REVIEWS OF MODERN PHYSICS, VOLUME 75, JULY 2003

Decoherence, einselection, and the quantum origins of the classical

Wojciech Hubert Zurek

Theory Division, LANL, Mail Stop B210, Los Alamos, New Mexico 87545

(Published 22 May 2003)

The manner in which states of some quantum systems become effectively classical is of great significance for the foundations of quantum physics, as well as for problems of practical interest such as quantum engineering. In the past two decades it has become increasingly clear that many (perhaps all) of the symptoms of classicality can be induced in quantum systems by their environments. Thus decoherence is caused by the interaction in which the environment in effect monitors certain observables of the system, destroying coherence between the *pointer states* corresponding to their eigenvalues. This leads to environment-induced superselection or *einselection*, a quantum process associated with selective loss of information. Einselected pointer states are stable. They can retain correlations with the rest of the universe in spite of the environment. Einselection enforces classicality by imposing an effective ban on the vast majority of the Hilbert space, eliminating especially the flagrantly nonlocal "Schrödinger-cat states." The classical structure of phase space emerges from the quantum Hilbert space in the appropriate macroscopic limit. Combination of einselection with dynamics leads to the idealizations of a point and of a classical trajectory. In measurements, einselection replaces quantum entanglement between the apparatus and the measured system with the classical correlation. Only the preferred pointer observable of the apparatus can store information that has predictive power. When the measured quantum system is microscopic and isolated, this restriction on the predictive utility of its correlations with the macroscopic apparatus results in the effective "collapse of the wave packet." The existential interpretation implied by einselection regards observers as open quantum systems, distinguished only by their ability to acquire, store, and process information. Spreading of the correlations with the effectively classical pointer states throughout the environment allows one to understand "classical reality" as a property based on the relatively objective existence of the einselected states. Effectively classical pointer states can be "found out" without being re-prepared, e.g, by intercepting the information already present in the environment. The redundancy of the records of pointer states in the environment (which can be thought of as their "fitness" in the Darwinian sense) is a measure of their classicality. A new symmetry appears in this setting. Environment-assisted invariance or *envariance* sheds new light on the nature of ignorance of the state of the system due to quantum correlations with the environment and leads to Born's rules and to reduced density matrices, ultimately justifying basic principles of the program of decoherence and einselection.

CONTENTS

I. INTRODUCTION

The interpretation of quantum theory has been an issue ever since its inception. Its tone was set by the discussions of Schrödinger (1926, 1935a, 1935b), Heisenberg (1927), and Bohr (1928, 1949; see also Jammer, 1974; Wheeler and Zurek, 1983). Perhaps the most incisive critique of the (then new) theory was that of Einstein, who, searching for inconsistencies, distilled the essence of the conceptual difficulties of quantum mechanics through ingenious gedanken experiments. We owe to him and Bohr clarification of the significance of quantum indeterminacy in the course of the Solvay Congress debates (see Bohr, 1949) and elucidation of the nature of quantum entanglement (Bohr, 1935; Einstein, Podolsky, and Rosen, 1935; Schrödinger, 1935a, 1935b). The issues they identified then are still a part of the subject.

Within the past two decades, the focus of research on the fundamental aspects of quantum theory has shifted from esoteric and philosophical to more down to earth as a result of three developments. To begin with, many of the old gedanken experiments [such as the Einstein-Podolsky-Rosen (EPR) "paradox"] became compelling demonstrations of quantum physics. More or less simultaneously the role of decoherence began to be appreciated and einselection was recognized as key to the emergence of classicality. Last but not least, various developments led to a new view of the role of information in physics. This paper reviews progress in the field with a focus on decoherence, einselection, and the emergence of classicality, and also attempts to offer a preview of the future of this exciting and fundamental area.

A. The problem: Hilbert space is big

The interpretation problem stems from the vastness of Hilbert space, which, by the principle of superposition, admits arbitrary linear combinations of any states as a possible quantum state. This law, thoroughly tested in the microscopic domain, bears consequences that defy classical intuition: It appears to imply that the familiar classical states should be an exceedingly rare exception. And, naively, one may guess that the superposition principle should always apply literally: Everything is ultimately made out of quantum "stuff." Therefore there is no *a priori* reason for macroscopic objects to have definite position or momentum. As Einstein noted[1] localization with respect to macrocoordinates is not just independent of, but incompatible with, quantum theory. How, then, can one establish a correspondence between the quantum and the familiar classical reality?

1. Copenhagen interpretation

Bohr's solution was to draw a border between the quantum and the classical and to keep certain objects—especially measuring devices and observers—on the classical side (Bohr, 1928, 1949). The principle of superposition was suspended "by decree" in the classical domain. The exact location of this border was difficult to pinpoint, but measurements "brought to a close" quantum events. Indeed, in Bohr's view the classical domain was more fundamental. Its laws were self-contained (they could be confirmed from within) and established the framework necessary to define the quantum.

The first breach in the quantum-classical border appeared early: In the famous Bohr-Einstein double-slit debate, quantum Heisenberg uncertainty was invoked by Bohr at the macroscopic level to preserve wave-particle duality. Indeed, since the ultimate components of classical objects are quantum, Bohr emphasized that the boundary must be moveable, so that even the human nervous system could be regarded as quantum, provided that suitable classical devices to detect its quantum features were available. In the words of Wheeler (1978, 1983), who has elucidated Bohr's position and decisively contributed to the revival of interest in these matters, "No [quantum] phenomenon is a phenomenon until it is a recorded (observed) phenomenon."

[1]In a letter dated 1954, Albert Einstein wrote to Max Born, "Let ψ_1 and ψ_2 be solutions of the same Schrödinger equation.... When the system is a macrosystem and when ψ_1 and ψ_2 are 'narrow' with respect to the macrocoordinates, then in by far the greater number of cases this is no longer true for $\psi = \psi_1 + \psi_2$. Narrowness with respect to macrocoordinates is not only *independent* of the principles of quantum mechanics, but, moreover, *incompatible* with them." [The translation from Born (1969) quoted here is due to Joos (1986), p. 7].

This is a pithy summary of a point of view—known as the Copenhagen interpretation (CI)—that has kept many a physicist out of despair. On the other hand, as long as no compelling reason for the quantum-classical border could be found, the CI universe would be governed by two sets of laws, with poorly defined domains of jurisdiction. This fact has kept many students, not to mention their teachers, awake at night (Mermin 1990a, 1990b, 1994).

2. Many-worlds interpretation

The approach proposed by Everett (1957a, 1957b) and elucidated by Wheeler (1957), DeWitt (1970), and others (see Zeh, 1970, 1973; DeWitt and Graham, 1973; Geroch, 1984; Deutsch, 1985, 1997; Deutsch *et al.*, 2001) was to enlarge the quantum domain. Everything is now represented by a unitarily evolving state vector, a gigantic superposition splitting to accommodate all the alternatives consistent with the initial conditions. This is the essence of the many-worlds interpretation (MWI). It does not suffer from the dual nature of the Copenhagen interpretation. However, it also does not explain the emergence of classical reality.

The difficulty many have in accepting the many-worlds interpretation stems from its violation of the intuitively obvious "conservation law"—that there is just one universe, the one we perceive. But even after this question is dealt with, many a convert from the Copenhagen interpretation (which claims the allegiance of a majority of physicists) to the many-worlds interpretation (which has steadily gained popularity; see Tegmark and Wheeler, 2001, for an assessment) eventually realizes that the original many-worlds interpretation does not address the preferred-basis question posed by Einstein (see footnote 1) (see Bell, 1981, 1987; Wheeler, 1983; Stein, 1984; Kent, 1990, for critical assessments of the many-worlds interpretation). And as long as it is unclear what singles out preferred states, perception of a unique outcome of a measurement and, hence, of a single universe cannot be explained either.[2]

In essence, the many-worlds interpretation does not address, but only postpones, the key question. The quantum-classical boundary is pushed all the way towards the observer, right against the border between the material universe and the consciousness, leaving it at a very uncomfortable place to do physics. The many-worlds interpretation is incomplete: it does not explain what is effectively classical and why. Nevertheless, it was a crucial conceptual breakthrough. Everett reinstated quantum mechanics as a basic tool in the search for its interpretation.

[2]DeWitt, in the many-worlds reanalysis of quantum measurements, makes this clear: in DeWitt and Graham (1973), the last paragraph of p. 189, he writes about the key "remaining problem." "Why is it so easy to find apparata in states [with a well defined value of the pointer observable]? In the case of macroscopic apparata it is well known that a small value for the mean square deviation of a macroscopic observable is a fairly stable property of the apparatus. But how does the mean square deviation become so small in the first place? Why is a large value of the mean-square deviation of a macroscopic observable virtually never, in fact, encountered in practice? ... a proof of this does not yet exist. It remains a program for the future."

B. Decoherence and einselection

Environment can destroy coherence between the states of a quantum system. This is *decoherence*. According to quantum theory, every superposition of quantum states is a legal quantum state. This egalitarian quantum principle of superposition applies in isolated systems. However, not all quantum superpositions are treated equally by decoherence. Interaction with the environment will typically single out a preferred set of states. These *pointer states* remain untouched in spite of the environment, while their superpositions lose phase coherence and decohere. Their name—pointer states—originates from the context of quantum measurements, where they were originally introduced (Zurek, 1981). They are the preferred states of the pointer of the apparatus. They are stable and, hence, retain a faithful record of and remain correlated with the outcome of the measurement in spite of decoherence.

Einselection is this decoherence-imposed selection of the preferred set of pointer states that remain stable in the presence of the environment. As we shall see, einselected pointer states turn out to have many classical properties. *Einselection* is an accepted nickname for environment-induced superselection (Zurek, 1982).

Decoherence and einselection are two complementary views of the consequences of the same process of environmental monitoring. Decoherence is the destruction of quantum coherence between preferred states associated with the observables monitored by the environment. Einselection is its consequence—the *de facto* exclusion of all but a small set, a *classical domain* consisting of pointer states—from within a much larger Hilbert space. Einselected states are distinguished by their resilience—stability in spite of the monitoring environment.

The idea that the "openness" of quantum systems might have anything to do with the transition from quantum to classical was ignored for a very long time, probably because in classical physics problems of fundamental importance were always settled in isolated systems. In the context of measurements, Gottfried (1966) anticipated some of the later developments. The fragility of energy levels of quantum systems was emphasized by the seminal papers of Zeh (1970, 1973), who argued [inspired by remarks relevant to what would be called today "deterministic chaos" (Borel, 1914)] that macroscopic quantum systems are in effect impossible to isolate.

The understanding of how the environment distills the classical essence from quantum systems is more recent (Zurek, 1981, 1982, 1993a). It combines two observations: (1) In quantum physics, "reality" can be attributed to the measured states. (2) Information transfer usually associated with measurements is a common result of almost any interaction of a system with its environment.

Some quantum states are resilient to decoherence. This is the basis of einselection. Using Darwinian analogy, one might say that pointer states are the most "fit." They survive monitoring by the environment to leave "descendants" that inherit their properties. The classical domain of pointer states offers a static summary of the result of quantum decoherence. Save for classical dynamics, (almost) nothing happens to these einselected states, even though they are immersed in the environment.

It is difficult to catch einselection in action. Environment has little effect on the pointer states, since they are already classical. Therefore it is easy to miss the decoherence-driven dynamics of einselection by taking for granted its result—existence of the classical domain and a ban on arbitrary quantum superpositions. Macroscopic superpositions of einselected states disappear rapidly. Einselection creates effective superselection rules (Wick, Wightman, and Wigner, 1952, 1970; Wightman, 1995). However, in the microscopic domain, decoherence can be slow in comparison with the dynamics.

Einselection is a quantum phenomenon. Its essence cannot even be motivated classically. In classical physics, arbitrarily accurate measurements (also by the environment) can, in principle, be carried out without disturbing the system. Only in quantum mechanics acquisition of information inevitably brings the risk of altering—of repreparation of the state of the system.

The quantum nature of decoherence and the absence of classical analogs are a source of misconceptions. For instance, decoherence is sometimes equated with relaxation or classical noise that can be introduced by the environment. Indeed, all of these effects often appear together and as a consequence of "openness." The distinction between them can be briefly summed up: Relaxation and noise are caused by the environment perturbing the system, while decoherence and einselection are caused by the system perturbing the environment.

Within the past few years decoherence and einselection have become familiar to many. This does not mean that their implications are universally accepted (see comments in the April 1993 issue of *Physics Today*; d'Espagnat, 1989, 1995; Bub, 1997; Leggett, 1998, 2002; the exchange of views between Anderson, 2001, and Adler, 2003; Stapp, 2002). In a field where controversy has reigned for so long this resistance to a new paradigm is no surprise.

C. The nature of the resolution and the role of envariance

Our aim is to explain why the quantum universe appears classical when it is seen "from within." This question can be motivated only in the context of a universe divided into systems, and must be phrased in the language of the correlations between systems. The Schrödinger equation dictates deterministic evolution;

$$|\Psi(t)\rangle = \exp(-iHt/\hbar)|\Psi(0)\rangle, \tag{1.1}$$

and, in the absence of systems, the problem of interpretation seems to disappear.

There is simply no need for "collapse" in a universe with no systems. Our experience of the classical reality does not apply to the universe as a whole, seen from the outside, but to the systems within it. Yet, the division into systems is imperfect. As a consequence, the universe is a collection of open (interacting) quantum systems. Since the interpretation problem does not arise in quantum theory unless interacting systems exist, we shall also feel free to assume that an environment exists when looking for a resolution.

Decoherence and einselection fit comfortably in the context of the many-worlds interpretation in which they define the "branches" of the universal state vector. Decoherence makes the many-worlds interpretation complete: It allows one to analyze the universe as it is seen by an observer, who is also subject to decoherence. Einselection justifies elements of Bohr's Copenhagen interpretation by drawing the border between the quantum and the classical. This natural boundary can sometimes be shifted. Its effectiveness depends on the degree of isolation and on the manner in which the system is probed, but it is a very effective quantum-classical border nevertheless.

Einselection fits either the MWI or the CI framework. It sets limits on the extent of the quantum jurisdiction, delineating how much of the universe will appear classical to observers who monitor it from within, using their limited capacity to acquire, store, and process information. It allows one to understand classicality as an idealization that holds in the limit of macroscopic open quantum systems.

The environment imposes superselection rules by preserving part of the information that resides in the correlations between the system and the measuring apparatus (Zurek, 1981, 1982). The observer and the environment compete for information about the system. The environment—because of its size and its incessant interaction with the system—wins that competition, acquiring information faster and more completely than the observer. Thus a record useful for the purpose of prediction must be restricted to observables that are already monitored by the environment. In that case, the observer and the environment no longer compete and decoherence becomes unnoticeable. Indeed, typically observers use the environment as a communication channel, and monitor it to find out about the system.

The spreading of information about the system through the environment is ultimately responsible for the emergence of "objective reality." The objectivity of a state can be quantified by the redundancy with which it is recorded throughout the universe. Intercepting fragments of the environment allows observers to identify (pointer) states of the system without perturbing it (Zurek, 1993a, 1998a, 2000; see especially Sec. VII of this paper for a preview of this new "environment as a witness" approach to the interpretation of quantum theory).

When an effect of a transformation acting on a system can be undone by a suitable transformation acting on the environment, so that the joint state of the two remains unchanged, the transformed property of the system is said to exhibit "environment-assisted invariance" or *envariance* (Zurek, 2003b). The observer must obviously be ignorant of the envariant properties of the system. Pure entangled states exhibit envariance. Thus, in quantum physics, perfect information about the joint state of the system-environment pair can be used to prove ignorance of the state of the system.

Envariance offers a new fundamental insight into what is information and what is ignorance in the quantum world. It leads to Born's rule for the probabilities and justifies the use of reduced density matrices as a description of a part of a larger combined system. Decoherence and einselection rely on reduced density matrices. Envariance provides a fundamental resolution of many of the interpretational issues. It will be discussed in Sec. VI.D.

D. Existential interpretation and quantum Darwinism

What the observer knows is inseparable from what the observer is: the physical state of his memory implies his information about the universe. The reliability of this information depends on the stability of its correlation with external observables. In this very immediate sense decoherence brings about the apparent collapse of the wave packet: after a decoherence time scale, only the einselected memory states will exist and retain useful correlations (Zurek, 1991, 1998a, 1998b; Tegmark, 2000). The observer described by some specific einselected state (including a configuration of memory bits) will be able to access ("recall") only that state. The collapse is a consequence of einselection and of the one-to-one correspondence between the state of the observer's memory and of the information encoded in it. Memory is simultaneously a description of the recorded information and part of an "identity tag," defining the observer as a physical system. It is as inconsistent to imagine the observer perceiving something other than what is implied by the stable (einselected) records in his possession as it is impossible to imagine the same person with a different DNA. Both cases involve information encoded in a state of a system inextricably linked with the physical identity of an individual.

Distinct memory/identity states of the observer (which are also his "states of knowledge") cannot be superposed. This censorship is strictly enforced by decoherence and the resulting einselection. Distinct memory states label and inhabit different branches of Everett's many-worlds universe. The persistence of *correlations* between the records (data in the possession of the observers) and the recorded states of macroscopic systems is all that is needed to recover "familiar reality." In this manner, the distinction between ontology and epistemology—between what is and what is known to be—is dissolved. In short (Zurek, 1994), there can be *no information without representation*.

There is usually no need to trace the collapse of the wave packet all the way to the observer's memory. It suffices that the states of a decohering system quickly evolve into mixtures of preferred (pointer) states. All that can be known, in principle, about a system (or even, introspectively, about an observer by himself) is its decoherence-resistant identity tag—a description of its einselected state.

Apart from this essentially negative function as a censor, the environment also plays a very different role as a broadcasting agent, relentlessly cloning the information about the einselected pointer states. This role of the environment as a witness in determining what exists was not appreciated until very recently. Over the past two decades, the study of decoherence has focused on the effect of the environment on the system. This led to a multitude of technical advances, which we shall review, but it also missed one crucial point of paramount conceptual importance: observers monitor systems indirectly, by intercepting small fractions of their environments (e.g., a fraction of the photons that have been reflected or emitted by the object of interest). Thus, if an understanding of why we perceive the quantum universe as classical is the principal aim, our study should focus on the information spread throughout the environment. This leads one away from the models of measurement inspired by the von Neumann chain (von Neumann, 1932) to studies of information transfer involving conditional dynamics and the resulting branching and "fan-out" of information throughout the environment (Zurek, 1983, 1998a, 2000). This view of the role of the environment, known as "quantum Darwinism" because of the analogy between the selective amplification of the information concerning pointer observables and the reproduction which is key to natural selection, is complementary to the usual image of the environment as a source of perturbations that destroy the quantum coherence of the system. It suggests that the redundancy of the imprint of a system in the environment may be a quantitative measure of its relative *objectivity* and hence of the classicality of quantum states. Quantum Darwinism is discussed in Sec. VII of this review.

The benefits of recognizing the role of environment include not just an operational definition of the objective existence of the einselected states, but—as is also detailed in Sec. VI—a clarification of the connection between quantum amplitudes and probabilities. Einselection converts arbitrary states into mixtures of well-defined possibilities. Phases are envariant. Appreciation of envariance as a symmetry tied to ignorance about the state of the system was the missing ingredient in the attempts of no-collapse derivation of Born's rule and its probability interpretation. While both envariance and quantum Darwinism are only beginning to be investigated, the extension of the program of einselection they offer allows one to understand the emergence of "classical reality" from the quantum substrate as a fundamental consequence of quantum laws and goes far beyond the "for all practical purposes" only view of the role of the environment.

II. QUANTUM MEASUREMENTS

The need for a transition from the quantum determinism of the global state vector to the classical definiteness of states of individual systems is traditionally illustrated by the example of quantum measurements. The outcome of a generic measurement of the state of a quantum system is *not* deterministic. In the textbook discussions, this random element is blamed on the "collapse of the wave packet," invoked whenever a quantum system comes into contact with a classical apparatus. In a fully quantum discussion this issue still arises, in spite of (or rather because of) the overall deterministic quantum evolution of the state vector of the universe. As pointed out by von Neumann (1932), there is no room for a real collapse in the purely unitary models of measurements.

A. Quantum conditional dynamics

To illustrate the difficulties, consider a quantum system $\mathcal{S}$ initially in a state $|\psi\rangle$ interacting with a quantum apparatus $\mathcal{A}$ initially in a state $|A_0\rangle$:

$$|\Psi_0\rangle=|\psi\rangle|A_0\rangle=\left(\sum_i a_i|s_i\rangle\right)|A_0\rangle \rightarrow\sum_i a_i|s_i\rangle|A_i\rangle=|\Psi_t\rangle. \tag{2.1}$$

Above, $\{|A_i\rangle\}$ and $\{|s_i\rangle\}$ are states in the Hilbert spaces of the apparatus and of the system, respectively, and a_i are complex coefficients. The conditional dynamics of such *premeasurement*, as the step achieved by Eq. (2.1) is often called, can be accomplished by means of a unitary Schrödinger evolution. Yet it is not enough to claim that a measurement has been achieved. Equation (2.1) leads to an uncomfortable conclusion: $|\Psi_t\rangle$ is an EPR-like entangled state. Operationally, this EPR nature of the state emerging from the premeasurement can be made more explicit by rewriting the sum in a different basis:

$$|\Psi_t\rangle=\sum_i a_i|s_i\rangle|A_i\rangle=\sum_i b_i|r_i\rangle|B_i\rangle. \tag{2.2}$$

This freedom of basis choice—basis ambiguity—is guaranteed by the principle of superposition. Therefore, if one were to associate states of the apparatus (or the observer) with decompositions of $|\Psi_t\rangle$, then even before inquiring about the specific outcome of the measurement one would have to decide on the decomposition of $|\Psi_t\rangle$; a change of the basis redefines the measured quantity.

1. Controlled NOT and bit-by-bit measurement

The interaction required to entangle a measured system and the measuring apparatus, Eq. (2.1), is a generalization of the basic logical operation known as a "controlled NOT" or a c-NOT. A classical c-NOT changes the state a_t of the target when the control is 1, and does nothing otherwise:

$$0_c a_t\rightarrow 0_c a_t;\quad 1_c a_t\rightarrow 1_c\neg a_t. \tag{2.3}$$

The quantum c-NOT is a straightforward quantum version of Eq. (2.3). It was known as a "bit-by-bit measurement" (Zurek, 1981, 1983) and already used to elucidate the connection between entanglement and premeasurement before it acquired its present name and significance in the context of quantum computation (see, for example, Nielsen and Chuang, 2000). Arbitrary superpositions of the control bit and of the target bit states are allowed:

$$(\alpha|0_c\rangle+\beta|1_c\rangle)|a_t\rangle\rightarrow\alpha|0_c\rangle|a_t\rangle+\beta|1_c\rangle|\neg a_t\rangle. \tag{2.4}$$

Here negation, $|\neg a_t\rangle$, of a state is basis dependent:

$$\neg(\gamma|0_t\rangle+\delta|1_t\rangle)=\gamma|1_t\rangle+\delta|0_t\rangle. \tag{2.5}$$

With $|A_0\rangle=|0_t\rangle$, $|A_1\rangle=|1_t\rangle$ we have an obvious analogy between a c-NOT and a premeasurement.

In the classical controlled NOT, the direction of information transfer is consistent with the designations of the two bits. The state of the control remains unchanged while it influences the target, Eq. (2.3). Classical measurement need not influence the system. Written in the logical basis $\{|0\rangle,|1\rangle\}$, the truth table of the quantum c-NOT is essentially—save for the possibility of superpositions—the same as Eq. (2.3). One might have anticipated that the direction of information transfer and the designations (control/system and target/apparatus) of the two qubits would also be unambiguous, as in the classical case. This expectation is incorrect. In the conjugate basis $\{|+\rangle,|-\rangle\}$ defined by

$$|\pm\rangle=(|0\rangle\pm|1\rangle)/\sqrt{2}, \tag{2.6}$$

the truth table, Eq. (2.3) (as such equations providing a map from the inputs to the outputs of the logic gates are known), along with Eq. (2.6), lead to a new complementary truth table:

$$|\pm\rangle|+\rangle\rightarrow|\pm\rangle|+\rangle, \tag{2.7}$$

$$|\pm\rangle|-\rangle\rightarrow|\mp\rangle|-\rangle. \tag{2.8}$$

In the complementary basis $\{|+\rangle,|-\rangle\}$, the roles of the control and of the target are reversed. The former target (basis $\{|0\rangle, |1\rangle\}$)—represented by the second ket above—remains unaffected, while the state of the former control (the first ket) is conditionally flipped.

In the bit-by-bit case the measurement interaction is

$$H_{int}=g|1\rangle\langle 1|_{\mathcal{S}}|-\rangle\langle -|_{\mathcal{A}} =\frac{g}{2}|1\rangle\langle 1|_{\mathcal{S}}\otimes[\mathbf{1}-(|0\rangle\langle 1|+|1\rangle\langle 0|)]_{\mathcal{A}}. \tag{2.9}$$

Here g is a coupling constant, and the two operators refer to the system (i.e., to the former control) and to the apparatus pointer (the former target), respectively. It is easy to see that the states $\{|0\rangle,|1\rangle\}_{\mathcal{S}}$ of the system are unaffected by H_{int}, since

$$[H_{int}, e_0|0\rangle\langle 0|_{\mathcal{S}}+e_1|1\rangle\langle 1|_{\mathcal{S}}]=0. \tag{2.10}$$

The measured observable $\hat{\epsilon}=e_0|0\rangle\langle 0|+e_1|1\rangle\langle 1|$ is a constant of motion under H_{int}. The c-NOT requires interaction time t such that $gt=\pi/2$.

The states $\{|+\rangle,|-\rangle\}_{\mathcal{A}}$ of the apparatus encode information about phases between the logical states. They have exactly the same "immunity:"

$$[H_{int}, f_+|+\rangle\langle+|_{\mathcal{A}}+f_-|-\rangle\langle-|_{\mathcal{A}}]=0, \tag{2.11}$$

where $f_\pm$ are the eigenvalues of this "phase observable." Hence, when the apparatus is prepared in a definite phase state (rather than in a definite pointer/logical state), it will pass its phase on to the system, as Eqs. (2.7) and (2.8) show. Indeed, H_{int} can be written as

$$H_{int}=g|1\rangle\langle 1|_{\mathcal{S}}|-\rangle\langle-|_{\mathcal{A}}$$
$$=\frac{g}{2}[\mathbf{1}-(|-\rangle\langle+|+|+\rangle\langle-|)]_{\mathcal{S}}\otimes|-\rangle\langle-|_{\mathcal{A}} \tag{2.12}$$

making this immunity obvious.

This basis-dependent direction of information flow in a quantum c-NOT (or in a premeasurement) is a consequence of complementarity. While the information about the observable with the eigenstates $\{|0\rangle,|1\rangle\}$ travels from the system to the measuring apparatus, in the complementary $\{|+\rangle,|-\rangle\}$ basis it seems that the apparatus is measured by the system. This observation (Zurek 1998a, 1998b; see also Beckman *et al.*, 2001) clarifies the sense in which phases are inevitably "disturbed" in measurements. They are not really destroyed, but rather, as the apparatus measures a certain observable of the system, the system simultaneously "measures" phases between the possible outcome states of the apparatus. This leads to loss of phase coherence. Phases become "shared property," as we shall see in more detail in the discussion of envariance.

The question "what measures what?" (decided by the direction of the information flow) depends on the initial states. In the classical practice this ambiguity does not arise. Einselection limits the set of possible states of the apparatus to a small subset.

2. Measurements and controlled shifts

The truth table of a whole class of c-NOT-like transformations that includes general premeasurement, Eq. (2.1), can be written as

$$|s_j\rangle|A_k\rangle\rightarrow|s_j\rangle|A_{k+j}\rangle. \tag{2.13}$$

Equation (2.1) follows when $k=0$. One can therefore model measurements as controlled shifts—c-shifts—or generalizations of the c-NOT. In the bases $\{|s_j\rangle\}$ and $\{|A_k\rangle\}$, the direction of information flow appears to be unambiguous—from the system $\mathcal{S}$ to the apparatus $\mathcal{A}$. However, a complementary basis can be readily defined (Ivanovic, 1981; Wootters and Fields, 1989):

$$|B_k\rangle=N^{-1/2}\sum_{l=0}^{N-1}\exp\left(\frac{2\pi i}{N}kl\right)|A_l\rangle. \tag{2.14a}$$

Above, N is the dimensionality of the Hilbert space. An analogous transformation can be carried out on the basis $\{|s_i\rangle\}$ of the system, yielding states $\{|r_j\rangle\}$.

Orthogonality of $\{|A_k\rangle\}$ implies

$$\langle B_l|B_m\rangle=\delta_{lm}, \tag{2.15}$$

$$|A_k\rangle=N^{-1/2}\sum_{l=0}^{N-1}\exp\left(-\frac{2\pi i}{N}kl\right)|B_l\rangle, \tag{2.14b}$$

the inverse of the transformation of Eq. (2.14a). Hence

$$|\psi\rangle=\sum_l \alpha_l|A_l\rangle=\sum_k \beta_k|B_k\rangle, \tag{2.16}$$

where the coefficients β_k are

$$\beta_k=N^{-1/2}\sum_{l=0}^{N-1}\exp\left(-\frac{2\pi i}{N}kl\right)\alpha_l. \tag{2.17}$$

The Hadamard transform of Eq. (2.6) is a special case of the more general transformation considered here.

To implement the truth tables involved in premeasurements, we define observable $\hat{A}$ and its conjugate:

$$\hat{A}=\sum_{k=0}^{N-1}k|A_k\rangle\langle A_k|, \tag{2.18a}$$

$$\hat{B}=\sum_{l=0}^{N-1}l|B_l\rangle\langle B_l|. \tag{2.18b}$$

The interaction Hamiltonian

$$H_{int}=g\hat{s}\hat{B} \tag{2.19}$$

is an obvious generalization of Eqs. (2.9) and (2.12), with g the coupling strength and

$$\hat{s}=\sum_{l=0}^{N-1}l|s_l\rangle\langle s_l|. \tag{2.20}$$

In the $\{|A_k\rangle\}$ basis, $\hat{B}$ is a shift operator,

$$\hat{B}=\frac{iN}{2\pi}\frac{\partial}{\partial\hat{A}}. \tag{2.21}$$

To show how H_{int} works, we compute

$$\exp(-iH_{int}t/\hbar)|s_j\rangle|A_k\rangle$$
$$=|s_j\rangle N^{-1/2}\sum_{l=0}^{N-1}\exp[-i(jgt/\hbar+2\pi k/N)l]|B_l\rangle. \tag{2.22}$$

We now adjust the coupling g and the duration of the interaction t so that the action ι expressed in Planck units $2\pi\hbar$ is a multiple of $1/N$:

$$\iota=gt/\hbar=G*2\pi/N. \tag{2.23a}$$

For an integer G, Eq. (2.22) can be readily evaluated:

$$\exp(-iH_{int}t/\hbar)|s_j\rangle|A_k\rangle=|s_j\rangle|A_{\{k+G*j\}_N}\rangle. \tag{2.24}$$

This is a shift of the apparatus state by an amount $G*j$ proportional to the eigenvalue j of the state of the system. G plays the role of gain. The index $\{k+G*j\}_N$ is evaluated mod$N$, where N is the number of possible outcomes, that is, the dimensionality of the Hilbert space of the apparatus pointer $\mathcal{A}$. When $G*j>N$, the pointer will just rotate through the initial zero. The truth table for $G=1$ defines a c-shift, Eq. (2.13), and with $k=0$ leads to a premeasurement, Eq. (2.1).

The form of the interaction, Eq. (2.19), in conjunction with the initial state, decides the direction of information transfer. Note that—as was the case with the c-NOT's—the observable that commutes with the interaction Hamiltonian will not be perturbed:

$$[H_{int},\hat{s}]=0. \tag{2.25}$$

$\hat{s}$ commutes with H_{int} and is therefore a nondemolition observable (Braginsky, Vorontsov, and Thorne, 1980; Caves *et al.*, 1980; Bocko and Onofrio, 1996).

3. Amplification

Amplification has often been regarded as the process forcing quantum potentialities to become classical reality. An example of it is the extension of the measurement model described above.

Assume the Hilbert space of the apparatus pointer is large compared with the space spanned by the eigenstates of the measured observable $\hat{s}$:

$$N=\dim(\mathcal{H}_{\mathcal{A}})\gg\dim(\mathcal{H}_{\mathcal{S}})=n. \tag{2.26}$$

Then one can increase ι to an integer multiple G of $2\pi/N$, Eqs. (2.23a) and (2.24). However, larger ι will lead to redundancy only when the Hilbert space of the apparatus has many more dimensions than possible outcomes. Otherwise, only "wrapping" of the same record will ensue. The simplest example of such wrapping, $(\text{c-NOT})^2$, is the identity operation. For $N\gg n$, however, one can attain gain:

$$G=Ngt/2\pi\hbar. \tag{2.23b}$$

The outcomes are now separated by $G-1$ empty eigenstates of the record observable. In this sense, $G\gg1$ achieves redundancy, providing that wrapping of the record is avoided. This is guaranteed when

$$nG<N. \tag{2.27}$$

Amplification is useful in the presence of noise. For example, it may be difficult to initiate the apparatus in $|A_0\rangle$, so the initial state may be a superposition:

$$|a_l\rangle=\sum_k \alpha_l(k)|A_k\rangle. \tag{2.28a}$$

Indeed, typically a mixture of such superpositions,

$$\rho_{\mathcal{A}}^0=\sum_i w_i|a_i\rangle\langle a_i|, \tag{2.28b}$$

may be the starting point for a premeasurement. Then

$$|s_k\rangle\langle s_{k'}|\rho_{\mathcal{A}}=|s_k\rangle\langle s_{k'}|\sum_l w_l|a_l\rangle\langle a_l|$$
$$\rightarrow|s_k\rangle\langle s_{k'}|\sum_l w_l|a_l+Gk\rangle\langle a_{l+Gk'}|, \tag{2.29}$$

where $|a_{l+Gk}\rangle$ obtains from $|a_l\rangle$, Eq. (2.28a), through

$$|a_{l+Gk}\rangle=\sum_j \alpha_l(j)|A_{j+Gk}\rangle, \tag{2.30}$$

and the simplifying assumption about the coefficients,

$$\alpha_l(j)=\alpha(j-l), \tag{2.31}$$

has been made. The aim of this simplification is to focus on the case when the apparatus states are peaked around a certain value l [e.g., $\alpha_l(j)\sim\exp\{-(j-l)^2/2\Delta^2\}$], and where the form of their distribution over $\{|A_k\rangle\}$ does not depend on l.

A good measurement allows one to distinguish states of the system. Hence it must satisfy

$$|\langle a_{l+Gk}|a_{l+Gk'}\rangle|^2=|\sum_j \alpha[j+G(k-k')]\alpha^*(j)|^2$$
$$\approx\delta_{k',k}. \tag{2.32}$$

States of the system that need to be distinguished should rotate the pointer of the apparatus to the correlated outcome states that are approximately orthogonal. When the coefficients $\alpha(k)$ are peaked around $k=0$ with dispersion Δ, this implies

$$\Delta\ll G. \tag{2.33}$$

In the general case of an initial mixture, Eq. (2.29), one can evaluate the dispersion of the expectation value of the record observable $\hat{A}$ as

$$\langle\hat{A}^2\rangle-\langle\hat{A}\rangle^2=\mathrm{Tr}\rho_{\mathcal{A}}^0\hat{A}^2-(\mathrm{Tr}\rho_{\mathcal{A}}^0\hat{A})^2. \tag{2.34}$$

The outcomes are distinguishable when

$$\langle\hat{A}^2\rangle-\langle\hat{A}\rangle^2\ll G. \tag{2.35}$$

Interaction with the environment yields a mixture of the form of Eq. (2.29). Amplification can protect measurement outcomes from noise through redundancy.[3]

[3]The above model of amplification is unitary. Yet it contains seeds of irreversibility. The reversibility of a c-shift is evident: as the interaction continues, the two systems will eventually disentangle. For instance, it takes $t_e=2\pi\hbar/(gN)$ [see Eq. (2.23b) with $G=1$] to entangle $\mathcal{S}[\dim(\mathcal{H}_{\mathcal{S}})=n]$ with an $\mathcal{A}$ with $\dim(\mathcal{H}_{\mathcal{A}})=N\gg n$ pointer states. However, as the interaction continues, $\mathcal{A}$ and $\mathcal{S}$ disentangle. For a c-shift, this recurrence time scale is $t_{Rec}=Nt_e=2\pi\hbar/g$. It corresponds to a gain $G=N$. Thus, for an instant of less than t_e at $t=t_{Rec}$, the apparatus disentangles from the system, as $\{k+N*j\}_N=k$. Reversibility results in recurrences of the initial state, but for $N\gg1$, they are rare.

For less regular interactions (e.g, involving the environment) the recurrence time is much longer. In that case, t_{Rec} is, in effect, a Poincaré time: $t_{Rec}\sim t_{Poincaré}\approx N!t_e$. In any case $t_{Rec}\gg t_e$ for large N. Undoing entanglement in this manner would be exceedingly difficult because one would need to know precisely when to look and because one would need to isolate the apparatus or the immediate environment from other degrees of freedom—their environments.

The price of letting the entanglement undo itself by waiting for an appropriate time interval is at the very least given by the cost of storing the information based on how long it is necessary to wait. In the special c-shift case this is proportional to

B. Information transfer in measurements

1. Reduced density matrices and correlations

Information transfer is the objective of the measurement process. Yet quantum measurements have only rarely been analyzed from that point of view. As a result of the interaction of the system $\mathcal{S}$ with the apparatus $\mathcal{A}$, their joint state is still pure $|\Psi_t\rangle$, Eq. (2.1), but each of the subsystems is in a mixture:

$$\rho_{\mathcal{S}}=\mathrm{Tr}_{\mathcal{A}}|\Psi_t\rangle\langle\Psi_t|=\sum_{i=0}^{N-1}|a_i|^2|s_i\rangle\langle s_i|, \tag{2.36a}$$

$$\rho_{\mathcal{A}}=\mathrm{Tr}_{\mathcal{S}}|\Psi_t\rangle\langle\Psi_t|=\sum_{i=0}^{N-1}|a_i|^2|A_i\rangle\langle A_i|. \tag{2.36b}$$

The partial trace leads to reduced density matrices, here $\rho_{\mathcal{S}}$ and $\rho_{\mathcal{A}}$, which are important for what follows. They describe subsystems to the observer who, before the premeasurement, knew pure states of the system and of the apparatus, but who has access to only one of them afterwards.

The reduced density matrix is a technical tool of paramount importance. It was introduced by Landau (1927) as the only density matrix that gives rise to the correct measurement statistics given the usual formalism that includes Born's rule for calculating probabilities (see, for example, p. 107 of Nielsen and Chuang, 2000, for an insightful discussion). This remark will come to haunt us later when in Sec. VI we consider the relation between decoherence and probabilities. In order to derive Born's rule it will be important not to assume it in some guise.

Following premeasurement, the information about the subsystems available to the observer locally decreases. This is quantified by the increase of the entropies:

$$H_{\mathcal{S}}=-\mathrm{Tr}\rho_{\mathcal{S}}\ln\rho_{\mathcal{S}}=-\sum_{i=0}^{N-1}|a_i|^2\ln|a_i|^2$$
$$=-\mathrm{Tr}\rho_{\mathcal{A}}\ln\rho_{\mathcal{A}}=H_{\mathcal{A}}. \tag{2.37}$$

Since the evolution of the whole $\mathcal{SA}$ is unitary, the increase of entropies in the subsystems is compensated for by the buildup of correlations, and the resulting increase in mutual information:

$$\mathcal{I}(\mathcal{S}:\mathcal{A})=H_{\mathcal{S}}+H_{\mathcal{A}}-H_{\mathcal{SA}}=-2\sum_{i=0}^{N-1}|a_i|^2\ln|a_i|^2. \tag{2.38}$$

This has been used in quantum theory as a measure of entanglement (Zurek, 1983; Barnett and Phoenix, 1989).

2. Action per bit

An often raised question concerns the price of information in units of some other "physical currency" (Brillouin, 1962, 1964; Landauer, 1991). Here we shall establish that the least action necessary to transfer one bit is of the order of a fraction of $\hbar$ for quantum systems with two-dimensional Hilbert spaces. Information transfer can be made cheaper on the "wholesale" level, when the systems involved have large Hilbert spaces.

Consider Eq. (2.1). It evolves the initial product state of the two subsystems into a superposition of product states, $(\Sigma_j\alpha_j|s_j\rangle)|A_0\rangle\rightarrow\Sigma_j\alpha_j|s_j\rangle|A_j\rangle$. The expectation value of the action involved is no less than

$$I=\sum_{j=0}^{N-1}|\alpha_j|^2\arccos|\langle A_0|A_j\rangle|. \tag{2.39}$$

When $\{|A_j\rangle\}$ are mutually orthogonal, the action is

$$I=\pi/2 \tag{2.40}$$

in Planck ($\hbar$) units. This estimate can be lowered by using as the initial $|A_0\rangle$ a superposition of the outcomes $|A_j\rangle$. In general, an interaction of the form

$$H_{\mathcal{SA}}=ig\sum_{k=0}^{N-1}|s_k\rangle\langle s_k|\sum_{l=0}^{N-1}(|A_k\rangle\langle A_l|-\mathrm{H.c.}), \tag{2.41}$$

where H.c. is the Hermitian conjugate, saturates the lower bound given by

$$I=\arcsin\sqrt{1-1/N}. \tag{2.42}$$

For a two-dimensional Hilbert space the average action can be thus brought down to $\pi\hbar/4$ (Zurek, 1981, 1983).

As the size of the Hilbert space increases, the action involved approaches the asymptotic estimate of Eq. (2.40). The entropy of entanglement can be as large as $\log N$ where N is the dimension of the Hilbert space of

---

$\log N$ memory bits. In situations when eigenvalues of the interaction Hamiltonian are not commensurate, it will be more like $\sim\log N!\approx N\log N$, since the entanglement will get undone only after a Poincaré time. Both classical and quantum cases can be analyzed using algorithmic information. For related discussions see Zurek (1989, 1998b), Caves (1994), and Schack and Caves (1996).

Amplified correlations are hard to contain. The return to purity after t_{Rec} in the manner described above can be hoped for only when the apparatus or the immediate environment $\mathcal{E}$ (i.e., the environment directly interacting with the system) cannot "pass on" the information to their more remote environments $\mathcal{E}'$. The degree of isolation required puts a stringent limit on the coupling $g_{\mathcal{E}\mathcal{E}'}$ between the two environments. Return to purity can be accomplished in this manner only if $t_{Rec}<t_{e'}=2\pi\hbar/(N'g_{\mathcal{E}\mathcal{E}'})$, where N' is the dimension of the Hilbert space of the environment $\mathcal{E}'$. Hence the two estimates of t_{Rec} translate into $g_{\mathcal{E}\mathcal{E}'}<g/N'$ for the regular spectrum and the much tighter $g_{\mathcal{E}\mathcal{E}'}<g/N!N'$ for the random case more relevant for decoherence.

In short, once information "leaks" into the correlations between the system and the apparatus or the environment, keeping it from spreading further ranges between very hard and next to impossible. With the exception of very special cases (small N, regular spectrum), the strategy of enlarging the system, so that it includes the environment—occasionally mentioned as an argument against decoherence—is doomed to fail, unless the universe as a whole is included. This is a questionable setting (since the observers are inside this "isolated" system) and in any case makes the relevant Poincaré time absurdly long.

the smaller of the two systems. Thus the least action per bit of information decreases with the increase of N:

$$\iota=\frac{I}{\log_2 N}\approx\frac{\pi}{2\log_2 N}. \tag{2.43}$$

This may be one reason why information appears "free" in the macroscopic domain, but expensive (close to $\hbar$/bit) in the quantum case of small Hilbert spaces.

C. "Collapse" analog in a classical measurement

Definite outcomes that we perceive appear to be at odds with the principle of superposition. They can nevertheless also occur in quantum physics when the initial state of the measured system is—already before the measurement—in one of the eigenstates of the measured observable. Then Eq. (2.1) will deterministically rotate the pointer of the apparatus to the appropriate record state. The result of such a measurement can be predicted by an *insider*—an observer aware of the initial state of the system. This *a priori* knowledge can be represented by the preexisting record $|A_i\rangle$, which is only corroborated by an additional measurement:

$$|A_i\rangle|A_0\rangle|\sigma_i\rangle\rightarrow|A_i\rangle|A_i\rangle|\sigma_i\rangle. \tag{2.44a}$$

In classical physics complete information about the initial state of an isolated system always allows for an exact prediction of its future state. A well-informed observer will even be able to predict the future of the classical universe as a whole ("Laplace's demon"). Any element of surprise (any use of probabilities) must therefore be blamed on partial ignorance. Thus, when the information available initially does not include the exact initial state of the system, the observer can use an ensemble described by $\rho_{\mathcal{S}}$—by a list of possible initial states $\{|\sigma_i\rangle\}$ and their probabilities p_i. This is the *ignorance interpretation* of probabilities. We shall see in Sec. VI that—using envariance—one can justify ignorance about a part of the system by relying on perfect knowledge of the whole.

Through measurement the observer finds out which of the potential outcomes consistent with his prior (incomplete) information actually happens. This act of information acquisition changes the physical state of the observer, the state of his memory. The initial memory state containing a description $A_{\rho_{\mathcal{S}}}$ of an ensemble and a "blank" A_0, $|A_{\rho_{\mathcal{S}}}\rangle\langle A_{\rho_{\mathcal{S}}}||A_0\rangle\langle A_0|$, is transformed into a record of a specific outcome: $|A_{\rho_{\mathcal{S}}}\rangle\langle A_{\rho_{\mathcal{S}}}||A_i\rangle\langle A_i|$. In quantum notation this process will be described by such a *discoverer* as a random "collapse:"

$$|A_{\rho_{\mathcal{S}}}\rangle\langle A_{\rho_{\mathcal{S}}}||A_0\rangle\langle A_0|\sum_i p_i|\sigma_i\rangle\langle\sigma_i| \rightarrow|A_{\rho_{\mathcal{S}}}\rangle\langle A_{\rho_{\mathcal{S}}}||A_i\rangle\langle A_i||\sigma_i\rangle\langle\sigma_i|. \tag{2.44b}$$

This is only the description of what happens as reported by the discoverer. Deterministic representation of this very same process by Eq. (2.44a) is still possible. In other words, in classical physics the discoverer can always be convinced that the system was in a state $|\sigma_i\rangle$ already before he has measured it in accord with Eq. (2.44b).

This sequence of events as seen by the discoverer looks like a collapse (see Zurek, 1998a, 1998b). For instance, an insider who knew the state of the system before the discoverer carried out his measurement need not notice any change of that state when he makes further confirmatory measurements. This property is the cornerstone of the "reality" of classical states—they need not ever change as a consequence of measurements. We emphasize, however, that while the state of the system may remain unchanged, the state of the observer must change to reflect the acquired information.

Last but not least, an *outsider*—someone who knows about the measurement, but (in contrast to the insider) not about the initial state of the system nor (in contrast to both the insider and the discoverer) about the outcome of the measurement—will describe the same process still differently:

$$|A_{\rho_{\mathcal{S}}}\rangle\langle A_{\rho_{\mathcal{S}}}||A_0\rangle\langle A_0|\sum_i p_i|\sigma_i\rangle\langle\sigma_i| \rightarrow|A_{\rho_{\mathcal{S}}}\rangle\langle A_{\rho_{\mathcal{S}}}|\left(\sum_i p_i|A_i\rangle\langle A_i||\sigma_i\rangle\langle\sigma_i|\right). \tag{2.44c}$$

This view of the outsider, Eq. (2.44c), combines a one-to-one classical correlation of the states of the system and the records with the indefiniteness of the outcome.

We have just seen three distinct quantum-looking descriptions of the very same classical process (see Zurek, 1989 and Caves, 1994 for previous studies of the insider-outsider theme). They differ only in the information available *ab initio* to the observer. The information in the possession of the observer prior to the measurement determines in turn whether—to the observer—the evolution appears to be (a) a confirmation of the preexisting data, Eq. (2.44a); (b) a collapse associated with the information gain, Eq. (2.44b), and with the entropy decrease translated into algorithmic randomness of the acquired data (Zurek, 1989, 1998b); or (c) an entropy-preserving establishment of a correlation, Eq. (2.44c). All three descriptions are classically compatible, and can be implemented by the same (deterministic and reversible) dynamics.

In classical physics the insider view always exists, in principle. In quantum physics it does not. Every observer in a classical universe could, in principle, aspire to be an ultimate insider. The fundamental contradiction between every observer's knowing precisely the state of the rest of the Universe (including the other observers) can be swept under the rug (if not really resolved) in a universe where the states are infinitely precisely determined and the observer's records (as a consequence of the $\hbar\rightarrow 0$ limit) may have an infinite capacity for information storage. However, given a set value of $\hbar$, the information storage resources of any finite physical system are finite. Hence, in quantum physics, observers remain largely ignorant of the detailed state of the universe, since there can be no information without representation (Zurek, 1994).

Classical collapse is described by Eq. (2.44b). The observer discovers the state of the system. From then on, the state of the system will remain correlated with his record, so that all future outcomes can be predicted, in effect by iterating Eq. (2.44a). This disappearance of all the potential alternatives save for one that becomes a "reality" is the essence of the collapse. There need not be anything quantum about it.

Einselection in the observer's memory provides many of the ingredients of classical collapse in the quantum context. In the presence of einselection, a one-to-one correspondence between the state of the observer and his knowledge about the rest of the universe can be firmly established, and (at least, in principle) operationally verified. One could measure bits in the observer's memory or even the "imprint" of their state on the environment and determine what he knows without altering his records—without altering his state. After all, one can do so with a classical computer. The existential interpretation recognizes that the information possessed by the observer is reflected in his einselected state, explaining his perception of a single branch—"his" classical universe.

III. CHAOS AND LOSS OF CORRESPONDENCE

The study of the relationship between the quantum and the classical has been, for a long time, focused almost entirely on measurements. However, the problem of measurement is difficult to discuss without observers. And once the observer enters, it is often hard to avoid its ill-understood anthropic attributes such as consciousness, awareness, and the ability to perceive.

We shall sidestep these "metaphysical" problems and focus on the information-processing underpinnings of observership. It is nevertheless fortunate that there is another problem with the quantum-classical correspondence that leads to interesting questions not motivated by measurements. As was anticipated by Einstein (1917) before the advent of modern quantum theory, chaotic motion presents such a challenge. The full implications of classical dynamical chaos were understood much later. The concern about the quantum-classical correspondence in this modern context dates to Berman and Zaslavsky (1978) and Berry and Balazs (1979) (see Haake, 1991 and Casati and Chirikov, 1995a, for references). It has even led some to question the validity of quantum theory (Ford and Mantica, 1992).

A. Loss of quantum-classical correspondence

The interplay between quantum interference and chaotic exponential instability leads to the rapid loss of quantum-classical correspondence. Chaos in dynamics is characterized by the exponential divergence of the classical trajectories. As a consequence, a small patch representing the probability density in phase space is exponentially stretching in unstable directions and exponentially compressing in stable directions. The rates of stretching and compression are given by positive and negative Lyapunov exponents Λ_i. Hamiltonian evolution demands that the sum of all the Lyapunov exponents be zero. In fact, they appear in $\pm\Lambda_i$ pairs.

Loss of correspondence in chaotic systems is a consequence of the exponential stretching of the effective support of the probability distribution in the unstable direction (say, x) and its exponential narrowing in the complementary direction (Zurek and Paz, 1994; Zurek, 1998b). As a consequence, the classical probability distribution will develop structures on the scale

$$\Delta p \sim \Delta p_0 \exp(-\Lambda t). \tag{3.1}$$

Above, Δp_0 is the measure of the initial momentum spread and Λ is the net rate of contraction in the direction of momentum given by the Lyapunov exponents (but see Boccaletti, Farini, and Arecchi, 1997). In a real chaotic system, stretching and narrowing of the probability distribution in both x and p occur simultaneously, as the initial patch is rotated and folded. Eventually, the envelope of its effective support will swell to fill in the available phase space, resulting in a wave packet that is coherently spread over a spatial region of no less than

$$\Delta x \sim (\hbar/\Delta p_0)\exp(\Lambda t). \tag{3.2}$$

until it becomes confined by the potential, while the small-scale structure will continue to descend to ever smaller scales (Fig. 1). Breakdown of the quantum-classical correspondence can be understood in two complementary ways, either as a consequence of small Δp (see the discussion of the Moyal bracket below) or as a result of large Δx.

Coherent exponential spreading of the wave packet—large Δx—must cause problems with correspondence. This is inevitable, since classical evolution appeals to the idealization of a point in phase space acted upon by a force given by the gradient $\partial_x V$ of the potential $V(x)$ evaluated at that point. But the quantum wave function can be coherent over a region larger than the nonlinearity scale χ over which the gradient of the potential changes significantly. χ can usually be estimated by

$$\chi \simeq \sqrt{\partial_x V/\partial_{xxx} V}, \tag{3.3}$$

and is typically of the order of the size L of the system:

$$L \sim \chi. \tag{3.4}$$

An initially localized state evolving in accord with Eqs. (3.1) and (3.2) will spread over such scales after

$$t_\hbar \simeq \Lambda^{-1} \ln \frac{\Delta p_0 \chi}{\hbar}. \tag{3.5}$$

It is then impossible to tell what force is acting upon the system, since it is not located in any specific x. This estimate of what can be thought of as Ehrenfest time, the time over which a quantum system that has started in a localized state will continue to be sufficiently localized for the quantum corrections to the equations of motion

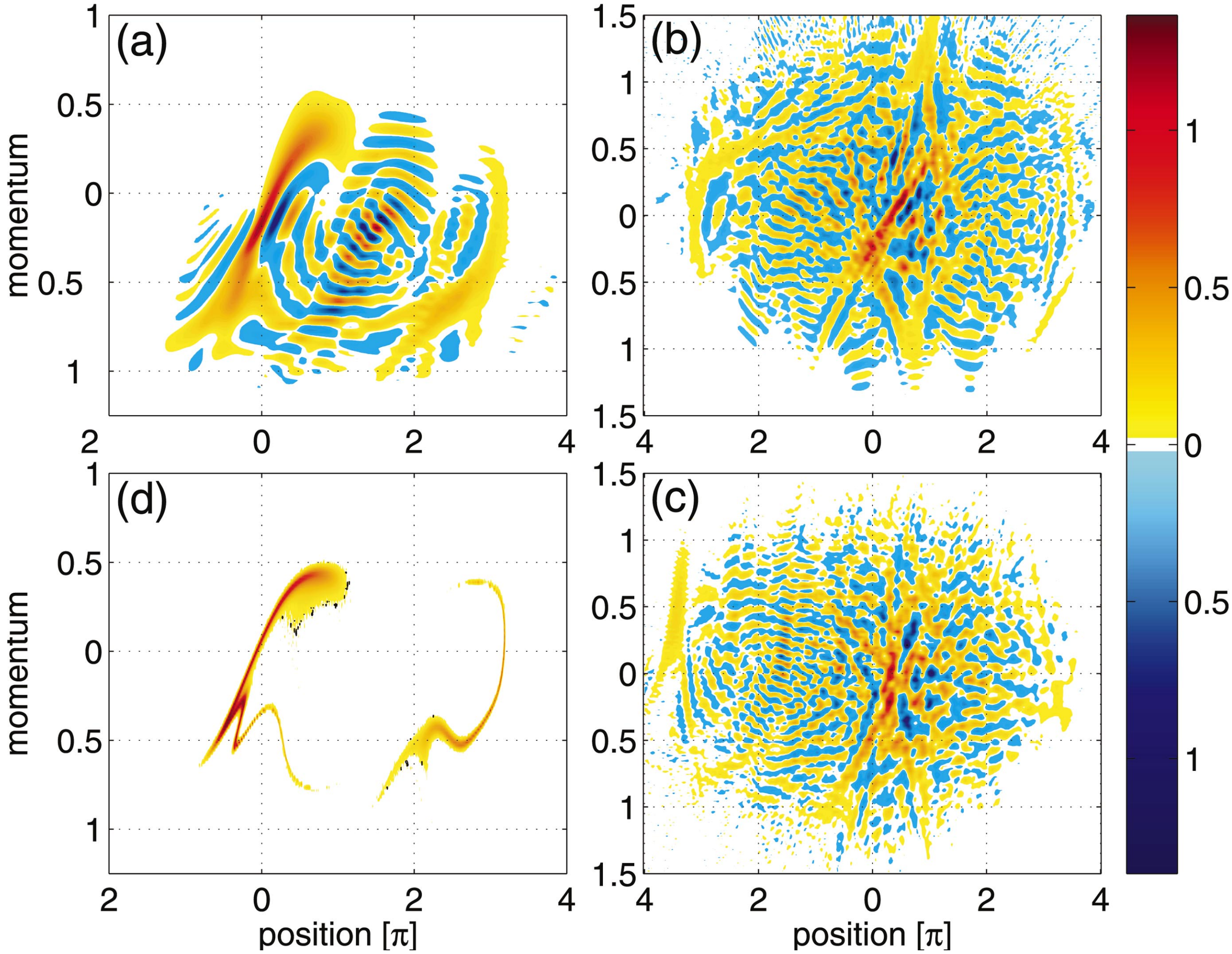


FIG. 1. Chaotic evolution generated from the same initial Gaussian by the Hamiltonian $H = p^2/2m - \kappa \cos(x - l \sin t) + ax^2/2$. (a)–(c) Snapshots of the quantum ($\hbar = 0.16$) Wigner function; (d) classical probability distribution in phase space. For $m = 1$, $\kappa = 0.36$, $l = 3$, and $a = 0 - 0.01$ the Hamiltonian exhibits chaos with the Lyapunov exponent $\Lambda = 0.2$ (Karkuszewski, Zakrzewski, and Zurek, 2002). Quantum (a) and classical (d) are obtained at the same instant, $t = 20$. They exhibit some similarities [i.e., the shape of the regions of significant probability density, "ridges" in the topographical maps of (a) and (d)], but the difference—the presence of the interference patterns with $W(x,p)$ assuming negative values (marked with blue)—is striking. Saturation of the size of the smallest patches is anticipated already at this early time, and indeed the ridges of the classical probability density are narrower than in the corresponding quantum features. Saturation is even more visible in (b) taken at $t = 60$ and (c), $t = 100$ [note change of scale from (a) and (d)]. Sharpness of the classical features makes simulations going beyond $t = 20$ unreliable, but quantum simulations can be effectively carried out much further, since the necessary resolution can be anticipated in advance from Eqs. (3.14)–(3.16) (Color).

obeyed by its expectation values to be negligible (Gottfried, 1966), is valid for chaotic systems. Logarithmic dependence is the result of inverting the exponential sensitivity. In the absence of exponential instability ($\Lambda = 0$), divergence of trajectories is typically polynomial and leads to a power-law dependence, $t_\hbar \sim (I/\hbar)^\alpha$, where I is the classical action. Thus macroscopic (large-I) integrable systems can follow classical dynamics for a very long time, providing they were initiated in a localized state. For chaotic systems $t_\hbar$ also becomes infinite in the limit $\hbar \to 0$, but that happens only logarithmically slowly. As we shall see below, in the context of quantum-classical correspondence this is too slow for comfort.

Another way of describing the root cause of a breakdown of correspondence is to note that after a time scale of the order of $t_\hbar$, the quantum wave function of the system would have spread over all of the available space and would be forced to fold onto itself. Fragments of the wave packet arrive at the same location (although with different momenta, and having followed different paths). The ensuing evolution depends critically on whether they have retained phase coherence. When coherence persists, a complicated interference event decides the subsequent evolution. And, as can be anticipated from the double-slit experiment, there is a big difference between coherent and incoherent folding in the configuration space. This translates into a loss of correspondence, which sets in surprisingly quickly, at $t_\hbar$.

To find out how quickly, we estimate $t_\hbar$ for an obviously macroscopic object, Hyperion, a chaotically tumbling moon of Saturn (Wisdom, 1985). Hyperion has the prolate shape of a potato and moves on an eccentric orbit with a period $t_O = 21$ days. Interaction between its gravitational quadrupole and the tidal field of Saturn leads to chaotic tumbling with Lyapunov time $\Lambda^{-1} \simeq 42$ days.

To estimate the time over which the orientation of Hyperion becomes delocalized, we use a formula (Berman and Zaslavsky, 1978; Berry and Balazs, 1979):

$$t_r = \Lambda^{-1} \ln \frac{LP}{\hbar} = \Lambda^{-1} \ln \frac{I}{\hbar}. \tag{3.6}$$

Above L and P give the range of values of the coordinates and momenta in phase space of the system. Since $L \simeq \chi$ and $P > \Delta p_0$, it follows that $t_r \geqslant t_\hbar$. On the other hand, $LP \simeq I$, the classical action of the system.

The advantage of Eq. (3.6) is its insensitivity to initial conditions and the ease with which the estimate can be obtained. For Hyperion, a generous overestimate of the classical action I can be obtained from its binding energy E_B and its orbital time t_O:

$$I/\hbar \simeq E_B t_O / \hbar \simeq 10^{77}. \tag{3.7}$$

The above estimate (Zurek, 1998b) is "astronomically" large. However, in the calculation of the loss of correspondence, Eq. (3.6), only the logarithm of I enters. Thus

$$t_r^{Hyper} \simeq 42\ [\text{days}] \ln 10^{77} \simeq 20\ \ [\text{yr}]. \tag{3.8}$$

After approximately 20 yr Hyperion would be in a coherent superposition of orientations that differ by 2π.

We conclude that after a relatively short time an obviously macroscopic chaotic system becomes forced into a flagrantly nonlocal "Schrödinger-cat" state. In the original discussion (Schrödinger, 1935a, 1935b) an intermediate step in which the decay products of the nucleus were measured to determine the fate of the cat was essential. Thus it was possible to maintain that the preposterous superposition of the dead and live cat could be avoided, providing that quantum measurement (with the collapse it presumably induces) was properly understood.

This cannot be the resolution for chaotic quantum systems. They can evolve, as the example of Hyperion demonstrates, into states that are nonlocal and, therefore, extravagantly quantum, simply as a result of exponentially unstable dynamics. Moreover, this happens surprisingly quickly, even for very macroscopic examples. Hyperion is not the only chaotic system. There are asteroids that have chaotically unstable orbits (e.g., Chiron), and even indications that the solar system as a whole is chaotic (Laskar, 1989; Sussman and Wisdom, 1992). In all of these cases straightforward estimates of $t_\hbar$ yield answers much smaller than the age of the solar system. Thus, if unitary evolution of closed subsystems was a complete description of planetary dynamics, planets would be delocalized along their orbits.

B. Moyal bracket and Liouville flow

Heuristic arguments about the breakdown of quantum-classical correspondence can be made more rigorous with the help of the Wigner function. We start with the von Neumann equation

$$i\hbar \dot{\rho} = [H, \rho]. \tag{3.9}$$

This can be transformed into the equation for the Wigner function W, which is defined in phase space as

$$W(x,p) = \frac{1}{2\pi\hbar} \int \exp\left(\frac{ipy}{\hbar}\right) \rho\left(x - \frac{y}{2}, x + \frac{y}{2}\right) dy. \tag{3.10}$$

The result is

$$\dot{W} = \{H, W\}_{MB}. \tag{3.11}$$

Here $\{\ldots,\ldots\}_{MB}$ stands for the Moyal bracket, the Wigner transform of the von Neumann bracket (Moyal, 1949).

The Moyal bracket can be expressed in terms of the Poisson bracket $\{\ldots,\ldots\}$, which generates Liouville flow in classical phase space, by the formula

$$i\hbar \{\ldots,\ldots\}_{MB} = \sin(i\hbar\{\ldots,\ldots\}). \tag{3.12}$$

When the potential $V(x)$ is analytic, the Moyal bracket can be expanded (Hillery *et al.*, 1984) in powers of $\hbar$:

$$\dot{W} = \{H, W\} + \sum_{n \geqslant 1} \frac{\hbar^{2n}(-)^n}{2^{2n}(2n+1)!} \partial_x^{2n+1} V \partial_p^{2n+1} W. \tag{3.13}$$

The first term is just the Poisson bracket. Alone, it would generate classical motion in phase space. However, when the evolution is chaotic, quantum corrections (proportional to the odd-order momentum derivatives of the Wigner function) will eventually dominate the right-hand side of Eq. (3.10). This is because the exponential squeezing of the initially regular patch in phase space (which begins its evolution in the classical regime, where the Poisson bracket dominates) leads to an exponential explosion of the momentum derivatives. Consequently, after a time logarithmic in $\hbar$ [Eqs. (3.5) and (3.6)], the Poisson bracket will cease to be a good estimate of the right-hand side of Eq. (3.13).

The physical reason for the ensuing breakdown of the quantum-classical correspondence has already been explained: exponential instability of the chaotic evolution delocalizes the wave packet. As a result, the force acting on the system is no longer given by the gradient of the potential evaluated at the location of the system. It is not even possible to say where the system is, since it is in a superposition of many distinct locations. Consequently, the phase-space distribution and even the expectation values of the observables of the system differ noticeably when evaluated classically and quantum mechanically (Haake, Kuś, and Sharf, 1987; Habib, Shi-

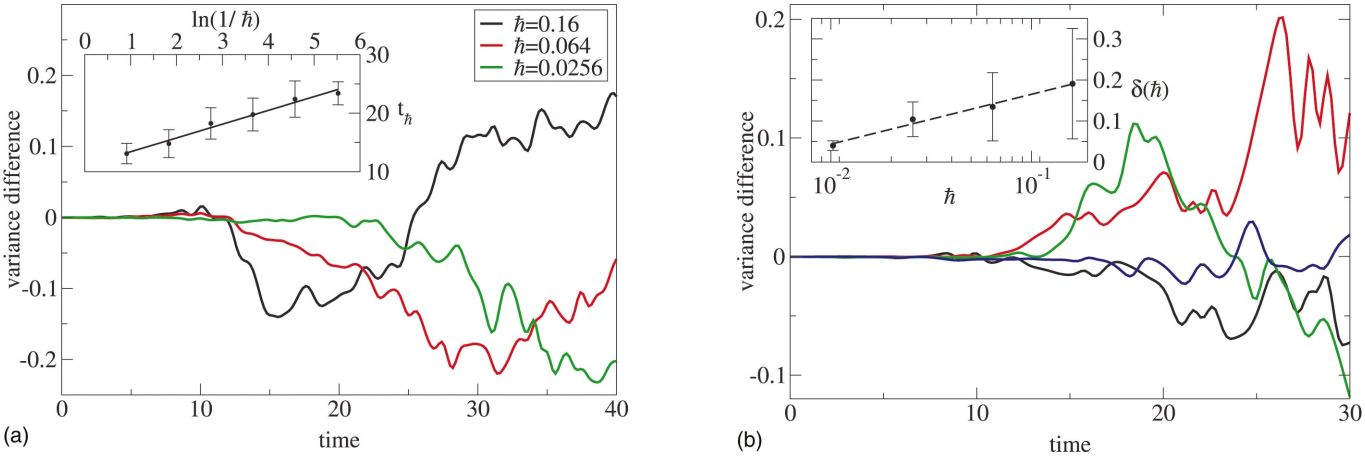


FIG. 2. Difference between the classical and quantum averages of the dispersion of momentum $\Delta^2 = \langle p^2 \rangle - \langle p \rangle^2$ plotted for (a) same initial condition, but three different values of $\hbar$ in the model defined in Fig. 1, with the parameter $a=0$. The instant when the difference between the classical and quantum averages becomes significant varies with $\hbar$ in a manner anticipated from Eqs. (3.5) and (3.6) for the Ehrenfest time, as can be seen in the inset; (b) same value of $\hbar$, but for four different initial conditions. Inset appears to indicate that the typical variance difference δ varies only logarithmically with decreasing $\hbar$, although the large error bars (tied to the large systematic changes of behavior for different initial conditions) preclude one from arriving at a firmer conclusion. (See Karkuszewski, Zakrzewski, and Zurek, 2002, for further details and discussion.) (Color).

zume, and Zurek, 1998; Karkuszewski, Zakrzewski, and Zurek, 2002). Moreover, this will happen after an uncomfortably short time $t_\hbar$.

C. Symptoms of correspondence loss

The wave packet becomes rapidly delocalized in a chaotic system, and the correspondence between classical and quantum is quickly lost. Flagrantly nonlocal Schrödinger-cat states appear no later than $t_\hbar$, and this is the overarching interpretational as well as physical problem. In the familiar real world we never seem to encounter such smearing of the wave function even in the examples of chaotic dynamics where it is predicted by quantum theory.

1. Expectation values

Measurements usually average out fine phase-space interference structures, which may be a striking, but experimentally inaccessible symptom of the breakdown of correspondence. Thus one might hope that when interference patterns in the Wigner function are ignored by looking at the coarse-grained distribution, the quantum results should be in accord with the classical. This would not exorcise the "chaotic cat" problem. Moreover, the breakdown of correspondence can also be seen in the expectation values of quantities that are smooth in phase space.

Trajectories diverge exponentially in a chaotic system. A comparison between expectation values for a single trajectory and for a delocalized quantum state (which is how the Ehrenfest theorem mentioned above is usually stated) would clearly lead to a rapid loss of correspondence. However, one may rightly object to the use of a single trajectory and argue that both the quantum and the classical state should be prepared and accessed only through measurements that are subject to Heisenberg indeterminacy. Still, it should be fair to compare averages over an evolving Wigner function with an initially identical classical probability distribution (Haake, Kuś, and Sharf, 1987; Ballentine, Yang, and Zibin, 1994; Fox and Elston, 1994a, 1994b; Miller, Sarkar, and Zarum, 1998). These are shown in Fig. 2 for an example of a driven chaotic system. Clearly, there is reason for concern. Figure 2 (corroborated by other studies—see Karkuszewski, Zakrzewski, and Zurek, 2002, for references) demonstrates that not just the phase-space portrait but also the averages diverge at a time $\sim t_\hbar$.

In integrable systems, the rapid loss of correspondence between the quantum and the classical expectation values may still occur, but only for very special initial conditions, due to the local instability in phase space. Indeed, a double-slit experiment is an example of a regular system in which a local instability (splitting of the paths) leads to correspondence loss, but only for judiciously selected initial conditions. Thus one may dismiss such a breakdown as a consequence of a rare pathological starting point and argue that the conditions that lead to discrepancies between classical and quantum behavior exist, but are of measure zero in the classical limit.

In the chaotic case the loss of correspondence is typical. As shown in Fig. 2, it happens after a disturbingly short $t_\hbar$ for generic initial conditions. The time at which the quantum and classical expectation values diverge in the example studied here is consistent with the estimates of $t_\hbar$, Eq. (3.5), but exhibits a significant scatter. This is not too surprising—exponents characterizing local instability vary noticeably with location in phase space. Hence stretching and contraction in phase space will occur at a rate that depends on the specific trajectory. The dependence of a typical magnitude on $\hbar$ is still not clear. Emerson and Ballentine (2001a, 2001b) studied coupled spins and argued that it is of the order of $\hbar$, but Fig. 2 suggests it decreases more slowly than that, and that it may be only logarithmic in $\hbar$ (Karkuszewski *et al.*, 2002).

2. Structure saturation

Evolution of the Wigner function leads to rapid buildup of interference fringes. These fringes become progressively smaller, until saturation, when the wave packet is spread over the available phase space. At that time their scales in momentum and position are typically given by

$$dp = \hbar / L, \tag{3.14}$$

$$dx = \hbar / P, \tag{3.15}$$

where $L(P)$ defines the range of positions (momenta) of the effective support of W in phase space.

Hence the smallest structures in the Wigner function occur (Zurek, 2001) on scales corresponding to an action

$$a = dx\, dp = \hbar \times \hbar / LP = \hbar^2 / I, \tag{3.16}$$

where $I \simeq LP$ is the classical action of the system. Action $a \ll \hbar$ for macroscopic I.

Sub-Planck structure is a kinematic property of quantum states. It helps determine their sensitivity to perturbations and has applications outside quantum chaos or decoherence. For instance, a Schrödinger-cat state can be used as a weak force detector (Zurek, 2001), and its sensitivity is determined by Eqs. (3.14)–(3.16).

Structure saturation on scale a is an important distinction between the quantum and the classical. In chaotic systems, the smallest structures in classical probability density exponentially shrink with time, in accord with Eq. (3.1) (see Fig. 1). Equation (3.16) has implications for decoherence, since a controls the sensitivity of systems as well as of environments (Zurek, 2001; Karkuszewski, Jarzynski, and Zurek, 2002). As a result of the smallness of a, Eq. (3.16), and as anticipated by Peres (1993), quantum systems are more sensitive to perturbations when their classical counterparts are chaotic (see also Jalabert and Pastawski, 2001). But in contrast to classical chaotic systems they are not exponentially sensitive to infinitesimally small perturbations. Rather, the smallest perturbations that can be effective are set by Eq. (3.16).

The emergence of Schrödinger-cat states through dynamics is a challenge to quantum-classical correspondence. It is not yet clear to what extent one should be concerned about the discrepancies between quantum and classical averages. The size of this discrepancy may or may not be negligible. But in the original Schrödinger-cat problem, quantum and classical expectation values (for the survival of the cat) were in accord. In both cases it is ultimately the state of the cat that is most worrisome.

Note that we have not dealt with dynamical localization (Casati and Chirikov, 1995a). This is because it appears after too long a time ($\sim \hbar^{-1}$) to be a primary concern in the macroscopic limit and is quite sensitive to small perturbations of the potential (Karkuszewski, Zakrzewski, and Zurek, 2002).

IV. ENVIRONMENT-INDUCED SUPERSELECTION

The principle of superposition applies only when the quantum system is closed. When the system is open, interaction with the environment results in an incessant monitoring of some of its observables. As a result, pure states turn into mixtures that become rapidly diagonal in einselected states. These pointer states are chosen with the help of the interaction Hamiltonian and are independent of the initial state of the system. Their predictability is key to the effective classicality (Zurek, 1993a; Zurek, Habib, and Paz, 1993).

Environments can be external (such as particles of the air or photons that scatter off, say, the apparatus pointer) or internal (collections of phonons or other internal excitations). Often, environmental degrees of freedom emerge from a split of the original set of degrees of freedom into a "system of interest," which may be a collective observable (e.g., an order parameter in a phase transition) and a "microscopic remainder."

The set of einselected states is called *the pointer basis* (Zurek, 1981) in recognition of its role in measurements. The criterion for the einselection of states goes well beyond the often repeated characterizations based on the instantaneous eigenstates of the density matrix. What is of the essence is the ability of the einselected states to survive monitoring by the environment. This heuristic criterion can be made rigorous by quantifying the predictability of the evolution of the candidate states, or of the associated observables. Einselected states provide optimal initial conditions. They can be employed for the purpose of prediction better than other Hilbert-space alternatives—they retain correlations in spite of their immersion in the environment.

Three quantum systems—the measured system $\mathcal{S}$, the apparatus $\mathcal{A}$, and the environment $\mathcal{E}$—and the correlations between them are the subject of our study. In premeasurements $\mathcal{S}$ and $\mathcal{A}$ interact. Their resulting entanglement transforms into an effectively classical correlation as a result of the interaction between $\mathcal{A}$ and $\mathcal{E}$.

This $\mathcal{SAE}$ triangle helps us to analyze decoherence and study its consequences. By keeping all three corners of this triangle in mind, one can avoid confusion and maintain focus on the correlations between, for example, the memory of the observer and the state of the measured system. The evolution from a quantum entanglement to a classical correlation we are about to discuss may be the easiest relevant aspect of the quantum-to-classical transition to define operationally. In the language of the last part of Sec. II, we are about to justify the "outsider" point of view, Eq. (2.44c), before considering the measurement from the vantage point of the "discoverer," Eq. (2.44b), and before tackling the issue of collapse. In spite of this focus on correlations, we shall often suppress one of the corners of the $\mathcal{SAE}$ triangle to simplify notation. All three parts will, however, play a role in formulating questions and in motivating the criteria for classicality.

A. Models of einselection

The simplest case of a single act of decoherence involves just three one-bit systems (Zurek, 1981, 1983). They are denoted by $\mathcal{S}$, $\mathcal{A}$, and $\mathcal{E}$ in an obvious reference to their roles. The measurement starts with the interaction of the measured system with the apparatus

$$|\uparrow\rangle|A_0\rangle\rightarrow|\uparrow\rangle|A_1\rangle, \tag{4.1a}$$

$$|\downarrow\rangle|A_0\rangle\rightarrow|\downarrow\rangle|A_0\rangle, \tag{4.1b}$$

where $\langle A_0|A_1\rangle=0$. For a general state,

$$(\alpha|\uparrow\rangle+\beta|\downarrow\rangle)|A_0\rangle\rightarrow\alpha|\uparrow\rangle|A_1\rangle+\beta|\downarrow\rangle|A_0\rangle=|\Phi\rangle. \tag{4.2}$$

These formulas represents a c-NOT implementation of the premeasurement discussed in Sec. II.

The basis ambiguity [that is, the ability to rewrite $|\Phi\rangle$, Eq. (4.2), in any basis of, say, the system, with the principle of superposition guaranteeing the existence of the corresponding pure states of the apparatus] disappears when an additional system, $\mathcal{E}$, performs a premeasurement on $\mathcal{A}$:

$$(\alpha|\uparrow\rangle|A_1\rangle+\beta|\downarrow\rangle|A_0\rangle)|\varepsilon_0\rangle$$
$$\rightarrow\alpha|\uparrow\rangle|A_1\rangle|\varepsilon_1\rangle+\beta|\downarrow\rangle|A_0\rangle|\varepsilon_0\rangle=|\Psi\rangle. \tag{4.3}$$

A collection of three correlated quantum systems is no longer subject to the basis ambiguity we have pointed out in connection with the EPR-like state $|\Phi\rangle$, Eq. (4.2). This is especially true when states of the environment are correlated with the simple products of the states of the apparatus-system combination (Zurek, 1981; Elby and Bub, 1994). In Eq. (4.3) this can be guaranteed (irrespective of the values of α and β) providing that

$$\langle\varepsilon_0|\varepsilon_1\rangle=0. \tag{4.4}$$

When this orthogonality condition is satisfied, the state of the $\mathcal{A}$-$\mathcal{S}$ pair is given by a reduced density matrix

$$\rho_{\mathcal{AS}}=\mathrm{Tr}_{\mathcal{E}}|\Psi\rangle\langle\Psi|$$
$$=|\alpha|^2|\uparrow\rangle\langle\uparrow||A_1\rangle\langle A_1|+|\beta|^2|\downarrow\rangle\langle\downarrow||A_0\rangle\langle A_0| \tag{4.5a}$$

containing only classical correlations.

If the condition of Eq. (4.4) did not hold, that is, if the orthogonal states of the environment were not correlated with the apparatus in the basis in which the original premeasurement was carried out, then the eigenstates of the reduced density matrix $\rho_{\mathcal{AS}}$ would be sums of products rather than simply products of states of $\mathcal{S}$ and $\mathcal{A}$. An extreme example of this situation is the predecoherence density matrix of the pure state

$$|\Phi\rangle\langle\Phi|=|\alpha|^2|\uparrow\rangle\langle\uparrow||A_1\rangle\langle A_1|+\alpha\beta^*|\uparrow\rangle\langle\downarrow||A_1\rangle\langle A_0|$$
$$+\alpha^*\beta|\downarrow\rangle\langle\uparrow||A_0\rangle\langle A_1|+|\beta|^2|\downarrow\rangle\langle\downarrow||A_0\rangle\langle A_0|. \tag{4.5b}$$

The single eigenstate of this density matrix is $|\Phi\rangle$. When expanded, $|\Phi\rangle\langle\Phi|$ contains terms that are off diagonal when expressed in any of the natural bases consisting of the tensor products of states in the two systems. Their disappearance as a result of tracing over the environment removes the basis ambiguity. Thus, for example, the reduced density matrix $\rho_{\mathcal{AS}}$, Eq. (4.5a), has the same form as the outsider description of the classical measurement, Eq. (2.44c).

In our simple model, pointer states are easy to characterize. To leave pointer states untouched, the Hamiltonian of interaction $H_{\mathcal{AE}}$ should have the same structure as that for the c-NOT, Eqs. (2.9) and (2.10). It should be a function of the pointer observable, $\hat{A}=a_0|A_0\rangle\langle A_0|+a_1|A_1\rangle\langle A_1|$ of the apparatus. Then the states of the environment will bear an imprint of the pointer states $\{|A_0\rangle,|A_1\rangle\}$. As noted in Sec. II,

$$[H_{\mathcal{AE}},\hat{A}]=0 \tag{4.6}$$

immediately implies that $\hat{A}$ is a control, and its eigenstates will be preserved.

1. Decoherence of a single qubit

An example of continuous decoherence is afforded by two-state apparatus $\mathcal{A}$ interacting with an environment of N other spins (Zurek, 1982). The two apparatus states are $\{|\Uparrow\rangle,|\Downarrow\rangle\}$. For the simplest, yet already interesting example, the self-Hamiltonian of the apparatus disappears, $H_{\mathcal{A}}=0$, and the interaction has the form

$$H_{\mathcal{AE}}=(|\Uparrow\rangle\langle\Uparrow|-|\Downarrow\rangle\langle\Downarrow|)\otimes\sum_k g_k(|\uparrow\rangle\langle\uparrow|-|\downarrow\rangle\langle\downarrow|)_k. \tag{4.7}$$

Under the influence of this Hamiltonian the initial state,

$$|\Phi(0)\rangle=(a|\Uparrow\rangle+b|\Downarrow\rangle)\prod_{k=1}^{N}(\alpha_k|\uparrow\rangle_k+\beta_k|\downarrow\rangle_k), \tag{4.8}$$

evolves into

$$|\Phi(t)\rangle=a|\Uparrow\rangle|\mathcal{E}_{\Uparrow}(t)\rangle+b|\Downarrow\rangle|\mathcal{E}_{\Downarrow}(t)\rangle; \tag{4.9}$$

$$|\mathcal{E}_{\Uparrow}(t)\rangle=\prod_{k=1}^{N}(\alpha_k e^{ig_k t}|\uparrow\rangle_k+\beta_k e^{-ig_k t}|\downarrow\rangle_k)=|\mathcal{E}_{\Downarrow}(-t)\rangle. \tag{4.10}$$

The reduced density matrix is

$$\rho_{\mathcal{A}}=|a|^2|\Uparrow\rangle\langle\Uparrow|+ab^*r(t)|\Uparrow\rangle\langle\Downarrow|+a^*br^*(t)|\Downarrow\rangle\langle\Uparrow|$$
$$+|b|^2|\Downarrow\rangle\langle\Downarrow|. \tag{4.11}$$

The coefficient $r(t)=\langle\mathcal{E}_{\Uparrow}|\mathcal{E}_{\Downarrow}\rangle$ determines the relative size of the off-diagonal terms. It is given by

$$r(t)=\prod_{k=1}^{N}[\cos 2g_k t+i(|\alpha_k|^2-|\beta_k|^2)\sin 2g_k t]. \tag{4.12}$$

For large environments consisting of many (N) spins at large times the off-diagonal terms are typically small:

$$|r(t)|^2\simeq 2^{-N}\prod_{k=1}^{N}[1+(|\alpha_k|^2-|\beta_k|^2)^2]. \tag{4.13}$$

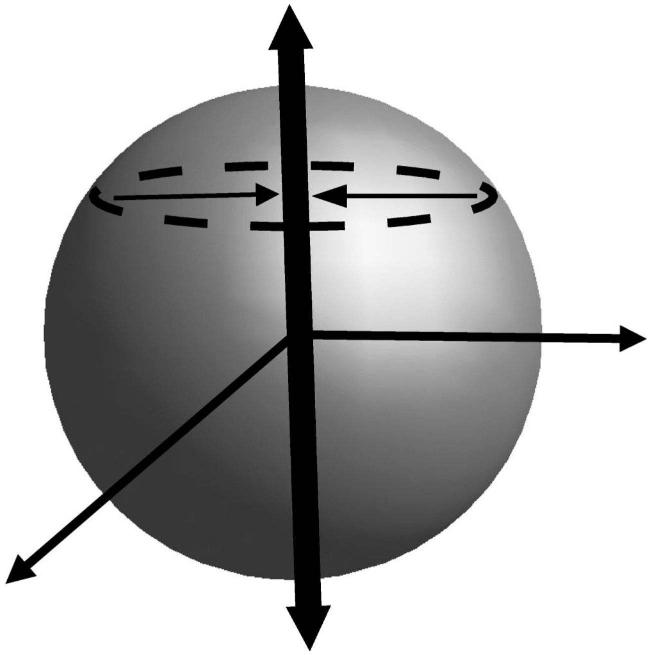


FIG. 3. Schematic representation of the effect of decoherence on a Bloch sphere. When interaction with the environment singles out pointer states located at the poles of the Bloch sphere, pure states (which lie on its surface) will evolve towards the vertical axis. This classical core is a set of all the mixtures of the pointer states.

The density matrix of any two-state system can be represented by a point in three-dimensional space. In terms of the coefficients a, b, and $r(t)$, the coordinates of the point representing it are $z=(|a|^2-|b|^2)$, $x=\mathrm{Re}(ab^*r)$, and $y=\mathrm{Im}(ab^*r)$, the real and imaginary parts of the complex ab^*r. When the state is pure, $x^2+y^2+z^2=1$. Pure states lie on the surface of the Bloch sphere (Fig. 3).

Any conceivable (unitary or nonunitary) quantum evolution can be thought of as a transformation of the surface of the pure states into the ellipsoid contained inside the Bloch sphere. Deformation of the Bloch sphere surface caused by decoherence is a special case of such general evolutions (Zurek, 1982, 1983; Berry, 1995). Decoherence does not affect $|a|$ or $|b|$. Hence evolution due to decoherence alone occurs in the $z=$const plane. Such a slice through the Bloch sphere would show the point representing the state at a fraction $|r(t)|$ of its maximum distance. The complex $r(t)$ can be expressed as the sum of the complex phase factors rotating with the frequencies given by the differences $\Delta\omega_j$ between the energy eigenvalues of the interaction Hamiltonian, weighted with the probabilities of finding them in the initial state:

$$r(t)=\sum_{j=1}^{2^N} p_j \exp(-i\Delta\omega_j t). \tag{4.14}$$

The index j denotes the environment part of the energy eigenstates of the interaction Hamiltonian, Eq. (4.7), for example: $|j\rangle=|\uparrow\rangle_1\otimes|\downarrow\rangle_2\otimes\cdots\otimes|\uparrow\rangle_N$. The corresponding differences between the energies of the eigenstates $|\Uparrow\rangle|j\rangle$ and $|\Downarrow\rangle|j\rangle$ are $\Delta\omega_j=\langle\Uparrow|\langle j|H_{\mathcal{AE}}|j\rangle|\Downarrow\rangle$. There are 2^N distinct $|j\rangle$'s, and, barring degeneracies, the same number of different $\Delta\omega_j$'s. Probabilities p_j are given by

$$p_j=|\langle j|\mathcal{E}(t=0)\rangle|^2, \tag{4.15}$$

which is in turn easily expressed in terms of the appropriate squares of α_k and β_k.

The evolution of $r(t)$, Eq. (4.14), is a consequence of the rotations of the complex vectors $p_k\exp(-i\Delta\omega_j t)$ with different frequencies. The resultant $r(t)$ will then start with the amplitude 1 and, as is anticipated by Eq. (4.13), quickly "crumble" to

$$\langle|r(t)|^2\rangle\sim\sum_{j=1}^{2^N} p_j^2\sim 2^{-N}. \tag{4.16}$$

In this sense, decoherence is exponentially effective. The magnitude of the off-diagonal terms decreases exponentially fast, with the physical size N of the environment effectively coupled to the state of the system.

We note that the effectiveness of einselection depends on the initial state of the environment. When $\mathcal{E}$ is in the kth eigenstate of $H_{\mathcal{AE}}$, $p_j=\delta_{jk}$, the coherence in the system will be retained. This special environment state is, however, unlikely in realistic circumstances.

2. The classical domain and quantum halo

The geometry of flows induced by decoherence in a Bloch sphere exhibits characteristics encountered in general:

(i) A classical set of the einselected pointer states ($\{|\Uparrow\rangle,|\Downarrow\rangle\}$ in our case). Pointer states are the pure states least affected by decoherence.

(ii) A classical domain consisting of all the pointer states and their mixtures. In Fig. 3 this corresponds to the section $[-1,+1]$ of the z axis.

(iii) The quantum domain, the rest of the volume of the Bloch sphere, consisting of more general density matrices.

Visualizing the decoherence-induced decomposition of Hilbert space may be possible only in the simple case studied here, but whenever decoherence leads to classicality, the emergence of generalized and often approximate versions of the elements (i)–(iii) is expected.

As a result of decoherence the part of Hilbert space outside the classical domain is ruled out by einselection. The severity of the prohibition on its states varies. One may measure the nonclassicality of (pure or mixed) states by quantifying their distance from the classical domain with the rate of entropy production and comparing it to the much lower rate in the classical domain. Classical pointer states would then be enveloped by a "quantum halo" (Anglin and Zurek, 1996) of nearby, relatively decoherence-resistant but still somewhat quantum states, with more flagrantly quantum (and more fragile) Schrödinger-cat states further away.

By the same token, one can define an einselection-induced metric in the classical domain, with the distance

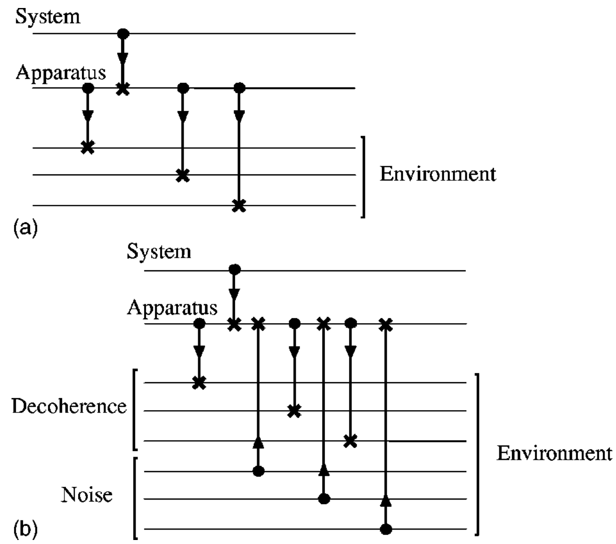


FIG. 4. Information transfer in a c-NOT or a c-shift "caricature" of measurement, decoherence, and decoherence with noise. Bit-by-bit measurement is shown on the top. This diagram is the fundamental logic circuit used to represent decoherence affecting the measuring apparatus. Note that the direction of the information flow in decoherence, from the decohering apparatus and to the environment, differs from the information flow associated with noise. In short, as a result of decoherence, environment is perturbed by the state of the system. Noise is, by contrast, perturbation inflicted by the environment. Preferred pointer states are selected so as to minimize the effect of the environment—to minimize the number of c-NOT's pointing from the environment at the expense of those pointing towards it.

between two pointer states given by the rate of entropy production of their superposition. This is not the only way to define a distance. As we shall see in Sec. VII, the redundancy of the record of a state imprinted on the environment is a very natural measure of its classicality. In the course of decoherence, pointer states tend to be recorded redundantly and can be deduced by intercepting a very small fraction of the environment (Zurek, 2000; Dalvit, Dziarmaga, and Zurek, 2001; Ollivier, Poulin, and Zurek, 2002).

3. Einselection and controlled shifts

Discussion of decoherence can be generalized to the situation in which the system, the apparatus, and the environment have many states, and their interactions are complicated. Here we assume that the system is isolated, and that it interacts with the apparatus in the c-shift manner discussed in Sec. II. As a result of that interaction the state of the apparatus becomes entangled with the state of the system: $(\Sigma_i \alpha_i |s_i\rangle)|A_0\rangle \rightarrow \Sigma_i \alpha_i |s_i\rangle |A_i\rangle$. This state suffers from basis ambiguity: the entanglement of $\mathcal{S}$ and $\mathcal{A}$ implies that for any state of either there exists a corresponding pure state of its partner. Indeed, when the initial state of $\mathcal{S}$ is chosen to be one of the eigenstates of the conjugate basis,

$$|r_l\rangle = N^{-1/2} \sum_{k=0}^{N-1} \exp(2\pi i k l/N)|s_k\rangle, \tag{4.17}$$

the c-shift could equally well represent a measurement of the apparatus (in the basis conjugate to $\{|A_k\rangle\}$) by the system. Thus it is not just the basis that is ambiguous, but also the roles of the control (system) and of the target (apparatus), which can be reversed when the conjugate basis is selected. These ambiguities can be removed by recognizing the role of the environment.

Figure 4 captures the essence of an idealized decoherence process that allows the apparatus to be—in spite of its interaction with the environment—a noiseless classical communication channel (Schumacher, 1996; Lloyd, 1997). This is possible because the $\mathcal{A}$-$\mathcal{E}$ c-shifts do not disturb the pointer states.

The advantage of this idealization of the decoherence process as a sequence of c-shifts lies in its simplicity. However, the actual process of decoherence is usually continuous (so that it can only be approximately broken up into discrete c-shifts). Moreover, in contrast to the c-NOT's used in quantum logic circuits, the record inscribed in the environment is usually distributed over many degrees of freedom. Last but not least, the observable of the apparatus (or any other open system) may be subject to noise (and not just decoherence), or its self-Hamiltonian may rotate instantaneous pointer states into their superpositions. These very likely complications will be investigated in specific models below.

Decoherence is caused by a premeasurementlike process carried out by the environment $\mathcal{E}$:

$$|\Psi_{\mathcal{SA}}\rangle|\varepsilon_0\rangle = \left(\sum_j \alpha_j |s_j\rangle|A_j\rangle\right)|\varepsilon_0\rangle \rightarrow \sum_j \alpha_j |s_j\rangle|A_j\rangle|\varepsilon_j\rangle$$
$$= |\Phi_{\mathcal{SAE}}\rangle. \tag{4.18}$$

Decoherence leads to einselection when the states of the environment $|\varepsilon_j\rangle$ corresponding to different pointer states become orthogonal:

$$\langle \varepsilon_i | \varepsilon_j \rangle = \delta_{ij}. \tag{4.19}$$

Then the Schmidt decomposition of the state vector $|\Phi_{\mathcal{SAE}}\rangle$ into composite subsystems $\mathcal{SA}$ and $\mathcal{E}$ yields product states $|s_j\rangle|A_j\rangle$ as partners of the orthogonal environment states. The decohered density matrix describing the $\mathcal{SA}$ pair is then diagonal in product states:

$$\rho_{\mathcal{SA}}^D = \sum_j |\alpha_j|^2 |s_j\rangle\langle s_j||A_j\rangle\langle A_j| = \mathrm{Tr}_{\mathcal{E}} |\Phi_{\mathcal{SAE}}\rangle\langle\Phi_{\mathcal{SAE}}|. \tag{4.20}$$

For simplicity we shall often omit reference to the object that does not interact with the environment (here, the system $\mathcal{S}$). Nevertheless, preservation of the $\mathcal{SA}$ correlations is the criterion defining the pointer basis. Invoking it would eliminate much confusion (see, for example, discussions in Halliwell, Perez-Mercader, and Zurek, 1994; Venugopalan, 1994, 2000). The density matrix of a single object in contact with the environment will always be diagonal in an (instantaneous) Schmidt basis. This instantaneous diagonality should not be used as the sole criterion for classicality (although see Zeh, 1973, 1990;

Albrecht, 1992, 1993). Rather, the ability of certain states to retain correlations in spite of coupling to the environment is decisive.

When the interaction with the apparatus has the form

$$H_{\mathcal{A}\mathcal{E}}=\sum_{k,l,m} g_{klm}^{\mathcal{A}\mathcal{E}}|A_k\rangle\langle A_k||\varepsilon_l\rangle\langle\varepsilon_m|+\text{H.c.}, \tag{4.21}$$

the basis $\{|A_k\rangle\}$ is left unperturbed and any correlation with the states $\{|A_k\rangle\}$ is preserved. But, by definition, pointer states preserve correlations in spite of decoherence, so that any observable $\hat{A}$ codiagonal with the interaction Hamiltonian will be pointer observable. For when the interaction is a function of $\hat{A}$, it can be expanded in $\hat{A}$ as a power series, so it commutes with $\hat{A}$:

$$[H_{\mathcal{A}\mathcal{E}}(\hat{A}),\hat{A}]=0. \tag{4.22}$$

The dependence of the interaction Hamiltonian on the observable is an obvious precondition for the monitoring of that observable by the environment. This admits the existence of degenerate pointer eigenspaces of $\hat{A}$.

B. Einselection as the selective loss of information

The establishment of a measurementlike correlation between the apparatus and the environment changes the density matrix from the premeasurement $\rho_{\mathcal{S}\mathcal{A}}^P$ to the decohered $\rho_{\mathcal{S}\mathcal{A}}^D$, Eq. (4.20). For the initially pure $|\Psi_{\mathcal{S}\mathcal{A}}\rangle$, Eq. (4.18), this transition is represented by

$$\rho_{\mathcal{S}\mathcal{A}}^P=\sum_{i,j}\alpha_i\alpha_j^*|s_i\rangle\langle s_j||A_i\rangle\langle A_j|$$
$$\rightarrow\sum_i|\alpha_i|^2|s_i\rangle\langle s_i||A_i\rangle\langle A_i|=\rho_{\mathcal{S}\mathcal{A}}^D. \tag{4.23}$$

Einselection is accompanied by an increase of entropy,

$$\Delta H(\rho_{\mathcal{S}\mathcal{A}})=H(\rho_{\mathcal{S}\mathcal{A}}^D)-H(\rho_{\mathcal{S}\mathcal{A}}^P)\geqslant 0, \tag{4.24}$$

and by the disappearance of the ambiguity in what was measured (Zurek, 1981, 1993a). Thus, before decoherence, the conditional density matrices of the system, $\rho_{\mathcal{S}|C_j\rangle}$, are pure for any state $|C_j\rangle$ of the apparatus pointer. They are defined using the unnormalized

$$\tilde{\rho}_{\mathcal{S}|\Pi_j}=\text{Tr}_{\mathcal{A}}\Pi_j\rho_{\mathcal{S}\mathcal{A}}, \tag{4.25}$$

where in the simplest case $\Pi_j=|C_j\rangle\langle C_j|$ projects onto a pure state of the apparatus.[4]

[4]This can be generalized to projections onto multidimensional subspaces of the apparatus. In that case, the purity of the conditional density matrix will usually be lost during the trace over the states of the pointer. This is not surprising. When the observer reads off the pointer of the apparatus only in a coarse-grained manner, he will forgo part of the information about the system. The amplification we have considered before can prevent some of this loss of resolution due to coarse graining in the apparatus. Generalizations to density matrices that are conditioned upon projection-operator-valued measures (Kraus, 1983) are also possible.

Normalized $\rho_{\mathcal{S}|\Pi_j}$ can be obtained by using the probability of the outcome:

$$\rho_{\mathcal{S}|\Pi_j}=p_j^{-1}\tilde{\rho}_{\mathcal{S}|\Pi_j}, \tag{4.26}$$
$$p_j=\text{Tr}\tilde{\rho}_{\mathcal{S}|\Pi_j}.$$

The conditional density matrix represents the description of the system $\mathcal{S}$ available to the observer who knows that the apparatus $\mathcal{A}$ is in a subspace defined by Π_j.

1. Conditional state, entropy, and purity

Before decoherence, $\rho_{\mathcal{S}|C_j\rangle}^P$ is pure for any state $|C_j\rangle$,

$$(\rho_{\mathcal{S}|\Pi_j}^P)^2=\rho_{\mathcal{S}|\Pi_j}^P\ \forall|C_j\rangle; \tag{4.27a}$$

providing the initial premeasurement state, Eq. (4.23), was pure as well. It follows that

$$H(\rho_{\mathcal{S}\mathcal{A}|C_j\rangle}^P)=0\ \forall|C_j\rangle. \tag{4.28a}$$

For this same case given by the initially pure $\rho_{\mathcal{S}\mathcal{A}}^P$ of Eq. (4.23), conditional density matrices obtained from the decohered $\rho_{\mathcal{S}\mathcal{A}}^D$ will be pure if and only if they are conditioned upon the pointer states $\{|A_k\rangle\}$:

$$(\rho_{\mathcal{S}|C_j\rangle}^D)^2=\rho_{\mathcal{S}|C_j\rangle}^P=|s_k\rangle\langle s_k|\Leftrightarrow|C_j\rangle=|A_j\rangle; \tag{4.27b}$$

$$H(\rho_{\mathcal{S}|A_j\rangle}^D)=H(\rho_{\mathcal{S}|A_j\rangle}^P). \tag{4.28b}$$

This last equation is valid even when the initial states of the system and of the apparatus are not pure. Thus only in the pointer basis will the predecoherence strength of the $\mathcal{S}\mathcal{A}$ correlation be maintained. In all other bases

$$\text{Tr}(\rho_{\mathcal{S}|C_j\rangle}^D)^2<\text{Tr}\rho_{\mathcal{S}|C_j\rangle}^D;\quad |C_j\rangle\notin\{|A_j\rangle\}, \tag{4.27c}$$

$$H(\rho_{\mathcal{S}|C_j\rangle}^P)<H(\rho_{\mathcal{S}|C_j\rangle}^D);\quad |C_j\rangle\notin\{|A_j\rangle\}. \tag{4.28c}$$

In particular, in the basis $\{|B_j\rangle\}$ conjugate to the pointer states $\{|A_j\rangle\}$, Eq. (2.14), there is no correlation left with the state of the system. That is,

$$\rho_{\mathcal{S}|B_j\rangle}^D=N^{-1}\sum_k|s_k\rangle\langle s_k|=\mathbf{1}/N, \tag{4.29}$$

where $\mathbf{1}$ is a unit density matrix. Consequently

$$(\rho_{\mathcal{S}|B_j\rangle}^D)^2=\rho_{\mathcal{S}|B_j\rangle}/N, \tag{4.27d}$$

$$H(\rho_{\mathcal{S}|B_j\rangle}^D)=H(\rho_{\mathcal{S}|B_j\rangle}^P)-\ln N=-\ln N. \tag{4.28d}$$

Note that, initially, the conditional density matrices were also pure in the conjugate (and any other) basis, provided that the initial state was the pure entangled projection operator $\rho_{\mathcal{S}\mathcal{A}}^P=|\Psi_{\mathcal{S}\mathcal{A}}\rangle\langle\Psi_{\mathcal{S}\mathcal{A}}|$, Eq. (4.23).

2. Mutual information and discord

Selective loss of information everywhere except in the pointer states is the essence of einselection. It is reflected in the change of the mutual information which starts from

$$\mathcal{I}^P(\mathcal{S}{:}\mathcal{A}) = H(\rho_\mathcal{S}^P) + H(\rho_\mathcal{A}^P) - H(\rho_{\mathcal{S},\mathcal{A}}^P)$$
$$= -2\sum_i |\alpha_i|^2 \ln|\alpha_i|^2. \tag{4.30a}$$

As a result of einselection, for initially pure cases, this decreases to at most, half its initial value:

$$\mathcal{I}^D(\mathcal{S}{:}\mathcal{A}) = H(\rho_\mathcal{S}^D) + H(\rho_\mathcal{A}^D) - H(\rho_{\mathcal{S},\mathcal{A}}^D)$$
$$= -\sum_i |\alpha_i|^2 \ln|\alpha_i|^2. \tag{4.30b}$$

This level is reached when the pointer basis coincides with the Schmidt basis of $|\Psi_{\mathcal{S}\mathcal{A}}\rangle$. The decrease in mutual information is due to the increase of the joint entropy $H(\rho_{\mathcal{S},\mathcal{A}})$:

$$\Delta\mathcal{I}(\mathcal{S}{:}\mathcal{A}) = \mathcal{I}^P(\mathcal{S}{:}\mathcal{A}) - \mathcal{I}^D(\mathcal{S}{:}\mathcal{A})$$
$$= H(\rho_{\mathcal{S},\mathcal{A}}^D) - H(\rho_{\mathcal{S},\mathcal{A}}^P) = \Delta H(\rho_{\mathcal{S},\mathcal{A}}). \tag{4.31}$$

Classically, an equivalent definition of the mutual information obtains from the asymmetric formula

$$\mathcal{J}_\mathcal{A}(\mathcal{S}{:}\mathcal{A}) = H(\rho_\mathcal{S}) - H(\rho_{\mathcal{S}|\mathcal{A}}), \tag{4.32}$$

with the help of the conditional entropy $H(\rho_{\mathcal{S}|\mathcal{A}})$. Above, the subscript $\mathcal{A}$ indicates the member of the correlated pair that will be the source of the information about its partner. A symmetric counterpart of the above equation, $\mathcal{J}_\mathcal{S}(\mathcal{S}{:}\mathcal{A}) = H(\rho_\mathcal{A}) - H(\rho_{\mathcal{A}|\mathcal{S}})$, can also be written.

In the quantum case, the definition of Eq. (4.32) is so far incomplete, since a quantum analog of the classical conditional information has not yet been specified. Indeed, Eqs. (4.30a) and (4.32) jointly imply that in the case of entanglement a quantum conditional entropy $H(\rho_{\mathcal{S}|\mathcal{A}})$ would have to be negative. For in that case,

$$H(\rho_{\mathcal{S}|\mathcal{A}}) = \sum_i |\alpha_i|^2 \ln|\alpha_i|^2 < 0 \tag{4.33}$$

would be required to allow for $\mathcal{I}(\mathcal{S}{:}\mathcal{A}) = \mathcal{J}_\mathcal{A}(\mathcal{S}{:}\mathcal{A})$. Various quantum redefinitions of $\mathcal{I}(\mathcal{S}{:}\mathcal{A})$ or $H(\rho_{\mathcal{S}|\mathcal{A}})$ have been proposed to address this (Lieb, 1975; Schumacher and Nielsen, 1996; Cerf and Adami, 1997; Lloyd, 1997). We shall simply regard this fact as an illustration of the strength of the quantum correlations (i.e., entanglement) that allow $\mathcal{I}(\mathcal{S}{:}\mathcal{A})$ to violate the inequality

$$\mathcal{I}(\mathcal{S}{:}\mathcal{A}) \leq \min(H_\mathcal{S}, H_\mathcal{A}). \tag{4.34}$$

This inequality follows directly from Eq. (4.32) and the non-negativity of classical conditional entropy (see, for example, Cover and Thomas, 1991).

Decoherence decreases $\mathcal{I}(\mathcal{S}{:}\mathcal{A})$ to this allowed level (Zurek, 1983). Moreover, now the conditional entropy can be defined in the classical pointer basis as the average of partial entropies computed from the conditional $\rho_{\mathcal{S}|A_i\rangle}^D$ over the probabilities of different outcomes:

$$H(\rho_{\mathcal{S}|\mathcal{A}}) = \sum_i p_{|A_i\rangle} H(\rho_{\mathcal{S}|A_i\rangle} D). \tag{4.35}$$

Prior to decoherence, the use of probabilities would not have been legal.

For the case considered here, Eq. (4.18), the conditional entropy $H(\rho_{\mathcal{S}|\mathcal{A}}) = 0$. In the pointer basis there is a perfect correlation between the system and the apparatus, providing that the premeasurement Schmidt basis and the pointer basis coincide. Indeed, it is tempting to define a good apparatus or a classical correlation by insisting on such a coincidence.

The difficulties with conditional entropy and mutual information are a symptom of the quantum nature of the problem. The trouble with $H(\rho_{\mathcal{S}|\mathcal{A}})$ arises for states that exhibit quantum correlations—entanglement of $|\Psi_{\mathcal{S}\mathcal{A}}\rangle$ being an extreme example—and thus do not admit an interpretation based on probabilities. A useful sufficient condition for the classicality of correlations is then the existence of an apparatus basis that allows quantum versions of the two classically identical expressions for the mutual information to coincide: $\mathcal{I}(\mathcal{S}{:}\mathcal{A}) = \mathcal{J}_\mathcal{A}(\mathcal{S}{:}\mathcal{A})$ (Zurek, 2000, 2003a; Ollivier and Zurek, 2002; Vedral, 2003). Equivalently, the *discord*

$$\delta\mathcal{I}_\mathcal{A}(\mathcal{S}|\mathcal{A}) = \mathcal{I}(\mathcal{S}{:}\mathcal{A}) - \mathcal{J}_\mathcal{A}(\mathcal{S}{:}\mathcal{A}) \tag{4.36}$$

must vanish. Unless $\delta\mathcal{I}_\mathcal{A}(\mathcal{S}|\mathcal{A}) = 0$, probabilities for the distinct apparatus pointer states cannot exist.

We end this subsection with part summary, part anticipatory remarks. Pointer states retain undiminished correlations with the measured system $\mathcal{S}$, or with any other system, including observers. The loss of information caused by decoherence is given by Eq. (4.31). This loss is precisely such as to lift conditional information from the paradoxical (negative) values, Eq. (4.33), to the classically allowed level. This is equal to the information gained by the observer when he consults the apparatus pointer. This is no accident—the environment has "measured" (become correlated with) the apparatus in the very same pointer basis at which observers have to access $\mathcal{A}$ to take advantage of the remaining (classical) correlation between the pointer and the system. Only when observers and the environment monitor codiagonal observables do they not get in each other's way.

In the idealized case, the preferred basis was distinguished by its ability to retain perfect correlations with the system in spite of decoherence. This remark will serve as a guide in other situations. It will lead to a criterion—the predictability sieve—used to identify preferred states in less idealized circumstances. For example, when the self-Hamiltonian of the system is nontrivial, or when the commutation relation, Eq. (4.22), does not hold exactly for any observable of $\mathcal{S}$, we shall seek states that are best in retaining correlations with the other systems.

C. Decoherence, entanglement, dephasing, and noise

In the symbolic representation of Fig. 4, noise is the process in which the environment acts as a control, inscribing information about its state on the state of the system, which assumes the role of the target. However, the direction of the information flow in c-NOT's and c-shifts depends on the choice of initial states. Control and target switch roles when, for a given interaction Hamiltonian, one prepares the input of the c-NOT in the basis conjugate to the logical pointer states. Einselected states correspond to the set of states that, when used in c-NOT's or c-shifts, minimizes the effect of interactions directed from the environment to the system.

Einselection is caused by the premeasurement carried out by the environment on the pointer states. Decoherence follows from Heisenberg's indeterminacy. Pointer observable is measured by the environment. Therefore the complementary observable must become at least as indeterminate as is demanded by Heisenberg's principle. As the environment and the systems entangle through an interaction that favors a set of pointer states, their phases become indeterminate [see Eq. (4.29) and the discussion of envariance in Sec. VI]. Decoherence can be thought of as the resulting loss of phase relations.

Observers can be ignorant of phases for reasons that do not lead to an imprint of the state of the system on the environment. Classical noise can cause such *dephasing* when the observer does not know the time-dependent classical perturbation Hamiltonian responsible for this unitary, but unknown, evolution. For example, in the predecoherence state vector, Eq. (4.18), random-phase noise will cause a transition:

$$|\Psi_{\mathcal{SA}}\rangle = \left(\sum_j \alpha_j |s_j\rangle|A_j\rangle\right) \to \sum_j \alpha_j \exp(i\phi_j^{(n)})|s_j\rangle|A_j\rangle = |\Psi_{\mathcal{SA}}^{(n)}\rangle. \tag{4.37}$$

A dephasing Hamiltonian acting either on the system or on the apparatus can lead to such an effect. In this second case its form could be

$$H_d^{(n)} = \sum_j \dot{\phi}_j^{(n)}(t)|A_j\rangle\langle A_j|. \tag{4.38}$$

In contrast to interactions causing premeasurements, entanglement, and decoherence, H_d cannot influence the nature or the degree of the $\mathcal{SA}$ correlations. H_d does not imprint the states of $\mathcal{S}$ or $\mathcal{A}$ anywhere else in the universe. For each individual realization n of the phase noise [each selection of $\{\phi_j^{(n)}(t)\}$ in Eq. (4.37)] the state $|\Phi_{\mathcal{SA}}^{(n)}\rangle$ remains pure. Given only $\{\phi_j^{(n)}\}$ one could restore the predephasing state on a case-by-case basis. However, in the absence of such detailed information, one is often forced to represent $\mathcal{SA}$ by the density matrix averaged over the ensemble of noise realizations:

$$\begin{aligned}\overline{\rho_{\mathcal{SA}}} &= \langle |\Psi_{\mathcal{SA}}\rangle\langle\Psi_{\mathcal{SA}}|\rangle \\ &= \sum_j |\alpha_j|^2 |s_j\rangle\langle s_j||A_j\rangle\langle A_j| \\ &\quad + \sum_{j,k}\sum_n e^{i[\phi_j^{(n)} - \phi_k^{(n)}]}\alpha_j\alpha_k|s_j\rangle\langle s_k||A_j\rangle\langle A_k|.\end{aligned} \tag{4.39}$$

In this phase-averaged density matrix off-diagonal terms will be suppressed and may even completely disappear. Nevertheless, each member of the ensemble may exist in a state as pure as it was before dephasing. Nuclear magnetic resonance (NMR) offers examples of dephasing (which can be reversed using spin echo). Dephasing is a loss of phase coherence between members of the ensemble due to differences in the noise in phases each member experiences. It does not result in an information transfer to the environment.

Dephasing cannot be used to justify the existence of a preferred basis in individual quantum systems. Nevertheless, the ensemble as a whole may obey the same master equation as do individual systems entangling with the environment. Indeed, many of the symptoms exhibited by, e.g., the expectation values for a single decohering system can be reproduced by ensemble averages in this setting. In spite of the light shed on this issue by the discussion of simple cases (Wootters and Zurek, 1979; Stern, Aharonov, and Imry, 1989), more remains to be understood, perhaps by considering the implications of envariance (see Sec. VI).

Noise is an even more familiar and less subtle effect represented by transitions that break the one-to-one correspondence in Eq. (4.39). Noise in the apparatus would cause a random rotation of states $|A_j\rangle$. It could be modeled by a collection of Hamiltonians similar to $H_d^{(n)}$ but not codiagonal with the observable of interest. Then, after an ensemble average similar to Eq. (4.39), the one-to-one correspondence between $\mathcal{S}$ and $\mathcal{A}$ would be lost. However, as before, the evolution is unitary for each n, and the unperturbed state could be reconstructed from information the observer could have in advance.

Hence, in the case of dephasing or noise, information about the cause obtained either in advance, or afterwards, suffices to undo the effect. Decoherence relies on entangling interactions [although, strictly speaking, it need not involve entanglement (Eisert and Plenio, 2002)]. Thus neither prior nor posterior knowledge of the state of the environment is enough. Transfer of information about a decohering system to the environment is essential, and plays a key role in the interpretation.

We note that, while the nomenclature used here seems the most sensible to this author and is widely used, it is unfortunately not universal. For example, in the context of quantum computation "decoherence" is sometimes used to describe any process that can cause errors (but see related discussion in Nielsen and Chuang, 2000).

D. Predictability sieve and einselection

The evolution of a quantum system prepared in a classical state should emulate classical evolution that can be idealized as a "trajectory"—a predictable sequence of objectively existing states. For a purely unitary evolution, all of the states in the Hilbert space retain their purity and are therefore equally predictable. However, in the presence of an interaction with the environment, a generic superposition representing correlated states of the system and of the apparatus will decay into a mixture diagonal in pointer states, Eq. (4.23). Only when the predecoherence state of $\mathcal{SA}$ is a product of a single apparatus pointer state $|A_i\rangle$ with the corresponding outcome state of the system (or a mixture of such product states) does decoherence have no effect:

$$\rho^P_{\mathcal{SA}} = |s_i\rangle\langle s_i||A_i\rangle\langle A_i| = \rho^D_{\mathcal{SA}}. \tag{4.40}$$

A correlation of a pointer state with any state of an isolated system is untouched by the environment. By the same token, when the observer prepares $\mathcal{A}$ in the pointer state $|A_i\rangle$, he can count on its remaining pure. One can even think of $|s_i\rangle$ as the record of the pointer state of $\mathcal{A}$. Einselected states are predictable: they preserve correlations and hence are effectively classical.

In the above idealized cases, the predictability of some states follows directly from the structure of the relevant Hamiltonians (Zurek, 1981). A correlation with a subspace associated with a projection operator P_A will be immune to decoherence providing that

$$[H_{\mathcal{A}} + H_{\mathcal{AE}}, P_A] = 0. \tag{4.41}$$

In more realistic cases it is difficult to demand the exact conservation guaranteed by such a commutation condition. Looking for approximate conservation may still be a good strategy. The various densities used in hydrodynamics are one obvious choice (see, e.g., Gell-Mann and Hartle, 1990, 1993).

In general, it is useful to invoke a more fundamental predictability criterion (Zurek, 1993a). One can measure the loss of predictability caused by evolution for every pure state $|\Psi\rangle$ by von Neumann entropy or some other measure of predictability such as the purity:

$$\mathrm{s}_\Psi(t) = \mathrm{Tr}\rho^2_\Psi(t). \tag{4.42}$$

In either case, predictability is a function of time and a functional of the initial state as $\rho_\Psi(0) = |\Psi\rangle\langle\Psi|$. Pointer states are obtained by maximizing the predictability functional over $|\Psi\rangle$. When decoherence leads to classicality, good pointer states exist, and the answer is robust.

A *predictability sieve* sifts all of Hilbert space, ordering states according to their predictability. The top of the list will be the most classical. This point of view allows for unification of the simple definition of the pointer states in terms of the commutation relation, Eq. (4.41), with the more general criteria required to discuss classicality in other situations. The eigenstates of the exact pointer observable are selected by the sieve. Equation (4.41) guarantees that they will retain their purity in spite of the environment and are (somewhat trivially) predictable.

The predictability sieve can be generalized to situations where the initial states are mixed (Paraoanu and Scutera, 1998; Paraoanu, 2002). Often whole subspaces emerge from the predictability sieve, naturally leading to decoherence-free subspaces (see, for example, Lidar *et al.*, 1999) and can be adapted to yield "noiseless subsystems" (which are a non-Abelian generalization of pointer states; see, for example, Knill, Laflamme, and Viola, 2000; Zanardi, 2001). However, calculations are in general quite difficult even for the initial pure-state cases.

The idea of the sieve selecting candidates for the classical states is one decade old, but still only partly explored. We shall see it in action below. We have outlined two criteria for sifting through the Hilbert space in search of classicality; von Neumann entropy and purity define, after all, two distinct functionals. Entropy is arguably an obvious information-theoretic measure of predictability loss. Purity is much easier to compute. It is often used as a "cheap substitute" and has a physical significance of its own. It seems unlikely that pointer states selected by the predictability and purity sieves could differ substantially. After all,

$$-\mathrm{Tr}\rho \ln \rho = \mathrm{Tr}\rho\{(\mathbf{1} - \rho) - (\mathbf{1} - \rho)^2/2 + \cdots, \tag{4.43}$$

so that one can expect the most predictable states to also remain the purest (Zurek, 1993a). However, the expansion, Eq. (4.43), is very slowly convergent. Therefore a more mathematically satisfying treatment of the differences between the states selected by these two criteria would be desirable, especially in cases where (as we shall see in the next section for the harmonic oscillator) the preferred states are coherent (Zurek, Habib, and Paz, 1993), and hence the classical domain forms a relatively broad "mesa" in Hilbert space.

The possible discrepancy between the states selected by sieves based on predictability and those based on purity raises a more general question. Will all the sensible criteria yield identical answers? After all, one can imagine other reasonable criteria for classicality, such as the yet-to-be-explored "distinguishability sieve" of Schumacher (1999), which picks out states whose descendants are most distinguishable in spite of decoherence. Moreover, as we shall see in Sec. VII (also, Zurek, 2000), one can ascribe classicality to the states that are most redundantly recorded by the environment. The menu of various classicality criteria already contains several positions, and more may be added in the future. There is no *a priori* reason to expect that all of these criteria will lead to identical sets of preferred states. It is nevertheless reasonable to hope that, in the macroscopic limit in which classicality is indeed expected, differences between various sieves should be negligible. The same stability in the selection of the classical domain is expected with respect to changes of, say, the time of evolution from the initial pure state. Reasonable changes of such details within the time interval in which einselection is expected to be effective should lead to more or less similar preferred states, and certainly to preferred states contained within each other's "quantum halo" (Anglin and Zurek, 1996). As noted above, this seems to be the case in the examples explored to date. It remains to be seen whether all criteria will agree in other situations of interest.

V. EINSELECTION IN PHASE SPACE

Einselection in phase space is a special, yet very important, topic. It should lead to phase-space points and trajectories and to classical (Newtonian) dynamics. The special role of position in classical physics can be traced to the nature of interactions that depend on distance (Zurek, 1981, 1982, 1991) and therefore commute with position [see Eq. (4.22)]. Evolution of open systems includes, however, the flow in phase space induced by the self-Hamiltonian. Consequently a set of preferred states turns out to be a compromise, localized in both position and momentum, localized in phase space.

Einselection is responsible for the classical structure of phase space. States selected by the predictability sieve become phase-space "points," and their time-ordered sequences turn into trajectories. In underdamped, classically regular systems one can recover this phase-space structure along with (almost) reversible evolution. In chaotic systems there is a price to be paid for classicality: combination of decoherence with the exponential divergence of classical trajectories (which is the defining feature of chaos) leads to entropy production at a rate given—in the classical limit—by the sum of positive Lyapunov exponents. Thus the second law of thermodynamics can emerge from the interplay of classical dynamics and quantum decoherence, with entropy production caused by information "leaking" into the environment (Zurek and Paz, 1994, 1995a; Zurek, 1998b; Paz and Zurek, 2001).

A. Quantum Brownian motion

The quantum Brownian motion model consists of an environment $\mathcal{E}$—a collection of harmonic oscillators (coordinates q_n, masses m_n, frequencies ω_n, and coupling constants c_n)—interacting with the system $\mathcal{S}$ (coordinate x), with a mass M and a potential $V(x)$. We shall often consider harmonic $V(x)=M\Omega^2x^2/2$ so that the whole $\mathcal{SE}$ is linear and one can obtain an exact solution. This assumption will be relaxed later.

The Lagrangian of the system-environment entity is

$$L(x,q_n)=L_{\mathcal{S}}(x)+L_{\mathcal{SE}}(x,\{q_n\}); \tag{5.1}$$

the system alone has the Lagrangian

$$L_S(x)=\frac{M}{2}\dot{x}^2-V(x)=\frac{M}{2}(\dot{x}^2-\Omega^2x^2). \tag{5.2}$$

The effect of the environment is modeled by the sum of the Lagrangians of individual oscillators and of the system-environment interaction terms:

$$L_{\mathcal{SE}}=\sum_n \frac{m_n}{2}\left[\dot{q}_n^2-\omega_n^2\left(q_n-\frac{c_n x}{m_n\omega_n^2}\right)^2\right]. \tag{5.3}$$

This Lagrangian takes into account the renormalization of the potential energy of the Brownian particle. The interaction depends (linearly) on the position x of the harmonic oscillator. Hence we expect x to be an instantaneous pointer observable. In combination with the harmonic evolution this leads to Gaussian pointer states, well localized in both x and p. An important characteristic of the model is the spectral density of the environment:

$$C(\omega)=\sum_n \frac{c_n^2}{2m_n\omega_n}\delta(\omega-\omega_n). \tag{5.4}$$

The effect of the environment can be expressed through the propagator J acting on the reduced $\rho_{\mathcal{S}}$:

$$\rho_{\mathcal{S}}(x,x',t)=\int dx_0dx_0'J(x,x',t|x_0,x_0',t_0)\rho_{\mathcal{S}}(x_0,x_0',t_0). \tag{5.5}$$

We focus on the case in which the system and the environment are initially statistically independent, so that their density matrices start from a product state:

$$\rho_{\mathcal{SE}}=\rho_{\mathcal{S}}\rho_{\mathcal{E}}. \tag{5.6}$$

This is a restrictive assumption. One can try to justify it as an idealization of a measurement that correlates $\mathcal{S}$ with the observer and destroys correlations of $\mathcal{S}$ with $\mathcal{E}$, but that is only an approximation, since realistic measurements leave partial correlations with the environment intact. Fortunately, such preexisting correlations lead to only minor differences in the salient features of the subsequent evolution of the system (Anglin, Paz, and Zurek, 1997; Romero and Paz, 1997).

The evolution of the whole $\rho_{\mathcal{SE}}$ can be represented as

$$\begin{aligned}\rho_{\mathcal{SE}}(x,q,x',q',t)&\\ =\int dx_0dx_0'dq_0dq_0'\rho_{\mathcal{SE}}(x_0,q_0,x_0',q_0',t_0)&\\ \times K(x,q,t,x_0,q_0)K^*(x',q',t,x_0',q_0').&\end{aligned} \tag{5.7}$$

Above, we suppress the sum over the indices of the individual environment oscillators. The evolution operator $K(x,q,t,x_0,q_0)$ can be expressed as a path integral

$$K(x,q,t,x_0,q_0)=\int DxDq\,\exp\left(\frac{i}{\hbar}I[x,q]\right), \tag{5.8}$$

where $I[x,q]$ is the action functional that depends on the trajectories x and q. The integration must satisfy the boundary conditions

$$x(0)=x_0;\quad x(t)=x;\quad q(0)=q_0;\quad q(t)=q. \tag{5.9}$$

The expression for the propagator of the density matrix can now be written in terms of actions corresponding to the two Lagrangians, Eqs. (5.1)–(5.3):

$$\begin{aligned}J(x,x',t|x_0,x_0',t_0)&\\ =\int DxDx'\,\exp\left(\frac{i}{\hbar}(I_{\mathcal{S}}[x]-I_{\mathcal{S}}[x'])\right)&\\ \times\int dqdq_0dq_0'\rho_{\mathcal{E}}(q_0,q_0')&\\ \times\int DqDq'\,\exp\left(\frac{i}{\hbar}(I_{\mathcal{SE}}[x,q]-I_{\mathcal{SE}}[x',q'])\right).&\end{aligned} \tag{5.10}$$

The separability of the initial conditions, Eq. (5.6), was used to make the propagator depend only on the initial conditions of the environment. Collecting all terms containing integrals over $\mathcal{E}$ in the above expression leads to the influence functional (Feynman and Vernon, 1963)

$$F(x,x')=\int dq\,dq_0\,dq_0'\rho_{\mathcal{E}}(q_0,q_0') \times\int Dq\,Dq'\exp\left(\frac{i}{\hbar}(I_{\mathcal{SE}}[x,q]-I_{\mathcal{SE}}[x',q'])\right). \tag{5.11}$$

Influence functional can be evaluated explicitly for specific models of the initial density matrix of the environment. An environment in thermal equilibrium provides a useful and tractable model for the initial state. The density matrix of the nth mode of the thermal environment is

$$\rho_{\mathcal{E}_n}(q,q')=\frac{m_n\omega_n}{2\pi\hbar\sinh\left(\frac{\hbar\omega_n}{k_BT}\right)} \times\exp-\left\{\frac{m_n\omega_n}{2\pi\hbar\sinh\left(\frac{\hbar\omega_n}{k_BT}\right)} \times\left[(q_n^2+q_n'^2)\cosh\left(\frac{\hbar\omega_n}{k_BT}\right)-2q_nq_n'\right]\right\}. \tag{5.12}$$

The influence functional F can be written as (Grabert, Schramm, and Ingold, 1988)

$$i\ln F(x,x')=\int_0^t ds(x-x')(s)\int_0^s du[\eta(s-s')(x+x')(s') -i\nu(s-s')(x-x')(s')], \tag{5.13}$$

where $\nu(s)$ and $\eta(s)$ are known as the dissipation and noise kernels, respectively, and are defined in terms of the spectral density:

$$\nu(s)=\int_0^\infty d\omega C(\omega)\coth(\hbar\omega\beta/2)\cos(\omega s); \tag{5.14}$$

$$\eta(s)=\int_0^\infty d\omega C(\omega)\sin(\omega s). \tag{5.15}$$

With the assumption of thermal equilibrium at $k_BT=1/\beta$, and in the harmonic-oscillator case $V(x)=M\Omega^2x^2/2$, the integrand of Eq. (5.10) for the propagator is Gaussian. The integral can be computed exactly and should also have a Gaussian form. The result can be conveniently written in terms of the diagonal and off-diagonal coordinates of the density matrix in the position representation, $X=x+x', Y=x-x'$:

$$J(X,Y,t|X_0,Y_0,t_0) =\frac{b_3}{2\pi}\frac{\exp[i(b_1XY+b_2X_0Y-b_3XY_0-b_4X_0Y_0)]}{\exp(a_{11}Y^2+2a_{12}YY_0+a_{22}Y_0^2)}. \tag{5.16}$$

The time-dependent coefficients b_k and a_{ij} are computed from the noise and dissipation kernels, which reflect properties of the environment. They obtain from the solutions of the equation

$$\ddot{u}(s)+\Omega^2u(s)+2\int_0^s ds\,\eta(s-s')u(s')=0, \tag{5.17}$$

where Ω is the "bare frequency" of the oscillator. Two such solutions that satisfy the boundary conditions $u_1(0)=u_2(t)=1$ and $u_1(t)=u_2(0)=0$ can be used for this purpose. They yield the coefficients of the Gaussian propagator through

$$b_{1(2)}(t)=\dot{u}_{2(1)}(t)/2,\quad b_{3(4)}(t)=\dot{u}_{2(1)}(0)/2, \tag{5.18a}$$

$$a_{ij}(t)=\frac{1}{1+\delta_{ij}}\int_0^t ds\int_0^t ds'u_i(s)u_j(s')\nu(s-s'). \tag{5.18b}$$

The master equation can now be obtained by taking the time derivative of Eq. (5.5), which in effect reduces to the computation of the derivative of the propagator, Eq. (5.16), above:

$$\dot{J}=\{\dot{b}_3/b_3+i\dot{b}_1XY+i\dot{b}_2X_0Y-i\dot{b}_3XY_0-i\dot{b}_4X_0Y_0 -\dot{a}_{11}Y^2-\dot{a}_{12}YY_0-\dot{a}_{22}Y_0^2\}J. \tag{5.19}$$

The time derivative of $\rho_{\mathcal{S}}$ can be obtained by multiplying the operator on the right-hand side by an initial density matrix and integrating over the initial coordinates X_0, Y_0. Given the form of Eq. (5.19), one may expect that this procedure will yield an integro-differential (nonlocal in time) evolution operator for $\rho_{\mathcal{S}}$. However, the time dependence of the evolution operator disappears as a result of the two identities satisfied by the propagator:

$$Y_0J=\left(\frac{b_1}{b_3}Y+\frac{i}{b_3}\partial_X\right)J, \tag{5.20a}$$

$$X_0J=\left[-\frac{b_1}{b_2}X-\frac{i}{b_2}\partial_Y-i\left(\frac{2a_{11}}{b_2}+\frac{a_{12}b_1}{b_2b_3}\right)Y +\frac{a_{12}}{b_2b_3}\partial_X\right]J. \tag{5.20b}$$

After the appropriate substitutions, the resulting equation with renormalized Hamiltonian H_{ren} has the form

$$\dot{\rho}_{\mathcal{S}}(x,x',t)=-\frac{i}{\hbar}\langle x|[H_{ren}(t),\rho_{\mathcal{S}}]|x'\rangle-[\gamma(t)(x-x') \times(\partial_x+\partial_{x'})-D(t)(x-x')^2]\rho_{\mathcal{S}}(x,x',t) -if(t)(x-x')(\partial_x+\partial_{x'})\rho_{\mathcal{S}}(x,x',t). \tag{5.21}$$

The calculations leading to this master equation are nontrivial. They involve the use of relations between the coefficients b_k and a_{ij}. The final result leads to explicit formulae for these coefficients:

$$\tilde{\Omega}^2(t)/2=b_1\dot{b}_2/b_2-\dot{b}_1, \tag{5.22a}$$

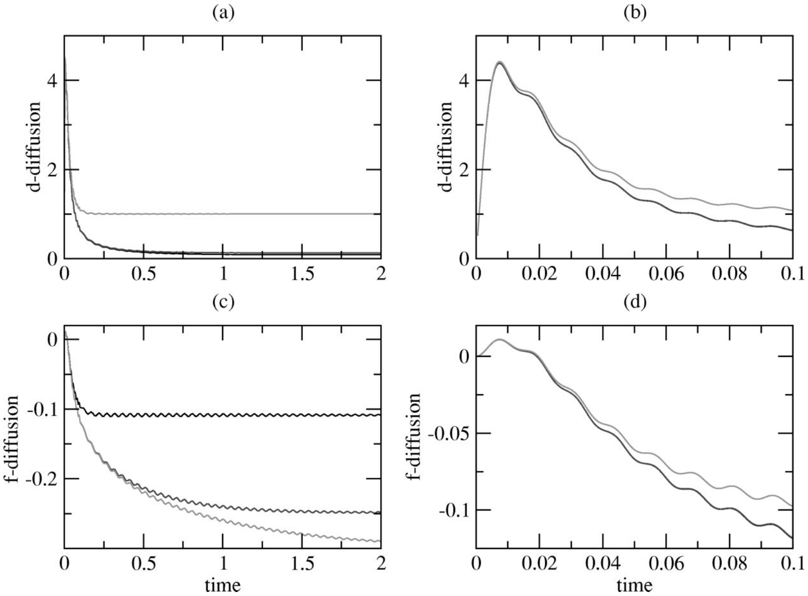


FIG. 5. Time-dependent coefficients of the perturbative master equation for quantum Brownian motion. The parameters used in these plots (where the time is measured in units of Ω^{-1}) are $\gamma/\Omega=0.05$, $\Gamma/\Omega=100$, $k_B T/\hbar\Omega=10$, 1, and 0.1. Plots on the right show the initial portion of the plots on the left—the initial transient—illustrating its independence of temperature (although higher temperatures produce higher final values of the coefficients). Plots on the right show that the final values of the coefficients strongly depend on temperature, and that anomalous diffusion is of importance only for very low temperatures.

$$\gamma(t)=-b_1-\dot{b}_2/2b_2, \tag{5.23a}$$

$$D(t)=\dot{a}_{11}-4a_{11}b_1+\dot{a}_{12}b_1/b_3 -\dot{b}_2(2a_{11}+a_{12}b_1/b_3)/b_2, \tag{5.24a}$$

$$2f(t)=\dot{a}_{12}/b_3-\dot{b}_2 a_{12}/(b_2 b_3)-4a_{11}. \tag{5.25a}$$

The fact that the exact master equation, Eq. (5.21), is local in time for an arbitrary spectrum of the environment is remarkable. This was demonstrated by Hu, Paz, and Zhang (1992) following discussions carried out under more restrictive assumptions by Caldeira and Leggett (1983); Haake and Reibold (1985); Grabert, Schramm, and Ingold (1988); and Unruh and Zurek (1989). It is the linearity of the problem that allows one to anticipate the (Gaussian) form of the propagator.

The above derivation of the exact master equation used the method of Paz (1994; see also Paz and Zurek, 2001). Explicit formulas for the time-dependent coefficients can be obtained when one focuses on the perturbative master equation. The formulas can be derived *ab initio* (see Paz and Zurek, 2001) but can also be obtained from the above results by finding a perturbative solution to Eq. (5.17) and then substituting it in Eqs. (5.22a)–(5.25a). The resulting master equation in the operator form is

$$\dot{\rho}_{\mathcal{S}}=-\frac{i}{\hbar}[H_{\mathcal{S}}+M\tilde{\Omega}(t)^2x^2/2,\rho_{\mathcal{S}}]-\frac{i\gamma(t)}{\hbar}[x,\{p,\rho_{\mathcal{S}}\}] -D(t)[x,[x,\rho_{\mathcal{S}}]]-\frac{f(t)}{\hbar}[x,[p,\rho_{\mathcal{S}}]]. \tag{5.26}$$

Coefficients such as the frequency renormalization $\tilde{\Omega}$, the relaxation coefficient $\gamma(t)$, and the normal and anomalous diffusion coefficients $D(t)$ and $f(t)$ are given by

$$\tilde{\Omega}^2(t)=-\frac{2}{M}\int_0^t ds\cos(\Omega s)\,\eta(s), \tag{5.22b}$$

$$\gamma(t)=\frac{2}{M\Omega}\int_0^t ds\sin(\Omega s)\,\eta(s), \tag{5.23b}$$

$$D(t)=\frac{1}{\hbar}\int_0^t ds\cos(\Omega s)\,\nu(s), \tag{5.24b}$$

$$f(t)=-\frac{1}{M\Omega}\int_0^t ds\sin(\Omega s)\,\eta(s). \tag{5.25b}$$

These coefficients can be made even more explicit when a convenient specific model is adopted for the spectral density:

$$C(\omega)=2M_{\gamma_0}\frac{\omega}{\pi}\frac{\Gamma^2}{\Gamma^2+\omega^2}. \tag{5.27}$$

Above, γ_0 characterizes the strength of the interaction, and Γ is the high-frequency cutoff. Then

$$\tilde{\Omega}^2=-\frac{2\gamma_0\Gamma^3}{\Gamma^2+\Omega^2}\left[1-\left(\cos\Omega t-\frac{\Omega}{\Gamma}\sin\Omega t\right)e^{-\Gamma t}\right]; \tag{5.22c}$$

$$\gamma(t)=\frac{\gamma_0\Gamma^2}{\Gamma^2+\Omega^2}\left[1-\left(\cos\Omega t-\frac{\Gamma}{\Omega}\sin\Omega t\right)e^{-\Gamma t}\right]. \tag{5.23c}$$

Note that both of these coefficients are initially zero. They grow to their asymptotic values on a time scale set by the inverse of the cutoff frequency Γ.

The two diffusion coefficients can also be studied, but it is more convenient to evaluate them numerically. In Fig. 5 we show their behavior. The normal diffusion coefficient quickly settles into its long-time asymptotic value:

$$D_\infty=M\gamma_0\Omega\hbar^{-1}\coth(\hbar\Omega\beta/2)\Gamma^2/(\Gamma^2+\Omega^2). \tag{5.28}$$

The anomalous diffusion coefficient $f(t)$ also approaches its asymptotic value. For high temperature it is suppressed by a cutoff Γ with respect to D_∞, but the approach to f_∞ is more gradual, algebraic rather than exponential. Environments with different spectral content exhibit different behavior (Hu, Paz, and Zhang, 1992; Paz, Habib, and Zurek, 1993; Paz, 1994; Anglin, Paz, and Zurek, 1997).

B. Decoherence in quantum Brownian motion

The coefficients of the master equation we have just derived can be computed under a variety of different assumptions. The two obvious characteristics of the environment that one can change are its temperature T and its spectral density $C(\omega)$. In the case of high temperatures, $D(t)$ tends to a temperature-dependent constant and dominates over $f(t)$. Indeed, in this case all of the coefficients settle to asymptotic values after an initial transient. Thus

$$\dot{\rho}_{\mathcal{S}}=-\frac{i}{\hbar}[H_{ren},\rho_{\mathcal{S}}]-\gamma(x-x')(\partial_x-\partial_{x'})\rho_{\mathcal{S}} -\frac{2M\gamma k_B T}{\hbar^2}(x-x')^2\rho_{\mathcal{S}}. \tag{5.29}$$

This master equation for $\rho(x,x')$ obtains in the unrealistic but convenient limit known as the *high-temperature approximation*, which is valid when $k_B T$ is much higher than all the other relevant energy scales, including the energy content of the initial state and the frequency cutoff in $C(\omega)$ (see Caldeira and Leggett, 1983). However, when these restrictive conditions hold, Eq. (5.29) can be written for an arbitrary $V(x)$. To see why, we give a derivation patterned on that of Hu, Paz, and Zhang (1993).

We start with the propagator, Eq. (5.5), $\rho_{\mathcal{S}}(x,x',t)=J(x,x',t|x_0,x_0',t_0)\rho_{\mathcal{S}}(x_0,x_0',t_0)$, which we shall treat as if it were an equation for a state vector of the two-dimensional system with coordinates x,x'. The propagator is then given by the high-temperature version of Eq. (5.10):

$$J(x,x',t|x_0,x_0',t_0)=\int DxDx'\exp\left(\frac{i}{\hbar}\{I_R(x)-I_R(x')\}\right)\times e^{-M\gamma(\int_0^t ds[x\dot{x}-x'\dot{x}'+x\dot{x}'-x'\dot{x}]+[2k_BT/\hbar^2][x-x']^2)}. \tag{5.30}$$

The term in the exponent can be interpreted as the effective Lagrangian of a two-dimensional system:

$$L_{\rm eff}(x,x')=M\dot{x}^2/2-V_R(x)-M\dot{x}'^2/2+V_R(x')+\gamma(x-x')(\dot{x}+\dot{x}')+i\frac{2M\gamma k_B T}{\hbar^2}(x-x')^2. \tag{5.31}$$

One can readily obtain the corresponding Hamiltonian,

$$H_{\rm eff}=\dot{x}\partial L_{\rm eff}/\partial\dot{x}+\dot{x}'\partial L_{\rm eff}/\partial\dot{x}'-L_{\rm eff}. \tag{5.32}$$

Conjugate momenta $p=p_x=M\dot{x}+\gamma(x-x')$ and $p'=p_{x'}=-M\dot{x}'+\gamma(x-x')$ are used to express the kinetic term of $H_{\rm eff}$. After evaluating $\dot{x}$ and $\dot{x}'$ in terms of p and p' in the expression for $H_{\rm eff}$ one obtains

$$H_{\rm eff}=[p-\gamma(x-x')]^2/2M-[p'-\gamma(x-x')]^2/2M+V(x)-V(x')-i2M\gamma k_BT(x'-x)^2/\hbar^2. \tag{5.33}$$

This expression yields the operator that generates the evolution of the density matrix, Eq. (5.29).

The coefficients of Eq. (5.21) approach their high-temperature values quickly (see Fig. 5). Already for T well below what the rigorous derivation would demand, the high-temperature limit appears to be an excellent approximation. The discrepancy is manifested by symptoms such as some of the diagonal terms of $\rho_{\mathcal{S}}(x',x)$ assuming negative values when the evolution starts from an initial state that is so sharply localized in position as to have kinetic energy in excess of the values allowed by the high-temperature approximation. However, this is limited to the initial instant of order $1/\Gamma$, and is known to be essentially unphysical for other reasons (Unruh and Zurek, 1989; Ambegoakar, 1991; Anglin, Paz, and Zurek, 1997; Romero and Paz, 1997). This short-time anomaly is closely tied to the fact that Eq. (5.33) (and, indeed, many of the exact or approximate master equations derived to date) does not have the Lindblad form (Kossakowski, 1973; Lindblad, 1976; see also Gorini, Kossakowski, and Sudarshan, 1976; Alicki and Lendi, 1987) of a dynamical semigroup.

The high-temperature master equation (5.29) is a good approximation in a wider range of circumstances than the one for which it was derived (Feynman and Vernon, 1963; Dekker, 1977; Caldeira and Leggett, 1983). Moreover, our key qualitative conclusion—rapid decoherence in the macroscopic limit—does not crucially depend on the approximations leading to Eq. (5.29). We shall therefore use it in our further studies.

1. Decoherence time scale

In the macroscopic limit [that is, when $\hbar$ is small compared to other quantities with dimensions of action, such as $\sqrt{2Mk_BT\langle(x-x')^2\rangle}$ in the last term] the high-temperature master equation is dominated by

$$\partial_t\rho_{\mathcal{S}}(x,x',t)=-\gamma\left\{\frac{(x-x')}{\lambda_T}\right\}^2\rho_{\mathcal{S}}(x,x',t). \tag{5.34}$$

Above,

$$\lambda_T=\frac{\hbar}{\sqrt{2Mk_BT}} \tag{5.35}$$

is the thermal de Broglie wavelength. Thus the density matrix loses off-diagonal terms in position representation:

$$\rho_{\mathcal{S}}(x,x',t)=\rho_{\mathcal{S}}(x,x',0)e^{-\gamma t(x-x'/\lambda_T)^2}, \tag{5.36}$$

while the diagonal $(x=x')$ remains untouched.

Quantum coherence decays exponentially at a rate given by the relaxation rate times the square of the distance, measured in units of thermal de Broglie wavelength (Zurek, 1984a). Position is the instantaneous pointer observable. If Eq. (5.36) was always valid, eigenstates of position would attain classical status.

The importance of position can be traced to the nature of the interaction Hamiltonian between the system and the environment. According to Eq. (5.3)

$$H_{\mathcal{SE}}=x\sum_n c_n q_n . \tag{5.37}$$

This form of $H_{\mathcal{SE}}$ is motivated by physics (Zurek, 1982, 1991). Interactions depend on the distance. However, had we endeavored to find a situation in which a different form of the interaction Hamiltonian—say, a momentum-dependent interaction—was justified, the form and consequently the predictions of the master equation would have been analogous to Eq. (5.36), but with a substitution of the relevant observable monitored by the environment for x. Such situations may be experimentally accessible (Poyatos, Cirac, and Zoller, 1996), providing a test of one of the key ideas of einselection: the relation between the form of interaction and the preferred basis.

The effect of the evolution, Eqs. (5.34)–(5.36), on the density matrix in the position representation is easy to envisage. Consider a superposition of two minimum-uncertainty Gaussians. Off-diagonal peaks represent coherence. They decay on a decoherence time scale τ_D, or with a decoherence rate (Zurek, 1984a, 1991)

$$\tau_D^{-1}=\gamma\left(\frac{x-x'}{\lambda_T}\right)^2. \tag{5.38}$$

The thermal de Broglie wavelength λ_T is microscopic for massive bodies and for the environment at reasonable temperatures. For a mass of 1 g at room temperature and for the separation $x'-x=1$ cm, Eq. (5.38) predicts a decoherence rate approximately 10^{40} times faster than relaxation. Even the cosmic microwave background suffices to cause rapid loss of quantum coherence in objects as small as dust grains (Joos and Zeh, 1985). These estimates for the rates of decoherence and relaxation should be taken with a grain of salt. Often the assumptions that have led to the simple high-temperature master equation, Eq. (5.29), are not valid (Gallis and Fleming, 1990; Gallis, 1992; Anglin, Paz, and Zurek, 1997). For example, the decoherence rate cannot be faster than the inverse of the spectral cutoff in Eq. (5.27), nor than the rate with which the superposition is created. Moreover, for large separations the quadratic dependence of the decoherence rate may saturate (Gallis and Fleming, 1990; Anglin, Paz, and Zurek, 1997), as seen in the simulated decoherence experiments of Cheng and Raymer (1999). Nevertheless, in the macroscopic domain decoherence of widely delocalized Schrödinger-cat states will occur very much faster than relaxation, which proceeds at the rate given by γ.

2. Phase-space view of decoherence

A useful alternative way of illustrating decoherence is afforded by the Wigner function representation

$$W(x,p)=\frac{1}{2\pi\hbar}\int_{-\infty}^{+\infty}dy\,e^{(ipy/\hbar)}\rho\left(x+\frac{y}{2},x-\frac{y}{2}\right). \tag{5.39}$$

The evolution equation followed by the Wigner function obtains through the Wigner transform of the corresponding master equation. In the high-temperature limit, Eq. (5.29) (valid for general potentials) yields

$$\partial_t W=\{H_{ren},W\}_{MB}+2\gamma\partial_p(pW)+D\,\partial_{pp}W. \tag{5.40}$$

The first term, the Moyal bracket, is the Wigner transform of the von Neumann equation (see Sec. III). In the linear case it reduces to the Poisson bracket. The second term is responsible for relaxation. The last diffusive term is responsible for decoherence.

Diffusion in momentum occurs at the rate set by $D=2M\gamma k_B T$. Its origin can be traced to the continuous measurement of the position of the system by the environment. In accord with Heisenberg indeterminacy, measurement of the position results in an increase of the uncertainty in the momentum (see Sec. IV).

Decoherence in phase space can be explained through the example of a superposition of two Gaussian wave packets. The Wigner function in this case is given by

$$W(x,p)=G(x+x_0,p)+G(x-x_0,p)+(\pi\hbar)^{-1}\times\exp(-p^2\xi^2/\hbar^2-x^2/\xi^2)\cos(\Delta x p/\hbar), \tag{5.41}$$

where

$$G(x\pm x_0,p-p_0)=\frac{e^{-(x\mp x_0)^2/\xi^2-(p-p_0)^2\xi^2/\hbar^2}}{\pi\hbar}. \tag{5.42}$$

We have assumed that the Gaussians are not moving ($p_0=0$).

The oscillatory term in Eq. (5.41) is the signature of superposition. The frequency of the oscillations is proportional to the distance between the peaks. When the separation is only in position x, this frequency is

$$f=\Delta x/\hbar=2x_0/\hbar. \tag{5.43}$$

Ridges and valleys of the interference pattern are parallel to the separation between the two peaks. This, and the fact that $\hbar$ appears in the interference term in W, is important for the phase-space derivation of the decoherence time. We focus on the dominant effect and direct our attention to the last term of Eq. (5.40). Its effect on a rapidly oscillating interference term will be very different from its effect on the two Gaussians. The interference term is dominated by the cosine:

$$W_{int}\sim\cos\left(\frac{\Delta x}{\hbar}p\right). \tag{5.44}$$

This is an eigenfunction of the diffusion operator. The decoherence time scale emerges (Zurek, 1991) from the corresponding eigenvalue

$$\dot{W}_{int}\approx-\{D\Delta x^2/\hbar^2\}\times W_{int}. \tag{5.45}$$

We have recovered the formula for τ_D, Eq. (5.38), from a different-looking argument. Equation (5.40) has no ex-

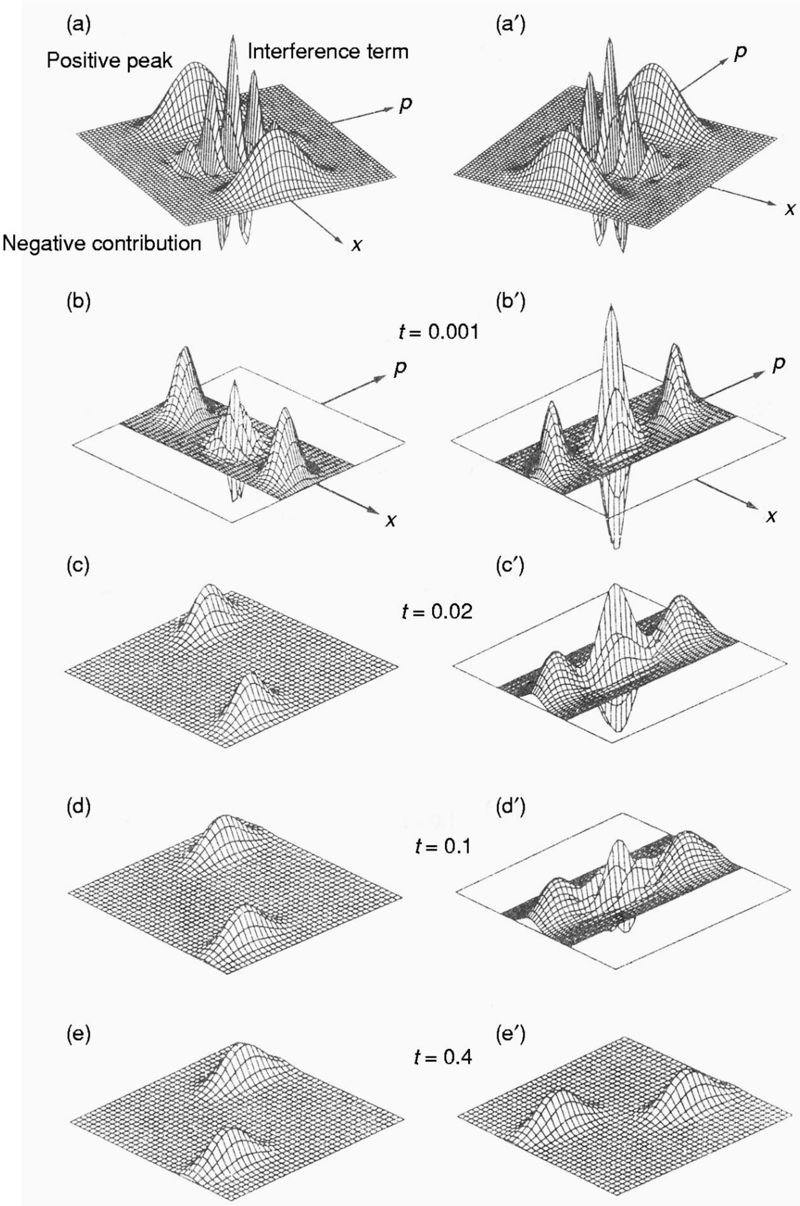


FIG. 6. Evolution of the Wigner function of a decohering harmonic oscillator. Note the difference between the rate at which the interference term disappears for the initial superposition of two minimal uncertainty Gaussians in position and in momentum.

plicit dependence on $\hbar$ for linear potentials (in the nonlinear case $\hbar$ enters through the Moyal bracket). Yet the decoherence time scale contains $\hbar$ explicitly. $\hbar$ enters through Eq. (5.43), that is, through its role in determining the frequency of the interference pattern W_{int}.

The evolution of a pure initial state of the type considered here is shown in Fig. 6. There we illustrate the evolution of the Wigner function for two initial pure states: superposition of two positions and superposition of two momenta. There is a noticeable difference between these two cases in the rate at which the interference term disappears. This was anticipated. The interaction in Eq. (5.3) is a function of x. Therefore x is monitored by the environment directly, and the superposition of positions decoheres almost instantly. By contrast, the superposition of momenta is initially insensitive to monitoring by the environment—the corresponding initial state is already well localized in the observable singled out by the interaction. However, a superposition of momenta leads to a superposition of positions, and hence to decoherence, albeit on a dynamical (rather than τ_D) time scale.

An intriguing example of a long-lived superposition of two seemingly distant Gaussians was pointed out by Braun, Braun, and Haake (2000) in the context of superradiance. As they note, the relevant decohering interaction cannot distinguish between some such superpositions, leading to a Schrödinger-cat pointer subspace.

C. Predictability sieve in phase space

Decoherence rapidly destroys nonlocal superpositions. Obviously, states that survive must be localized. However, they cannot be localized to a point in x, since this would imply—by Heisenberg's indeterminacy—an infinite range of momenta and hence of velocities. As a result, a wave function localized too well at one instant would become very nonlocal a moment later.

Einselected pointer states minimize the damage done by decoherence over the time scale of interest (usually associated with predictability or with dynamics). They can be found through the application of a predictability sieve as outlined at the end of Sec. IV. To implement it, we compute entropy increase or purity loss for all initially pure states in the Hilbert space of the system under the cumulative evolution caused by the self-Hamiltonian and by the interaction with the environment. It would be a tall order to carry out the requisite calculations for an arbitrary quantum system interacting with a general environment. We focus on an exactly solvable case.

In the high-temperature limit the master equations (5.26) and (5.29) can be expressed in the operator form

$$\dot{\rho}=\frac{1}{i\hbar}[H_{ren},\rho]+\frac{\gamma}{i\hbar}[\{p,x\},\rho]-\frac{\eta k_B T}{\hbar^2}[x,[x,\rho]] -\frac{i\gamma}{\hbar}([x,\rho p]-[p,\rho x]). \tag{5.46}$$

Above, $\eta=2M\gamma$ is the viscosity. Only the last two terms can change entropy. Terms of the form

$$\dot{\rho}=[\hat{O},\rho], \tag{5.47}$$

where $\hat{O}$ is the Hermitian, leave the purity $\varsigma=\mathrm{Tr}\rho^2$ and the von Neumann entropy $H=-\mathrm{Tr}\rho\ln\rho$ unaffected. This follows from the cyclic property of the trace:

$$\frac{d}{dt}\mathrm{Tr}\rho^N=\sum_{k=1}^{N}(\mathrm{Tr}\rho^{k-1}[\hat{O},\rho]\rho^{N-k})=0. \tag{5.48}$$

Constancy of $\mathrm{Tr}\rho^2$ is obvious, while for $\mathrm{Tr}\rho\ln\rho$ it follows when the logarithm is expanded in powers of ρ.

Equation (5.46) leads to the loss of purity at the rate (Zurek, 1993a)

$$\frac{d}{dt}\mathrm{Tr}\rho^2=-\frac{4\eta k_B T}{\hbar^2}\mathrm{Tr}[\rho^2x^2-(\rho x)^2]+2\gamma\mathrm{Tr}\rho^2. \tag{5.49}$$

The second term increases purity—or decreases entropy—as the system is damped from an initial highly mixed state. For the predictability sieve this term is usually unimportant, since for a vast majority of initially pure states its effect will be negligible when compared to the first decoherence-related term. Thus, in the case of pure initial states,

$$\frac{d}{dt}\mathrm{Tr}\rho^2=-\frac{4\eta k_B T}{\hbar^2}(\langle x^2\rangle-\langle x\rangle^2). \tag{5.50}$$

Therefore the instantaneous loss of purity is minimized for perfectly localized states (Zurek, 1993a). The second term of Eq. (5.49) allows for equilibrium. Nevertheless, early on, and for very localized states, its presence causes an (unphysical) increase of purity to above unity. This is a well-known artifact of the high-temperature approximation [see discussion following Eq. (5.33)].

To find the most predictable states relevant for dynamics, we consider the increase in entropy over an oscillation period. For a harmonic oscillator with mass M and frequency Ω, one can compute the purity loss averaged over $\tau=2\pi/\Omega$:

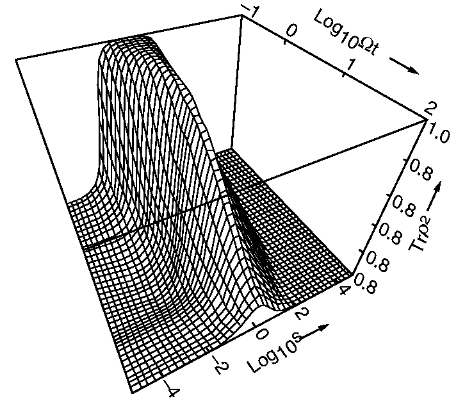


FIG. 7. Predictability sieve in action. The plot shows purity $\mathrm{Tr}\rho^2$ for mixtures that have evolved from initial minimum-uncertainty wave packets with different squeeze parameters s in an underdamped harmonic oscillator with $\gamma/\omega=10^{-4}$. Coherent states, which have the same spread in position as in momentum, $s=1$, are clearly most predictable.

$$\Delta\varsigma|_0^{2\pi/\Omega}=-2D[\Delta x^2+\Delta p^2/(M\Omega)^2]. \tag{5.51}$$

Above, Δx and Δp are dispersions of the state at the initial time. By Heisenberg indeterminacy, $\Delta x\Delta p\geqslant\hbar/2$. The loss of purity will be smallest when

$$\Delta x^2=\hbar/2M\Omega,\quad \Delta p^2=\hbar M\Omega/2. \tag{5.52}$$

Coherent quantum states are selected by the predictability sieve in an underdamped harmonic oscillator (Zurek, 1993a; Zurek, Habib, and Paz, 1993; Tegmark and Shapiro, 1994; Gallis, 1996; Wiseman and Vaccaro, 1998; Paraoanu, 1999, 2002). Rotation induced by the self-Hamiltonian turns preference for states localized in position into preference for localization in phase space. This is illustrated in Fig. 7.

We conclude that for an underdamped harmonic oscillator coherent Gaussians are the best quantum theory has to offer as an approximation to a classical point. Similar localization in phase space should be obtained in the reversible classical limit in which the familiar symptoms of the openness of the system, such as the finite relaxation rate $\gamma=\eta/2M$, become vanishingly small. This limit can be attained for large mass $M\rightarrow\infty$, while the viscosity η remains fixed and sufficiently large to assure localization (Zurek, 1991, 1993a). This is, of course, not the only possible situation. Haake and Walls (1987) discussed the overdamped case, in which pointer states are still localized, but become relatively narrower in position. On the other hand, an "adiabatic" environment with high-frequency cutoff large compared to the level spacing in the system enforces einselection in energy eigenstates (Paz and Zurek, 1999).

D. Classical limit in phase space

There are three strategies that allow one to simultaneously recover the classical phase-space structure and the classical equations of motion from quantum dynamics and decoherence.

1. Mathematical approach ($\hbar \to 0$)

This mathematical classical limit could not be implemented without decoherence, since the oscillatory terms associated with interference do not have an analytic $\hbar \to 0$ limit (see, for example, Peres, 1993). However, in the presence of the environment, the relevant terms in the master equations increase as $\mathcal{O}(\hbar^{-2})$ and make the nonanalytic manifestations of interference disappear. Thus phase-space distributions can always be represented by localized coherent state points, or by distributions over the basis consisting of such points.

This strategy is easiest to implement starting from the phase-space formulation. It follows from Eq. (5.45) that the interference term in Eq. (5.41) will decay (Paz, Habib, and Zurek, 1993) over the time interval Δt as

$$W_{int} \sim \exp\left(-\Delta t \frac{D \Delta x^2}{\hbar^2} \right) \cos\left(\frac{\Delta x}{\hbar} p \right). \tag{5.53}$$

As long as Δt is large compared to the decoherence time scale $\tau_D \simeq \hbar^2 / D \Delta x^2$, oscillatory contributions to the Wigner function $W(x,p)$ should disappear as $\hbar \to 0$. Simultaneously, Gaussians representing likely locations of the system become narrower, approaching Dirac δ functions in phase space. For instance, in Eq. (5.42),

$$\lim_{\hbar \to 0} G(x - x_0, p - p_0) = \delta(x - x_0, p - p_0), \tag{5.54}$$

providing that half-widths of the coherent states in x and p decrease to zero as $\hbar \to 0$. This would be assured when, for instance, in Eqs. (5.41) and (5.42),

$$\xi^2 \sim \hbar. \tag{5.55}$$

Thus individual coherent-state Gaussians approach phase-space points. This behavior indicates that in a macroscopic open system nothing but probability distributions over localized phase-space points can survive in the $\hbar \to 0$ limit for any time of dynamical or predictive significance. [Coherence between immediately adjacent points separated only by $\sim \xi$, Eq. (5.55), can last longer. This is no threat to the classical limit. Small-scale coherence is a part of a quantum halo of the classical pointer states (Anglin and Zurek, 1996).]

The mathematical classical limit implemented by letting $\hbar \to 0$ becomes possible in the presence of decoherence. It is tempting to carry this strategy to its logical conclusion and represent every probability density in phase space in the point(er) basis of narrowing coherent states. Such a program is beyond the scope of this review, but the reader should by now be convinced that it is possible. Indeed, Perelomov (1986) showed that a general quantum state could be represented in a sparse basis of coherent states that occupy the sites of a regular lattice, providing that the volume per coherent state point was no more than $(2\pi\hbar)^d$ in the d-dimensional configuration space. In the presence of decoherence arising from a coordinate-dependent interaction, evolution of a general quantum superposition should be, after a few decoherence times, well approximated by a probability distribution over such Gaussian points.

2. Physical approach: The macroscopic limit

The possibility of the $\hbar \to 0$ classical limit in the presence of decoherence is of interest. But $\hbar = 1.054\,59 \times 10^{-27}$ erg s. Therefore a physically more reasonable approach increases the size of the object, and, hence, its susceptibility to decoherence. This strategy can be implemented starting with Eq. (5.40). Reversible dynamics obtains as $\gamma \to 0$ while $D = 2M\gamma k_B T = \eta k_B T$ increases.

The decrease of γ and the simultaneous increase of $\eta k_B T$ can be anticipated with the increase of the size and mass. Assume that the density of the object is independent of its size R, and that the environment quanta scatter from its surface (as would photons or air molecules). Then $M \sim R^3$ and $\eta \sim R^2$. Hence

$$\eta \sim \mathcal{O}(R^2) \to \infty, \tag{5.56}$$

$$\gamma = \eta / 2M \sim \mathcal{O}(1/R) \to 0, \tag{5.57}$$

as $R \to \infty$. Localization in phase space and reversibility can be simultaneously achieved in a macroscopic limit.

The existence of a macroscopic classical limit in simple cases was pointed out some time ago (Zurek, 1984a, 1991; Gell-Mann and Hartle, 1993). We shall analyze it in the next section in a more complicated chaotic setting, where reversibility can no longer be taken for granted. In the harmonic-oscillator case, approximate reversibility is effectively guaranteed, since the action associated with the 1-σ contour of the Gaussian state increases with time at the rate (Zurek, Habib, and Paz, 1993)

$$\dot{I} = \gamma \frac{k_B T}{\hbar \Omega}. \tag{5.58}$$

Action I is a measure of the lack of information about phase-space location. Hence its rate of increase is a measure of the rate of predictability loss. The trajectory is a limit of the "tube" swept through phase space by the moving contour representing the instantaneous uncertainty of the observer about the state of the system. Evolution is approximately deterministic when the area of this contour is nearly constant. In accord with Eqs. (5.56) and (5.57) $\dot{I}$ tends to zero in the reversible macroscopic limit:

$$\dot{I} \sim \mathcal{O}(1/R). \tag{5.59}$$

The existence of an approximately reversible trajectory-like thin tube provides an assurance that, having localized the system within a regular phase-space volume at $t = 0$, we can expect to find it later inside the Liouville-transported contour of nearly the same measure. Similar conclusions follow for integrable systems.

3. Ignorance inspires confidence in classicality

Dynamical reversibility can be achieved with einselection in the macroscopic limit. Moreover, $\dot{I}/I$ or other measures of predictability loss decrease with the increase of I. This is especially dramatic when quantified in terms of the von Neumann entropy, that, for Gaussian states, increases at the rate (Zurek, Habib, and Paz, 1993)

$$\dot{H}=\dot{I}\ln\frac{I+1}{I-1}. \quad (5.60a)$$

The resulting $\dot{H}$ is infinite for pure coherent states ($I=1$), but quickly decreases with increasing I. Similarly, the rate of purity loss for Gaussians is

$$\dot{\varsigma}=\dot{I}/I^2. \quad (5.60b)$$

Again, it tapers off for more mixed states.

This behavior is reassuring. It leads us to conclude that irreversibility quantified through, say, von Neumann entropy production, Eq. (5.60a), will approach $\dot{H}\approx 2\dot{I}/I$, vanishing in the limit of large I. When, in the spirit of the macroscopic limit, we do not insist on the maximal resolution allowed by quantum indeterminacy, the subsequent predictability losses measured by the increase of entropy or through the loss of purity will diminish. Illusions of reversibility, determinism, and exact classical predictability become easier to maintain in the presence of ignorance about the initial state!

To think about phase-space points one may not even need to invoke a specific quantum state. Rather, a point can be regarded as the limit of an abstract recursive procedure in which the phase-space coordinates of the system are determined better and better in a succession of increasingly accurate measurements. One may be tempted to extrapolate this limiting process *ad infinitessimum*, which would lead beyond Heisenberg's indeterminacy principle and to a false conclusion that idealized points and trajectories exist objectively, and that the insider view of Sec. II can always be justified. While in our quantum universe this conclusion is wrong, and the extrapolation described above illegal, the presence, within Hilbert space, of localized wave packets near the minimum-uncertainty end of such imagined sequences of measurements is reassuring. Ultimately, the ability to represent motion in terms of points and their time-ordered sequences (trajectories) is the essence of classical mechanics.

E. Decoherence, chaos, and the second law

The breakdown of correspondence in this chaotic setting was described in Sec. III. It is anticipated to occur in all nonlinear systems, since stretching of the wave packet by the dynamics is a generic feature, absent only from a harmonic oscillator. However, the exponential instability of chaotic dynamics implies rapid loss of quantum-classical correspondence after the Ehrenfest time, $t_\hbar=\Lambda^{-1}\ln \chi\Delta p/\hbar$. Here Λ is the Lyapunov exponent, while $\chi=\sqrt{V_x/V_{xxx}}$ characterizes the dominant scale of nonlinearities in the potential $V(x)$, and Δp gives the coherence scale in the initial wave packet. The above estimate, Eq. (3.5), depends on the initial conditions. It is smaller than, but typically close to, $t_r=\Lambda^{-1}\ln I/\hbar$, Eq. (3.6), where I is the characteristic action of the system. By contrast, phase-space patches of regular systems undergo stretching with a power of time. Consequently, loss of correspondence occurs only over a much longer $t_r\sim(I/\hbar)^\alpha$, which depends polynomially on $\hbar$.

1. Restoration of correspondence

Exponential instability spreads the wave packet to a paradoxical extent at the rate given by the positive Lyapunov exponents $\Lambda_+^{(i)}$. Einselection attempts to enforce localization in phase space by tapering off interference terms at a rate given by the inverse of the decoherence time scale, $\tau_D=\gamma^{-1}(\lambda_T/\Delta x)^2$. The two processes reach *status quo* when the coherence length ℓ_c of the wave packet makes their rates comparable, that is,

$$\tau_D\Lambda_+\simeq 1. \quad (5.61)$$

This yields an equation for the steady-state coherence length and for the corresponding momentum dispersion:

$$\ell_c\simeq\lambda_T\sqrt{\Lambda_+/2\gamma}; \quad (5.62)$$

$$\sigma_c=\hbar/\ell_c=\sqrt{2D/\Lambda_+}. \quad (5.63)$$

Above, we have quoted results (Zurek and Paz, 1994) that follow from a more rigorous derivation of the coherence length ℓ_c than the rough and ready approach that led to Eq. (5.61). They embody the same physical argument, but seek asymptotic behavior of the Wigner function that evolves according to the equation

$$\dot{W}=\{H,W\}+\sum_{n\geqslant 1}\frac{\hbar^{2n}(-)^n}{2^{2n}(2n+1)!}\partial_x^{2n+1}V\partial_p^{2n+1}W + D\partial_p^2 W. \quad (5.64)$$

The classical Liouville evolution generated by the Poisson bracket ceases to be a good approximation of the decohering quantum evolution when the leading quantum correction becomes comparable to the classical force:

$$\frac{\hbar^2}{24}V_{xxx}W_{ppp}\approx\frac{\hbar^2}{24}\frac{V_x}{\chi^2}\frac{W_p}{\sigma_c^2}. \quad (5.65)$$

The term V_xW_p represents the classical force in the Poisson bracket. Quantum corrections are small when

$$\sigma_c\chi\gg\hbar. \quad (5.66)$$

Equivalently, the Moyal bracket generates approximately Liouville flow when the coherence length satisfies

$$\ell_c\ll\chi. \quad (5.67)$$

This last inequality has an obvious interpretation: it is a condition for localization to within a region ℓ_c that is

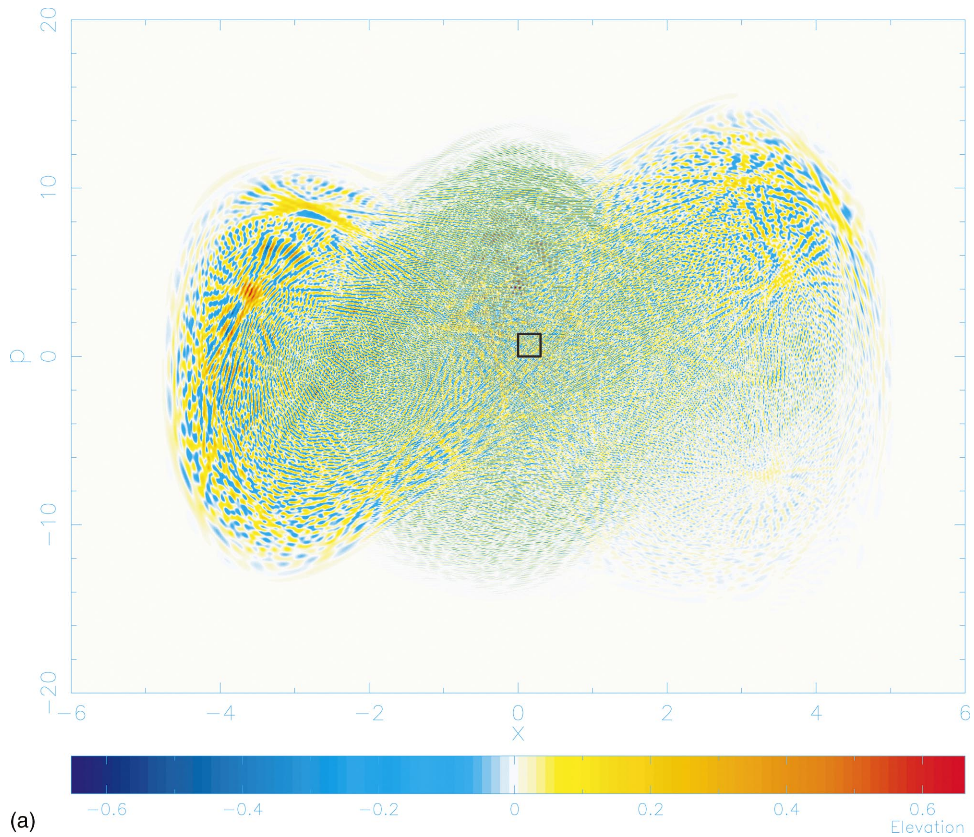


FIG. 8. Snapshots of a chaotic system with a double-well potential: $H=p^2/2m+Ax^4-Bx^2+Cx\cos(ft)$. In the example discussed here $m=1$, $A=0.5$, $B=10$, $f=6.07$ and $C=10$, yielding the Lyapunov exponent $\Lambda\approx 0.45$ (see Habib, Shizume, and Zurek, 1998). All figures were obtained after approximately eight periods of the driving force. The evolution started from the same minimum-uncertainty Gaussian and proceeded according to (a) the quantum Moyal bracket, (b) the Poisson bracket, and (c) the Moyal bracket with decoherence [constant $D=0.025$ in Eq. (5.64)]. In the quantum cases $\hbar=0.1$, which corresponds to the area of the rectangle in the image of the Wigner function above. Interference fringes are clearly visible in (a), and the Wigner function shown there is only vaguely reminiscent of the classical probability distribution in (b). Even modest decoherence [$D=0.25$ used to get (c) corresponds to coherence length $\ell_c=0.3$] dramatically improves the correspondence between the quantum and the classical. The remaining interference fringes appear on relatively large scales, which implies small-scale quantum coherence (Color).

small compared to the scale χ of the nonlinearities of the potential. When this condition holds, classical force will dominate over quantum corrections.

Restoration of correspondence is illustrated in Fig. 8 where Wigner functions are compared with classical probability distributions in a chaotic system. The difference between the classical and quantum expectation values in the same chaotic system is shown in Fig. 9. Even relatively weak decoherence suppresses the discrepancy, helping reestablish the correspondence: $D=0.025$ translates through Eq. (5.62) into coherence over $\ell_c\simeq 0.3$, not much smaller than the nonlinearity scale $\chi\simeq 1$ for the investigated Hamiltonian of Fig. 8.

2. Entropy production

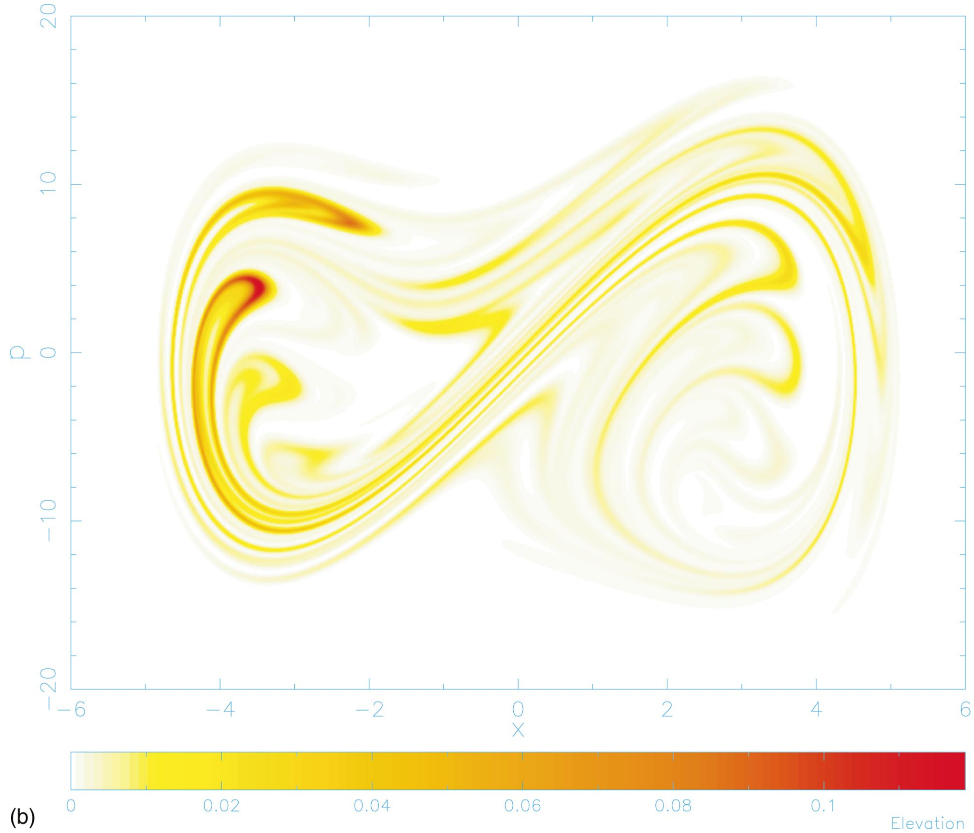
Irreversibility is the price for the restoration of quantum-classical correspondence in chaotic dynamics. It can be quantified through the entropy production rate. The simplest argument recognizes that decoherence restricts spatial coherence to ℓ_c. Consequently, as the exponential instability stretches the size $L^{(i)}$ of the distribution in directions corresponding to the positive Lyapunov exponents $\Lambda_+^{(i)}$, with $L^{(i)}\sim\exp[\Lambda_+^{(i)}t]$, the squeezing mandated by the Liouville theorem in the complementary directions corresponding to $\Lambda_-^{(i)}$ will halt at $\sigma_c^{(i)}$, Eq. (5.63). In this limit, the number of pure states needed to represent the resulting mixture increases exponentially:

$$N^{(i)}\simeq L^{(i)}/\ell_c^{(i)} \tag{5.68}$$

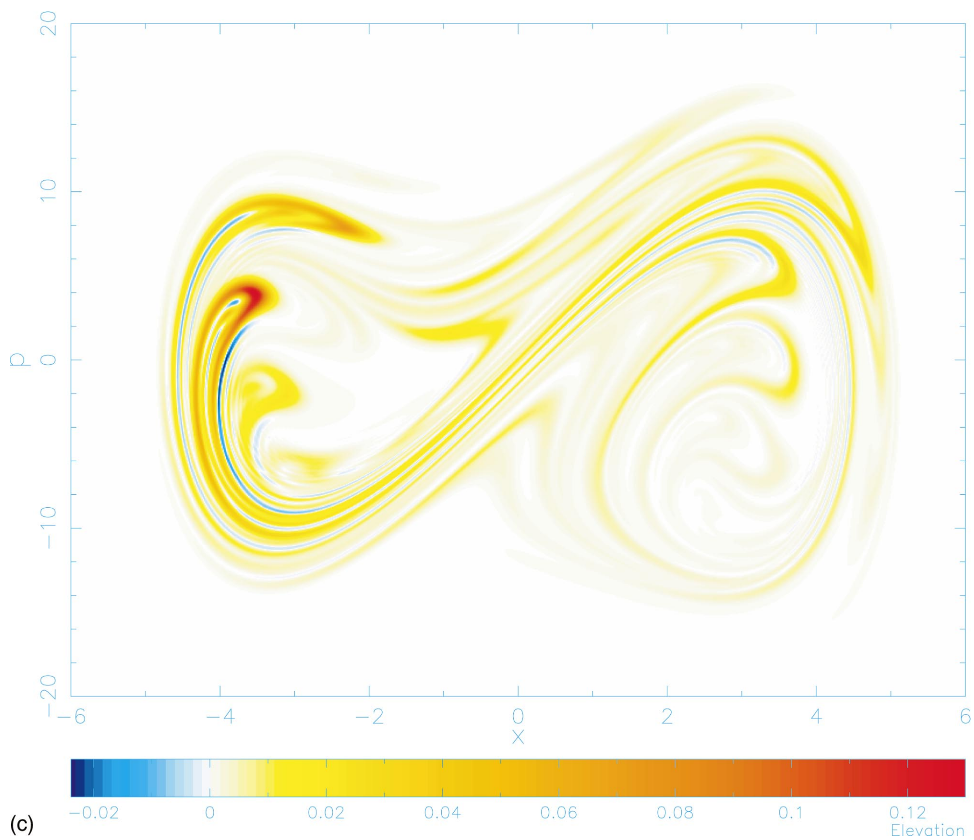
in each dimension. The least number of pure states overlapped by W will then be $\mathcal{N}=\Pi_i N^{(i)}$. This implies

$$\dot{H}\simeq\partial_t\ln\mathcal{N}\simeq\sum_i\Lambda_+^{(i)}. \tag{5.69}$$

This estimate for the entropy production rate becomes accurate as the width of the Wigner function reaches saturation at $\sigma_c^{(i)}$. When a patch in phase space corresponding to the initial W is regular and smooth on scales large compared to $\sigma_c^{(i)}$, evolution will start nearly reversibly (Zurek and Paz, 1994). However, as squeezing brings the extent of the effective support of W close to



(b)



(c)

FIG. 8. (*Continued*.)

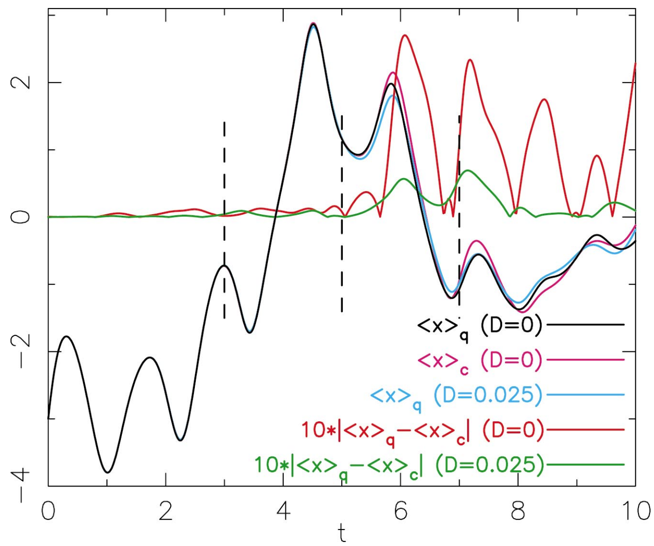


FIG. 9. Classical and quantum expectation values of position $\langle x \rangle$ as a function of time for an example of Fig. 8. Evolution started from a minimum-uncertainty Gaussian. Noticeable discrepancies between the quantum and classical averages appear on a time scale consistent with the Ehrenfest time $t_\hbar$. Decoherence, even in modest doses, dramatically decreases differences between the expectation values (Color).

$\sigma_c^{(i)}$, diffusion bounds from below the size of the smallest features of W. Stretching in the unstable directions continues unabated. As a consequence, the volume of the support of W will grow exponentially, resulting in an entropy production rate set by Eq. (5.69), that is, by the sum of the classical Lyapunov exponents. Yet, it has an obviously quantum origin in decoherence. This quantum origin may be apparent initially, since the rate of Eq. (5.69) will be approached from above when the initial state is nonlocal. On the other hand, in a multidimensional system different Lyapunov exponents may begin to contribute to entropy production at different instants [since the saturation condition, Eq. (5.61), may not be met simultaneously for all $\Lambda_+^{(i)}$]. Then the entropy production rate can accelerate, before subsiding as a consequence of approaching equilibrium.

The time scales on which this estimate of entropy production applies are still subject to investigation (Zurek and Paz, 1995a; Zurek, 1998b; Monteoliva and Paz, 2000) and even controversy (Casati and Chirikov, 1995b; Zurek and Paz, 1995b). The instant when Eq. (5.69) becomes a good approximation corresponds to the moment when the exponentially unstable evolution forces the Wigner function to develop small phase-space structures on the scale of the effective decoherence-imposed coarse graining, Eq. (5.63). Equation (5.69) will be a good approximation until the time t_{EQ} at which equilibrium sets in. Both $t_\hbar$ and t_{EQ} have a logarithmic dependence on the corresponding (initial and equilibrium) phase-space volumes I_0 and I_{EQ}, so the validity of Eq. (5.69) will be limited to $t_{EQ} - t_\hbar \simeq \Lambda^{-1} \ln I_{EQ}/I_0$.

There is a simple and conceptually appealing way to extend the interval over which entropy is produced at the rate given by Eq. (5.69). Imagine an observer monitoring a decohering chaotic system, recording its state at time intervals small compared to Λ^{-1}, but large compared to the decoherence time scale. One can show (Zurek, 1998b) that the average increase in the size of the algorithmically compressed records of measurement of a decohering chaotic system (that is, the algorithmic randomness of the acquired data; see, for example, Cover and Thomas, 1991) is given by Eq. (5.69). This conclusion holds, providing that the effect of the collapses of the wave packet caused by the repeated measurements is negligible, i.e., the observer is "skillful." A possible strategy that the skillful observer may adopt is that of indirect measurements, of monitoring a fraction of the environment responsible for decoherence to determine the state of the system. As we shall see in more detail in the following sections of the paper, this is a very natural strategy, often employed by observers.

A classical analog of Eq. (5.69) was obtained by Kolmogorov (1960) and Sinai (1960) starting from very different, mathematical arguments that in effect relied on an arbitrary but fixed coarse graining imposed on phase space (see Wehrl, 1978). Decoherence leads to a similar-looking *quantum* result in a very different fashion: Effective "coarse graining" is imposed by coupling to the environment, but only in the sense implied by einselection. Its graininess (resolution) is set by the accuracy of the monitoring environment. This is especially obvious when the indirect monitoring strategy mentioned immediately above is adopted by the observers. Preferred states will be partly localized in x and p, but (in contrast to the harmonic-oscillator case with its coherent states) details of this environment-imposed coarse graining will likely depend on phase-space location, the precise nature of the coupling to the environment, etc. Yet, in the appropriate limit, Eqs. (5.66) and (5.67), the asymptotic entropy production rate defined with the help of the algorithmic contribution discussed above [i.e., in the manner of *physical entropy*, that is, the sum of the measure of ignorance given by the von Newmann entropy and the algorithmic randomness of the records, Zurek (1989)] does not depend on the strength or nature of the coupling, but is instead given by the sum of the positive Lyapunov exponents.

von Neumann entropy production consistent with the above discussion has now been seen in numerical studies of decohering chaotic systems (Shiokawa and Hu, 1995; Furuya, Nemes, and Pellegrino, 1998; Schack, 1998; Miller and Sarkar, 1999; Monteoliva and Paz, 2000). Extensions to situations in which relaxation matters, as well as in the opposite direction to where decoherence is relatively gentle have also been discussed (Brun, Percival, and Schack, 1996; Miller, Sarkar, and Zarum, 1998; Pattanayak, 2000). A related development is the experimental study of the Loschmidt echo using NMR techniques (Levstein, Usaj, and Pastawski, 1998; Levstein *et al.*, 2000; Jalabert and Pastawski, 2001), which sheds new light on the irreversibility in decohering complex dynamical systems. We shall return briefly to this subject in Sec. VIII.

3. Quantum predictability horizon

The cross section I of the trajectorylike tube containing the state of the harmonic oscillator in phase space increases only slowly, Eq. (5.58), at a rate which—once the limiting Gaussian is reached—does not depend on I. By contrast, in chaotic quantum systems this rate is

$$\dot{I} \simeq I \sum_i \Lambda_+^{(i)}. \tag{5.70}$$

A fixed rate of entropy production implies an exponential increase of the cross section of the tube of, say, the 1-σ contour containing points consistent with the initial conditions. Phase-space support expands exponentially.

This quantum view of chaotic evolution can be compared with the classical deterministic chaos. In both cases, in the appropriate classical limit, which may involve either mathematical $\hbar \to 0$, or a macroscopic limit, the future state of the system can, in principle, be predicted to a set accuracy for an arbitrarily long time. However, such predictability can be accomplished only when the initial conditions are given with the resolution that increases exponentially with the time interval over which the predictions are to be valid. Given the fixed value of $\hbar$, there is therefore a *quantum predictability horizon* after which the Wigner function of the system starting from an initial minimum-uncertainty Gaussian becomes stretched to a size of the order of the characteristic dimensions of the system (Zurek, 1998b). The ability to predict the location of the system in phase space is then lost after $t \sim t_\hbar$, Eq. (3.5), regardless of whether evolution is generated by the Poisson or Moyal bracket or, indeed, whether the system is closed or open.

The case of regular systems is closer to that of a harmonic oscillator. The rate at which the cross section of the phase-space trajectory tube increases, consistent with the initial patch in phase space, will asymptote to $\dot{I} \simeq$ const.

$$\dot{H} = \dot{I}/I \sim 1/t. \tag{5.71}$$

Thus initial conditions allow one to predict the future of a regular system for time intervals that are exponentially longer than those in the chaotic case. The rate of entropy production of an open quantum system is therefore a very good indicator of its dynamics, as was conjectured some time ago (Zurek and Paz, 1995a), and as seems borne out in the numerical simulations (Shiokawa and Hu, 1995; Miller, Sarkar, and Zarum, 1998; Miller and Sarkar, 1999; Monteoliva and Paz, 2000).

VI. EINSELECTION AND MEASUREMENTS

It is often said that quantum states play only an epistemological role, describing the observer's knowledge about past measurement outcomes that have prepared the system (Jammer, 1974; d'Espagnat, 1976, 1995; Fuchs and Peres, 2000). In particular—and this is a key argument against their objective existence (against their ontological status)—it is impossible to determine what the state of an isolated quantum system is without prior information about the observables used to prepare it. Measurements of observables that do not commute with this original set will inevitably create a different state.

The continuous monitoring of the einselected observables by the environment allows pointer states to exist in much the same way as do classical states. This ontological role of the einselected quantum states can be justified operationally, by showing that in the presence of einselection one can find out what the quantum state is, without inevitably re-preparing it by the measurement. Thus einselected quantum states are no longer just epistemological. In a system monitored by the environment, *what is* the einselected states coincides with *what is known to be*—what is recorded by the environment (Zurek, 1993a, 1993b, 1998a).

The conflict between the quantum and the classical was originally noted and discussed almost exclusively in the context of quantum measurements.[5] Here I shall consider measurements, and, more to the point, acquisition of information in quantum theory from the point of view of decoherence and einselection.

A. Objective existence of einselected states

To demonstrate the objective existence of einselected states we now develop an operational definition of existence and show how, in an open system, one can find out what the state *was* and *is*, rather than just prepare it. This point has been made before (Zurek, 1993a, 1998a), but this is the first time I shall discuss it in more detail.

The objective existence of states can be defined operationally by considering two observers. The first of them is the *record keeper* **R**. He prepares the states with the original measurement and will use his records to determine if they were disturbed by measurements carried out by other observers, e.g., the *spy* **S**. The goal of **S** is to discover the state of the system without perturbing it. When an observer can consistently determine the state of a system without changing it, that state, by our operational definition, will be said to *exist objectively*.

In the absence of einselection the situation of **S** is hopeless: **R** prepares states by measuring sets of commuting observables. Unless **S** picks, by sheer luck, the same observables in the case of each state, his measurements will re-prepare the states of the system. Thus, when **R** remeasures using the original observables, he will likely find answers different from his records of preparatory measurements. The spy **S** will "get caught" because it is impossible to find out an initially unknown state of an isolated quantum system.

[5]See, for example, Bohr, 1928; Mott, 1929; von Neumann, 1932; Dirac, 1947; Zeh, 1971, 1973, 1993; d'Espagnat, 1976, 1995; Zurek, 1981, 1982, 1983, 1991, 1993a 1993b, 1998a; Omnès, 1992, 1994; Elby, 1993, 1998; Donald, 1995; Butterfield, 1996; Giulini *et al.*, 1996; Bub, 1997; Bacciagaluppi and Hemmo, 1998; Healey, 1998; Healey and Hellman, 1998; Saunders, 1998.

In the presence of environmental monitoring the nature of the game between **R** and **S** is dramatically altered. Now it is no longer possible for **R** to prepare an arbitrary pure state that will persist or predictably evolve without losing purity. Only einselected states that are already monitored by the environment, that are selected by the predictability sieve, will survive. By the same token, **S** is no longer clueless about the observables he can measure without leaving incriminating evidence. For example, he can independently prepare and test the survival of various states in the same environment to establish which states are einselected, and then measure appropriate pointer observables. Better yet, **S** can forgo direct measurements of the system and gather information indirectly, by monitoring the environment.

This last strategy may seem contrived, but indirect measurements—acquisition of information about the system by examining fragments of the environment that have interacted with it—is in fact more or less the only strategy employed by observers. Our eyes, for example, intercept only a small fraction of the photons that scatter from various objects. The rest of the photons constitute the environment, which retains at least as complete a record of the same einselected observables as we can obtain (Zurek, 1993a, 1998a).

The environment $\mathcal{E}$ acts as a persistent observer, dominating the game with frequent questions, always about the same observables, compelling both **R** and **S** to focus on the einselected states. Moreover, $\mathcal{E}$ can be persuaded to share its records of the system. This accessibility of the einselected states is not a violation of the basic tenets of quantum physics. Rather, it is a consequence of the fact that the data required to turn a quantum state into an ontological entity, an einselected pointer state, are abundantly supplied by the environment.

We emphasize the operational nature of this criterion for existence. There may, in principle, be a pure state of the universe including the environment, the observer, and the measured system. While this may matter to some (Zeh, 2000), real observers are forced to perceive the universe the way we do: We are a part of the universe, observing it from within. Hence, for us, *environment-induced superselection specifies what exists.*

Predictability emerges as a key criterion of existence. The only states **R** can rely on to store information are the pointer states. They are also the obvious choice for **S** to measure. Such measurements can be accomplished without danger of re-preparation. Einselected states are insensitive to measurement of the pointer observables—they have already been measured by the environment. Therefore additional projections P_i onto the einselected basis will not perturb the density matrix (Zurek, 1993a); it will be the same before and after the measurement:

$$\rho^D_{after}=\sum_i P_i\rho^D_{before}P_i . \tag{6.1}$$

Correlations with the einselected states will be left intact (Zurek, 1981, 1982).

Superselection for the observable $\hat{A}=\Sigma_i\lambda_iP_i$ with essentially arbitrary nondegenerate eigenvalues λ_i and eigenspaces P_i can be expressed (Bogolubov *et al.*, 1990) through Eq. (6.1). Einselection attains this, guaranteeing the diagonality of density matrices in the projectors P_i corresponding to pointer states. These are sometimes called *decoherence-free subspaces* when they are degenerate (compare also the non-Abelian case of *noiseless subsystems* discussed in quantum computation; see Zanardi and Rasetti, 1997; Duan and Guo, 1998; Zanardi, 1998, 2001; Lidar, Bacon, and Whaley, 1999; Blanchard and Olkiewicz, 2000; Knill, Laflamme, and Viola, 2000).

B. Measurements and memories

The memory of a measuring device or of an observer can be modeled as an open quantum apparatus $\mathcal{A}$, interacting with $\mathcal{S}$ through a Hamiltonian explicitly proportional to the measured observable[6] $\hat{s}$:

$$H_{int}=-g\hat{s}\hat{B}\sim\hat{s}\frac{\partial}{\partial\hat{A}}. \tag{6.2}$$

von Neumann (1932) considered an apparatus isolated from the environment. At the instant of the interaction between the apparatus and the measured system this is a convenient assumption. For us it suffices to assume that, at that instant, the interaction Hamiltonian between the system and the apparatus dominates. This can be accomplished by taking the coupling g in Eq. (6.2) to be $g(t)\sim\delta(t-t_0)$. Premeasurement happens at t_0:

$$\left(\sum_i \alpha_i|s_i\rangle\right)|A_0\rangle\rightarrow\sum_i \alpha_i|s_i\rangle|A_i\rangle. \tag{6.3}$$

In practice the action is usually large enough to accomplish amplification. As we have seen in Sec. II, all this can be done without an appeal to the environment.

For a real apparatus, interaction with the environment is inevitable. Idealized effectively classical memory will retain correlations, but will be subject to einselection. Only the einselected memory states (rather than their superpositions) will be useful for (or, for that matter, accessible to) the observer. The decoherence time scale is very short compared to the time after which memory states are typically consulted (i.e., copied or used in information processing), which is in turn much shorter than the relaxation time scale, on which memory "forgets."

Decoherence leads to classical correlation,

---

[6]The observable $\hat{s}$ of the system and $\hat{B}$ of the apparatus memory need not be discrete with a simple spectrum as was previously assumed. Even when $\hat{s}$ has a complicated spectrum, the outcome of the measurement can be recorded in the eigenstates of the memory observable $\hat{A}$, the conjugate of $\hat{B}$, Eq. (2.21). For the case of discrete $\hat{s}$ the necessary calculations that attain premeasurement—the quantum correlation that is the first step in the creation of the record—were already carried out in Sec. II. For the other situations they are quite similar. In either case, they follow the general outline of von Neumann's (1932) discussion.

$$\rho_{\mathcal{SA}}^{P}=\sum_{i,j}\alpha_i\alpha_j^*|s_i\rangle\langle s_j||A_i\rangle\langle A_j| \rightarrow\sum_i|\alpha_i|^2|s_i\rangle\langle s_i||A_i\rangle\langle A_i|=\rho_{\mathcal{SA}}^{D}, \quad (6.4)$$

following an entangling premeasurement. The left-hand side of Eq. (6.4) coincides with Eq. (2.44c), the outsider's view of the classical measurement. We shall see how and to what extent its other aspects, including the insider's Eq. (2.44a) and the discoverer's Eq. (2.44b), can be understood through einselection.

C. Axioms of quantum measurement theory

Our goal is to establish whether the above model can fulfill the requirements expected from measurement in textbooks (which are, essentially without exception, written in the spirit of the Copenhagen interpretation). There are several equivalent textbook formulations of the axioms of quantum theory. We shall (approximately) follow Farhi, Goldstone, and Gutmann (1989) and posit them as follows:

- (i) The states of a quantum system $\mathcal{S}$ are associated with the vectors $|\psi\rangle$, which are the elements of the Hilbert space $\mathcal{H}_{\mathcal{S}}$ that describes $\mathcal{S}$.
- (ii) The states evolve according to $i\hbar|\dot{\psi}\rangle=H|\psi\rangle$, where H is Hermitian.
- (iii a) Every observable O is associated with a Hermitian operator $\hat{O}$.
- (iii b) The only possible outcome of a measurement of O is an eigenvalue o_i of $\hat{O}$.
- (iv) Immediately after a measurement that yields the value o_i the system is in the eigenstate $|o_i\rangle$ of $\hat{O}$.
- (v) If the system is in a normalized state $|\psi\rangle$, then a measurement of $\hat{O}$ will yield the value o_i with the probability $p_i=|\langle o_i|\psi\rangle|^2$.

The first two axioms make no reference to measurements. They state the formalism of the theory. Axioms (iii)–(v) are, on the other hand, at the heart of the present discussion. In spirit, they go back to Bohr and Born. In letter, they follow von Neumann (1932) and Dirac (1947). The two key issues are the *projection postulate*, implied by a combination of (iv) with (iii b), and the *probability interpretation*, axiom (v).

To establish (iii b), (iv), and (v) we shall interpret in operational terms statements such as "the system is in the eigenstate" and "measurement will yield value ... with the probability ..." by specifying what these statements mean for the sequences of records made and maintained by an idealized, but physical memory.

We note that the above Copenhagen-like axioms presume the existence of quantum systems and of classical measuring devices. This (unstated) axiom (ø) complements axioms (i)–(v). Our version of axiom (ø) posits that the universe consist of quantum systems, and asserts that a composite system can be described by a tensor product of the Hilbert spaces of the constituent systems. Some quantum systems can be measured, and others can be used as measuring devices and/or memories and as quantum environments that interact with either or both.

Axioms (iii)–(v) contain many idealizations. For instance, in real life or in laboratory practice measurements have errors (and hence can yield outcomes other than the eigenvalues o_i). Moreover, only rarely do they prepare the system in the eigenstate of the observable they are designed to measure. Furthermore, coherent states—often an outcome of measurements, e.g., in quantum optics—form an overcomplete basis. Thus their detection does not correspond to a measurement of a Hermitian observable. Last but not least, the measured quantity may be inferred from some other quantity (e.g., beam deflection in the Stern-Gerlach experiment). Yet, we shall not go beyond the idealizations of (i)–(v) above. Our goal is to describe measurements in a quantum theory without collapse, to use axioms (ø), (i), and (ii) to understand the origin of the other axioms. Nonideal measurements are a fact of life incidental to this goal.

1. Observables are Hermitian—axiom (iii a)

In the model of measurement considered here the observables are Hermitian as a consequence of an assumed premeasurement interaction, e.g., Eq. (2.24). In particular, H_{int} is a product of the to-be-measured observable of the system and of the "shift operator" in the pointer of the apparatus or in the record state of the memory. Interactions involving non-Hermitian operators (e.g., $H_{int}\sim a^{\dagger}b+ab^{\dagger}$) may, however, also be considered.

It is tempting to speculate that one could dispose of the observables [and hence of the postulate (iii a)] altogether in the formulation of the axioms of quantum theory. The only ingredients necessary to describe measurements are then the effectively classical, but ultimately quantum, apparatus and the measured system. Observables emerge as a derived concept, as a useful idealization, ultimately based on the structure of the Hamiltonians. Their utility relies on the conservation laws, which relate the outcomes of several measurements. The most basic of these laws states that the system that did not (have time to) evolve will be found in the same state when it is remeasured. This is the content of axiom (iv). Other conservation laws are also reflected in the patterns of correlation in the measurement records, which must in turn arise from the underlying symmetries of the Hamiltonians.

Einselection should be included in this program, as it decides which observables are accessible and useful—which are effectively classical. It is conceivable that the fundamental superselection may also emerge in this manner (see Zeh, 1970 and Zurek, 1982, for early speculations; see Giulini, Kiefer, and Zeh, 1995; Kiefer, 1996, and Giulini, 2000, for the present status of this idea).

2. Eigenvalues as outcomes—axiom (iii b)

This axiom is the first part of the collapse postulate. Given einselection, axiom (iii b) is easy to justify: we need to show that only the records inscribed in the einselected states of the apparatus pointer can be read off, and that, in a well-designed measurement, they correlate with the eigenstates (and therefore, eigenvalues) of the measured observable $\hat{s}$.

With Dirac (1947) and von Neumann (1932) we assume that the apparatus is built so that it satisfies the obvious truth table when the eigenstates of the measured observable are at the input:

$$|s_i\rangle|A_0\rangle \rightarrow |s_i\rangle|A_i\rangle. \tag{6.5}$$

To assure this one can implement the interaction in accord with Eq. (6.2) and the relevant discussion in Sec. II. This is not to say that there are no other ways: Aharonov, Anandan, and Vaidman (1993); Braginski and Khalili (1996); and Unruh (1994) have all considered "adiabatic measurements" that correlate the apparatus with the discrete energy eigenstates of the measured system, nearly independently of the structure of H_{int}.

The truth table of Eq. (6.5) does not require collapse; for any initial $|s_i\rangle$ it represents a classical measurement in quantum notation in the sense of Sec. II. However, Eq. (6.5) typically leads to a superposition of outcomes. This is the "measurement problem." To address it, we assume that the record states $\{|A_i\rangle\}$ are einselected. This has two related consequences: (i) Following the measurement, the joint density matrix of the system and the apparatus decoheres, Eq. (6.3), so that it satisfies the superselection condition, Eq. (6.1), for $P_i = |A_i\rangle\langle A_i|$. (ii) Einselection restricts states that can be read off as if they were classical to pointer states.

Indeed, following decoherence only the pointer states $\{|A_i\rangle\}$ of the memory can be measured without diminishing the correlation with the states of the system. Without decoherence, as we have seen in Sec. II, one could use the entanglement between $\mathcal{S}$ and $\mathcal{A}$ to end up with almost arbitrary superposition states of either and hence to violate the letter and the spirit of axiom (iii b).

Outcomes are restricted to the eigenvalues of measured observables because of einselection. Axiom (iii b) is then a consequence of the effective classicality of the pointer states, the only ones that can be found out without being disturbed. They can be consulted repeatedly and remain unaffected under the joint scrutiny of the observers and of the environment (Zurek, 1981, 1993a, 1998a).

3. Immediate repeatability, axiom (iv)

This axiom supplies the second half of the collapse postulate. It asserts that in the absence of (the time for) evolution the quantum system will remain in the same state, and its remeasurement will lead to the same outcome. Hence, once the system is found out to be in a certain state, it really is there. As in Eq. (2.44b) the observer perceives potential options collapse to a single actual outcome. [The association of axiom (iv) with the collapse advocated here seems obvious, but it is not common. Rather, some form of our axiom (iii b) is usually regarded as the sole collapse postulate.]

Immediate repeatability for Hermitian observables with discrete spectra is straightforward to justify on the basis of the Schrödinger evolution generated by H_{int} of Eq. (6.2) alone, although its implications depend on whether the premeasurement is followed by einselection. Everett (1957a, 1957b) used the "no-decoherence" version as a foundation of his relative-state interpretation. On the other hand, without decoherence and einselection one could postpone the choice of what was actually recorded by taking advantage of the entanglement between the system and the apparatus and the resulting basis ambiguity, as is evident on the right-hand side of Eq. (6.3). For instance, a measurement carried out on the apparatus in a basis different from $\{|A_i\rangle\}$ would also exhibit a one-to-one correlation with the system: $\Sigma_i \alpha_i |s_i\rangle|A_i\rangle = \Sigma_k \beta_k |r_k\rangle|B_k\rangle$. This flexibility to rewrite wave functions in different bases comes at the price of relaxing the demand that the outcome states $\{|r_k\rangle\}$ be orthogonal (so that there would be no associated Hermitian observable). However, as was already noted, coherent states associated with a non-Hermitian annihilation operator can also be an outcome of a measurement. Therefore [and in spite of the strict interpretation of axiom (iii a)] this is not a very serious restriction.

In the presence of einselection the basis ambiguity disappears. Immediate repeatability would apply only to the records made in the einselected states. Other apparatus observables lose correlation with the state of the system on the decoherence time scale. In the effectively classical limit it is natural to demand repeatability extending beyond that very brief moment. This demand makes the role of einselection in establishing axiom (iv) evident. Indeed, such repeatability is, albeit in a more general context, the motivation for the predictability sieve.

4. Probabilities, einselection, and records

Density matrix alone, without the preferred set of states, does not suffice as a foundation for a probability interpretation. For, any mixed-state density matrix $\rho_{\mathcal{S}}$ can be decomposed into sums of density matrices that add up to the same resultant $\rho_{\mathcal{S}}$, but need not share the same eigenstates. For example, consider $\rho_{\mathcal{S}}^a$ and $\rho_{\mathcal{S}}^b$, representing two different preparations (i.e., involving the measurement of two distinct, noncommuting observables) of two ensembles, each with multiple copies of a system $\mathcal{S}$. When they are randomly mixed in proportions p^a and p^b, the resulting density matrix

$$\rho_{\mathcal{S}}^{a\vee b} = p^a \rho_{\mathcal{S}}^a + p^b \rho_{\mathcal{S}}^b$$

is the complete description of the unified ensemble (see Schrödinger, 1936; Jaynes, 1957).

Unless $[\rho_{\mathcal{S}}^a, \rho_{\mathcal{S}}^b] = 0$, the eigenstates of $\rho_{\mathcal{S}}^{a\vee b}$ do not coincide with the eigenstates of the components. This feature makes it difficult to regard any density matrix in terms of probabilities and ignorance. Such ambiguity would be especially troubling if it arose in the description of an observer (or, for that matter, of any classical system). The ignorance interpretation, i.e., the idea that probabilities are the observer's way of dealing with uncertainty about the outcome which we have briefly explored in the discussion of the insider-outsider dichotomy, Eqs. (2.44), requires at the very least that the set of events ("the sample space") exists independently of the information at hand, that is, independently of p^a and p^b in the example above. Eigenstates of the density matrix do not supply such events, since the additional loss of information associated with mixing of the ensembles alters the candidate events.

Basis ambiguity would be disastrous for record states. Density matrices describing a joint state of the memory $\mathcal{A}$ and of the system $\mathcal{S}$

$$\rho_{\mathcal{AS}}^{a\vee b} = p_a \rho_{\mathcal{AS}}^a + p_b \rho_{\mathcal{AS}}^b$$

would have to be considered. In the absence of einselection the eigenstates of such $\rho_{\mathcal{AS}}^{a\vee b}$ need not even be associated with a fixed set of record states of the presumably classical $\mathcal{A}$. Indeed, in general $\rho_{\mathcal{AS}}^{a\vee b}$ has a nonzero discord,[7] and its eigenstates are entangled (even when the above $\rho_{\mathcal{AS}}^{a\vee b}$ is separable, and can be expressed as a mixture of matrices that have no entangled eigenstates). This would imply an ambiguity of what the record states are precluding a probability interpretation of measurement outcomes.

The observer may nevertheless have records of a system that is in the ambiguous situation described above. Thus

$$\rho_{\mathcal{AS}}^{a\vee b} = \sum_k w_k |A_k\rangle\langle A_k| (p_k^a \rho_{\mathcal{S}_k}^a + p_k^b \rho_{\mathcal{S}_k}^b)$$

is admissible for an effectively classical $\mathcal{A}$ correlated with a quantum $\mathcal{S}$. Now the discord $\delta_{\mathcal{A}}\mathcal{I}(\mathcal{S}|\mathcal{A}) = 0$.

Mixing of ensembles of pairs of correlated systems, one of which is subject to einselection, does not lead to the ambiguities discussed above. The discord $\delta_{\mathcal{A}}(\mathcal{S}|\mathcal{A})$ disappears in the einselected basis of $\mathcal{A}$, and the eigenvalues of the density matrices can behave as classical probabilities associated with events with the records. The menu of possible events, in the sample space, e.g., records in memory, is fixed by einselection. Whether one can really justify this interpretation of the eigenvalues of the reduced density matrix is a separate question we are about to address.

D. Probabilities from envariance

The view of the emergence of the classical based on environment-induced superselection has been occasionally described as "for all practical purposes only" (see, for example, Bell, 1990), to be contrasted with the more fundamental (if nonexistent) solutions of the problem one could imagine (i.e., by modifying quantum theory; see Bell, 1987, 1990). This attitude can be traced in part to the reliance of einselection on reduced density matrices. For even when explanations of all aspects of the effectively classical behavior are accepted in the framework of, say, Everett's many-worlds interpretation, and after the operational approach to the objectivity and perception of unique outcomes based on the existential interpretation explained earlier is adopted, one major gap remains: Born's rule—axiom (v) connecting probabilities with amplitudes, $p_k = |\psi_k|^2$—has to be postulated in addition to axioms (ø)–(ii). True, one can show that within the framework of einselection Born's rule emerges naturally (Zurek, 1998a). Decoherence is, however, based on reduced density matrices. Since their introduction by Landau (1927), it is known that a partial trace leading to reduced density matrices is predicated on Born's rule (see Nielsen and Chuang, 2000, for a discussion). Thus derivations of Born's rule that employ reduced density matrices are open to the charge of circularity (Zeh, 1997). Moreover, repeated attempts to justify $p_k = |\psi_k|^2$ within the no-collapse many-worlds interpretation (Everett, 1957a, 1957b; DeWitt, 1970; De Witt and Graham, 1973; Geroch, 1984) have failed (see, for example, Stein, 1984; Kent, 1990; Squires, 1990). The problem is their circularity. An appeal to the connection (especially in certain limiting procedures) between the smallness of the amplitude and the vanishing of the probabilities has to be made to establish that the relative frequencies of events averaged over branches of the universal state vector are consistent with Born's rule. In particular, one must a claim that "maverick" branches of the MWI state vector that have "wrong" relative frequencies have a vanishing probability because their Hilbert-space measures are small. This is circular, as noted even by the proponents (DeWitt, 1970).

My aim here is to look at the origin of ignorance, information, and, therefore, probabilities from a very quantum and fundamental perspective. Rather than fo-

[7]As we have seen in Sec. IV, Eqs. (4.30)–(4.36), *discord* $\delta\mathcal{I}_{\mathcal{A}}(\mathcal{S}|\mathcal{A}) = \mathcal{I}(\mathcal{S}{:}\mathcal{A}) - \mathcal{J}_{\mathcal{A}}(\mathcal{S}{:}\mathcal{A})$ is a measure of the "quantumness" of correlations. It should disappear as a result of the classical equivalence of two definitions of the mutual information, but is in general positive for quantum correlations, including, in particular, predecoherence $\rho_{\mathcal{SA}}$. Discord is asymmetric, $\delta\mathcal{I}_{\mathcal{A}}(\mathcal{S}|\mathcal{A}) \neq \delta\mathcal{I}_{\mathcal{S}}(\mathcal{A}|\mathcal{S})$. The vanishing of $\delta\mathcal{I}_{\mathcal{A}}(\mathcal{S}|\mathcal{A})$ [i.e., of the discord in the direction exploring the classicality of the states of $\mathcal{A}$, on which $H(\rho_{\mathcal{S}|\mathcal{A}})$ in the asymmetric $J_{\mathcal{A}}(\mathcal{S}{:}\mathcal{A})$, Eq. (4.32), is conditioned] is necessary for the classicality of the measurement outcome (Ollivier and Zurek, 2002; Zurek, 2000, 2003a). $\delta\mathcal{I}_{\mathcal{A}}(\mathcal{S}|\mathcal{A})$ can disappear as a result of decoherence in the einselected basis of the apparatus. Following einselection it is then possible to ascribe probabilities to the pointer states. In perfect measurements of Hermitian observables discord vanishes "both ways": $\delta\mathcal{I}_{\mathcal{A}}(\mathcal{S}|\mathcal{A}) = \delta\mathcal{I}_{\mathcal{S}}(\mathcal{A}|\mathcal{S}) = 0$ for the pointer basis and for the eigenbasis of the measured observable correlated with it. Nevertheless, it is possible to encounter situations when vanishing of the discord in one direction is not accompanied by its vanishing "in reverse." Such correlations are "classical one way" (Zurek, 2003a).

This asymmetry between classical $\mathcal{A}$ and quantum $\mathcal{S}$ arises from the einselection. Classical record states are not arbitrary superpositions. The observer accesses his memory in the basis in which it is monitored by the environment. The information stored is effectively classical because it is being widely disseminated. States of the observer's memory exist objectively; they can be determined through their imprints in the environment.

cus on probabilities for an individual isolated system I shall—in the spirit of einselection, but without employing its usual tools such as trace or reduced density matrices—consider what the observer can (and cannot) know about a system entangled with its environment. Within this context I shall demonstrate that Born's rule follows from the very quantum fact that one can know precisely the state of the composite system and yet be provably ignorant of the state of its components. This is due to environment-assisted invariance or envariance, a hitherto unrecognized symmetry I am about to describe.

Envariance of pure states is a symmetery conspicuously missing from classical physics. It allows one to define ignorance as a consequence of invariance, and thus to understand the origin of Born's rule, the probabilities, and ultimately the origin of information through arguments based on assumptions different from Gleason's (1957) famous theorem. Rather, it is based on the Machian idea of the relativity of quantum states, suggested by this author two decades ago (see p. 772 of Wheeler and Zurek, 1983), but not developed until now. Envariance (Zurek, 2003b) addresses the question of the meaning of these probabilities by defining "ignorance" and leads to correct relative frequencies.

1. Envariance

Environment-assisted invariance is a symmetry exhibited by a system $\mathcal{S}$ correlated with another system (which we shall call the environment $\mathcal{E}$). When a state of the composite $\mathcal{SE}$ can be transformed by $u_{\mathcal{S}}$ acting solely on the Hilbert space $\mathcal{H}_{\mathcal{S}}$, but the effect of this transformation can be undone with an appropriate $u_{\mathcal{E}}$ acting only on $\mathcal{H}_{\mathcal{E}}$, so that the joint state $|\psi_{\mathcal{SE}}\rangle$ remains unaltered, so that

$$u_{\mathcal{E}} u_{\mathcal{S}} |\psi_{\mathcal{SE}}\rangle = u_{\mathcal{E}} |\eta_{\mathcal{SE}}\rangle = |\psi_{\mathcal{SE}}. \tag{6.6}$$

Such a $|\psi_{\mathcal{SE}}\rangle$ is envariant under $u_{\mathcal{S}}$. The generalization to mixed $\rho_{\mathcal{SE}}$ is obvious, but we shall find it easier to assume that $\mathcal{SE}$ has been purified in the usual fashion, i.e., by enlarging the environment.

Envariance is best elucidated by considering an example, an entangled state of $\mathcal{S}$ and $\mathcal{E}$. It can be expressed in the Schmidt basis as

$$|\psi_{\mathcal{SE}}\rangle = \sum_k \alpha_k |s_k\rangle |\varepsilon_k\rangle, \tag{6.7}$$

where α_k are complex, while $\{|s_k\rangle\}$ and $\{|\varepsilon_k\rangle\}$ are orthonormal. For $|\psi_{\mathcal{SE}}\rangle$ (and hence, given our above remark about purification, for any system correlated with the environment) it is easy to demonstrate the following.

Lemma 6.1: Unitary transformations codiagonal with the Schmidt basis of $|\psi_{\mathcal{SE}}\rangle$ leave it envariant.

The proof relies on the form of such transformations:

$$u_{\mathcal{S}}^{\{|s_k\rangle\}} = \sum_k e^{i\sigma_k} |s_k\rangle\langle s_k|, \tag{6.8}$$

where σ_k is a phase. Hence

$$u_{\mathcal{S}}^{\{|s_k\rangle\}} |\psi_{\mathcal{SE}}\rangle = \sum_k \alpha_k e^{i\sigma_k} |s_k\rangle |\varepsilon_k\rangle \tag{6.9}$$

can be undone by

$$u_{\mathcal{E}}^{\{|\varepsilon_k\rangle\}} = \sum_k e^{i\epsilon_k} |\varepsilon_k\rangle\langle \varepsilon_k|, \tag{6.10}$$

providing that $\epsilon_k = 2\pi l_k - \sigma_k$ for some integer l_k. QED.

Thus phases associated with the Schmidt basis are envariant. We shall see below that they are the only envariant property of entangled states. The transformations defined by Eq. (6.8) are rather specific—they share (Schmidt) eigenstates. Still, their existence leads us to

Theorem 6.1: A local description of the system $\mathcal{S}$ entangled with a causally disconnected environment $\mathcal{E}$ must not depend on the phases of the coefficients α_k in the Schmidt decomposition of $|\psi_{\mathcal{SE}}\rangle$.

It follows that all the measurable properties of $\mathcal{S}$ are completely specified by the list of pairs $\{|\alpha_k|;|s_k\rangle\}$. A different way of establishing this phase envariance theorem appeals even more directly to causality. Phases of $|\psi_{\mathcal{SE}}\rangle$ can be arbitrarily changed by acting on $\mathcal{E}$ alone [e.g, by the local Hamiltonian with eigenstates $|\varepsilon_k\rangle$, generating evolution of the form of Eq. (6.9)]. But causality prevents faster-than-light communication. Hence no measurable property of $\mathcal{S}$ can be effected by acting on $\mathcal{E}$. Clearly, there is an intimate connection between envariance and causality. Independence of the local state of $\mathcal{S}$ from the phases of the Schmidt coefficients α_k follows from envariance alone, but it could be also established through an appeal to causality. The situation is similar as with "no cloning theorem." It was proved using linearity of quantum theory, but one could have also inferred impossibility of cloning an unknown quantum state from special-relativistic causality. The proof based on linearity is "less expensive," as it does not require ingredients that go beyond quantum theory.

Phase envariance theorem will turn out to be the crux of our argument. It relies on an input—entanglement and envariance—which has not been employed to date in discussions of the origin of probabilities. In particular, this input is different and more "physical" than that of the successful derivation of Born's rules by Gleason (1957).

We also note that information contained in the "database" $\{|\alpha_k|;|s_k\rangle\}$ implied by Theorem 6.1 is the same as in the reduced density matrix of the system $\rho_{\mathcal{S}}$. Although we do not yet know the probabilities of various $|s_k\rangle$, the preferred basis of $\mathcal{S}$ has been singled out; Schmidt states (sometimes regarded as instantaneous pointer states; see, for example, Albrecht, 1992, 1993) play a special role as the eigenstates of envariant transformations. Moreover, probabilities can depend on $|\alpha_k|$ (but not on the phases). We still do not know that $p_k = |\alpha_k|^2$.

The causality argument we could have used to establish Theorem 6.1 applies of course to arbitrary transformations one could perform on $\mathcal{E}$. However, such transformations would in general not be envariant (i.e., could not be undone by acting on $\mathcal{S}$ alone). Indeed, this is one way to see that causality is a more potent ingredient than envariance. In particular, all envariant transformation have a fairly restricted form as follows.

Lemma 6.2: All of the unitary envariant transformations of $|\psi_{\mathcal{SE}}\rangle$ have Schmidt eigenstates.

The proof relies on the fact that other unitary transformations would rotate the Schmidt basis, $|s_k\rangle \rightarrow |\tilde{s}_k\rangle$. The rotated basis becomes a new "Schmidt," and this fact cannot be affected by unitary transformations of $\mathcal{E}$, by state rotations in the environment. But a state that has a different Schmidt decomposition from the original $|\psi_{\mathcal{SE}}\rangle$ is different. Hence a unitary transformation must be codiagonal with the Schmidt basis of $\psi_{\mathcal{SE}}$ to leave it envariant. QED.

2. Born's rule from envariance

When absolute values of some of the coefficients in Eq. (6.7) are equal, any orthonormal basis is Schmidt in the corresponding subspace of $\mathcal{H}_{\mathcal{S}}$. This implies envariance of a more general nature, e.g., under a *swap*:

$$u_{\mathcal{S}}(k \leftrightarrow j) = e^{i\phi_{kj}}|s_k\rangle\langle s_j| + \text{H.c.} \tag{6.11}$$

A swap can be generated by a phase rotation, Eq. (6.8), but in a basis complementary to the one swapped. Its envariance does not contradict Lemma 6.2, as any orthonormal basis in this case is also Schmidt. So when $|\alpha_k| = |\alpha_j|$, the effect of a swap on the system can be undone by an obvious counterswap in the environment:

$$u_{\mathcal{S}}(k \leftrightarrow j) = e^{-i(\phi_{kj}+\phi_k-\phi_j+2\pi l_{kj})}|\varepsilon_k\rangle\langle\varepsilon_j| + \text{H.c.} \tag{6.12}$$

A swap can be applied to states that do not have equal absolute values of the coefficients, but in that case it is no longer envariant. Partial swaps can also be generated, for example, by underrotating or by a $u_{\mathcal{S}}^{\{|r_i\rangle\}}$, Eq. (6.8), but with the eigenstates $\{|r_i\rangle\}$ intermediate between those of the swapped and the complementary (Hadamard) basis. A swap followed by a counterswap exchanges coefficients of the swapped states in the Schmidt expansion, Eq. (6.7). Hence, $\psi_{\mathcal{SE}}$ is envariant under swaps $u_{\mathcal{S}}(j \leftrightarrow k)$ only when $|\alpha_k| = |\alpha_j|$.

States of correlated classical systems can also exhibit something akin to envariance under a classical version of swaps. For instance, a correlated state of a system and an apparatus described by $\rho_{\mathcal{SA}} \sim |s_k\rangle\langle s_k||A_k\rangle\langle A_k| + |s_j\rangle\langle s_j||A_j\rangle\langle A_j|$ can be swapped and counterswapped. The corresponding transformations would be still given by, in effect, Eqs. (6.11) and (6.12), but without phases, and swaps could no longer be generated by rotations around the complementary basis. This situation corresponds to the outsider's view of the measurement process, Eq. (2.44c). The outsider can be aware of the correlation between the system and the apparatus, but ignorant of their individual states. This connection between ignorance and envariance will be exploited below.

Envariance based on ignorance may be found in the classical setting, but envariance of pure states is purely quantum. Observers can know perfectly the quantum joint state of $\mathcal{SE}$, yet be provably ignorant of $\mathcal{S}$. Consider a measurement carried out on the state vector of $\mathcal{SE}$ from the point of view of envariance:

$$|A_0\rangle \sum_{k=1}^{N} |s_k\rangle|\varepsilon_k\rangle \rightarrow \sum_{k=1}^{N} |A_k\rangle|s_k\rangle|\varepsilon_k\rangle \sim |\Phi_{\mathcal{SAE}}\rangle. \tag{6.13}$$

Above, we have assumed that the absolute values of the coefficients are equal (and omitted them for notational simplicity). We have also ignored phases (which need not be equal) since by the phase envariance theorem they will not influence the state (and hence, the probabilities) associated with $\mathcal{S}$.

Before the measurement the observer with access to $\mathcal{S}$ cannot notice swaps in the states [such as Eq. (6.13)] with equal absolute values of the Schmidt coefficients. This follows from the envariance of the premeasurement $|\psi_{\mathcal{SE}}\rangle$ under swaps, Eq. (6.11).

One could argue this point in more detail by comparing what happens for two very different input states: an entangled $|\psi_{\mathcal{SE}}\rangle$ with equal absolute values of Schmidt coefficients and a product state:

$$|\varphi_{\mathcal{SE}}\rangle = |s_J\rangle|\varepsilon_J\rangle.$$

When the observer knows he is dealing with $\varphi_{\mathcal{SE}}$, he knows the system is in the state $|s_J\rangle$, and can predict the outcome of the corresponding measurement on $\mathcal{S}$. The Schrödinger equation or just the resulting truth table, Eq. (6.5), implies with certainty that his state—the future state of his memory—will be $|A_J\rangle$. Moreover, swaps involving $|s_J\rangle$ are not envariant for $\varphi_{\mathcal{SE}}$. They just swap the outcomes [i.e., when $u_{\mathcal{S}}(J \leftrightarrow L)$ precedes the measurement, memory will end up in $|A_L\rangle$].

By contrast,

$$|\psi_{\mathcal{SE}}\rangle \sim \sum_{k=1}^{N} e^{i\phi_k}|s_k\rangle|\varepsilon_k\rangle$$

is envariant under swaps. This allows the observer (who knows the joint state of $\mathcal{SE}$ exactly) to conclude that the probabilities of all the envariantly swappable outcomes must be the same. The observer cannot predict his memory state after the measurement of $\mathcal{S}$ because he knows too much: the exact combined state of $\mathcal{SE}$.

For completeness, we note that when there are system states that are absent from the above sum, i.e., states that appear with zero amplitude, they cannot be envariantly swapped with the states present in the sum. Of course, the observer can predict with certainty that he will not detect any of the corresponding zero-amplitude outcomes. For, following the measurement that correlates memory of the observer with the basis $\{|s_k\rangle\}$ of the system, there will be simply no terms describing observer with the record of such nonexistent states of $\mathcal{S}$. This argument about the ignorance of the observer concerning his future state, concerning the outcome of the measurement he is about to perform, is based on his perfect knowledge of a joint state of $\mathcal{SE}$.

Probabilities refer to the guess the observer makes on the basis of his information before the measurement about the state of his memory—the future outcome—after the measurement. Since the left-hand side of Eq. (6.13) is envariant under swaps of the system states, the probabilities of all the states must be equal. Thus, by normalization,

$$p_k = 1/N. \tag{6.14}$$

Moreover, the probability of n mutually exclusive events that all appear in Eq. (6.13) with equal coefficients must be

$$p_{k_1 \vee k_2 \vee \cdots \vee k_n} = n/N. \tag{6.15}$$

This concludes the discussion of the equal probability case. Our case rests on the independence of the state of $\mathcal{S}$ entangled with $\mathcal{E}$ from the phases of the coefficients in the Schmidt representation, the phase envariance theorem 6.1, which in the case of equal coefficients, Eq. (6.13), allows envariant swapping, and yields Eqs. (6.14) and (6.15).

After a measurement the situation changes. In accord with our preceding discussion we interpret the presence of the term $|A_k\rangle$ in Eq. (6.13) as evidence that an outcome $|s_k\rangle$ can be (or indeed has been—the language here is somewhat dependent on the interpretation) recorded. Conversely, the absence of some $|A_{k'}\rangle$ in the sum above implies that the outcome $|s_{k'}\rangle$ cannot occur. After a measurement the memory of the observer who has detected $|s_k\rangle$ will contain the record $|A_k\rangle$. Further measurements of the same observable on the same system will confirm that $\mathcal{S}$ is in indeed in the state $|s_k\rangle$.

This postmeasurement state is still envariant, but only under swaps that involve jointly the state of the system and the correlated state of the memory:

$$u_{\mathcal{AS}}(k \leftrightarrow j) = e^{i\phi_{kj}} |s_k, A_k\rangle\langle s_j, A_j| + \text{H.c.} \tag{6.16}$$

Thus if another observer ("Wigner's friend") was getting ready to find out, either by direct measurement of $\mathcal{S}$ or by communicating with observer $\mathcal{A}$, the outcome of $\mathcal{A}$'s measurement, he would be (on the basis of envariance) provably ignorant of the outcome $\mathcal{A}$ has detected, but could be certain of the $\mathcal{AS}$ correlation. We shall employ this joint envariance in the discussion of the case of unequal probabilities immediately below.

Note that our reasoning does not really appeal to the information lost in the environment in the sense in which this phrase is often used. Perfect knowledge of the combined state of the system and the environment is the basis of the argument for the ignorance of $\mathcal{S}$ alone. For entangled $\mathcal{SE}$, perfect knowledge of $\mathcal{SE}$ is incompatible with perfect knowledge of $\mathcal{S}$. This is really a consequence of indeterminacy; joint observables with entangled eigenstates such as $\psi_{\mathcal{SE}}$ simply do not commute (as the reader is invited to verify) with the observables of the system alone. Hence ignorance associated with envariance is ultimately mandated by Heisenberg indeterminacy.

The case of unequal coefficients can be reduced to the case of equal coefficients. This can be done in several ways, of which we choose one that makes use of the preceding discussion of the envariance of the postmeasurement state. We start with

$$|\Phi_{\mathcal{SAE}}\rangle \sim \sum_{k=1}^{N} \alpha_k |A_k\rangle |s_k\rangle |\varepsilon_k\rangle, \tag{6.17}$$

where $\alpha_k \sim \sqrt{m_k}$ and m_k is a natural number (and, by Theorem 6.1, we drop the phases). To get an envariant state we increase the resolution, of $\mathcal{A}$ by assuming that

$$|A_k\rangle = \sum_{j_k=1}^{m_k} |a_{j_k}\rangle / \sqrt{m_k}. \tag{6.18}$$

An increase of resolution is a standard trick, used in classical probability theory "to even the odds." Note that we assume that basis states such as $|A_k\rangle$ are normalized (as they must be in a Hilbert space). This leads to

$$|\Phi_{\mathcal{SAE}}\rangle \sim \sum_{k=1}^{N} \sqrt{m_k} \frac{\sum_{j_k=1}^{m_k} |a_{j_k}\rangle}{\sqrt{m_k}} |s_k\rangle |\varepsilon_k\rangle. \tag{6.19}$$

We now assume that $\mathcal{A}$ and $\mathcal{E}$ interact (e.g., through a c-shift of Sec. II, with a truth table $|a_{j_k}\rangle|\varepsilon_k\rangle \rightarrow |a_{j_k}\rangle|e_{j_k}\rangle$ where $\{|e_{j_k}\rangle\}$ are all orthonormal). After simplifying and rearranging terms we get a sum, over a new fine-grained index, with the states of $\mathcal{S}$ that remain the same within coarse-grained cells, with the cell size measured by m_k:

$$|\tilde{\Phi}_{\mathcal{SAE}}\rangle \sim \sum_{k=1}^{N} |s_k\rangle \left(\sum_{j_k=1}^{m_k} |a_{j_k}\rangle |e_{j_k}\rangle \right) = \sum_{j=1}^{M} |s_{k(j)}\rangle |a_j\rangle |e_j\rangle. \tag{6.20}$$

Above, $M = \Sigma_{k=1}^{N} m_k$, $k(j)=1$ for $j \leq m_1$, $k(j)=2$ for $m_1 < j \leq m_1 + m_2$, etc. The above state is envariant under combined swaps:

$$u_{\mathcal{SA}}(j \leftrightarrow j') = \exp(i\phi_{jj'}) |s_{k(j)}, a_j\rangle\langle a_{j'}, s_{k(j')}| + \text{h.c.}$$

Suppose that an additional observer measures $\mathcal{SA}$ in the obviously swappable joint basis. By our equal-coefficients argument, Eq. (6.14), we get

$$p(s_{k(j)}, a_j) = 1/M.$$

But the observer can ignore states a_j. Then the probability of different Schmidt states of $\mathcal{S}$ is, by Eq. (6.15),

$$p(s_k) = m_k / M = |\alpha_k|^2. \tag{6.21}$$

This is Born's rule.

The case with coefficients that do not lead to commensurate probabilities can be treated by assuming continuity of probabilities as a function of the amplitudes, and taking appropriate (and obvious) limits. This can be physically motivated: One would not expect probabilities to change drastically depending on infinitesimal changes of state. One can also extend the strategy outlined above to deal with probabilities (and probability densities) in cases such as $|s(x)\rangle$, i.e., when the index of the state vector changes continuously. This can be accomplished by discretizing it [so that the measurement of Eq. (6.17) correlates different apparatus states with small intervals of x] and then repeating the strategy of Eqs. (6.17)–(6.21). The wave function $s(x)$ should be sufficiently smooth for this strategy to succeed.

We note that the increase of resolution we have exploited, Eqs. (6.18)–(6.21), need not be physically implemented for the argument to proceed. The very possibility of carrying out these steps within the quantum formalism forces one to adopt Born's rule. For example, if the apparatus did not have the requisite extra resolution, Eq. (6.18), the interaction of the environment with a still different "counterweight" system $\mathcal{C}$ that yields

$$|\Psi_{\mathcal{SAEC}}\rangle=\sum_{k=1}^{N}\sqrt{m}_k|s_k\rangle|A_k\rangle|\varepsilon_k\rangle|C_k\rangle \tag{6.22}$$

would lead one to Born's rule through steps similar to these that we have invoked before, providing that $\{|C_k\rangle\}$ has the requisite resolution, $|C_k\rangle=\Sigma_{j_k=1}^{m_k}|c_{j_k}\rangle/\sqrt{m}_k$. An interaction resulting in a correlation, Eq. (6.22), can occur between $\mathcal{E}$ and $\mathcal{C}$, and happen far from the system of interest or from the apparatus. Thus it will not influence the probabilities of the outcomes of measurements carried out on $\mathcal{S}$ or of the records made by $\mathcal{A}$. Yet, the fact that it can happen leads us to the desired conclusion.

3. Relative frequencies from envariance

Relative frequency is a common theme in studies that aim to elucidate the physical meaning of probabilities in quantum theory (Everett, 1957a, 1957b; Hartle, 1968; DeWitt, 1970; Graham, 1970; Farhi, Goldstone, and Gutmann, 1989; Aharonov and Reznik, 2002). In particular, in the context of the no-collapse many-worlds interpretation relative frequency seems to offer the best hope of arriving at Born's rule and elucidating its physical significance. Yet, it is generally acknowledged that the MWI derivations offered to date have failed to attain this goal (Kent, 1990).

We postpone a brief discussion of these efforts to the next section, and describe an approach to relative frequencies based on envariance. Consider an ensemble of many ($\mathcal{N}$) distinguishable systems prepared in the same initial state:

$$|\sigma_{\mathcal{S}}\rangle=\alpha|0\rangle+\beta|1\rangle. \tag{6.23}$$

We focus on the two-state case to simplify the notation. We also assume that $|\alpha|^2$ and $|\beta|^2$ are commensurate, so that the state vector of the whole ensemble of correlated triplets $\mathcal{SAE}$ after the requisite increases of resolution [see Eqs. (6.18)–(6.20) above] is given by

$$|\Phi_{\mathcal{SAE}}^{\mathcal{N}}\rangle\sim\left(\sum_{j=1}^{m}|0\rangle|a_j\rangle|e_j\rangle+\sum_{j=m+1}^{M}|1\rangle|a_j\rangle|e_j\rangle\right)^{\otimes\mathcal{N}}, \tag{6.24}$$

save for the obvious normalization. This state is envariant under swaps of the joint states $|s,a_j\rangle$, as they appear with the same (absolute value) of the amplitude in Eq. (6.24). (By Theorem 6.1 we can omit phases.)

After the exponentiation is carried out, and the resulting product states are sorted by the number of 0's and 1's in the records, we can calculate the number of terms with exactly n 0's: $\nu_{\mathcal{N}}(n)=\binom{\mathcal{N}}{n}m^n(M-m)^{\mathcal{N}-n}$. To get the probability, we normalize:

$$p_{\mathcal{N}}(n)=\binom{\mathcal{N}}{n}\frac{m^n(M-m)^{\mathcal{N}-n}}{M^{\mathcal{N}}}=\binom{\mathcal{N}}{n}|\alpha|^{2n}|\beta|^{2(\mathcal{N}-n)}. \tag{6.25}$$

This is the distribution one would expect from Born's rule. To establish the connection with relative frequencies we appeal to the de Moivre–Laplace theorem (Gnedenko, 1982), which allows one to approximate above $p_{\mathcal{N}}(n)$ with a Gaussian:

$$p_{\mathcal{N}}(n)\simeq\frac{1}{\sqrt{2\pi\mathcal{N}}|\alpha\beta|}\exp\left\{-\frac{1}{2}\left[\frac{n-\mathcal{N}|\alpha|^2}{\sqrt{\mathcal{N}}|\alpha\beta|}\right]^2\right\}. \tag{6.26}$$

This last step requires large $\mathcal{N}$, but our previous discussion including Eq. (6.25) is valid for arbitrary $\mathcal{N}$. Indeed, Eq. (6.21) can be regarded as the $\mathcal{N}=1$ case.

Nevertheless, for large $\mathcal{N}$ the relative frequency is sharply peaked around the expected $\langle n\rangle=\mathcal{N}|\alpha|^2$. Indeed, in the limit $\mathcal{N}\to\infty$ the appropriately rescaled $p_{\mathcal{N}}(n)$ tends to a Dirac $\delta(v-|\alpha|^2)$ in the relative frequency $v=n/\mathcal{N}$. This justifies the relative-frequency interpretation of the squares of amplitudes as probabilities in the MWI context. Maverick universes with different relative frequencies exist, but have a vanishing probability (and not just a vanishing Hilbert-space measure) for large $\mathcal{N}$.

Our derivation of the physical significance of the probabilities, while it led to relative frequencies, was based on a very different set of assumptions than previous derivations. The key idea behind it is the connection between symmetry (envariance) and ignorance (impossibility of knowing something). The unusual feature of our argument is that this ignorance (for an individual system $\mathcal{S}$) is demonstrated by appealing to the perfect knowledge of the larger joint system that includes $\mathcal{S}$ as a subsystem.

We emphasize that one could not carry out the basic step of our argument—the proof of the independence of the likelihoods from the phases of the Schmidt expansion coefficients—for an equal-amplitude pure state of a single, isolated system. The problem with $|\psi\rangle=N^{-1/2}\Sigma_k^N\exp(i\phi_k)|k\rangle$ is the accessibility of the phases. Consider, for instance, $|\psi\rangle\sim|0\rangle+|1\rangle-|2\rangle$ and $|\psi'\rangle\sim|2\rangle+|1\rangle-|0\rangle$. In the absence of decoherence the swapping of k's is detectable. Interference measurements (i.e., measurements of the observables with phase-dependent eigenstates $|1\rangle+|2\rangle$, $|1\rangle-|2\rangle$, etc.) would have revealed the difference between $|\psi\rangle$ and $|\psi'\rangle$. Indeed, given an ensemble of identical pure states an observer will simply determine what they are. Loss of phase coherence is essential to allow for the shuffling of the states and coefficients.

Note that in our derivation the environment and einselection play an additional, more subtle role. Once a measurement has taken place, i.e., a correlation with the apparatus or with the memory of the observer has been established, one would hope that the records would retain validity over a long time, well beyond the decoherence time scale. This is a precondition for axiom (iv). Thus a collapse from a multitude of possibilities to a single reality can be confirmed by subsequent measurements only in the einselected pointer basis.

4. Other approaches to probabilities

Gnedenko (1982), in his classic textbook, lists three classical approaches to probability:

(a) Definitions that appeal to the relative frequency of occurrence of events in a large number of trials.

(b) Definitions of probability as a measure of the certainty of the observer.

(c) Definitions that reduce probability to the more primitive notion of equal likelihood.

In the quantum setting, the relative-frequency approach has been to date the most popular, especially in the context of the no-collapse many-worlds interpretation (Everett, 1957a, 1957b; DeWitt, 1970; Graham, 1970). Counting the number of "clicks" seems most directly tied to the experimental manifestations of probability. Yet, the Everett interpretation versions were generally found lacking (Kent, 1990; Squires, 1990), since they relied on circular reasoning, invoking, without physical justification, an abstract measure of Hilbert space to obtain a physical measure (frequency). Some of the criticisms seem relevant also for the versions of this approach that allow for the measurement postulates (iii) and (iv) (Hartle, 1968; Farhi, Goldstone, and Guttmann, 1989). Nevertheless, for the infinite ensembles considered in the above references (where, in effect, the Hilbert-space measure of the many-worlds interpretation branches that violate relative-frequency predictions is zero) the eigenvalues of the frequency operator acting on a large or infinite ensemble of identical states will be consistent with the (Born formula) prescription for probabilities.

However, the infinite size of the ensemble necessary to prove this point is troubling (and unphysical) and taking the limit starting from a finite case is difficult to justify (Stein, 1984; Kent, 1990; Squires, 1990). Moreover, the frequency operator is a collective observable of the whole ensemble. It may be possible to relate observables defined for such an infinite ensemble supersystem to the states of individual subsystems, but the frequency operator does not do this. This is well illustrated by the gedanken experiment envisaged by Farhi *et al.* (1989). To provide a physical implementation of the frequency operator they consider a version of the Stern-Gerlach experiment where all the spins are attached to a common lattice. Thus during the passage through the inhomogeneity of the magnetic field, the center of mass of the whole lattice is deflected by an angle proportional to the projection of the net magnetic moment associated with the spins on the direction defined by the field gradient. The deflection is proportional to the eigenvalue of the frequency operator, which is then a collective observable—states of individual spins remain in superpositions, uncorrelated with anything outside. This difficulty can be addressed with the help of decoherence (Zurek, 1998a), but using decoherence without justifying Born's formula first is fraught with the danger of circularity.

The measure of certainty seems to be a rather vague concept. Yet Cox (1946) has demonstrated that Boolean logic leads, after the addition of a few reasonable assumptions, to the definition of probabilities that, in a sense, appear as an extension of the logical truth values. However, the rules of symbolic logic that underlie Cox's theorems are classical. One can adopt this approach (Zurek, 1998a) to probabilities in quantum physics only after decoherence intervenes, restoring the validity of the distributive law, which is not valid in quantum physics (Birkhoff and von Neumann, 1936).

One can carry out the equal-likelihood approach in the context of decoherence (Zurek, 1998a). The problems are, as pointed out before, the use of the trace and the dangers of circularity. An attempt to pursue a strategy akin to equal likelihood in the quantum setting at the level of the pure states of individual systems has also been made by Deutsch in his (unpublished) "signaling" approach to probabilities. The key idea is to consider a source of pure states, and to find out when the permutations of a set of basis states can be detected, and therefore, used for communication. When permutations are undetectable, the probabilities of the permuted set of states are declared equal. The problem with this idea (or with its more formal version described by DeWitt, 1998) is that it works only for superpositions that have all the coefficients identical, including their phases. Thus, as we have already noted, for closed systems, phases matter and there is no invariance under swapping. In a recent paper Deutsch (1999) adopted a different approach based on decision theory. The basic argument focuses again on individual states of quantum systems, but, as noted in the critical comment by Barnum *et al.* (2000), seems to appeal to some of the aspects of decision theory that depend on probabilities. In my view, it also leaves the problem of phase dependence of the coefficients unaddressed.

Among other approaches, the recent work of Gottfried (2000) shows that in a discrete quantum system coupled with a continuous quantum system Born's formula follows from the demand that the continuous system should follow classical mechanics in the appropriate limit. A somewhat different strategy, with a focus on the coincidences of the expected magnitude of fluctuations, was proposed by Aharonov and Reznik (2002).

In comparison with all of the above strategies, "probabilities from envariance" is the most radically quantum, in that it ultimately relies on entanglement (which is still sometimes regarded as "a paradox," and "to be explained"; I have used it as an explanation). This may be the reason why it has not been discovered until now. The insight offered by envariance into the nature of ignorance and information sheds new light on probabilities in physics. The (very quantum) ability to prove the ignorance of a part of a system by appealing to perfect knowledge of the whole may resolve some of the difficulties of the classical approaches.

VII. ENVIRONMENT AS A WITNESS

The emergence of classicality can be viewed either as a consequence of the widespread dissemination of the information about the pointer states through the environment, or as a result of the censorship imposed by decoherence. So far I have focused on this second view, defining existence as persistence—predictability in spite of the environmental monitoring. The predictability sieve is a way of discovering states that are classical in this sense (Zurek, 1993a, 1993b; Zurek, Habib, and Paz, 1993; Gallis, 1996).

A complementary approach focuses not on the system, but on the records of its state spread throughout the environment. Instead of seeking the least-perturbed states one can ask what states of the system are easiest to discover by looking at the environment. Thus the environment is no longer just a source of decoherence, but acquires the role of a communication channel with basis-dependent noise that is minimized by the preferred pointer states.

This approach can be motivated by the old dilemma: On one hand, quantum states of isolated systems are purely "epistemic" (see, for example, Peres, 1993; Fuchs and Peres, 2000). Quantum cryptography (Bennett and DiVincenzo, 2000; Nielsen and Chuang, 2000, and references therein) uses this impossibility of determining the unknown state of an isolated quantum system. On the other hand, classical reality seems to be made up of quantum building blocks: States of macroscopic systems exist objectively; they can be determined by many observers independently, without being destroyed or reprepared. So the question arises: How can objective existence—the "reality" of the classical states—emerge from purely epistemic wave functions?

There is not much one can do about this in the case of a single state of an isolated quantum system. But open systems are subject to einselection and can bridge the chasm dividing their epistemic and ontic roles. The most direct way to see this arises from the recognition of the fact that we never directly observe any system. Rather, we discover states of macroscopic systems from the imprints they make on the environment: A small fraction of the photon environment intercepted by our eyes is often all that is needed. States that are recorded most redundantly in the rest of the universe (Zurek, 1983, 1998a, 2000) are also the easiest to discover. They can be found out indirectly, from multiple copies of the evidence imprinted in the environment, without a threat to their existence. Such states exist and are real; they can be found out without being destroyed as if they were really classical.

Environmental monitoring creates an ensemble of "witness states" in the subsystems of the environment that allows one to invoke some of the methods of the statistical interpretation (Ballentine, 1970) while subverting its ideology—to work with an ensemble of objective evidence of a state of a single system. From this ensemble of witness states one can infer the state of the quantum system that has led to such "advertising." This can be done without disrupting the einselected states.

The predictability sieve selects states that entangle least with the environment. Questions about predictability simultaneously lead to states that are most redundantly recorded in the environment. Indeed, this idea is the essence of the "quantum Darwinism" we alluded to in the Introduction. The einselected pointer states are not only best at surviving the environment, they also broadcast information about themselves spreading out their "copies" throughout the rest of the universe. Amplified information is easiest to amplify. This leads to analogies with "fitness" in the Darwinian sense, and suggests looking at einselection as a sort of natural selection.

A. Quantum Darwinism

Consider the "bit-by-byte" example of Sec. IV. Spin system $\mathcal{S}$ is correlated with the environment:

$$\begin{aligned}|\psi_{\mathcal{SE}}\rangle &= a|\uparrow\rangle|00\ldots0\rangle + b|\downarrow\rangle|11\ldots1\rangle\\ &= a|\uparrow\rangle|\mathcal{E}_\uparrow\rangle + b|\downarrow\rangle|\mathcal{E}_\downarrow\rangle. \qquad (7.1)\end{aligned}$$

The basis $\{|\uparrow\rangle,|\downarrow\rangle\}$ of $\mathcal{S}$ is singled out by the redundancy of the record. This can be illustrated by rewriting the same $|\psi_{\mathcal{SE}}\rangle$,

$$\begin{aligned}|\psi_{\mathcal{SE}}\rangle &= |\odot\rangle(a|00\ldots0\rangle + b|11\ldots1\rangle)\sqrt{2}\\ &\quad + |\otimes\rangle(a|00\ldots0\rangle - b|11\ldots1\rangle)/\sqrt{2}\\ &= (|\odot\rangle|\mathcal{E}_\odot\rangle + |\otimes\rangle|\mathcal{E}_\otimes\rangle)/\sqrt{2}, \qquad (7.2)\end{aligned}$$

in terms of the Hadamard transformed $\{|\odot\rangle,|\otimes\rangle\}$.

One can find out whether $\mathcal{S}$ is $|\uparrow\rangle$ or $|\downarrow\rangle$ from a small subset of the environment bits. By contrast, states $\{|\odot\rangle,|\otimes\rangle\}$ cannot be easily inferred from the environment. States $\{|\mathcal{E}_\odot\rangle,|\mathcal{E}_\otimes\rangle\}$ are typically not even orthogonal, $\langle\mathcal{E}_\odot|\mathcal{E}_\otimes\rangle = |a|^2 - |b|^2$. And even when $|a|^2 - |b|^2 = 0$, the record in the environment is fragile. Only one relative phase distinguishes $|\mathcal{E}_\odot\rangle$ from $|\mathcal{E}_\otimes\rangle$ in that case, in contrast with multiple records of the pointer states in $|\mathcal{E}_\uparrow\rangle$ and $|\mathcal{E}_\downarrow\rangle$. Remarks that elaborate this observation follow. They correspond to several distinct measures of the analogs of the Darwinian fitness of the states.

1. Consensus and algorithmic simplicity

From the state vector $|\psi_{\mathcal{SE}}\rangle$, Eqs. (7.1) and (7.2), the observer can find the state of the quantum system just by looking at the environment. To accomplish this, the total N of the environment bits can be divided into samples of n bits each, with $1 \ll n \ll N$. These samples can then be measured using observables that are the same within each sample, but that differ between samples. They may correspond, for example, to different antipodal points in the Bloch spheres of the environment bits. In the basis $\{|0\rangle,|1\rangle\}$ (or bases closely aligned with it) the record inferred from the bits of information scattered in the environment will be easiest to come by. Thus, starting from the environment part of $|\psi_{\mathcal{SE}}\rangle$, Eq. (7.1), the observer can find out, with no prior knowledge, the state of the system. Redundancy of the record in the environment allows for a trial-and-error indirect approach while leaving the system untouched.

In particular, measurement of n environment bits in a Hadamard transform of the basis $\{|0\rangle,|1\rangle\}$, Eq. (7.2), yields a random-looking sequence of outcomes (i.e., $\{|+\rangle_1,|-\rangle_2,\ldots,|-\rangle_n\}$). This record is algorithmically random. Its algorithmic complexity is of the order of its length (Li and Vitànyi, 1993):

$$K(\langle \mathcal{E}_n | +, - \rangle) \simeq n. \tag{7.3}$$

By contrast, the algorithmic complexity of the measurement outcomes in the $\{|0\rangle, |1\rangle\}$ basis will be small:

$$K(\langle \mathcal{E}_n | 0,1 \rangle) \ll n, \tag{7.4}$$

since the outcomes will be either 00 . . . 0 or 11 . . . 1. The observer seeking the preferred states of the system by looking at the environment should then search for the minimal record size and thus, for the maximum redundancy in the environmental record. States of the system that are recorded redundantly in the environment must have survived repeated instances of environment monitoring, and are obviously robust and predictable.

The predictability we have utilized before to devise a sieve to select preferred states is used here again, but in a different guise. Rather than search for predictable sets of states of the system, we are now looking for the records of the states of the system in the environment. Sequences of states of environment subsystems correlated with pointer states are mutually predictable and hence, collectively algorithmically simple. States that are predictable in spite of interactions with the environment are also easiest to predict from their impact on its state.

The state of the form of Eq. (7.1) can serve as an example of amplification. The generation of redundancy through amplification brings about the objective existence of the otherwise subjective quantum states. States $|\uparrow\rangle$ and $|\downarrow\rangle$ of the system can be determined reliably from a small fraction of the environment. By contrast, to determine whether the system was in a state $|\odot\rangle$ or $|\otimes\rangle$ one would need to detect all of the environment. Objectivity can be defined as the ability of many observers to reach consensus independently. Such consensus concerning states $|\uparrow\rangle$ and $|\downarrow\rangle$ is easily established—many $(\sim N/n)$ observers can independently measure fragments of the environment.

2. Action distance

One measure of the robustness of environmental records is the *action distance* (Zurek, 1998a). It is given by the total action necessary to undo the distinction between the states of the environment corresponding to different states of the system, subject to the constraints arising from the fact that the environment consists of subsystems. Thus to obliterate the difference between $|\mathcal{E}_\uparrow\rangle$ and $|\mathcal{E}_\downarrow\rangle$ in Eq. (7.1), one needs to "flip" one by one N subsystems of the environment. That implies an action, i.e., the least total angle by which a state must be rotated, see Sec. II.B, of

$$\Delta(|\mathcal{E}_\uparrow\rangle, |\mathcal{E}_\downarrow\rangle) = N\left[\frac{\pi}{2}\cdot\hbar\right]. \tag{7.5}$$

By contrast a flip of phase of just one bit will reverse the correspondence between the states of the system and those of the environment superpositions that make up $|\mathcal{E}_\odot\rangle$ and $|\mathcal{E}_\otimes\rangle$ in Eq. (7.2). Hence

$$\Delta(|\mathcal{E}_\odot\rangle, |\mathcal{E}_\otimes\rangle) = 1\left[\frac{\pi}{2}\cdot\hbar\right]. \tag{7.6}$$

Given a fixed division of the environment into subsystems the action distance is a metric on the Hilbert space (Zurek, 1998a). That is,

$$\Delta(|\psi\rangle, |\psi\rangle) = 0, \tag{7.7}$$

$$\Delta(|\psi\rangle, |\varphi\rangle) = \Delta(|\varphi\rangle, |\psi\rangle) \geqslant 0, \tag{7.8}$$

and the triangle inequality

$$\Delta(|\psi\rangle, |\varphi\rangle) + \Delta(|\varphi\rangle, |\gamma\rangle) \geqslant \Delta(|\psi\rangle, |\gamma\rangle) \tag{7.9}$$

are all satisfied.

In defining Δ it is essential to restrict rotations to the subspaces of the subsystems of the whole Hilbert space, and to insist that the unitary operations used in defining distance act on these subspaces. It is possible to relax constraints on such unitary operations by allowing, for example, pairwise or even more complex interactions between subsystems. Clearly, in the absence of any restrictions the action required to rotate any $|\psi\rangle$ into any $|\varphi\rangle$ would be no more than $(\pi/2)\hbar$. Thus the constraints imposed by the natural division of the Hilbert space of the environment into subsystems play an essential role. The preferred states of the system can be sought by extremizing the action distance between the corresponding record states of the environment. In simple cases [e.g., see "bit-by-byte," Eq. (4.7), and below] the action distance criterion for preferred states coincides with the predictability sieve definition (Zurek, 1998a).

3. Redundancy and mutual information

The most direct measure of the reliability of the environment as a witness is the information-theoretic redundancy of einselection itself. When the environment monitors the system (see Fig. 4), the information about its state will spread to more and more subsystems of the environment. This can be represented by the state vector $|\psi_{\mathcal{S}\mathcal{E}}\rangle$, Eq. (7.1), with increasingly long sequences of 0's and 1's in the record states. The record size, the number N of the subsystems of the environment involved, does not affect the density matrix of the system $\mathcal{S}$. Yet, it obviously changes the accessibility and robustness of the information analogs of the Darwinian fitness. As an illustration, let us consider c-shifts. One subsystem of the environment (say, $\mathcal{E}_1$) with the dimension of the Hilbert space no less than that of the system,

$$\dim \mathcal{H}_{\mathcal{E}_1} \geqslant \dim \mathcal{H}_{\mathcal{S}},$$

suffices to eradicate the off-diagonal elements of $\rho_{\mathcal{S}}$ in the control basis. On the other hand, when N subsystems of the environment correlate with the same set of states of $\mathcal{S}$, the information about these states is simultaneously accessible more widely. While $\rho_{\mathcal{S}}$ is no longer changing, spreading of the information makes the existence of the pointer states of $\mathcal{S}$ more objective—they are easier to discover without being perturbed.

Information-theoretic *redundancy* is defined as the difference between the least number of bits needed to uniquely specify the message and the actual size of the encoded message. Extra bits allow for detection and correction of errors (Cover and Thomas, 1991). In our case, the message is the state of the system, and the channel is the environment. The information about the system will often spread over all of the Hilbert space $\mathcal{H}_{\mathcal{E}}$, which is enormous compared to $\mathcal{H}_{\mathcal{S}}$. The redundancy of the record of the pointer observables of selected systems can also be huge. Moreover, typical environments consist of obvious subsystems (i.e., photons, atoms, etc.). It is then useful to define the redundancy of the record by the number of times the information about the system has been copied, or by how many times it can be independently extracted from the environment.

In the simple example of Eq. (7.1) such a *redundancy ratio* $\mathcal{R}$ for the $\{|\uparrow\rangle,|\downarrow\rangle\}$ basis will be given by N, the number of environment bits perfectly correlated with the obviously preferred basis of the system. More generally, but in the same case of perfect correlation, we obtain

$$\mathcal{R}=\frac{\ln(\dim\mathcal{H}_{\mathcal{E}})}{\ln(\dim\mathcal{H}_{\mathcal{S}})}=\log_{\dim \mathcal{H}_{\mathcal{S}}}\dim\mathcal{H}_{\mathcal{E}}=N, \tag{7.10}$$

where $\mathcal{H}_{\mathcal{E}}$ is the Hilbert space of the environment perfectly correlated with the pointer states of the system.

On the other hand, with respect to the $\{|\odot\rangle,|\otimes\rangle\}$ basis, the redundancy ratio for $|\psi_{\mathcal{SE}}\rangle$ of Eq. (7.2) is only ~ 1 (see also Zurek, 1983, 2000). Redundancy measures the number of errors that can obliterate the difference between two records, and in this basis one phase flip is clearly enough. This basis dependence of redundancy suggests an alternative strategy to seeking preferred states.

To define $\mathcal{R}$ in general we can start with mutual information between the subsystems of the environment $\mathcal{E}_k$ and the system $\mathcal{S}$. As we have already seen in Sec. IV, the definition of mutual information in quantum mechanics is not straightforward. The basis-independent formula

$$\mathcal{I}_k=\mathcal{I}(\mathcal{S}:\mathcal{E}_k)=H(\mathcal{S})+H(\mathcal{E}_k)-H(\mathcal{S},\mathcal{E}_k) \tag{7.11}$$

is simple to evaluate [although it does have some strange features; see Eqs. (4.30)–(4.36)]. In the present context it involves the joint density matrix

$$\rho_{\mathcal{S}\mathcal{E}_k}=\mathrm{Tr}_{\mathcal{E}/\mathcal{E}_k}\rho_{\mathcal{SE}}, \tag{7.12}$$

where the trace is carried out over all of the environment except for its singled-out fragment $\mathcal{E}_k$. In the example of Eq. (7.1), for any of the environment bits,

$$\rho_{\mathcal{S}\mathcal{E}_k}=|a|^2|\uparrow\rangle\langle\uparrow||0\rangle\langle 0|+|b|^2|\downarrow\rangle\langle\downarrow||1\rangle\langle 1|.$$

Given the partitioning of the environment into subsystems, the redundancy ratio can be defined as

$$\mathcal{R}_{\mathcal{I}(\{\otimes\mathcal{H}_{\mathcal{E}_k}\})}=\sum_k \mathcal{I}(\mathcal{S}:\mathcal{E}_k)/H(\mathcal{S}). \tag{7.13}$$

When $\mathcal{R}$ is maximized over all of the possible partitions,

$$\mathcal{R}_{\mathcal{I}\max}=\max_{\{\otimes\mathcal{H}_{\mathcal{E}_k}\}}\mathcal{R}_{\{\otimes\mathcal{H}_{\mathcal{E}_k}\}} \tag{7.14}$$

is obtained. Roughly speaking, $\mathcal{R}_{\mathcal{I}\max}$ is the total number of copies of the information about (the optimal basis of) $\mathcal{S}$ that exist in $\mathcal{E}$. The maximal redundancy ratio $\mathcal{R}_{\mathcal{I}\max}$ is of course basis independent.

The information defined through the symmetric $\mathcal{I}_k$, Eq. (7.11), is in general inaccessible to observers who interrogate the environment one subsystem at a time (Zurek, 2003a). It therefore makes sense to consider the basis-dependent locally accessible information and define the corresponding redundancy ratio $\mathcal{R}_{\mathcal{J}}$ using

$$\mathcal{J}_k=\mathcal{J}(\mathcal{S}:\mathcal{E}_k)=H(\mathcal{S})+H(\mathcal{E}_k)-[H(\mathcal{S})+H(\mathcal{E}_k|\mathcal{S})]. \tag{7.15}$$

The conditional entropy must be computed in a specific basis of the system [see Eq. (4.32)]. All of the other steps that have led to the definition of $\mathcal{R}_{\mathcal{I}\max}$ can now be repeated using $\mathcal{J}_k$. In the end, a basis dependent

$$\mathcal{R}_{\mathcal{J}}(\{|s\rangle\})=\mathcal{R}_{\mathcal{J}}(\otimes\mathcal{H}_{\mathcal{E}_k}) \tag{7.16}$$

is obtained. $\mathcal{R}_{\mathcal{J}}(\{|s\rangle\})$ quantifies the mutual information between the collection of subsystems $\mathcal{H}_{\mathcal{E}_k}$ of the environment and the basis $\{|s\rangle\}$ of the system. We note that the condition of nonoverlapping partitions guarantees that all of the corresponding measurements commute, and that the information can indeed be extracted independently from each environment fragment $\mathcal{E}_k$.

The preferred basis of $\mathcal{S}$ can now be defined by maximizing $\mathcal{R}_{\mathcal{J}}(\{|s\rangle\})$ with respect to the selection of $\{|s\rangle\}$:

$$\mathcal{R}_{\mathcal{J}\max}=\max_{\{|s\rangle\};\{\otimes\mathcal{H}_{\mathcal{E}_k}\}}\mathcal{R}_{\mathcal{J}}(\{|s\rangle\}). \tag{7.17}$$

This maximum can be sought either by varying the basis of the system only or (as is indicated above) by varying both the basis and the partition of the environment.

It remains to be seen whether and under what circumstances the pointer basis "stands out" through its definition in terms of $\mathcal{R}_{\mathcal{J}}$. The criterion for a well-defined set of pointer states $\{|p\rangle\}$ would be

$$\mathcal{R}_{\mathcal{J}\max}=\mathcal{R}_{\mathcal{J}}(\{|p\rangle\})\gg\mathcal{R}_{\mathcal{J}}(\{|s\rangle\}), \tag{7.18}$$

where $\{|s\rangle\}$ are typical superpositions of states belonging to different pointer eigenstates.

This definition of preferred states directly employs the notion of multiplicity of records available in the environment. Since $\mathcal{J}\leq\mathcal{I}$, it follows that

$$\mathcal{R}_{\mathcal{J}\max}\leq\mathcal{R}_{\mathcal{I}\max}. \tag{7.19}$$

The important feature of either version of $\mathcal{R}$ that makes them useful for our purpose is their independence on $H(\mathcal{S})$. The dependence on $H(\mathcal{S})$ is in effect normalized out of $\mathcal{R}$. $\mathcal{R}$ characterizes the fan-out of information about the preferred basis throughout the environment, without reference to what is known about the system. The usual redundancy (in bits) is then $\sim\mathcal{R}\cdot H(\mathcal{S})$, although other implementations of this program (Ollivier, Poulin, and Zurek, 2002) employ different measures of redundancy, which may be even more accurate than the redundancy ratio we have described above. Indeed, what is important here is the general idea of measuring the classicality of quantum states through the number of copies they imprint throughout the universe. This is a very Darwinian approach. We define classicality related to einselection in ways reminiscent of "fitness" in natural selection: states that spawn most of the (information-theoretic) progeny are the most classical.

4. Redundancy ratio rate

The rate of change of redundancy is of interest as another measure of "fitness," perhaps closest to the definitions of fitness used in modeling natural selection. Redundancy can increase either as a result of interactions between the system and the environment, or because the environment already correlated with $\mathcal{S}$ is passing on the information to more distant environments. In this second case "genetic information" is passed on by the "progeny" of the original state. Even an observer consulting the environment becomes a part of such a more-distant environment. The redundancy rate is defined as

$$\dot{\mathcal{R}}=\frac{d}{dt}\mathcal{R}. \tag{7.20}$$

Either basis-dependent or basis-independent versions of $\dot{\mathcal{R}}$ may be of interest.

In general, it may not be easy to compute either $\mathcal{R}$ or $\dot{\mathcal{R}}$ exactly. This is nevertheless possible in models [such as those leading to Eqs. (7.1) and (7.2)]. The simplest illustrative example corresponds to the c-NOT model of decoherence in Fig. 4. One can imagine that the consecutive record bits get correlated with the two branches (corresponding to $|0\rangle$ and $|1\rangle$ in the control) at discrete moments of time. $\mathcal{R}(t)$ would then be the total number of c-NOT's that have acted over time t, and $\dot{\mathcal{R}}$ is the number of new c-NOT's added per unit time.

The redundancy rate measures information flow from the system to the environment. Note that, after the first c-NOT in the example of Eqs. (7.1) and (7.2), $\mathcal{R}_{\mathcal{I}}$ will jump immediately from 0 to 2 bits, while the basis specific $\mathcal{R}_{\mathcal{J}}$ will increase from 0 to 1. In our model this initial discrepancy [which reflects quantum discord, Eq. (4.36), between $\mathcal{I}$ and $\mathcal{J}$] will disappear after the second c-NOT.

Finally, we note that $\mathcal{R}$ and, especially, $\dot{\mathcal{R}}$ can be used to introduce new predictability criteria: The states (or the observables) that are being recorded most redundantly are the obvious candidates for the objective states, and therefore for the classical states.

B. Observers and the existential interpretation

von Neumann (1932), London and Bauer (1939), and Wigner (1963) have all appealed to the special role of the conscious observer. Consciousness was absolved from following unitary evolution, and thus, could collapse the wave packet. Quantum formalism has led us to a different view that nevertheless allows for a compatible conclusion. In essence, macroscopic systems are open, and their evolution is almost never unitary. Records maintained by the observers are subject to einselection. In a binary alphabet decoherence will allow for only the two logical states and prohibit their superpositions (Zurek, 1991). For human observers, neurons conform to this binary convention, and the decoherence times are short (Tegmark, 2000). Thus, even if a cell of the observer entangles through a premeasurement with a pure quantum state, the record will become effectively classical almost instantly. As a result, it will be impossible to "read it off" in any basis except for the einselected one. This censorship of records is the key difference between the existential interpretation and Everett's original many-worlds interpretation.

Decoherence treats the observer as any other macroscopic quantum system. There is, however, one feature distinguishing observers from the rest of the universe: They are aware of the content of their memory. Here we are using "aware" in a down-to-earth sense: Quite simply, observers know what they know. Their information-processing machinery (which must underlie higher functions of the mind such as "consciousness") can readily consult the contents of their memory.

The information stored in memory comes with strings attached. The physical state of the observer is described in part by the data in his records. There is no information without representation. The information the observer has could be, in principle, deduced from his physical state. The observer is, in part, information. Moreover, this information encoded in states of macroscopic quantum systems (neurons) is by no means secret. As a result of the lack of isolation, the environment, having redundant copies of the relevant data, knows in detail everything the observer knows. Configurations of neurons in our brains, while at present undecipherable, are, in principle, as objective and as widely accessible as the information about the states of other macroscopic objects.

The observer is what he knows. In the unlikely case of a flagrantly quantum input the physical state of the observer's memory will decohere, resulting almost instantly in the einselected alternatives, each of them representing simultaneously both the observer and his memory. The "advertising" of this state throughout the environment makes it effectively objective.

An observer perceiving the universe from within is in a very different position than an experimental physicist studying a state vector of a quantum system. In a laboratory, the Hilbert space of the investigated system is typically tiny. Such systems can be isolated, so that often the information loss to the environment can be prevented. Then the evolution is unitary. The experimentalist can know everything there is to know about it.

Common criticisms of the approach advocated in this paper are based on an unjustified extrapolation of the above laboratory situation to the case of the observer who is a part of the universe. Critics of decoherence often note that the differences between the laboratory example above and the case of the rest of the universe are merely quantitative: the system under investigation is bigger, etc. So why cannot one analyze, they ask, interactions of the observer and the rest of the universe as before, for a small isolated quantum system?

In the context of the existential interpretation the analogy with the laboratory is, in effect, turned upside down: For, now the observer (or the apparatus, or anything effectively classical) is continuously monitored by the rest of the universe. Its state is repeatedly forced into the einselected states, and very well (very redundantly) known to the rest of the universe.

The higher functions of observers, e.g., consciousness, etc., may be at present poorly understood, but it is safe to assume that they reflect physical processes in the information-processing hardware of the brain. Hence mental processes are in effect objective, since they must reflect conditional quantum dynamics of open system—observer's network of neurons—and, hence leave an indelible imprint on the environment. The observer has no chance of perceiving either his memory, or any other macroscopic part of the universe in some arbitrary superposition. Moreover, the memory capacity of observers is miniscule compared to the information content of the universe. So, while observers may know the exact state of the laboratory systems, their records of the universe will be very fragmentary. By contrast, the universe has enough memory capacity to acquire and maintain detailed records of the states of macroscopic systems and their histories. Thus, indeed, it appears that consciousness does not follow a unitary quantum evolution, as the conditional dynamics that implements such "higher functions" must be subject to decoherence and einselection (see also Tegmark, 2000). As promised, we have in a sense recovered postulates of von Neumann, London and Bauer, and Wigner, and we have done that without involing any "extraphysical" postulates.

C. Events, records, and histories

Suppose that instead of a monotonous record sequence in the environment basis corresponding to the pointer states of the system $\{|\uparrow\rangle,|\downarrow\rangle\}$ implied by Eq. (7.1) the observer looking at the environment detects

$$000\ldots0111\ldots1000\ldots0111\ldots.$$

Given the appropriate additional assumptions, such sequences consisting of long stretches of record 0's and 1's justify inference of the history of the system. Let us further assume that the observer's records come from intercepting a small fragment of the environment. Other observers will then be able to consult their independently accessible environmental records, and will infer (more or less) the same history. Thus, in view of the preponderance of evidence, history defined as a sensible inference from the available records can be probed by many observers independently, and can be regarded as classical and objective.

The redundancy ratio of the records $\mathcal{R}$ is a measure of this objectivity. Note that this relatively objective existence (Zurek, 1998a) is an operational notion, quantified by the number of times the state of the system can be determined independently, and not some absolute objectivity. However, and in a sense that can be rigorously defined, relative objectivity tends to absolute objectivity in the limit $\mathcal{R}\rightarrow\infty$. For example, cloning of unknown states becomes possible (Bruss, Ekert, and Macchiavello, 1999; Jozsa, 2002) in spite of the no-cloning theorem (Dieks 1982; Wootters and Zurek, 1982). In that limit, and given the same reasonable constraints on the nature of the interactions and on the structure of the environment that underlies that definition of $\mathcal{R}$, it would take infinite resources such as action, Eqs. (7.5)–(7.9), to hide or subvert evidence of such an objective history.

There are differences and parallels between the relatively objective histories introduced here and the consistent histories proposed by Griffiths (1984, 1996), and investigated by Gell-Mann and Hartle (1990, 1993, 1997), Omnès (1988, 1992, 1994), Halliwell (1999), and others (Dowker and Kent, 1996; Kiefer, 1996). Such histories are defined as time-ordered sequences of projection operators $P^1_{\alpha_1}(t_1),P^2_{\alpha_2}(t_2),\ldots,P^n_{\alpha_n}(t_n)$ and are abbreviated $[P_\alpha]$. Consistency is achieved when they can be combined into coarse-grained sets (where the projectors defining a coarse-grained set are given by the sums of the projectors in the original set) while obeying probability sum rules: The probability of a bundle of histories should be a sum of the probabilities of the constituent histories. The corresponding condition can be expressed in terms of the *decoherence functional* (Gell-Mann and Hartle, 1990):

$$D([P_\alpha],[P_\beta]) = \mathrm{Tr}[P^n_{\alpha_n}(t_n)\ldots P^1_{\alpha_1}(t_1)\rho P^1_{\beta_1}(t_1)\ldots P^n_{\beta_n}(t_n)]. \tag{7.21}$$

Above, the state of the system of interest is described by the density matrix ρ. Griffiths' condition is equivalent to the vanishing of the real part of the expression above, $\mathrm{Re}\{D([P_\alpha],[P_\beta])\}=p_\alpha\delta_{\alpha,\beta}$. As Gell-Mann and Hartle (1990) emphasize, it is more convenient, and in the context of an emergent classicality more realistic, to require instead that $D([P_\alpha],[P_\beta])=p_\alpha\delta_{\alpha,\beta}$. Both weaker and stronger conditions for the consistency of histories were considered (Goldstein and Page, 1995; Gell-Mann and Hartle, 1997). The problem with all of them is that the resulting histories are very subjective: Given an initial density matrix of the universe it is in general quite easy to specify many different, mutually incompatible consistent sets of histories. This subjectivity leads to serious interpretational problems (d'Espagnat, 1989, 1995; Dowker and Kent, 1996). Thus a demand for exact consistency as one of the conditions for classicality is both uncomfortable (overly restrictive) and insufficient (since the resulting histories are very nonclassical). Moreover, coarse grainings that help secure approximate consistency have to be, in effect, guessed at.

The attitudes adopted by the practitioners of the consistent-histories approach in view of its unsuitability for the role of the cornerstone of emergent classicality differ. Initially, before difficulties became apparent, it was hoped that such an approach would answer all of the interpretational questions, perhaps when supplemented by a subsidiary condition, i.e., some assumption about favored coarse grainings. At present, some still aspire to the original goals of deriving classicality from consistency alone. Others may uphold the original aims of the program, but they also generally rely on environment-induced decoherence, using in calculations variants of models we have presented in this paper. This strategy has been quite successful—after all, decoherence leads to consistency. For instance, the special role of the hydrodynamic observables (Gell-Mann and Hartle, 1990; Dowker and Halliwell, 1992; Brun and Hartle, 1999; Halliwell, 1999) can be traced to their predictability, or to their approximate commutativity with the total Hamiltonian [see Eq. (4.41)]. On the other hand, the original goals of Griffiths (1984, 1996) have been more modest. Using consistent histories, one can discuss the sequences of events in an evolving quantum system without logical contradictions. The "golden middle" is advocated by Griffiths and Omnès (1999) who regard consistent histories as a convenient language, rather than as an explanation of classicality.

The origin of effective classicality can be traced to decoherence and einselection. As was noted by Gell-Mann and Hartle (1990), and elucidated by Omnès (1992, 1994) decoherence suffices to ensure approximate consistency. But consistency is both not enough and too much; it is too easy to accomplish, and does not necessarily lead to classicality (Dowker and Kent, 1996). What is needed instead is the objectivity of events and their time-ordered sequences—their histories. As we have seen above, both can appear as a result of einselection.

We have already provided an operational definition of the relatively objective existence of quantum states. It is easy to apply it to events and histories: When many observers can independently gather compatible evidence concerning an event, we call it relatively objective. Relatively objective history is then a time-ordered sequence of relatively objective events.

Monitoring of the system by the environment leads to decoherence and einselection. It will also typically lead to redundancy and hence to an effectively objective classical existence in the sense of quantum Darwinism. Observers can independently access redundant records of events and histories imprinted in the environmental degrees of freedom. The number of observers who can examine evidence etched in the environment can be of the order of, and is bounded from above by, $\mathcal{R}_{\mathcal{J}}$. Redundancy is a measure of this objectivity and classicality.

As observers record their data, $\mathcal{R}_{\mathcal{J}}$ changes. Consider an observer who measures the "right observable" of $\mathcal{E}$ [i.e., the one with the eigenstates $|0\rangle,|1\rangle$ in the example of Eq. (7.1)]. Then his records and, as his records decohere, also their environment, become a part of the evidence, and are correlated with the preferred basis of the system. Consequently $\mathcal{R}_{\mathcal{J}}$ computed from Eq. (7.14) increases. Every interaction that increases the number of records also increases $\mathcal{R}_{\mathcal{J}}$. This is obvious for the "primary" interactions with the system, but it is also true for the secondary, tertiary, etc., acts of replication of the information obtained from the observers who recorded the primary state of the system, from the environment, from the environment of the environment, and so on.

A measurement reveals to the observer his branch of the universal state vector. The correlations established alter the observer's state, his records, and "attach" him to this branch. He will share it with other observers who examined the same set of observables, and who have recorded compatible results.

It is also possible to imagine a stubborn observer who insists on measuring either the relative phase between the two obvious branches of the environment in Eq. (7.2), or the state of the environment in the Hadamard-transformed basis $\{|+\rangle,|-\rangle\}$. In either case the distinction between the two outcomes could determine the state of the spin in the $\{|\odot\rangle,|\otimes\rangle\}$ basis. However, in that basis $\mathcal{R}_{\mathcal{J}}=1$. Hence, while, in principle, these measurements can be carried out and yield the correct result, the information concerning the $\{|\odot\rangle,|\otimes\rangle\}$ basis is not redundant and therefore not objective: Only one stubborn observer can access it directly. As a result $\mathcal{R}_{\mathcal{J}}$ will decrease. Whether $\mathcal{R}_{\mathcal{J}}(\{|\odot\rangle,|\otimes\rangle\})$ will become larger than $\mathcal{R}_{\mathcal{J}}(\{|\uparrow\rangle,|\downarrow\rangle\})$ was before the measurement of the stubborn observer will depend on a detailed comparison of the initial redundancy with the amplification involved, the decoherence, etc.

There is a further significant difference between the two stubborn observers considered above. When the observer measures the phase between the two sequences of 0's and 1's in Eq. (7.2), correlations between the bits of the environment remain. Thus, even after his measurement, one could find relatively objective evidence of the past event—the past state of the spin—and, in more complicated cases, of the history. On the other hand, measurement of all the environment bits in the $\{|+\rangle,|-\rangle|$ basis will obliterate evidence of such a past.

The relatively objective existence of events is the strongest condition we have considered here. It is a consequence of the existence of multiple records of the same set of states of the system. It allows for such manifestations of classicality as unimpeded cloning. It implies einselection of states most closely monitored by the environment. Decoherence is clearly weaker and easier to accomplish.

"The past exists only insofar as it is recorded in the present" (a dictum often repeated by Wheeler) may the best summary of the above discussion. The relatively objective reality of a few selected observables in our familiar universe is measured by their "Darwinian" fitness—by the redundancy with which they are recorded in the environment. This multiplicity of available copies of the same information can be regarded as a consequence of amplification, and as a cause of indelibility. Multiple records safeguard the objectivity of our past.

VIII. DECOHERENCE IN THE LABORATORY

The biggest obstacle in the experimental study of decoherence is, paradoxically, its effectiveness. In the macroscopic domain only the einselected states survive. Their superpositions are next to impossible to prepare. In the mesoscopic regime one may hope to adjust the size of the system, and thus, interpolate between quantum and classical. The strength of the coupling to the environment is the other parameter one may employ to control the decoherence rate.

One of the key consequences of monitoring by the environment is the inevitable introduction of the Heisenberg uncertainty into the observable complementary to the one that is monitored. One can simulate such uncertainty without any monitoring environment by introducing classical noise. In each specific run of the experiment, for each realization of time-dependent noise, the quantum system will evolve deterministically. However, after averaging over different noise realizations, as it is discussed in Sec. IV.C, the evolution of the density matrix describing an ensemble of systems may approximate decoherence due to an entangling quantum environment. In particular, the master equation may be essentially the same as that for true decoherence, although the interpretational implications are more limited. Yet, using such strategies one can simulate much of the dynamics of open quantum systems.

The strategy of simulating decoherence can be taken further: Not just the effect of the environment, but also the dynamics of the quantum system can be simulated by classical means. This can be accomplished when classical wave phenomena follow equations of motion related to the Schrödinger equation. We shall discuss experiments that fall into all of the above categories.

Last but not least, while decoherence—through einselection—helps solve the measurement problem, it is also a major obstacle to quantum information processing. We shall thus end this section briefly describing strategies that may allow one to tame decoherence.

A. Decoherence due to entangling interactions

Several experiments fit this category, and more have been proposed. Decoherence due to emission or scattering of photons has been investigated by the MIT group of David Pritchard (Chapman *et al.*, 1995) using atomic interferometry. Emission or scattering deposits a record in the environment. It can store information about the path of the atom providing the photon wavelength is shorter than the separation between two of the atoms. In the case of emission this record is not redundant, since the atom and photon are simply entangled, $\mathcal{R}_{\mathcal{J}} \sim 1$, in any basis. Scattering may involve more photons, and a recent careful experiment (Kokorowski *et al.*, 2001) has confirmed the saturation of decoherence rate at distances in excess of the photon wavelength (Gallis and Fleming, 1990; Anglin, Paz, and Zurek, 1997).

There is an intimate connection between interference and complementarity in the two-slit experiment on one hand, and entanglement on the other (Wootters and Zurek, 1979). Consequently, appropriate measurements of the photon allow one to restore interference fringes in the conditional subensembles corresponding to a definite phase between the two photon trajectories (see especially Chapman *et al.*, 1995, as well as Kwiat, Steinberg, and Chiao, 1992; Pfau *et al.*, 1994; Herzog, Kwiat, Weinfurter, and Zeilinger, 1995, for implementations of this "quantum erasure" trick of Hillery and Scully, 1983). Similar experiments have also been carried out using neutron interferometry (see, for example, Rauch, 1998).

In all of these experiments one is dealing with a very simplified situation involving a single microsystem and a single "unit" of decoherence ($\mathcal{R}_{\mathcal{J}} \sim 1$) caused by a single quantum of the environment. Experiments on a mesoscopic system monitored by the environment are obviously much harder to devise. Nevertheless, Haroche, Raimond, Brune, and their colleagues at the Ecole Normale Supérieure (Brune *et al.*, 1996; Haroche, 1998; Raimond, Brune, and Haroche, 2001) have carried out a spectacular experiment of this type, yielding solid evidence in support of the basic tenets of the environment-induced transition from quantum to classical. Their system is a microwave cavity. It starts in a coherent state with an amplitude corresponding to a few photons.

A Schrödinger-cat state is created by introducing an atom in a superposition of two Rydberg states, $|+\rangle = |0\rangle + |1\rangle$: The atom passing through the cavity puts its refractive index in a superposition of two values. Hence the phase of the coherent state shifts by the amount correlated with the state of the atom, creating an entangled state:

$$|\rightarrow\rangle(|0\rangle+|1\rangle) \Rightarrow |\nearrow\rangle|0\rangle + |\searrow\rangle|1\rangle = |\vartheta\rangle. \tag{8.1}$$

Arrows indicate relative phase-space locations of coherent states. States of the atom are $|0\rangle$ and $|1\rangle$. The "Schrödinger kitten" is prepared from this entangled state by measuring the atom in the $\{|+\rangle,|-\rangle\}$ basis:

$$|\vartheta\rangle = (|\nearrow\rangle + |\searrow\rangle)|+\rangle + (|\nearrow\rangle - |\searrow\rangle)|-\rangle. \tag{8.2}$$

Thus the atom in the state $|+\rangle$ implies preparation of a "positive cat" $|\uplus\rangle = |\nearrow\rangle + |\searrow\rangle$ in the cavity. Such superpositions of coherent states could survive forever if there was no decoherence. However, radiation leaks out of the cavity. Hence the environment acquires information about the state inside. Consequences are tested by passing another atom in the state $|+\rangle = |0\rangle + |1\rangle$ through the cavity. In the absence of decoherence the state would evolve as

$$\begin{aligned} |\uplus\rangle|+\rangle &= (|\nearrow\rangle + |\searrow\rangle)(|0\rangle + |1\rangle) \\ &\Rightarrow (|\uparrow\rangle|0\rangle + |\rightarrow\rangle|1\rangle) \\ &\quad + (|\rightarrow\rangle|0\rangle + |\downarrow\rangle|1\rangle) \\ &= (|\uparrow\rangle|0\rangle + |\downarrow\rangle|1\rangle) + \sqrt{2}|\rightarrow\rangle|+\rangle. \end{aligned} \tag{8.3}$$

Above we have omitted the overall normalization, but retained the (essential) relative amplitude.

For the above state, detection of $|+\rangle$ in the first (preparatory) atom implies the conditional probability of detection of $|+\rangle$, $p_{+|+} = 3/4$, for the second (test) atom. Decoherence will suppress the off-diagonal terms of the density matrix so that, some time after the preparation, ρ_{cavity} that starts, say, as $|\uplus\rangle\langle\uplus|$ becomes

$$\begin{aligned} \rho_{cavity} &= (|\nearrow\rangle\langle\nearrow| + |\searrow\rangle\langle\searrow|)/2 \\ &\quad + z(|\nearrow\rangle\langle\searrow| + |\searrow\rangle\langle\nearrow|)/2. \end{aligned} \tag{8.4}$$

When $z=0$ the conditional probability is $p(+|+)=1/2$.

In the intermediate cases intermediate values of this and other relevant conditional probabilities are predicted. The rate of decoherence, and consequently, the time-dependent value of z can be estimated from the cavity quality factor Q, and from the data about the coherent state initially present in the cavity. The decoherence rate is a function of the separation of the two components of the cat $|\uplus\rangle$. Experimental results agree with predictions.

The discussion above depends on the special role of coherent states. Coherent states are einselected in harmonic oscillators, and hence, in underdamped bosonic fields (Anglin and Zurek, 1996). Thus they are the pointer states of the cavity. Their special role is recognized implicitly above: If number eigenstates were einselected, predictions would obviously be quite different. Therefore, while the Ecole Normale Supérieure experiment is focused on the decoherence rate, confirmation of the predicted special role of coherent states in bosonic fields is its important (albeit implicit) corollary.

B. Simulating decoherence with classical noise

From the fundamental point of view, the distinction between cases in which decoherence is caused by entangling interactions with the quantum state of the environment and in which it is simulated by classical noise in the observable complementary to the pointer is essential. However, from the engineering point of view (adopted, for example, by the practitioners of quantum computation, see Nielsen and Chuang, 2000, for a discussion) this may not matter. For instance, quantum error-correction techniques (Shor, 1995; Steane, 1996; Preskill, 1999) are capable of dealing with either. Moreover, experimental investigations of this subject often involve both.

The classic experiment in this category was carried out recently by Wineland, Monroe, and their collaborators (Myatt *et al.*, 2000; Turchette *et al.*, 2000). They used an ion trap to study the behavior of individual ions in a Schrödinger-cat state (Monroe *et al.*, 1996) under the influence of injected classical noise. They also embarked on a preliminary study of "environment engineering."

Superpositions of two coherent states as well as of number eigenstates were subjected to simulated high-temperature amplitude and phase "reservoirs." This was done through time-dependent modulation of the self-Hamiltonian of the system. For the amplitude noise these are, in effect, random fluctuations of the location of the minimum of the harmonic trap. Phase noise corresponds to random fluctuations of the trap frequency.

In either case, the resulting loss of coherence is well described by the exponential decay with time, with an exponent that scales with the square of the separation between the two components of the macroscopic quantum superposition [e.g., Eq. (5.34)]. The case of the amplitude noise approximates decoherence in quantum Brownian motion in that the coordinate is monitored by the environment, and hence, the momentum is perturbed. (Note that in an underdamped harmonic oscillator the rotating-wave approximation blurs the distinction between x and p, leading to einselection of coherent states.) The phase noise would arise in an environment monitoring the number operator, thus leading to uncertainty in phase. Consequently, number eigenstates are einselected.

The applied noise is classical, and the environment does not acquire any information about the ion ($\mathcal{R}_{\mathcal{I}}=0$). Thus, following a particular realization of the noise, the state of the system is still pure. Nevertheless, an ensemble average over many noise realizations is represented by the density matrix that follows an appropriate master equation. Thus, as Wineland, Monroe, and their colleagues note, decoherence simulated by classical noise could be in each individual case—for each realization—reversed by simply measuring the corresponding time-dependent noise either beforehand or afterwards, and then applying the appropriate unitary transformation to the state of the system. By contrast, in the case of entangling interactions, two measurements, one preparing the environment before the interaction with the environment, the other following it, would be the least required for a chance of undoing the effect of decoherence.

The same two papers study the decay of a superposition of number eigenstates $|0\rangle$ and $|2\rangle$ due to an indirect coupling with the vacuum. This proceeds through entanglement with the first-order environment (that, in effect, consists of the other states of the harmonic oscillator) and a slower transfer of information to the distant environment. Dynamics involving the system and its first-order environment lead to nonmonotonic behavior of the off-diagonal terms representing coherence. Further studies of decoherence in the ion-trap setting are likely to follow, since this is an attractive implementation of the quantum computer (Cirac and Zoller, 1995).

1. Decoherence, noise, and quantum chaos

Following a proposal of Graham, Schlautmann, and Zoller (1993) Raizen and his group (Moore *et al.*, 1994) used a one-dimensional (1D) optical lattice to implement a variety of 1D chaotic systems including the "standard map." Various aspects of the behavior expected from a quantized version of a classically chaotic system were subsequently found, including, in particular, dynamical localization (Reichl, 1992; Casati and Chirikov, 1995a).

Dynamical localization establishes, in a class of driven quantum chaotic systems, a saturation of momentum dispersion, and leads to a characteristic exponential form of its distribution (Casati and Chirikov, 1995a). Localization is obviously a challenge to the quantum-classical correspondence, since in these very same systems the classical prediction has the momentum dispersion growing unbounded, more or less with the square root of time. However, localization sets in after $t_L \sim \hbar^{-\alpha}$, where $\alpha \sim 1$ (rather than on the much shorter

$t_\hbar \sim \ln \hbar^{-1}$ that we have discussed in Sec. III) so it can be ignored for macroscopic systems. On the other hand, its signature is easy to detect.

Demonstration of dynamical Anderson localization in the optical-lattice implementation of the δ-kicked rotor and related studies of quantum chaos have been a significant success (Moore *et al.*, 1994). More recently, the attention of both Raizen and his group in Texas as well as of Christensen and his group in New Zealand has shifted towards the effect of decoherence on quantum chaotic evolution (Ammann *et al.*, 1998; Klappauf *et al.*, 1998).

In all of the above studies the state of the chaotic system (δ-kicked rotor) was perturbed by spontaneous emission from the trapped atoms, which was induced by decreasing the detuning of the lasers used to set up the optical lattice. In addition, noise was occasionally introduced into the potential. Both groups found that, as a result of spontaneous emission, localization disappears, although the two studies differ in some of the details. More experiments, including some that allow gentler forms of monitoring by the environment (rather than spontaneous emission noise) appear to be within reach.

In all of the above cases one deals, in effect, with a large ensemble of identical atoms. While each atom suffers repeated disruptions of its evolution due to spontaneous emission, the ensemble evolves smoothly and in accord with the appropriate master equation. The situation is reminiscent of decoherence simulated by noise. Indeed, experiments that probed the effect of classical noise on chaotic systems were carried out earlier (Koch, 1995). They were, however, analyzed from a point of view that does not readily shed light on decoherence.

A novel experimental approach to decoherence and to irreversibility in open complex quantum systems has been pursued by Levstein, Pastawski, and their colleagues (Levstein, Usaj, and Pastawski, 1998; Levstein *et al.*, 2000). Using NMR techniques they investigated the reversibility of dynamics by implementing a version of spin echo. This promising "Loschmidt echo" approach has led to renewed interest in the issues that touch on quantum chaos, decoherence, and related subjects (see, for example, Jacquod, Silvestrov, and Beenakker, 2001; Jalabert and Pastawski, 2001; Gorin and Seligman, 2002; Prosen and Seligman, 2002).

2. Analog of decoherence in a classical system

Both the system and the environment are effectively classical in the last category of experiments, represented by the work of Cheng and Raymer (1999). They have investigated the behavior of transverse spatial coherence during the propagation of an optical beam through a dense, random dielectric medium. This problem can be modeled by a Boltzmann-like transport equation for the Wigner function of the wave field, and exhibits a characteristic increase of decoherence rate with the square of the spatial separation, followed by saturation at sufficiently large distances. This saturation contrasts with the simple models of decoherence in quantum Brownian motion that are based on a dipole approximation. However, it is in good accord with more sophisticated discussions that recognize that, for separations of the order of the prevalent wavelength in the environment, the dipole approximation fails and other more complicated behaviors can set in (Gallis and Fleming, 1990; Anglin, Paz, and Zurek, 1997; Paz and Zurek, 1999). A similar result in a completely quantum case was obtained by Kokorowski *et al.* (2001) using atomic interferometry.

C. Taming decoherence

In many of the applications of quantum mechanics the quantum nature of the information stored or processed needs to be protected. Thus decoherence is an enemy. Quantum computation is an example of this situation. A quantum computer can be thought of as a sophisticated interference device that works by performing in parallel a coherent superposition of a multitude of classical computations. Loss of coherence would disrupt the quantum parallelism essential for the expected speedup.

In the absence of the ideal—a completely isolated absolutely perfect quantum computer, something easy for a theorist to imagine but impossible to attain in the laboratory—one must deal with imperfect hardware "leaking" some of its information to the environment. And maintaining isolation while simultaneously achieving a reasonable "clock time" for the quantum computer is likely to be difficult since both are in general controlled by the same interaction [although there are exceptions; for example, in the ion-trap proposal of Cirac and Zoller (1995) interaction is in a sense "on demand," and is turned on by the laser coupling the internal states of ions with the vibrational degree of freedom of the ion chain].

The need for error correction in quantum computation was realized early on (Zurek, 1984b) but methods for accomplishing this goal have evolved dramatically from the Zeno effect suggested then to the very sophisticated (and much more effective) strategies in recent years. This is fortunate. Without error correction even fairly modest quantum computations (such as factoring the number 15 in an ion trap with imperfect control of the duration of the laser pulses) go rapidly astray as a consequence of relatively small imperfections (Miquel, Paz, and Zurek, 1997).

Three different, somewhat overlapping, approaches that aim to control and tame decoherence, or to correct errors caused by decoherence or by the other imperfections of the hardware, have been proposed. We summarize them very briefly, spelling out main ideas and pointing out references that discuss them in greater detail.

1. Pointer states and noiseless subsystems

The most straightforward strategy to suppress decoherence is to isolate the system of interest (e.g., the quantum computer). Failing that, one may try to isolate some of its observables with degenerate pointer subspaces, which then constitute niches in the Hilbert space of the information-processing system that do not get disrupted in spite of the coupling to the environment. Decoherence-free subspaces are thus identical in conception with the pointer subspaces introduced some time ago (Zurek, 1982), and satisfy (exactly or approximately) the same Eqs. (4.22) and (4.41) or their equivalents [given, for example, in terms of "Kraus operators" (Kraus, 1983)] that represent the nonunitary consequences of the interaction with the environment in the Lindblad (1976) form of the master equation. Decoherence-free subspaces were (re)discovered in the context of quantum information processing. They appear as a consequence of an exact or approximate symmetry of the Hamiltonians that govern the evolution of the system and its interaction with the environment (Zanardi and Rasetti, 1997; Duan and Guo, 1998; Lidar *et al.*, 1999; Zanardi, 1998, 2001).

An active extension of this approach aimed at finding quiet corners of the Hilbert space is known as dynamical decoupling. There the effectively decoupled subspaces are induced by time-dependent modifications of the evolution of the system deliberately introduced from the outside by time-dependent Hamiltonians and/or measurements (see, for example, Viola and Lloyd, 1998; Zanardi, 2001). A further generalization and unification of various techniques leads to the concept of noiseless quantum subsystems (Knill, Laflamme, and Viola, 2000; Zanardi, 2001), which may be regarded as a non-Abelian (and quite nontrivial) generalization of pointer subspaces.

A sophisticated and elegant strategy that can be regarded as a version of the decoherence-free approach was devised independently by Kitaev (1997a, 1997b). He has advocated using states that are topologically stable, and thus, can successfully resist arbitrary interactions with the environment. The focus here (in contrast to much of the decoherence-free subspace work) is on devising a system with a self-Hamiltonian that—as a consequence of the structure of the gap in its energy spectrum relate to the "cost" of topologically nontrivial excitations—acquires a subspace isolated *de facto* from the environment. This approach has been further developed by Bravyi and Kitaev (1998) and by Freedman and Meyer (2001).

2. Environment engineering

This strategy involves altering the (effective) interaction Hamiltonian between the system and the environment or influencing the state of the environment to selectively suppress decoherence. There are many ways to implement it, and we shall describe under this label a variety of proposed techniques (some of which are not all that different from the strategies we have just discussed) that aim to protect the quantum information stored in selected subspaces of the Hilbert space of the system, or even to exploit the pointer states induced or redefined in this fashion.

The basic question that started this line of research, whether one can influence the choice of the preferred pointer states, arose in the context of the ion-trap quantum computer proposed by Cirac and Zoller (1995). The answer given by the theory is, of course, that the choice of the einselected basis is predicated on the details of the situation and, in particular, on the nature of the interaction between the system and the environment (Zurek, 1981, 1982, 1993a). Yet Poyatos, Cirac, and Zoller (1996) have suggested a scheme suitable for implementation in an ion trap, in which interaction with the environment, and, in accord with Eq. (4.41), the pointer basis itself, can be adjusted. The key idea is to recognize that the effective coupling between the vibrational degrees of freedom of an ion (the system) and the laser light (which plays the role of the environment) is given by

$$H_{int}=\frac{\Omega}{2}(\sigma_+ e^{-i\omega_L t}+\sigma_- e^{i\omega_L t})\sin[\kappa(a+a^+)+\phi]. \tag{8.5}$$

Above, Ω is the Rabi frequency, ω_L the laser frequency, ϕ is related to the relative position of the center of the trap with respect to the laser standing wave, and κ is the Lamb-Dicke parameter of the transition, while $\sigma_-(\sigma_+)$ and $a(a^+)$ are the annihilation and creation operators of the atomic transition and of the harmonic oscillator representing ion in the trap, respectively.

By adjusting ϕ and ω_L and adopting the appropriate set of approximations (which includes the elimination of the internal degrees of freedom of the atom) one is led to the master equation for the system, i.e, for the density matrix of the vibrational degree of freedom,

$$\dot{\rho}=\gamma(2f\rho f^+ - f^+ f\rho - \rho f^+ f). \tag{8.6}$$

Above, f is an operator with a form that depends on the adjustable parameters ϕ and ω_L in H_{int}, while γ is a constant that also depends on Ω and η. As Poyatos *et al.* show, one can alter the effective interaction between the slow degree of freedom (the oscillator) and the environment (laser light) by adjusting the parameters of the actual H_{int}.

The first steps towards realization of these "environment engineering" proposals were taken by the NIST group (Myatt *et al.*, 2000; Turchette *et al.*, 2000). Similar techniques can be employed to protect deliberately selected states from decoherence (Carvalho *et al.*, 2001).

Other ideas aimed at controling and even at exploiting decoherence have also been explored in contexts that range from quantum information processing (Beige *et al.*, 2000) to preservation of Schrödinger cats in Bose-Einstein condensates (Dalvit, Dziarmaga, and Zurek, 2000).

3. Error correction and resilient quantum computing

This strategy is perhaps the most sophisticated and comprehensive, and capable of dealing with the greatest variety of errors in the most hardware-independent manner. It is a direct descendant of the error-correction techniques employed in dealing with classical information based on redundancy. Multiple copies of the information are made, and the errors are found and corrected by sophisticated "majority voting" techniques.

One might have thought that implementing error correction in the quantum setting would be difficult for two reasons. To begin with, quantum states and hence, quantum information cannot be cloned (Dieks, 1982; Wootters and Zurek, 1982). Moreover, quantum information is very private, and the measurement that is involved in majority voting would infringe on this privacy and destroy quantum coherence, making quantum information classical. Fortunately, both of these difficulties can be simultaneously overcome by encoding quantum information in entangled states of several qubits. Cloning turns out not to be necessary. And measurements can be carried out in a way that identifies errors while keeping quantum information untouched. Moreover, error correction is discrete; measurements that reveal error syndromes have "yes-no" outcomes. Thus, even though the information stored in a qubit represents a continuum of possible quantum states (e.g., corresponding to a surface of the Bloch sphere) error correction is discrete, allaying one of the earliest worries concerning the feasibility of quantum computation—the unchecked "drift" of the quantum state representing the information (Landauer, 1995).

This strategy [discovered by Shor (1995) and Steane (1996)] has been since investigated by many (Bennett *et al.*, 1996; Ekert and Macchiavello, 1996; Laflamme *et al.*, 1996) and codified into a mathematically appealing formalism (Gottesman, 1996; Knill and Laflamme, 1997). Moreover, the first examples of successful implementation (see, for example, Cory *et al.*, 1999) are already at hand.

Error correction allows one, at least, in principle, to compute forever, providing that the errors are suitably small ($\sim 10^{-4}$ per computational step seems to be the error-probability threshold sufficient for most error-correction schemes). Strategies that accomplish this encode qubits in already encoded qubits (Aharonov and Ben-Or, 1996; Knill, Laflamme, and Zurek, 1996, 1998a, 1998b; Kitaev, 1997c; Preskill, 1998). The number of layers of such concatenations necessary to achieve fault tolerance—the ability to carry out arbitrarily long computations—depends on the size (and the character) of the errors, and on the duration of the computation, but when the error probability is smaller than the threshold, that number of layers is finite. Overviews of fault-tolerant computation are already at hand (Preskill, 1999; Nielsen and Chuang, 2000, and references therein).

An interesting subject related to the above discussion is quantum process tomography, anticipated by Jones (1994), and described in the context of quantum information processing by Chuang and Nielsen (1997) and by Poyatos, Cirac, and Zoller (1997). The aim here is to completely characterize a process, such as a quantum logical gate, and not just a state. The first deliberate implementation of this procedure (Nielsen, Knill, and Laflamme, 1998) has also demonstrated experimentally that einselection is indeed equivalent to an unread measurement of the pointer basis by the environment, and can be regarded as such from the standpoint of applications (e.g., NMR teleportation in the example above).

IX. CONCLUDING REMARKS

Decoherence, einselection, pointer states, and even the predictability sieve have become familiar to many in the past decade. The first goal of this paper was to review these advances and to survey, and—where possible, to address—the remaining difficulties. The second related aim was to preview future developments. This has led to considerations involving information, as well as to the operational, physically motivated discussions of seemingly esoteric concepts such as objectivity. Some of the material presented (including the Darwinian view of the emergence of objectivity through redundancy, as well as the discussion of envariance and probabilities) is rather new, and a subject of research, hence the word "preview" applies here.

New paradigms often take a long time to gain ground. The atomic theory of matter (which, until the early 20th century, was "just an interpretation") is a case in point. Some of the most tangible applications and consequences of new ideas are difficult to recognize immediately. In the case of atomic theory, Brownian motion is a good example. Even when the evidence is available, it is often difficult to decode its significance.

Decoherence and einselection are no exception. They have been investigated for about two decades. They are the only explanation of classicality that does not require modifications of quantum theory, as do the alternatives (Bohm, 1952; Pearle, 1976, 1993; Leggett, 1980, 1998, 2002; Ghirardi, Rimini, and Weber, 1986, 1987; Penrose, 1986, 1989; Gisin and Percival, 1992, 1993a, 1993b, 1993c; Holland, 1993; Goldstein, 1998). Ideas based on the immersion of the system in the environment have recently gained enough support to be described (by skeptics) as "the new orthodoxy" (Bub, 1997). This is a dangerous characterization, since it suggests that the interpretation based on the recognition of the role of the environment is both complete and widely accepted. Certainly neither is the case.

Many conceptual and technical issues (such as what constitutes a system) are still open. As for the breadth of acceptance, "the new orthodoxy" seems to be an optimistic (mis)characterization of decoherence and einselection, especially since this explanation of the transition from quantum to classical has (with very few exceptions) not made it into the textbooks. This is intriguing, and may be as much a comment on the way in which quantum physics has been taught, especially on the undergraduate level, as on the status of the theory we have reviewed and its level of acceptance among physicists.

Quantum mechanics has been to date, by and large, presented in a manner that reflects its historical development. That is, Bohr's planetary model of the atom is still often the point of departure, Hamilton-Jacobi equations are used to "derive" the Schrödinger equation, and an oversimplified version of the quantum-classical relationship (attributed to Bohr, but generally not doing justice to his much more sophisticated views) with the correspondence principle, kinship of commutators and Poisson brackets, the Ehrenfest theorem, some version of the Copenhagen interpretation, and other evidence that quantum theory is really not all that different from classical—especially when systems of interest become macroscopic, and all one cares about are averages—is presented.

The message seems to be that there is really no problem and that quantum mechanics can be "tamed" and confined to the microscopic domain. Indeterminacy and the double-slit experiment are of course discussed, but to prove peaceful coexistence within the elbow room assured by Heisenberg's principle and complementarity. Entanglement is rarely explored. This is quite consistent with the aim of introductory quantum-mechanics courses, which has been (only slightly unfairly) summed up by the memorable phrase "shut up and calculate." Discussion of measurement is either dealt with through models based on the Copenhagen interpretation "old orthodoxy" or not at all. An implicit (and sometimes explicit) message is that those who ask questions that do not lend themselves to an answer through laborious, preferably perturbative calculations are "philosophers" and should be avoided.

The above description is of course a caricature. But given that the calculational techniques of quantum theory needed in atomic, nuclear, particle, or condensed-matter physics are indeed difficult to master, and given that, to date, most of the applications had nothing to do with the nature of quantum states, entanglement, and such, the attitude of avoiding the most flagrantly quantum aspects of quantum theory is easy to understand.

Yet, novel applications force one to consider questions about the information content, the nature of the quantum, and the emergence of the classical much more directly, with a focus on states and correlations, rather than on the spectra, cross sections, and the expectation values. Hence problems that are usually bypassed will come to the fore. It is hard to brand Schrödinger cats and entanglement as exotic and make them the centerpiece of a marketable device. I believe that as a result decoherence will become part of textbook lore. Indeed, at the graduate level there are already some notable exceptions among monographs (Peres, 1993) and specialized texts (Walls and Milburn, 1994; Nielsen and Chuang, 2000).

Moreover, the range of subjects already influenced by decoherence and einselection—by the ideas originally motivated by the quantum theory of measurements—is beginning to extend way beyond its original domain. In addition to atomic physics, quantum optics, and quantum information processing (which were all mentioned throughout this review) it stretches from material sciences (Karlsson, 1998; Chatzidimitriou-Dreismann *et al.*, 1997, 2001), surface science, where it seems to be an essential ingredient explaining the emission of electrons (Brodie, 1995; Durakiewicz *et al.*, 2001), through heavy-ion collisions (Krzywicki, 1993) to quantum gravity and cosmology (Zeh, 1986, 1988, 1992; Kiefer, 1987; Halliwell, 1989; Barvinsky and Kamenshchik, 1990, 1995; Brandenberger, Laflamme, and Mijic, 1990; Paz and Sinha, 1991, 1992; Castagnino *et al.*, 1993; Kiefer and Zeh, 1995; Mensky and Novikov, 1996) and even (quantum) robotics (Benioff, 1988). Given the limitations of space we have not done justice to most of these subjects, focusing instead on issues of principle. In some areas reviews already exist. Thus Giulini *et al.* (1996) is a valuable collection of essays, where, for example, decoherence in field theories is addressed. The dissertation of Wallace (2002) offers a good (if somewhat philosophical) summary of the role of decoherence with a rather different emphasis on similar field-theoretic issues. Conference proceedings edited by Blanchard *et al.* (2000) and, especially, an extensive historical overview of the foundation of quantum theory from the modern perspective by Auletta (2000) are also recommended. More specific technical issues with implications for decoherence and einselection have also been reviewed. For example, on the subject of master equations there are several reviews with very different emphases including Alicki and Lendi (1987); Grabert, Schramm, and Ingold (1988); Namiki and Pascazio (1993); as well as—more recently—Paz and Zurek (2001). In some areas, such as atomic Bose-Einstein condensation, the study of decoherence has only started (Anglin, 1997; Dalvit, Dziarmaga, and Zurek, 2001). In many situations (e.g., quantum optics) a useful supplement to the decoherence view of the quantum-classical interface is afforded by quantum trajectories—a study of the state of the system inferred from the intercepted state of the environment (see Carmichael, 1993; Gisin and Percival, 1993a, 1993b, 1993c; Wiseman and Milburn, 1993). This approach "unravels" the evolving density matrices of open systems into trajectories conditioned upon the measurement carried out on the environment, and may have—especially in quantum optics—intriguing connections with the "environment as a witness" point of view (see Dalvit, Dziarmaga, and Zurek, 2001). In other areas, such as condensed matter, decoherence phenomena have so many variations and are so pervasive that a separate "decoherent review" may be in order, especially as intriguing experimental puzzles seem to challenge the theory (Mohanty and Webb, 1997; Kravtsov and Altshuler, 2000). Indeed, perhaps the most encouraging development is the increase of interest in experiments that test validity of quantum physics beyond the microscopie domain (see, for example, Folman, Krüger, Schmiedmayer, *et al.*, 2002; Marshall, Simon, Penrose, and Bouwmeester, 2002).

The physics of information and computation is a special case. Decoherence is obviously a key obstacle in the implementation of information-processing hardware that takes advantage of the superposition principle. While we have not focused on quantum information processing, the discussion has often been couched in language inspired by information theory. This is no accident. It is the belief of this author that many of the remaining gaps in our understanding of quantum physics and its relation to the classical domain—such as the definition of systems, or the still mysterious aspects of collapse—will follow the pattern of the predictability sieve and be expanded into new areas of investigation by considerations that simultaneously elucidate the nature of the quantum and of information.

ACKNOWLEDGMENTS

John Archibald Wheeler—a quarter century ago—taught a course on the subject of quantum measurements at the University of Texas in Austin. The questions raised then have since evolved into ideas presented here, partly in collaboration with Juan Pablo Paz, and through interactions with many colleagues, including Andreas Albrecht, James Anglin, Charles Bennett, Robin Blume-Kohout, Carlton Caves, Isaac Chuang, Diego Dalvit, David Deutsch, David Divincenzo, Jacek Dziarmaga, Richard Feynman, Murray Gell-Mann, Daniel Gottesmann, Robert Griffiths, Salman Habib, Jonathan Halliwell, Serge Haroche, James Hartle, Christopher Jarzynski, Erich Joos, Emmanuel Knill, Raymond Laflamme, Rolf Landauer, Anthony Leggett, Seth Lloyd, Gerard Milburn, Michael Nielsen, Harold Ollivier, Asher Peres, David Poulin, Rüdiger Schack, Benjamin Schumacher, Kosuke Shizume, William Unruh, David Wallace, Eugene Wigner, William Wootters, Dieter Zeh, Anton Zeilinger, and Peter Zoller. Moreover, Serge Haroche, Mike Nielsen, Harold Ollivier, Juan Pablo Paz, and David Wallace have provided me with extensive written comments on earlier versions of the manuscript. Its preparation was assisted by the author's participation in two Institute for Theoretical Physics programs on decoherence-related subjects, and was in part supported by a grant from the National Security Agency. Last but not least, this paper has evolved in the course of over a dozen years, along with the field, under the watchful eye of a sequence of increasingly impatient editors of RMP, their feelings shared by my family. The perseverance of all afflicted was very much appreciated by the author (if thoroughly tested by the process).

REFERENCES

Adler, S. L., 2003, Stud. Hist. Philos. Mod. Phys. **34B**, 135.
Aharonov, D., and M. Ben-Or, 1996, *Proceedings of the 29th Annual ACM Symposium on the Theory of Computing* (ACM Press, New York), p. 176.
Aharonov, Y., J. Anandan, and L. Vaidman, 1993, Phys. Rev. A **47**, 4616.
Aharonov, Y., and B. Reznik, 2003, Phys. Rev. A **65**, 052116.
Albrecht, A., 1992, Phys. Rev. D **46**, 5504.
Albrecht, A., 1993, Phys. Rev. D **48**, 3768.
Alicki, R., and K. Lendi, 1987, *Quantum Dynamical Semigroups and Applications*, Lecture Notes in Physics No. 286 (Springer, Berlin).
Ambegaokar, V., 1991, Ber. Bunsenges. Phys. Chem. **95**, 400.
Ammann, H., R. Gray, I. Shvarchuk, and N. Christensen, 1998, Phys. Rev. Lett. **80**, 4111.
Anderson, P. W., 2001, Stud. Hist. Philos. Mod. Phys. **32**, 487.
Anglin, J. R., 1997, Phys. Rev. Lett. **79**, 6.
Anglin, J. R., J. P. Paz, and W. H. Zurek, 1997, Phys. Rev. A **55**, 4041.
Anglin, J. R., and W. H. Zurek, 1996, Phys. Rev. D **53**, 7327.
Auletta, G., 2000, *Foundations and Interpretation of Quantum Theory* (World Scientific, Singapore).
Bacciagaluppi, G., and M. Hemmo, 1998, in *Quantum Measurements Beyond Paradox*, edited by R. A. Healey and G. Hellman (University of Minnesota, Minneapolis), p. 95.
Ballentine, L. E., 1970, Rev. Mod. Phys. **42**, 358.
Ballentine, L. E., Y. Yang, and J. P. Zibin, 1994, Phys. Rev. A **50**, 2854.
Barnett, S. M., and S. J. D. Phoenix, 1989, Phys. Rev. A **40**, 2404.
Barnum, H., C. M. Caves, J. Finkelstein, C. A. Fuchs, and R. Schack, 2000, Proc. R. Soc. London, Ser. A **456**, 1175.
Barvinsky, A. O., and A. Yu. Kamenshchik, 1990, Class. Quantum Grav. **7**, 2285.
Barvinsky, A. O., and A. Yu. Kamenshchik, 1995, Phys. Rev. D **52**, 743.
Beckman, D., D. Gottesman, M. A. Nielsen, and J. Preskill, 2001, Phys. Rev. A **64**, 052309.
Beige, A., D. Braun, D. Tregenna, and P. L. Knight, 2000, Phys. Rev. Lett. **85**, 1762.
Bell, J. S., 1981, in *Quantum Gravity 2*, edited by C. Isham, R. Penrose, and D. Sciama (Clarendon, Oxford, England), p. 611.
Bell, J. S., 1987, *Speakable and Unspeakable in Quantum Mechanics* (Cambridge University, Cambridge, England).
Bell, J. S., 1990, in *Sixty-Two Years of Uncertainty*, edited by A. I. Miller (NATO Series, Plenum, New York), p. 17.
Benioff, P., 1998, Phys. Rev. A **58**, 893.
Bennett, C. H., and D. P. DiVincenzo, 2000, Nature (London) **404**, 247.
Bennett, C. H., D. P. DiVincenzo, J. Smolin, and W. K. Wootters, 1996, Phys. Rev. A **54**, 3824.
Berman, G. P., and G. M. Zaslavsky, 1978, Physica A **91**, 450.
Berry, M. V., 1995, in *Fundamental Problems in Quantum Theory*, edited by D. M. Greenberger and A. Zeilinger (N.Y. Academy of Sciences, New York), p. 303.
Berry, M. V., and N. L. Balazs, 1979, J. Phys. A **12**, 625.
Birkhoff, G., and J. von Neumann, 1936, Ann. Math. **37**, 823.
Blanchard, Ph., D. Giulini, E. Joos, C. Kiefer, and I.-O. Stamatescu, 2000, Eds., *Decoherence: Theoretical, Experimental, and Conceptual Problems* (Springer, Berlin).
Blanchard, Ph., and R. Olkiewicz, 2000, Phys. Lett. A **273**, 223.
Boccaletti, S., A. Farini, and F. T. Arecchi, 1997, Phys. Rev. E **55**, 4979.
Bocko, M. F., and R. Onofrio, 1996, Rev. Mod. Phys. **68**, 755.
Bogolubov, N. N., Jr., A. A. Logunov, A. I. Oksak, and I. T. Todorov, 1990, *General Principles of Quantum Field Theory* (Kluwer, Dordrecht).
Bohm, D., 1952, Phys. Rev. **85**, 166. Reprinted 1983 in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek (Princeton University, Princeton, NJ), p. 369.
Bohr, N., 1928, Nature (London) **121**, 580. Reprinted 1983 in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek (Princeton University, Princeton, NJ), p. 87.
Bohr, N., 1935, Phys. Rev. **48**, 696. Reprinted 1983 in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek (Princeton University, Princeton, NJ), p. 144.
Bohr, N., 1949, in *Albert Einstein: Philosopher-Scientist*, edited by P. A. Schilpp (Open Court, Evanston), p. 200. Reprinted 1983 in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek (Princeton University, Princeton, NJ), p. 9.
Borel, E., 1914, *Introduction Géométrique à Quelques Théories Physiques* (Gauthier-Villars, Paris).

Born, M., 1969, *Albert Einstein, Hedwig und Max Born, Briefwechsel 1916–1955* (Nymphenburger, München).
Braginsky, V. B., and F. Ya. Khalili, 1996, Rev. Mod. Phys. **68**, 1.
Braginski, V. B., Y. I. Vorontsov, and K. S. Thorne, 1980, Science **209**, 547. Reprinted 1983 in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek (Princeton University, Princeton, NJ), p. 749.
Brandenberger, R., R. Laflamme, and M. Mijic, 1990, Mod. Phys. Lett. A **5**, 2311.
Braun, D., P. A. Braun, and F. Haake, 2000, Opt. Commun. **179**, 411.
Bravyi, S. B., and A. Y. Kitaev, 1998, e-print quant-ph/9811052.
Brillouin, L., 1962, *Science and Information Theory* (Academic, New York).
Brillouin, L., 1964, *Scientific Uncertainty, and Information* (Academic, New York).
Brodie, I., 1995, Phys. Rev. B **51**, 13 660.
Brun, T., and J. B. Hartle, 1999, Phys. Rev. D **60**, 1165.
Brun, T., I. C. Percival, and R. Schack, 1996, J. Phys. A **29**, 2077.
Brune, M., E. Hagley, J. Dreyer, X. Maître, C. Wunderlich, J.-M. Raimond, and S. Haroche, 1996, Phys. Rev. Lett. **77**, 4887.
Bruss, D., A. Ekert, and C. Macchiavello, 1999, Phys. Rev. Lett. **81**, 2598.
Bub, J., 1997, *Interpreting the Quantum World* (Cambridge University, Cambridge, England).
Butterfield, J., 1996, Br. J. Philos. Sci. **47**, 200.
Caldeira, A. O., and A. J. Leggett, 1983, Physica A **121**, 587.
Carmichael, H. J., 1993, *An Open Systems Approach to Quantum Optics* (Springer, Berlin).
Carvalho, A. R. R., P. Milman, R. L. de Matos Filho, and L. Davidovich, 2001, Phys. Rev. Lett. **86**, 4988.
Casati, G., and B. V., Chirikov, 1995a, *Quantum Chaos: Between Order and Disorder* (Cambridge University, Cambridge, England).
Casati, G., and B. V. Chirikov, 1995b, Phys. Rev. Lett. **75**, 2507.
Castagnino, M. A., A. Gangui, F. D. Mazzitelli, and I. I. Tkachev, 1993, Class. Quantum Grav. **10**, 2495.
Caves, C. M., 1994, in *Physical Origins of Time Asymmetry*, edited by J. J. Halliwell, J. Perez-Mercader, and W. H. Zurek (Cambridge University, Cambridge, England), p. 47.
Caves, C. M., K. S. Thorne, R. W. P. Drever, V. P. Sandberg, and M. Zimmerman, 1980, Rev. Mod. Phys. **52**, 341.
Cerf, N., and C. Adami, 1997, Phys. Rev. A **55**, 3371.
Chapman, M. S., T. D. Hammond, A. Lenef, J. Schmiedmayer, R. A. Rubenstein, E. Smith, and D. E. Pritchard, 1995, Phys. Rev. Lett. **75**, 3783.
Chatzidimitriou-Dreismann, C. A., T. Abdul-Redah, and B. Kolaric, 2001, J. Am. Chem. Soc. **123**, 11 945.
Chatzidimitriou-Dreismann, C. A., T. Abdul Redah, R. M. Streffer, and J. Mayers, 1997, Phys. Rev. Lett. **79**, 2839.
Cheng, C. C., and M. G. Raymer, 1999, Phys. Rev. Lett. **82**, 4807.
Chuang, I. L., and M. A. Nielsen, 1997, J. Mod. Opt. **44**, 2455.
Cirac, J. I., and P. Zoller, 1995, Phys. Rev. Lett. **74**, 4091.
Cory, D. G., W. Mass, M. Price, E. Knill, R. Laflamme, W. H. Zurek, T. F. Havel, and S. S. Somaroo, 1999, Phys. Rev. Lett. **81**, 2152.
Cover, T. M., and J. A. Thomas, 1991, *Elements of Information Theory* (Wiley, New York).
Cox, R. T., 1946, Am. J. Phys. **14**, 1.
Dalvit, D., J. Dziarmaga, and W. H. Zurek, 2000, Phys. Rev. A **62**, 013607.
Dalvit, D., J. Dziarmaga, and W. H. Zurek, 2001, Phys. Rev. Lett. **86**, 373.
Dekker, H., 1977, Phys. Rev. A **16**, 2126.
d'Espagnat, B., 1976, *Conceptual Foundations of Quantum Mechanics* (Benjamin, Reading, MA).
d'Espagnat, B., 1989, J. Stat. Phys. **58**, 747.
d'Espagnat, B., 1995, *Veiled Reality* (Addison-Wesley, Reading, MA).
Deutsch, D., 1985, Int. J. Theor. Phys. **24**, 1.
Deutsch, D., 1997, *The Fabric of Reality* (Penguin, New York).
Deutsch, D., 1999, Proc. R. Soc. London, Ser. A **455**, 3129.
Deutsch, D., 2002, Proc. R. Soc. London, Ser. A **458**, 2911.
DeWitt, B. S., 1970, Phys. Today **23**, 30.
DeWitt, B. S., 1998, in *Frontiers of Quantum Physics*, edited by S. C. Lim, R. Abd-Shukor, and K. H. Kwek (Springer, New York), p. 39.
DeWitt, B. S., and N. Graham, 1973, *The Many-Worlds Interpretation of Quantum Mechanics* (Princeton University, Princeton, NJ).
Dieks, D., 1982, Phys. Lett. **92A**, 271.
Dirac, P. A. M., 1947, *The Principles of Quantum Mechanics* (Clarendon, Oxford, England).
Donald, M. J., 1995, Found. Phys. **25**, 529.
Dowker, H. F., and J. J. Halliwell, 1992, Phys. Rev. D **46**, 1580.
Dowker, H. F., and A. Kent, 1996, J. Stat. Phys. **82**, 1575.
Duan, L.-M., and G.-C. Guo, 1998, Phys. Rev. A **57**, 737.
Durakiewicz, T., S. Halas, A. Arko, J. J. Joyce, and D. P. Moore, 2001, Phys. Rev. B **64**, 045101.
Einstein, A., 1917, Verh. Dtsch. Phys. Ges. **19**, 82.
Einstein, A., B. Podolsky, and N. Rosen, 1935, Phys. Rev. **47**, 777. Reprinted 1983 in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek (Princeton University, Princeton, NJ), p. 138.
Eisert, J., and M. B. Plenio, 2002, Phys. Rev. Lett. **89**, 137902.
Ekert, A., and C. Macchiavello, 1996, Phys. Rev. Lett. **77**, 2585.
Elby, A., 1993, in *Symposium on the Foundations of Modern Physics*, edited by P. Busch, P. Lahti, and P. Mittelstaedt (World Scientific, Singapore), p. 168.
Elby, A., 1998, in *Quantum Measurement: Beyond Paradox*, edited by R. A. Healey and G. Hellman (University of Minnesota, Minneapolis), p. 87.
Elby, A., and J. Bub, 1994, Phys. Rev. A **49**, 4213.
Emerson, J., and L. E. Ballentine, 2001a, Phys. Rev. A **63**, 052103.
Emerson, J., and L. E. Ballentine, 2001b, Phys. Rev. E **64**, 026217.
Everett, H., III, 1957a, Rev. Mod. Phys. **29**, 454. Reprinted 1983 in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek (Princeton University, Princeton, NJ), p. 315.
Everett, H., III, 1957b, Ph.D. thesis (Princeton University). Reprinted 1973 in *The Many-Worlds Interpretation of Quantum Mechanics*, by B. S. Dewitt and N. Graham (Princeton University, Princeton, NJ), p. 3.
Farhi, E., J. Goldstone, and S. Guttmann, 1989, Ann. Phys. (N.Y.) **192**, 368.
Feynman, R. P., and F. L. Vernon, Jr., 1963, Ann. Phys. (N.Y.) **24**, 118.
Folman, R., P. Krüger, J. Schmiedmayer, J. Denschlag, and C. Henkel, 2002, Adv. At., Mol. Opt. Phys. **48**, 263.
Ford, J., and G. Mantica, 1992, Am. J. Phys. **60**, 1086.

Fox, R. F., and T. C. Elston, 1994a, Phys. Rev. E **49**, 3683.
Fox, R. F., and T. C. Elston, 1994b, Phys. Rev. E **50**, 2553.
Freedman, M. H., and D. A. Meyer, 2001, Found. Comp. Math. **1**, 325.
Furuya, K., M. C. Nemes, and G. Q. Pellegrino, 1998, Phys. Rev. Lett. **80**, 5524.
Fuchs, C. A., and A. Peres, 2000, Phys. Today **53**, 70.
Gallis, M. R., 1992, Phys. Rev. A **45**, 47.
Gallis, M. R., 1996, Phys. Rev. A **53**, 655.
Gallis, M. R., and G. N. Fleming, 1990, Phys. Rev. A **42**, 38.
Gell-Mann, M., and J. B. Hartle, 1990, in *Complexity Entropy, and the Physics of Information*, edited by W. H. Zurek (Addison-Wesley, Reading, MA), p. 425.
Gell-Mann, M., and J. B. Hartle, 1993, Phys. Rev. D **47**, 3345.
Gell-Mann, M., and J. B. Hartle, 1997, in *Quantum-Classical Correspondence*, edited by D. H. Feng and B. L. Hu (International, Cambridge, MA), p. 3.
Geroch, R., 1984, Noûs **XVIII**, 617.
Ghirardi, G. C., A. Rimini, and T. Weber, 1986, Phys. Rev. D **34**, 470.
Ghirardi, G. C., A. Rimini, and T. Weber, 1987, Phys. Rev. D **36**, 3287.
Gisin, N., and I. C. Percival, 1992, J. Phys. A **25**, 5677.
Gisin, N., and I. C. Percival, 1993a, J. Phys. A **26**, 2233.
Gisin, N., and I. C. Percival, 1993b, J. Phys. A **26**, 2245.
Gisin, N., and I. C. Percival, 1993c, Phys. Lett. A **175**, 144.
Giulini, D., 2000, in *Relativistic Quantum Measurement and Decoherence*, edited by H.-P. Breuer and F. Petruccione (Springer, Berlin), p. 67.
Giulini, D., E. Joos, C. Kiefer, J. Kupsch, I.-O. Stamatescu, and H. D. Zeh, 1996, *Decoherence and the Appearance of a Classical World in Quantum Theory* (Springer, Berlin).
Giulini, D., C. Kiefer, and H. D. Zeh, 1995, Phys. Lett. A **199**, 291.
Gleason, A. M., 1957, J. Math. Mech. **6**, 885.
Gnedenko, B. V., 1982, *The Theory of Probability* (MIR, Moscow).
Goldstein, S., 1998, Phys. Today **52** (3), 42 (part I) and (4), 38 (part II).
Goldstein, S., and D. N. Page, 1995, Phys. Rev. Lett. **74**, 3715.
Gorin, T., and T. H. Seligman, 2002, J. Opt. B: Quantum Semiclassical Opt. **4**, S386.
Gorini, V., A. Kossakowski, and E. C. G. Sudarshan, 1976, J. Math. Phys. **17**, 821.
Gottesman, D., 1996, Phys. Rev. A **54**, 1862.
Gottfried, K., 1966, *Quantum Mechanics* (Benjamin, New York), Sec. 20.
Gottfried, K., 2000, Nature (London) **405**, 533.
Grabert, H., P. Schramm, and G.-L. Ingold, 1988, Phys. Rep. **168**, 115.
Graham, N., 1970, *The Everett Interpretation of Quantum Mechanics* (University of North Carolina, Chapel Hill).
Graham, R., M. Schlautmann, and P. Zoller, 1992, Phys. Rev. A **45**, R19.
Griffiths, R. B., 1984, J. Stat. Phys. **36**, 219.
Griffiths, R. B., 1996, Phys. Rev. A **54**, 2759.
Griffiths, R. B., and R. Omnès, 1999, Phys. Today **52**, 26.
Haake, F., 1991, *Quantum Signatures of Chaos* (Springer, Berlin).
Haake, F., M. Kuś, and R. Sharf, 1987, Z. Phys. B: Condens. Matter **65**, 381.
Haake, F., and R. Reibold, 1985, Phys. Rev. A **32**, 2462.
Haake, F., and D. F. Walls, 1987, Phys. Rev. A **36**, 730.
Habib, S., K. Shizume, and W. H. Zurek, 1998, Phys. Rev. Lett. **80**, 4361.
Halliwell, J. J., 1989, Phys. Rev. D **39**, 2912.
Halliwell, J. J., 1999, Phys. Rev. D **60**, 105031.
Halliwell, J. J., J. Pérez Mercader, and W. H. Zurek, 1994, *Physical Origins of Time Asymmetry* (Cambridge University, Cambridge, England).
Haroche, S., 1998, Phys. Scr., T **76**, 159.
Hartle, J. B., 1968, Am. J. Phys. **36**, 704.
Healey, R. A., 1998, in *Quantum Measurement: Beyond Paradox*, edited by R. A. Healey and G. Hellman (University of Minnesota, Minneapolis), p. 52.
Healey, R. A., and G. Hellman, 1998, Eds., *Quantum Measurement: Beyond Paradox* (University of Minnesota, Minneapolis).
Heisenberg, W., 1927, Z. Phys. **43**, 172. English translation 1983 in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek (Princeton University, Princeton, NJ), p. 62.
Herzog, T. J., P. G. Kwiat, H. Weinfurter, and A. Zeilinger, 1995, Phys. Rev. Lett. **75**, 3034.
Hillery, M. A., R. F. O'Connell, M. O. Scully, and E. P. Wigner, 1984, Phys. Rep. **106**, 121.
Hillery, M. A., and M. O. Scully, 1983, in *Quantum Optics, Experimental Gravitation, and Measurement Theory*, edited by P. Meystre and M. O. Scully (Plenum, New York), p. 65.
Holland, P. R., 1993, *The Quantum Theory of Motion* (Cambridge University, Cambridge, England).
Hu, B. L., J. P. Paz, and Y. Zhang, 1992, Phys. Rev. D **45**, 2843.
Hu, B. L., J. P. Paz, and Y. Zhang, 1993, Phys. Rev. D **47**, 1576.
Ivanovic, I. D., 1981, J. Phys. A **14**, 3241.
Jacquod, P., P. G. Silvestrov, and C. W. J. Beenakker, 2001, Phys. Rev. E **64**, 055203(R).
Jalabert, R. A., and H. M. Pastawski, 2001, Phys. Rev. Lett. **86**, 2490.
Jammer, M., 1974, *The Philosophy of Quantum Mechanics* (Wiley, New York).
Jaynes, E. T., 1957, Phys. Rev. **106**, 620.
Jones, K. R. W., 1994, Phys. Rev. A **50**, 3682.
Joos, E., 1986, Ann. N.Y. Acad. Sci. **480**, 6.
Joos, E., and H. D. Zeh, 1985, Z. Phys. B: Condens. Matter **59**, 223.
Jozsa, R., 2002, e-print quant-ph/02044153.
Karkuszewski, Z. P., C. Jarzynski, and W. H. Zurek, 2002, Phys. Rev. Lett. **89**, 170405.
Karkuszewski, Z. P., J. M. Zakrzewski, and W. H. Zurek, 2002, Phys. Rev. A **65**, 042113.
Karlsson, E. B., 1998, Phys. Scr., T **76**, 179.
Kent, A., 1990, Int. J. Mod. Phys. A **5**, 1745.
Kiefer, C., 1987, Class. Quantum Grav. **4**, 1369.
Kiefer, C., 1996, *Decoherence and the Appearance of a Classical World in Quantum Theory* (Springer, Berlin), p. 157 [D. Giulini *et al.* (1996)].
Kiefer, C., and H. D. Zeh, 1995, Phys. Rev. D **51**, 4145.
Kitaev, A. Y., 1997a, e-print quant-ph/9707021.
Kitaev, A. Y., 1997b, Russ. Math. Surveys **52**, 1191.
Kitaev, A. Y., 1997c, in *Quantum Communication, Computing, and Measurement*, edited by A. S. Holevo, O. Hirota, and C. M. Caves (Kluwer Academic, New York), p. 181.
Klappauf, B. G., W. H. Oskay, D. A. Steck, and M. G. Raizen, 1998, Phys. Rev. Lett. **81**, 1203; **82**, 241(E) (1999).
Knill, E., and R. Laflamme, 1997, Phys. Rev. A **55**, 900.

Knill, E., R. Laflamme, and L. Viola, 2000, Phys. Rev. Lett. **84**, 2525.
Knill, E., R. Laflamme, and W. H. Zurek, 1996, e-print quant-ph/96010011.
Knill, E., R. Laflamme, and W. H. Zurek, 1998a, Science **279**, 324.
Knill, E., R. Laflamme, and W. H. Zurek, 1998b, Proc. R. Soc. London, Ser. A **454**, 365.
Koch, P. M., 1995, Physica D **83**, 178.
Kokorowski, D., A. D. Cronin, T. D. Roberts, and D. E. Pritchard, 2001, Phys. Rev. Lett. **86**, 2191.
Kolmogorov, A. N., 1960, Math. Rev. **21**, 2035.
Kossakowski, A., 1973, Bull. Acad. Pol. Sci., Ser. Sci., Math. Astron. Phys. **21**, 649.
Kraus, K., 1983, *States, Effects, and Operations* (Springer, Berlin).
Kravtsov, V. E., and B. L. Altshuler, 2000, Phys. Rev. Lett. **84**, 3394.
Krzywicki, A., 1993, Phys. Rev. D **48**, 5190.
Kwiat, P. G., A. M. Steinberg, and R. Y. Chiao, 1992, Phys. Rev. A **45**, 7729.
Laflamme, R., C. Miquel, J. P. Paz, and W. H. Zurek, 1996, Phys. Rev. Lett. **77**, 198.
Landau, L., 1927, Z. Phys. **45**, 430.
Landauer, R., 1991, Phys. Today **44**, 23.
Landauer, R., 1995, Philos. Trans. R. Soc. London **353**, 367.
Laskar, J., 1989, Nature (London) **338**, 237.
Leggett, A. J., 1980, Suppl. Prog. Theor. Phys. **69**, 80.
Leggett, A. J., 1998, in *Quantum Measurement*, edited by R. A. Healey and G. Hellman (University of Minnesota, Minneapolis), p. 1.
Leggett, A. J., 2002, Phys. Scr., T **102**, 69.
Levstein, P. R., G. Usaj, H. M. Pastawski, J. Raya, and J. Hirschinger, 2000, J. Chem. Phys. **113**, 6285.
Levstein, P. R., G. Usaj, and H. M. Pastawski, 1998, J. Chem. Phys. **108**, 2718.
Li, M., and P. Vitányi, 1993, *An Introduction to Kolmogorov Complexity and Its Applications* (Springer, New York).
Lidar, D. A., D. Bacon, and K. B. Whaley, 1999, Phys. Rev. Lett. **82**, 4556.
Lieb, E. H., 1975, Bull. Am. Math. Soc. **81**, 1.
Lindblad, G., 1976, Commun. Math. Phys. **48**, 119.
Lloyd, S., 1997, Phys. Rev. A **55**, 1613.
London, F., and E., Bauer, 1939, *La Théorie de l'Observation en Méchanique Quantique* (Hermann, Paris). Translation 1983 in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek (Princeton University, Princeton, NJ), p. 217.
Marshall, W., C. Simon, R. Penrose, and D. Bouwmeester, 2002, e-print quant-ph/0210001.
Mensky, M. B., and I. D. Novikov, 1996, Int. J. Mod. Phys. D **5**, 1.
Mermin, N. D., 1990a, Phys. Today **43**, 9.
Mermin, N. D., 1990b, Am. J. Phys. **58**, 731.
Mermin, N. D., 1994, Phys. Today **47**, 9.
Miquel, C., J. P. Paz, and W. H. Zurek, 1997, Phys. Rev. Lett. **78**, 3971.
Miller, P. A., and S. Sarkar, 1999, Phys. Rev. E **60**, 1542.
Miller, P. A., S. Sarkar, and R. Zarum, 1998, Acta Phys. Pol. B **29**, 4643.
Mohanty, P., and R. A. Webb, 1997, Phys. Rev. B **55**, R13 452.
Monroe, C., D. M. Meekhof, B. E. King, and D. J. Wineland, 1996, Science **272**, 1131.
Monteoliva, D., and J.-P. Paz, 2000, Phys. Rev. Lett. **85**, 3373.
Moore, F. L., *et al.*, 1994, Phys. Rev. Lett. **73**, 2974.
Mott, N. F., 1929, Proc. R. Soc. London, Ser. A **126**, 79. Reprinted 1983 in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek (Princeton University, Princeton, NJ), p. 129.
Moyal, J. E., 1949, Proc. Cambridge Philos. Soc. **45**, 99.
Myatt, C. J., *et al.*, 2000, Nature (London) **403**, 269.
Namiki, M., and S. Pascazio, 1993, Phys. Rep. **232**, 301.
Nielsen, M. A., and I. L. Chuang, 2000, *Quantum Computation and Quantum Information* (Cambridge University, Cambridge, England).
Nielsen, M. A., E. Knill, and R. Laflamme, 1998, Nature (London) **396**, 52.
Ollivier, H., D. Poulin, and W. H. Zurek, 2002, Los Alamos preprint LAUR-025844.
Ollivier, H., and W. H. Zurek, 2002, Phys. Rev. Lett. **88**, 017901.
Omnès, R., 1988, J. Stat. Phys. **53**, 893.
Omnès, R., 1992, Rev. Mod. Phys. **64**, 339.
Omnès, R., 1994, *The Interpretation of Quantum Mechanics* (Princeton University, Princeton, NJ).
Paraoanu, G. S., 1999, Europhys. Lett. **47**, 279.
Paraoanu, G.-S., 2002, e-print quant-ph/0205127.
Paraoanu, G. S., and H. Scutari, 1998, Phys. Lett. A **238**, 219.
Pattanayak, A. K., 2000, Phys. Rev. Lett. **83**, 4526.
Paz, J.-P., 1994, in *Physical Origins of Time Asymmetry*, edited by J. J. Halliwell, J. Perez-Mercader, and W. H. Zurek (Cambridge University, Cambridge, England), p. 213.
Paz, J.-P., S. Habib, and W. H. Zurek, 1993, Phys. Rev. D **47**, 488.
Paz, J.-P., and S. Sinha, 1991, Phys. Rev. D **44**, 1038.
Paz, J.-P., and S. Sinha, 1992, Phys. Rev. D **45**, 2823.
Paz, J.-P., and W. H. Zurek, 1999, Phys. Rev. Lett. **82**, 5181.
Paz, J.-P., and W. H. Zurek, 2001, in *Coherent Atomic Matter Waves, Les Houches Lectures*, edited by R. Kaiser, C. Westbrook, and F. David (Springer, Berlin), p. 533.
Pearle, P., 1976, Phys. Rev. D **13**, 857.
Pearle, P., 1993, Phys. Rev. A **48**, 913.
Penrose, R., 1986, in *Quantum Concepts in Space and Time*, edited by C. J. Isham and R. Penrose (Clarendon, Oxford, England), p. 129.
Penrose, R., 1989, *The Emperor's New Mind* (Oxford University, Oxford, England).
Perelomov, A., 1986, *Generalized Coherent States and Their Applications* (Springer, Berlin).
Peres, A., 1993, *Quantum Theory: Concepts and Methods* (Kluwer, Dordrecht).
Pfau, T., S. Spälter, C. Kurtsiefer, C. R. Ekstrom, and J. Mlynek, 1994, Phys. Rev. Lett. **73**, 1223.
Poyatos, J. F., J. I. Cirac, and P. Zoller, 1996, Phys. Rev. Lett. **77**, 4728.
Poyatos, J. F., J. I. Cirac, and P. Zoller, 1997, Phys. Rev. Lett. **78**, 390.
Preskill, J., 1998, Proc. R. Soc. London, Ser. A **454**, 385.
Preskill, J., 1999, Phys. Today **52** (6), 24.
Prosen, T., and T. H. Seligman, 2002, J. Phys. A **35**, 4707.
Raimond, J. M., M. Brune, and S. Haroche, 2001, Rev. Mod. Phys. **73**, 565.
Rauch, H., 1998, Phys. Scr., T **76**, 24.
Reichl, L. E., 1992, *Transition to Chaos in Conservative Classical Systems: Quantum Manifestations* (Springer, Berlin).
Romero, L. D., and J.-P. Paz, 1997, Phys. Rev. A **55**, 4070.

Saunders, S., 1998, Synthesis **114**, 373.
Schack, R., 1998, Phys. Rev. A **57**, 1634.
Schack, R., and C. M. Caves, 1996, Phys. Rev. E **53**, 3387.
Schrödinger, E., 1926, Naturwissenschaften **14**, 664.
Schrödinger, E., 1935a, Naturwissenschaften **23**, 807–812, 823–828, 844–849. English translation 1983 in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek (Princeton University, Princeton, NJ), p. 152.
Schrödinger, E., 1935b, Proc. Cambridge Philos. Soc. **31**, 555.
Schrödinger, E., 1936, Proc. Cambridge Philos. Soc. **32**, 446.
Schumacher, B. W., 1996, Phys. Rev. A **54**, 2614.
Schumacher, B. W., and M. A. Nielsen, 1996, Phys. Rev. A **54**, 2629.
Schumacher, B. W., 1999, private communication concerning *distinguishability sieve*.
Shiokawa, K., and B. L. Hu, 1995, Phys. Rev. E **52**, 2497.
Shor, P. W., 1995, Phys. Rev. A **52**, R2493.
Sinai, Ya., 1960, Math. Rev. **21**, 2036.
Squires, E. J., 1990, Phys. Lett. A **145**, 67.
Stapp, H. P., 2002, Can. J. Phys. **80**, 1043.
Steane, A. M., 1996, Proc. R. Soc. London, Ser. A **452**, 2551.
Stein, H., 1984, Noûs **XVIII**, 635.
Stern, A., Y. Aharonow, and Y. Imry, 1989, Phys. Rev. A **41**, 3436.
Sussman, G. J., and J. Wisdom, 1992, Science **257**, 56.
Tegmark, M., 2000, Phys. Rev. E **61**, 4194.
Tegmark, M., and H. S. Shapiro, 1994, Phys. Rev. E **50**, 2538.
Tegmark, M., and J. A. Wheeler, 2001, Sci. Am. **284**, 68.
Turchette, Q. A., *et al.*, 2000, Phys. Rev. A **62**, 053807.
Unruh, W. G., 1994, Phys. Rev. A **50**, 882.
Unruh, W. G., and W. H. Zurek, 1989, Phys. Rev. D **40**, 1071.
Vedral, V., 2003, Phys. Rev. Lett. **90**, 050401.
Venugopalan, A., 1994, Phys. Rev. A **50**, 2742.
Venugopalan, A., 2000, Phys. Rev. A **6**, 012102.
Viola, L., and S. Lloyd, 1998, Phys. Rev. A **58**, 2733.
von Neumann, J., 1932, *Mathematische Grundlagen der Quantenmechanik* (Springer, Berlin); reprinted 1981; English translation by R. T. Beyer, 1955: *Mathematical Foundations of Quantum Mechanics* (Princeton University, Princeton, NJ).
Wallace, D., 2002, *Issues in the Foundations of Relativistic Quantum Theory* (Merton College, University of Oxford, Oxford, England).
Walls, D. F., and G. J., Milburn, 1994, *Quantum Optics* (Springer, Berlin).
Wehrl, A., 1978, Rev. Mod. Phys. **50**, 221.
Wheeler, J. A., 1957, Rev. Mod. Phys. **29**, 463.
Wheeler, J. A., 1978, in *Mathematical Foundations of Quantum Theory*, edited by A. R. Marlow (Academic, New York), p. 9.
Wheeler, J. A., 1983, in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek (Princeton University, Princeton, NJ), p. 182.
Wheeler, J. A., and W. H. Zurek, 1983, Eds., *Quantum Theory and Measurement* (Princeton University, Princeton, NJ).
Wick, G. C., A. S. Wightman, and E. P. Wigner, 1952, Phys. Rev. **88**, 101.
Wick, G. C., A. S. Wightman, and E. P. Wigner, 1970, Phys. Rev. D **1**, 3267.
Wightman, A. S., 1995, Nuovo Cimento B **110**, 751.
Wigner, E. P., 1963, Am. J. Phys. **31**, 6. Reprinted 1983 in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek (Princeton University, Princeton, NJ), p. 324.
Wisdom, J., 1985, Icarus **63**, 272.
Wiseman, H. M., and G. J. Milburn, 1993, Phys. Rev. Lett. **70**, 548.
Wiseman, H. M., and J. A. Vaccaro, 1998, Phys. Lett. A **250**, 241.
Wootters, W. K., and B. D. Fields, 1989, Ann. Phys. (N.Y.) **191**, 363.
Wootters, W. K., and W. H. Zurek, 1979, Phys. Rev. D **19**, 473.
Wootters, W. K., and W. H. Zurek, 1982, Nature (London) **299**, 802.
Zanardi, P., 1998, Phys. Rev. A **57**, 3276.
Zanardi, P., 2001, Phys. Rev. A **63**, 012301.
Zanardi, P., and M. Rasetti, 1997, Phys. Rev. Lett. **79**, 3306.
Zeh, H. D., 1970, Found. Phys. **1**, 69. Reprinted 1983 in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek (Princeton University, Princeton, NJ), p. 342.
Zeh, H. D., 1971, in *Foundations of Quantum Mechanics*, edited by B. d'Espagnat (Academic, New York), p. 263.
Zeh, H. D., 1973, Found. Phys. **3**, 109.
Zeh, H. D., 1986, Phys. Lett. A **116**, 9.
Zeh, H. D., 1988, Phys. Lett. A **172**, 311.
Zeh, H. D., 1990, in *Complexity, Entropy, and the Physics of Information*, edited by W. H. Zurek (Addison-Wesley, Palo Alto), p. 405.
Zeh, H. D., 1992, *The Physical Basis of the Direction of Time* (Springer, Berlin).
Zeh, H. D., 1993, Phys. Lett. A **172**, 189.
Zeh, H. D., 1997, in *New Developments on Fundamental Problems in Quantum Physics*, edited by M. Ferrero and A. van der Merwe (Kluwer, Dordrecht).
Zeh, H. D., 2000, in *Decoherence: Theoretical, Experimental, and Conceptual Problems*, edited by Ph. Blanchard *et al.* (Springer, Berlin), p. 19.
Zurek, W. H., 1981, Phys. Rev. D **24**, 1516.
Zurek, W. H., 1982, Phys. Rev. D **26**, 1862.
Zurek, W. H., 1983, in *Quantum Optics, Experimental Gravitation, and Measurement Theory*, edited by P. Meystre and M. O. Scully (Plenum, New York), p. 87.
Zurek, W. H., 1984a, in *Frontiers in Nonequilibrium Statistical Physics*, edited by G. T. Moore and M. T. Scully (Plenum, New York, 1986), p. 145.
Zurek, W. H., 1984b, Phys. Rev. Lett. **53**, 391.
Zurek, W. H., 1989, Phys. Rev. A **40**, 4731.
Zurek, W. H., 1991, Phys. Today **44** (10), 36.
Zurek, W. H., 1993a, Prog. Theor. Phys. **89**, 281.
Zurek, W. H., 1993b, Phys. Today **46** (4), 13.
Zurek, W. H., 1994, in *From Statistical Physics to Statistical Inference and Back*, edited by P. Grassberger and J.-P. Nadal (Plenum, New York), p. 341.
Zurek, W. H., 1998a, Philos. Trans. R. Soc. London, Ser. A **356**, 1793.
Zurek, W. H., 1998b, Phys. Scr., T **76**, 186.
Zurek, W. H., 2000, Ann. Phys. (Leipzig) **9**, 855.
Zurek, W. H., 2001, Nature (London) **412**, 712.
Zurek, W. H., 2003a, Phys. Rev. A **67**, 012320.
Zurek, W. H., 2003b, Phys. Rev. Lett. **90**, 120404.
Zurek, W. H., S. Habib, and J.-P. Paz, 1993, Phys. Rev. Lett. **70**, 1187.
Zurek, W. H., and J.-P. Paz, 1994, Phys. Rev. Lett. **72**, 2508.
Zurek, W. H., and J.-P. Paz, 1995a, Physica D **83**, 300.
Zurek, W. H., and J.-P. Paz, 1995b, Phys. Rev. Lett. **72**, 2508.