha)\omega_n(\alpha) = \omega_n^\flat$$

We then have the equality

$$S_n\{\alpha\} = S(\omega_n^\flat) = \sum_\alpha p_n(\alpha) S(\omega_n(\alpha)|\omega_n^\flat)$$

and we find from Proposition 1

$$S_n\{\alpha\} \geq \sum_\alpha p_n(\alpha) S(\sigma_n(\alpha)|\rho) = I_n\{\alpha\}$$

The second part of the inequality of the statement follows from the fact that a deletion of the offdiagonal matrix elements cannot decrease the entropy. In fact, such a deletion $\mu \mapsto \mu \circ E$ is an idempotent CP map and one finds that $\ln(\mu \circ E) = E[\ln(\mu \circ E)]$ and hence that

$$S(\mu \circ E|\mu) = -S(\mu) - \mathrm{Tr}[\mu \ln(\mu \circ E)] = S(\mu \circ E) - S(\mu) \geq 0$$

where the equality holds precisely when $\mu \circ E = \mu$.

Proposition 5 S_n *and* I_n *are nondecreasing in* n*:*

$$S_n\{\alpha\} \leq S_{n+1}\{\alpha\}$$
$$I_n\{\alpha\} \leq I_{n+1}\{\alpha\}$$

Proof. In the proof of Proposition 4 the state $\omega_{n+1}^\flat$ is obtained from $\omega_n^\flat$ through composing a unitary transformation, which leaves the entropy invariant, with the map

$$\rho \circ T = \sum_\alpha P(\alpha) \rho P(\alpha)$$

where $\{P(\alpha)\}$ satisfy (2.1). This map satisfies (A1) and leaves the tracial state μ invariant: $\mu \circ T = \mu$. From (A2) follows that $S(\rho \circ T|\mu) \leq S(\rho|\mu)$. But a simple calculation gives $S(\rho|\mu) = S(\rho) - S(\mu)$ and consequently that $S(\rho) \leq S(\rho \circ T)$, which proves the first statement. For the second, consider a set of density operators of the form

$$p(\alpha, \beta)\sigma(\alpha, \beta) = \sqrt{\rho} V(\alpha)^\dagger V(\beta)^\dagger V(\beta) V(\alpha) \sqrt{\rho}$$

and sum over the later outcomes

$$\sum_\beta p(\alpha, \beta)\sigma(\alpha, \beta) = \sqrt{\rho} V(\alpha)^\dagger V(\alpha) \sqrt{\rho} = p(\alpha)\sigma(\alpha)$$

It is then found that

$$p(\alpha) S(\sigma(\alpha)) - \sum_\beta p(\alpha, \beta) S(\sigma(\alpha, \beta)) = \sum_\beta p(\alpha, \beta) S(\sigma(\alpha, \beta)|\sigma(\alpha)) \geq 0$$

and consequently

$$S(\rho) - \sum_\alpha p(\alpha) S(\sigma(\alpha)) \leq S(\rho) - \sum_{\alpha,\beta} p(\alpha, \beta) S(\sigma(\alpha, \beta)|\sigma(\alpha))$$

which is the first result.

Proposition 6 *With the notation of section II*

$$S_n\{\alpha\} \leq S(\rho) + S(\rho_n)$$

Proof. We first note that the map $\rho \mapsto V(\alpha)\rho V(\beta)$ is an isometric map from density operators in $\mathcal{K}$ to density operators in a Hilbert space $\mathcal{K} \otimes h_{n,\mathcal{I}}$, so it preserves the entropy. Partial traces then gives $\mathcal{D}_n$ and ρ_n as partial states, and from the triangle inequality (A3) the statement follows.

Proposition 7 *For any convex decomposition of a state* $\rho = \sum p_k \rho_k$ *it holds that* -

$$\sum p_k S(\rho_k) \leq S(\rho) \leq \sum p_k S(\rho_k) + H\{p_k\} \tag{A5}$$

It follows that, with the definition (3.4) and suppressing the index n*,*

$$\sum_\alpha p(\alpha) S(\rho(\alpha)) \leq S(\rho) \leq \sum_\alpha p(\alpha) S(\rho(\alpha)) + H\{\alpha\} \tag{A6}$$

Proof. The proof of (A5) can be found in section II.B of [28]. We remark that the first equality holds if and only if all $\rho_k = \rho$, while the second equality holds if and only if the ρ_k are orthogonal. For the second statement we construct from ρ a pure state ω as described in Proposition 3, and define the pure state $\omega(\alpha)$ by

$$p(\alpha)\omega(\alpha) = V(\alpha)\omega V(\alpha)^\dagger$$

The partial states are $\rho(\alpha)$ and the transpose (in the basis used in (A4)) of the density operator $\sigma(\alpha)$ defined in analogy with (3.5). But a transposition leaves the entropy invariant, so by Proposition 3 $S(\rho(\alpha)) = S(\sigma(\alpha))$. Now, from (3.6) we can use (A5) to obtain (A6).

---

[1] R. Omnès: *Consistent interpretations of quantum mechanics.* Rev. Mod. Phys. **64**(2), 339–382 (1992)
[2] M. Gell-Mann and J.B. Hartle: *Quantum mechanics in the light of quantum cosmology.* In: W. Zurek (ed), *Complexity, entropy and the physics of information.* Redwood City: Addison-Wesley 1990, pp 425–458
[3] M. Gell-Mann and J.B. Hartle: *Classical Equations for Quantum Systems.* Phys. Rev. D **47**(8), 3345–3382 (1993)
[4] J.B. Hartle: *Spacetime quantum mechanics and the quantum mechanics of spacetime.* Lectures given at the 1992 Les Houches École d'été, Gravitation et Quantifications, July 9–17, 1992
[5] H.F. Dowker and J.J. Halliwell: *Quantum mechanics of history: The decoherence functional in quantum mechanics.* Phys. Rev. D **46**(4) 1580–1609 (1992)
[6] R.B. Griffiths: *Consistent histories and the interpretation of quantum mechanics.* J. Stat. Phys. **36**(1-2), 219–272 (1984)
[7] R.B. Griffiths: *Consistent interpretation of quantum mechanics using quantum trajectories.* Phys. Rev. Lett. **70**(15), 2201–04 (1993)
[8] E. Joos and H.D. Zeh: *The emergence of classical properties through interaction with the environment.* Z. Phys. B **59**, 223–243 (1985)
[9] W.H. Zurek: *Environment-induced superselection rules.* Phys. Rev. D **26**(8) 1862–1880 (1982)
[10] W.H. Zurek: *Decoherence and the transition from quantum to classical.* Physics Today **44**(10), 36–44 (1991), and the resulting correspondence published in the April 1993 issue of Physics Today
[11] J.P. Paz, S. Habib and W.H. Zurek: *Reduction of the wave packet: Preferred observables and decoherence time scale.* Phys. Rev. D **47**(2), 488–501 (1993)
[12] J.P. Paz and W.H. Zurek: *Environment-induced decoherence, classicality and consistency of quantum histories.* Phys. Rev. D **48**(6), 2728–2738 (1993)
[13] W.H. Zurek: *Preferred states, predictability, classicality and the environment-induced decoherence.* Progr. Theor. Phys. **89**(2), 281–312 (1993)
[14] K. Kraus: *States, Effects and Operations.* Berlin: Springer-Verlag 1983
[15] V.I. Paulsen: *Completely bounded maps and dilations.* Longman Scientific & Technical 1986
[16] E.B. Davies: *Quantum theory of open systems.* New York, London: Academic Press 1976
[17] G. Lindblad: *Non-Markovian quantum stochastic processes and their entropy.* Commun. Math. Phys. **65**, 281–294 (1979)
[18] M. Ozawa: *On information gain by quantum measurements of continuous observables.* J. Math. Phys. **27**(3), 759–763 (1986)
[19] C.W. Helstrom: *Quantum Detection and Estimation Theory.* New York: Academic Press 1976
[20] A.S. Holevo: *Probabilistic and Statistical Aspects of Quantum Theory.* Amsterdam: North-Holland 1982
[21] R.F. Streater and A.S. Wightman: *PCT, Spin & Statistics and All That.* New York: W.A. Benjamin 1964
[22] R. Haag: *Local Quantum Physics.* Berlin: Springer-Verlag 1992
[23] R.G. Glauber: *Coherent and incoherent states of the radiation field.* Phys. Rev. **131**, 2766–88 (1963)
[24] C.W. Gardiner: *Quantum Noise.* Berlin: Springer-Verlag 1991
[25] G. Lindblad: *Determinism and randomness in quantum dynamics.* J. Phys. A. **26**, 7193–7211 (1993)
[26] G. Lindblad: *On the generators of quantum dynamical semigroups.* Commun. Math. Phys. **48**, 119–130 (1976)
[27] G. Lindblad: *Completely positive maps and entropy inequalities.* Commun. Math. Phys. **40**, 147–151 (1975)
[28] A. Wehrl: *General properties of entropy.* Rev. Mod. Phys. **50**, 221–260 (1978)
[29] M. Ohya and D. Petz: *Quantum Entropy and its Use.* Berlin: Springer-Verlag 1993
[30] K. Petersen: *Ergodic Theory.* Cambridge University Press 1983
[31] B. Sz.-Nagy and C. Foias: *Harmonic Analysis of Operators on Hilbert Space.* Amsterdam: North-Holland 1970

[32] B.V. Chirikov: *The problem of quantum chaos.* In: W.D. Heiss (ed), *Chaos and quantum chaos.* Lecture Notes in Physics **411**, Berlin: Springer-Verlag 1992, pp 1–56
[33] M.L. Mehta: *Random matrices*, 2nd edition. San Diego: Academic Press 1991
[34] P.A. Mello: *Averages on the unitary group and applications to the problem of disordered conductors.* J. Phys. A **23**, 4061–4080 (1990)
[35] W. Feller: *An introduction to probability theory and applications.* Vol 2, 2nd edition. New York: Wiley 1971
[36] A. Erdelyi (ed): *Higher Transcendental Functions*, Vol 1. New York: MacGraw-Hill 1953
[37] R.G. Woolley: *Quantum theory and molecular structure.* Adv. Phys. **25**(1), 27–52 (1976)
[38] H. Primas: *Chemistry, Quantum Mechanics and Reductionism.* Berlin: Springer-Verlag 1981
[39] P. Claverie and G. Jona-Lasinio: *Instability of tunneling and the concept of molecular structure in quantum mechanics: The case of pyramidal molecules and the enantiomer problem.* Phys. Rev. A **33**(4), 2245–53 (1986)
[40] W. Gans, A. Blumen and A. Amann (eds): *Large-scale molecular systems : quantum and stochastic aspects - beyond the simple molecular picture.* NATO ASI B258, Plenum Press 1991
[41] O. Bratteli, D. Robinson: *Operator Algebras and Quantum Statistical Mechanics.* Vol I. New York: Springer-Verlag 1981
[42] H. Kosaki: *Relative entropy of states: a variational expression.* J. Operator Theory **16**, 123–140 (1986)
[43] E.H. Lieb: *Convex trace functions and the Wigner-Yanase-Dyson conjecture.* Adv. Math. **11**, 267–288(1973)
[44] E.H. Lieb and M.B. Ruskai: *Proof of the strong subadditivity of quantum-mechanical entropy.* J. Math. Phys. **14**, 1938–41 (1973)
[45] D. Petz: *Properties of quantum entropy*, in: L. Accardi and W. von Waldenfels (eds), *Quantum Probability and Applications II.* Lecture Notes in Mathematics **1136**), Berlin: Springer-Verlag 1985, pp 428–441
[46] D. Petz: *Sufficient subalgebras and the relative entropy of states on a von Neumann algebra.* Commun. Math. Phys. **105**, 123–131 (1986)