Amit Hagar

Statement

and

Readings

Decoherence and Entanglement

Amit Hagar

Since the inception of quantum mechanics entanglement has been acknowledged, as Schrödinger had put it, as "not one, but The" characteristic feature of the theory that sets it apart from classical physics. Decoherence, on the other hand, was recognized only recently as the "new orthodoxy", crucial to the consistency of quantum theory with our everyday notions of the classical world. In this short introduction I shall present the basic concepts underlying this "new orthodoxy", and place these in the broader context of the philosophy of physics, touching upon topics such as (1) the methodological role of decoherence in the philosophy of quantum theory, (2) the conceptual and historical relation between the foundations of statistical mechanics and decoherence, and (3) the impact of decoherence on quantum information theory, especially on the question of the feasibility of large-scale and computationally superior quantum information processing devices. Finally, I shall also try to raise some doubts about the claim that decoherence plays a role in the so-called "emergence of the classical world".

Readings:

Bacciagaluppi, Guido, "The Role of Decoherence in Quantum Mechanics", The Stanford Encyclopedia of Philosophy (Fall 2008 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2008/entries/qm-decoherence/

Joos, Eric, "Elements of Environmental Decoherence", arXiv:quant-ph/9908008v1.

STANFORD ENCYCLOPEDIA OF PHILOSOPHY

OPEN ACCESS TO THE ENCYCLOPEDIA HAS BEEN MADE POSSIBLE, IN PART, WITH A FINANCIAL CONTRIBUTION FROM THE INDIANA UNIVERSITY LIBRARIES. WE GRATEFULLY ACKNOWLEDGE THIS SUPPORT.

The Role of Decoherence in Quantum Mechanics

First published Mon Nov 3, 2003; substantive revision Thu Aug 23, 2007

Interference phenomena are a well-known and crucial feature of quantum mechanics, the two-slit experiment providing a standard example. There are situations, however, in which interference effects are (artificially or spontaneously) suppressed. We shall need to make precise what this means, but the *theory of decoherence* is the study of (spontaneous) interactions between a system and its environment that lead to such suppression of interference. This study includes detailed modelling of system-environment interactions, derivation of equations ('master equations') for the (reduced) state of the system, discussion of time-scales etc. A discussion of the concept of suppression of interference and a simplified survey of the theory is given in Section 2, emphasising features that will be relevant to the following discussion (and restricted to standard non-relativistic particle quantum mechanics.[1] A partially overlapping field is that of *decoherent histories*, which proceeds from an abstract definition of loss of interference, but which we shall not be considering in any detail.

Decoherence is relevant (or is claimed to be relevant) to a variety of questions ranging from the measurement problem to the arrow of time, and in particular to the question of whether and how the 'classical world' may emerge from quantum mechanics. This entry mainly deals with the role of decoherence in relation to the main problems and approaches in the foundations of quantum mechanics. Section 3 analyses the claim that decoherence solves the measurement problem, as well as the broadening of the problem through the inclusion of environmental interactions, the idea of emergence of classicality, and the motivation for discussing decoherence together with approaches to the foundations of quantum mechanics. Section 4 then reviews the relation of decoherence with some of the main foundational approaches. Finally, in Section 5 we mention suggested applications that would push the role of decoherence even further.

Suppression of interference has of course featured in many papers since the beginning of quantum mechanics, such as Mott's (1929) analysis of alpha-particle tracks. The modern beginnings of decoherence as a subject in its own right are arguably the papers by H. D. Zeh of the early 1970s (Zeh 1970; 1973). Very well known are also the papers by W. Zurek from the early 1980s (Zurek 1981; 1982). Some of these earlier examples of decoherence

(e.g., suppression of interference between left-handed and right-handed states of a molecule) are mathematically more accessible than more recent ones. A concise and readable introduction to the theory is provided by Zurek in *Physics Today* (1991). This article was followed by publication of several letters with Zurek's replies (1993), which highlight controversial issues. More recent surveys are Zeh 1995, which devotes much space to the interpretation of decoherence, and Zurek 2003. The textbook on decoherence by Giulini *et al*. (1996) and the very recent book by Schlosshauer (2007) are also highly recommended.[2]

2. Basics of Decoherence

2.1 Interference and suppression of interference

The two-slit experiment is a paradigm example of an *interference* experiment. One repeatedly sends electrons or other particles through a screen with two narrow slits, the electrons impinge upon a second screen, and we ask for the probability distribution of detections over the surface of the screen. In order to calculate this, one cannot just take the probabilities of passage through the slits, multiply with the probabilities of detection at the screen conditional on passage through either slit, and sum over the contributions of the two slits.[3] There is an additional so-called interference term in the correct expression for the probability, and this term depends on *both* wave components that pass through the slits.

Thus, the experiment shows that the correct description of the electron in terms of quantum wave functions is indeed one in which the wave passes through both slits. The quantum state of the electron is not given by a wave that passes through the upper slit *or* a wave that passes through the lower slit, not even with a probabilistic measure of ignorance.

There are, however, situations in which this interference term is not observed, i.e., in which the classical probability formula applies. This happens for instance when we perform a detection at the slits, whether or not we believe that measurements are related to a 'true' collapse of the wave function (i.e., that only *one* of the components survives the measurement and proceeds to hit the screen). The disappearence of the interference term, however, can happen also spontaneously, even when no 'true collapse' is presumed to happen, namely if some other systems (say, sufficiently many stray cosmic particles scattering off the electron) suitably interact with the wave between the slits and the screen. In this case, the interference term is not observed, because the electron has become *entangled* with the stray particles (see the entry on quantum entanglement and information).[4] The phase relation between the two components which is responsible for interference is well-defined only at the level of the larger system composed of electron and stray particles, and can produce interference only in a suitable experiment including the larger system. Probabilities for results of measurements are calculated *as if* the wave function had collapsed to one or the other of its two components, but the phase relations have merely been distributed over a larger system.

It is this phenomenon of suppression of interference through suitable interaction with the environment that we refer to by 'suppression of interference', and that is studied in the

theory of decoherence.[5] For completeness, we mention the overlapping but distinct concept of *decoherent* (or *consistent*) *histories*. Decoherence in the sense of this abstract formalism is defined simply by the condition that (quantum) probabilities for wave components at a later time may be calculated from those for wave components at an earlier time and the (quantum) conditional probabilities, according to the standard classical formula, i.e., as if the wave had collapsed. There is some controversy, which we leave aside, as to claims surrounding the status of this formalism as a foundational approach in its own right. Without these claims, the formalism is interpretationally neutral and can be useful in describing situations of suppression of interference. Indeed, the abstract definition has the merit of bringing out two conceptual points that are crucial to the idea of decoherence and that will be emphasised in the following: that wave components can be reidentified over time, and that if we do so, we can formally identify 'trajectories' for the system.[6]

2.2 Features of decoherence

The *theory of decoherence* (sometimes also referred to as 'dynamical' decoherence) studies concrete spontaneous interactions that lead to suppression of interference.

Several features of interest arise in models of such interactions (although by no means are all such features common to all models):

- Suppression of interference can be an extremely fast process, depending on the system and the environment considered.[7]
- The environment will tend to couple to and suppress interference between a preferred set of states, be it a discrete set (left- and right- handed states in models of chiral molecules) or some continuous set ('coherent' states of a harmonic oscillator).
- These preferred states can be characterised in terms of their 'robustness' or 'stability' with respect to the interaction with the environment. Roughly speaking, while the system gets entangled with the environment, the states between which interference is suppressed are the ones that get *least* entangled with the environment themselves under further interaction. This point leads us to various further (interconnected) aspects of decoherence.
- First of all, an intuitive picture of the interaction between system and environment can be provided by the analogy with a measurement interaction (see the entries on quantum mechanics and measurement in quantum theory): the environment is 'monitoring' the system, it is spontaneously 'performing a measurement' (more precisely letting the system undergo an interaction as in a measurement) of the preferred states. The analogy to the standard idealised quantum measurements will be very close in the case of, say, the chiral molecule. In the case, say, of the coherent states of the harmonic oscillator, one should think instead of *approximate* measurements of position (or in fact of approximate joint measurements of position and momentum, since information about the time of flight is also recorded in the environment).
- Secondly, the robustness of the preferred states is related to the fact that information

about them is stored in a *redundant* way in the environment (say, because the Schrödinger cat has interacted with so many stray particles — photons, air molecules, dust). This can later be accessed by an observer without further disturbing the system (we measure — however that may be interpreted — whether the cat is alive or dead by intercepting on our retina a small fraction of the light that has interacted with the cat).

- Thirdly, one often says in this context that decoherence induces 'effective superselection rules'. The concept of a (strict) superselection rule is something that requires a generalisation of the formalism of quantum mechanics, and means that there are some observables — called 'classical' in technical terminology — that *commute* with all observables (for a review, see Wightman 1995). Intuitively, these observables are infinitely robust, since no possible interaction can disturb them (at least as long as the interaction Hamiltonian is considered to be an observable). By an effective superselection rule one means that, roughly analogously, certain observables (e.g., chirality) will not be disturbed by the interactions that actually take place. (See also the comments on the charge superselection rule in Section 5 below.)
- Fourthly and perhaps most importantly, robustness has to do with the possibility or reidentifying a component of the wave over time, and thus talking about *trajectories*, whether spatial or not (the component of the electron's wave that goes through the upper slit hits the screen at a particular place with a certain probability; the left-handed component of the state of a chiral molecule at some time t evolves into the left-handed component of the perhaps slightly altered state of the molecule at some later time t'). Notice that in many of the early papers on decoherence the emphasis is on the preferred states themselves, or on how the (reduced) state of the system evolves: notably on how the state of the system becomes approximately diagonal in the basis defined by the preferred states. This emphasis on (so to speak) *kinematical* aspects must not mislead one: the *dynamical* aspects of reidentification over time and trajectory formation are just as important if not the *most* important for the concept of decoherence and its understanding.
- In the case of decoherence interactions of the form of approximate joint position and momentum measurements, the preferred states are obviously Schrödinger waves localised (narrow) in both position and momentum (essentially the 'coherent states' of the system). Indeed, they can be *very* narrow. A speck of dust of radius $a = 10^{-5}$cm floating in the air will have interference suppressed between (position) components with a width ('coherence length') of 10^{-13}cm.[8]
- In this case, the trajectories at the level of the components (the trajectories of the preferred states) will approximate surprisingly well the corresponding classical (Newtonian) trajectories. Intuitively, one can explain this by noting that if the preferred states, which are 'wave packets' that are both narrow in position and remaining narrow (because narrow in momentum), tend to get entangled least with the environment, they will tend to follow more or less undistrubed the Schrödinger equation. But in fact, narrow wave packets will follow approximately Newtonian trajectories (if the external potentials in which they move are uniform enough along the width of the packets: results of this kind are known as 'Ehrenfest theorems'.) Thus, the resulting 'histories' will be close to Newtonian ones (on the relevant

scales).[9] The most intuitive physical example for this are the observed trajectories of alpha particles in a bubble chamber, which are indeed extremely close to Newtonian ones, except for additional tiny 'kinks'.[10]

None of these features are claimed to obtain in all cases of interaction with some environment. It is a matter of detailed physical investigation to assess which systems exhibit which features, and how general the lessons are that we might learn from studying specific models. In particular one should beware of common overgeneralisations. For instance, decoherence does *not* affect only and all 'macroscopic systems'. True, middle-sized objects, say, on the Earth's surface will be very effectively decohered by the air in the atmosphere, and this is an excellent example of decoherence at work. On the other hand, there are also very good examples of decoherence-like interactions affecting microscopic systems, such as in the interaction of alpha particles with the gas in a bubble chamber. And further, there are arguably macroscopic systems for which interference effects are not suppressed. For instance, it has been shown to be possible to sufficiently shield SQUIDS (a type of superconducting devices) from decoherence for the purpose of observing superpositions of different macroscopic currents — contrary to what one had expected (see e.g., Leggett 1984; and esp. 2002, Section 5.4). Anglin, Paz and Zurek (1997) examine some less well-behaved models of decoherence and provide a useful corrective as to the limits of decoherence.

3. Conceptual Appraisal

3.1 Solving the measurement problem?

The fact that interference is typically very well suppressed between localised states of macroscopic objects suggests that it is relevant to why macroscopic objects in fact appear to us to be in localised states. A stronger claim is that decoherence is not only relevant to this question but by itself already provides the complete answer. In the special case of measuring apparatus, it would explain why we never observe an apparatus pointing, say, to two different results, i.e., decoherence would provide a solution to the measurement problem. As pointed out by many authors, however (recently e.g., Adler 2003; Zeh 1995, pp. 14-15), this claim is not tenable.

The measurement problem, in a nutshell, runs as follows. Quantum mechanical systems are described by wave-like mathematical objects (vectors) of which sums (superpositions) can be formed (see the entry on quantum mechanics). Time evolution (the Schrödinger equation) preserves such sums. Thus, if a quantum mechanical system (say, an electron) is described by a superposition of two given states, say, spin in x-direction equal +1/2 and spin in x-direction equal -1/2, and we let it interact with a measuring apparatus that couples to these states, the final quantum state of the composite will be a sum of two components, one in which the apparatus has coupled to (has registered) x-spin = +1/2, and one in which the apparatus has coupled to (has registered) x-spin = -1/2. The problem is that while we may accept the idea of microscopic systems being described by such sums, we cannot even begin to imagine what it would mean for the (composite of electron and) apparatus to be so

described.

Now, what happens if we include decoherence in the description? Decoherence tells us, among other things, that there are plenty of interactions in which differently localised states of macroscopic systems couple to different states of their environment. In particular, the differently localised states of the macroscopic system could be the states of the pointer of the apparatus registering the different *x*-spin values of the electron. By the same argument as above, the composite of electron, apparatus and environment will be a sum of a state corresponding to the environment coupling to the apparatus coupling in turn to the value +1/2 for the spin, and of a state corresponding to the environment coupling to the apparatus coupling in turn to the value -1/2 for the spin. So again we cannot imagine what it would mean for the composite system to be described by such a sum.

We are left with the following choice *whether or not* we include decoherence: either the composite system is not described by such a sum, because the Schrödinger equation actually breaks down and needs to be modified, or it is, but then we need to understand what that means, and this requires giving an appropriate interpretation of quantum mechanics. Thus, decoherence as such does not provide a solution to the measurement problem, at least not unless it is combined with an appropriate interpretation of the wave function. And indeed, as we shall see, some of the main workers in the field such as Zeh (2000) and Zurek (1998) suggest that decoherence is most naturally understood in terms of Everett-like interpretations (see below Section 4.3, and the entries on Everett's relative-state interpretation and on the many-worlds interpretation).

Unfortunately, naive claims of the kind above are still somewhat part of the 'folklore' of decoherence, and deservedly attract the wrath of physicists (e.g., Pearle 1997) and philosophers (e.g., Bub 1999, Chap. 8) alike. (To be fair, this 'folk' position has the merit of attempting to subject measurement interactions to further physical analysis, without assuming that measurements are a fundamental building block of the theory.)

3.2 Compounding the measurement problem

Decoherence is clearly neither a dynamical evolution contradicting the Schrödinger equation, nor a new interpretation of the wave function. As we shall discuss, however, it does both reveal important dynamical effects *within* the Schrödinger evolution, and may be *suggestive* of possible interpretations of the wave function.

As such it has other things to offer to the philosophy of quantum mechanics. At first, however, it seems that discussion of environmental interactions even exacerbates the problems. Intuitively, if the environment is carrying out, without our intervention, lots of approximate position measurements, then the measurement problem ought to apply more widely, also to these spontaneously occurring measurements.

Indeed, while it is well-known that localised states of macroscopic objects spread very slowly under the free Schrödinger evolution (i.e., if there are no interactions), the situation turns out to be different if they are in interaction with the environment. Although the

different components that couple to the environment will be individually incredibly localised, collectively they can have a spread that is many orders of magnitude larger. That is, the state of the object and the environment could be a superposition of zillions of very well localised terms, each with slightly different positions, and which are collectively spread over a *macroscopic distance*, even in the case of everyday objects.[11]

Given that everyday macroscopic objects are particularly subject to decoherence interactions, this raises the question of whether quantum mechanics can account for the appearance of the everyday world even beyond the measurement problem in the strict sense. To put it crudely: if everything is in interaction with everything else, everything is entangled with everything else, and that is a worse problem than the entanglement of measuring apparatuses with the measured probes. And indeed, discussing the measurement problem without taking decoherence (fully) into account may not be enough, as we shall illustrate by the case of some versions of the modal interpretation in Section 4.4.

3.3 Emergence of classicality

What suggests that decoherence may be relevant to the issue of the classical appearance of the everyday world is that *at the level of components* the quantum description of decoherence phenomena can display tantalisingly classical aspects. The question is then whether, if viewed in the context of any of the main foundational approaches to quantum mechanics, these classical aspects can be taken to explain corresponding classical aspects of the phenomena. The answer, perhaps unsurprisingly, turns out to depend on the chosen approach, and in the next section we shall discuss in turn the relation between decoherence and several of the the main approaches to the foundations of quantum mechanics.

Even more generally, one could ask whether the results of decoherence could thus be used to explain the emergence of the entire *classicality of the everyday world*, i.e., to explain both kinematical features such as macroscopic localisation and dynamical features such as approximately Newtonian or Brownian trajectories, *whenever they happen to be* phenomenologically adequate descriptions. As we have mentioned, there are cases in which a classical description is not a good description of a phenomenon, even if the phenomenon involves macroscopic systems. There are also cases, notably quantum *measurements*, in which the classical aspects of the everyday world are only kinematical (definiteness of pointer readings), while the dynamics is highly non-classical (indeterministic response of the apparatus). In a sense, the everyday world is the world of classical concepts as presupposed by Bohr (see the entry on the Copenhagen interpretation) in order to describe in the first place the ‘quantum phenomena’, which *themselves* would thus become a consequence of decoherence (Zeh 1995, p. 33; see also Bacciagaluppi 2002, Section 6.2). The question of explaining the classicality of the everyday world becomes the question of whether one can *derive* from within quantum mechanics the conditions necessary to *discover and practise* quantum mechanics itself, and thus, in Shimony's (1989) words, closing the circle.

In this generality the question is clearly too hard to answer, depending as it does on how far the physical *programme of decoherence* (Zeh 1995, p. 9) can be successfully developed.

We shall thus postpone the (partly speculative) discussion of how far the programme of decoherence might go until Section 5.

4. Decoherence and Approaches to Quantum Mechanics

There is a wide range of approaches to the foundations of quantum mechanics. The term 'approach' here is more appropriate than the term 'interpretation', because several of these approaches are in fact *modifications* of the theory, or at least introduce some prominent new theoretical aspects. A convenient way of classifying these approaches is in terms of their strategies for dealing with the measurement problem.

Some approaches, so-called collapse approaches, seek to modify the Schrödinger equation, so that superpositions of different 'everyday' states do not arise or are very unstable. Such approaches may have intuitively little to do with decoherence since they seek to suppress precisely those superpositions that are created by decoherence. Nevertheless their relation to decoherence is interesting. Among collapse approaches, we shall discuss (in Section 4.1) von Neumann's collapse postulate and theories of spontaneous localisation (see the entry on collapse theories).

Other approaches, known as 'hidden variables' approaches, seek to explain quantum phenomena as equilibrium statistical effects arising from a theory at a deeper level, rather strongly in analogy with attempts at understanding thermodynamics in terms of statistical mechanics (see the entry on philosophy of statistical mechanics). Of these, the most developed are the so-called pilot-wave theories, in particular the theory by de Broglie and Bohm (see the entry on Bohmian mechanics), whose relation to decoherence we discuss in Section 4.2.

Finally, there are approaches that seek to solve the measurement problem strictly by providing an appropriate *interpretation* of the theory. Slightly tongue in cheek, one can group together under this heading approaches as diverse as Everett interpretations (see the entries on Everett's relative-state interpretation and on the many-worlds interpretation), modal interpretations and Bohr's Copenhagen interpretation (Sections 4.3, 4.4 and 4.5, respectively).

We shall be analysing these approaches specifically in their relation to decoherence. For further details and more general assessment or criticism we direct the reader to the relevant entries.

4.1 Collapse approaches

4.1.1 Von Neumann

It is notorious that von Neumann (1932) proposed that the observer's consciousness is somehow related to what he called Process I, otherwise known as the collapse postulate or the projection postulate, which in his book is treated on a par with the Schrödinger equation (his Process II). There is some ambiguity in how to interpret von Neumann. He

may have been advocating some sort of special access to our own consciousness that makes it appear to us that the wave function has collapsed, thus justifying a phenomenological reading of Process I. Alternatively, he may have proposed that consciousness plays some causal role in precipitating the collapse, in which case Process I is a physical process fully on a par with Process II.[12]

In either case, von Neumann's interpretation relies on the insensitivity of the final predictions (for what we consciously record) to exactly where and when Process I is used in modelling the evolution of the quantum system. This is often referred to as the *movability of the von Neumann cut* between the subject and the object, or some similar phrase. Collapse could occur when a particle impinges on a screen, or when the screen blackens, or when an automatic printout of the result is made, or in our retina, or along the optic nerve, or when ultimately consciousness is involved. Before and after the collapse, the Schrödinger equation would describe the evolution of the system.

Von Neumann shows that all of these models are equivalent, as far as the final predictions are concerned, so that he can indeed maintain that collapse is related to consciousness, while in practice applying the projection postulate at a much earlier (and more practical) stage in the description. What allows von Neumann to derive this result, however, is the assumption of *absence of interference* between different components of the wave function. Indeed, if interference were otherwise present, the timing of the collapse would influence the final statistics, just as it would in the case of the two-slit experiment (collapse behind the slits or at the screen). Thus, although von Neumann's is (at least on some readings) a true collapse approach, its reliance on decoherence is in fact crucial.

4.1.2 Spontaneous collapse theories

The best known theory of spontaneous collapse is the so-called GRW theory (Ghirardi Rimini & Weber 1986), in which a material particle spontaneously undergoes *localisation* in the sense that at random times it experiences a collapse of the form used to describe approximate position measurements.[13] In the original model, the collapse occurs independently for each particle (a large number of particles thus 'triggering' collapse much more frequently); in later models the frequency for each particle is weighted by its mass, and the overall frequency for collapse is thus tied to mass density.[14]

Thus, formally, the effect of spontaneous collapse is the same as in some of the models of decoherence, at least for one particle.[15] Two crucial differences on the other hand are that we have 'true' collapse instead of suppression of interference (see above Section 2), and that spontaneous collapse occurs *without* there being any interaction between the system and anything else, while in the case of decoherence suppression of interference obviously arises through interaction with the environment.

Can decoherence be put to use in GRW? The situation may be a bit complex when the decoherence interaction does not approximately privilege position (e.g., currents in a SQUID instead), because collapse and decoherence might actually 'pull' in different directions.[16] But in those cases in which the main decoherence interaction also takes the

form of approximate position measurements, the answer boils down to a quantitative comparison. If collapse happens faster than decoherence, then the superposition of components relevant to decoherence will not have time to arise, and insofar as the collapse theory is successful in recovering classical phenomena, decoherence plays no role in this recovery. Instead, if decoherence takes place faster than collapse, then (as in von Neumann's case) the collapse mechanism can find 'ready-made' structures onto which to truly collapse the wave function. This is indeed borne out by detailed comparison (Tegmark 1993, esp. Table 2). Thus, it seems that decoherence does play a role also in spontaneous collapse theories.

A related point is whether decoherence has implications for the *experimental testability* of spontaneous collapse theories. Indeed, provided decoherence can be put to use also in no-collapse approaches such as pilot-wave or Everett (possibilities that we discuss in the next sub-sections), then in all cases in which decoherence is faster than collapse, what might be interpreted as evidence for collapse could be reinterpreted as 'mere' suppression of interference (think of definite measurement outcomes!), and only cases in which the collapse theory predicts collapse but the system is shielded from decoherence (or perhaps in which the two pull in different directions) could be used to test collapse theories experimentally.

One particularly bad scenario for experimental testability is related to the speculation (in the context of the 'mass density' version) that the cause of spontaneous collapse may be connected with gravitation. Tegmark 1993 (Table 2) quotes some admittedly uncertain estimates for the suppression of interference due to a putative quantum gravity, but they are quantitatively very close to the rate of destruction of interference due to the GRW collapse (at least outside of the microscopic domain). Similar conclusions are arrived at by Kay (1998). If there is indeed such a quantitative similarity between these possible effects, then it would become extremely difficult to distinguish between the two (with the above proviso). In the presence of gravitation, any positive effect could be interpreted as support for either collapse or decoherence. And in those cases in which the system is effectively shielded from decoherence (say, if the experiment is performed in free fall), if the collapse mechanics is indeed triggered by gravitational effects, then no collapse might be expected either. The relation between decoherence and spontaneous collapse theories is thus indeed far from straightforward.

4.2 Pilot-wave theories

Pilot-wave theories are no-collapse formulations of quantum mechanics that assign to the wave function the role of determining the evolution of ('piloting', 'guiding') the variables characterising the system, say particle configurations, as in de Broglie's (1928) and Bohm's (1952) theory, or fermion number density, as in Bell's (1987, Chap. 19) 'beable' quantum field theory, or again field confugurations, as in Valentini's proposals for pilot-wave quantum field theories (Valentini, in preparation; see also Valentini 1996).

De Broglie's idea had been to modify classical Hamiltonian mechanics in such a way as to make it analogous to classical wave optics, by substituting for Hamilton and Jacobi's action

function the phase *S* of a physical wave. Such a 'wave mechanics' of course yields non-classical motions, but in order to understand how de Broglie's dynamics relates to typical quantum phenomena, we must include Bohm's (1952, Part II) analysis of the appearance of collapse. In the case of measurements, Bohm argued that the wave function evolves into a superposition of components that are and remain separated in the total configuration space of measured system and apparatus, so that the total configuration is 'trapped' inside a *single component* of the wave function, which will guide its further evolution, as if the wave had collapsed ('effective' wave function). This analysis allows one to recover qualitatively the measurement collapse and by extension typical quantum features such as the uncertainty principle and the perfect correlations in an EPR experiment (we are ignoring here the well developed quantitative aspects of the theory).

A natural idea is now that this analysis should be extended from the case of measurements induced by an apparatus to that of the 'spontaneous measurements' performed by the environment in the theory of decoherence, thus applying the same strategy for recovering both quantum and classical phenomena. The resulting picture is one in which de Broglie-Bohm theory, in cases of decoherence, would describe the motion of particles that are trapped inside one of the extremely well localised components selected by the decoherence interaction. Thus, de Broglie-Bohm trajectories will partake of the classical motions on the level defined by decoherence (the width of the components). This use of decoherence would arguably resolve the puzzles discussed e.g., by Holland (1996) with regard to the possibility of a 'classical limit' of de Broglie's theory. One baffling problem is for instance that possible trajectories in de Broglie-Bohm theory differing in their initial conditions cannot cross, because the wave guides the particles by way of a first-order equation, while Newton's equations are second-order, as well-known, and possible trajectories do cross. However, the non-interfering components produced by decoherence can indeed cross, and so will the trajectories of particles trapped inside them.

The above picture is natural, but it is not obvious. De Broglie-Bohm theory and decoherence contemplate two a priori *distinct* mechanisms connected to apparent collapse: respectively, separation of components in configuration space and suppression of interference. While the former obviously implies the latter, it is equally obvious that decoherence need not imply separation in configuration space. One can expect, however, that decoherence interactions of the form of approximate position measurements will.

If the main instances of decoherence are indeed coextensive with instances of separation in configuration, de Broglie-Bohm theory can thus *use* the results of decoherence relating to the formation of classical structures, while providing an interpretation of quantum mechanics that explains why these structures are indeed observationally relevant. The question that arises for de Broglie-Bohm theory is then the extension of the well-known question of whether all apparent *measurement* collapses can be associated with separation in configuration (by arguing that at some stage all measurement results are recorded in macroscopically different configurations) to the question of whether *all* appearance of classicality can be associated with separation in configuration space.[17]

A discussion of the role of decoherence in pilot-wave theory in the form suggested above is

still largely outstanding. An informal discussion is given in Bohm and Hiley (1993, Chap. 8), partial results are given by Appleby (1999), and a different approach is suggested by Allori (2001; see also Allori & Zanghì 2001). Appleby discusses trajectories in a model of decoherence and obtains approximately classical trajectories, but under a special assumption.[18] Allori investigates in the first place the 'short wavelength' limit of de Broglie-Bohm theory (suggested by the analogy to the geometric limit in wave optics). The role of decoherence in her analysis is crucial but limited to *maintaining* the classical behaviour obtained under the appropriate short wavelength conditions, because the behaviour would otherwise break down after a certain time.

4.3 Everett interpretations

Everett interpretations are very diverse, and possibly only share the core intuition that a *single* wave function of the universe should be interpreted in terms of a *multiplicity* of 'realities' at some level or other. This multiplicity, however understood, is formally associated with *components* of the wave function in some decomposition.[19]

Various Everett interpretations, roughly speaking, differ as to how to *identify* the relevant components of the universal wave function, and how to *justify* such an identification (the so-called problem of the 'preferred basis' — although this may be a misnomer), and differ as to how to *interpret* the resulting multiplicity (various 'many-worlds' or various 'many-minds' interpretations), in particular with regard to the interpretation of the (emerging?) *probabilities* at the level of the components (problem of the 'meaning of probabilities').

The last problem is perhaps the most hotly debated aspect of Everett. Clearly, decoherence enables reidentification over time of both observers and of results of repeated measurement and thus definition of empirical frequencies. In recent years progress has been made especially along the lines of interpreting the probabilities in decision-theoretic terms for a 'splitting' agent (see in particular Wallace 2003b, and its longer preprint, Wallace 2002).[20]

The most useful application of decoherence to Everett, however, seems to be in the context of the problem of the preferred basis. Decoherence seems to yield a (maybe partial) solution to the problem, in that it naturally identifies a class of 'preferred' states (not necessarily an orthonormal basis!), and even allows to reidentify them over time, so that one can identify 'worlds' with the trajectories defined by decoherence (or more abstractly with decoherent histories).[21] If part of the aim of Everett is to interpret quantum mechanics without introducing extra structure, in particular without *postulating* the existence of some preferred basis, then one will try to identify structure that is already present in the wave function at the level of components (see e.g., Wallace, 2003a). In this sense, decoherence is an ideal candidate for identifying the relevant components.

A *justification* for this identification can then be variously given by suggesting that a 'world' should be a *temporally extended* structure and thus reidentification over time will be a necessary condition for identifying worlds, or similarly by suggesting that in order for observers to *evolve* there must be *stable records* of past events (Saunders 1993, and the

unpublished Gell-Mann & Hartle 1994 (see the Other Internet Resources section below), or that observers must be able to access *robust states*, preferably through the existence of redundant information in the environment (Zurek's 'existential interpretation', 1998).

In alternative to some global notion of 'world', one can look at the components of the (mixed) state of a (local) system, either from the point of view that the different components defined by decoherence will separately affect (different components of the state of) another system, or from the point of view that they will separately underlie the conscious experience (if any) of the system. The former sits well with Everett's (1957) original notion of relative state, and with the relational interpretation of Everett preferred by Saunders (e.g., 1993) and, it would seem, Zurek (1998). The latter leads directly to the idea of many-minds interpretations (see the entry on Everett's relative-state interpretation and the website on 'A Many-Minds Interpretation of Quantum Theory' referenced in the Other Internet Resources). If one assumes that mentality can be associated only with certain decohering structures of great complexity, this might have the advantage of further reducing the remaining ambiguity about the preferred 'basis'.

The idea of many minds was suggested early on by Zeh (2000; also 1995, p. 24). As Zeh puts it, von Neumann's motivation for introducing collapse was to save what he called psycho-physical parallelism (arguably supervenience of the mental on the physical: only one mental state is experienced, so there should be only one corresponding component in the physical state). In a decohering no-collapse universe one can instead introduce a *new* psycho-physical parallelism, in which individual minds supervene on each non-interfering component in the physical state. Zeh indeed suggests that, given decoherence, this is the most natural interpretation of quantum mechanics.[22]

4.4 Modal interpretations

Modal interpretations originated with Van Fraassen (1973, 1991) as pure reinterpretations of quantum mechanics (other later versions coming to resemble more hidden variables theories). Van Fraassen's basic intuition was that the quantum state of a system should be understood as describing a collection of possibilities, represented by components in the (mixed) quantum state. His proposal considers only decompositions at single instants, and is agnostic about reidentification over time. Thus, it can directly exploit only the fact that decoherence produces descriptions in terms of classical-like states, which will count as possibilities in Van Fraassen's interpretation. This ensures 'empirical adequacy' of the quantum description (a crucial concept in Van Fraassen's philosophy of science). The dynamical aspects of decoherence can be exploited indirectly, in that single-time components will exhibit *records* of the past, which ensure adequacy with respect to observations, but about whose veridicity Van Fraassen remains agnostic.

A different strand of modal interpretations is loosely associated with the (distinct) views of Kochen (1985), Healey (1989) and Dieks and Vermaas (e.g., 1998). We focus on the last of these to fix the ideas. Van Fraassen's possible decompositions are restricted to one singled out by a mathematical criterion (related to the so-called biorthogonal decomposition theorem), and a dynamical picture is explicitly sought (and was later developed). In the

case of an ideal (non-approximate) quantum measurement, this special decomposition coincides with that defined by the eigenstates of the measured observable and the corresponding pointer states, and the interpretation thus appears to solve the measurement problem (in the strict sense).

At least in Dieks's original intentions, however, the approach was meant to provide an attractive interpretation of quantum mechanics also in the case of decoherence interactions, since at least in simple models of decoherence the same kind of decomposition singles out more or less also those states between which interference is suppressed (with a proviso about very degenerate states).

However, this approach fails badly when applied to other models of decoherence, e.g., that in Joos and Zeh (1985, Section III.2). Indeed, it appears that in general the components singled out by this version of the modal interpretation are given by *delocalised* states, as opposed to the components arising naturally in the theory of decoherence (Bacciagaluppi 2000; Donald 1998). Notice that van Fraassen's original interpretatioin is untouched by this problem, and so are possibly some more recent modal or modal-like interpretations by Spekkens and Sipe (2001), Bene and Dieks (2002) and Berkovitz and Hemmo (in preparation).

Finally, some of the views espoused in the decoherent histories literature could be considered as cognate to Van Fraassen's views, identifying possibilities, however, at the level of possible courses of world history. Such ‘possible worlds’ would be those temporal sequences of (quantum) propositions that satisfy the decoherence condition and in this sense support a description in terms of a probabilistic evolution. This view would be using decoherence as an essential ingredient, and in fact may turn out to be the most fruitful way yet of implementing modal ideas; a discussion in these terms still needs to be carried out in detail, but see Hemmo (1996).

4.5 Bohr's Copenhagen interpretation

It appears that Bohr held more or less the following view. Everyday concepts, in fact the concepts of classical physics, are indispensable to the description of any physical phenomena (in a way — and terminology — much reminiscent of Kant's transcendental arguments). However, experimental evidence from atomic phenomena shows that classical concepts have fundamental limitations in their applicability: they can only give partial (complementary) pictures of physical objects. While these limitations are quantitatively negligible for most purposes in dealing with macroscopic objects, they apply also at that level (as shown by Bohr's willingness to apply the uncertainty relations to parts of the experimental apparatus in the Einstein-Bohr debates), and they are of paramount importance when dealing with microscopic objects. Indeed, they shape the characteristic features of quantum phenomena, e.g., indeterminism. The quantum state is not an ‘intuitive’ (*anschaulich*, also translated as ‘visualisable’) representation of a quantum object, but only a ‘symbolic’ representation, a shorthand for the quantum phenomena constituted by applying the various complementary classical pictures.

While it is difficult to pinpoint exactly what Bohr's views were (the concept and even the term ‘Copenhagen interpretation’ appear to be a later construct; see Howard 2003), it is clear that according to Bohr, classical concepts are autonomous from, and indeed conceptually prior to, quantum theory. If we understand the theory of decoherence as pointing to how classical concepts might in fact emerge from quantum mechanics, this seems to undermine Bohr's basic position. Of course it would be a mistake to say that decoherence (a part of quantum theory) *contradicts* the Copenhagen approach (an interpretation of quantum theory). However, decoherence does suggest that one might want to adopt alternative interpretations, in which it is the quantum concepts that are prior to the classical ones, or, more precisely, the classical concepts at the everyday level emerge from quantum mechanics (irrespectively of whether there are even more fundamental concepts, as in pilot-wave theories). In this sense, if the programme of decoherence is successful as sketched in Section 3.3, it will indeed be a blow to Bohr's *interpretation* coming from quantum physics itself.

On the other hand, Bohr's *intuition* that quantum mechanics as practised requires a classical domain would in fact be *confirmed* by decoherence, if it turns out that decoherence is indeed the basis for the phenomenology of quantum mechanics, as the Everettian and possibly the Bohmian analysis suggest. As a matter of fact, Zurek (2003) locates his existential interpretation half-way between Bohr and Everett. It is perhaps a gentle irony that in the wake of decoherence, the foundations of quantum mechanics might end up re-evaluating this part of Bohr's thinking.

5. Scope of Decoherence

We have already mentioned in Section 2.2 that some care has to be taken lest one overgeneralise conclusions based on examining only well-behaved models of decoherence. On the other hand, in order to assess the programme of explaining the emergence of classicality using decoherence (together with appropriate foundational approaches), one has to probe *how far* the applications of decoherence can be pushed. In this final section, we survey some of the further applications that have been proposed for decoherence, beyond the easier examples we have seen such as chirality or alpha-particle tracks. Whether decoherence can indeed be successfully applied to all of these fields will be in part a matter for further assessment, as more detailed models are proposed.

A straightforward application of the techniques allowing one to derive Newtonian trajectories at the level of components has been employed by Zurek and Paz (1994) to derive *chaotic trajectories* in quantum mechanics. The problem with the quantum description of chaotic behaviour is that *prima facie* there should be none. Chaos is characterised roughly as extreme sensitivity in the behaviour of a system on its initial conditions, where the distance between the trajectories arising from different initial conditions increases exponentially in time. Since the Schrödinger evolution is *unitary*, it preserves all scalar products and all distances between quantum state vectors. Thus, it would seem, close initial conditions lead to trajectories that are uniformly close throughout all of time, and no chaotic behaviour is possible (‘problem of quantum chaos’). The crucial

point that enables Zurek and Paz' analysis is that the relevant trajectories in decoherence theory are at the level of *components* of the state of the system. Unitarity is preserved because the vectors in the environment to which these different components are coupled, are and remain orthogonal: how the components themselves evolve is immaterial. Explicit modelling yields a picture of quantum chaos in which different trajectories branch (a feature absent from classical chaos, which is deterministic) and then indeed diverge exponentially. As with the crossing of trajectories in de Broglie-Bohm theory (Section 4.2), one has behaviour at the level of components that is qualitatively different from the behaviour derived from wave functions of an isolated system.

The idea of effective superselection rules was mentioned in Section 2.2. As pointed out by Giulini, Kiefer and Zeh (1995, see also Giulini *et al*. 1996, Section 6.4), the justification for the (strict) superselection rule for charge in quantum field theory can also be phrased in terms of decoherence. The idea is simple: an electric charge is surrounded by a Coulomb field (which electrostatically is infinitely extended; the argument can also be carried through using the retarded field, though). States of different electric charge of a particle are thus coupled to different, presumably orthogonal, states of its electric field. One can consider the far-field as an effectively uncontrollable environment that decoheres the particle (and the near-field), so that superpositions of different charges are indeed never observed.

Another claim about the significance of decoherence relates to time asymmetry (see e.g., the entries on time asymmetry in thermodynamics and philosophy of statistical mechanics), in particular of whether decoherence can explain the apparent time-directedness in our (classical) world. The issue is again one of time-directedness at the level of components emerging from a time-symmetric evolution at the level of the universal wave function (presumably with special initial conditions). Insofar as (apparent) collapse is indeed a time-directed process, decoherence will have direct relevance to the emergence of this 'quantum mechanical arrow of time' (for a spectrum of discussions, see Zeh 2001, Chap. 4; Hartle 1998, and references therein; and Bacciagaluppi 2002, Section 6.1). Whether decoherence is connected to the other familiar arrows of time is a more specific question, various discussions of which are given, e.g., by Zurek and Paz (1994), Hemmo and Shenker (2001) and the unpublished Wallace (2001) (see the Other Internet Resources Section below).

In a recent paper, Zeh (2003) argues from the notion that decoherence can explain 'quantum phenomena' such as *particle detections* that the concept of a particle in quantum field theory is itself a consequence of decoherence. That is, only fields need to be included in the fundamental concepts, and 'particles' are a derived concept, unlike what is suggested by the customary introduction of fields through a process of 'second quantisation'. Thus decoherence seems to provide a further powerful argument for the conceptual primacy of fields over particles in the question of the interpretation of quantum field theory.

Finally, it has been suggested that decoherence could be a useful ingredient in a theory of quantum gravity, for two reasons. First, because a suitable generalisation of decoherence theory to a full theory of quantum gravity should yield suppression of interference between

different classical spacetimes (Giulini *et al*. 1996, Section 4.2). Second, it is speculated that decoherence might solve the so-called *problem of time*, which arises as a prominent puzzle in (the 'canonical' approach to) quantum gravity. This is the problem that the candidate fundamental equation (in this approach) — the Wheeler-DeWitt equation — is an analogue of a time-*independent* Schrödinger equation, and does not contain time at all. The problem is thus simply: where does time come from? In the context of decoherence theory, one can construct toy models in which the analogue of the Wheeler-DeWitt wave function decomposes into non-interfering components (for a suitable sub-system) each satisfying a time-*dependent* Schrödinger equation, so that decoherence appears in fact as the source of time.[23] An accessible introduction to and philosophical discussion of these models is given by Ridderbos (1999), with references to the original papers.

Bibliography

- Adler, S. L. (2003), 'Why Decoherence has not Solved the Measurement Problem: A Response to P. W. Anderson', *Studies in History and Philosophy of Modern Physics* **34B**, 135-142. [Preprint available online]
- Albert, D., and Loewer, B. (1988), 'Interpreting the Many Worlds Interpretation', *Synthese* **77**, 195-213.
- Allori, V. (2001), *Decoherence and the Classical Limit of Quantum Mechanics*, Ph.D. Thesis, Università di Genova, Dipartimento di Fisica.
- Allori, V., and Zanghì, N. (2001), 'On the Classical Limit of Quantum Mechanics', *International Journal of Theoretical Physics*, forthcoming. [Preprint available online]
- Anglin, J. R., Paz, J. P., and Zurek, W. H. (1997), 'Deconstructing Decoherence', *Physical Review* **A 55**, 4041-4053. [Preprint available online]
- Appleby, D. M. (1999), 'Bohmian Trajectories Post-Decoherence', *Foundations of Physics* **29**, 1885-1916. [Preprint available online]
- Bacciagaluppi, G. (2000), 'Delocalized Properties in the Modal Interpretation of a Continuous Model of Decoherence', *Foundations of Physics* **30**, 1431-1444.
- Bacciagaluppi, G. (2002), 'Remarks on Space-Time and Locality in Everett's Interpretation', in T. Placek and J. Butterfield (eds), *Non-Locality and Modality*, NATO Science Series, II. Mathematics, Physics and Chemistry, Vol. 64 (Dordrecht: Kluwer), pp. 105-122. [Preprint available online]
- Barbour, J. (1999), *The End of Time* (London: Weidenfeld and Nicolson).
- Bell, J. S. (1987), *Speakable and Unspeakable in Quantum Mechanics* (Cambridge: Cambridge University Press).
- Bene, G., and Dieks, D. (2002), 'A Perspectival Version of the Modal Interpretation of Quantum Mechanics and the Origin of Macroscopic Behavior', *Foundations of Physics* **32**, 645-672. [Preprint available online]
- Berkovitz, J., and Hemmo, M. (in preparation), 'Modal Interpretations and Relativity: A Reconsideration'.
- Broglie, L. de (1928), 'La nouvelle dynamique des quanta', in H. Lorentz (ed.), *Électrons et Photons: Rapports et Discussions du Cinquième Conseil de Physique [...] Solvay* (Paris: Gauthiers-Villars).
- Bohm, D. (1952), 'A Suggested Interpretation of the Quantum Theory in Terms of

"Hidden" Variables. 'I' and 'II', *Physical Review* **85**, 166-179 and 180-193.
- Bohm, D., and Hiley, B. (1993), *The Undivided Universe* (London: Routledge).
- Bub, J. (1999), *Interpreting the Quantum World* (Cambridge: Cambridge University Press, second edition).
- Cushing, J. T., Fine, A., and Goldstein, S. (1996), *Bohmian Mechanics and Quantum Theory: An Appraisal* (Dordrecht: Kluwer).
- DeWitt, B. S. (1971), 'The Many-Universes Interpretation of Quantum Mechanics', in B. d'Espagnat (ed.), *Foundations of Quantum Mechanics*, Proceedings of the International School of Physics 'Enrico Fermi', Vol. 49 (New York: Academic Press). Reprinted in B. S. DeWitt and N. Graham (eds), *The Many-Worlds Interpretation of Quantum Mechanics* (Princeton: Princeton University Press, 1973), pp. 167-218.
- Dieks, D., and Vermaas, P. E. (eds) (1998), *The Modal Interpretation of Quantum Mechanics* (Dordrecht: Kluwer).
- Donald, M. (1998) 'Discontinuity and Continuity of Definite Properties in the Modal Interpretation', in Dieks and Vermaas (1998), pp. 213-222. [Preprint available online in PDF]
- Dowker, F., and Kent, A. (1995), 'Properties of Consistent Histories', *Physical Review Letters* **75**, 3038-3041. [Preprint available online]
- Epstein, S. T. (1953), 'The Causal Interpretation of Quantum Mechanics', *Physical Review* **89**, 319.
- Everett, H. III (1957), '"Relative-State" Formulation of Quantum Mechanics', *Reviews of Modern Physics* **29**, 454-462. Reprinted in Wheeler and Zurek (1983), pp. 315-323.
- Fraassen, B. van (1973), 'Semantic Analysis of Quantum Logic', in C. A. Hooker (ed.), *Contemporary Research in the Foundations and Philosophy of Quantum Theory* (Dordrecht: Reidel), pp. 180-213.
- Fraassen, B. van (1991), *Quantum Mechanics: An Empiricist View* (Oxford: Clarendon Press).
- Ghirardi, G., Rimini, A., and Weber, T. (1986), 'Unified Dynamics for Microscopic and Macroscopic Systems', *Physical Review* **D 34**, 470-479.
- Giulini, D., Joos, E., Kiefer, C., Kupsch, J., Stamatescu, I.-O., and Zeh, H. D. (1996), *Decoherence and the Appearance of a Classical World in Quantum Theory* (Berlin: Springer; second revised edition, 2003).
- Halliwell, J. J. (1995), 'A Review of the Decoherent Histories Approach to Quantum Mechanics', *Annals of the New York Academy of Sciences* **755**, 726-740. [Preprint available online]
- Halliwell, J. J., and Thorwart, J. (2002), 'Life in an Energy Eigenstate: Decoherent Histories Analysis of a Model Timeless Universe', *Physical Review* **D 65**, 104009-104027. [Preprint available online]
- Hartle, J. B. (1998), 'Quantum Pasts and the Utility of History', *Physica Scripta* **T 76**, 67-77. [Preprint available online]
- Healey, R. (1989), *The Philosophy of Quantum Mechanics: An Interactive Interpretation* (Cambridge: Cambridge University Press).
- Hemmo, M. (1996), *Quantum Mechanics Without Collapse: Modal Interpretations,*

Histories and Many Worlds, Ph.D. Thesis, University of Cambridge, Department of History and Philosophy of Science.
- Hemmo, M. and Shenker, O. (2001) 'Can we Explain Thermodynamics by Quantum Decoherence?', *Studies in History and Philosophy of Modern Physics* **32 B**, 555-568.
- Holland, P. R. (1996), 'Is Quantum Mechanics Universal?', in Cushing, Fine and Goldstein (1996), pp. 99-110.
- Howard, D. (2003), 'Who Invented the Copenhagen Interpretation? A Study in Mythology', talk given at the One-Day Conference in Memory of Jim Cushing, Faculty of Philosophy, Oxford, 26 June 2003.
- Joos, E. and Zeh, H. D. (1985), 'The Emergence of Classical Properties through Interaction with the Environment', *Zeitschrift für Physik* **B 59**, 223-243.
- Kay, B. S. (1998), 'Decoherence of Macroscopic Closed Systems within Newtonian Quantum Gravity', *Classical and Quantum Gravity* **15**, L89-L98. [Preprint available online]
- Kochen, S. (1985), 'A new Interpretation of Quantum Mechanics', in P. Mittelstaedt and P. Lahti (eds), *Symposium on the Foundations of Modern Physics 1985* (Singapore: World Scientific), pp. 151-169.
- Leggett, A. J. (1984), 'Schrödinger's Cat and her Laboratory Cousins', *Contemporary Physics* **25**, 583-594.
- Leggett, A. J. (2002), 'Testing the Limits of Quantum Mechanics: Motivation, State of Play, Prospects', *Journal of Physics* **C 14**, R415-R451.
- Mott, N. F. (1929), 'The Wave Mechanics of α-ray Tracks', *Proceedings of the Royal Society of London* **A 126** (1930, No. 800 of 2 December 1929), 79-84.
- Neumann, J. von (1932), *Mathematische Grundlagen der Quantenmechanik* (Berlin: Springer). Translated as *Mathematical Foundations of Quantum Mechanics* (Princeton: Princeton University Press, 1955).
- Pearle, P. (1997), 'True Collapse and False Collapse', in Da Hsuan Feng and Bei Lok Hu (eds), *Quantum Classical Correspondence: Proceedings of the 4th Drexel Symposium on Quantum Nonintegrability, Philadelphia, PA, USA, September 8-11, 1994* (Cambridge, MA: International Press), pp. 51-68. [Preprint available online]
- Pearle, P. (1989), 'Combining Stochastic Dynamical State-vector Reduction with Spontaneous Localization', *Physical Review* **A 39**, 2277-2289.
- Pearle, P., and Squires, E. (1994), 'Bound-State Excitation, Nucleon Decay Experiments, and Models of Wave-Function Collapse', *Physical Review Letters*, **73**, 1-5.
- Ridderbos, K. (1999), 'The Loss of Coherence in Quantum Cosmology', *Studies in History and Philosophy of Modern Physics* **30 B** , 41-60.
- Saunders, S. (1993), 'Decoherence, Relative States, and Evolutionary Adaptation', *Foundations of Physics* **23**, 1553-1585.
- Saunders, S. (1999), 'The "Beables" of Relativistic Pilot-Wave Theory', in J. Butterfield and C. Pagonis (eds), *From Physics to Philosophy* (Cambridge: Cambridge University Press), pp. 71-89.
- Saunders, S. (2004), 'Operational Derivation of the Born Rule', *Proceedings of the Royal Society of London* **460**, 1-18. [Preprint available online]
- Schlosshauer, M. (2007), *Decoherence and the Quantum-to-Classical Transition*

(Springer: Heidelberg/Berlin, 1st ed.).
- Shimony, A. (1989), 'Search for a Worldview which can Accommodate our Knowledge of Microphysics', in J. T. Cushing and E. McMullin (eds), *Philosophical Consequences of Quantum Theory* (Notre Dame, Indiana: University of Notre Dame Press). Reprinted in A. Shimony, *Search for a Naturalistic Worldview*, Vol. 1 (Cambridge: Cambridge University Press, 1993), pp. 62-76.
- Spekkens, R. W., and Sipe, J. E. (2001), 'A Modal Interpretation of Quantum Mechanics based on a Principle of Entropy Minimization', *Foundations of Physics* **31**, 1431-1464.
- Tegmark, M. (1993), 'Apparent Wave Function Collapse Caused by Scattering', *Foundations of Physics Letters* **6**, 571-590. [Preprint available online]
- Valentini, A. (1996), 'Pilot-Wave Theory of Fields, Gravitation and Cosmology', in Cushing, Fine and Goldstein (1996), pp. 45-66.
- Valentini, A. (in preparation), *Pilot-Wave Theory* (Cambridge: Cambridge University Press).
- Van Fraassen, B., see Fraassen, B. van.
- Von Neumann, J., see Neumann, J. von.
- Wallace, D. (2003a), 'Everett and Structure', *Studies in History and Philosophy of Modern Physics* **34 B**, 87-105. [Preprint available online]
- Wallace, D. (2003b), 'Everettian Rationality: Defending Deutsch's Approach to Probability in the Everett Interpretation', *Studies in History and Philosophy of Modern Physics* **34 B**, 415-439. [Preprint available online] [See also the longer, unpublished version titled 'Quantum Probability and Decision Theory, Revisited' referenced in the Other Internet Resources.]
- Wheeler, J. A., and Zurek, W. H. (1983) (eds), *Quantum Theory and Measurement* (Princeton: Princeton University Press).
- Wightman, A. S. (1995), 'Superselection Rules; Old and New', *Il Nuovo Cimento* **110 B**, 751-769.
- Zeh, H. D. (1970), 'On the Interpretation of Measurement in Quantum Theory', *Foundations of Physics* **1**, 69-76. Also reprinted in Wheeler and Zurek (1983), pp. 342-349.
- Zeh, H. D. (1973), 'Toward a Quantum Theory of Observation', *Foundations of Physics* **3**, 109-116.
- Zeh, H. D. (1995), 'Basic Concepts and Their Interpretation'. Revised edition of Chapter 2 of Giulini *et al*. (1996). [Page numbers refer to the preprint available online, entitled 'Decoherence: Basic Concepts and Their Interpretation'.]
- Zeh, H. D. (2000), 'The Problem of Conscious Observation in Quantum Mechanical Description', *Foundations of Physics Letters* **13**, 221-233. [Preprint available online]
- Zeh, H. D. (2001), *The Physical Basis of the Direction of Time* (Berlin: Springer, 4th ed.).
- Zeh, H. D. (2003), 'There is no "First" Quantization', *Physics Letters* **A 309**, 329-334. [Preprint available online]
- Zurek, W. H. (1981), 'Pointer Basis of Quantum Apparatus: Into what Mixture does the Wave Packet Collapse?', *Physical Review* **D 24**, 1516-1525.
- Zurek, W. H. (1982), 'Environment-Induced Superselection Rules', *Physical Review*

D 26, 1862-1880.
- Zurek, W. H. (1991), 'Decoherence and the Transition from Quantum to Classical', *Physics Today* **44** (October), 36-44. [Abstract and updated (2003) version available online, under the title 'Decoherence and the Transition from Quantum to Classical — Revisited'.]
- Zurek, W. H. (1993), 'Negotiating the Tricky Border Between Quantum and Classical', *Physics Today* **46** (April), 84-90.
- Zurek, W. H. (1998), 'Decoherence, Einselection, and the Existential Interpretation (The Rough Guide)', *Philosophical Transactions of the Royal Society of London* **A 356**, 1793-1820. [Preprint available online]
- Zurek, W. H. (2003), 'Decoherence, Einselection, and the Quantum Origins of the Classical', *Reviews of Modern Physics* **75**, 715-775. [Page numbers refer to the preprint available online.]
- Zurek, W. H., and Paz, J.-P. (1994), 'Decoherence, Chaos, and the Second Law', *Physical Review Letters* **72**, 2508-2511.

Other Internet Resources

- Gell-Mann, M. (Santa Fe Institute), and Hartle, J.B. (UC/Santa Barbara), 1994, 'Equivalent Sets of Histories and Multiple Quasiclassical Realms', available online in the arXiv.org e-Print archive.
- Wallace, D. (Oxford University), 2000, 'Implications of Quantum Theory in the Foundations of Statistical Mechanics', available online in the Pittsburgh Phil-Sci Archive.
- Wallace, D. (Oxford University), 2002, 'Quantum Probability and Decision Theory, Revisited', available online in the arXiv.org e-Print archive. This is a longer version of Wallace (2003b).
- The arXiv.org e-Print archive, formerly the Los Alamos archive. This is the main physics preprint archive; most of the links above are to this archive.
- The Pittsburgh Phil-Sci Archive. This is the main philosophy of science preprint archive; some of the links above are to this archive.
- A Many-Minds Interpretation Of Quantum Theory, maintained by Matthew Donald (Cavendish Lab, Physics, University of Cambridge). This page contains details of his many-minds interpretation, as well as discussions of some of the books and papers quoted above (and others of interest). Follow also the link to the 'Frequently Asked Questions', some of which (and the ensuing dialogue) contain useful discussion of decoherence.
- Quantum Mechanics on the Large Scale, maintained by Philip Stamp (Physics, University of British Columbia). This page has links to the available talks from the Vancouver workshop mentioned in footnote 2; see especially the papers by Tony Leggett and by Philip Stamp.
- Decoherence Website, maintained by Erich Joos. This is a site with information, references and further links to people and institutes working on decoherence, especially in Germany and the rest of Europe.

Related Entries

Einstein, Albert: Einstein-Bohr debates | quantum mechanics | quantum mechanics: Bohmian mechanics | quantum mechanics: collapse theories | quantum mechanics: Copenhagen interpretation of | quantum mechanics: Everett's relative-state formulation of | quantum mechanics: many-worlds interpretation of | quantum theory: measurement in | quantum theory: quantum entanglement and information | quantum theory: quantum field theory | quantum theory: quantum gravity | quantum theory: the Einstein-Podolsky-Rosen argument in | statistical physics: philosophy of statistical mechanics | time: thermodynamic asymmetry in | Uncertainty Principle

Acknowledgments

I wish to think many people in discussion with whom I have shaped my understanding of decoherence over the years, in particular Marcus Appleby, Matthew Donald, Beatrice Filkin, Meir Hemmo, Simon Saunders, David Wallace and Wojtek Zurek. For more recent discussions and correspondence relating to this article I wish to thank Valia Allori, Peter Holland, Martin Jones, Tony Leggett, Hans Primas, Alberto Rimini, Philip Stamp and Bill Unruh. I also gratefully acknowledge my debt to Steve Savitt and Philip Stamp for an invitation to talk at the University of British Columbia, and to Claudius Gros for an invitation to the University of the Saarland, and for the opportunities for discussion arising from these talks. Finally I wish to thank the referee of this entry, again David Wallace, for his clear and constructive commentary, my fellow subject editor John Norton, who corresponded with me extensively over a previous version of part of the material and whose suggestions I have taken to heart, my editor-in-chief Edward N. Zalta for his saintly patience, and my friend and predecessor as subject editor, the late Rob Clifton, who invited me to write on this topic in the first place.

Elements of Environmental Decoherence*

Erich Joos
Rosenweg 2, D-22869 Schenefeld, Germany

Abstract

In this contribution I give an introduction to the essential concepts and mechanisms of decoherence by the environment. The emphasis will be not so much on technical details but rather on conceptual issues and the impact on the interpretation problem of quantum theory.

1 What is decoherence?

Decoherence is the irreversible formation of quantum correlations of a system with its environment. These correlations lead to entirely new properties and behavior compared to that shown by isolated objects.

Whenever we have a product state of two interacting systems - a very special state - the unitary evolution according to the Schrödinger equation will lead to entanglement,

$$\begin{aligned} |\varphi\rangle|\Phi\rangle \quad &\stackrel{t}{\longrightarrow} \quad \sum_{n,m} c_{nm}|\varphi_n\rangle|\Phi_m\rangle \\ &= \quad \sum_n \sqrt{p_n(t)}|\tilde{\varphi}_n(t)\rangle|\tilde{\Phi}_n(t)\rangle. \end{aligned} \tag{1}$$

The rhs of (1) can no longer be written as a single product in the general case. This can also be described by using the Schmidt representation, shown in the second line, where the presence of more than one component is equivalent to the existence of quantum correlations.

If many degrees of freedom are involved in this process, this entanglement will become practically irreversible, except for very special situations. Decoherence is thus a quite normal and, moreover, ubiquitous, quantum mechanical process. Historically, the important observation was that this de-separation of quantum states happens extremely fast for macroscopic objects [17]. The natural environment cannot simply be ignored or treated as a classical background in this case.

Equation (1) shows that there is an intimate connection to the theory of irreversible processes. However, decoherence must not be identified or confused with

*To be published in the proceedings of the Bielefeld conference on "Decoherence: Theoretical, Experimental, and Conceptual Problems", edited by P. Blanchard, D. Giulini, E. Joos, C. Kiefer, and I.-O. Stamatescu (Springer 1999).

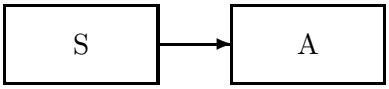
dissipation: decoherence precedes dissipation by acting on a much faster timescale, while requiring initial conditions which are essentially the same as those responsible for the thermodynamic arrow of time [18].

When we consider observations at one of the two systems, we see various consequences of this entanglement. First of all, our considered subsystem will no longer obey a Schrödinger equation, the local dynamics is in general very complicated, but can often be approximated by some sort of master equation (The Schmidt decomposition is directly related to the subsystem density matrices). The most important effect is the disappearance of phase relations (i.e., interference) between certain subspaces of the Hilbert space of the system. Hence the resulting superselection rules can be understood as emerging from a dynamical, approximate and time-directed process. If the coupling to the environment is very strong, the internal dynamics of the system may become slowed down or even frozen. This is now usually called the quantum Zeno effect, which apparently does not occur in our macroscopic world.

The details of the dynamics depend on the kind of coupling between the system we consider and its environment. In many cases – especially in the macroscopic domain – this coupling leads to an evolution similar to a measurement process. Therefore it is appropriate to recall the essential elements of the quantum theory of measurement.

1.1 Dynamical Description of Measurement

The standard description of measurement was laid down by von Neumann already in 1932 [15]. Consider a set of system states $|n\rangle$ which our apparatus is built to discriminate.



Original form of the von Neumann measurement model. Information about the state of the measured system S is transferred to the measuring apparatus A.

For each state $|n\rangle$ we have a corresponding pointer state $|\Phi_n\rangle$ (more precisely, for each "quantum number" n there exists a large set of macrostates $|\Phi_n^{(\alpha)}\rangle$, α describing microscopic degrees of freedom). If the measurement is repeatable or ideal the dynamics of the measurement interaction must look like

$$|n\rangle|\Phi_0\rangle \quad \stackrel{t}{\longrightarrow} \quad |n\rangle|\Phi_n(t)\rangle \ . \tag{2}$$

From linearity we can immediately see what happens for a general initial state of the measured system,

$$\left(\sum_n c_n|n\rangle\right)|\Phi_0\rangle \quad \stackrel{t}{\longrightarrow} \quad \sum_n c_n|n\rangle|\Phi_n(t)\rangle \ . \tag{3}$$

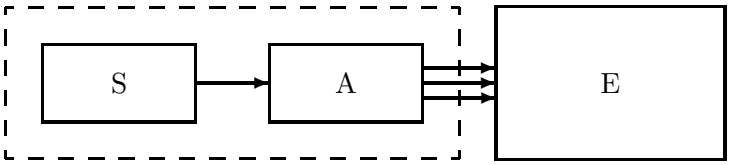
We do not find a certain measurement result, but a superposition. Through unitary evolution, a correlated (and still pure) state results, which contains all possible

results as components. Of course such a superposition must not be interpreted as an ensemble. The transition from this superposition to a single component – which is what we observe – constitutes the quantum measurement problem. As long as there is no collapse we have to deal with the whole superposition – and it is well known that a superposition has very different properties compared to any of its components. Quantum correlations are often misinterpreted as (quantum) noise. This is wrong, however: Noise would mean that the considered system is in a certain state, which may be unknown and/or evolve in a complicated way. Such an interpretation is untenable and contradicts all experiments which show the nonlocal features of quantum-correlated (entangled) states.

Von Neumann's treatment, as described so far, is unrealistic since it does not take into account the essential openness of macroscopic objects. This deficiency can easily be remedied by extending the above scheme.

1.2 Classical Properties through Decoherence

If one takes into account that the apparatus A is coupled to its environment E, which also acts like a measurement device, the phase relations are (extremely fast) further dislocalized into the total system – finally the entire universe, according to



Realistic extension of the von Neumann measurement model. Information about the state of the measured system S is transferred to the measuring apparatus A and then very rapidly sent to the environment E. The back-reaction on the (local) system S+A originates entirely from quantum nonlocality.

$$\left(\sum_n c_n|n\rangle|\Phi_n\rangle\right)|E_0\rangle \quad \stackrel{t}{\longrightarrow} \quad \sum_n c_n|n\rangle|\Phi_n\rangle|E_n\rangle. \tag{4}$$

The behavior of system+apparatus is then described by the density matrix

$$\rho_{SA} \approx \sum_n |c_n|^2 |n\rangle\langle n| \otimes |\Phi_n\rangle\langle\Phi_n| \qquad \text{if} \qquad \langle E_n|E_m\rangle \approx \delta_{nm} \tag{5}$$

which is identical to that of an ensemble of measurement results $|n\rangle|\Phi_n\rangle$.

Of course, this does not resolve the measurement problem! This density matrix describes only an "improper" ensemble, i.e., with respect to all possible observations at S+A it *appears* that a certain measurement result has been achieved. Again, classical notions like noise or recoil are not appropriate: A acts dynamically on E, but the back-action arises entirely from quantum nonlocality (as long as the measurement is "ideal", that is, (4) is a good approximation). Nevertheless, the system S+A acquires classical behavior, since interference terms are absent with respect to local observations if the above process is irreversible [19, 10].

Needless to say, the interference terms still exist globally in the total (pure) state, although they are unobservable at either system alone – a situation which may be characterized by the statement

The interference terms still exist, but they are not there.[10]

2 Do we need observables?

In most treatments of quantum mechanics the notion of an observable plays a central role. Do observables represent a fundamental concept or can they be derived? If we describe a measurement as a certain kind of interaction, then observables should not be required as an essential ingredient of quantum theory. In a sense this was also done by von Neumann, but not used later very much because of restrictions enforced by the Copenhagen school (e.g., the demand to describe a measurement device in classical terms instead of seeking for a consistent treatment in terms of wave functions).

Two elements are necessary to derive an observable that discriminates certain (orthogonal) system states $|n\rangle$. First, one needs an appropriate interaction which is diagonal in the eigenstates of the measured "observable" and is able to "move the pointer", so that we have as above

$$|n\rangle|\Phi_0\rangle \quad \stackrel{H_{int}}{\longrightarrow} \quad |n\rangle|\Phi_n\rangle \ . \tag{6}$$

This can be achieved by Hamiltonians of the form

$$H_{int} = \sum_n |n\rangle\langle n| \otimes \hat{A}_n \tag{7}$$

with appropriate $\hat{A}_n$ leading to orthogonal pointer states (Note that (6) defines only the eigenbasis of an observable; the eigenvalues represent merely scale factors and are therefore of minor importance). The second condition that must be fulfilled is dynamical stability of pointer states against decoherence, that is, the pointer states must only be passively recognized by the environment according to,

$$|\Phi_n\rangle|E_0\rangle \quad \stackrel{decoherence}{\longrightarrow} \quad |\Phi_n\rangle|E_n\rangle \ . \tag{8}$$

Both conditions must be fulfilled. For example, a measurement device which acts according to (6) would be totally useless, if it were not stable against decoherence: Consider a Schrödinger cat state as pointer state! The *same* basis states $|\Phi_n\rangle$ must be distinguished as dynamically relevant in (6) as well as in (8).*

*This explains *dynamically* why certain observables may "not exist" operationally. For a general discussion of the relation between quantum states and observables see Sect. 2.2 of [5]. Arguments along these lines lead to the conclusion that one should not attribute a fundamental status to the Heisenberg picture – contrary to widespread belief – despite its *phenomenological* equivalence with the Schrödinger picture.

3 Do we need superselection rules?

What is a superselection rule? One way to define a superselection rule is to say, that certain states $|\Psi_1\rangle$, $|\Psi_2\rangle$ are found in nature, but never general superpositions $|\Psi\rangle = \alpha|\Psi_1\rangle + \beta|\Psi_2\rangle$. This means that all observations can be described by a density matrix of the form $\rho = p_1|\Psi_1\rangle\langle\Psi_1| + p_2|\Psi_2\rangle\langle\Psi_2|$. Clearly such a density matrix is exactly what is obtained through decoherence in appropriate situations.

3.1 Approximate superselection rules

There are many examples, where it is hard to find certain superpositions in the real world. The most famous example has been given by Schrödinger: A superposition of a dead and an alive cat

$$|\Psi\rangle = |\text{dead cat}\rangle + |\text{alive cat}\rangle \tag{9}$$

is never observed, contrary to what should be possible according to the superposition principle (and, in fact, *must* necessarily occur according to the Schrödinger equation). Another drastic situation is given by a state like

$$|\Psi\rangle = |\text{cat}\rangle + |\text{dog}\rangle \ . \tag{10}$$

Such a superposition looks truly absurd, but only because we never observe states of this kind! (The obvious objection that one cannot superpose states of "different systems" seems to be inappropriate. For example, nobody hesitates to superpose states with different numbers of particles.) A more down-to-earth example is given by the position of large objects, which are never found in states

$$|\Psi\rangle = |\text{here}\rangle + |\text{there}\rangle \ , \tag{11}$$

with "here" and "there" macroscopically distinct. Under realistic circumstances such objects are always well described by a localized density matrix $\rho(x, x') \approx p(x)\delta(x - x')$. A special case of this localization occurs in molecules (except the very small ones), which show a well-defined spatial structure. The Born-Oppenheimer approximation is not sufficient to explain this fact.

Quite generally we have an approximate superselection rule whenever we describe the dynamics of a dynamical variable by some rate equation (that is, without interference) instead of the Schrödinger equation.

3.2 Exact superselection rules

Strict absence of interference can only be expected for discrete quantities. One important example is electric charge. Can this be understood via decoherence? We know from Maxwell's theory, that every charge carries with itself an associated electric field, so that a superposition of charges may be written in the form [16]

$$\begin{aligned} \sum_q c_q|\Psi_q^{total}\rangle &= \sum_q c_q|\chi_q^{bare}\rangle|\Psi_q^{field}\rangle \\ &= \sum_q c_q|\chi_q^{local}\rangle|\Psi_q^{farfield}\rangle \ . \end{aligned} \tag{12}$$

Since we can only observe the local dressed charge, it has to be described by the density matrix

$$\rho = \sum_q |c_q|^2 |\chi_q^{local}\rangle\langle\chi_q^{local}| \tag{13}$$

If the far fields are orthogonal (distinguishable), coherence would be absent locally. So the question arises: Is the Coulomb field only part of the kinematics (implemented via the Gauss constraint) or does it represent a quantum dynamical degree of freedom so that we have to consider decoherence via a retarded Coulomb field? For an attempt to understand part of the Coulomb field as dynamical see [4].

What do experiments tell us? A superposition of the form (11) can be observed for charged particles (cf. the contribution by Hasselbach[6]). On the other hand, the classical (retarded) Coulomb field would contain information about the path of the charged particle, destroying coherence. The situation does not appear very clear-cut. Hence one essential question remains:

What is the *quantum* physical role of the Coulomb field?

A similar situation arises in quantum gravity, where we can expect that superpositions of different masses (energies) are decohered by the spatial curvature.

Another important "exact" superselection rule forbids superposing states with integer and half-integer spin, for example

$$|\Psi\rangle = |\text{ spin } 1\rangle + |\text{ spin } {}^1\!/_2\rangle \ , \tag{14}$$

which would transform under a rotation by 2π into

$$|\Psi_{2\pi}\rangle = |\text{ spin } 1\rangle - |\text{ spin } {}^1\!/_2\rangle \ , \tag{15}$$

clearly a different state because of the different relative phase. If one *demands* that such a rotation should not change anything, such a state must be excluded. This is one standard argument in favor of the "univalence" superselection rule. On the other hand, one *has* observed the sign-change of spin ${}^1\!/_2$ particles under a (relative) rotation by 2π in *certain* experiments. Hence we are left with two options: Either we view the group SO(3) as the proper rotation group also in quantum theory. Then nothing must change if we rotate the system by an angle of 2π. Hence we can derive this superselection rule from symmetry. But this may merely be a classical prejudice. The other choice is to use SU(2) instead of SO(3) as rotation group. Then we are in need of explaining why those strange superpositions never occur. This last choice amounts to keeping the superposition principle as the fundamental principle of quantum theory. In more technical terms we should then avoid using groups with non-unique ("ray" [¶]) representations, such as SO(3). In supersymmetric theories, bosons and fermions are treated on an equal footing, so it would be natural to superpose their states (what is apparently never done in particle theory).

[¶] The widely used argument that physical states are to be represented by rays, not vectors, in Hilbert space because the phase of a state vector cannot be observed, is misleading. Since relative phases are certainly relevant, one should prefer a vector as a *fundamental* physical state concept, rather than a ray. Rays cannot even be superposed without (implicitly) using vectors.

In a similar manner one could undermine the well-known argument leading from the Galilean symmetry of nonrelativistic quantum mechanics to the mass superselection rule. In this case we could maintain the superposition principle and replace the Galilei group by a larger group. How this can be done is shown by Domenico Giulini[4].

The final open question for this section then is:

Can *all* superselection rules be understood as decoherence effects?

4 Examples

4.1 Localization

The by now standard example of decoherence is the localization of macroscopic objects. Why do macroscopic objects always appear localized in space? Coherence between macroscopically different positions is destroyed *very* rapidly because of the strong influence of scattering processes. The formal description may proceed as follows. Let $|x\rangle$ be the position eigenstate of a macroscopic object, and $|\chi\rangle$ the state of the incoming particle. Following the von Neumann scheme (2), the scattering of such particles off an object located at position x may be written as

$$|x\rangle|\chi\rangle \stackrel{t}{\longrightarrow} |x\rangle|\chi_x\rangle = |x\rangle S_x|\chi\rangle \ , \tag{16}$$

where the scattered state may conveniently be calculated by means of an appropriate S-matrix. For the more general initial state of a wave packet we have then

$$\int d^3x \ \varphi(x)|x\rangle|\chi\rangle \stackrel{t}{\longrightarrow} \int d^3x \ \varphi(x)|x\rangle S_x|\chi\rangle \ . \tag{17}$$

Therefore, the reduced density matrix describing our object changes into

$$\rho(x,x') = \varphi(x)\varphi^*(x') \left\langle \chi|S_{x'}^\dagger S_x|\chi\right\rangle \ . \tag{18}$$

Of course, a single scattering process will usually not resolve a small distance, so in most cases the matrix element on the right-hand side of (18) will be close to unity. If we add the contributions of many scattering processes, an exponential damping of spatial coherence results:

$$\rho(x,x',t) = \rho(x,x',0) \exp\left\{-\Lambda t(x-x')^2\right\} \ . \tag{19}$$

The strength of this effect is described by a single parameter Λ that may be called "localization rate". It is given by

$$\Lambda = \frac{k^2 N v \sigma_{eff}}{V} \ . \tag{20}$$

Here, k is the wave number of the incoming particles, Nv/V the flux, and σ_{eff} is of the order of the total cross section (for details see [10] or Sect. 3.2.1 and Appendix 1 of [5]). Some values of Λ are given in the table.

Localization rate Λ in $\mathrm{cm}^{-2}\mathrm{s}^{-1}$ for three sizes of "dust particles" and various types of scattering processes (from [10]). This quantity measures how fast interference between different positions disappears as a function of distance in the course of time.

	$a = 10^{-3}$cm dust particle	$a = 10^{-5}$cm dust particle	$a = 10^{-6}$cm large molecule
Cosmic background radiation	10^{6}	10^{-6}	10^{-12}
300 K photons	10^{19}	10^{12}	10^{6}
Sunlight (on earth)	10^{21}	10^{17}	10^{13}
Air molecules	10^{36}	10^{32}	10^{30}
Laboratory vacuum (10^3 particles/cm^3)	10^{23}	10^{19}	10^{17}

Most of the numbers in the table are quite large, showing the extremely strong coupling of macroscopic objects, such as dust particles, to their natural environment. Even in intergalactic space, the 3K background radiation cannot simply be neglected.
Hence the main lesson is:

Macroscopic objects are not even approximately isolated.

A consistent unitary description must therefore include the environment and finally the whole universe.*

If we combine this damping of coherence with the "free" Schrödinger dynamics we arrive at an equation of motion for the density matrix that to a good approximation simply adds these two contributions,

$$i\frac{\partial\rho}{\partial t} = [H_{internal}, \rho] + i\frac{\partial\rho}{\partial t}\Big|_{scatt.} \quad . \tag{21}$$

In the position representation this equation reads in one space dimension

$$i\frac{\partial\rho(x,x',t)}{\partial t} = \frac{1}{2m}\left(\frac{\partial^2}{\partial x'^2} - \frac{\partial^2}{\partial x^2}\right)\rho - i\Lambda(x-x')^2\rho \ . \tag{22}$$

Solutions of this equation can easily be found (see, e.g.[5])

*One of the first stressing the importance of the dynamical coupling of macro-objects to their environment was Dieter Zeh, who wrote in his 1970 Found. Phys. paper [17]: "Since the interactions between macroscopic systems are effective even at astronomical distances, the only 'closed system' is the universe as a whole. ... It is of course very questionable to describe the universe by a wavefunction that obeys a Schrödinger equation. Otherwise, however, there is no inconsistency in measurement, as there is no theory."

This is now more or less commonplace, but this was not the case some 30 years ago, when he sent an earlier version of this paper to the journal Il Nuovo Cimento. I quote from the referee's reply: "The paper is completely senseless. It is clear that the author has not fully understood the problem and the previous contributions in this field." (H.D. Zeh, private communication)

So far this treatment represents *pure* decoherence, following directly the von Neumann scheme. If recoil is added as a next step, we arrive at models including friction, that is, quantum Brownian motion. There are several models for the quantum analogue of Brownian motion, some of which are even older than the first decoherence studies. Early treatments did not, however, draw a distinction between decoherence and friction (decoherence alone does *not* imply friction.). As an example, consider the equation of motion derived by Caldeira and Leggett [2],

$$i\frac{\partial\rho}{\partial t} = [H,\rho] + \frac{\gamma}{2}[x,\{p,\rho\}] - im\gamma k_B T[x,[x,\rho]] \tag{23}$$

which reads for a "free" particle

$$\begin{aligned} i\frac{\partial\rho(x,x',t)}{\partial t} &= \left[\frac{1}{2m}\left(\frac{\partial^2}{\partial x'^2} - \frac{\partial^2}{\partial x^2}\right) - i\Lambda(x-x')^2\right. \\ &\quad \left. + i\gamma(x-x')\left(\frac{\partial}{\partial x'} - \frac{\partial}{\partial x}\right)\right]\rho(x,x',t)\ , \end{aligned} \tag{24}$$

where γ is the damping constant, and here $\Lambda = m\gamma k_B T$.

If one compares the effectiveness of the two terms representing decoherence and relaxation, one finds that their ratio is given by

$$\frac{\text{decoherence rate}}{\text{relaxation rate}} = mk_B T(\delta x)^2 \propto \left(\frac{\delta x}{\lambda_{th}}\right)^2\ , \tag{25}$$

where λ_{th} denotes the thermal de Broglie wavelength of the considered object. This ratio has for a typical macroscopic situation ($m = 1$g, $T = 300$K, $\delta x = 1$cm) the enormous value of about 10^{40}! This shows that in these cases decoherence is *far more important* than dissipation.

Not only the center-of-mass position of dust particles becomes "classical" via decoherence. The spatial structure of molecules represents another most important example. Consider a simple model of a chiral molecule.

Right- and left-handed versions both have a rather well-defined spatial structure, whereas the ground state is – for symmetry reasons – a superposition of both chiral states. These chiral configurations are usually separated by a tunneling barrier, which is so high that under normal circumstances tunneling is very improbable, as was already shown by Hund in 1929. But this alone does not explain why chiral (and, indeed, most) molecules are never found in energy eigenstates!

In a simplified model with low-lying nearly-degenerate eigenstates $|1\rangle$ and $|2\rangle$, the right- and left-handed configurations may be given by

$$\begin{aligned} |L\rangle &= \frac{1}{\sqrt{2}}(|1\rangle + |2\rangle) \\ |R\rangle &= \frac{1}{\sqrt{2}}(|1\rangle - |2\rangle)\ . \end{aligned} \tag{26}$$

Because the environment recognizes the spatial structure via scattering processes, only chiral states are stable against decoherence,

$$|R,L\rangle|\Phi_0\rangle \stackrel{t}{\longrightarrow} |R,L\rangle|\Phi_{R,L}\rangle \ . \tag{27}$$

The dynamical instability of energy (i.e., parity) eigenstates of molecules represents a typical example of "spontaneous symmetry breaking" induced by decoherence. Additionally, transitions between spatially oriented states are suppressed by the quantum Zeno effect, described below.

4.2 Quantum Zeno Effect

The most dramatic consequence of a strong measurement-like interaction of a system with its environment is the quantum Zeno effect. It has been discovered several times and is also sometimes called "watchdog effect" or "watched pot behavior", although most people now use the term Zeno effect. It is surprising only if one sticks to a classical picture where observing a system and just verifying its state should have no influence on it. Such a prejudice is certainly formed by our everyday experience, where observing things in our surroundings does not change their properties. As is known since the early times of quantum theory, observation can drastically change the observed system.

The essence of the Zeno effect can easily be shown as follows. Consider the "decay" of a system which is initially prepared in the "undecayed" state $|u\rangle$. The probability to find the system undecayed, i.e., in the same state $|u\rangle$ at time t is for small time intervals given by

$$\begin{aligned} P(t) &= |\langle u|\exp(-iHt)|u\rangle|^2 \\ &= 1-(\Delta H)^2t^2+\mathcal{O}(t^4) \end{aligned} \tag{28}$$

with

$$(\Delta H)^2 = \langle u|H^2|u\rangle - \langle u|H|u\rangle^2 \ . \tag{29}$$

If we consider the case of N measurements in the interval $[0,t]$, the non-decay probability is given by

$$P_N(t) \approx \left[1-(\Delta H)^2\left(\frac{t}{N}\right)^2\right]^N > 1-(\Delta H)^2t^2 = P(t) \ . \tag{30}$$

This is always larger than the single-measurement probability given by (28). In the limit of arbitrary dense measurements, the system no longer decays,

$$P_N(t) = 1-(\Delta H)^2\frac{t^2}{N}+\ldots \stackrel{N\to\infty}{\longrightarrow} 1 \ . \tag{31}$$

Hence we find that repeated measurements can completely hinder the natural evolution of a quantum system. Such a result is clearly quite distinct from what is observed for classical systems. Indeed, the paradigmatic example for a classical stochastic process, exponential decay,

$$P(t) = \exp(-\Gamma t) \ , \tag{32}$$

is not influenced by repeated observations, since for N measurements we simply have

$$P_N(t) = \left(\exp\left(-\Gamma\frac{t}{N}\right)\right)^N = \exp(-\Gamma t) \ . \tag{33}$$

So far we have treated the measurement process in our discussion of the Zeno effect in the usual way by assuming a collapse of the system state onto the subspace corresponding to the measurement result. Such a treatment can be extended by employing a von Neumann model for the measurement process, e.g., by coupling a pointer to a two-state system. A simple toy model is given by the Hamiltonian

$$\begin{aligned} H &= H_0 + H_{int} \\ &= V(|1\rangle\langle 2| + |2\rangle\langle 1|) + E|2\rangle\langle 2| + \gamma\hat{p}(|1\rangle\langle 1| - |2\rangle\langle 2|) \ , \end{aligned} \tag{34}$$

where transitions between states $|1\rangle$ and $|2\rangle$ (induced by the "perturbation" V) are monitored by a pointer (coupling constant γ). This model already shows all the typical features mentioned above.

The transition probability starts for small times always quadratically, according to the general result (28). For times, where the pointer resolves the two states, a behavior similar to that found for Markow processes appears: The quadratic time-dependence changes to a linear one. For strong coupling the transitions are suppressed. This clearly shows the dynamical origin of the Zeno effect.

An extension of the above model allows an analysis of the transition from the Zeno effect to master behavior (described by transition *rates* as was first studied in quantum mechanics by Pauli in 1928). It can be shown that for many (micro-)states which are not sufficiently resolved by the environment, Fermi's Golden Rule can be recovered, with transition rates which are no longer reduced by the Zeno effect. Nevertheless, interference between macrostates is suppressed very rapidly [7].

4.3 Decoherence of Fields

In QED we find two (related) situations,

- "Measurement" of charges by fields;
- "Measurement" of fields by charges.

In both cases, the entanglement between charge and field states leads to decoherence as already described above in the discussion of superselection rules, see also [5] and references therein.

In recent quantum optics experiments it is possible to prepare and study superpositions of different classical field states, quantum-mechanically represented by coherent states, for example Schrödinger cat states of the form

$$|\Psi\rangle = N(|\alpha\rangle + |-\alpha\rangle) \tag{35}$$

which can be realized as field states in a cavity. In these experiments (see [1]) decoherence can be turned on gradually by coupling the cavity to a reservoir. Typical decoherence times are in the range of about 100 μs.

For *true* cats the decoherence time is much shorter (in particular, it is *very much* shorter than the lifetime of a cat!). This leads to the appearance of *quantum jumps*, although all underlying processes are smooth in principle since they are governed by the Schrödinger equation.

In experimental situations of this kind we find a gradual transition from a superposition of different decay times (seen in "collapse and revival" experiments) to a local mixture of decay times (leading to "quantum jumps") according to the following scheme.

theory	experiment
superposition of different decay times	collapse and revivals
$\Downarrow$	$\Downarrow$
local mixture of different decay times	quantum jumps

4.4 Spacetime and Quantum Gravity

In quantum theories of the gravitational field, no classical spacetime exists at the most fundamental level. Since it is generally assumed that the gravitational field has to be quantized, the question again arises how the corresponding classical properties can be understood.

Genuine quantum effects of gravity are expected to occur for scales of the order of the Planck length $\sqrt{G\hbar/c^3}$. It is therefore often argued that the spacetime structure at larger scales is automatically classical. However, this Planck scale argument is as insufficient as the large mass argument in the evolution of free wave packets. As long as the superposition principle is valid (and even superstring theory leaves this untouched), superpositions of different metrics should occur at any scale.

The central problem can already be demonstrated in a simple Newtonian model[8]. Consider a cube of length L containing a homogeneous gravitational field with a quantum state ψ such that at some initial time $t = 0$

$$|\psi\rangle = c_1|g\rangle + c_2|g'\rangle \,, \tag{36}$$

where g and g' correspond to two different field strengths. A particle with mass m in a state $|\chi\rangle$, which moves through this volume, "measures" the value of g, since its trajectory depends on the acceleration g:

$$|\psi\rangle|\chi^{(0)}\rangle \to c_1|g\rangle|\chi_g(t)\rangle + c_2|g'\rangle|\chi_{g'}(t)\rangle \,. \tag{37}$$

This correlation destroys the coherence between g and g', and the reduced density matrix can be estimated to assume the following form after many such interactions are taken into account:

$$\rho(g, g', t) = \rho(g, g', 0) \exp\left(-\Gamma t (g - g')^2\right), \tag{38}$$

where

$$\Gamma = nL^4 \left(\frac{\pi m}{2k_B T}\right)^{3/2}$$

for a gas with particle density n and temperature T. For example, air under ordinary conditions, $L = 1$ cm, and $t = 1$ s yields a remaining coherence width of $\Delta g/g \approx 10^{-6}$[8].

Thus, matter does not only tell space to curve but also to behave classically. This is also true in full quantum gravity.

In a fully quantized theory of gravity, for example in the canonical approach described by the Wheeler-deWitt equation,

$$H|\Psi(\Phi, {}^{(3)}\mathcal{G})\rangle = 0 \ , \tag{39}$$

where Φ describes matter and ${}^{(3)}\mathcal{G}$ is the three-metric, everything is contained in the "wave function of the universe" Ψ. Here we encounter new problems: There is neither an external time parameter, nor is there an external observer. How these problems can be tackled is described in Claus Kiefer's contribution[12].

5 Lessons

What insights can be drawn from decoherence studies? It should be emphasized that decoherence derives from a straightforward application of standard quantum theory to realistic situations. It seems to be a historical accident, that the importance of the interaction with the natural environment was overlooked for such a long time. Certainly the still prevailing (partly philosophical) attitudes enforced by the Copenhagen school played a (negative) role here, for example by outlawing a physical analysis of the measurement process in quantum-mechanical terms.

Because of the strong coupling of macroscopic objects, a quantum description of macroscopic objects *requires* the inclusion of the natural environment. A fully unitary quantum theory is only consistent if applied to the whole universe. This does not preclude local phenomenological descriptions. However, their derivation from a universal quantum theory and the interpretation assigned to such descriptions have to be analyzed very carefully.

We have seen that typical classical properties, such as localization in space, are *created* by the environment in an irreversible process, and are therefore not inherent attributes of macroscopic objects. The features of the interaction define *what* is classical by selecting a certain basis in Hilbert space. Hence superselection sectors emerge from the dynamics. In all "classical" situations, the relevant decoherence time is extremely short, so that the smooth Schrödinger dynamics leads to apparent discontinuities like "events", "particles" or "quantum jumps".

There are certain ironies in this situation. *Local* classical properties find their explanation in the *nonlocal* features of quantum states. Usually quantum objects are considered as fragile and easy to disturb, whereas macroscopic objects are viewed as the rock-solid building blocks of empirical reality. However, the opposite is true: macroscopic objects are extremely sensitive and immediately decohered.

On the practical side, decoherence also has its disadvantages. It makes testing alternative theories difficult (more on that below), and it represents a major obstacle for people trying to construct a quantum computer. Building a really big one may well turn out to be as difficult as detecting other Everett worlds!

5.1 Does decoherence solve the measurement problem?

Clearly not. What decoherence tells us, is that certain objects *appear* classical when they are observed. But what is an observation? At some stage, we still have to apply the usual probability rules of quantum theory. These are hidden in density matrices, for example.

5.2 Which interpretations make sense?

One could also ask: what interpretations are left from the many that have been proposed during the decades since the invention of quantum theory? I think, we do not have much of a choice at present*, *if* we restrict ourselves to use only wavefunctions as kinematical concepts (that is, we ignore hidden-variable theories, for example).

There seem to be only the two possibilities either (1) to alter the Schrödinger equation to get something like a "real collapse" [3, 13], or (2) to keep the theory unchanged and try to establish some variant of the Everett interpretation. Both approaches have their pros and cons, some of them are listed in the following table.

Clearly collapse models face the immediate question of how, when and where a collapse takes place. If a collapse occurs before the information enters the consciousness of an observer, one can maintain some kind of psycho-physical parallelism by assuming that what is experienced subjectively is parallel to the physical state of certain objects, e.g., parts of the brain. The last resort is to view consciousness as *causing* collapse, an interpretation which can more or less be traced back to von Neumann. In any case, the collapse happens with a certain probability (and with respect to a certain basis in Hilbert space) and this element of the theory comprises an *additional* axiom.

How would we want to test such theories? One would look for collapse-like deviations from the unitary Schrödinger dynamics. However, similar *apparent* deviations are also produced by decoherence, in particular in the relevant meso- and macroscopic range. So it is hard to discriminate these *true* changes to the Schrödinger equation from the *apparent* deviations brought about by decoherence[9].

Everett interpretations lead into rather similar problems. Instead of specifying the collapse one has to define precisely how the wavefunction is to be split up into branches. Decoherence can help here by selecting certain directions in Hilbert space as dynamically stable (and others as extremely fragile – branches with macroscopic objects in nonclassical states immediately decohere), but the location of the observer in the holistic quantum world is always a decisive ingredient. It must be assumed that what is subjectively experienced is parallel to certain states (observer states) in a certain *component* of the global wave function. The probabilities (frequencies)

*The following owes much to discussions with Dieter Zeh, who finally convinced me that the Everett interpretation *could* perhaps make sense at all.

collapse models	Everett
traditional psycho-physical parallelism: What is perceived is parallel to the observer's physical *state*	new form of psycho-physical parallelism: Subjective perception is parallel to the observer state in a *component* of the universal wave function
probabilities put in by hand	probabilities must also be postulated (existing "derivations" are circular)
problems with relativity	peaceful coexistence with relativity
experimental check: look for collapse-like deviations from the Schrödinger equation ⇓ hard to test because of decoherence	experimental check: look for macroscopic superpositions ⇓ hard to test because of decoherence

we observe in repeated measurements form also an additional axiom [§]. The peaceful coexistence with relativity seems not to pose much problems, since no collapse ever happens and all interactions are local in (high-dimensional) configuration space. But testing Everett means testing the Schrödinger equation in particular with respect to macroscopic superpositions, and this again is precluded by decoherence.

So it seems that both alternatives still have conceptual problems and both are hard to test because of decoherence. We should not be surprised, however, if it finally turned out that we do not know enough about consciousness and its relation to the physical world to solve the quantum mystery [14].

References

[1] Brune, M., Hagley, E., Dreyer, J., Maître, X., Maali, A., Wunderlich, C., Raimond, J.M., Haroche, S. (1996): Observing the Progressive Decoherence of the "Meter" in a Quantum Measurement. Phys. Rev. Lett. **77**, 4887–4890.

[2] Caldeira, A.O., Leggett, A.J. (1983): Path integral approach to quantum Brownian motion. Physica **121A**, 587–616.

[§]There exist several claims in the literature that probabilities can be derived in the Everett interpretation. I think these proofs are circular. Consider a sequence of N measurements on copies of a two-state system, all prepared in the initial state $a|1\rangle + b|2\rangle$. The resulting correlated state contains 2^N components, where each pointer state shows one of the 2^N possible sequences of measurement results (e.g. as a computer printout). But these pointer states are *always the same*, independently of the values of a and b! Only if each branch is given a *weight* involving $|a|^2$ and $|b|^2$ one may recover the correct frequencies. See also [11]. In addition, deep (and partially very old) questions about the meaning of probabilities seem to reappear in the framework of Everett interpretations.

[3] Ghirardi, G.C., Rimini, A. and Weber, T. (1986): Unified dynamics for microscopic and macroscopic systems. Phys. Rev. **D34**, 470–491.

[4] Giulini, D.: States, Symmetries and Superselection, contribution to this volume.

[5] Giulini, D., Joos, E., Kiefer, C., Kupsch, J., Stamatescu, I.-O., Zeh, H.D. (1996): *Decoherence and the Appearance of a Classical World in Quantum Theory* (Springer, Berlin).

[6] Hasselbach. F., Kiesel, H., and Sonnentag, P.: Exploration of the Fundamentals of Quantum Mechanics by Charged Particle Interferometry, contribution to this volume.

[7] Joos, E. (1984): Continuous measurement: Watchdog effect versus golden rule. Phys. Rev. **D29**, 1626–1633.

[8] Joos, E. (1986): Why do we observe a classical spacetime? Phys. Lett. **A116**, 6–8.

[9] Joos, E. (1987): Comment on 'Unified dynamics for microscopic and macroscopic systems', Phys. Rev. **D36**, 3285–3286.

[10] Joos, E., Zeh, H.D. (1985): The emergence of classical properties through interaction with the environment. Z. Phys. **B59**, 223–243.

[11] Kent, A. (1990): Against Many-World Interpretations. Int. J. Mod. Phys. **A5**, 1745–1762. Also available as eprint gr-qc/9703089.

[12] Kiefer, C.: Decoherence in Situations Involving the Gravitational Field, contribution to this volume.

[13] Pearle, P. (1999): Collapse Models. eprint quant-ph/9901077.

[14] Squires, E. (1990): *Conscious Mind in the Physical World* (IOP Publishing, Bristol, Philadelphia).

[15] von Neumann, J. (1932): *Mathematische Grundlagen der Quantentheorie* (Springer, Berlin).

[16] Zeh, H.D.: The Meaning of Decoherence, contribution to this volume.

[17] Zeh, H.D. (1970): On the interpretation of measurement in quantum theory. Found. Phys. **1**, 69–76.

[18] Zeh, H.D. (1999): *The Physical Basis of the Direction of Time.* 3^{rd} edn. (Springer).

[19] Zurek, W.H. (1981): Pointer basis of quantum apparatus: Into what mixture does the wave packet collapse? Phys. Rev. **D24**, 1516–1525.

William G. Unruh

Statement

and

Readings

Decoherence and Entanglement

William G. Unruh

Decoherence is the obverse side of entanglement, the peculiarly quantum nature of correlations between systems. By Bell's theorem, we know that entanglement has a number of non-intuitive properties, implying that quantum correlations can in some cases be stronger than classical, and in some cases violate transitivity ($A \Rightarrow B$, $B \Rightarrow C$, $C \Rightarrow D$ but *not* $A \Rightarrow D$, where $\Rightarrow$ is implication). If we disregard these correlations, looking at only one of the systems on its own, the statistical properties of that system suffer decoherence. Interference terms which would in general be present for a quantum system with a variety of possible values for some attribute, are not present.

This decoherence has be argued to solve a variety of problems including the measurement problem in quantum mechanics.

However it is also true that the presence or absense of decoherence is far more subtle than usually described. A system, quantum correlated with an "environment" (another quantum system), can for certain measurements appear to be highly decohered, while still exhibiting interference between the apparently decohered values with suitable, long time, experiments.

False loss of coherence

W. G. Unruh
CIAR Cosmolgy Program, Dept. of Physics
University of B.C.
Vancouver, Canada V6T 1Z1

email:unruh@physics.ubc.ca

Abstract

The loss of coherence of a quantum system coupled to a heat bath as expressed by the reduced density matrix is shown to lead to the miss-characterization of some systems as being incoherent when they are not. The spin boson problem and the harmonic oscillator with massive scalar field heat baths are given as examples of reduced incoherent density matrices which nevertheless still represent perfectly coherent systems.

I. MASSIVE FIELD HEAT BATH AND A TWO LEVEL SYSTEM

How does an environment affect the quantum nature of a system? The standard technique is to look at the reduced density matrix, in which one has traced out the environment variables. If this changes from a pure state to a mixed state (entropy $Tr\rho\ln\rho$ not equal to zero) one argues that the system has lost quantum coherence, and quantum interference effects are suppressed . However this criterion is too strong. There are couplings to the environment which are such that this reduced density matrix has a high entropy, while the system alone retains virtually all of its original quantum coherence certain experiments.

The key idea is that the external environment can be different for different states of the system. There is a strong correlation between the system and the environment. As usual, such correlations lead to decoherence in the reduced density matrix. However, the environment in these cases is actually tied to the system, and is adiabatically dragged along by the system. Thus although the state of the environment is different for the two states, one can manipulate the system alone so as to cause these apparently incoherent states to interfere with each other. One simply causes a sufficiently slow change in the system so as to drag the environment variables into common states so the quantum interference of the system can again manifest itself.

An example is if one looks at an electron with its attached electromagnetic field. Consider the electron at two different positions. The static coulomb field of the two charges differ, and thus the states of the electromagnetic field differ with the electron in the two positions. These differences can be sufficient to cause the reduced electron wave function loose coherence for a state which is a coherent sum of states located at these two positions. However, if one causes the system to evolve so as to cause the electron in those two positions to come together (

eg, by having a force field such that the electron in both positions to be brought together at some central point for example), those two apparently incoherent states will interfere, demonstrating that the loss of coherence was not real.

Another example is light propagating through a slab of glass. If one simply looks at the electromagnetic field, and traces out over the states of the atoms in the glass, the light beams travelling through two separate regions of the glass will clearly decohere– the reduced density matrix for the electromagnetic field will lose coherence in postition space– but those two beams of light will also clearly interfer when they exit the glass or even when they are within the glass.

The above is not to be taken as proof, but as a motivation for the further investigation of the problem. The primary example I will take will be of a spin $\frac{1}{2}$ particle (or other two level system). I will also examine a harmonic oscillator as the system of interest. In both cases, the heat bath will be a massive one dimensional scalar field. This heat bath is of the general Caldera Leggett type [1](and in fact is entirely equivalent to that model in general). The mass of the scalar field will be taken to be larger than the inverse time scale of the dynamical behaviour of the system. This is not to be taken as an attempt to model some real heat bath, but to display the phenomenon in its clearest form. Realistic heat baths will in general also have low frequency excitations which will introduce other phenomena like damping and genuine loss of coherence into the problem.

II. SPIN-$\frac{1}{2}$ SYSTEM

Let us take as our first example that of a spin-$\frac{1}{2}$ system coupled to an external environment. We will take this external environment to be a one dimensional massive scalar field. The coupling to the spin system will be via purely the 3 component of the spin. I will use the velocity coupling which I have used elsewhere as a simple example of an environment (which for a massless field is completely equivalent to the Caldera Leggett model). The Lagrangian is

$$L = \int \frac{1}{2}((\dot{\phi}(x))^2 - (\phi(x)')^2 + m^2\phi(x)^2 + 2\epsilon\dot{\phi}(x)h(x)\sigma_3)dx \tag{1}$$

which gives the Hamiltonian

$$H = \int \frac{1}{2}((\pi(x) - \epsilon h(x)\sigma_3)^2 + (\phi(x)')^2 + m^2\phi(x)^2)dx \tag{2}$$

$h(x)$ is the interaction range function, and its Fourier transform is related to the spectral response function of Leggett and Caldera.

This system is easily solvable. I will look at this system in the following way. Start initially with the field in its free ($\epsilon = 0$) vacuum state, and the system is in the $+1$ eigenstate of σ_1. I will start with the coupling ϵ initially zero and gradually increase it to some large value. I will look at the reduced density matrix for the system, and show that it reduces one which is almost the identity matrix (the maximally incoherent density matrix) for strong coupling. Now I let ϵ slowly drop to zero again. At the end of the procedure, the state of the system will again be found to be in the original eigenstate of σ_1. The intermediate maximally incoherent density matrix would seem to imply that the system no longer has any quantum

coherence. However this lack of coherence is illusionary. Slowly decoupling the system from the environment should in the usual course simply maintain the incoherence of the system Yet here, as if by magic, an almost completely incoherent density matrix magically becomes coherent when the system is decoupled from the environment.

In analyzing the system, I will look at the states of the field corresponding to the two possible σ_3 eigenstates of the system. These two states of the field are almost orthogonal for strong coupling. However they correspond to fields tightly bound to the spin system. As the coupling is reduced, the two states of the field adiabatically come closer and closer together until finally they coincide when ϵ is again zero. The two states of the environment are now the same, there is no correlation between the environment and the system, and the system regains its coherence.

The density matrix for the spin system can always be written as

$$\rho(t) = \frac{1}{2}(1 + \vec{\rho}(t) \cdot \vec{\sigma}) \tag{3}$$

where

$$\vec{\rho}(t) = Tr(\vec{\sigma}\rho(t)) \tag{4}$$

We have

$$\vec{\rho}(t) = Tr\left(\vec{\sigma}\mathcal{T}[e^{-i\int_0^t H dt}]\frac{1}{2}(1 + \vec{\rho}(0) \cdot \vec{\sigma})R_0\mathcal{T}[e^{-i\int H dt}]^\dagger\right) \tag{5}$$

where R_0 is the initial density matrix for the field (assumed to be the vacuum), and $\mathcal{T}[]$ is the time ordering operator. (Because ϵ and thus H is time dependent, the H at different times do not commute. this leads to requirement for the time ordering in the expression. As usual, the time ordered integral is the way of writing the time ordered product $\prod_n e^{-iH(t_n)dt} = e^{-iH(t)dt}e^{-iH(t-dt)dt}....e^{-iH(0)dt}$.)

Let us first calculate $\rho_3(t)$. We have

$$\rho_3(t) = Tr\left(\sigma_3\mathcal{T}[e^{-i\int_0^t H dt}]\tfrac{1}{2}(1 + \vec{\rho}(0) \cdot \vec{\sigma})R_0\mathcal{T}[e^{-i\int H dt}]^\dagger\right) \tag{6}$$

$$= Tr\left(\mathcal{T}[e^{-i\int_0^t H dt}]\sigma_3\tfrac{1}{2}(1 + \vec{\rho}(0) \cdot \vec{\sigma})R_0\mathcal{T}[e^{-i\int H dt}]^\dagger\right) \tag{7}$$

$$= Tr\left(\sigma_3\tfrac{1}{2}(1 + \vec{\rho}(0) \cdot \vec{\sigma})R_0\right) \tag{8}$$

$$= \rho_3(0) \tag{9}$$

because σ_3 commutes with $H(t)$ for all t. We now define

$$\sigma_+ = \frac{1}{2}(\sigma_1 + i\sigma_2) = |+><-|; \qquad \sigma_- = \sigma_+^\dagger \tag{10}$$

Using $\sigma_+\sigma_3 = -\sigma_+$ and $\sigma_3\sigma_+ = \sigma_+$ we have

$$\begin{aligned} Tr & \left(\sigma_+\mathcal{T}[e^{-i\int_0^t H dt}]\frac{1}{2}(1 + \vec{\rho}(0) \cdot \vec{\sigma})R_0\mathcal{T}[e^{-i\int H dt}]^\dagger\right) \\ = & \; Tr_\phi\left(\mathcal{T}[e^{-i\int (H_0 - \epsilon(t)\int \pi(x)h(x)dx)dt}]^\dagger\right. \\ & \left.\mathcal{T}[e^{-i\int (H_0 + \epsilon(t)\int \pi(x)h(x)dx)dt}]\right) < -|\frac{1}{2}(1 + \vec{\rho}(0) \cdot \vec{\sigma})|+> \end{aligned} \tag{11}$$

$$= (\rho_1(0) + i\rho_2(0))J(t)$$

where H_0 is the Hamiltonian with $\epsilon = 0$, i.e., the free Hamiltonian for the massless scalar field and

$$J(t) = Tr_\phi \left(\mathcal{T}[e^{-i\int (H_0 - \epsilon(t)\int \pi(x)h(x)dx)dt}]^\dagger \mathcal{T}[e^{-i\int (H_0+\epsilon(t)\int \pi(x)h(x)dx)dt}]R_0 \right) \tag{12}$$

Breaking up the time ordered product in the standard way into a large number of small time steps, using the fact that $e^{-i\epsilon(t)\int h(x)\phi(x)dx}$ is the displacement operator for the field momentum through a distance of $\epsilon(t)h(x)$, and commuting the free field Hamiltonian terms through, this can be written as

$$\begin{aligned} J(t) = Tr_\phi \Bigg(& e^{-i\epsilon(0)\Phi(0)} \prod_{n=1}^{t/dt} \left[e^{-i(\epsilon(t_n)-\epsilon(t_{n-1})\Phi(t_n)} \right] \\ & e^{i\epsilon(t)\Phi(t)} e^{i\epsilon(t)\Phi(t)} \prod_{n=t/dt}^{1} \left[e^{i\epsilon(t_n - \epsilon(t_{n-1}))\Phi(t_n)} \right] e^{i\epsilon(0)\Phi(0)} R_0 \Bigg) \end{aligned} \tag{13}$$

where $t_n = ndt$ and dt is a very small value, $\Phi(t) = \int h(x)\phi(t,x)dx$ and $\phi_0(t,x)$ is the free field Heisenberg field operator. Using the Campbell-Baker-Hausdorff formula, realizing that the commutators of the Φs are c-numbers, and noticing that these c-numbers cancel between the two products, we finally get

$$J(t) = Tr_\phi \left(e^{2i(\epsilon(t)\Phi(t) - \epsilon(0)\Phi(0) + \int_0^t \dot{\epsilon}(t')\Phi(t')dt')} R_0 \right) \tag{14}$$

from which we get

$$ln(J(t)) = -2Tr_\phi \left(R_0 \left(\epsilon(t)\Phi(t) - \epsilon(0)\Phi(0) + \int_0^t \dot{\epsilon}(t')\Phi(t')dt' \right)^2 \right) \tag{15}$$

I will assume that $\epsilon(0) = 0$, and that $\dot{\epsilon}(t)$ is very small, and that it can be neglected. (The neglected terms are of the form

$$\int\int \dot{\epsilon}^2 < \Phi(t')\Phi(t") > dt'dt" \approx \dot{\epsilon}^2 t\tau < \Phi(0)^2 >$$

which for a massive scalar field has τ, the coherence time scale, $\approx 1/m$. Thus, as we let $\dot{\epsilon}$ go to zero these terms go to zero.)

We finally have

$$\begin{aligned} ln(J(t)) &= -2\epsilon(t)^2 < \Phi(t)^2 > \\ &= -2\epsilon(t)^2 \int |\hat{h}(k)|^2 \frac{1}{\sqrt{(k^2+m^2)}} dk \end{aligned} \tag{16}$$

Choosing $\hat{h}(k) = e^{-\Gamma|k|/2}$, we finally get

$$ln(J(t)) = -4\int_0^\infty \epsilon(t)^2 \frac{e^{-\Gamma|k|}dk}{sqrt(k^2+m^2)} \tag{17}$$

This goes roughly as $\ln(\Gamma m)$ for small Γm, (which I will assume is true). For Γ sufficiently small, this makes J very small, and the density matrix reduces to essentially diagonal form ($\rho_z(t) \approx \rho_y(t) \approx 0$, $\rho_z(t) = \rho_z(0)$.)

However it is clear that if $\epsilon(t)$ is now lowered slowly to zero, the decoherence factor J goes back to unity, since it depends only on $\epsilon(t)$. The density matrix now has exactly its initial form again. The loss of coherence at the intermediate times was illusionary. By decoupling the system from the environment after the coherence had been lost, the coherence is restore. this is in contrast with the naive expectation in which the loss of coherence comes about because of the correlations between the system and the environment. Decoupling the system from the environment should not in itself destroy that correlation, and should not reestablish the coherence.

The above approach, while giving the correct results, is not very transparent in explaining what is happening. Let us therefor take a different approach. Let us solve the Heisenberg equations of motion for the field $\phi(t,x)$. The equations are (after eliminating π)

$$\partial_t^2 \phi(t,x) - \partial_x^2 \phi(t,x) + m^2 \phi(t,x) = -\dot{\epsilon}(t)\sigma_3 h(x) \tag{18}$$

$$\pi(t,x) = \dot{\phi}(t,x) + \epsilon(t)h(x)\sigma_3 \tag{19}$$

If ϵ is slowly varying in time, we can solve this approximately by

$$\phi(t,x) = \phi_0(t,x) + \dot{\epsilon}(t)\int \frac{1}{2m} e^{-m|x-x'|} h(x')dx'\sigma_3 + \psi(t,x)\epsilon(0)\sigma_3 \tag{20}$$

$$\pi(t,x) = \dot{\phi}_0(t,x) + \epsilon(t)h(x)\sigma_3 + \dot{\psi}(t,x)\epsilon(0)\sigma_3 \tag{21}$$

where $\phi_0(t,x)$ and $\pi_0(t,x)$ are free field solution to the equations of motion in absence of the coupling, with the same initial conditions

$$\dot{\phi}_0(0,x) = \pi(0,x) \tag{22}$$

$$\phi_0(0,x) = \phi(0,x) \tag{23}$$

, while ψ is also a solution of the free field equations but with initial conditions

$$\psi(0,x) = 0 \tag{24}$$

$$\dot{\psi}(0,x) = -h(x). \tag{25}$$

If we examine this for the two possible eigenstates of σ_3, we find the two solutions

$$\phi_\pm(t,x) \approx \phi_0(t,x) \pm (\dot{\epsilon}(t)\int \frac{1}{2m} e^{-m|x-x'|} h(x')dx' + \psi(t,x)) \tag{26}$$

$$\pi_\pm(t,x) \approx \dot{\phi}_0(t,x) + O(\dot{\epsilon}) \pm (\epsilon(t)h(x) + \epsilon(0)\dot{\psi}(t,x)) \tag{27}$$

These solutions neglect terms of higher derivatives in ϵ. The state of the field is the vacuum state of ϕ_0, π_0. $\phi_\pm$ and $\pi_\pm$ are equal to this initial field plus c number fields. Thus in terms of the $\phi_\pm$ and $\pi_\pm$, the state is a coherent state with non-trivial displacement from the vacuum. Writing the fields in terms of their creation and annihilation operators,

$$\phi_\pm(t,x) = \int A_{k\pm}(t)e^{ikx} + A^\dagger_{k\pm}e^{-ikx}\frac{dk}{\sqrt{2\pi\omega_k}} \tag{28}$$

$$\pi_\pm(t,x) = i\int A_{k\pm}(t)e^{ikx} - A^\dagger_{k\pm}e^{-ikx}\sqrt{\frac{k^2+m^2}{2\pi}}dk \tag{29}$$

we find that we can write $A_{k\pm}$ in terms of the initial operators A_{k0} as

$$A_{k\pm}(t) \approx A_{k0}e^{-i\omega_k t} \pm \frac{1}{2}i(\epsilon(t) - \epsilon(0)e^{-i\omega_k t})(h(k)/\sqrt{\omega_k} + O(\dot{\epsilon}(t))) \tag{30}$$

where $\omega_k = \sqrt{k^2 + m^2}$. Again I will neglect the terms of order $\dot{\epsilon}$ in comparison with the ϵ terms. Since the state is the vacuum state with respect to the initial operators A_{k0}, it will be a coherent state with respect to the operators $A_{k\pm}$, the annihilation operators for the field at time t. We thus have two possible coherent states for the field, depending on whether the spin is in the upper or lower eigenstate of $/sigma_3$. But these two coherent states will have a small overlap. If $A|\alpha >= \alpha|\alpha >$ then we have

$$|\alpha >= e^{\alpha A^\dagger - |\alpha|^2/2}|0 > \tag{31}$$

Furthermore, if we have two coherent states $|\alpha >$ and $|\alpha' >$, then the overlap is given by

$$< \alpha|\alpha' >=< 0|e^{\alpha^* A - |\alpha|^2/2}e^{\beta A^\dagger - |\beta|^2/2}|0 >= e^{\alpha^*\beta - (|\alpha|^2 + |\beta|^2)/2} \tag{32}$$

In our case, taking the two states $|\pm_\phi >$, these correspond to coherent states with

$$\alpha = -\alpha' = \frac{1}{2}i(\epsilon(t) - \epsilon(0)e^{-i\omega_k t}) = \frac{1}{2}i\epsilon(t)h(k)/\sqrt{\omega_k} \tag{33}$$

Thus we have

$$< +_\phi, t|-_\phi, t >= \prod_k e^{-\epsilon(t)^2|h(k)|^2/(k^2+m^2)} = e^{-\epsilon(t)^2 \int \frac{|h(k)|^2}{\omega_k} dk} = J(t). \tag{34}$$

Let us assume that we began with the state of the spin as $\frac{1}{\sqrt{2}}(|+ > +|- >)$. The state of the system at time t in the Schroedinger representation is $\frac{1}{\sqrt{2}}(|+ > |+_\phi(t) > +|- > |-_\phi >)$ and the reduced density matrix is

$$\rho = \frac{1}{2}(|+ >< +| + |- >< -| + J^*(t)|+ >< -| + J(t)|- >< +|). \tag{35}$$

The off diagonal terms of the density matrix are suppressed by the function $J(t)$. $J(t)$ however depends only on $\epsilon(t)$ and thus , as long as we keep $\dot{\epsilon}$ small, the loss of coherence represented by J can be reversed simply by decoupling the system from the environment slowly.

The apparent decoherence comes about precisely because the system in either the two eigenstates of σ_3 drives the field into two different coherent states. For large ϵ, these two states have small overlap. However, this distortion of the state of the field is tied to the system. π changes only locally, and the changes in the field caused by the system do not radiate away. As ϵ slowly changes, this bound state of the field also slowly changes in concert . However if one examines only the system, one sees a loss of coherence because the field states have only a small overlap with each other.

The behaviour is very different if the system or the interaction changes rapidly. In that case the decoherence can become real. As an example, consider the above case in which $\epsilon(t)$ suddenly is reduced to zero. In that case, the field is left as a free field, but a free field whose state (the coherent state) depends on the state of the system. In this case the field radiates away as real (not bound) excitations of the scalar field. The correlations with the system are carried away, and even if the coupling were again turned on, the loss of coherence would be permanent.

III. OSCILLATOR

For the harmonic oscillator coupled to a heat bath, the Hamiltonian can be taken as

$$H = \frac{1}{2}\int (\pi(x) - \epsilon(t)q(t)\tilde{h}(x))^2 + (\partial_x\phi(x))^2 + m^2\phi(t,x)^2 dx + \frac{1}{2}(p^2 + \Omega^2 q^2) \tag{36}$$

Let us assume that m is much larger than Ω or that the inverse time rate of change of ϵ. The solution for the field is given by

$$\phi(t,x) \approx \phi_0(t,x) + \psi(t,x)\epsilon(0)q(0) - \dot{\overline{\epsilon(t)q(t)}}\int \frac{e^{-m|x-x'|}}{2m}h(x')dx' \tag{37}$$

$$\pi(t,x) \approx \dot{\phi}(t,x) + \dot{\psi}(t,x)\epsilon(0)q(0) - \ddot{\overline{\epsilon(t)q(t)}}\int \frac{e^{-m|x-x'|}}{2m}h(x')dx' + \epsilon(t)q(t)h(x) \tag{38}$$

where again ϕ_0 is the free field operator, ψ is a free field solution with $\psi(0) = 0$, $\dot{\psi}(0) = -h(x)$. Retaining terms only of the lowest order in ϵ

$$\psi(t,x) \approx \phi_0(t,x) \tag{39}$$

$$\pi(t,x) \approx \dot{\phi}(t,x) + \epsilon(t)q(t)h(x) \tag{40}$$

The equation of motion for q is

$$\dot{q}(t) = p(t) \tag{41}$$

$$\dot{p}(t) = -\Omega^2 q + \epsilon(t)\dot{\Phi}(t) \tag{42}$$

where $\Phi(t) = \int h(x)\phi(t,x)dx$. Substitution in the expression for ϕ, we get

$$\ddot{q}(t) + \Omega^2 q(t) \approx \epsilon\dot{(}\Phi_0(t)) + \epsilon(t)\ddot{\overline{\epsilon(t)q(t)}}\int\int h(x)h(x')\frac{e^{-m|x-x'|}}{2m}dxdx' \tag{43}$$

Neglecting the derivatives of ϵ (i.e., assuming that ϵ changes slowly even on the time scale of $1/\Omega$), this becomes

$$\left(1 + \epsilon(t)^2\int\int h(x)h(x')\frac{e^{-m|x-x'|}}{2m}dxdx'\right)\ddot{q} + \Omega^2 q = \partial_t(\epsilon(t)\Phi(t)) \tag{44}$$

The interaction with the field thus renormalizes the mass of the oscillator to

$$M = \left(1 + \epsilon(t)^2\int\int h(x)h(x')\right)$$

The solution for q is thus

$$q(t) \approx q(0)\cos(\int_0^t \tilde{\Omega}(t)dt) + \frac{1}{\tilde{\Omega}}\sin(\int_0^t \tilde{\Omega}(t)dt)p(0) + \frac{1}{\tilde{\Omega}}\int_0^t \sin(\int_{t'}^t \tilde{\Omega}(t)dt)\partial_t(\epsilon(t')\dot{\overline{\epsilon(t)\Phi_0(t')}})dt' \tag{45}$$

where $\tilde{\Omega}(t) \approx \Omega/\sqrt{M(t)}$.

The important point is that the forcing term dependent on Φ_0 is a rapidly oscillating term of frequency at least m. Thus if we look for example at $< q^2 >$, the deviation from the free evolution of the oscillator (with the renormalized mass) is of the order of $\int \sin(\tilde{\Omega}t - t')\sin(\omega(t - t")) < \dot{\Phi}_0(t')\dot{\Phi}_0(t") > dt'dt"$. But $< \dot{\Phi}_0(t')\dot{\Phi}_0(t") >$ is a rapidly oscillating function of frequency at least m, while the rest of the integrand is a slowly varying function with frequency much less than m, Thus this integral will be very small (at least $\tilde{\Omega}/m$ but typically much smaller than this depending on the time dependence of ϵ). Thus the deviation of $q(t)$ from the free motion will in general be very very small, and I will neglect it.

Let us now look at the field. The field is put into a coherent state which depends on the value of q, because $\pi(t,x) \approx \dot{\phi}_0(t,x) + \epsilon(t)q(t)h(x)$ Thus

$$A_k(t) \approx a_{0k}e^{-i\omega_k t} + i\frac{1}{2}\hat{h}(k)\epsilon(t)q(t)/\omega_k \tag{46}$$

The overlap integral for these coherent states with various values of q is

$$\prod_k < i\frac{1}{2}\hat{h}(k)\epsilon(t)q/\omega_k | i\frac{1}{2}\hat{h}(k)\epsilon(t)q'/\omega_k >= e^{-\frac{1}{8}\int |\hat{h}(k)|^2 dk (q-q')^2} \tag{47}$$

The density matrix for the Harmonic oscillator is thus

$$\rho(q,q') = \rho_0(t,q,q')e^{-\frac{1}{8}\int |\hat{h}(k)|^2 dk (q-q')^2} \tag{48}$$

where ρ_0 is the density matrix for a free harmonic oscillator (with the renormalized mass).

Ie, we see a strong loss of coherence of the off diagonal terms of the density matrix. However this loss of coherence is false. If we take the initial state for example with two packets widely separated in space, these two packets will loose their coherence. However, as time proceeds, the natural evolution of the Harmonic oscillator will bring those two packets together ($q - q'$ small across the wave packet). For the free evolution they would then interfere. They still do. The loss of coherence which was apparent when the two packets were widely separated disappears, and the two packets interfere just as if there were no coupling to the environment. The effect of the particular environment used is thus to renormalise the mass, and to make the density matrix appear to loose coherence.

IV. SPIN BOSON PROBLEM

Let us now complicate the spin problem in the first section by introducing into the system a free Hamiltonian for the spin as well as the coupling to the environment. Following the example of the spin boson problem, let me introduce a free Hamiltonian for the spin of the form $\frac{1}{2}\Omega\sigma_1$, whose effect is to rotate the σ_3 states (or to rotate the vector $\vec{\rho}$ in the $2-3$ plane with frequency Ω.

The Hamiltonian now is

$$H = \frac{1}{2}\left(\int (\pi(t,x) - \epsilon(t)h(x)\sigma_3)^2 + (\partial_x\phi(x))^2 + m^2\phi(t,x)^2 dx + \Omega\sigma_1\right) \tag{49}$$

where again $\epsilon(t)$ is a slowly varying function of time. We will solve this in the manner of the second part the first section.

If we let Ω be zero, then the eigenstates of σ_z are eigenstates of the Hamiltonian. The field Hamiltonian (for constant ϵ) is given by

$$H_\pm = \frac{1}{2}\int (\pi - (\pm\epsilon(t)h(x)))^2 + (\partial_x\phi)^2 dx. \tag{50}$$

Defining $\tilde{\pi} = \pi - (\pm h(x))$, $\tilde{\pi}$ has the same commutation relations with π and ϕ as does π. Thus in terms of $\tilde{\pi}$ we just have the Hamiltonian for the free scalar field. The instantaneous minimum energy state is therefor the ground state energy for the free scalar field for both $H_\pm$. Thus the two states are degenerate in energy. In terms of the operators π and ϕ, these ground states are coherent states with respect to the vacuum state of the original uncoupled ($\epsilon = 0$) free field, with the displacement of each mode given by

$$a_k|\pm> = \pm i\epsilon(t)\frac{h(k)}{\sqrt{\omega_k}}|\pm> \tag{51}$$

or

$$|\pm> = \prod_k |\pm\alpha_k> |\pm>_{\sigma_3} \tag{52}$$

where the $|\alpha_k>$ are coherent states for the k^{th} modes with coherence parameter $\alpha_k = i\epsilon(t)\frac{h(k)}{\sqrt{\omega_k}}$, and the states $|\pm>_{\sigma_3}$ are the two eigenstates of σ_3. (In the following I will eliminate the $\prod_k$ symbol.) The energy to the next excited state in each case is just m, the mass of the free field.

We now introduce the $\Omega\sigma_x$ as a perturbation parameter. The two lowest states (and in fact the excited states) are two fold degenerate. Using degenerate perturbation theory to find the new lowest energy eigenstates, we must calculate the overlap integral of the perturbation between the original degenerate states and must then diagonalise the resultant matrix to lowest order in Ω. The perturbation is $\frac{1}{2}\Omega\sigma_1$. All terms between the same states are zero, because of the $<\pm|_{\sigma_3}\sigma_1|\pm>_{\sigma_3} = 0$. Thus the only terms that survive for determining the lowest order correction to the lowest energy eigenvalues are

$$\frac{1}{2}<+|\Omega\sigma_1|->\ = \frac{1}{2}<-|\Omega\sigma_1|+>^* \tag{53}$$

$$= \frac{1}{2}\Omega\prod_k <\alpha_k|-\alpha_k> = \frac{1}{2}\Omega\prod_k e^{-2|\alpha_k|^2} \tag{54}$$

$$= \frac{1}{2}\Omega e^{-2\int \epsilon(t)^2|h(k)|^2/\omega_k dk} = \frac{1}{2}\Omega J(t) \tag{55}$$

The eigenstates of energy thus have energy of $E(t)_\pm = E_0 \pm \frac{1}{2}\Omega J(t)$, and the eigenstates are $\sqrt{\frac{1}{2}}(|+>\pm|->)$ If epsilon varies slowly enough, the instantaneous energy eigenstates will be the actual adiabatic eigenstates at all times, and the time evolution of the system will just be in terms of these instantaneous energy eigenstates. Thus the system will evolve as

$$|\psi(t)> = \sqrt{\tfrac{1}{2}}e^{-iE_0t}\Big((c_+ + c_-)e^{-i\int \frac{1}{2}\Omega_t J(t)dt}(|+>+|->) \tag{56}$$

$$+ (c_- - c_+)e^{+i\int \frac{1}{2}\Omega_t J(t)dt}(|+>-|->)\Big) \tag{57}$$

where the c_+ and c_- are the initial amplitudes for the $|+>_{\sigma_3}$ and $|->_{\sigma_3}$ states. The reduced density matrix for the spin system in the σ_3 basis can now be written as

$$\vec{\rho}(t) = (J(t)\rho_{01}(t), J(t)\rho_{02}(t), \rho_{03}(t)) \tag{58}$$

where $\vec{\rho_0}(t)$ is the density matrix that one would obtain for a free spin half particle moving under the Hamiltonian $J(t)\Omega\sigma_1$.

$$\begin{aligned} \rho_{01}(t) &= \rho_1(0) \\ \rho_{02}(t) &= \rho_2(0)\cos(\Omega\int J(t')dt') + \rho_3(0)\sin(\Omega\int J(t')dt') \\ \rho_{03}(t) &= \rho_3(0)\cos(\Omega\int J(t')dt') - \rho_2(0)\sin(\Omega\int J(t')dt \end{aligned} \tag{59}$$

Thus if $J(t)$ is very small (ii.e., ϵ large) , we have a renormalized frequency for the spin system, and the the off diagonal terms (in the σ_3 representation) of the density matrix are strongly suppressed by a factor of $J(t)$. Thus if we begin in an eigenstate of σ_3 the density matrix will begin with the vector $\vec{\rho}$ as a unit vector pointing in the 3 direction. As time goes on the 3 component gradually decreases to zero, but the 2 component increases only to the small value of $J(t)$. The system looks almost like a completely incoherent state, with almost the maximal entropy that the spin system could have. However as we wait longer, the 3 component of the density vector reappears and grows back to its full unit value in the opposite direction, and the entropy drop to zero again. This cycle repeats itself endlessly with the entropy oscillating between its minimum and maximum value forever.

The decoherence of the density matrix (the small off diagonal terms) obviously represent a false loss of coherence. It represents a strong correlation between the system and the environment. However the environment is bound to the system, and essentially forms a part of the system itself, at least as long as the system moves slowly. However the reduced density matrix makes no distinction between whether or not the correlations between the system and the environment are in some sense bound to the system, or are correlations between the system and a freely propagating modes of the medium in which case the correlations can be extremely difficult to recover, and certainly cannot be recovered purely by manipulations of the system alone.

V. INSTANTANEOUS CHANGE

In the above I have assumed throughout that the system moves slowly with respect to the excitations of the heat bath. Let us now look at what happens in the spin system if we rapidly change the spin of the system. In particular I will assume that the system is as in section 1, a spin coupled only to the massive heat bath via the component σ_3 of the spin. Then at a time t_0, I instantly rotate the spin through some angle θ about the 1 axis. In this case we will find that the environment cannot adjust rapidly enough, and at least a part of the loss of coherence becomes real, becomes unrecoverable purely through manipulations of the spin alone.

The Hamiltonian is

$$H = \frac{1}{2}\int \left((\pi(t,x) - \epsilon(t)h(x)\sigma_3)^2 + (\partial_x\phi(t,x)^2 + m^2\phi(t,x)\right)dx + \theta/2\delta(t-t_0)\sigma_1 \tag{60}$$

Until the time t_0 σ_3 is a constant of the motion, and similarly afterward. Before the time t_0, the energy eigenstates state of the system are as in the last section given by

$$|\pm, t >= \{|+>_{\sigma_3} |\alpha_k(t)> \text{or}\{|->_{\sigma_3} |-\alpha_k(t)>\} \tag{61}$$

An arbitrary state for the spin–environment system is given by

$$|\psi >= c_+|+> +c_-|-> \tag{62}$$

Now, at time t_0, the rotation carries this to

$$\begin{aligned} |\phi(t_0)>= & \quad c_+(\cos(\theta/2)|+>_{\sigma_3} + i\sin(\theta/2)|->_{\sigma_3} |\alpha_k(t)> \\ & + c_-(\cos(\theta/2)|->_{\sigma_3} + i\sin(\theta/2)|+>_{\sigma_3})|-\alpha_k(t)> \\ = & \quad \cos(\theta/2)\,(c_+|+> +c_-|->) \\ & + i\sin(\theta/2)(c_+|->_{\sigma_3} |\alpha_k(t)> - c_-|+>_{\sigma_3} |-\alpha_k(t)> \end{aligned} \tag{63}$$

The first term is still a simple sum of eigenvectors of the Hamiltonian after the interaction. The second term, however, is not. We thus need to follow the evolution of the two states $|->_{\sigma_3} |\alpha_k(t_0)>$ and $|+>_{\sigma_3} |-\alpha_k(t_0)>$. Since σ_3 is a constant of the motion after the interaction again, the evolution takes place completely in the field sector. Let us look at the first state first. (The evolution of the second can be derived easily from that for the first because of the symmetry of the problem.)

I will again work in the Heisenberg representation. The field obeys

$$\dot{\phi}_-(t,x) = \pi_-(t,x) + \epsilon(t)h(x) \tag{64}$$

$$\dot{\pi}_-(t,x) = \partial_x^2\phi_-(t,x) - m^2\phi_-(t,x) \tag{65}$$

with solution At the time t_0 the field is in the coherent state $|\alpha_k>$. This can be represented by taking the field operator to be of the form

$$\phi_-(t_0,x) = \phi_0(t_0,x) \tag{66}$$

$$\pi_-(t_0,x) = \dot{\phi}_0(t_0,x) + \epsilon(t_0)h(x) \tag{67}$$

whee the state $|\alpha_k>$ is the vacuum state for the free field ϕ_0.. We can now solve the equations of motion for ϕ_- and obtain (again assuming that $\epsilon(t)$ is slowly varying)

$$\phi_-(t,x) = \phi_0(t,x) + 2\psi(t,x)\epsilon(t_0) \tag{68}$$

$$\pi_-(t,x) = \dot{\phi}_0(t,x) + 2\psi(t,x)\epsilon(t_0) - \epsilon(t)h(x) \tag{69}$$

where $\psi(t_0,x) = 0$ and $\dot{\psi}(t_0,x) = h(x)$. Thus again, the field is in a coherent state set by both $2\epsilon(t_0)\psi$ and $\epsilon(t)h(x)$. The field ψ propagates away from the interaction region determined by $h(x)$, and I will assume that I am interested in times t a long time after the time t_0. At these times I will assume that $\int h(x)\psi(t,x)dx = 0$. (This overlap dies out as $1/\sqrt{mt}$. The calculations can be carried out for times nearer t_0 as well— the expressions are just messier and not particularly informative.)

Let me define the new coherent state as $|-\alpha_k(t) + \beta_k(t)>$, where α_k is as before and

$$\beta_k(t) = 2\epsilon(t_0)\omega_k\tilde{\psi}(t,k) = 2i\epsilon(t_0)e^{i\omega_k t}\tilde{h}(k)/\omega_k \tag{70}$$

(The assumption regarding the overlap of $h(x)$ and $\psi(t)$ corresponds to the assumption that $\int \alpha_k^*(t)\beta_k(t)dk = 0$). Thus the state $|- >_{\sigma_3} |\alpha_k >$ evolves to the state $|- >_{\sigma_3} |-\alpha_k + \beta_k(t) >$. Similarly, the state $|+ >_{\sigma_3} |-\alpha_k >$ evolves to $|+ >_{\sigma_3} |\alpha_k - \beta_k(t) >$.

We now calculate the overlaps of the various states of interest.

$$< \alpha_k|\alpha_k \pm \beta_k >=< -\alpha_k|-\alpha_k \pm \beta_k >= e^{-\int |\beta_k|^2 dk} = J(t_0) \tag{71}$$

$$< -\alpha_k|\alpha_k \pm \beta_k >=< \alpha_k|-\alpha_k \pm \beta_k >= J(t)J(t_0) \tag{72}$$

$$< -\alpha_k + \beta_k|\alpha_k - \beta_k >=< -\alpha_k - \beta_k|\alpha_k + \beta_k >= J(t)J(t_0)^4 \tag{73}$$

The density matrix becomes

$$\rho_3 = \cos(\theta)\rho_{03} + \sin(\theta)J(t_0)\rho_{02} \tag{74}$$

$$\rho_1 = J(t)\left(\cos(\theta) + J^4(t_0)\sin(\theta)\right)\rho_{01} \tag{75}$$

$$\rho_2(t) = J(t)\left(-\sin(\theta)\rho_{03} + (\cos(\theta/2) - J^4(t_0)\sin(\theta))\rho_{02}\right) \tag{76}$$

where

$$\rho_{03} = \frac{1}{2}(|c_+|^2 - |c_-|^2) \tag{77}$$

$$\rho_{01} = Re(c_+c_-^*) \tag{78}$$

$$\rho_{02} = Im(c_+c_-^*) \tag{79}$$

If we now let $\epsilon(t)$ go slowly to zero again (to find the 'real' loss of coherence), we find that unless $\rho_{01} = \rho_{02} = 0$ the system has really lost coherence during the sudden transition. The maximum real loss of coherence occurs if the rotation is a spin flip ($\theta = \pi$) and ρ_{03} was zero. In that case the density vector dropped to $J(t_0)^4$ of its original value. If the density matrix was in an eigenstate of σ_3 on the other hand, the density matrix remained a coherent density matrix, but the environment was still excited by the spin.

We can use the models of a fast or a slow spin flip interaction to discuss the problem of the tunneling time. As Leggett et al argue [3], the spin system is a good model for the consideration of the behaviour of a particle in two wells, with a tunneling barrier between the two wells. One view of the transition from one well to the other is that the particle sits in one well for a long time. Then at some random time it suddenly jumps through the barrier to the other side. An alternative view would be to see the particle as if it were a fluid, with a narrow pipe connecting it to the other well- the fluid slowly sloshing between the two wells. The former is supported by the fact that if one periodically observes which of the two wells the particle is in, one sees it staying in one well for a long time, and then between two observations, suddenly finding it in the other well. This would, if one regarded it as a classical particle imply that the whole tunneling must have occurred between the two observations. It is as if the system were in an eigenstate and at some random time an interaction flipped the particle from one well to the other. However, this is not a good picture. The environment is continually observing the system. It it really moved rapidly from one to the other, the environment would see the rapid change, and would radiate. Instead, left on its own, the environment in this problem (with a mass much greater than the frequency of transition of the system) simply adjust continually to the changes in the system. The tunneling thus seems to take place continually and slowly.

VI. DISCUSSION

The high frequency modes of the environment lead to a loss of coherence (decay of the off-diagonal terms in the density matrix) of the system, but as long as the changes in the system are slow enough this decoherence is false– it does not prevent the quantum interference of the system. The reason is that the changes in the environment caused by these modes are essentially tied to the system, they are adiabatic changes to the environment which can easily be adiabatically reversed. Loosely one can say that coherence is lost by the transfer of information (coherence) from the system to the environment. However in order for this information to be truly lost, it must be carried away by the environment, separated from the system by some mechanism or another so that it cannot come back into the system. In the environment above, this occurs when the information travels off to infinity. Thus the loss of coherence as represented by the reduced density matrix is in some sense the maximum loss of coherence of the system. Rapid changes to the system, or rapid decoupling of the system from the environment, will make this a true decoherence. However, gradual changes in the system or in the coupling to the external world can cause the environment to adiabatically track the system and restore the coherence apparently lost.

This is of special importance to understanding the effects of the environmental cutoff in many environments [3]. For "ohmic" or "superohmic" environments (where h does not fall off for large arguments), one has to introduce a cutoff into the calculation for the reduced density matrix. This cutoff has always been a bit mysterious, especially as the loss of coherence depends sensitively on the value of this cutoff. If one imagines the environment to include say the electromagnetic field, what is the right value for this cutoff? Choosing the Plank scale seems silly, but what is proper value? The arguments of this paper suggest that in fact the cutoff is unnecessary except in renormalising the dynamics of the system. The behaviour of the environment at frequencies much higher than the inverse time scale of the system leads to a false loss of coherence, a loss of coherence which does not affect the actual coherence (ability to interfere with itself) of the system. Thus the true coherence is independent of cutoff.

As far as the system itself is concerned, one should regard it as "dressed" with a polarization of the high frequency components of the environment. One should regard not the system itself as important for the quantum coherence, but a combination of variables of the system plus the environment.What is difficult is the dependence of which the degrees of freedom of the environment are simply dressing and which are degrees of freedom which can lead to loss of coherence depends crucially on the motion and the interactions of the system itself. They are history dependent, not simply state dependent. This make it very difficult to simply find some transformation which will express the system plus environment in terms of variables which are genuinely independent, in the sense that if the new variable loose coherence, then that loss is real.

These observations emphasis the importance of not making too rapid conclusions from the decoherence of the system. This is especially true in cosmology, where high frequency modes of the cosmological system are used to decohere low frequency quantum modes of the universe. Those high frequency modes are likely to behave adiabatically with respect to the low frequency behaviour of the universe. Thus although they will lead to a reduced density matrix for the low frequency modes which is apparently incoherent, that incoherence

is likely to be a false loss of coherence.

ACKNOWLEDGMENTS

I would like to thank the Canadian Institute for Advanced Research for their support of this research. This research was carried out under an NSERC grant 580441.

REFERENCES

[1] A. O. Caldeira, A. J. Leggett Physica **121A** 587(1983), Phys Rev **A31** 1057 (1985) . See also the paper by W. Unruh, W. Zurek, Phys Rev **D40**1071(1989) where a field model for coherence instead of the oscillator model for calculating the density matrix of an oscillator coupled to a heat bath.

[2] Many of the points made here have also been made by A. Leggett. See for example A. J. Leggett in *Applications of Statistical and Field Theory Methods to Condensed Matter*(Proc. 1989 Nato Summer School, Evora, Portugal), ed D. Baeriswyl, A.R. Bishop, and J. Carmelo. Plenum Press (1990) and **Macroscopic Realism: What is it, and What do we know about it from Experiment** in *Quantum Measurement: Beyond Paradox* ed R. A. Healey, and G. Hellman, U. Minnesota Press (Minneapolis, 1998)

[3] See for example the detailed analysis of the density matrix of a spin 1/2 system in an oscillator heat bath, where the so called superohmic coupling to the heat bath leads to a rapid loss of coherence due to frequencies in the bath much higher than the frequency of the system under study. A.J. Leggett et al Rev. Mod.Phys **59** 1 (1987)

[4]

Thoughts on Non-Locality and Quantum Mechanics

W. Unruh

Program in Cosmology and Gravity of CIAR

Dept Physics and Astronomy

University of B.C.

Vancouver, Canada V6T 1Z1

Abstract

The debate about the non-locality of quantum mechanics is old, but still lively. Numerous people use non-locality as a (bad) shorthand for quantum entanglement. But some have a long standing commitment to the validity of this characterisation. This paper examines two separate streams in this debate. The first is the arguments of Stapp, and especially his recent paper where he simplifies his contractually argument in the Hardy situation to argue for the non-locality of quantum mechanics. He has maintained his contention that an analysis of a Hardy type correlation between two spatially separated observers proves that quantum mechanics itself is non-local, with out any additional assumption of realism or hidden variables.

In the second section I try to carefully examine the Bell argument, in the CHSH variant to see where the difference between the quantum and classical situations differ.

Asher Peres was one of the great physicists on the late 20th century, especially in his intense concern with the fundamental nature of quantum mechanics. His courage in devoting his life to an area many considered "philosophical" (ie non-physical) paved the way for the rest of us to reveal our interests and confusions about this area. I am not sure that he would agree with everything in this paper, but I offer it as a tribute to him.

I. STAPP

Stapp[1] has long maintained the position that quantum mechanics must be considered to be a non-local theory in its own right. He believes that the the assumption of "hidden variables" or local realism in Bell's argument is unnecessary, and that no local theory or any form could mimic quantum mechanics. It is not that any hidden variable theory, or locally realistic theory must be non-local in order to mimic quantum mechanics, as Bell showed. It is that quantum mechanics itself is non-local.

In much of the popular vocabulary of physicists, his war has been won. Many physicists, including many of those with an interest in the foundational issues of quantum theory, refer to quantum mechanics as non-local– using Bell's arguments as a justification. By this they usually mean that quantum mechanical entanglement has non-classical features and when pushed, they will back off and agree that that non-locality is not really what Bell's arguments mean. However, they stubbornly insist on using the terminology. (Names or references are purposely omitted to protect the guilty).

Stapp would however like to put this popular misnaming onto a firm footing. Despite a large amount of criticism, he still insists that his analysis of a Hardy type experiment shows that quantum mechanics itself is non-local. Unfortunately, in the face of this criticism, his claims have become more and more diluted.

He has recently published another paper in the American Journal of Physics [1] with new arguments on the non-locality of quantum mechanics. The end of the paper states "This conclusion represents *some sort of failure* of the notion that no influence *of any kind* can act over a space-like interval". "Some sort of failure" is so vague that almost anything can be subsumed under its mantle. Meanwhile "no influence of any kind" is so strong that many innocuous aspects of both classical and quantum physics can fall under this rubric.

Of course neither quantum mechanics nor classical mechanics has never argued that

no influence of any kind can act over a space-like interval. The existence of correlations between widely separated bodies could be taken to imply some sort of action over space-like intervals. A measurement operation, in which the measuring apparatus is only read when widely separated from the object could be taken to act over space-like intervals, since the value of the variable measured on the system in question has changed from unknown to known when the measuring apparatus is read. This is especially true in quantum mechanics where one cannot regard the system in question as having a value for the quantity of interest even in the absence of measurement. Ie, this sentence makes it unclear as to what Stapp is claiming. With a suitably diluted notion of non-locality, any theory could be said to be non-local.

The above paragraph may be taken as unfairly using his infelicitous language to erect and demolish a straw man. Let us therefor look a bit more closely at his argument.

He uses a Hardy-type experiment in his argument. The quantum Hardy-type experiment has been extensively described and generalised. It is a thought experiment in which two (spatially separated) physical systems are described by some state which is weakly entangled between the two systems. The weaker the entanglement, the more striking is the violation of the classical expectations, although the more rare the conditions under which it applies. We can consider the two systems to each be a two level system, and the state to be any state which is not a product state. For any such state, one can find a set of two dynamical variables for each sub-system, call them **L1** and **L2** for the one sub-system, and **R1** and **R2** for the other, with each variable having a pair of eigenvalues, denoted by + and **-**. These attributes have the following four properties in the given state.

In all experiments with the system in that given initial state and in which L1 and R1 are measured, and L1 is found to have value +, then R1 always has value +.

If R1 and L2 are measured, and R1 has value +, then L2 always has value +.

If L2 and R2 are measured, and L2 has value +, then R2 always has value +.

If L1 and R2 are measured, and L1 has value +, then R2 has value - with a probability which approaches unity as the state approaches a product state. This is clearly in conflict with the logical chain

$$L1 = + \Rightarrow R1 = + \Rightarrow L2 = + \Rightarrow R2 = + \tag{1}$$

which one would naively deduce from the chain of bipartite measurements.

To make the above more definite, consider the two systems to be two two level systems, with the usual Pauli matrices . Assume that in the σ_z basis for each the state of the system is

$$|\Psi >= sin(\phi)|++> +cos(\phi)|--> \tag{2}$$

Ie, if ϕ is small, this state is almost a product state. Take L1 to be $cos(2\mu)\sigma_{Lz}+sin(2\mu)\sigma_{Lx}$ where $tan(\mu) = tan(\phi)^2$, and take R1 to be $\cos(2\phi)\sigma_{Rz} + \sin(2\phi)\sigma_{Rx}$, Choosing L2 to be $\frac{1}{\sqrt{2}}(\sigma_{Lz} + \sigma_{Lx})$ and $R2$ to be $\sin(2\phi)\sigma_{Rz} + \cos(2\phi)\sigma_{Rx}$, where These operators obey the above conditions. This choice comes very close to maximizing the probability, $\cos(2\phi)^2$, that if L1 and R2 are measured and L1 is +, then R2 is -. For $\phi << 1$, this probability becomes very close to unity. Note that attribute R2 is almost exactly the negative of R1, and its + eigenvector is almost exactly the − eigenvector of R1. Ie, L2 having value + implies R1 has value + while L1 having value + implies that R2 has value - with high probability. However, for any value of ϕ except 0 (no entanglement but the probability of L1=+ is zero) or $\pi/2$ (maximum entanglement) these operators obey the conditions of this generalised Hardy system.

For any classical system, the first three properties would imply that if L1 has value + then R2 must have value +. The fourth property contradicts this. Stapp's argument is that this chain of reasoning also applies in quantum mechanics. The argument is subtle and uses the language of counterfactuals.

Counterfactual arguments are tricky (see for example Shimony's criticism of this paper by Stapp which is similar to my criticism)[2], and are invariably heavily theory laden. They are not statements about the world, but rather about one's theory of the world. This is especially clear in the example which could be called the argument of Peres's mother [3].

When young his mother asked herself the counterfactual question of whose child, her mother's or her father's, she would have been if her mother and father had each married different people. While she ultimately decided the question was meaningless, it is clear that it would not have been meaningless, and would furthermore have had a definite answer, had her theory of human essence rested upon matrilinear reincarnation. Furthermore, had she asked instead whether her father's or mother's child would have had her blue eyes, we would have had no difficulty giving an answer based on our theories of genetic inheritance. Ie, the meaningfulness and answer to a counterfactual question depends crucially on the theoretical

context in which it is embedded.

The central point of Stapp's argument rests on a proposition which he calls SR. This proposition is (translated to the notation I am using)

If R1 is measured and gives outcome +, then if, instead, R2 had been measured, the outcome would have been +.

By "instead" he does not mean in some other experiment, but means a counterfactual replacement of the measurement of $R1$ by a measurement of $R2$ in the same experiment in which $R1$ was measured. As a counterfactual statement, it can of course never actually be tested by experiment in the real world. As with all counterfactual statements, it is a statement made within the context of a theoretical framework. As such one must be careful to ensure that the replacement makes sense within the context of the theory. Within quantum theory this becomes especially ticklish, since the attribute R2 does not commute with R1, and quantum mechanics thus rules out any interpretation of "instead of" which makes it synonymous with "as well as ". Ie, the measurement of $R1$ inherently destroys the probability structure of the outcomes for $R2$ and interferes with any measurement of $R2$.

The first question to ask is whether, within the theoretical context of quantum mechanics, the statement makes any sense. The statement assumes a number of other postulates– namely that the state of the system before any measurement is the Hardy state, a state which explicitly refers to both L and R. One can certainly argue that in fact, as in the case of Peres' mother, this statement does not make any sense withing the context of quantum mechanics. Because attributes do not have values in the absence of measurement, because the values found in a measurement occur without sufficient cause, are generated out of thin air by the measurement itself, the question of what quantum mechanics would have to say about the counterfactual replacement of R1 with R2 is "nothing" in the absence of any other conditions. But let us push the analysis a little bit further.

Within quantum mechanics, the validity of this counterfactual replacement hinges on whether or not L2 was actually measured. If it was measured, then, because of the prior condition that the state is the Hardy state and the assumption about the measured value of R1, it is a fact that both R1 and L2 have values +. The validity of the counterfactual replacement of R1 with R2 giving the value + then rests on the reality of the measurement of L2. If, on the other hand, L2 was not actually measured, then the validity of the argument rests on a double counterfactual– namely that if instead of not being measured at all, L2

had been measured, it would have had value +, and thus on the second counterfactual substitution of R2 for R1, R2 would have value +. There is no reason to believe that quantum physics makes any sense at all out of such a double counterfactual substitution.

Stapp argues that if, one had, in the past (but space-like separated from) of the measurement of either R1 or R2 referred to in this statement, L2 had been measured, then this statement is true. The measurement result of + for the measurement of R1 would ensure through the correlations inherent in the state that the outcome of the L2 measurement in the past must have been + as well. But, since that result is surely independent of whether or not R1 or R2 were measured in the future, it would still have had outcome + if the experimenter in R had decided to measure R2 instead, and thus, because of the correlations in the state, R2 would then have had value + as well. Thus, given only the knowledge that L2 was measured, the statement SR is true. Of course it is true only because of the existence of the measurement of L2. Without the existence of that measurement, the statement SR is nonsense (ie, untrue).

However Stapp here uses the fee will of the experimentalist and his notion of locality to argue that, as a statement about region R, SR must surely be independent of what experiment was carried out in region L, since, it being space-like separated from R, one can consider the measurement in region L to occur after that in R. Thus SR should continue to be true if L2 were replaced with L1, in which case however, the inference of SR does not follow (and is in fact negated with high probability if the outcome of the L1 measurement is +).

However this notion of locality is strange. SR, is not a statement about region R, rather it is a statement about two different counterfactual worlds, the one in which R1 was measured and the other where R2 was measured. There seems to me to be no argument from locality or anything else which could demand that such a counterfactual relationship should be independent of the actions in region L. The existence of the measurement of L2 plays a crucial role in the establishment of the truth of SR, and there is no reason why that truth should be independent of that measurement. IF SR refereed to some actual state of affairs in a single world (established even by counterfactual reasoning) then such a locality requirement might be reasonable. But as I have stated, the assumption that SR says something about the single real world is a form of realism.

II. VON NEUMAN MEASUREMENT

This Hardy type system can also be used to point out some features and limitations of the von Neuman description of measurement. In establishing the logical consistency of quantum mechanics and in particular of the measurement hypothesis, von Neuman introduces a measurement hypothesis. A measurement on a system could be regarded as a primate operation on that system. Alternatively, one could introduce a measuring apparatus which was itself a quantum system, and whose interactions with the system were fully governed by the laws of quantum mechanics. The measurement process on the original system was now regarded as the establishment of correlations between some dynamic attribute of the apparatus with the "measured" attribute of the system. The measurement, in the primate sense, on this pointer attribute of the apparatus, could be used to infer, by means of the correlations between the two systems, a value for the attribute of the system. He argued that regarding a measurement on the system either as a primitive, or as being inferred from a measurement on an apparatus, are consistent, and equivalent.

However, this model demonstrates limitations of this equivalence in some situations. Because of the correlations inherent in this Hardy state, one can regard the either the system on the left or on the right as the system of interest and the other to be a measuring apparatus. The correlations created by the interaction which placed the system into the partially entangled state are of the kind discussed by von Neuman. In particular, a measurement, in the primitive sense, of R1 giving value + is perfectly correlated with L2 having value +. Ie, a primitive measurement of R1 giving value + is a measurement in the von-Neuman sense of L2 giving value +. (The primitive measurement of R1 giving any value is not equivalent to a generic von Neuman measurement of L2, since the correlation is not valid for R1 having value -.) Now, the primitive measurement of L2, giving value + can also be regarded as a measurement in the von Neuman sense of R2 giving value +. But von Neuman also insisted that there is no difference between a von Neuman and a primitive measurement as far as the system is concerned. Thus, we can take the primate measurement of R1 with value + to be equivalent to the measurement of L2 referred to the above, which was also a measurement of R2 with value +. Ie, by the double application of von Neuman's argument the (primitive) measurement of R1 giving value + can apparently also be regarded as a von Neuman measurement of R2 giving value +.

One might make two objections. The first is that the measurement of R1 destroys the probability distribution of R2, leaving R2 with an entirely different probability distribution. What is R1 then measuring? However, physicists have long engaged in measurements which destroy the system being measured even more completely. When a photon impinges on a photographic plate or a CCD, the fact that the photon is completely destroyed in the process does not change physicists' notion that the photographic plate has measured the position of the photon. Yes, it is destroyed, but just before the destruction the photon had that position.

The second possible objection is that the measurement is very indirect. After all we are operating through the intermediary of L2. Without L2, the measurement of R1 would not allow anything to be inferred about the value of R2. But again, this possibility was already envisioned by von Neuman, who discussed a whole chain of measuring apparatuses. One could "measure" the pointer of the apparatus, either as a primitive operation, or by coupling it again to another super-apparatus, whose pointer we correlated with the pointer of the first apparatus. This chain could be as long as one wished, as long as one had established the chain of correlations between the various pointers and the original attribute in the system to be measured. Ie, there is nothing in the von Neuman equivalence which limits our right to regard R1 having value + as being a measurement of R2.

Note of course that this is a system to which we cannot apply the arguments of "Wigner's Friend". Ie, a separate attempt to measure R2 either by coupling it to some other apparatus , or via a primitive measurement will not give the same result as the result inferred from the measurement of R1. But nowhere in the naive von Neuman analysis is there any requirement that the "Wigner's Friend" argument apply.

But, of course, if one does allow the measurement of R1 with value + to be a valid measurement of R2, the plot grows even more convoluted. One could regard the measurement of L1, giving value +, to be a measurement of R1 (with value +) which is a measurement of L2 (with value +) which is a measurement of R2 (with value +). Again the fact that L2 is destroyed in the primitive measurement of L1 would seems to be irrelevant.

But this leads to a contradiction. For exactly the same correlated state between the measuring apparatus L1 and the system attribute R2 allows one to assume that if L1 has value +, R2 almost certainly has value -. Ie, the equivalence between primitive measurements and von Neuman causal chain measurements fails spectacularly. At the same time it is not

clear exactly where it fails.

Ie, it would seem that one needs to restrict the von Neuman measurement chain such that at each step one can apply a "Wigner's Friend" argument to obtain the same outcome for the measurement as the one inferred from the von Neuman chain. Or equivalently one must restrict the measurement chain so that at no point can a measuring apparatus be regarded as measuring itself.

Bell's theorem and Quantum Systems

Ultimately all arguments for the non-locality of quantum mechanics can be traced back to Bell's arguments [4] in establishing his theorem for "Locally realistic" systems. It seems to be because of the powerful fascination of realism that the violation Bell's inequality for quantum mechanics and for the real physical world is interpreted as a violation of locality. It is worth looking in more detail at Bell's argument and at the differences between quantum and classical systems for each step in the argument. In the following I will use the name Bell to refer to the Clauser, Horn, Shimony and Holt [5] version of the argument.(See also Jarret[6] for a discussion of the experiment).

The setup is that we have two attributes L1 and L2 on the left and R1 and R2 on the right. (these are not the same as the attributes above in the Stapp argument.) Each takes values of ± 1. In the quantum system we will take L1 and L2 to be maximally non-commuting attributes, and can take them as σ_1 and σ_2, the two Pauli spin matrices, and R1 and R2 are also the two sigma matrices for another two level system.The system is set up in a correlated state, and a sequence of measurements are made on the L and R systems. In particular L1 or L2 is measured on the left and R1 or R2 on the right. In each measurement only one of the pair are measured. After the measurements have all been made, a set of correlation functions is measured. Namely

$$[L1\ R1] = \sum_{11} L1_i R1_i / \sum_{11} 1 \tag{3}$$

$$[L1\ R2] = \sum_{12} L1_i R2_i / \sum_{12} 1 \tag{4}$$

$$[L2\ R1] = \sum_{21} L2_i R1_i / \sum_{21} 1 \tag{5}$$

$$[L2\ R2] = \sum_{22} L2_i R2_i / \sum_{22} 1 \tag{6}$$

where in each case terms like $L1_i$ refer to the value obtained for $L1$ in the i^{th} trial and the sum over i is over all instances in which the corresponding attributes were measured. (ie,

$\sum_{12}$ is the sum over all instances in which $L1$ and $R2$ were measured.)

Now, of course each of these correlation function is taken over disjoint sets. It is never the case that both $R1$ and $R2$ were measured in the same instance, and similarly no case where $L1$ and $L2$ were measured in the same instance.

The critical procedure in Bell's proof is to argue using local realism, that even though they were not measured in any instance, all of the operators $L1,\ L2,\ R1,\ R2$ actually have values in each of the instances of measurement. Furthermore, he uses locality to argue that if this is true, then the measured correlation function $< La\ Rb >$, with a, b both taking values 1 and 2 is a good estimator of the (counterfactual) correlator

$$< La\ Rb > \approx \sum La_j Rb_j / \sum_j 1 \tag{7}$$

where this time the sum is taken over all instances in which any measurement was taken. If we assume that the sets are or roughly equal size, in 1/4 of the values of j, these correspond to real values for La and Rb and in 3/4 of the cases at least one of them is the value assumed to exist by counterfactual realism.

Furthermore, locality is used to argue that we can write

$$[L1\ R1] + [L1\ R2] + [L2\ R1] - [L2\ R2] =< L1(R1 + R2) > + < L2(R1 - R2) >=< L1(R1 + R2) + L2(R$$

This is the critical relation. Ie, the whole use of locality and local realism is to argue that the sum of the correlators is equal to the correlation of the sum of the operators.

What is of course interesting about quantum mechanics is this property comes free. If we define **La** and **Ra** as the quantum operators and the expectation values as the quantum expectation values, then quantum mechanics gives us, for free, that

$$\begin{aligned} [L1\ R1] + [L1\ R2] + [L2\ R1] - [L2\ R2] &= \langle\psi|\mathbf{L1}(\mathbf{R1} + \mathbf{R2})|\psi\rangle + \langle\psi|\mathbf{L2}(\mathbf{R1} - \mathbf{R2})|\psi\rangle \quad (9) \\ &= \langle\psi|\mathbf{L1}(\mathbf{R1} + \mathbf{R2}) + \mathbf{L2}(\mathbf{R1} - \mathbf{R2})|\psi\rangle \quad (10) \end{aligned}$$

Since the use of locality in the classical case is solely to demonstrate the truth of something which quantum mechanics apparently gives us for free, the question now arises as to where the difference between the quantum and classical resides.

The first instance is when we examine the meaning of these expectation values. In the classical case, for example $< L1(R1 + R2) >$ is taken to mean something different from $\langle\psi|\mathbf{L1}(\mathbf{R1} + \mathbf{R2})|(\rangle\psi)$. In the classical case, Bell took $R1 + R2$ in each instance to be the

sum of the values of $R1$ and $R2$ for each particular instance. Since by assumption Ra took values of ± 1, $R1+R2$ has values of ± 2 or 0. However, a critical feature of quantum mechanics is that **R1+R2** is an operator, and attribute in its own right, and will take values of $\pm\sqrt{2}$. Furthermore, in all situations in which the operators $R1$ and $R2$ are measured separately, or their sum is measured,

$$|\psi\rangle \mathbf{M}\ \mathbf{R1}\langle\psi| + |\psi\rangle \mathbf{M}\ \mathbf{R2}\langle\psi| = |\psi\rangle \mathbf{M}(\ \mathbf{R1} + \ \mathbf{R2})\langle\psi| \tag{11}$$

where M is any operator which commutes with **Ra**. Ie, measured separately or measured as a sum, these two correlators are identical.

If the classical system is to mimic the quantum system, this must also be true of the classical system. In general since $R1 + R2$ has different values than $\mathbf{R1} + \mathbf{R2}$ (namely $\pm 2, 0$ instead of $\pm\sqrt{2}$ this mimicking is difficult for the classical system to maintain.

Secondly, Bell makes use of another feature. Both of the attributes $R1 + R2$ and $R1 - R2$ are assumed to have possible values of $\pm 2, 0$. Furthermore they are perfectly anti-correlated in that one and only of of the two ever has the value 0 in any one instance of the experiment. Thus in each element of the sum, either $R1 + R2$ or $R1 - R2$ is zero. Since. $L1$ and $L2$ have values of ± 1 we immediately get Bell's theorem, namely that

$$-2 \leq [L1\ R1] + [L1\ R2] + [L2\ R1] - [L2\ R2] \leq 2 \tag{12}$$

The quantum violation comes about by noting that we can find a state, $|\psi\rangle$such that L1 and (R1+R2) are maximally correlated– ie every-time L1 has value +1, R1+R2 has value $+\sqrt{2}$ and every time L1 has value -1, R1+R2 has value $+\sqrt{2}$. That same state $|\psi\rangle$ can be chosen so that L2 and R1-R2 are also maximally correlated. This immediately leads to the quantum correlation

$$[\mathbf{L1\ R1}] + [\mathbf{L1\ R2}] + [\mathbf{L2\ R1}] - [\mathbf{L2\ R2}] = 2\sqrt{2} \tag{13}$$

Where can one locate the difference between the quantum and classical case. A key location is the assumption that the values of $R1 + R2$ take values of $\pm 2, 0$ rather than the $\pm\sqrt{2}$ of the quantum system. Ie, in quantum mechanics the sum of the values is not the same as the values of the sum. This is clearly crucial in Bell's argument.

WE can express this in a slightly different way. If we look at the correlation $< (R1 + R2)(R1 - R2) >$ for the classical system, it is crucial to Bell's argument that this is zero.

Quantum mechanically of course, this expression would not necessarily be zero and in fact in the quantum state under consideration it is non-zero.

The second point, related to the first, is that R1+R2 is anti-correlated with R1-R2 in that a non-zero value of one is perfectly correlated with a zero value of the other. Clearly if the values are not 0 and ± 2 this correlation between the two makes little sense. Finally, the perfect correlations between **R1+R2** and **L1** at the same time as a perfect correlation between **R1-R2** and **L2** obtains in the quantum system is also critical to the possibility of its violating the classical limits. Can a classical system be set up so as to have this same correlation? The answer is of course yes. We take $R1 \pm R2$ to have values $\pm\sqrt{2}$ as for the quantum system. Set up the four states $\{+1, +1, +\sqrt{2}, +\sqrt{2}$, $\{+1, -1, +\sqrt{2}, -\sqrt{2}$, $\{-1, +1, -\sqrt{2}, +\sqrt{2}$, and $\{-1, -1, -\sqrt{2}, -\sqrt{2}$ where these four values are the classical values of L1,L2,(R1+R2), and (R1-R2) respectively. The classical state is now defined by taking each of these states with probability of 1/4. Thus we see that the critical difference between quantum and classical system is in the fact that the sum of values is not the same as the values of the sum. Classically, the values of R1+R2 are just the values of $R1$ added to those of $R2$, namely $\pm 2, 0$ while quantum mechanically they are just $\pm\sqrt{2}$.

We note that the locality has played a weak role. It has acted to allow us to argue that for the classical system, the correlations behave in just the way we would expect the quantum system to behave– namely that the sum of the correlators is just the theoretical correlation of the sum.

Acknowledgments

I would like to thank A. Shimony for showing me a copy of his paper before publication, and drawing my attention to Stapp's latest paper. I would also thank the Canadian Institute for Advanced Research, and the NSERC for their support during this research.

---

[1] H.P. Stapp, Am. J. Phys **71**, 30 (2004).

[2] A. Shimony, Found. Phys (to Appear as a special edition for A. Peres on his 70th birthday) (2005) See also the extensive bibliography in this paper.

[3] W. Unruh, Phys Rev **A 59**, 126 (1999).

[4] See the reprints of the original papers is J.S.Bell **Speakable and Unspeakable in Quantum Mechanics** Cambridge U Press (Cambridge, 1987)

[5] J. F. Clauser, M.A. Horne, A. Shimony and R. A. Holt, Phys. Rev. Lett. 23, 880-884 (1969),

[6] Jon Jarrett "Bell's Theorem: A Guide to the Implications" in J. Cushing, E. McMullin, eds **Philosophical Consequences of Quantum Theory** University of Notre Dame Press (1989)

David Wallace

Statement

and

Readings

Abstract

David Wallace

I'll try to clarify just what decoherence has to do with the emergence of multiple quasiclassical dynamical processes. In particular, I'll try to give an account of the the significance of decoherence for the dynamical evolution of systems, and of what decoherence adds to older and more elementary arguments for classicality—notably, Ehrenfest's theorem. I'll be pretty light on mathematical detail in the talk (the details are filled out more in my contribution to the reader), and I'll confine my attention to the interpretation of unitary quantum mechanics without hidden variables—to Everett's approach to quantum mechanics, in effect.

Decoherence and Ontology (or: How I learned to stop worrying and love FAPP)

David Wallace

July 15, 2009

> The form of a philosophical theory, often enough, is: *Let's try looking over here.*
>
> (Fodor 1985, p. 31)

1 Introduction: taking physics seriously

NGC 1300 (shown in figure 1) is a spiral galaxy 65 million light years from Earth.[1] We have never been there, and (although I would love to be wrong about this) we will never go there; all we will ever know about NGC 1300 is what we can see of it from sixty-five million light years away, and what we can infer from our best physics.

Fortunately, "what we can infer from our best physics" is actually quite a lot. To take a particular example: our best theory of galaxies tells us that that hazy glow is actually made up of the light of hundreds of billions of stars; our best theories of planetary formation tell us that a sizable fraction of those stars

[1] Source: `http://leda.univ-lyon1.fr/`. This photo taken from `http://hubblesite.org/gallery/album/galaxy/pr2005001a/`. [NB: issue of getting credit here.]

Figure 1: The spiral galaxy NGC 1300

have planets circling them, and our best theories of planetology tells us that some of those planets have atmospheres with such-and-such properties. And because I think that those "best theories" are actually pretty *good* theories, I regard those inferences as fairly *reliable*. That is: I think there actually *are* atmospheres on the surfaces of some of the planets in NGC 1300, with pretty much the properties that our theories ascribe to them. That is: I think that those atmospheres *exist*. I think that they are *real*. I *believe* in them. And I do so despite the fact that, at sixty-five million light years' distance, the chance of directly observing those atmospheres is nil.

I present this example for two reasons. The first is to try to demystify — deflate, if you will — the superficially "philosophical" — even "metaphysical" — talk that inevitably comes up in discussions of "the ontology of the Everett interpretation". Talk of "existence" and "reality" can sound too abstract to be relevant to physics (talk of "belief" starts to sound downright theological!) but in fact, when I say that "I believe such-and-such is real" I intend to mean no more than that it is on a par, evidentially speaking, with the planetary atmospheres of distant galaxies.

The other reason for this example brings me to the main claim of this paper. For the form of reasoning used above goes something like this: we have good grounds to take such-and-such physical theory seriously; such-and-such physical theory, taken literally, makes such-and-such ontological claim; therefore, such-and-such ontological claim is to be taken seriously.[2]

Now, if the mark of a serious scientific theory is its breadth of application, its explanatory power, its quantitative accuracy, and its ability to make novel predictions, then it is hard to think of a theory more "worth taking seriously" than quantum mechanics. So it seems entirely apposite to ask what ontological claims quantum mechanics makes, if taken literally, and to take those claims seriously in turn.

And quantum mechanics, taken literally, claims that we are living in a multiverse: that the world we observe around us is only one of countless quasi-classical universes ("branches") all coexisting. In general, the other branches are no more observable than the atmospheres of NGC 1300's planets, but the theory claims that they exist, and so if the theory is worth taking seriously, we should take the branches seriously too. To belabour the point:

According to our best current physics, branches are real.

Everett was the first to recognise this, but for much of the ensuing fifty years it was overlooked: Everett's claim to be "interpreting" existing quantum mechanics, and de Witt's claim that "the quantum formalism is capable of yielding its own interpretation" were regarded as too simplistic, and much discussion on the Everett interpretation (even that produced by advocates such as Deutsch

[2]Philosophers of science will recognise that, for reasons of space, and to avoid getting bogged down, I gloss over some subtle issues in the philosophy of science; the interested reader is invited to consult, e. g. , Newton-Smith (1981), Psillos (1999), or Ladyman and Ross (2007) for more on this topic.

(1985)) took as read that the "preferred basis problem" — the question of how the "branches" were to be defined — could be solved only by adding something additional to the theory. Sometimes that "something" was additional physics, adding a multiplicity of worlds to the unitarily-evolving quantum state (Deutsch (1985, Bell (1981, Barrett (1999)). Sometimes it was a purpose-built theory of consciousness: the so-called "many-minds theories" (Lockwood (1989, Albert and Loewer (1988)). But whatever the details, the end result was a replacement of quantum mechanics by a new theory, and furthermore a new theory constructed specifically to solve the quantum measurement problem. No wonder interest in such theories was limited: if the measurement problem really does force us to change physics, hidden-variables theories like the de Broglie-Bohm theory[3] or dynamical-collapse theories like the GRW theory[4] seem to offer less extravagantly science-fictional options.

It now seems to be widely recognised that if Everett's idea really is worth taking seriously, it must be taken on Everett's own terms: as an understanding of what (unitary) quantum mechanics *already* claims, not as a proposal for how to amend it. There is precedent for this: mathematically complex and conceptually subtle theories do not always wear their ontological claims on their sleeves. In general relativity, it took decades fully to understand that the existence of gravity waves and black holes really is a claim of the theory rather than some sort of mathematical artifact.

Likewise in quantum physics, it has taken the rise of decoherence theory to illuuminate the structure of quantum physics in a way which makes the reality of the branches apparent. But twenty years of decoherence theory, together with the philosophical recognition that to be a "world" is not necessarily to be part of a theory's fundamental mathematical framework, now allow us to resolve — or, if you like, to dissolve — the preferred basis problem in a perfectly satisfactory way, as I shall attempt to show in the remainder of the paper.

2 Emergence and Structure

It is not difficult to see why Everett and de Witt's literalism seemed unviable for so long. The axioms of unitary quantum mechanics say nothing of "worlds" or "branches": they speak only of a unitarily-evolving quantum state, and however suggestive it may be to write that state as a superposition of (what appear to be) classically definite states, we are not justified in speaking of those states as "worlds" unless they are somehow added into the formalism of quantum mechanics. As Adrian Kent put it in his influential (1990) critique of Many-Worlds interpretations:

> ... one can perhaps intuitively view the corresponding components [of the wave function] as describing a pair of independent worlds. But this intuitive interpretation goes beyond what the axioms justify: the

[3] SeeCushing, Fine, and Goldstein (1996) and references therein for more information.
[4] See Bassi and Ghirardi (2003) and references therein for more information.

> axioms say nothing about the existence of multiple physical worlds corresponding to wave function components.

And so it appears that the Everettian has a dilemma: either the axioms of the theory must be modified to include explicit mention of "multiple physical worlds", or the existence of these multiple worlds must be some kind of illusion. But the dilemma is false. It is simply untrue that any entity not directly represented in the basic axioms of our theory is an illusion. Rather, science is replete with perfectly respectable entities which are nowhere to be found in the underlying microphysics. Douglas Hofstader and Daniel Dennett make this point very clearly:

> Our world is filled with things that are neither mysterious and ghostly nor simply constructed out of the building blocks of physics. Do you believe in voices? How about haircuts? Are there such things? What are they? What, in the language of the physiicist, is a hole - not an exotic black hole, but just a hole in a piece of cheese, for instance? Is it a physical thing? What is a symphony? Where in space and time does "The Star-Spangled Banner" exist? Is it nothing but some ink trails in the Library of Congress? Destroy that paper and the anthem would still exist. Latin still *exists* but it is no longer a living language. The language of the cavepeople of France no longer exists at all. The game of bridge is less than a hundred years old. What sort of a thing is it? It is not animal, vegetable, or mineral.
>
> These things are not physical objects with mass, or a chemical composition, but they are not purely abstract objects either - objects like the number pi, which is immutable and cannot be located in space and time. These things have birthplaces and histories. They can change, and things can happen to them. They can move about - much the way a species, a disease, or an epidemic can. We must not suppose that science teaches us that every *thing* anyone would want to take seriously is identifiable as a collection of particles moving about in space and time. Hofstadter and Dennett (1981, pp. 6–7)

The generic philosophy-of-science term for entities such as these is *emergent*: they are not directly definable in the language of microphysics (try defining a haircut within the Standard Model!) but that does not mean that they are somehow independent of that underlying microphysics. To look in more detail at a particularly vivid example,[5] consider Figure 2.[6] Tigers are (I take it!) unquestionably real, objective physical objects, but the Standard model contains quarks, electrons and the like, but no tigers. Instead, tigers should be understood as patterns, or structures, *within* the states of that microphysical theory.

[5] I first presented this example in Wallace (2003).

[6] Photograph @ Philip Wallace, 2007. Reproduced with permission.

Figure 2: An object not among the basic posits of the Standard Model

To see how this works in practice, consider how we could go about studying, say, tiger hunting patterns. In principle — and only in principle — the most reliable way to make predictions about these would be in terms of atoms and electrons, applying molecular dynamics directly to the swirl of molecules which make up, say, the Kanha National Park (one of the sadly diminishing places where Bengal tigers can be found). In practice, however (even ignoring the measurement problem itself!) this is clearly insane: no remotely imaginable computer would be able to solve the 10^{35} or so simultaneous dynamical equations which would be needed to predict what the tigers would do.

Actually, the problem is even worse than this. For in a sense, we *do* have a computer capable of telling us how the positions and momentums of all the molecules in the Kanha National Park change over time. It is called the Kanha National Park. (And it runs in real time!) Even if, *per impossibile*, we managed to build a computer simulation of the Park accurate down to the last electron, it would tell us no more than what the Park itself tells us. It would provide no explanation of any of its complexity. (It would, of course, be a superb vindication of our extant microphysics.)

If we want to understand the complex phenomena of the Park, and not just reproduce them, a more effective strategy can be found by studying the structures observable at the multi-trillion-molecule level of description of this 'swirl of molecules'. At this level, we will observe robust — though not 100% reliable — regularities, which will give us an alternative description of the tiger in a language of cell membranes, organelles, and internal fluids. The principles by which these interact will be derivable from the underlying microphysics, and will involve various assumptions and approximations; hence very occasionally they will be found to fail. Nonetheless, this slight riskiness in our description is overwhelmingly worthwhile given the enormous gain in usefulness of this new description: the language of cell biology is both explanatorily far more powerful, and practically far more useful, than the language of physics for describing tiger behaviour.

Nonetheless it is still ludicrously hard work to study tigers in this way. To reach a really practical level of description, we again look for patterns and

regularities, this time in the behaviour of the cells that make up individual tigers (and other living creatures which interact with them). In doing so we will reach yet another language, that of zoology and evolutionary adaptationism, which describes the system in terms of tigers, deer, grass, camouflage and so on. This language is, of course, the norm in studying tiger hunting patterns, and another (in practice very modest) increase in the riskiness of our description is happily accepted in exchange for another phenomenal rise in explanatory power and practical utility.

The moral of the story is: there are structural facts about many microphysical systems which, although perfectly real and objective (try telling a deer that a nearby tiger is not objectively real) simply cannot be seen if we persist in describing those systems in purely microphysical language. Talk of zoology is of course grounded in cell biology, and cell biology in molecular physics, but the entities of zoology cannot be discarded in favour of the austere ontology of molecular physics alone. Rather, those entities are structures instantiated within the molecular physics, and the task of almost all science is to study structures of this kind.

Of *which* kind? (After all, "structure" and "pattern" are very broad terms: almost any arrangement of atoms might be regarded as some sort of pattern.) The tiger example suggests the following answer, which I have previously Wallace (2003, p.93) called "Dennett's criterion" in recognition of the very similar view proposed by Daniel Dennett (Dennett 1991):

> **Dennett's criterion:** A macro-object is a pattern, and the existence of a pattern as a real thing depends on the usefulness — in particular, the explanatory power and predictive reliability — of theories which admit that pattern in their ontology.

Dennett's own favourite example is worth describing briefly in order to show the ubiquity of this way of thinking: if I have a computer running a chess program, I can in principle predict its next move from analysing the electrical flow through its circuitry, but I have no chance of doing this in practice, and anyway it will give me virtually no understanding of that move. I can achieve a vastly more effective method of predictions if I know the program and am prepared to take the (very small) risk that it is being correctly implemented by the computer, but even this method will be practically very difficult to use. One more vast improvement can be gained if I don't concern myself with the details of the program, but simply assume that whatever they are, they cause the computer to play good chess. Thus I move successively from a language of electrons and silicon chips, through one of program steps, to one of intentions, beliefs, plans and so forth — each time trading a small increase in risk for an enormous increase in predictive and explanatory power.[7]

[7] It is, of course, highly contentious to suppose that a chess-playing computer *really* believes, plans etc. Dennett himself would embrace such claims (see Dennett (1987) for an extensive discussion), but for the purposes of this section there is no need to resolve the issue: the computer can be taken only to 'pseudo-plan', 'pseudo-believe' and so on, without reducing the explanatory importance of a description in such terms.

Nor is this account restricted to the relation between physics and the rest of science: rather, it is ubiquitous within physics itself. Statistical mechanics provides perhaps the most important example of this: the temperature of bulk matter is an emergent property, salient because of its explanatory role in the behaviour of that matter. (It is a common error in textbooks to suppose that statistical-mechanical methods are used only because in practice we cannot calculate what each atom is doing separately: even if we could do so, we would be missing important, objective properties of the system in question if we abstained from statistical-mechanical talk.) But it is somewhat unusual because (unlike the case of the tiger) the principles underlying statistical-mechanical claims are (relatively!) straightforwardly derivable from the underlying physics.

For an example from physics which is closer to the cases already discussed, consider the case of quasi-particles in solid-state physics. As is well known, vibrations in a (quantum-mechanical) crystal, although they can in principle be described entirely in terms of the individual crystal atoms and their quantum entanglement with one another, are in practice overwhelmingly simpler to describe in terms of 'phonons' — collective excitations of the crystal which behave like 'real' particles in most respects. And furthermore, this sort of thing is completely ubiquitous in solid-state physics, with different sorts of excitation described in terms of different sorts of "quasi-particle" — crystal vibrations are described in terms of phonons; waves in the magnetisation direction of a ferromagnet are described in terms of magnons, collective waves in a plasma are described in terms of plasmons, etc.

Are quasi-particles real? They can be created and annihilated; they can be scattered off one another; they can be detected (by, for instance, scattering them off "real" particles like neutrons); sometimes we can even measure their time of flight; they play a crucial part in solid-state explanations. We have no more evidence than this that "real" particles exist, and so it seems absurd to deny that quasi-particles exist — and yet, they consist only of a certain pattern within the constituents of the solid-state system in question.

When *exactly* are quasi-particles present? The question has no precise answer. It is essential in a quasi-particle formulation of a solid-state problem that the quasi-particles decay only slowly relative to other relevant timescales (such as their time of flight) and when this criterion (and similar ones) are met then quasi-particles are definitely present. When the decay rate is much too high, the quasi-particles decay too rapidly to behave in any 'particulate' way, and the description becomes useless explanatorily; hence, we conclude that no quasi-particles are present. It is clearly a mistake to ask *exactly* when the decay time is short enough ($2.54 \times$ the interaction time?) for quasi-particles not to be present, but the somewhat blurred boundary between states where quasi-particles exist and states when they don't should not undermine the status of quasi-particles as real, any more than the absence of a precise boundary to a mountain undermines the existence of mountains.

One more point about emergence will be relevant in what follows. In a certain sense emergence is a bottom-up process: knowledge of all the microphysical facts about the tiger and its environment suffices to derive all the tiger-level facts

(in principle, and given infinite computing power). But in another sense it is a top-down process: no *algorithmic* process, applied to a complex system, will tell us what higher-level phenomena to look for in that system. What makes it true that (say) a given lump of organic matter has intentions and desires is not something derivable algorithmically from that lump's microscopic constituents; it is the fact that, when it occurs to us to try interpreting its behaviour in terms of beliefs and desires, that strategy turns out to be highly effective.

3 Decoherence and quasiclassicality

We now return to quantum mechanics, and to the topic of decoherence. In this section I will briefly review decoherence theory, in a relatively simple context (that of non-relativistic particle mechanics) and in the environment-induced framework advocated by, e. g. , Joos, Zeh, Kiefer, Giulini, Kubsch, and Stametescu (2003) and Zurek (1991, 2003). (An alternative formalism — the "decoherent histories" framework advocated by, e. g. , Gell-Mann and Hartle (1990) and Halliwell (1998) — is presented in the Introduction to this volume and in Halliwell's contribution to this volume.)

The basic setup is probably familiar to most readers. We assume that the Hilbert space $\mathcal{H}$ of the system we are interested in is factorised into "system" and "environment" subsystems, with Hilbert spaces $\mathcal{H}_S$ and $\mathcal{H}_E$ respectively —

$$\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_E. \tag{1}$$

Here, the "environment" might be a genuinely external environment (such as the atmosphere or the cosmic microwave background); equally, it might be an "internal environment", such as the microscopic degrees of freedom of a fluid. For decoherence to occur, there needs to be some basis $\{|\alpha\rangle\}$ of $\mathcal{H}_S$ such that the dynamics of the system-environment interaction give us

$$|\alpha\rangle \otimes |\psi\rangle \longrightarrow |\alpha\rangle \otimes |\psi; \alpha\rangle \tag{2}$$

and

$$\langle \psi; \alpha | \psi; \beta \rangle \simeq \delta(\alpha - \beta). \tag{3}$$

on timescales much shorter than those on which the system itself evolves. (Here I use α as a "schematic label". In the case of a discrete basis $\delta(\alpha - \beta)$ is a simple Kronecker delta; in the case of a continuous basis, such as a basis of wavepacket states, then (3) should be read as requiring $\langle \alpha | \beta \rangle \simeq 0$ unless $\alpha \simeq \beta$.) In other words, the environment effectively "measures" the state of the system and records it. (The orthogonality requirement can be glossed as "record states are distinguishable", or as "record states are dynamically sufficiently different", or as "record states can themselves be measured"; all, mathematically, translate into a requirement of orthogonality). Furthermore, we require that this measurement happens quickly: quickly, that is, relative to other relevant dynamical timescales for the system. (I use "decoherence timescale" to refer to the characteristic timescale on which the environment measures the system.)

Decoherence has a number of well-known consequences. Probably the best-known is diagonalisation of the system's density operator. Of course, *any* density operator is diagonal in some basis, but decoherence guarantees that the system density operator will rapidly become diagonal in the $\{|\alpha\rangle\}$ basis, independently of its initial state: any initially non-diagonalised state will rapidly have its non-diagonal elements decay away.

Diagonalisation is a synchronic result: a constraint on the system at all times (or at least, on all time-intervals of order the decoherence timescale). But the more important consequence of decoherence is diachronic, unfolding over a period of time much longer than the decoherence timescale. Namely: because the environment is constantly measuring the system in the $\{|\alpha\rangle\}$ basis, any interference between distinct terms in this basis will be washed away. This means that, in the presence of decoherence, the system's dynamics is *quasi-classical* in an important sense. Specifically: if we want to know the expectation value of any measurement on the system at some future time, it suffices to know what it would be were the system prepared in each particular $|\alpha\rangle$ at the present time (that is, to start the system in the state $|\alpha\rangle \otimes |\psi\rangle$ (for some environment state $|\psi\rangle$ whose exact form is irrelevant within broad parameters) and evolve it forwards to the future time), and then take a weighted sum of the resultant values. Mathematically speaking, this is equivalent to treating the system as though it were in some definite but unknown $|\alpha\rangle$.

Put mathematically: suppose that the superoperator $\mathcal{R}$ governs the evolution of density operators over some given time interval, so that if the system intially has density operator ρ then it has density operator $\mathcal{R}(\rho)$ after that time interval. Then in the presence of decoherence,

$$\mathcal{R}(\rho) = \int \mathrm{d}\alpha \ \langle\alpha|\,\rho\,|\alpha\rangle\, \mathcal{R}(|\alpha\rangle\,\langle\alpha|). \tag{4}$$

(Again: this integral is meant schematically, and should be read as a sum or an integral as appropriate.)

And of course, quasi-classicality is rather special. The reason, in general, that the quantum state cannot *straightforwardly* be regarded as a probabilistic description of a determinate underlying reality is precisely that interference effects prevent the dynamics being quasi-classical. In the presence of decoherence, however, those interference effects are washed away.

4 The significance of decoherence

It might then be thought — perhaps, at one point, it was thought — that decoherence alone suffices to solve the measurement problem. For if decoherence picks out a certain basis for a system, and furthermore has the consequence that the dynamics of that system are quasi-classical, then — it might seem — we can with impunity treat the system not just as *quasi*-classical but straightforwardly as classical. In effect, this would be to use decoherence to give a precise and

observer-independent definition of the collapse of the wavefunction: the quantum state evolves unitarily as long as superpositions which are not decohered from one another do not occur; when such superpositions do occur, the quantum state collapses instantaneously into one of them. To make this completely precise would require us to discretize the dynamics so that the system evolves in discrete time steps rather than continuously The decoherent-histories formalism mentioned earlier is a rather more natural mathematical arena to describe this than the continuous formalism I developed in section 3, but the result is the same in any case: decoherence allows us to extract from the unitary dynamics a space of *histories* (strings of projectors onto decoherence-preferred states) and to assign probabilities to each history in a consistent way (i. e. , without interference effects causing the probability calculus to be violated.

From a conceptual point of view there is something a bit odd about this strategy. Decoherence is a dynamical process by which two components of a complex entity (the quantum state) come to evolve independently of one another, and it occurs due to rather high-level, emergent consequences of the particular dynamics and initial state of our Universe. Using this rather complex high-level process as a criterion to define a new fundamental law of physics is, at best, an exotic variation of normal scientific practice. (To take a philosophical analogy, it would be as if psychologists constructed a complex theory of the brain, complete with a physical analysis of memory, perception, reasoning and the like — and then decreed that, as a new fundamental law of physics (and not a mere definition), a system was conscious if and only if it had those physical features.[8])

Even aside from such conceptual worries, however, a pure-decoherence solution to the measurement problem turns out to be impossible on technical grounds: the decoherence criterion is both too strong, and too weak, to pick out an appropriate set of classical histories from the unitary quantum dynamics.

That decoherence is too *strong* a condition should be clear from the language of section 3. Everything there was approximate, effective, for-all-practical-purposes: decoherence occurs on short timescales (not instantaneously); it causes interference effects to become negligible (not zero); it approximately diagonalises the density operator (not exactly); it approximately selects a preferred basis (not precisely). And while approximate results are fine for calculational shortcuts or for emergent phenomena, they are most unwelcome when we are trying to define new fundamental laws of physics. (Put another way, a theory cannot be 99.99804% conceptually coherent.)

That it is too *weak* is more subtle, but ultimately even more problematic. There are simply *far too many* bases picked out by decoherence — in the language of section 3 there are far too many system-environment splits which give rise to an approximately decoherent basis for the system; in the language of decoherent histories, there are far too many choices of history that lead to consistent classical probabilities. Worse, there are good reasons (cf Dowker and

[8] As it happens, this is not a straw man: David Chalmers has proposed something rather similar. See Chalmers (1996) for an exposition, and Dennett (2001) for some sharp criticism.

Kent (1996))to think that many, many of these histories are wildly non-classical.

What can be done? Well, if we turn away from the abstract presentation of decoherence theory, and look at the concrete models (mathematical models and computer simulations) to which decoherence has been applied, and if, in those models, we make the sort of system/environment split that fits our natural notion of environment (so that we take the environment, as suggested previously, to be — say — the microwave background radiation, or the residual degrees of freedom of a fluid once its bulk degrees of freedom have been factored out), then we find two things.

Firstly: The basis picked out by decoherence is approximately a coherent-state basis: that is, it is a basis of wave-packets approximately localised in both position and momentum. And secondly: The dynamics is quasi-classical not just in the rather abstract, bloodless sense used in section 3, but in the sense that the behaviour of those wave-packets approximates the behaviour predicted by classical mechanics.

In more detail: let $|q,p\rangle$ denote a state of the system localised around phase-space point (q,p). Then decoherence ensures that the state of the system+environment at any time t can be written as

$$|\Psi\rangle = \int \mathrm{d}q\,\mathrm{d}p\,\alpha q,p;t\,|q,p\rangle\otimes|\epsilon(q,p)\rangle \tag{5}$$

with $\langle\epsilon(q,p)|\epsilon(q',p')\rangle = 0$ unless $q \simeq q'$ and $p \simeq p'$. The conventional (i. e. , textbook) interpretation of quantum mechanics tells us that $|\alpha(q,p)|^2$ is the probability density for finding the system in the vicinity of phase-space point (q,p).[9] Then in the presence of decoherence, $|\alpha|^2(q,p)$ evolves, to a good approximation, like a *classical* probability density on phase space: it evolves, approximately, under the Poisson equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(|\alpha(q,p)|^2\right) \simeq \frac{\partial H}{\partial q}\frac{\partial|\alpha(q,p)|^2}{\partial p} - \frac{\partial H}{\partial p}\frac{\partial|\alpha(q,p)|^2}{\partial q} \tag{6}$$

where $H(q,p)$ is the Hamiltonian.

On the assumption that the system is classically non-chaotic (chaotic systems add a few subtleties), this is equivalent to the claim that each individual wave-packet follows a classical trajectory on phase space. Structurally speaking, the dynamical behaviour of each wave-packet is the same as the behaviour of a macroscopic classical system. And if there are multiple wave-packets, the system is dynamically isomorphic to a collection of independent classical systems.

(*Caveat*: this does not mean that the wave-packets are actually evolving on phase space. If phase space is understood as the position-momentum space of a collection of classical point particles, then *of course* the wave-packets are

[9] At a technical level, this requires the use of phase-space POVMs (i. e. , positive operator valued measures, a generalisation of the standard projection-valued measures; see, e. g. , Nielsen and Chuang (2000) for details): for instance, the continuous family $\{N\,|q,p\rangle\,\langle q,p|\}$ is an appropriate POVM for suitably-chosen normalisation constant N. Of course, this or any phase-space POVM can only be defined for measurements of accuracy $\leq \hbar$.

not evolving on phase space. They are evolving on a space isomorphic to phase space. Henceforth when I speak of phase space, I mean this space, not the "real" phase space.)

So: if we pick a particular choice of system-environment split, we find a "strong" form of quasi-classical behaviour: we find that the system is isomorphic to a collection of dynamically independent simulacra of a classical system. We did not find this isomorphism by some formal algorithm; we found it by making a fairly unprincipled choice of system-environment split and then noticing that that split led to interesting behaviour. The interesting behaviour is no less real for all that.

We can now see that all three of the objections at the start of this section point at the same — fairly obvious — fact: decoherence is an emergent process occurring *within* an already-stated microphysics: unitary quantum mechanics. It is not a mechanism to define a part *of* that microphysics. If we think of quasiclassical histories as emergent in this way, then

- The "conceptual mystery" dissolves: we are not using decoherence to define a dynamical collapse law, we are just using it as a (somewhat pragmatic) criterion for when quantum systems display quasiclassical behaviour.
- There is nothing problematic about the approximateness of the decoherence process: as we saw in section 2, this is absolutely standard features of emergence.
- Similarly, the fact that we had no algorithmic process to tell us in a bottom-up way what system-environment splits would lead to the discovery of interesting structure is just a special case of section 2's observation that emergence is in general a somewhat top-down process.

Each decoherent history is an emergent structure within the underlying quantum state, on a par with tigers, tables, and the other emergent objects of section 2 — that is, on a par with practically all of the objects of science, and no less real for it.

But the price we pay for this account is that, if the fundamental dynamics are unitary, at the fundamental level there is no collapse of the quantum state. There is just a dynamical process — decoherence — whereby certain components of that state become dynamically autonomous of one another. Put another way: if each decoherent history is an emergent structure within the underlying microphysics, and if the underlying microphysics doesn't do anything to prioritise one history over another (which it doesn't) then all the histories exist. That is: a unitary quantum theory with emergent, decoherence-defined quasi-classical histories is a many-worlds theory.

5 Simulation or reality?

At this point, a skeptic might object:

> All you have shown is that certain features of the unitarily-evolving quantum state are isomorphic to a classical world. If that's true, the most it shows that the quantum state is running a simulation of the classical world. But I didn't want to recover a *simulation* of the world. I wanted to recover *the world*.

I rather hope that this objection is a straw man: as I attempted to illustrate in section 2, this kind of structural story about higher-level ontology (the classical world is a structure instantiated in the quantum state) is totally ubiquitous in science. But it seems to be a common enough thought (at least in philosophical circles) to be worth engaging with in more detail.

Note firstly that the very assumption that a certain entity which is structurally like our world is not *our world* is manifestly question-begging. How do we know that space is three-dimensional? We look around us. How do we know that we are seeing something fundamental rather than emergent? We don't; all of our observations (*pace* Maudlin, this volume) are structural observations, and only the sort of aprioristic knowledge now fundamentally discredited in philosophy could tell us more.

Furthermore, physics itself has always been totally relaxed about this sort of possibility. A few examples will suffice:

- Solid matter — described so well, and in such accord with our observations, in the language of continua — long ago turned out to be only emergently continuous, only emergently solid.

- Just as solid state physics deals with emergent quasi-particles, so — according to modern "particle physics" — elementary particles themselves turn out to be emergent from an underlying quantum field. Indeed, the "correct" — that is, most explanatorily and predictively useful — way of dividing up the world into particles of different types turns out to depend on the energy scales at which we are working.[10]

- The idea that particles should be emergent from some field theory is scarcely new: in the 19th century there was much exploration of the idea that particles were topological structures within some classical continuum (cf Epple (1998)), and later, Wheeler (1962) proposed that matter was actually just a structural property of a very complex underlying spacetime. Neither proposal eventually worked out, but for technical reasons: the proposals themselves were seen as perfectly reasonable.

- The various proposals to quantize gravity have always been perfectly happy with the idea that space itself would turn out to be emergent.

[10] The best known example of this phenomenon occurs in quantum chromodynamics: treating the quark field in terms of approximately-free quarks works well at very high energies, but at lower energies the appropriate particle states are hadrons and mesons; see, e. g. , Cheng and Li (1984) and references therein for details. For a more mathematically tractable example (in which even the correct choice of whether particles are fermionic or bosonic is energy-level-dependent), see chapter 5 of Coleman (1985), esp. pp. 246–253.

From Borel dust to non-commutative geometry to spin foam, program after program has been happy to explore the possibility that spacetime is only emergently a four-dimensional continuum.[11]

- String theory, currently the leading contender for a quantum theory of gravity, regards spacetime as fundamentally high-dimensional and only emergently four-dimensional, and the recent development of the theory makes the nature of that emergence more and more indirect (it has been suggested, for instance, that the "extra" dimensions may be several centimetres across[12]). The criterion for emergence, here as elsewhere, are dynamical: if the functional integrals that define the cross-sections have the approximate functional form of functional integrals of fields on four-dimensional space, that is regarded as sufficient to establish emergence.

Leaving aside these sorts of naturalistic[13] considerations, we might ask: *what* distinguishes a simulation of a thing from the thing itself? It seems to me that there are two relevant distinctions:

Dependency: Tigers don't interact with simulations of tigers; they interact with the computers that run those simulations. The simulations are instantiated in "real" things, and depend on them to remain in existence.

Parochialism: Real things have to be made of a certain sort of stuff, and/or come about in a certain sort of way. Remarkably tiger-like organisms in distant galaxies are not tigers; synthetic sparkling wine, however much it tastes like champagne, is not champagne unless its origins and makeup fit certain criteria.

Now, these considerations are themselves problematic. (Is a simulation of a person themselves a person? — see (Hofstadter 1981) for more thoughts on these matters). But, as I hope is obvious, both considerations are question-begging in the context of the Everett interpretation: only if we begin with the assumption that our world is instantiated in a certain way can we argue that Everettian branches are instantiated in a relevantly different way.

6 How many worlds?

We are now in a position to answer one of the most commonly asked questions about the Everett interpretation,[14] namely: how much branching actually happens? As we have seen, branching is caused by any process which magnifies microscopic superpositions up to the level where decoherence kicks in, and there are basically three such processes:

[11] For the concept of Borel dust, see Misner, Thorne, and Wheeler (1973, p.1205); for references on non-commutative geometry, see http://www.alainconnes.org/en/downloads.php; for references on spin foam, see Rovelli (2004).

[12] For a brief introduction to this proposal, see Dine (2007, chapter 29).

[13] I use "naturalism" in Quine's sense ((Quine 1969)): a naturalistic philosophy is one which regards our best science as the only good guide to our best epistemology,

[14] Other than "and you believe this stuff?!", that is.

1. Deliberate human experiments: Schrödinger's cat, the two-slit experiment, Geiger counters, and the like.

2. "Natural quantum measurements", such as occur when radiation causes cell mutation.

3. Classically chaotic processes, which cause small variations in initial conditions to grow exponentially, and so which cause quantum states which are initially spread over small regions in phase space to spread over macroscopically large ones. (See Zurek and Paz (1994) for more details; I give a conceptually oriented introduction in Wallace (2001).)

The first is a relatively recent and rare phenomenon, but the other two are ubiquitous. Chaos, in particular, is everywhere, and where there is chaos, there is branching (the weather, for instance, is chaotic, so there will be different weather in different branches). Furthermore, there is no sense in which these phenomena lead to a naturally *discrete* branching process. Quantum chaos gives rise to macroscopic superpositions, and so to decoherence and to the emergence of a branching structure, but that structure has no natural "grain". To be sure, by choosing a certain discretisation of (phase-)space and time, a discrete branching structure will emerge, but a finer or coarser choice would also give branching. And there is no "finest" choice of branching structure: as we fine-grain our decoherent history space, we will eventually reach a point where interference between branches ceases to be negligible, but there is no precise point where this occurs. As such, the question "how many branches are there?" does not, ultimately, make sense.

This may seem paradoxical — certainly, it is not the picture of "parallel universes" one obtains from science fiction. But as we have seen in this chapter, it is commonplace in emergence for there to be some indeterminacy (recall: when *exactly* are quasi-particles of a certain kind present?) And nothing prevents us from making statements like:

> Tomorrow, the branches in which it is sunny will have combined weight 0.7

— the combined weight of all branches having a certain macroscopic property is very (albeit not precisely) well-defined. It is only if we ask: "*how many* branches are there in which it is sunny", that we end up asking a question which has no answer.

This bears repeating, as it is central to some of the arguments about probability in the Everett interpretation:

> Decoherence causes the Universe to develop an emergent branching structure. The existence of this branching is a robust (albeit emergent) feature of reality; so is the mod-squared amplitude for any *macroscopically described* history. But there is *no* non-arbitrary decomposition of macroscopically-described histories into "finest grained" histories, and *no* non-arbitrary way of counting those histories.

(Or, put another way: asking how many worlds there are is like asking how many experiences you had yesterday, or how many regrets a repentant criminal has had. It makes perfect sense to say that you had many experiences or that he had many regrets; it makes perfect sense to list the most important categories of either; but it is a non-question to ask *how many*.)

If this picture of the world seems unintuitive, a metaphor may help.

1. Firstly, imagine a world consisting of a very thin, infinitely long and wide, slab of matter, in which various complex internal processes are occurring — up to and including the presence of intelligent life, if you like. In particular one might imagine various forces acting in the plane of the slab, between one part and another.

2. Now, imagine stacking many thousands of these slabs one atop the other, but without allowing them to interact at all. If this is a "many-worlds theory", it is a many-worlds theory only in the sense of the philosopher David Lewis (Lewis 1986): none of the worlds are dynamically in contact, and no (putative) inhabitant of any world can gain empirical evidence about any other.

3. Now introduce a weak force normal to the plane of the slabs — a force with an effective range of 2-3 slabs, perhaps, and a force which is usually very small compared to the intra-slab force. Then other slabs will be detectable from within a slab but will not normally have much effect on events within a slab. If this is a many-worlds theory, it is a science-fiction-style many-worlds theory (or maybe a Phillip Pullman or C.S. Lewis many-worlds theory[15]): there are many worlds, but each world has its own distinct identity.

4. Finally, turn up the interaction sharply: let it have an effective range of several thousand slabs, and let it be comparable in strength (over that range) with characteristic short-range interaction strengths within a slab. Now, dynamical processes will not be confined to a slab but will spread over hundreds of adjacent slabs; indeed, *evolutionary* processes will not be confined to a slab, so living creatures in this universe will exist spread over many slabs. At this point, the boundary between slabs becomes epiphenomenal. Nonetheless, this theory is *stratified* in an important sense: dynamics still occurs predominantly along the horizontal axis and events hundreds of thousands of slabs away from a given slab are dynamically irrelevant to that slab.[16] One might well, in studying such a system, divide it into layers thick relative to the range of the inter-slab force — and emergent dynamical processes in those layers would be no less real just because the exact choice of layering is arbitrary.

[15] See, for instance, Pullman's *Northern Lights* or Lewis's *The Magician's Nephew*.

[16] Obviously there would be ways of constructing the dynamics so that this was not the case: if signals could easily propagate vertically, for instance, the stratification would be lost. But it's only a thought experiment, so we can construct the dynamics how we like.

Ultimately, though, that a theory of the world is "unintuitive" is no argument against it, provided it can be cleanly described in mathematical language. Our intuitions about what is "reasonable" or "imaginable" were designed to aid our ancestors on the savannahs of Africa, and the Universe is not obliged to conform to them.

7 Conclusion

The claims of the Everett interpretation are:

- At the most fundamental level, the quantum state is all there is – quantum mechanics is about the structure and evolution of the quantum state in the same way that (e.g.) classical field theory is about the structure and evolution of the fields.
- As such, the "Everett interpretation of quantum mechanics" is just quantum mechanics itself, taken literally (or, as a philosopher of science might put it, Realist-ically) as a description of the Universe. De Witt has been widely criticized for his claim that "the formalism of quantum mechanics yields its own interpretation" (DeWitt 1970), but there is nothing mysterious or Pythagorean about it: *every* scientific theory yields its own interpretation, or rather (cf David Deutsch's contribution to this volume) the idea that one can divorce a scientific theory from its interpretation is confused.
- "Worlds" are mutually dynamically isolated structures instantiated within the quantum state, which are structurally and dynamically "quasiclassical".
- The existence of these "worlds" is established by decoherence theory.

No *postulates* about the worlds have needed to be added: the question of whether decoherence theory does indeed lead to the emergence of a quasiclassical branching structure is (at least in principle) settled *a priori* for any particular quantum theory once we know the initial state. It is not even a *postulate* that decoherence is the source of all "worlds"; indeed, certain specialised experiments — notably, some algorithms on putative quantum computers — would also give rise to multiple quasiclassical worlds at least locally; cf. Deutsch (1997).[17]

[17]Since much hyperbole and controversy surrounds claims about Everett and quantum computation, let me add two deflationary comments:

1. There is no particular reason to assume that *all* or even *most* interesting quantum algorithms operate by any sort of "quantum parallelism" (that is: by doing different classical calculations in a large number of terms in a superposition and then interfering them). Indeed, Grover's algorithm does not seem open to any such analysis. But Shor's algorithm, at least, does seem to operate in this way.
2. The correct claim to make about Shor's algorithm is not (*pace* (Deutsch 1997)) that the calculations *could not* have been done other than by massive parallelism, but simply that the actual explanation of how they *were* done — that is, the workings of Shor's

I will end this discussion on a lighter note, aimed at a slightly different audience. I have frequently talked to physicists who accept Everett's interpretation, accept (at least when pressed!) that this entails a vast multiplicity of quasiclassical realities, but reject the "many-worlds" label for the interpretation — thhey prefer to say that there is only one world but it contains many non- or hardly-interacting quasiclassical parts.

But, as I hope I have shown, the "many worlds" of Everett's many-worlds interpretation are not fundamental additions to the theory. Rather, they are emergent entities which, according to the theory, are present in large numbers. In this sense, the Everett interpretation is a "many-worlds theory" in just the same sense as African zoology is a "many-hippos theory": that is, there are entities whose existence is entailed by the theory which deserve the name "worlds". So, to Everettians cautious about the "many-worlds" label, I say: come on in, the water's lovely.

References

Albert, D. Z. and B. Loewer (1988). Interpreting the Many Worlds Interpretation. *Synthese 77*, 195–213.

Barrett, J. A. (1999). *The quantum mechanics of minds and worlds*. Oxford: Oxford University Press.

Bassi, A. and G. Ghirardi (2003). Dynamical reduction models. *Physics Reports 379*, 257. Available online at http://arxiv.org/abs/quant-ph/0302164.

Bell, J. S. (1981). Quantum Mechanics for Cosmologists. In C. J. Isham, R. Penrose, and D. Sciama (Eds.), *Quantum Gravity 2: a second Oxford Symposium*, Oxford. Clarendon Press. Reprinted in Bell (1987), pp. 117–138.

Bell, J. S. (1987). *Speakable and Unspeakable in Quantum Mechanics*. Cambridge: Cambridge University Press.

Chalmers, D. J. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press.

Cheng, T.-P. and L.-F. Li (1984). *Gauge Theory of Elementary Particle Physics*. Oxford, UK: Oxford University Press.

Coleman, S. (1985). *Aspects of Symmetry*. Cambridge, UK: Cambridge University Press.

Cushing, J. T., A. Fine, and S. Goldstein (Eds.) (1996). *Bohmian Mechanics and Quantum Theory: An Appraisal*, Dordrecht. Kluwer Academic Publishers.

algorithm — does involve massive parallelism.

For some eloquent (albeit, in my view, mistaken) criticisms of the link between quantum computation and the Everett interpretation, see Steane (2003).

Dennett, D. C. (1987). *The intentional stance*. Cambridge, Mass.: MIT Press.

Dennett, D. C. (1991). Real patterns. *Journal of Philosophy 87*, 27–51. Reprinted in *Brainchildren*, D. Dennett, (London: Penguin 1998) pp. 95–120.

Dennett, D. C. (2001). The fantasy of first-person science. Available online at http://ase.tufts.edu/cogstud/papers/chalmersdeb3dft.htm.

Deutsch, D. (1985). Quantum Theory as a Universal Physical Theory. *International Journal of Theoretical Physics 24*(1), 1–41.

Deutsch, D. (1997). *The Fabric of Reality*. London: Penguin.

DeWitt, B. (1970). Quantum Mechanics and Reality. *Physics Today 23*(9), 30–35. Reprinted in (DeWitt and Graham 1973).

DeWitt, B. and N. Graham (Eds.) (1973). *The many-worlds interpretation of quantum mechanics*. Princeton: Princeton University Press.

Dine, M. (2007). *Supersymmetry and String Theory: Beyond the Standard Model*. Cambridge: Cambridge University Press.

Dowker, F. and A. Kent (1996). On the consistent histories approach to quantum mechanics. *Journal of Statistical Physics 82*, 1575–1646.

Epple, M. (1998). Topology, matter and space, i: Topological notions in 19th-century natural philosophy. *Archive for History of Exact Sciences 52*, 297–392.

Fodor, J. A. (1985). Fodor's guide to mental representation: the intelligent auntie's vade-mecum. *Mind 94*, 76–100. Reprinted in Jerry A. Fodor, *A Theory of Content and Other Essays* (MIT Press, 1992).

Gell-Mann, M. and J. B. Hartle (1990). Quantum Mechanics in the Light of Quantum Cosmology. In W. H. Zurek (Ed.), *Complexity, Entropy and the Physics of Information*, pp. 425–459. Redwood City, California: Addison-Wesley.

Halliwell, J. J. (1998). Decoherent histories and hydrodynamic equations. *Physical Review D 35*, 105015.

Hofstadter, D. R. (1981, May). Metamagical themas: A coffeehouse conversation on the Turing test to determine if a machine can think. *Scientific American*, 15–36. Reprinted as 'The Turing Test: A Coffeehouse Conversation' in Hofstadter and Dennett (1981).

Hofstadter, D. R. and D. C. Dennett (Eds.) (1981). *The Mind's I: Fantasies and Reflections on Self and Soul*. London: Penguin.

Joos, E., H. D. Zeh, C. Kiefer, D. Giulini, J. Kubsch, and I. O. Stametescu (2003). *Decoherence and the Appearence of a Classical World in Quantum Theory* (2nd Edition ed.). Berlin: Springer.

Kent, A. (1990). Against Many-Worlds Interpretations. *International Journal of Theoretical Physics A5*, 1764. Available at http://www.arxiv.org/abs/gr-qc/9703089.

Ladyman, J. and D. Ross (2007). *Every Thing Must Go: Metaphysics Naturalized*. Oxford: Oxford University Press.

Lewis, D. (1986). *On the Plurality of Worlds*. Oxford: Basil Blackwell.

Lockwood, M. (1989). *Mind, Brain and the Quantum: the compound 'I'*. Oxford: Blackwell Publishers.

Misner, C. W., K. S. Thorne, and J. A. Wheeler (1973). *Gravitation*. New York: W.H. Freeman and Company.

Newton-Smith, W. S. (1981). *The Rationality of Science*. London: Routledge.

Nielsen, M. A. and I. L. Chuang (2000). *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press.

Psillos, S. (1999). *Scientific Realism: How Science Tracks Truth*. London: Routledge.

Quine, W. (1969). Epistemology naturalized. In *Ontological Relativity and Other Essays*. New York: Columbia University Press.

Rovelli, C. (2004). *Quantum Gravity*. Cambridge: Cambridge University Press.

Steane, A. (2003). A quantum computer only needs one universe. *Studies in the History and Philosophy of Modern Physics 34*, 469–478.

Wallace, D. (2001). Implications of Quantum Theory in the Foundations of Statistical Mechanics. Available online from http://philsci-archive.pitt.edu.

Wallace, D. (2003). Everett and Structure. *Studies in the History and Philosophy of Modern Physics 34*, 87–105. Available online at http://arxiv.org/abs/quant-ph/0107144 or from http://philsci-archive.pitt.edu.

Wheeler, J. A. (1962). *Geometrodynamics*. New York: Academic Press.

Zurek, W. H. (1991). Decoherence and the transition from quantum to classical. *Physics Today 43*, 36–44. Revised version available online at http://arxiv.org/abs/quant-ph/0306072.

Zurek, W. H. (2003). Decoherence, einselection, and the quantum origins of the classical. *Reviews of Modern Physics 75*, 715. Available online at http://arxiv.org/abs/quant-ph/0105127.

Zurek, W. H. and J. P. Paz (1994). Decoherence, chaos and the second law. *Physical Review Letters 72*(16), 2508–2511.

Chapter 3

Chaos, decoherence, and branching

> Classicality simply does not follow "as $\hbar \to 0$" in most *physically* interesting cases... The Planck constant is $\hbar = 1.05459 \times 10^{-27}$ erg s and — *licentia mathematica* to vary it notwithstanding — it is a *constant*.
>
> Wojciech Zurek and Juan Pablo Paz[1]

3.1 Emergent quasi-classicality in simple isolated systems

In chapter 2, we saw how, in outline, the quasi-classical "worlds" of the Everett interpretation emerge from the underlying quantum mechanics. They do so because

1. Certain quantum-mechanical histories of certain systems instantiate — simulate, if you like — a quasi-classical history.

2. Superpositions of those histories then instantiate multiple quasi-classical histories — always assuming that interference between histories can be neglected.

The purpose of this chapter is to go from this rather hand-waving description of emergence of worlds, to something much more quantitative and precise. We begin by considering the textbook example of emergent quasi-classicality in quantum physics: a single, isolated system whose characteristic action is large compared with $\hbar$.

[1] Zurek and Paz (1995b).

Consider, therefore, a massive point particle of mass m, moving in some potential $V(\mathbf{x})$. The Hamiltonian of this particle is then

$$\widehat{H} = \frac{\widehat{P}^2}{2m} + V(\widehat{X}). \tag{3.1}$$

Under what circumstances does this system behave approximately classically? That is (in the language of chapter 2): under what circumstances does it instantiate a classical dynamical system? There is a fairly standard answer: it does so when the state of the system is a wave-packet, reasonably localised in position and momentum, and when the centre of that wavepacket follows an approximately classical trajectory. In fact, since we know from Ehrenfest's theorem[2] that the expectation values of $\widehat{P}$ and $\widehat{X}$ evolve in the same way as their classical counterparts, the former condition — that the wave-packet remains localised — suffices to ensure the latter.

So far, so banal; but let us dwell on it a little longer. What justifies our regarding a localised wavepacket following an approximately classical trajectory as an approximately classical state? Sometimes it can seem that some sort of tacit "hidden variable" theory is present: that the state is approximately classical because the probabilities it predicts for particle location are highly peaked around a certain classical trajectory. But this will not do, of course (at least, not unless we are actually trying to develop that hidden-variable theory!) Rather, the real reason that we can regard the quantum state as approximately classical is that it is dynamically isomorphic, very nearly, to a system of a classical point particle.

It may help to consider in more detail how that isomorphism works. We could understand it in the position representation: the trajectory of the centre of a localised wave-packet defines a line in configuration space, and that line is (very nearly) a solution to the classical dynamical equations for a mass-m point particle. It is somewhat more perspicuous when viewed using one of the phase-space POVMs discussed in chapter 1: a wave-packet defines a small region (of area $\sim \hbar^3$) in phase space via this method, and because its average phase-space position evolves classically (by Ehrenfest's theorem) and its spread around that phase-space position remains small, the trajectory followed by that small region is itself a solution to the classical dynamical equations in Hamiltonian form. (I call this "more perspicuous" because it makes transparent the fact that an instantaneous quantum state suffices to pick out the corresponding classical trajectory; in the position representation the needed momentum information is unhelpfully encoded in the phase structure of the wavepacket.)

[2]For an account of Ehrenfest's theorem, see Joos *et al* (2003, pp. 87–88) or any textbook discussion, such as Cohen-Tannoudji, Diu, and Laloë (1977, pp. 240–245), Sakurai (1994, pp. 84–87), or Townsend (1992, pp. 153–156).

In either case, both rules:

$$\begin{gathered} |\langle \mathbf{x}|\psi\rangle|^2 \simeq 0 \text{ unless } \mathbf{x} \simeq \mathbf{q}(t) \\ \leftrightarrow \text{Wave-packet is centred at } \mathbf{q}(\mathbf{t}) \\ \leftrightarrow |\psi\rangle \text{ instantiates classical particle with trajectory } \mathbf{q}(t) \end{gathered} \tag{3.2}$$

and

$$\begin{gathered} |\langle \psi|\, \widehat{\Pi}(\mathbf{q},\mathbf{p})\, |\psi\rangle \simeq 0 \text{ unless } (\mathbf{x},\mathbf{p}_0) \simeq (\mathbf{q},\mathbf{p}) \\ \leftrightarrow \text{Wave-packet is centred at } (\mathbf{q},\mathbf{p}) \\ \leftrightarrow |\psi\rangle \text{ instantiates classical particle at phase-space location } (\mathbf{q},\mathbf{p})) \end{gathered} \tag{3.3}$$

ultimately pick out the same structure[3] in the quantum system. Notice also that we see again the emptiness of questions like "which is the correct phase-space POVM? Within broad limits, any such POVM will succeed in picking out the structure we are interested in (and, outside those broad limits, we simply are not using a POVM which makes manifest that structure; it's still *there*).

To see another important property of this emergent dynamics, let us consider a particular (overcomplete) basis $|\mathbf{q},\mathbf{p}\rangle$ of wavepacket states centred at phase-space point $|\mathbf{q},\mathbf{p}\rangle$, one of which is the actual wavepacket of the system. To a very good approximation, then, if the phase-space point $(\mathbf{q},\mathbf{p})$ evolves over time to $(\mathbf{q}(t),\mathbf{p}(t))$ then the corresponding quantum state evolves to $|\mathbf{q}(t),\mathbf{p}(t)\rangle$ over the same period. (Perhaps the wavepacket will spread out a little, so that it is not exactly any single element of the basis, but (we are assuming that) it remains reasonably localised.) This is a somewhat remarkable property of the phase-space basis: the dynamics takes elements of the basis to other elements of the basis. Fairly clearly, this can only occur exactly for an orthonormal basis in the trivial cases where that basis is an eigenbasis of the Hamiltonian; in this case, though, the overcompleteness of the basis (and, in most realistic situations, our willingness to settle for a very high but not 100% level of precision) allows basis preservation and nontriviality to coexist.

Because of the property of basis preservation, the various classical histories instantiated by different wave-packet states can coexist. To see this, suppose $|\psi_1(t)\rangle$ and $|\psi_2(t)\rangle$ each instantiate some classical history. The structures which make up

[3]Note for philosophers: I am helping myself here to something that was not actually developed in chapter 2: namely, an identity criterion for structures. Something like "two structures are the same when they are instantiated by precisely the same states of the instantiating theory" will probably do, but in practice I am again happy to fall back on the fact that in practice we have no trouble working out when two structures are really the same one differently described, and to leave the task of making this precise to future work in general philosophy of science .

those classical histories are, as we have seen, structures in the expectation values of the phase-space POVMs, and so a superposition

$$|\Psi(t)\rangle = \alpha\,|\psi_1(t)\rangle + \beta\,|\psi_2(t)\rangle \tag{3.4}$$

will instantiate both histories simultaneously provided that those structures are not erased by interference between the terms in the superposition.

The particular expectation values in this case are

$$\langle\Psi(t)|\,\widehat{\Pi}(\mathbf{q},\mathbf{p})\,|\Psi(t)\rangle = |\alpha|^2\,\langle\psi_1(t)|\,\widehat{\Pi}(\mathbf{q},\mathbf{p})\,|\psi_1(t)\rangle + |\beta|^2\,\langle\psi_2(t)|\,\widehat{\Pi}(\mathbf{q},\mathbf{p})\,|\psi_2(t)\rangle$$

$$+2\mathrm{Re}\left(\alpha^*\beta\,\langle\psi_1(t)|\,\widehat{\Pi}(\mathbf{q},\mathbf{p})\,|\psi_2(t)\rangle\right) \tag{3.5}$$

The first two terms are simply the weighted sum of the two expectation values of the original structures. The third term — the interference term — will vanish, to a very good approximation, at all times, because if $|\psi_1(t)\rangle$ and $|\psi_2(t)\rangle$ are instantiating different quasi-classical histories in the way described above, they will be localised at different phase-space points at all times (this is basis preservation in action: a superposition of two orthogonal terms in the basis will forever after remain a superposition of two orthogonal terms in the basis). So we are just left with the first two terms, and with the observation that the expectation values of the phase-space POVMs have the structure of two independent, non-interacting classical worlds.

Notice that it is not merely the linearity of quantum mechanics which allows us to interpret superpositions as instantiating multiple structures.[4] Rather, it is the disappearence of interference terms between the relevant terms in those superpositions. Basis preservation is a sufficient condition for this to occur; as we will shortly see, it is not a necessary condition.

So: in this simple model, we seem to have achieved emergent classicality — and to have achieved it in a way which leads to superpositions representing multiple quasi-classical worlds. Furthermore, nothing we did really relied on the system being a single particle: generalising to a system with N degrees of freedom, with some Hamiltonian like

$$\widehat{H} = \sum_i \frac{1}{2m_i}\hat{P}_i^2 + V(\hat{Q}_1,\ldots\hat{Q}_n) \tag{3.6}$$

is straightforward. (In realistic cases the degrees of freedom will normally be grouped into triples, of course, given the three-dimensional[5] nature of the universe we live

[4] Notwithstanding the overly simplistic claims of Wallace (2003a).

[5] A worry: is it really three-dimensional, given that the theory seems to be about the quantum state and not about entities in space at all? I address this question in chapter 8 of Wallace (2010c); for now, it suffices to note that the theory is *emergently* three-dimensional, that the emergent classical dynamics that it instantiates is on three-dimensional space.

in). Localised wavepackets of this system will now pick out trajectories in a high-dimensional space, and these trajectories will instantiate the dynamics of a classical theory with N degrees of freedom. Superficially, this seems to be everything that Everett-interpreted quantum mechanics needs.

We shall see shortly that in fact this account has a number of conceptual problems. However, there is a technical problem that is at least as severe: namely, we are relying on the assumption that the wave-packets of isolated macroscopic systems do, indeed, remain in fairly-well-localised states whose trajectories satisfy classical dynamics. As we shall see, things are not actually that simple.

3.2 Dynamical properties of isolated quantum systems

In this section I want to investigate how initially-localised quantum states actually do behave under different Hamiltonians. We can consider this under fairly general conditions: we will assume that the system has N degrees of freedom and that its Hamiltonian is of the form of equation 3.6: that is, the sum of a term in $\widehat{Q}_1, \ldots \widehat{Q}_N$ and of a quadratic term in each $\widehat{P}_i$. For convenience I will just write (q, p) to encode the $2N$ position and momentum coordinates in the system's phase space.

As we saw in Box 1.1, given a set of coherent (wave-packet) states $|q, p\rangle$, each one representing a Gaussian wavepacket localised around q in position space and p in momentum space, then the set of (improper) operators $|q, p\rangle \langle q, p|$ provides a satisfactory phase-space POVM for the system. It follows that the function

$$H\psi(q, p) = |\psi\rangle \langle q, p| \, |q, p\rangle \langle \psi| \tag{3.7}$$

(known as the *Husimi function*) expresses the phase-space structure of the quantum state $|\psi\rangle$. It can further be shown that, given the Husimi function, the state vector can be recovered (up to phase).[6]

Because the Husimi function is somewhat cumbersome to track, however, it will be useful to set out an alternative way of representing the phase-space structure of the state: the so-called Wigner function [7]

$$W(q, p) = \frac{1}{\pi^{N/2}} \int \mathrm{d}y \, \mathrm{e}^{-y^2/4\lambda^2} \mathrm{e}^{ipy} \langle q - y/2| \, \rho \, |q + y/2\rangle \,, \tag{3.8}$$

[6] The Husimi function was first introduced in Husimi (1940); see Hillery *et al* (1984) for a review of its properties.

[7] The Wigner function was first introduced in Wigner (1932) and explored further by Moyal (1949); see citeNhilleryetal for a review of its properties.

which is related to the Husimi function by

$$H(q,p) = \frac{1}{\pi^{N/2}} \int \mathrm{d}q'\, \mathrm{d}p'\, \mathrm{e}^{-(q-q')^2/\lambda^2} \mathrm{e}^{-(p-p')^2/\lambda^2} W(q',p'). \tag{3.9}$$

(That is, the Husimi function is obtained from the Wigner function by smearing it over a small region of phase space.)

It is sometimes said that the Wigner function "is not a probability distribution because it is not positive definite". This is misleading at best. It is indeed the case that the Wigner function is not guaranteed to be nonnegative, but the deeper reason why it is not a probability distribution is that (at the risk of being repetitive), if"phase space" means "space representing the positions and momenta of all the particles", then there is no phase space in quantum mechanics (except emergently), and the Husimi function, positive definite though it may be, is no more a probability distribution on phase space than the Wigner function. The only reason for using these "phase space" representations of the state at all is that we are interested in the emergent quasi-classical structures within the state, and these structures are most perspicuously identifiable in the phase-space representation.

The Wigner function is computationally somewhat more tractable than the Husimi function (being obtained rather more straightforwardly from the position representation of the state): its dynamics can be expressed in closed form as

$$\dot{W} = \{H,W\}_{MB} \equiv \frac{2i}{\hbar} \sin\left(\frac{\hbar}{2i}\{\cdot,\cdot\}_{PB}\right) \cdot (H,W), \tag{3.10}$$

where $\{\cdot,\cdot\}_{PB}$ is the classical Poisson bracket and $\{\cdot,\cdot\}_{MB}$ is known as the *Moyal bracket* (Moyal 1949). Less compactly but more illuminatingly, we can expand (3.10) as

$$\dot{W} = \{H,W\}_{PB} + \frac{\hbar^2}{24}\frac{\partial^3 V}{\partial q^3}\frac{\partial^3 W}{\partial p^3} + O(\hbar^4). \tag{3.11}$$

showing that the quantum dynamics is the classical dynamics plus correction terms in successively higher powers of $\hbar^2$. This seems very reassuring: as $\hbar \to 0$, we revert to classical dynamics. But as Zurek and Paz reminded us in the quotation at the start of this chapter, this formal mathematical limit is not directly physically relevant: what matters for emergent classicality is the behaviour of macroscopic systems for fixed $\hbar$.

The simplest such system is a free particle in one dimension. For this system, the higher-order terms in the Moyal bracket vanish, and classical dynamics holds exactly. The spread of a wavepacket in this situation is then a purely classical phenomenon: if the wavepacket has position spread Δq (and thus momentum spread at least $\sim \hbar/\Delta q$), over a time t the part of the packet with momentum $p + \hbar/\Delta q$ will travel

a distance $\hbar t/m\Delta q$ further than the part with momentum p, and so the position spread will increase to $\Delta q + \hbar t/m\Delta q$. Over a time t, then, the minimum size that a packet will obtain is

$$\Delta q(t) \sim \sqrt{\frac{\hbar t}{m}}. \tag{3.12}$$

Not only does this decrease to zero as $\hbar \to 0$, it does so satisfactorily fast. An invisibly small dust mote, for instance (ten microns across, say, with a mass of $\sim 10^{-12}$kg) , if evolving freely, could be prepared in a wavepacket state that remained of width $\leq$ 1cm for the age of the Universe; a bowling ball with a mass of $\sim 1kg$, could be similarly prepared in a state that remained of width $\leq 10^{-8}$m.

No real systems are entirely free, of course; but some real systems (sometimes called *regular*) share with free systems the property that phase-space distributions spread out at a rate linear in time. For these systems, (3.12) will remain a fairly good approximation for the minimum achievable spread of a classical distribution of area $\sim \hbar$. (I continue to work in one dimension for convenience; the generalisation is straightforward). Furthermore, the classical spread will be a good approximation to the quantum spread as long as the higher terms in the Moyal bracket are small. The first such term, evaluated for a wavepacket of initial size Δq, will be of order

$$\hbar^2 V'''(q) \times \left(\frac{1}{\Delta p}\right)^3 \sim \hbar^{1/2} V'''(q)(\Delta q)^3. \tag{3.13}$$

Again, this goes to zero as $\hbar \to 0$; again, it does it sufficiently quickly that, for systems of micron size or above, quantum corrections are utterly negligible.

So: regular, isolated systems do indeed instantiate quasi-classical dynamics if they are above a certain size. Unfortunately, most Hamiltonians do not give rise to regular dynamics. Much more commonly, a system is *chaotic*: phase-space regions in such systems spread out exponentially, not linearly. (Or, more accurately: they spread out exponentially in some directions and contract exponentially in others, so as to conserve phase-space volume.) In such a system, the spread of a classical packet of initial width Δq (and so of a quantum wavepacket of width Δq, as long as classical dynamics remains approximately valid for it) will be of the form[8]

$$\Delta q(t) \simeq \mathrm{e}^{t/\tau_L} \Delta q \tag{3.14}$$

where τ_L is the so-called Lyapunov exponent.[9] Since the wavepacket cannot be dramatically narrower than (3.12) on pain of being so delocalised in momentum space

[8]The results in this section are based on results in Berry and Balzas (1979), Zurek and Paz (1995a) and Zurek and Paz (1994).

[9]In the classical theory of chaos, a system is chaotic if (roughly) infinitesimally close points in phase space diverge exponentially in some directions; the Lyapunov exponent is the timescale of this exponential divergence. See (e. g.) Cvitanović *et al* (2009) for a formal definition.

that it rapidly spreads out anyway, a crude estimate for the minimum achievable wavepacket spread after time t is

$$\Delta q(t) \sim \mathrm{e}^{t/\tau_L}\sqrt{\frac{\hbar t}{m}}; \tag{3.15}$$

equivalently, we have

$$\ln \Delta q(t) \sim \frac{t}{\tau_L} + \ln\left(\frac{\hbar t}{m}\right) = \frac{t}{\tau_L} + \ln\left(\frac{\hbar \tau_L}{m}\right) + \ln\left(\frac{t}{\tau_L}\right) \tag{3.16}$$

or

$$\ln \Delta q(t) \sim \frac{t}{\tau_L} + \ln\left(\frac{\hbar t}{m}\right) \tag{3.17}$$

in the regime where $t \gg \tau_L$. If the packet becomes so spread that it samples regions of appreciably different potentials, it certainly will no longer instantiate a classical trajectory, so a criterion for emergent classicality (at least of the form we have so far discussed) is that $\Delta q(t)$ remains below the lengthscale on which this happens. Writing this lengthscale as L, we find that classicality fails once

$$t \geq \tau_L \ln\left(\frac{Lm}{\hbar \tau_L}\right). \tag{3.18}$$

The good news is: t does go to infinity as $\hbar \to 0$. The bad news is: thanks to the logarithm in (3.18), it does so alarmingly slowly. Suppose that our dust mote (mass $\sim 10^{-12}$ kg) is experiencing chaotic dynamics with a Lyapunov timescale of ~ 10 seconds in a region where the potential varies on a scale of ~ 10cm. (These numbers are off the top of my head; the logarithm means that (3.18) is enormously insensitive to the details.) Then classicality fails when

$$t \geq 10 \text{ s} \times \ln 10^{22}. \tag{3.19}$$

The logarithm of 10^{22} is about 50, so the system will cease to behave classically after about 500 seconds. This is uncomfortably short compared with, say, the age of the Universe. Nor does the problem go away for still larger systems. To borrow an example from Zurek and Paz (1995a), Saturn's moon Hyperion tumbles chaotically in its orbit on a Lyapunov timescale of about 20 days. Hyperion weighs $\sim 10^20$ kg and is $\sim 10^5$ m in size, so (if treating it as an isolated system were appropriate) its wavefunction would become highly nonclassical once

$$t \geq 20 \text{ days} \times \ln 10^{64} \sim 10 \text{ yrs}. \tag{3.20}$$

Since we are discussing a supposed *many-worlds theory*, one tempting idea is to say: this spreading out of the quantum state is exactly the branching of worlds that we were expecting to find. Whether or not this is conceptually appropriate, though (more on this later), it fails on technical grounds in this case, for presumably a necessary condition for the idea is that the phase-space distribution defined by the quantum state — localised or no — continues to follow, approximately, the classical dynamics. If not, the various parts of the wavefunction cannot suffice to instantiate dynamically independent worlds. And it turns out that classical dynamics, too, fail for chaotic systems. For consider the correction term (3.13), the leading-order correction to the classical dynamics. This term grows as $1/(\Delta p)^3$. But — thanks to the conservation of phase-space volume — generically we would expect Δp to shrink exponentially as Δq grows. (Chaos generally "fibrillates" systems, turning compact regions into long, thin ones.) In this case, the correction term will also grow exponentially, and so on a timescale which increases logarithmically with $1/\hbar$, but will in general still be uncomfortably short, we would expect classical dynamics to fail for the system's Wigner function.

To conclude: chaotic, isolated, unitarily evolving quantum systems cannot approximate classical ones on acceptably long timescales.

3.3 The need for decoherence

Leaving aside for the moment the technical problems with chaotic isolated systems, there remain severe *conceptual* problems with the naive recovery of quasi-classicality which was sketched in section 3.1. For a start, notice that we found the emergent structure in the quantum state not by any principled means, but by our pre-existing intuitions that those variables which we call "position" and "momentum" would indeed turn out to function like classical position and momentum. We might worry that, in fact, this supposed "structure" is an artefact of our choosing those variables, and that we might have found similar results in any number of alternative ways.

I think that this is more of a "niggling doubt" than it is a real worry. As chapter 2 stressed, emergent properties cannot be deductively found by applying any sort of algorithm to the instantiating theory (the fact that biology is instantiated by molecular physics is something we realised after the fact, not something we deduced from physics). If quasi-classical dynamics are present, then this is a real, objective fact about the system. Nonetheless, it would be more satisfactory if we were able to gain a better understanding of why the structures we seek are instantiated in the phase-space basis.

A much more serious reason to be unsatisfied is that we have assumed, with-

out any justification, that the system we are studying — consisting, recall, of the *macroscopic* degrees of freedom of some isolated system — can indeed be considered as isolated. For a system such as a rigid body, we know (from the translational invariance of the global Hamiltonian) that the centre-of-mass degrees of freedom are dynamically independent of the internal degrees of freedom, but we have no reason to assume that those centre-of-mass degrees of freedom are dynamically isolated from other systems. And in more general cases we cannot even neglect the internal degrees of freedom — in a fluid, for instance, the macroscopic coordinates would normally be taken to be spatial averages of fluid density and momentum over small regions, but there is no reason at all to suppose that those coordinates are dynamically independent of the remaining coordinates (no reason except, perhaps, classical intuition — but to invoke *that* would be to beg the question.) Indeed, even in the case of the "rigid body" we do not escape such worries — the very claim that the body is "rigid" cannot be taken as primitive, but must be regarded as something which ought to be derivable from the underlying physics of its constituents.

A further concern is that, if quantum systems always behave approximately classically, we would not have needed quantum mechanics! Obviously our theory must accommodate situations — such as quantum measurements — where classical mechanics breaks down even at the macroscopic scale. In these situations, we have as yet no solid reason to expect the "branching" behaviour which the Everett interpretation claims is the correct description of measurement.

To summarise, the main problems with directly reading off quasi-classical structure from the dynamics of isolated macroscopic systems are:

1. It is inaccurate, or at least question-begging, to treat the macroscopic degrees of freedom of a system as dynamically isolated from its residual degrees of freedom.

2. In chaotic systems, it is simply false that the system has any states which behave quasi-classically over acceptably long timescales.

3. In situations like quantum measurements where the dynamics are not even approximately classical, we have no reason to assume that a macroscopic quantum system remains treatable as a collection of non-interacting quasi-classical systems.

As we will see in the remainder of this chapter, all of these problems are satisfactorily solved once *decoherence* — the interaction of a system's macroscopic degrees of freedom with its internal and external environments — is properly allowed for.[10]

[10]There is a terminological issue here. Some authors (such as Wojciech Zurek, Erich Joos, and

Furthermore, this section's "niggling doubt" is also at least partially assuaged: decoherence provides at least a substantial part of the answer to the question of why it is the quasi-classical degrees of freedom which instantiate the interesting structures in macroscopic quantum systems.

3.4 Environment-induced decoherence: a simple model

"Decoherence" is the process by which the environment of a system continually interacts with, and becomes entangled with, that system. Its most well-known property is the suppression of coherence in coherent superpositions of states in that basis — hence the name — but, as we will see, its real significance is much greater. However, suppression of coherence is a convenient way to begin our investigations.[11]

Let us begin by considering a simple model: suppose that we have two one-particle systems, the first much heavier than the other and that the first system is prepared in a superposition of two localised wavepackets separated from one another by some distance large compared to the packet width. That is: let the first system be in state

$$|\psi\rangle = \alpha\,|\psi_{q_1}\rangle + \beta\,|\psi_{q_2}\rangle \tag{3.21}$$

where $|\psi_{q_i}\rangle$ is localised around q_i, and suppose for simplicity that $|\psi\rangle$ is stationary on relevant timescales. And suppose that the Hamiltonian of the system contains some interaction term

$$\widehat{H}_{int} = V(\widehat{X} - \widehat{x}) \tag{3.22}$$

where $\widehat{X}$ and $\widehat{x}$ are the position operator of the first and second particles respectively.

H. Dieter Zeh) use "decoherence" to mean specifically an *environment-induced* process. Others (such as Jonathan Halliwell, James Hartle and Murray Gell-Mann) use 'decoherence' to mean any process by which interference between quasi-classical histories is suppressed: to them, then, the evolution of the isolated regular system in section 3.1 is also decoherent. Halliwell (2010), in fact, calls this sort of decoherence "conservation-induced decoherence", and distinguishes it from "environment-induced decoherence". In this thesis, I largely follow the former authors' terminology, writing just 'decoherence' where Halliwell would write "environment-induced decoherence"; I do, however, follow standard terminology in referring to a *history space* (as discussed in section 3.8 and subsequently) as decoherent in the event that its decoherence functional vanishes.

[11] Here and subsequently I draw extensively on the discussions of decoherence by Zurek (1991, 1998, 2003), Joos *et al* (2003) and Schlosshauer (2007), and while my models and analyses are in many cases not explicitly lifted from any single source, I claim no particular originality for any of them.

If one of α or β is zero, then to a very good approximation this problem reduces to a standard piece of scattering theory: the second particle is scattering off a scattering centre at $x = q_i$, and (again, to a very good approximation) the first particle does not change at all. (See box 3.4 for a proof of this.)

So the dynamics is

$$|\psi_{q_i}\rangle \otimes |\phi_0\rangle \longrightarrow |\psi_{q_i}\rangle \otimes \left|\phi_i^+\right\rangle \tag{3.31}$$

where $\left|\phi_i^+\right\rangle$ is some post-scattering state: for instance, if $|\psi_0\rangle$ was a plane wave or nearly so, then $\left|\phi_i^+\right\rangle$ will be a superposition of a plane wave with an outgoing spherical wave centred on q_i. By the linearity of the Schrödinger equation, then, the general evolution has the form

$$|\psi\rangle \otimes |\phi_0\rangle \longrightarrow \alpha\, |\psi_{q_1}\rangle \otimes \left|\phi_1^+\right\rangle + \beta\, |\psi_{q_2}\rangle \otimes \left|\phi_2^+\right\rangle . \tag{3.32}$$

That is: in the case where the first particle is in a superposition, but not in the case where it is not, the scattering interaction causes the two particles to become entangled. We might even say (though nothing hangs on this way of talking) that the second particle has measured the position of the first.

The level of entanglement can be quantified by considering the density operator for the first particle in the $|\psi_{q_i}\rangle$ basis. If we idealise it as having exactly two possible position states, $|\psi_{q_1}\rangle$ and $|\psi_{q_1}\rangle$, then tracing over equation 3.32 tells us that the first particle's density operator evolves like

$$\rho_0 = |\alpha|^2\, |\psi_{q_1}\rangle \langle\psi_{q_1}| + |\beta|^2\, |\psi_{q_2}\rangle \langle\psi_{q_2}| + \alpha^*\beta\, |\psi_{q_2}\rangle \langle\psi_{q_1}| + \beta^*\alpha\, |\psi_{q_1}\rangle \langle\psi_{q_2}|$$

$$\Longrightarrow \rho_+ = |\alpha|^2\, |\psi_{q_1}\rangle \langle\psi_{q_1}| + |\beta|^2\, |\psi_{q_2}\rangle \langle\psi_{q_2}| + \alpha^*\beta \left\langle\phi_1^+|\phi_2^+\right\rangle |\psi_{q_2}\rangle \langle\psi_{q_1}| + \beta^*\alpha \left\langle\phi_2^+|\phi_1^+\right\rangle |\psi_{q_1}\rangle \langle\psi_{q_2}| \tag{3.33}$$

or, in matrix form,

$$\rho_0 = \begin{pmatrix} |\alpha|^2 & \alpha\beta^* \\ \alpha^*\beta & |\beta|^2 \end{pmatrix} \longrightarrow \rho_+ = \begin{pmatrix} |\alpha|^2 & \alpha\beta^* \left\langle\phi_2^+|\phi_1^+\right\rangle \\ \alpha^*\beta \left\langle\phi_1^+|\phi_2^+\right\rangle & |\beta|^2 \end{pmatrix} . \tag{3.34}$$

The off-diagonal terms provide a measure of the coherence between the two possible positions of the first particle: when they have magnitude equal to $|\alpha^*\beta|$, the first particle is in a pure state and so not at all entangled with the second particle; if they are equal to zero, then the entanglement is maximal (and, if we apply the quantum measurement algorithm, the first particle's state cannot be empirically distinguished from a probabilistic mixture of the two positions.)

Hence, if the scattering is very weak, or if the wavelength of the incoming particle is large compared with $q_2 - q_1$, then $\left\langle\phi_2^+|\phi_1^+\right\rangle \simeq 1$, and the systems become only

Box 3.1: Scattering of light particles off heavy ones

If two interacting particles have position operators $\widehat{X}_1$ and $\widehat{X}_2$ and Hamiltonian

$$\widehat{H} = \frac{1}{2m_1}\widehat{P}_1^2 + \frac{1}{2m_2}\widehat{P}_2^2 + V(\widehat{X}_2 - \widehat{X}_1), \tag{3.23}$$

we define the centre-of-mass coordinates by

$$\widehat{R} = \frac{m_1}{M}\widehat{X}_1 + \frac{m_2}{M}\widehat{X}_2; \quad \widehat{r} = \widehat{X}_2 - \widehat{X}_1 \tag{3.24}$$

where $M = m_1 + m_2$ is the total mass of the system, and the conjugate momenta by

$$\widehat{P} = \widehat{P}_1 + \widehat{P}_2; \quad \widehat{p} = \mu\left(\frac{\widehat{P}_2}{m_2} - \frac{\widehat{P}_1}{m_1}\right) \tag{3.25}$$

where $\mu = m_1 m_2/(m_1 + m_2)$ is the *reduced mass*. It is then easy to verify that $[r, P] = [R, p] = 0$ and $[r, p] = [R, P] = i\hbar$, and that the Hamiltonian can be rewritten as

$$\widehat{H} = \frac{1}{2M}\widehat{P}^2 + \frac{1}{2\mu}\widehat{p}^2 + V(\widehat{r}); \tag{3.26}$$

in other words, the system is mathematically equivalent to the tensor product of a free particle with mass M and a particle with mass μ interacting with a scattering centre at the origin.

We now shift to the position basis. If $\Psi(x_1, x_2; t)$ is the system's wavefunction, we will suppose that at time 0 it is factorised:

$$\Psi(x_1, x_2; 0) = \psi(x_1)\phi(x_2); \tag{3.27}$$

in the centre-of-mass coordinates, then, this is

$$\Psi(R, r; 0) = \psi(R - m_2 r/M)\phi(R + m_1 r/M). \tag{3.28}$$

We now assume that $M \gg m$. Then to a very good approximation, $m_2/M = 0$, m_1/M and we have

$$\Psi(R, r; 0) \simeq \psi(R)\phi(R + r). \tag{3.29}$$

If we further assume that ψ is tightly localised around $R = q$ then we can approximate this as

$$\Psi(R, r; 0) \simeq \psi(R)\phi(q + r) : \tag{3.30}$$

that is, the wavefunction factorises. Since there is no interaction between q and R, this remains the case over time: ϕ evolves as if scattering from a centre at $r = -q$, and ψ remains stationary (and, inter alia, justifies our continuing to assume it to be tightly peaked around $R = q$). Reversing the coordinate transformation at the end of the interaction process gives us our result.

slightly entangled. At the other extreme, if $|\psi\rangle$ is highly localised, incident on q_1, and strongly scattered, then $\left\langle \phi_2^+ | \phi_1^+ \right\rangle \simeq 0$, and entanglement is almost maximal.

So: prepare a heavy particle in a macroscopic superposition and expose it to a scattering environment, and that environment will become entangled with the particle, causing the coherence between the terms in the superposition to decay. If the environment consists of short-wavelength particles which interact strongly with the system, the coherence will be completely lost after a single scattering event. Even if the environment is not so constituted, sufficiently many scattering events will still suffice to remove the coherence: it can be shown (Joos *et al* 2003, pp. 64–67) that the rate is approximately

$$\langle q_1 | \, \rho(t) \, | q_2 \rangle = \langle q_1 | \, \rho(0) \, | q_2 \rangle \exp\left[-\Lambda t (q_1 - q_2)\right] \tag{3.35}$$

where

$$\Lambda \sim k^2 F \sigma, \tag{3.36}$$

where k is the wavenumber, F the incoming particle flux, and σ is the interaction cross-section.

In fact, it is by now well known that in realistic situations, coherence is lost very, very quickly. For a one-micron dust particle, the value of Λ due to the atmosphere is 10^{36}; the value due to sunlight is 10^{21}; even the value due to the cosmic background radiation is 10^6. The rates for larger objects are correspondingly more rapid: Schroödinger's cat, for instance, would endure in a coherent macroscopic superposition for only $\sim 10^{-35}$ seconds before the microwave background radiation — let alone the atmosphere — sufficed to destroy the coherence. (Of course, absent some non-unitary dynamical process of a kind for which we have no evidence, the cat-plus-environment system remains in a superposition of live-cat and dead-cat states. Decoherence, alone, does not solve the measurement problem.)

Furthermore, although these examples all involve an *external* environment, there is no need to make this restriction. There is, in fact, every reason to think that the microscopic degrees of freedom of even an isolated system suffice to destroy coherence between macroscopic superpositions of that systems macroscopic degrees of freedom.[12]The upshot, in either case, is that for systems above quite small lengthscales, coherent superpositions of states with macroscopically distinct positions rapidly become entangled with their environment. Conversely, though, if a macroscopic system

[12]For a concrete model, consider a solid-state system — a crystal, say — which is approximately but not exactly harmonic. The macroscopic degrees of freedom of the system correspond to the long-wavelength phonons; these will be decohered by scattering off the short-wavelength phonons in qualitatively the same way that massive particles are decohered by scattering off light particles. (Systems like this will also, in general, behave quasi-classically even absent the anharmonic terms, for the reasons explained in sections 3.1–3.2: they are regular. See Halliwell (1998, 2010) for a detailed analysis.)

is prepared in a state highly localised in spatial position, very little entanglement will occur.

3.5 Environment-induced decoherence: further details

So far we have been ignoring the dynamics of the system itself. Qualitatively, though, it is easy to see — at least, for regular systems — how this dynamics will proceed. Systems prepared in superpositions of macroscopically different positions will decohere on timescales much more swift than their characteristic dynamical timescales. Systems prepared in superpositions of macroscopically different *momentums* will quickly evolve into states with macroscopically different positions, and these too will swiftly decohere. But if the system is prepared in a state which is approximately localised in both position and momentum, then this state will undergo very little decoherence, and will simply be able to evolve under the system's own Hamiltonian. Since we already know that that evolution takes localised states to localised states — again, for regular systems — then this evolution will continue to be unaffected by decoherence.

Purely phenomenologically it is fairly straightforward to write down dynamical equations for the density operator of a decohering system: the exponential decay in equation (3.35), in particular, is generated by the equation[13]

$$\dot{\rho} = -\Lambda[X,[X,\rho]], \tag{3.37}$$

which suggests the equation

$$\dot{\rho} = -i[H,\rho] - \Lambda[X,[X,\rho]]. \tag{3.38}$$

A microphysical derivation of such an equation would require a specific model for the environment, and a number of such models have been analysed. One of the most well-studied is the Caldeira-Leggett model[14] in which a particle interacts linearly with an environment of harmonic oscillators; under appropriate simplifying conditions[15], this model yields an equation of the form

$$\dot{\rho} = -i[\widehat{H} + \frac{1}{2}m\Omega^2\widehat{X}^2,\rho] - \eta k_B T\Lambda[\widehat{X},[\widehat{X},\rho]] - i\frac{\eta}{2m}[\widehat{X},\{\widehat{P},\rho\}] \tag{3.39}$$

[13]For further discussion of this expression see Joos *et al* (2003, pp. 64-75) and references therein.

[14]The Caldeira-Leggett model was first analysed in Caldeira and Leggett (1983); see Schlosshauer (2007, pp. 71–74) for a discussion.

[15]The "appropriate simplifying conditions" are a nice example of the way theoretical physics works in practice. One of the assumptions is that the system's internal dynamics are harmonic — that is, that the internal potential is quadratic — and this is clearly much too strong to rigorously justify applying the Caldeira-Leggett equation to, e. g. , chaotic systems. On the other hand, any

Equations derived from different environments have the same general form, consisting of:

1. The system's unitary dynamics (which, generically, will turn superpositions in momentum into superpositions in position via wave-packet spreading)
2. A decoherence term which suppresses superpositions in the position basis
3. A dissipation term (the last term in the Caldeira-Leggett equation) corresponding to classical friction
4. A renormalisation term (the term proportional to Ω^2 in the Caldeira-Leggett equation).

In situations of the sort discussed earlier — a macroscopic system interacting relatively weakly with a microscopic environment — the dissipation and renormalisation terms are negligible compared with the other two terms, and the decoherence term suppresses macroscopic superpositions very quickly relative to the dynamical timescale of the unitary term.

In the Wigner-function representation, and ignoring renormalisation and dissipation, the Caldeira-Leggett equation (and, as noted, most realistic equations for decoherent systems) takes the form [p. 304](Zurek and Paz 1995a)

$$\dot{W} = \{H, W\}_{MB} + \Lambda \frac{\partial^2 W}{\partial p^2}. \tag{3.40}$$

It can readily be seen that the decoherence term is a diffusion term, which will cause W to spread out as long as it is sufficiently localised in momentum. For regular systems, this term will normally be negligible for quasi-classical states: the spread of such states in momentum space is such that the diffusion term is almost irrelevant.

Things are interestingly different for chaotic systems. Recall that for such systems, the fact that the system begins in a phase-space-localised state is insufficient to ensure that it remains in such a state. Instead, a state initially localised will

potential is approximately quadratic as long as we remain confined to a sufficiently small region of it. So, provided we are entitled to assume that the system is never in a coherent superposition which is large compared with the lengthscales on which the potential deviates from quadradicity, we can derive the equation on the basis of a quadratic potential. And what justifies *this* assumption? Earlier, qualitative arguments, of the form described above. The self-consistency of the whole thing can be seen when it is noted that Caldeira-Leggett dynamics do indeed suppress coherent superpositions on the required lengthscale. Philosophers of science take note: theoretical physics does this sort of thing all the time, and naturalistically inclined philosophers should be fine with this.

begin to spread out — and, as soon as it starts to spread out, the diffusion term will come into play (i. e. , the state will start to become entangled with its environment), so that the pure delocalised state becomes replaced by a mixed state which is an incoherent superposition of localised states. Each of *these* states will spread out under the chaotic dynamics, and so will be decohered in their turn ... and so on. At any given time, the density operator of the system will be a weighted sum of localised states, and because of the constant decoherence, each such state will evolve independently of all the others, even though it is constantly splitting into multiple states. So in the case of chaos, "worlds" — that is, emergent quasi-classical systems — are constantly splitting from one another.

Notice that the irreversibility induced by decoherence is of a very different character from that which would be induced by the dissipative term: there is no energy loss, no deviation from isolated classical dynamics on long lengthscales, and the process can occur — and occur extremely quickly — in cases where dissipation is negligible. (Consider Jupiter, for instance: the interstellar medium decoheres Jupiter essentially instantly, but friction between the medium and Jupiter is dynamically utterly irrelevant.) Nonetheless, decoherence *is* an irreversible process, and so the usual questions arise as to how this is compatible with an underlying reversible dynamics. I address this question in chapter 9 of Wallace (2010c); see Schlosshauer (2007, pp93–95) for more on the contrast between decoherence and dissipation.

3.6 Decoherent histories

Let us take stock. In section 3.3 I identified three problems with extracting quasi-classical behaviour from macroscopic quantum systems: (i) what justifies our treating the macroscopic degrees of freedom as dynamically isolated from the remainder of the system; (ii) why chaotic systems behave quasi-classically given that in isolation they evolve into non-quasi-classical states; (iii) why even when the dynamics of a system is not even approximately classical — such as in the case of quantum measurement — macroscopic systems still seem to stay in quasi-classical states

We can now see that decoherence provides an answer to all three worries. Firstly, it explains why, for the macroscopic degrees of freedom of regular systems, we are justified in ignoring the effects of the environment: the main effect of the environment is to measure the system in the position basis, and this has no effect on the system if it is already in a reasonably localised state.

Secondly, it explains how chaotic systems nonetheless evolve in a classical way, at least at the coarse-grained level: decoherence constantly transforms delocalised states into mixtures of localised states, and so prevents the system ever ending up

in a state so delocalised that the dynamics ceases to be approximately classical.

And as for non-classical events like quantum measurement: whatever state they put a system into, if that system's macroscopic degrees of freedom are not fairly localised in position then it will very rapidly become decohered: as such, it will evolve as a collection of non-interacting systems each of which is itself fairly localised in position.

Furthermore, decoherence at least helps to explain why it seems to be only phase-space local states which can instantiate emergent structure. For suppose some state like

$$\alpha \left|q_1, p_1\right\rangle + \beta \left|q_2, p_2\right\rangle \tag{3.41}$$

is supposed to instantiate a state of some emergent theory. Decoherence will wipe away any information contained in the relative phases: the system will almost immediately move into the mixed state

$$\left|\alpha\right|^2 \left|q_1, p_1\right\rangle \left\langle q_1, p_1\right| + \left|\beta\right|^2 \left|q_2, p_2\right\rangle \left\langle q_2, p_2\right| \tag{3.42}$$

which is simply a weighted sum of two independently evolving quasi-classical states. So the complete dynamical story of the system is known once we know its quasi-classical dynamics and the relative weights of the quasi-classical histories.

However, our analysis so far — which has been concentrated on the evolution of the system's density operator, and has invariably traced away the environment — makes it somewhat difficult to appreciate how exactly it is that the quantum state has the structure of a collection of quasi-classical *branching* worlds. We may have established that the density operator of such systems is diagonalised in a quasi-classical basis, but it is not immediately obvious how to read the branching structure off from this observation.

An example may help to see the difficulty — and to see how to surmount it. The orbit of the Earth around the Sun is chaotic: over timescales of a few million years it is impossible (using classical physics) to predict where in its orbit the planet may be found.[16] The earth is also (obviously!) very strongly decohered by its environment. The general considerations of section 3.5 tell us that the system's density operator will evolve, over the same timescales, to be a uniform mixture of states localised at all locations in the orbit, and will thereafter remain in that state indefinitely. That is: if $\left|\theta\right\rangle$ is a state of the Earth localised at a particular angular coordinate θ, after a few million years the Earth (or at least, its centre-of-mass degrees of freedom) will have state

$$\rho(t) = \int_0^{2\pi} \mathrm{d}\theta \ \left|\theta\right\rangle \left\langle\theta\right| . \tag{3.43}$$

[16] This example is discussed in detail in Zurek and Paz (1995a).

This stationary state does not look much like what the Everett interpretation predicts: a set of histories of the Earth's orbital position, each one evolving quasi-classically. Nor does it seem to match our own observations of the Earth as in motion.

However, this is an illusion caused by our failure to look at the overall state of the Earth-plus-environment system. The actual structure of the this state would be best written as

$$|\Psi(t)\rangle = \int \mathcal{D}\theta \Lambda[\theta]\, |\theta(t)\rangle \otimes |[\theta]\rangle\,, \tag{3.44}$$

where the integral is over all histories $\theta(\xi)$ of the angular coordinate of the Earth, and where states $|[\theta]\rangle$, $|[\theta']\rangle$ of the environment are orthogonal if $\theta(\xi)$ and $\theta'(\xi)$ differ significantly for any significant period of time. Each $|[\theta]\rangle$, in other words, encodes a different history of the Earth's location, and this is as we should expect: the position of the Earth at any time leaves an irreversible record in the pattern of light, gravitational waves, and neutrinos radiating outwards from the Solar system at that time. So despite the apparent stationarity of (3.43), actually the system is a superposition of quasi-classical states, each of which is evolving approximately classically but which is branching into multiple approximately-classical states on a long timescale.

For the rest of this chapter, I wish to explore the structure of the quantum state from this more "historical" perspective. I will begin by getting a little more precise about what it is to say that a system's state is "branching".

3.7 Analysing branching structure

What would it mean to say that a quantum state "has a branching structure"? Firstly, clearly that branching structure would have to be defined by the state *together with* other dynamical structures in the theory: a state, interpreted as a mere vector in a featureless Hilbert space, has no structure at all. Relative to a basis, on the other hand, it is comparatively clear to understand how a state could be branching: if the state evolves from a basis vector to a superposition of such basis vectors, and if each of *those* evolves into a superposition of *different* basis vectors so that no two such superpositions interfere with one another — then we would have branching (relative to that basis, at any rate).

To get rather more precise about this, suppose we have a physical system represented by some Hilbert space $\mathcal{H}$, evolving unitarily under some dynamics $\widehat{U}(t, t_0)$. Instead of restricting ourselves to a basis, we will consider a PVM $\widehat{P}_1, \ldots \widehat{P}_n$ (that is, a family of disjoint projectors whose sum is the identity but which need not be all of dimension one). At any given time (t), and for an initial state $|\psi\rangle$ (at time

t_0), the weight of projector $\widehat{P_j}$ is

$$\mathcal{W}_j(t) = \| \widehat{P_j}\, \widehat{U}(t,t_0)\, |\psi\rangle \, \|^2 \equiv \langle \psi |\, \widehat{U}^\dagger(t,t_0) \widehat{P_j}\, \widehat{U}(t,t_0)\, |\psi\rangle \,, \tag{3.45}$$

and the transition weight between $\widehat{P}_j$ at time t and $\widehat{P}_{j'}$ at time t' is

$$\mathcal{T}(j,t;j',t') = \frac{\| \widehat{P_{j'}} \widehat{U}(t',t) \widehat{P_j} \widehat{U}(t,t_0)\, |\psi\rangle \, \|^2}{\| \widehat{P_j} \widehat{U}(t,t_0)\, |\psi\rangle \, \|^2}$$

$$= \frac{\langle \psi |\, \widehat{U}^\dagger(t,t_0) \widehat{P_j} \widehat{U}^\dagger(t',t) \widehat{P_{j'}} \widehat{U}(t',t) \widehat{P_j} \widehat{U}(t,t_0)\, |\psi\rangle}{\langle \psi |\, \widehat{U}^\dagger(t,t_0) \widehat{P_j} \widehat{U}(t,t_0)\, |\psi\rangle}. \tag{3.46}$$

For convenience, define $\mathcal{T}(j,t;j',t') = 0$ whenever $\mathcal{W}_j(t) = 0$ (the above definition leaves it undefined).

When quantum mechanics is interpreted instrumentally, of course, the transition weights are supposed to be conditional probabilities and the absolute weights are supposed to be unconditional probabilities; in quantum mechanics interpreted realistically, though, they are just objective properties of the quantum-mechanical Universe.

As we have noted, "branching" (relative to a given basis) is just the absence of interference. This in turn occurs (between times t and t') when at most one component of the quantum state (in that basis) at time t contributes to the weight of any given component at time t'. In terms of transition weights, this is just to require that no two transition weights of transitions into a given projector are nonzero — that is, to require that

$$\mathcal{T}(j_1,t;j',t') \neq 0, \mathcal{T}(j_2,t;j',t') \neq 0 \Longrightarrow j_1 = j_2. \tag{3.47}$$

(To visualise this, think of "weight" as a fluid, redistributing itself across the projectors over time. (3.47) guarantees that each projector receives weight from exactly one previous projector. Less picturesquely, if (3.47) holds then there is a unique way to connect projectors at later times to projectors at earlier times: each projector's weight may determine the weight of many future projectors but its own weight is determined by exactly one past projector at any given past time.)

The importance of decoherence is: when it occurs, quantum-mechanical systems (approximately) develop a particularly natural branching structure. For decoherence is a process which constantly, and (on sub-Poincaré-recurrent timescales) irreversibly entangles the environment with the system so as to suppress interference between terms of the decoherence-preferred basis. (We might say that the environment constantly measures the system and records the result). If we idealise the dynamics as discrete, then at each branching event, the environment permanently records the

pre-branching state, so that at each time the universal state is a superposition of states each of which encodes a complete record of where "its weight" comes from.

Even if the dynamics is not itself discrete, a branching structure is still readily discernible in decohering systems. In the case of phase-space decoherence, in particular, we can in full generality write the total state of the system and environment at a given time as

$$|\Psi\rangle = \int \mathrm{d}p_0\, \mathrm{d}q_0\, \alpha(p_0, q_0)\, |p_0, q_0\rangle \otimes |\phi(p_0, q_0)\rangle \tag{3.48}$$

Because of decoherence, whatever initial state the system is prepared in, the total state will quickly evolve to one where $\langle \phi(p_0, q_0) | \phi(p_0', q_0')\rangle \simeq 0$ for sufficiently separated q', p' and q, p.

After some further time Δt, the state

$$|p_0, q_0\rangle \otimes |\phi(p_0, q_0)\rangle \tag{3.49}$$

will evolve to a state of form

$$|\psi(p_0, q_0)\rangle = \int \mathrm{d}p_{\,1} \mathrm{d}q_{\,1} \beta_1(p_1, q_1; p_0, q_0)\, |p_1, q_1\rangle \otimes |\phi(p_1, q_1, p_0, q_0)\rangle\,. \tag{3.50}$$

Again, decoherence ensures that $\langle \phi(p_1, q_1, p_0, q_0) | \phi(p_1', q_1', p_0, q_0)\rangle \simeq 0$ for sufficiently separated q_1', p_1' and q_1, p_1. But we would also expect, in general, to find that if $\langle \phi(p_0, q_0) | \phi(p_0', q_0')\rangle \simeq 0$, then $\langle \phi(p_1, q_1, p_0, q_0) | \phi(p_1', q_1', p_0', q_0')\rangle \simeq 0$ irrespective of the values of p_1, q_1, p_1', q_1'. For the information about the system recorded in the original decoherence process will be distributed very widely across the environment (think of our original example of decoherence by particle scattering: the initial particles that caused the decoherence are now a distance $\sim v\Delta t$ from the system). The total state at after time Δt is then

$$\begin{aligned} |\Psi(\Delta t)\rangle &\equiv \widehat{U}(\Delta t)\, |\Psi\rangle \\ &= \int\int \mathrm{d}p_0\, \mathrm{d}q_0\, \mathrm{d}p_1\, \mathrm{d}q_1\, \beta_1(p_1, q_1; p_0, q_0) \alpha(p_0, q_0)\, |p_0, q_0\rangle \otimes |\phi(p_1, q_1, p_0, q_0)\rangle \end{aligned} \tag{3.51}$$

Iterating, then (and writing $\mathbf{p}$, $\mathbf{q}$ to symbolise the N-tuples $p_0, \ldots p_N$, $q_0, \ldots q_N$), after a time $N\Delta t$ the system will have state

$$|\Psi(N\Delta t)\rangle \equiv \widehat{U}(N\Delta t)\, |\Psi\rangle = \int \cdots \int \mathrm{d}\mathbf{p}\, \mathrm{d}\mathbf{q}\, C_N(\mathbf{p}, \mathbf{q})\, |p_N, q_N\rangle \otimes |\phi_N(\mathbf{p}, \mathbf{q})\rangle\,, \tag{3.52}$$

where $\langle \phi_N(\mathbf{p}, \mathbf{q}) | \phi_N(\mathbf{p}', \mathbf{q}')\rangle \simeq 0$ if any of the (q_i, p_i) are sufficiently separated from the (q_i', p_i'). (For the example described by (3.35), for instance, this amounts to

requiring that the position-space width of the cell is much larger than $(\Lambda(t_{i+1} - t_i)^{-1/2}$.)Each dynamical step can be represented by

$$\widehat{U}(\Delta t)\,|p_N, q_N\rangle \otimes |\phi_N(\mathbf{p},\mathbf{q})\rangle$$

$$= \int \mathrm{d}p_{\,N+1}\mathrm{d}q_{\,N+1} B_N(q_{N+1}, p_{N+1}; \mathbf{q}, \mathbf{p})\,|p_{N+1}, q_{N+1}\rangle \otimes |\phi_{N+1}(\mathbf{p}\oplus p_{N+1}, \mathbf{q}\oplus q_{N+1})\rangle \quad (3.53)$$

where $\mathbf{q}\otimes q$ is the sequence obtained by appending q to the sequence $\mathbf{q}$ (and similarly for $\mathbf{p}\otimes p$).

Informally, it should be clear that a state whose dynamics take this form will have a branching structure relative to the basis of $|\mathbf{p},\mathbf{q}\rangle$ states at each time-step. To make this more rigorous, though, let us choose a partition Σ_i of phase space, and define the operators

$$\widehat{\Pi}^N_{\mathbf{i}} = \int_{\Sigma_{i_0}} \cdots \int_{\Sigma_{i_n}} \mathrm{d}\mathbf{q}\,\mathrm{d}\mathbf{p}\,\widehat{1} \otimes |\phi_N(\mathbf{p},\mathbf{q})\rangle\,\langle\phi_N(\mathbf{p},\mathbf{q})|\,. \quad (3.54)$$

If the cells of the partition are chosen to be sufficiently large (in the case described by (3.35), for instance, if they have spatial width $\gg (\Lambda\Delta t)^{-1/2}$ and an appropriate momentum-space width) then these operators will approximately define a PVM:

$$\widehat{\Pi}^N_{\mathbf{i}}\widehat{\Pi}^N_{\mathbf{j}} \simeq \delta_{\mathbf{i},\mathbf{j}}\widehat{\Pi}^N_{\mathbf{i}}. \quad (3.55)$$

Moreover, we have

$$\widehat{\Pi}^N_{\mathbf{i}}\widehat{U}(N\Delta t)\,|\Psi\rangle = \int_{\Sigma_{i_0}} \cdots \int_{\Sigma_{i_N}} \mathrm{d}\mathbf{p}\,\mathrm{d}\mathbf{q}\,C_N(\mathbf{p},\mathbf{q})\,|p_N, q_N\rangle \otimes |\phi_N(\mathbf{p},\mathbf{q})\rangle \quad (3.56)$$

and from this and (3.53) it can readily be seen that

$$\widehat{\Pi}^{N+1}_{\mathbf{i}'}\widehat{U}(\Delta t)\widehat{\Pi}^N_{\mathbf{i}}\widehat{U}(N\Delta t)\,|\Psi\rangle \simeq 0 \text{ unless } \mathbf{i} \text{ is the initial segment of } \mathbf{i}'. \quad (3.57)$$

That is: the structure of the quantum state relative to the family of PVMs $\{\widehat{\Pi}^N_{\mathbf{i}}\}$ (for each N) is branching.

Notice that although we have imposed a discrete structure on the system so as to make precise the claim that it branches, there is no intrinsic discreteness in the branching process. Less rigorously, but perhaps more perspicuously, we might rewrite (3.52) as

$$|\Psi(t))\rangle = \int \mathbf{D}[q(\xi)]\;C_t[q(\xi)]\;\;|p(t), q(t)\rangle \otimes |\phi[q(\xi)]\rangle \quad (3.58)$$

where the integral ranges over all classical trajectories defined up to time t and where $\langle\phi[q(\xi)]|\phi[q'(\xi)]\rangle \simeq 0$ if the trajectories $q(\xi)$ and $q'(\xi)$ are sufficiently different

for sufficiently long (if they differ by $\gg \Lambda\delta t)^{-1/2}$ over a period of $\sim \delta t$ in the case of (3.35), for instance). In this formalism, the state has branching structure because $|p(t), q(t)\rangle \otimes |\phi[q(\xi)]\rangle$ evolves over time Δt to

$$\int \mathbf{D}[q'(\xi)]\ B_{t,t+\Delta t}[q'(\xi)]\ \ |p(t), q(t)\rangle \otimes |\phi[q(\xi) \oplus q'(\xi)]\rangle \tag{3.59}$$

where the integral ranges over classical trajectories defined between times t and $t + \Delta t$, and where $q(\xi) \oplus q'(\xi)$ is the trajectory given by $q(\xi)$ up till $\xi = t$ and by $q'(\xi)$ thereafter.

3.8 The decoherent-histories framework

To talk more generally about the relation between branching and decoherence, and to help the reader to connect my discussion to the literature, it will be useful to develop a more sophisticated mathematical description of branching. We will consider a discrete set of times $t_0, \ldots t_n$, and will generalise our earlier description by allowing the PVMs used to define branching to vary from time to time; we will also (purely for mathematical convenience) switch to the Heisenberg picture. Then the spaces on which the branching structure is defined is just a time-indexed family of PVMs $\widehat{P}^i_j$ (with the superscript indicating that the operator is a member of the time-t_i PVM and the subscript indexing it within that PVM), and the transition weights are given by

$$\mathcal{T}(j, t_i; j'; t_{i'}) = \frac{\langle\psi|\, \widehat{P}^i_j \widehat{P}^{i'}_{j'} \widehat{P}^i_j\, |\psi\rangle}{\langle\psi|\, \widehat{P}^i_j\, |\psi\rangle}. \tag{3.60}$$

The branching criterion can then be succinctly expressed as: if $\widehat{P}^{i'}_{j'} \widehat{P}^i_{j_1} |\psi\rangle$ and $\widehat{P}^{i'}_{j'} \widehat{P}^i_{j_2}$ are both non-zero, then $j_1 = j_2$.

It is again useful to define a *history* as a sequence of projectors, one from each of the time-indexed PVMs: I call the set of such histories generated from some such sequence of PVMs a *history space*. Since a sequence of projectors can also be viewed as a function from histories to projectors, given a history α I write $\widehat{\alpha}(m)$ for the projector associated with time index m; each $\widehat{\alpha}(m)$ is specified uniquely by giving its index number in the time-t_m PVM, and I write this index number as α_m, so that

$$\widehat{\alpha}(m) = \widehat{P}^m_{\alpha_m}. \tag{3.61}$$

I call a history *realised* if $\mathcal{T}(\alpha_m, t_m; \alpha_{m+1}, t_{m+1}) \neq 0$ for all $m \leq n$. The branching criterion then guarantees that if two realised histories coincide at some time (that is, assign the same projector to that time) then they coincide at all earlier times, and

we will say that any set of histories with this property has a *branching structure*. Given two history spaces $\{\mathcal{P}^i\}$, $\{\mathcal{Q}^i\}$, $\{\mathcal{Q}^i\}$ is a *coarse-graining* of $\{\mathcal{P}^i\}$ if every projector in $\mathcal{Q}^i$ is a sum of projectors in $\mathcal{P}^i$.

Following Gell-Mann and Hartle (1990), we can define the *history operator* $\widehat{C}_\alpha$ of the history α by

$$\widehat{C}_\alpha = \widehat{\alpha}(n) \cdots \widehat{\alpha}(0), \tag{3.62}$$

and the *decoherence functional*, a complex function on pairs of histories, by

$$\mathcal{D}(\alpha, \beta) = \langle\psi| \widehat{C}_\alpha^\dagger \widehat{C}_\beta |\psi\rangle . \tag{3.63}$$

A history space is said to satisfy the *decoherence condition* or to be *decoherent*[17] if the decoherence functional between any two incompatible histories is zero. (Hence, implicitly a history space is only decoherent relative to a choice of state vector.)

The significance of all this formalism is summarised in the following theorem (first stated by Griffiths (1993), so far as I know).

> **Branching-Decoherence Theorem:** If $\mathcal{P} = \{\widehat{P}^i_j\}$ is a history space and $|\psi\rangle$ is a quantum state, then
>
> (i) If $|\psi\rangle$ has branching structure (relative to $\mathcal{P}$) and α is a history then $\widehat{C}_\alpha |\psi\rangle \neq 0$ iff α is realised (with respect to $|\psi\rangle$.
>
> (ii) If the set Hist of all historiesα such that $\widehat{C}_\alpha |\psi\rangle \neq 0$ has branching structure (that is, if no two histories in Hist agree on their nth index but not on all previous indices), then $|\psi\rangle$ also has branching structure (relative to $\mathcal{P}$), and the realised histories in that branching structure are just the histories in Hist.
>
> (iii) If $|\psi\rangle$ has branching structure (relative to $\mathcal{P}$), $\mathcal{P}$ satisfies the decoherence condition.
>
> (iv) If $\mathcal{P}$ satisfies the decoherence condition, it is a coarse-graining of a (decoherent) history space relative to which $|\psi\rangle$ has branching structure.

The proof of the Branching-Decoherence Theorem is straightforward but tedious and is relegated to Appendix A; however, the basic ideas behind it are easy to understand. The first two parts is just an iteration of the branching condition to apply

[17] Sometimes this condition is called *medium* decoherence, following Gell-Mann and Hartle (1990) and in contrast to *weak decoherence*, defined in the next section.

to sequences of more than two projectors, and the third part follows straightforwardly from the first two. The key to understanding the fourth part is to notice that it implies that the states

$$|\alpha\rangle = \widehat{C}_\alpha |\psi\rangle \tag{3.64}$$

are orthonormal. These states can be thought of as "record states", each recording the structure of an entire branch. The state of the system at a given time, then, is a superposition of all these histories, and the subsequent evolution of the system will not erase these histories; hence, the terms in the superposition cannot interfere with one another, and so the state has a branching structure.

3.9 Decoherence, records, and consistency

From the Everettian perspective, the decoherence functional is a purely technical tool: its significance comes from the Branching-Decoherence theorem, which tells us that the vanishing of the decoherence function between any two distinct histories is a necessary and sufficient condition for a history space to have a branching structure. An alternative perspective, however — developed by Robert Griffiths (1984, 1996, 2002), Roland Omnés (1988, 1992, 1994), and (from a rather different viewpoint) by Murray Gell-Mann and James Hartle (1990, 1993, 2007) — was historically important and remains frequently discussed in the literature, and is the subject of this section. For clarity, I follow Griffiths in calling this approach a *consistent histories* approach, though actual terminology has been somewhat varied.

This framework starts with the idea that quantum mechanics ought somehow to be interpreted as a stochastic theory. Doing this consistently would require the theory to specify a space of histories and some probability measure over those histories. Within quantum mechanics, the obvious mathematical representation of a history is that of the previous section: a string of time-indexed projectors (note that for the moment I do not assume that a history is part of some previously specified history *space*) . And the obvious probability to assign to a history α is

$$\Pr(\alpha) = \| \widehat{\alpha}_n \cdots \widehat{\alpha}_1 |\psi\rangle \|^2 \tag{3.65}$$

(that is, start with the quantum state, sequentially project it out by the projectors, and take the mod-squared amplitude of the resulting state — in the Schrödinger picture it would also be necessary to evolve the state unitarily between sequential projections). Using the history operator $\widehat{C}_\alpha$ and decoheence functional $\mathcal{D}(\alpha, \beta)$ defined in the previous section, we can write this succinctly as

$$\Pr(\alpha) = \langle\psi| \widehat{C}_\alpha^\dagger \widehat{C}_\alpha |\psi\rangle = \mathcal{D}(\alpha, \alpha). \tag{3.66}$$

The problem, of course, is the same problem that besets all attempts to interpret quantum mechanics probabilistically: interference. In this case, the mathematical representation of interference is as a failure of the probability calculus. Suppose, for instance, that α and β are histories with $\widehat{\alpha}(k) = \widehat{\beta}(k)$ for all time indexes k except some m, and that $\widehat{\alpha}(m)$ and $\widehat{\beta}(m)$ are orthogonal. If γ is defined by

$$\widehat{\gamma}(k) = \widehat{\alpha}(k) = \widehat{\beta(k)} \ (k \neq m)$$
$$\widehat{\gamma}(m) = \widehat{\alpha}(m) + \widehat{\beta}(m) \tag{3.67}$$

then the probability calculus would require that $\Pr(\gamma) = \Pr(\alpha) + \Pr(\beta)$. But this, of course, is generally not the case.

In the consistent-histories approach, this is solved by restricting the set of allowed histories. The starting point here is the *history space* of section 3.8, which was defined (recall) as the set of histories generated from a particular time-indexed family of PVMs. To allow for histories which are sums of other histories (as in the above case), we now permit histories which assign to a time t_i a sum of projectors (rather than just a single projector) in the time-t_i PVM. A history which assigns only one projector in the appropriate PVM to each time is called *atomic*. (In fact, once we generalise history spaces in this way, the notion of atomic histories becomes dispensible, as I explain in box 3.2, but for expository purposes it is convenient to retain them.)

Given histories α and β, I call α a *subhistory* of β iff $\widehat{\alpha(k} \subset \widehat{\beta}(k)$[18] for all k. And $\mathrm{Dec}(\alpha)$, the *decomposition* of α, is then the set of all atomic histories that are subhistories of α: in effect (if a stochastic interpretation is required) the various elements of the decomposition of α are the various ways of filling in the details of a system's history which α itself leaves unspecified.

A succinct way of writing the condition required by the probability calculus is then that for any history α,

$$\Pr(\alpha) = \sum_{\alpha_i \in \mathrm{Dec}(\alpha)} \Pr(\alpha_i), \tag{3.68}$$

or in terms of the history formalism,

$$\langle\psi|\, \widehat{C}^{\dagger}_{\alpha} \widehat{C}_{\alpha}\, |\psi\rangle = \sum_{\alpha_i \in \mathrm{Dec}(\alpha)} \langle\psi|\, \widehat{C}^{\dagger}_{\alpha_i} \widehat{C}_{\alpha_i}\, |\psi\rangle\,. \tag{3.69}$$

We can say that a history space is *consistent* if this condition holds; it follows that in general, consistency is relative to the quantum state.

[18]Recall that given projectors P, Q, then $P \subset Q$ iff the range of P is a subspace of the range of Q.

Box 3.2: Atomless history spaces

Given a Hilbert space, a Boolean algebra of projectors on that Hilbert space is just a set of projectors which contains the identity and is closed under taking countable sums and complements; such an algebra is *atomic* if there is a countable set of projectors such that all elements of the algebra are sums of elements of the set. I specify a *history algebra* $\{\mathcal{S}^i\}$ by assigning to each time index t_i a Boolean algebra $\mathcal{S}^i$ of projectors; the histories in that algebra are sequences of such projectors, and I call the history atomic iff all its Boolean algebras are atomic.The history operator and the decoherence functional can be defined as before; the probability of history α is by definition $\mathcal{D}(\alpha, \alpha)$.

Two histories α, β are *overlapping* if for each k, $\widehat{\alpha}(k)\widehat{\beta}(k) \neq 0$. Given a history α in $\{\mathcal{S}^i\}$, a *decomposition* of α is a set of histories specified by giving, for each k, a set of mutually orthogonal projectors $\widehat{P}_i^k \in \mathcal{S}^k$ whose sum is $\alpha(k)$; the histories in the refinement are exactly those histories constructed from projectors in this set.

A history *space* continues to be specified by a time-indexed sequence of sets of projectors; each history space determines an atomic history algebra in the obvious way, and conversely a history space is *contained within* a history algebra if all its histories are histories in the algebra. Given a history algebra, and two history spaces contained within it, the first is a *refinement* of the second iff each projector in each time-t_k projector set in the second space is the sum of projectors in the time-t_k projector set in the first space. (It follows that a history algebra is atomic iff it contains some history space with no proper refinements.)

We can then make the following definitions. Given a history algebra, then with respect to some state $|\psi\rangle$:

- the algebra is *branching* if it contains some history space relative to which $|\psi\rangle$ has branching structure.
- the algebra satisfies *decoherence* iff $\mathcal{D}(\alpha, \beta)$ vanishes whenever α, β are non-overlapping, and *weak decoherence* if the real part of $\mathcal{D}(\alpha, \beta)$ vanishes for non-overlapping α, β.
- the algebra is *consistent* iff for any history α, and any decomposition of that history, the probability of α is the sum of the probabilities of the histories in its decomposition.

It then follows that:

1. A history algebra is decoherent iff it is branching (atomless version of the Branching-Decoherence Theorem)
2. A history algebra is weakly decoherent iff it is consistent

The former is proved in appendix A; the latter is proved by the method used in section 3.9.

Now, since

$$\widehat{C}_\alpha = \sum_{\alpha_i \in \mathrm{Dec}(\alpha)} \widehat{C}_{\alpha_i} \tag{3.70}$$

we can rewrite the left hand side of (3.69) as

$$\langle\psi|\, \widehat{C}_\alpha^\dagger \widehat{C}_\alpha \,|\psi\rangle = \sum_{\alpha_i, \alpha_j \in \mathrm{Dec}(\alpha)} \langle\psi|\, \widehat{C}_{\alpha_j}^\dagger \widehat{C}_{\alpha_i} \,|\psi\rangle = \sum_{\alpha_i, \alpha_j \in \mathrm{Dec}(\alpha)} \mathcal{D}(\alpha_i, \alpha_j) \tag{3.71}$$

and the right hand side as

$$\sum_{\alpha_i \in \mathrm{Dec}(\alpha)} \mathcal{D}(\alpha_i, \alpha_i). \tag{3.72}$$

It follows that any history space which is decoherent — that is, which satisfies $\mathcal{D}(\alpha, \beta) = 0$ for $\alpha \neq \beta$ — is also consistent. Because $\mathcal{D}(\alpha, \beta) = \mathcal{D}(\beta, \alpha)^*$, a slightly weaker condition — that the real part of $\mathcal{D}(\alpha, \beta)$ vanishes for $\alpha \neq \beta$, suffices to guarantee that a history space is consistent; for this reason, Griffiths calls this condition *consistency*; Gell-Mann and Hartle call it *weak decoherence*. However, weak decoherence does not seem to have any dynamical significance (in the way that decoherence proper has been shown to have) and composite systems satisfying weak but not full decoherence have been shown to have various unsatisfactory properties (Diósi 2004). By the branching-decoherence theorem, it follows that any branching history space is consistent and that physically interesting consistent history spaces are coarse-grainings of branching history spaces.

Originally, it was possible to suppose that consistency, or decoherence, or some reasonable strengthening of these conditions, would suffice to pick out a *unique* history space; the measurement problem would thereby have been solved and quantum mechanics could have been interpreted as a stochastic theory. Unfortunately for the consistent-histories program, this turns out not to be the case: Fay Dowker and Adrian Kent demonstrated convincingly (Dowker and Kent 1996; Kent 1996) that an enormous number of consistent history spaces and that many of them are pathologically unlike the observed macroworld.

The responses[19] of Griffiths, Omnes, and Gell-Mann and Hartle to this problem differ interestingly. Griffiths and Omnes attempt to hold on to the idea of quantum mechanics as a stochastic theory of a single quasi-classical world, and in doing so end up advocating interpretations of quantum mechanics that offer "vestiges of reality" as I put it in section 1.6, but fall short of conventional scientific realism. Griffiths (2002), for instance, tries to regard different history spaces as different ways of describing the same underlying reality. But while in classical mechanics

[19] I do not want to make any *historical* claim here as to the influence or otherwise of Dowker and Kent's work on proponents of consistent-histories approaches: my account is intended to capture the *logic* of the situation, rather than its chronology.

such multiple descriptions can always be understood as coarse-grainings of a single exhaustive description (a principle which Griffiths dubs the *principle of unicity*), this fails in the consistent-histories setting:

> The principle of unicity does not hold: there is not a unique exhaustive description of a physical system or a physical process. Instead, reality is such that it can be described in various alternative, incompatible ways, using descriptions which cannot be combined or compared.

Approaches of this kind, of course, fall outside the scope of this thesis.

Gell-Mann and Hartle, on the other hand, rule out pathological history spaces by requiring histories to be "quasi-classical", which they define (consistently with my usage in this chapter) as histories

> such that the individual histories obey, with high probability, effective classical equations of motion interrupted continually by small fluctuations and occasionally by large ones.

This is not the kind of criterion which can be formalised as a new law of physics: it is a criterion for emergent structure of very much the same kind as I discussed in chapter 2. Gell-Mann and Hartle's exploration of consistent histories, in other words, can be understood as an exploration of those emergent structures which exist within the unitarily evolving state: that is, it can be understood as an exploration of Everettian quantum mechanics. (And indeed, this is how Hartle, at least, does understand it; see Hartle (2010)).

3.10 How many worlds?

We are finally in a position to answer one of the most commonly asked questions about the Everett interpretation,[20] namely: how much branching actually happens? As we have seen, branching is caused by any process which magnifies microscopic superpositions up to the level where decoherence kicks in, and there are basically three such processes:

1. Deliberate human experiments: Schrödinger's cat, the two-slit experiment, Geiger counters, and the like.

2. "Natural quantum measurements", such as occur when radiation causes cell mutation.

[20] Other than "and you believe this stuff?!", that is.

Box 3.3: A metaphor for indefinite branch number

1. Firstly, imagine a world consisting of a very thin, infinitely long and wide, slab of matter, in which various complex internal processes are occurring — up to and including the presence of intelligent life, if you like. In particular one might imagine various forces acting in the plane of the slab, between one part and another.

2. Now, imagine stacking many thousands of these slabs one atop the other, but without allowing them to interact at all. If this is a "many-worlds theory", it is a many-worlds theory only in the sense of the philosopher David Lewis (Lewis 1986a): none of the worlds are dynamically in contact, and no (putative) inhabitant of any world can gain empirical evidence about any other.

3. Now introduce a weak force normal to the plane of the slabs — a force with an effective range of 2-3 slabs, perhaps, and a force which is usually very small compared to the intra-slab force. Then other slabs will be detectable from within a slab but will not normally have much effect on events within a slab. If this is a many-worlds theory, it is a science-fiction-style many-worlds theory (or maybe a Phillip Pullman or C.S. Lewis many-worlds theory): there are many worlds, but each world has its own distinct identity.

4. Finally, turn up the interaction sharply: let it have an effective range of several thousand slabs, and let it be comparable in strength (over that range) with characteristic short-range interaction strengths within a slab. Now, dynamical processes will not be confined to a slab but will spread over hundreds of adjacent slabs; indeed, *evolutionary* processes will not be confined to a slab, so living creatures in this universe will exist spread over many slabs. At this point, the boundary between slabs becomes epiphenomenal. Nonetheless, this theory is *stratified* in an important sense: dynamics still occurs predominantly along the horizontal axis and events hundreds of thousands of slabs away from a given slab are dynamically irrelevant to that slab.[a] One might well, in studying such a system, divide it into layers thick relative to the range of the inter-slab force — and emergent dynamical processes in those layers would be no less real just because the exact choice of layering is arbitrary.

[a]Obviously there would be ways of constructing the dynamics so that this was not the case: if signals could easily propagate vertically, for instance, the stratification would be lost. But it's only a thought experiment, so we can construct the dynamics how we like.

3. Classically chaotic processes.

The first is a relatively recent and rare phenomenon, but the other two are ubiquitous. Chaos, in particular, is everywhere, and where there is chaos, there is branching (the weather, for instance, is chaotic, so there will be different weather in different branches). Furthermore, there is no sense in which these phenomena lead to a naturally *discrete* branching process: as we have seen in studying quantum chaos, while a branching structure can be discerned in such systems it has no natural "grain". To be sure, by choosing a certain discretisation of (configuration-))space and time, a discrete branching structure will emerge, but a finer or coarser choice would also give branching. And there is no "finest" choice of branching structure: as we fine-grain our decoherent history space, we will eventually reach a point where interference between branches ceases to be negligible, but there is no precise point where this occurs. As such, the question "how many branches are there?" does not, ultimately, make sense.

This may seem paradoxical — certainly, it is not the picture of "parallel universes" one obtains from science fiction. But as we have seen in chapter 2, it is commonplace in emergence for there to be some indeterminacy (recall: when *exactly* are quasi-particles of a certain kind present?) And nothing prevents us from making statements like:

> Tomorrow, the branches in which it is sunny will have combined weight 0.7

— the combined weight of all branches having a certain macroscopic property is very (albeit not precisely) well-defined. It is only if we ask: "*how many* branches are there in which it is sunny", that we end up asking a question which has no answer.

This bears repeating, as it will be central to some of the arguments of Part II:

> Decoherence causes the Universe to develop an emergent branching structure. The existence of this branching is a robust (albeit emergent) feature of reality; so is the mod-squared amplitude for any *macroscopically described* history. But there is *no* non-arbitrary decomposition of macroscopically-described histories into "finest grained" histories, and *no* non-arbitrary way of counting those histories.

(Or, put another way: asking how many worlds there are is like asking how many experiences you had yesterday, or how many regrets a repentant criminal has had. It makes perfect sense to say that you had many experiences or that he had many regrets; it makes perfect sense to list the most important categories of either; but it is a non-question to ask *how many*.)

If this picture of the world seems unintuitive, the metaphor in box 3.10 may help. Ultimately, though, that a theory of the world is "unintuitive" is no argument against it, provided it can be cleanly described in mathematical language.

CHAPTER 3: If we apply to quantum mechanics the same principles we apply right across science, we find that a multiplicity of quasi-classical worlds are emergent from the underlying quantum physics. These worlds are structures instantiated within the quantum state, but they are no less real for all that.

CHAPTER 4: Quantum mechanics is a probabilistic theory; how is this compatible with the Everett interpretation's deterministic dynamics?

Alistair I. M. Rae

Statement

and

Readings

Abstract

Alistair I. M. Rae

A central assumption of the conventional “Copenhagen” interpretation of quantum measurement is the “collapse” of the wavefunction, which is not predicted by the time-dependent Schrödinger equation. Alternative interpretations (notably “many worlds”) have been developed that claim to avoid this. I aim to develop an essentially simple, but I believe novel, argument to show that the quantum measurement process requires that some aspect of the measurement must be distinct from the quantum system being studied, so that no description in terms of the whole process in terms of Schrödinger evolution only is possible .

Seven Pines Symposium

Introductory notes to paper by Alastair Rae

Abstract

A central assumption of the conventional "Copenhagen" interpretation of quantum measurement is the "collapse" of the wavefunction, which is not predicted by the time-dependent Schrődinger equation. Alternative interpretations (notably "many worlds") have been developed that claim to avoid this. I aim to develop an essentially simple, but I believe novel, argument to show that the quantum measurement process requires that some aspect of the measurement must be distinct from the quantum system being studied, so that no description in terms of the whole process in terms of Schrődinger evolution only is possible .

I intend to initiate a discussion on the nature of quantum measurement. I shall develop an essentially simple, but I believe novel, argument showing that the measuring apparatus and observer must be in some way distinct from the quantum system being studied. This is set out briefly in the following notes; accompanying papers develop the argument further and include further criticisms of attempts to reconcile the predictions of the TDSE and the assignment of probabilities to branches.

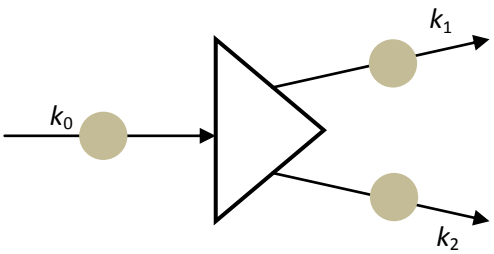
Consider an archetypical quantum measurement with two possible outcomes, such as the measurement of the spin component of a spin-half particle in a Stern-Gerlach experiment. Before the measurement, the wavefunction has the form

$$[\cos\theta|\alpha\rangle + \sin\theta|\beta\rangle]|k_0\rangle$$

where $|\alpha\rangle$ and $|\beta\rangle$ are eigenstates of the spin components parallel to the measurement direction of the apparatus; the angle between the incident spin and the measurement axis is 2θ; $|k_0\rangle$ describes a spatial wave packet travelling towards the measuring apparatus. After the measurement, the TDSE predicts that the state evolves to

$$\cos\theta|\alpha\rangle|k_1\rangle + \sin\theta|\beta\rangle|k_2\rangle$$

where $|k_1\rangle$ and $|k_2\rangle$ describe wave packets travelling away from the apparatus, in the directions k_1 and k_2 We assume that the size of the wavepacket is assumed to be much greater than the wavelength of the wave (so that spreading can be ignored) and significantly smaller than the dimensions of the experiment.



The measurement postulate (also known as the Born rule) states that after the completion of a measurement, the outcome will be $|\alpha\rangle|k_1\rangle$ with probability $\cos^2\theta$ **or** $|\beta\rangle|k_2\rangle$ with probability $\sin^2\theta$. The inconsistency between this and the predictions of the TDSE constitutes the quantum measurement problem.

Discussions of the measurement problem often stress the importance of the actual measurement process whereby the outcome is recorded on an apparatus whose final states, along with those of the associated environment, can be represented by, say, $|A_1\rangle$ and $|A_2\rangle$, so that the TDSE prediction for the whole set up becomes

$$\cos\theta|\alpha\rangle|k_1\rangle|A_1\rangle + \sin\theta|\beta\rangle|k_2\rangle|A_2\rangle$$

with a corresponding density operator given by

$$\cos^2\theta|\alpha\rangle\langle\alpha||k_1\rangle|\langle k_1||A_1\rangle\langle A_1| + \sin^2\theta|\beta\rangle\langle\beta||k_2\rangle|\langle k_2||A_2\rangle\langle A_2|$$

$$+ 2\cos\theta\sin\theta|\alpha\rangle\langle\beta||k_1\rangle|\langle k_2||A_1\rangle\langle A_2|$$

This is an entangled state of the system and the apparatus. However, to demonstrate this, say by constructing an interference experiment, is impossible in practice because of decoherence. For all practical purposes at least, we can assume that "collapse" has occurred into one or other of the component states of the supervision. The "holy grail" of quantum measurement theory is to provide a consistent self-contained interpretation whereby this random collapse and the relative probabilities of the possible outcomes is a natural part of our theoretical description rather than something that has to be added "by hand".

Let us consider the role of the amplitudes $\cos\theta$ and $\sin\theta$, which determine the probabilities, a little more carefully. We first note that under the given conditions illustrated in the figure, the outgoing states are spatially separated so the product $|k_1\rangle|\langle k_2|$ vanishes along with the product terms in the density operator which are therefore equal to zero. This happens without the need to invoke environmental decoherence. In contrast, if an interference experiment is carried out, the two components of the superposition are brought together and the measurement probabilities are determined by the modulus squared of the wavefunction with the interference pattern being generated by these off-diagonal terms and decoherence is essential if these are to be ignored. The role of decoherence is therefore quite incidental if no interference because the output states are always spatially separated, but is crucial in ensuring that the off-diagonal terms of the (now partially traced) density operator are zero in situations where interference patterns would otherwise be formed. Once the density operator contains only diagonal terms, their magnitudes can be interpreted as the probabilities of obtaining the corresponding results and these agree with the Born rule and experiment. However, this implies an additional assumption that the system has collapsed" into one or other of the outcome states. This statistical collapse cannot be a direct consequence of the linear TDSE.

We further note that, if the output states remain spatially separated, the form of the state vectors $|\alpha\rangle|k_1\rangle|A_1\rangle$ and $|\beta\rangle|k_2\rangle|A_2\rangle$ are independent of θ. The question arises as to how the probabilities of the measurement results, which are recorded by the apparatus (possibly including an observer) and therefore embedded in $|A_1\rangle$ and $|A_2\rangle$ can be affected by the value of θ. If there is only the TDSE, there would appear to be no means whereby an observer interacting with such an experiment could ever gain any knowledge of θ by simply observing the measurement outcomes. This is the central point of my argument, which is probably more clearly illustrated if we consider an ensemble

consisting of a large number (N) of identical measurements. The initial state is assumed to be the same in each case. There are then 2^N possible outcomes and in ${}^NC_M$ of these, M positive and $N - M$ negative spins are detected. The Born rule states that the most probable outcome is when $M/N = \cos^2\theta$. For large N, results different from this are very unlikely, so an observer can reliably deduce the value of θ by obtaining M from such an observation, without directly observing how the experiment is set up. However, this directly contradicts the earlier result where it was shown that the state of the apparatus, including the observer is the same, whatever the value of θ.

One possible counter to this argument is that, if all branches resulting from decoherence exist and the observers associated with them know about the Born rule, then they can each make a prediction of the value of θ, based on the Born rule. Some observers' predictions will be correct and others (in general the great majority) will be wrong: which is which depends on the value of θ, but the observer's experience is the same, whatever the value of θ and they can only obtain this by a direct observation. It would then be irrational of them to expect that they could deduce a value of θ from their observations and they could not logically believe in the Born rule.

These arguments are consistent with the standard view of quantum measurement as an operation that is performed from "outside" the system, which has always been the conventional "Copenhagen" interpretation of quantum measurement. However, I believe that they are incompatible with alternative interpretations – in particular "many-worlds" theories – where everything is described by the TDSE, and measurement becomes a choice between co-existing "branches" that are defined by the experimental set-up and the action of decoherence and interaction with the environment. The implications of these ideas to other approaches to the quantum measurement problem will be discussed in my presentation and will hopefully form a part of our discussions.

Red Hats and Ancillae

1. The Everett interpretation claims that all the observed features of quantum measurement are contained in the Schrődinger equation applied to the wavefunction, extended (by implication at least) to include some aspects of quantum field theory.

2. The metaphysical implications associated with the many-world features of the Everett approach are radical in the extreme, but this is not a sufficient reason to reject the interpretation. However, the onus should be on its supporters to show that observed behaviour does supervene on the Schrodinger equation and that any additional assumptions made are consistent with it.

3. The Copenhagen interpretation is widely considered unsatisfactory for several reasons, but its predictions agree with experiment and can be used as a yardstick for assessing the validity of alternatives. In particular, if we consider the standard Stern-Gerlach (SG) measurement of the spin of a spin-half particle, we expect the outcome to be in one or other channel with probabilities defined by the Born rule. Whatever these probabilities "really" are, they are confirmed or otherwise by observations of outcome frequencies. The quantum field point mentioned above is covered in the SG context by assuming that the system is in a superposition of two Foch states, each of which corresponds to a particle emerging through one of the possible output channels.

4. The first challenge to the Everett interpretation is the preferred basis problem: why do the detected states correspond to one or other of the output channels and not a linear combination of them. I assume that this has been resolved by the work surrounding decoherence, where the system can be shown to be extremely well approximated by a density matrix (DM) that is diagonal in a representation defined by the two output states. We note two points in passing: (i) Ignoring very small, (or even zero) elements of the DM is an assumption that assigns some significance of the wavefunction amplitude, which also enters the Born rule; (ii) DM diagonalisation is initially achieved when the wave packets associated with the two Foch states emerging from SG magnet are spatially separated and these states are chosen as a basis of the representation of Hilbert space: provided detection takes place without allowing the states to interfere, decoherence acts to confirm this initial diagonalisation and ensures that it is effectively permanent.

5. The remaining challenge is to show that Born-rule probabilities also supervene from the Schrodinger equation. This result is often thought to contradict "common sense" and a more intuitive expectation appears to be that, when a system including a detector (and possibly an observer) branches, there would seem to be no reason for preferring any one outcome over another, and therefore every branch should be equally probable. This of course is inconsistent with the Born rule and with experiment which has motivated a search for arguments that would go beyond common sense and reconcile the assignment of Born probabilities with the existence of multiple branches.

6. The first point to be addressed is how branches are to be defined and counted. Leaving aside the actual quantum context for the moment, consider the following example, which is based on one originally proposed by David Wallace

   Suppose I am part of a system that splits into two branches (A and B) at some time t_1, following which the observer in branch A is given a red hat to wear. At time t_2, A splits into two branches A_1 and A_2 while B remains unsplit. What initial probability should I assign before the first split occurs to having a hat after both splits have been completed? Between t_1 and t_2 there will be two branches, in one of which I get the hat; so the branch-counting rule says that the probability of getting it is 1/2. However, after t_2 there will be three branches, in two of which I get the hat; so the branch-counting rule says that the probability of getting it is 2/3. Which is right, or does the probability change with time? In fact, standard probability theory implies that if the probability equals 1/2 between t_1 and t_2, it will continue to have this value after t_2.

   We may be led to conclude from this example, that estimating probabilities on the basis of the number of branches created is likely to lead to ambiguities in the values of the defined probabilities. However, we should note that it is relevant to situation where there are two successive branching events and we are considering the probability as estimated at time t_0 where the first branching takes place. Unsurprisingly, branching that occurs after I have recorded a result (obtained a red hat) is irrelevant to the calculation of this probability, which depends only on the number of branches at the first node. I also note in passing that the same result would hold if red hats were awarded after rather than before the second bifurcation, provided this occurred in both the resulting branches.

7. Suppose now that, instead of successive bifurcations there was trifurcation from a single node into three branches, in two of which I am given a red hat. The branch-counting probability of getting a red hat would now be 2/3. (This begs the question of what determines the size of the node and whether there are situations where this is ambiguous, but there are many classical cases, such as throwing a fair die, where more than two outcomes follow a single event). We see that, although the triplet of final states is the same after both processes, the branch-counting probabilities are different, depending on whether the history of the process was a trifurcation or two successive bifurcations.

8. I develop this point further by returning to the original example with the addition of a second observer (Alice) who interacts with the system. If this happens after the first and before the second, branching (i.e. for $t_1 < t < t_2$) and if she has equal expectation of ending up in one or other of the two branches, I should expect the likelihood of her seeing me with a red hat to be ½. However, if she interacts with the system after the second branching ($t > t_2$) and now has equal expectation of emerging in one of the three branches, her

probability of seeing me with a red hat will be increased to 2/3, which is the same as we would expect in the case of three branches emerging from a single node. We appear to have a potential paradox: I know that the probability of my obtaining a red hat is 1/2, but I also believe that the probability of Alice seeing me with a red hat is 2/3.

Let us compare the above with what we should expect in a "collapse" context where there is only one outcome at each branch point. We shall assume equal likelihoods of possible outcomes at any branching. For example: at t_1 I *either* emerge in branch A and then acquire a red hat *or* I emerge in branch B with no hat; while at t_2 I proceed along with my hat to *either* A_1 *or* A_2. Given this my red-hat prediction before t_1 will be 1/2 as before and we now consider what Alice will experience in this case. She can see one of three things: either I am in branch B with no hat on or I am in one of the other two branches wearing a hat. Moreover, if she repeats the observation many times, she can deduce (by Baysian updating or otherwise) that the probabilities of my being found in B (without a hat) or in one of the A branches (with a hat) are equal, so she can conclude that my red-hat probability is 1/2, agreeing with my original estimate.

9. The source of the apparent inconsistencies in the splitting scenario arises from the fact that splitting implies the simultaneous creation of several possible outcomes, while the probability calculus applies to mutually exclusive events. Effectively, treat the second branching event is one in which I, along with my hat, am cloned into two copies of myself. This increases the likelihood that the Alice will see me with my red hat and the probability calculus cannot be applied to this situation unless it is modified to take this cloning into account. The situation is very similar to the following. Suppose I take two coins and place them on the table one showing heads and one showing tails; if Alice picks one of them at random she can expect to see heads or tails with equal probability. Suppose I now replace the coin showing heads with ten more coins, each of which also has heads up. Alice will now have a ten-to-one probability of seeing heads rather than tails.

10. I now consider how the cloning model would apply if the pattern of splitting were more complex. There might be a huge number of branches at any one time and this number might also be subject to wild, unpredictable fluctuations. Decoherence is generally believed to produce such a scenario. The first point is that, the complexity of the pattern of branching *subsequent to* an observation can have no effect on the prior estimates of the probabilities of the outcomes of that observation. Secondly, if an observation is made *after* complex branching has occurred, the observer will be unable to predict probabilities because she does not know the instantaneous number of branches. However, this does not *ipso facto* imply that there is no fact of the matter concerning the number of branches at that point. If the observation is repeated, the number of branches will be different on each occasion, leading to an apparently chaotic, unpredictable pattern of events.

11. I now turn to the quantum measurement problem in the light of the above. Consider the archetypical example of a spin-half particle initially in a state $|\alpha_z,k_0\rangle$ passing through a SG magnet, this will produce the output state

$$\cos\theta|\alpha\, k_1\rangle + \sin\theta|\beta,k_2\rangle \qquad (1)$$

where α and β are the spin eigenstates corresponding to the orientation of the SG magnet which is at an angle $\theta/2$ to z; k_o, k_1 and k_2 represent the directions of motion of the wave packets associated with the spins. Now suppose that the state $|\beta,k_2\rangle$ is further split by a partial reflector into states defined by k_{21} and k_{22}. The total state vector is now a sum of three spatially separated terms:

$$\cos\theta\ |\alpha\ k_1\rangle\ +\ \sin\theta\cos\phi\ |\beta\ k_{21}\rangle\ +\ \sin\theta\sin\phi\ |\beta\ k_{22}\rangle \qquad (2)$$

12. Some special values of θ and ϕ are of interest. First, suppose that $\cos\theta = \sin\theta = \cos\phi = \sin\phi = 2^{-1/2}$. Expressions (1) and (2) are then

$$2^{-1/2}|\alpha\ k_1\rangle + 2^{-1/2}|\beta,k_2\rangle \qquad (3)$$

$$2^{-1/2}\ |\alpha\ k_1\rangle\ +\ \tfrac{1}{2}|\beta\ k_{21}\rangle\ +\ \tfrac{1}{2}|\beta\ k_{22}\rangle \qquad (4)$$

respectively. Now consider the case where $\cos\theta = 3^{-1/2}$, $\sin\theta = (2/3)^{1/2}$ and $\cos\phi = \sin\phi = 2^{-1/2}$. Expressions (1) and (2) are then

$$(1/3)^{1/2}\ |\alpha\ k_1\rangle + (2/3)^{1/2}\ |\beta,k_2\rangle \qquad (5)$$

$$(1/3)^{1/2}\ [|\alpha\ k_1\rangle\ +\ |\beta\ k_{21}\rangle\ +\ |\beta\ k_{22}\rangle\ ] \qquad (6)$$

That is, all three have the same amplitude in the final state.

13. If we apply the Born rule to (3) and (4), the probabilities of the system being in the states $|\alpha\rangle$ and $|\beta\rangle$ are ½, and ½ in both cases, which are the same as the probabilities in the original example discussed above, provided we identify obtaining a red hat with the act of detecting the spin state to be $|\beta\rangle$. Moreover, if we apply the Born rule to (5) and (6), we find that each output channel has the same probability and that the total probability of being in state $|\beta\rangle$ is 2/3, which is the same as predicted in the case of a single node with three outputs, discussed in paragraph 8.

14. In a "collapse" scenario, where only one output survives every splitting event, the Born probabilities follow all the normal rules of probability, including the updating rule, but I contend that this should require justification if Everett branching occurs. I now wish to re-examine the DSW proof of the Born rule to identify what assumptions are made or implied in the cases where collapse or branching is assumed.

15. The first step in the proof is the case where there are two outputs with the same amplitude. Symmetry is taken to imply that the two outcomes should have equal probability and I accept that this follows in both the collapse and the branching case.

16. The second step is to extend the above result to the case of more than two outputs where the amplitudes are again the same. If these are from a single node, symmetry arguments similar to those used in the previous case, imply that each branch will have equal probability. If the multiple branches result from more than one branching event, the symmetry may appear to be broken (c.f. the red hat example) and the result is then less obvious. To consider this further, I return to the example of the state set out in (4), which was created by two bifurcations. A similar state would also result from a single equal-amplitude trifurcation, but the histories of the process are not the same and their symmetries differ (see diagram). It is *an assumption* that the resulting probabilities must be

the same in both cases.

17. If we do assume that the probabilities are the same in the two set-ups and the updating rule holds, we obtain the Born rule for the unequal probabilities of the two branches following the first bifurcation. This is essentially the DSW proof of the Born rule with the second bifurcation playing the role of the ancilla.

18. A further often unstated assumption underlies all attempts to reconcile probabilities with Everett. This is that it is actually possible to make sense of probabilities in this context. Once one assumes *ab initio* that a probabilistic model *must* supervene on the Schrődinger equation, that the probabilities *must be* a function only of the final state and that the standard updating rule *must* apply, everything else follows. But all these assumptions require justification: if supervenience holds, they need to be shown to be consistent with the Schrődinger equation and, ideally, we should be able to see how they emerge from it.

19. I now restate the argument in my paper that there is an actual inconsistency between the Born rule and the Everettian assumption that everything supervenes on Schrődinger. I consider the case where a significantly large number (N) of two-state systems have been measured using a SG apparatus or its equivalence. Assuming no collapse, branching occurs to produce one branch associated with each of the 2^N permutations of the final states of the N particles. The number of branches where M particles have been observed to have spin state $|\beta>$ equals ${}^{N}C_{M}$ and if this is combined with the Born-rule weights, the outcomes where M/N equals the Born probability are much more likely than the others when N and M are >> 1.

20. Given this, we consider a scenario where an observer (Alice again) has no prior information about the setting of the SG apparatus, but does know the value of N. After she observes M positive outcomes, she can deduce the likely setting of the apparatus from M/N using the Born rule. Moreover, if her state is completely described by the Schrődinger equation, all her properties, including this newly acquired knowledge, must be included in the part of the wavefunction associated with the branch she is in. The key point is that this is inconsistent with the linearity of the Schrődinger equation which requires that the form of the wavefunction associated with a branch is independent of the expansion coefficients that enter the Born rule.

21. A counter-argument to the above is to note that there are branches corresponding to every value of M, each containing a copy of Alice, who may attempt to use the Born rule to deduce the SG orientation. Some of these deductions will be right, but others (in general many more) will be wrong and they can only tell whether they are right or wrong by making a direct observation of the actual SG setting. Alice's "state of expectation" resulting from the observations would then be the same whatever the actual setting of the SG apparatus. However, the logical consequence of this is that if Alice understands these arguments, she should not expect to be able to acquire knowledge of the SG settings from her observation of the outputs and that her experience cannot be a function of the Born weights. She should therefore conclude that any probabilities or betting preferences she forms cannot be influenced by the Born weights and that the Born rule cannot apply. But of course, Alice's experience is the same as ours and is strongly governed by the Born rule.

Studies in History and Philosophy of Modern Physics 40 (2009) 243–250

Contents lists available at ScienceDirect

Studies in History and Philosophy of Modern Physics

journal homepage: www.elsevier.com/locate/shpsb

Everett and the Born rule

Alastair I.M. Rae

School of Physics and Astronomy, University of Birmingham, Birmingham B15 2TT, UK

ARTICLE INFO

Article history:
Received 6 October 2008
Received in revised form
10 February 2009

Keywords:
Everett
Born rule
Quantum measurement

ABSTRACT

During the last 10 years or so, derivations of the Born rule based on decision theory have been proposed and developed, and it is claimed that these are valid in the context of the Everett interpretation. This claim is critically assessed and it is shown that one of its key assumptions is a natural consequence of the principles underlying the Copenhagen interpretation, but constitutes a major additional postulate in an Everettian context. It is further argued that the Born rule, in common with any interpretation that relates outcome likelihood to the expansion coefficients connecting the wavefunction with the eigenfunctions of the measurement operator, is incompatible with the purely unitary evolution assumed in the Everett interpretation.

1. Introduction

The conventional (“Copenhagen”) interpretation of quantum mechanics states that the result of a measurement is one (and only one) of the eigenvalues belonging to the operator representing the measurement and that, following the measurement, the wavefunction “collapses” to become the corresponding eigenfunction (ignoring the possibility of degeneracy). According to the “Born rule”, the probability of any particular outcome is proportional to the squared modulus of the scalar product of this eigenfunction with the pre-measurement wavefunction. This analysis underlies many of the predictions of quantum mechanics that have been invariably confirmed by experiment. An alternative approach to quantum measurement is the Everett interpretation (also known as the “relative states” or the “many worlds” interpretation) which was proposed by Everett III (1957). The essence of this approach is that it assumes no collapse of the wavefunction associated with a measurement: instead, the time development of the state is everywhere governed by the time-dependent Schrödinger equation. After a “measurement-like” event, this results in a splitting of the wavefunction into a number of branches, which are then incapable of reuniting or communicating with each other in any way. This splitting occurs even when a human observer is part of the measurement chain: the resulting branches then each contain a copy of the observer, who is completely unaware of the existence of the others.

Since its inception, the Everett interpretation has been subject to considerable criticism—e.g. Kent (1990) and Squires (1990)—which has three main strands (or branches [*sic*]). First, there is its metaphysical extravagance. The continual evolution of the universe into a “multiverse” containing an immense number of branches would mean that the universe we observe should be accompanied by an immense number of parallel universes, which we do not observe and have no awareness of—surely such a postulate must be a gross breach of the principle of Occam’s razor! Everett himself was aware of this criticism and, in a footnote to his original paper, he compares the conceptual difficulties of accepting his interpretation with those encountered by Copernicus when the latter proposed the (in his time revolutionary) idea that the earth moves around the sun. However, the reason that arguments based on Occam’s razor have not led to the universal rejection of Everett’s ideas has less to do with the strength or otherwise of the Copernican analogy and is more a result of the fact that the branching of the universe into the multiverse is claimed to be a direct consequence of the time-dependent Schrödinger equation: no additional postulate, such as the collapse of the wavefunction, is required to explain the phenomenon of quantum measurement and the extravagance with universes may therefore be considered a price worth paying for the economy in postulates.

The second strand in the criticism of Everett is known as the “preferred basis” problem. This is because there is an apparent ambiguity in the way the branches are defined. Thus, if the wavefunction of a system has the form $\psi = A\psi_1 + B\psi_2$, then Everett suggests that a measurement should lead to two sets of

E-mail address: alastair@aarae.co.uk

1355-2198/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.shpsb.2009.06.001

branches, one associated with each of the states represented by ψ_1 and ψ_2. However, the original state could just as well be written as $\psi = C\phi_1 + D\phi_2$ where $\phi_1 = 2^{-1/2}(\psi_1 + \psi_2)$, $\phi_2 = 2^{-1/2}(\psi_1 - \psi_2)$, $C = 2^{-1/2}(A + B)$ and $D = 2^{-1/2}(A - B)$, so why should the branches not be just as well defined by ϕ_1 and ϕ_2—or indeed any other orthogonal pair of linear combinations of ψ_1 and ψ_2? This problem has been largely resolved by the appreciation of the importance of the effect of the environment on a quantum system and the associated "decoherence"—Zurek (2007) and Wallace (2002, 2003a). A quantum measurement is inevitably accompanied by complex, chaotic processes which act to pick out the particular basis defined by the eigenstates of the measurement operator. This basis is then the one "preferred" by the Everett interpretation and this supervenes on the Schrödinger wavefunction. This result is now generally accepted, although Baker (2006) argued that its derivation uses the Born rule so that there is a danger of circularity if it is then assumed as part of its proof.

The third criticism leveled at Everett is the problem of probabilities. The conventional (Copenhagen) interpretation states that, if the wavefunction before a measurement is $\psi = A\psi_1 + B\psi_2$, and if ψ_1 and ψ_2 are eigenstates of the measurement operator with eigenvalues q_1 and q_2, respectively, then the outcome will be *either* q_1 with probability $|A^2|$ *or* q_2 with probability $|B^2|$, where these probabilities reflect the frequencies of the corresponding outcomes after a large number of similar measurements. However, according to the Everett approach there is no "either–or" because both outcomes are manifest, albeit in different branches. Instead of a *disjunction* to which we can apply standard probability theory, we have a *conjunction*, where it is hard to see how probabilities can make any sense—Squires (1990), Graham (1973), and Lewis (2004). There have been several attempts to resolve this conundrum and to show how probability (or something else that is in practice equivalent to it) can be used in an Everettian context. David Wallace has proposed a principle that he calls "subjective uncertainty" in which he claims that a rational observer should expect to emerge in one branch after a measurement, even though she is also reproduced in the other branches—Wallace (2003b, 2007). Greaves (2004) has criticized this approach and suggested an alternative in which we have to take into account the observer's "descendants" in all the branches, but we should "care" more about some than others; the extent to which we should care is quantified by a "caring measure" that is proportional to the corresponding Born-rule weight. Both these approaches are designed to explain why some branches appear to be favored over others, but both attempt to do this without altering Everett's main principle that the quantum state evolves under the influence of the time-dependent Schrödinger equation with nothing else added, so that the Born rule supervenes on this. An alternative approach, which I shall not discuss any further in this paper, is to maintain most of the fundamental ideas of the Everettian interpretation, but add a further layer of "reality" to justify the use of probabilities; an example of this can be found in Lockwood (1989).

Interest in the Everett interpretation has been on the increase recently—particularly during 2007, which was the 50th anniversary of the publication of Everett's original paper (Everett III, 1957). Much of the renewed interest has developed from the work by Deutsch (1999) some eight years earlier, which was then developed by Wallace (2003b, 2007) and Saunders (2004). This program (which I refer to below by the initials DSW) aims to derive the Born rule from minimal postulates that are claimed to be consistent with the Everett interpretation, as well as with other approaches to the measurement problem. In fact, Deutsch (1999) made little reference to the Everett interpretation in his derivation of the Born rule, and Saunders (2004) emphasized and believed that his derivation is independent of any assumptions about the measurement process. However, Wallace (2007) assumed the Everett interpretation and claimed that his derivation shows the Born rule to be completely consistent with it. Gill (2005) examined Deutsch's derivation and sought to clarify the assumptions underlying it, again without referring to the Everett interpretation as such. A similar approach, but using slightly different assumptions, has been developed by Zurek (2007) and is set out in a recent review paper.

The present paper aims to show that some of the postulates underlying the above derivations arguments do not follow naturally from the Everett interpretation and may well not be consistent with it.

2. The DSW proof of the Born rule

This section sets out the DSW derivation of the Born rule by applying it to a particular example. The argument is deliberately kept as simple as possible and more general treatments can be found in the cited references. Consider the case of a spin-half particle, initially in an eigenstate of an operator representing a component of spin in a direction in the xz plane at an angle θ to the z axis, passing along the y axis through a Stern–Gerlach apparatus oriented to measure a spin component in the z direction.

Standard quantum mechanics tells us that the initial state α_θ can be written as a linear combination of the eigenstates of $\hat{S}_z$: α with eigenvalue $+1$ (in units of $\hbar/2$) and β with eigenvalue -1. We have

$$\alpha_\theta = c\alpha + s\beta \tag{1}$$

where $c = \cos(\theta/2)$ and $s = \sin(\theta/2)$. Particles emerge from the two channels of the Stern–Gerlach apparatus, with the upper and lower channels indicating $S_z = +1$ and -1, respectively, and are then detected. After they have entered the detectors, but before any collapse[1] associated with the measurement, the total wavefunction of the system is

$$\psi = c\alpha\chi_+ + s\beta\chi_- \tag{2}$$

where χ_+ (χ_-) is the wavefunction representing the detectors, including their environment, when a particle is detected in the positive (negative) channel. According to the Copenhagen interpretation, the corresponding probabilities for a positive or a negative outcome are given by the Born rule as c^2 and s^2, respectively. From the Everettian point of view, on the other hand, there is no collapse and the system is always in a state of the form ψ. However, because of the effects of the environment and decoherence, phase coherence between the two terms on the right-hand side of (2) is lost, so they can never in practice interfere. The wavefunction has therefore evolved into two "branches" which then develop independently.

The principle of the DSW approach is to describe the process being studied as a game, or series of games, where we receive rewards, or pay penalties (i.e. receive negative rewards) depending on the outcomes. The derivation proposed by Zurek (2007) is quite similar to this, although it does not use game theory.

Imagine a game where the player receives a reward depending on the outcome of the experiment. Assume that the value of θ is under our control and that, whenever the experimenter observes a

[1] At a number of points in this paper, I compare the predictions of the Everett model with those produced by the "Copenhagen interpretation", by which I mean a model in which the wavefunction collapses into one of the eigenstates of the measurement operator. This is assumed to occur early enough in the process for the outcomes to be the same as would be observed if particles were to emerge randomly from one or other output channel, with the relative probabilities of the two outcomes determined by the Born rule.

particle emerging from the positive or negative channel of the Stern–Gerlach apparatus, she receives a reward equal to x_+ or x_-, respectively; these values can be chosen arbitrarily by the experimenter. In the special cases where $\theta = 0$ or π, the initial spin state is an eigenstate of $\hat{S}_z$ with eigenvalues $+1$ and -1, respectively. The particle then definitely emerges from the corresponding channel of the Stern–Gerlach apparatus and the corresponding reward is paid.

In the general case, we define the "value"—$V(\theta)$—of the game as the minimum payment a rational player would accept not to play the game, and look for an expression for $V(\theta)$ of the form

$$V(\theta) = w_+(\theta)x_+ + w_-(\theta)x_- \tag{3}$$

where the ws are non-negative real numbers that we call "weights" and which are normalized so that their total is unity. We shall find that

$$w_+(\theta) = c^2 \quad \text{and} \quad w_-(\theta) = s^2 \tag{4}$$

which are the probabilities predicted by the Born rule for this setup.

First consider the effect on the wavefunction of rotating the SG magnet through 180° about the y axis. It follows from the symmetry of the Stern–Gerlach apparatus that spins that were previously directed into the upper channel will now be detected in the lower channel and vice versa. Thus

$$V(\theta + \pi) = w_+(\theta + \pi)x_+ + w_-(\theta + \pi)x_- = w_-(\theta)x_+ + w_+(\theta)x_- \tag{5}$$

From standard quantum mechanics, the effect of this rotation on wavefunction (2) is to transform it to

$$\psi = -s\alpha\chi_+ + c\beta\chi_- \tag{6}$$

We now proceed by considering a series of particular values of θ.

Case 1: The first case is where $\theta = 0$ so that the initial state, α_θ, is identical with α. As noted above, this state is unaffected by the measurement and the particle is always detected in the positive channel. Thus $V(0) = x_+$, $w_+(0) = 1$ and $w_-(0) = 0$. Similarly, $V(\pi) = x_-$, $w_+(\pi) = 0$ and $w_-(\pi) = 1$.

Case 2: In the second case, $\theta = \pi/2$ so that ψ is as in (2), but with $c = s = 2^{-1/2}$. Now consider the effect of rotating the Stern–Gerlach apparatus through an angle π. Using (5) and (6), we get the following expressions for V and ψ:

$$V(3\pi/2) = w_-(\pi/2)x_+ + w_+(\pi/2)x_- \tag{7}$$

$$\psi = 2^{-1/2}[-\alpha\chi_+ + \beta\chi_-] \tag{8}$$

The only change in the wavefunction is the change of sign in the term involving α. DSW point out that this sign, in common with any other phase factor, should not affect the value, because it can be removed by performing a unitary transformation on this part of the wavefunction only—e.g. by a rotation of the spin through 2π or by introducing an additional path length equal to half a wavelength. Moreover, Zurek (2007) showed that one of the effects of the interaction of the system with the environment is to remove any physical significance from these phase factors. It follows that the value should not be affected by the rotation so that $V(3\pi/2) = V(\pi/2)$, which leads directly to

$$w_+(\pi/2) = w_-(\pi/2) = 1/2 \quad \text{and} \quad V(\pi/2) = (x_+ + x_-)/2 \tag{9}$$

This result (which might be thought to be an inevitable consequence of symmetry) is considered by DSW to be the key point of the proof. We should note that, although it agrees with the Born rule, it would also be consistent with any alternative weighting scheme that predicted equal weights in this symmetric situation: in particular it is consistent with a model in which the weights were assumed to be independent of θ.

We now extend the result to the case where the number of output channels is M instead of two and the wavefunction is the sum of M terms, each of which corresponds to a different eigenstate of the measurement operator. In the case where the coefficients of this expansion are all equal, any action that has the effect of exchanging any two output channels (which are numbered 1 and 2) must leave the wavefunction unchanged apart from irrelevant changes in phase. The value is then also unchanged, but the roles of w_1 and w_2 are reversed. Hence

$$w_1x_1 + w_2x_2 = w_1x_2 + w_2x_1 \tag{10}$$

where x_i is the reward associated with the ith output channel. It follows that $w_1 = w_2$; consideration of other permutations immediately extends this result to all i and we have $w_i = N^{-1}$.

Case 3: In the third case, $\theta = \pi/3$ so that $\cos(\theta/2) = \sqrt{3}/2$ and $\sin(\theta/2) = 1/2$. We now assume that the system is modified so that, after emerging from the Stern–Gerlach magnet and before being detected, the outgoing particles interact with a separate quantum system that can exist in one of, or a linear combination of, four eigenstates ϕ_i. Following Zurek (2007), this is referred to as an "ancilla" from now on. The ancilla is designed so that, if $\theta = 0$ so that all spins emerge from the positive channel, the ancilla is placed in the state $3^{-1/2}\sum_{i=1,3}\phi_i$; while, if $\theta = \pi$ and all spins are negative, its state becomes ϕ_4. From linear superposition it follows that if the original spin is in a state of form (2) with $\theta = \pi/3$, the total wavefunction of the spin plus the ancilla is

$$\begin{aligned}\Psi &= 3^{-1/2}[\phi_1 + \phi_2 + \phi_3]\cos(\pi/6)\alpha + \phi_4\sin(\pi/6)\beta \\ &= \tfrac{1}{2}[\phi_1\alpha + \phi_2\alpha + \phi_3\alpha + \phi_4\beta]\end{aligned} \tag{11}$$

As the coefficients of each term in the above expansion are equal, it follows from the earlier discussion of case 2 that all four weights are equal to 0.25. If we were to measure on the ancilla a quantity whose eigenstates were one of the functions ϕ_1–ϕ_4, we should obtain a result equal to one of the corresponding eigenvalues. If the result corresponds to one of the first three eigenfunctions, we can conclude that if, instead, we had measured the spin directly, we would have got a positive result, while a result corresponding to ϕ_4 indicates a negative spin. As this is the only such state, it follows that the weight corresponding to a negative spin is $w_-(\pi/3) = 0.25$ and therefore, from normalization, that $w_+(\pi/3) = 0.75$. (The last step, which follows Zurek (2007), establishes these results without assuming that the weights are additive.) The value of the game therefore equals $0.75x_+ + 0.25x_-$. It can also be shown quite straight forwardly—Zurek (2007)—that, after a number of repeats of the experiment, the predicted distribution of the results is as observed experimentally.

Following DSW, the above argument can be extended to the case of a measurement made in the absence of the ancilla if we make a further assumption, known as "measurement neutrality". This states that the outcome of the game is independent of the details of the measurement process—i.e. the presence or absence of the ancilla—so that $w_+(\pi/3) = 0.75$ and $w_-(\pi/3) = 0.25$ in either case. These quantities are identical to $\cos^2(\theta/2)$ and $\sin^2(\theta/2)$, respectively, so the derived weights are the same as those predicted by the Born rule. By choosing an appropriate ancilla, the above argument can be directly extended to examples where the ratio of the weights is any rational number, and then to the general case by assuming that the weights are continuous functions of θ. Hence, the expression for the value is the same as that predicted by the Born rule:

$$V(\theta) = c^2x_+ + s^2x_- \tag{12}$$

Further generalization to experiments with more than two possible outcomes is reasonably straightforward and does not introduce any major new principles.

3. Discussion

There have been a number of criticisms of the DSW proof when applied to the Everett model in particular—e.g. Baker (2006), Barnum, Caves, Finkelstein, Fuchs, and Schack (2000), Lewis (2005, 2007), and Hemmo and Pitowski (2007); some of these even challenge result (9) for the symmetric case. I shall shortly develop arguments to show that, although the symmetric results appear to be consistent with the Everett model, this may not be so in the asymmetric case.

First consider how the above translates into predictions of experimental results. The game value is the minimum payment a rational observer would accept in order not to play the game. This means that after playing the game a number of times, a rational observer should expect to receive a set of rewards whose average is equal to the game value. Thus, if we consider a sequence of N such observations in which n_+ and $n_-(=N-n_+)$ particles are detected in the positive and negative channels, respectively, the total reward received will be $n_+x_+ + n_-x_-$, and this should equal $N(w_+x_+ + w_-x_-)$ implying that $w_+ = n_+/N$ and $w_- = n_-/N$. This, of course, is just what is observed in a typical experiment provided N is large enough for statistical fluctuations to be negligible. It should be noted that frequencies are *not* being used to *define* probabilities, but the derived weights are used to predict the results of experimental measurement of the frequencies.

The above results are of course consistent with the standard Copenhagen interpretation, whose fundamental mantra was set out by Bohr (1935): "… there is essentially the question of an influence on the very conditions which define the possible types of predictions regarding the future behavior of the system". In the present context, this means that, because an experiment designed to demonstrate interference would involve a different experimental arrangement, the experiment can be modelled as a classical stochastic system in which spins emerge from *either* the positive *or* the negative channel of the Stern–Gerlach apparatus. (It should be noted that this paper does not aim to justify the Copenhagen interpretation, but employs its results as a comparator with the Everettian case.)

Why should an Everettian observer have experiences such as those just described? In the Everett interpretation, the quantum state evolves deterministically and on first sight, there would appear to be no room for uncertainty. However, after a splitting has occurred, observers in different branches have the same memories of their state before the split, but undergo different experiences after it. Given this, it may be meaningful for an experimenter to have an opinion about the likelihood of becoming a particular one of her successors. This introduces a form of subjective uncertainty, and Wallace (2007) claimed that this plays a role in the Everett interpretation that is equivalent to that played by objective stochastic uncertainty in the Copenhagen case. However, we should note that such subjective uncertainty can only come into play at the point where the experimenter becomes aware of an experimental result, in contrast to the Copenhagen model where the splitting is assumed to occur as the particles emerge from the Stern–Gerlach magnet. I shall shortly proceed to compare and contrast the Copenhagen and Everettian interpretations of the different experiments discussed above. To help focus the discussion, I shall initially assume that in such experiments each particular result is associated with only one branch of the final wavefunction. This assumption has been strongly criticized by DSW and others and I shall return to the question of how it affects our conclusions at a later stage. I now analyze our earlier arguments step by step.

Case 1: *Copenhagen*: As the initial spin state is in an eigenstate of S_z, the result is completely determined. The probability of the result equalling the corresponding eigenvalue is 1 and the probability of the alternative is zero.

Case 1: *Everett*: There is only one branch and this contains the only copy of the observer who invariably records the appropriate eigenvalue.

There is therefore no difference between the observers' experiences in case 1 under the Copenhagen and Everettian interpretations.

Case 2: *Copenhagen*: The probabilities of positive and negative results are both 0.5. After a large number of repeats of the experiment, the experimenter will have recorded approximately equal numbers of positive and negative results, so her average reward will be $(x_1 + x_2)/2$, which is the same as the game value.

Case 2: *Everett*: The observer will split into two copies each time a spin is observed and the weights of the two branches are equal for the reasons discussed earlier. After a large number (N) of repeats of the experiment the vast majority of observers will have recorded close to $N/2$ positive and $N/2$ negative results and their average rewards will both equal the game value.

There is therefore no difference between the predictions of the Copenhagen and Everettian interpretations in case 2.

Case 3: *Copenhagen*: As emphasized above, this assumes that the experiment is a stochastic process in which a particle emerges from either the positive or the negative channel and the relative probabilities of the outcomes are equal to the Born weights. In the presence of the ancilla, a particle is detected in one (and only one) of the equally weighted states ϕ_1 to ϕ_4, and all four outcomes have equal probability. To have been observed in any of the first three states, the spin must have emerged from the Stern–Gerlach experiment through the positive channel, while if the final result corresponded to ϕ_4, it must have come through the negative channel. It follows directly that if the ancilla were absent, three times as many spins would be detected as positive than as negative. Thus, the principle of measurement neutrality, assumed in stage 3 of the earlier derivation, follows naturally from the assumptions underlying the Copenhagen interpretation.

Case 3: *Everett*: We first consider the situation where an ancilla is present so that the state is described by (11); there are therefore four equally weighted branches, one corresponding to each of the ϕ_i. The observer splits into four equally weighted copies and should expect her descendants to record an equal number of each of the four possible results and therefore conclude that there are three times as many positive as negative spins. However, in the absence of an ancilla, there are only two branches and the observer is split into two copies each time a result is obtained. To show that a typical Everettian observer should record results that are consistent with the Born weights, we again have to apply the principle of measurement neutrality. We saw above that this is a natural, if not inevitable, consequence of the Copenhagen interpretation, but we shall now demonstrate that this is not the case in an Everettian context.

Under the Copenhagen interpretation, particles are assumed to emerge from *either* the positive *or* the negative channel and then into one, *and only one*, of the states ϕ_i. This is not true in the case of the Everett interpretation, where the system evolves deterministically and the state is described by a linear combination of the wavefunctions associated with a particle being present in each channel. Apparent stochasticity, or subjective uncertainty, only enters the situation at the point where the experimenter observes the result and splits into a number of descendants—two in the absence of the ancilla and four if it is present. There is no requirement for the frequencies to be the same in both cases—i.e. no *a priori* reason to apply the principle of measurement neutrality. In the language of decision theory, the values of the two games are not necessarily the same, so a decision on whether or not to accept a payoff may depend on whether the game is

being played with or without an ancilla. Indeed, as in the absence of an ancilla there are only two branches, we might expect each observer's experience to be the same as in case 2, with equal numbers of positive and negative results and an equal reward for each outcome—i.e. the statistical outcomes would be independent of the weights. I shall argue later that this is a natural consequence of the Everettian interpretation, but at present simply emphasize that the principle of measurement neutrality is a self-evident consequence of the assumptions underlying the Copenhagen interpretation, but constitutes a major additional postulate in the context of Everett.

I further illustrate this last point by considering a simple classical example that consists of a box with two exit ports from each of which a series of balls emerges as in Fig. 1. The apparatus can be operated in one of two modes that we denote as "C" and "E". In the C mode, balls emerge one at a time from one of two output ports and, on average, three times as many come out of the upper port as from the lower. An experimenter observes the balls as they emerge and confirms this relative likelihood. Still in the C mode, the experiment is modified so that when a ball emerges from the upper port, it passes into a second, "ancillary" box and then emerges at random through one of three output channels before being detected. The experimenter now detects a ball either in one of these three channels or emerging from the lower port. Clearly the first of these results is three times as likely as the second, so the observed frequencies are independent of the presence or absence of the second box. Thus the equivalent of measurement neutrality holds in this case.

Now consider the game in the E mode, which is also illustrated in Fig. 1. In this case two balls emerge from the box simultaneously: a black ball from the upper port and a white ball from the lower. The two balls fall into a receptacle (not shown in the figure) and an experimenter draws one at random; after repeating the experiment a number of times she sees equal numbers of black and white balls. The experiment is now modified so that the black balls are directed into an ancillary box which now contains a device that releases three identical black balls, one through each of the three output ports, whenever one enters. These three balls along with the white one now fall into the receptacle and the observer again draws one at random; she now sees a black ball three times as often as a white ball. Thus, the relative likelihood of a black or a white ball depends on the presence or absence of the second box, and we can conclude that measurement neutrality is not necessarily preserved when the game is played in the E mode.

A more whimsical analogy follows the precedent set by Schrödinger's cat by using animals to illustrate our point. First consider Copenhagen rabbits. These come in two colors—black and white; they are all female and capable of giving birth to one (and only one) baby rabbit which is always of the same color as its mother. Let us suppose we have four Copenhagen rabbits, three black and one white in a hat and suppose that one, of them, chosen at random, is pregnant. We first play the game of "pick out the pregnant rabbit" by putting our hand in the hat, identifying and then pulling out the pregnant rabbit. We are paid different rewards (x_b and x_w) depending on whether the extracted rabbit is black or white. After playing the game a number of times, we find that we have pulled out three times as many black rabbits as white, so that the game value is $(3x_b + x_w)/4$. The second game is one where we wait until the pregnant rabbit has given birth and then pull out and identify the color of the baby. Clearly the results and the value are the same as in the first game.

Now consider Everettian rabbits, which are also either black or white. In contrast to the Copenhagen rabbits, they are capable of carrying and giving birth to more than one offspring. Suppose we have two pregnant Everettian rabbits: a white rabbit that is pregnant with a single offspring and a black rabbit that is expecting triplets. If we draw one of the two pregnant rabbits from the hat at random, the game value will be $(x_b + x_w)/2$. However, if, instead, we wait until after the rabbits have given birth and then draw out one of the offsprings at random, the game value will now be $(3x_b + x_w)/4$. Thus Copenhagen rabbits preserve measurement neutrality, but Everettian rabbits do not.

Given the assumptions underlying the Copenhagen interpretation, the first game in the C mode and the game with the Copenhagen rabbits form close parallels with the quantum example discussed earlier. Similarly, the first game in the E mode and the game with Everettian rabbits are closely parallel to the quantum case, provided we accept that random selection at the point where the observer becomes aware of the result is equivalent to subjective uncertainty in the quantum case.

In both these examples as well as in the quantum case, I have shown that measurement neutrality is not a necessary consequence of the principles underlying the Everett interpretation. However, in all the cases where it need not apply, the symmetry is broken in the sense that the weights associated with the different outcomes are not equal. It follows that measurement neutrality (or, indeed, some other quite different principle) could be restored in the classical examples by making additional assumptions: for example, it could be arranged that the ball emerging from the upper channel in the E game is heavier than that coming out of the lower, and that it is three times easier to find and extract a more massive ball when making the selection; similarly, it might be three times easier to catch a rabbit carrying triplets than one pregnant with a single offspring. However, such *ad hoc* rules would have to be built into the physics of the setup when it was designed and constructed. In the quantum case under the Everett interpretation, measurement neutrality therefore has to be an additional assumption, rather than following directly from the structure of the theory as in the Copenhagen case. Gill (2005) showed that measurement neutrality is equivalent to assuming that the measures of probability are invariant under functional transformations—i.e. the probability of obtaining a particular result when measuring a variable is the same as that pertaining when a function of the variable is measured. He considers that functional invariance in the case of one-to-one transformations is "more or less definitional", but is much less obvious in the

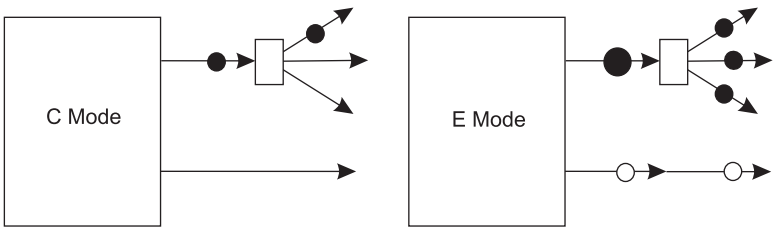


Fig. 1. In the C mode a ball is emitted from the first box through either the upper or the lower port and detected either before or after entering the second box; the figure shows one possible outcome. In the E mode, balls emerge from both ports and one of them is detected either before or after the second box, which releases three balls every time one enters.

many-to-one case, which is required for situations such as case 3. Gill's discussion relates to probabilities as conventionally defined and his paper makes no reference to the Everett interpretation. I believe that the above argument shows that many-to-one transformations are also "more or less definitional" under the Copenhagen interpretation, but not in the Everettian context.

Measurement neutrality and an associated principle that he calls "equivalence" have been argued for by Wallace in a number of papers—Wallace (2002, 2003a, 2003b, 2007). He considers games in which the measurement result is erased after it triggers an associated reward and before the experimenter has recorded the outcome. In the symmetric ($\theta = \pi/2$) case, the final states are independent of the pattern of rewards, which reinforces the arguments leading to (9). However, this is not an issue in the present discussion, which challenges the assumption of measurement neutrality only in the asymmetric case. Another point emphasized by Wallace (2007) is that the boundary between what is usually taken as preparation and what is part of the "actual" measurement is essentially arbitrary, particularly in the context of the Everett interpretation. However, the observation and recording of the result by a conscious observer is part of the measurement proper, and it is only at this point that subjective uncertainty or the relevance of a caring measure is introduced into the Everettian treatment of the Born rule.

Up to this point I have argued that the assumptions underlying the derivation of the Born rule, in particular measurement neutrality, are not necessary in an Everettian context, though they may be treated as added postulates. I now intend to go further and argue that there is an inconsistency between the assumptions underlying the Everett interpretation and the Born rule—or, indeed any rule that relates the likelihood of a measurement outcome to the amplitudes (c and s in the above example) associated with the branching of the wavefunction in a non-trivial way. I shall continue to use the example of the measurement of the spin component of a spin-half particle as a focus of the discussion.

The scenario I now discuss is one where an observer ("Bob") records the number of positive spins (M) in a set of measurements of the state of N identically prepared spins that have passed through a Stern–Gerlach apparatus. We consider the particular case where Bob *does not know the value of* θ *before he makes any measurements*; that is, he has not seen the apparatus or been told how the magnet is oriented, which means that his initial state is represented by a wavefunction which is independent of θ. However, if Bob knows the Born rule, he can estimate the value of θ as $2\cos^{-1}(M/N)^{1/2}$ and his confidence in this value will be the greater, the larger are M and N. As a result of this experience, Bob's state has been changed from one of ignorance to one where he has some knowledge of θ. This change must therefore have been reflected in Bob's quantum state, causing a modification to his wavefunction, which now depends on θ. To further emphasize this point, suppose that the value of θ can be changed without Bob's direct knowledge by another experimenter ("Alice") who has control of the Stern–Gerlach apparatus. If she does this and the experiment is repeated a number of times at the new setting, Bob will find that his expectations have been consistently wrong. He may initially attribute this to statistical fluctuation, but eventually he will amend his state of expectation to bring it into line with his experience. Indeed, Bob may know that Alice is able to do this, in which case he will be more likely to amend his state of expectation at an earlier stage. Alice could then send signals to Bob by transmitting sets of N particles using the same value of θ for each set, but changing it between sets. If the Born rule applies, Bob can deduce the values of θ that Alice has used from the relative numbers of positive and negative results, so Alice has again caused changes in Bob's state of expectation and therefore of his wavefunction.

It is one of the principles of the Everett interpretation that, once branching has occurred and the possibility of interference between branches has been eliminated, the wavefunction associated with a branch describes the "relative state" of the system contained in that branch, which cannot be influenced by the state of any other branch. Moreover, the form of the relative state functions, which represent the whole branch including the version of Bob associated with it, are the same whatever the values of the expansion coefficients c and s. This implies that the properties of a system represented by such a relative state are not affected by the measuring process. Thus, although these constants enter the expressions, they do so only as expansion coefficients, which have no effect on the wavefunctions of the relative states associated with the component branches. In particular, the observer's state of knowledge of the value of θ cannot be altered as a result of this process. This is in direct contradiction to the conclusion reached above, assuming that the Born rule holds. There is therefore an inconsistency between the principles underlying the Everett interpretation and the appearance of a correlation between the apparatus setting and the relative frequencies of the possible outcomes, such as is implied by the Born rule.

To develop this point further, consider the state of the whole system after N particles have passed through the apparatus, so that, according to the Everett interpretation, the wavefunction contains 2^N branches that correspond to all possible sequences of the results of the measurements performed so far. That is, using (2),

$$\prod_{i=1,N} \alpha_\theta(i)\chi_0 \longrightarrow \sum_{P_{s_i}} c^m s^{N-m} \Psi(s_1, s_2, \ldots, s_N) \tag{13}$$

where $\alpha_\theta(i)$ is the initial state of spin i and χ_0 refers to the initial state of the detecting apparatus, including the observer Bob, which is independent of θ, given the assumptions set out earlier. Each parameter s_i has two values, + and −; $\Psi(s_1, s_2, \ldots, s_N)$ represents the state of the whole system (i.e. spins, measuring apparatus and Bob) after the results s_i have been recorded in a measurements on spin i for all i from 1 to N; m equals the number of positive spins in this set; $\sum_{P_{s_i}}$ implies a summation over all 2^N permutations of s_i. Each term in the summation refers to a separate branch in the Everett interpretation.

It follows from (13) that the number of branches in which m positive results have been recorded is $N!/m!(N-m)!$ and the Born weight associated with this whole subset equals $c^{2m}s^{2(N-m)}$. Under the Copenhagen interpretation, the probability of observing m positive results is the product of these two quantities: this has a maximum value when $m = M = Nc^2$ ($= 3N/4$, if $\theta = \pi/3$ as in case 3) and a standard deviation of $|cs|N^{1/2}$ ($= \sqrt{N}/4$). Suppose now that the Everett assumptions hold so that there has been no collapse. After the measurement, wavefunction (13) will consist of a linear combination of branches, each of which contains a version of Bob who has recorded a value for m. If N is large, the vast majority of observers will observe approximately equal numbers of positive and negative results and a small minority will observe results in the vicinity of the ratio predicted by the Born rule. Repeating the experiment with a different value of θ does not change the number of observers recording any particular result, so, if this were all there were to it, Bob's experience would not correlate with the apparatus setting and he would be unable to deduce a reliable value of θ from his observations. However, the Everett interpretation only works if this is not all there is to it. Because of subjective uncertainty, an observer's successors in branches that have a high Born-rule weight are somehow favored over the others. How this can work is at the heart of the

difficulties many critics have with the Everett interpretation, but let us leave this on one side. The fact that these successors are so preferred means that they can with confidence deduce the value of θ from their observations of M and N. Acquiring this information must therefore have altered their reduced state, in contradiction to the Everettian assumptions set out above.

Several points should be noted about the above. First, the contradiction does not arise in the Copenhagen interpretation because, as noted earlier, this assumes that stochasticity arises at the point where the spin emerges from the Stern–Gerlach magnet. The information as to which branch is occupied by the spin is additional to that contained in the wavefunction and is obtained by Bob through the collapse process. Hence, no contradiction arises when this is used by the experimenter to guide his expectations about subsequent measurements.

Second, it should be emphasized that the argument applies only to information about the apparatus setting that is obtained by Bob *as a result of the measurement process*. He could of course have been told in advance how the apparatus was set up so, in this case, χ_0 would already be a function of θ. The latter argument could probably be extended to show that he should not be able to obtain further information about θ by the measurement process, but I believe it clarifies the discussion if we focus on the case where Bob has no prior knowledge of θ: to demonstrate inconsistency, it is only necessary to establish a contradiction in at least one particular case.

Third, although I have focussed on the Born rule, the above arguments would apply equally well to any model in which the outcome frequencies were assumed to depend systematically on the expansion coefficients. This is of rather marginal interest given that the Born rule is the one that is established by experiment.

If we accept the above, it follows that the only way probability should be able to enter the Everett interpretation is if all branches are assigned equal weight. Might it nevertheless be possible to reconcile this conclusion with experiment? Up to now, we have assumed one branch per outcome, without attempting to justify this. We now turn to the question of "branch counting", which means considering the number of branches associated with any given measurement outcome. If we accept the argument that the expansion coefficients play no role in determining the outcome likelihood in an Everettian context, then an experimenter's expectation of a particular outcome should be proportional to the number of branches associated with it. Such an assumption is similar to that made in statistical thermodynamics, where the ergodic hypothesis states that the result of averaging over an ensemble of systems is the same as the time average for a single system. When applied to the symmetric case, this is an essential part of the arguments leading to (9) and (10). However, branch counting has been strongly criticized by DSW on a number of grounds. Wallace (2007) considered a scenario in which extra branching is introduced into one (say the plus) channel by associating with it a device that displays one of, say, a million random numbers. He argued that this must be irrelevant to an experimenter who sees only the measurement result and is indifferent to the outcome of the randomizing apparatus. This is because "if we divide one outcome into equally valued sub-outcomes, that division is not decision-theoretically relevant". However, this argument does not fully take into account the Everettian context. Referring again to the classical game discussed earlier and illustrated in Fig. 1, we can consider the additional branching on the right of both setups as due to the presence of a randomizer with three possible outputs. In the case of the C game, these are indeed irrelevant to the expectation of the player, because a ball emerges from only one of the three channels and must therefore have passed through the upper channel at the previous stage. However, in the case of the E game, the chances of observing a black ball are enhanced (tripled) by the splitting and this would have to be taken into account by any rational player, even if the only result she sees is the color of the ball. Similarly, if we introduce a random number machine as Wallace suggests, then its state will be a linear combination of its million possible outcomes and all these will be associated with a positive value of spin. Given that there is only one branch associated with the alternative outcome, we could well expect the subjective likelihood of a positive result to be one million times greater than that for a negative outcome.

A second argument deployed to criticize branch counting is based on the fact that the interaction of a quantum system with its environment leads to an immensely complex branching structure. Indeed it is claimed by DSW that the number of branches is not only very large (possibly infinite), but is also subject to very large and rapid fluctuations before, during and after the observation of a result; which may mean that it is not meaningful to talk about even the approximate number of branches that exist at any time. This is adduced as a reason why a rational player should ignore the complexity of the branching structure and instead expect to observe results consistent with the Born rule. However, if the likelihood of observing a particular result is proportional to the number of associated branches, the complexity introduced by decoherence should actually result in the outcome of a measurement being completely unpredictable. The situation is similar to chaos in classical mechanics or to turbulence in hydrodynamics, whose onset certainly does not lead to increased predictability. In the arguments above, we assumed that each outcome was associated with a single branch, so what would be the likely consequences of a complex branch structure in an Everettian context? First, there may well be situations in which we could expect the number of branches associated with different outcomes to be equal, at least when averaged over a number of measurements, and in this case our earlier discussion would not be affected. However, we might be able to devise a situation (e.g. one in which a detector was placed in the positive output channel only) where the numbers of branches in the two channels would be expected to differ greatly. We should then expect to detect a larger number of (say) positive than negative results. This would be true even if the Stern–Gerlach apparatus were oriented symmetrically—i.e. with $\theta = \pi/2$, so the symmetry on which we based some of our earlier arguments would not hold. The complexity and fluctuations of the branch structure in the Everett case would render even the statistical results of a quantum measurement unpredictable. Such a situation is sometimes described as being "incoherent" and it has been argued that this would mean that the universe would be nothing like the one we experience. However, the obvious conclusion to draw from this is that the Everett assumptions are falsified, rather than that the Everett model is correct and the arguments based on it that lead to this incoherence must be wrong.

It might be thought that branch counting could restore the Born rule if the number of branches associated with a particular outcome were proportional to the Born weight. However, not only is there no obvious mechanism to achieve this, but it is also inconsistent with the Everett model for the same reasons as were set out earlier. The quantum description of the branch structure is contained within Ψ in (13) and therefore cannot depend on the expansion coefficients for the reasons argued above.

4. Conclusions

I have argued that attempts to prove the Born rule make assumptions that are essentially self-evident in the context of the Copenhagen interpretation, but not with the Everett model of

measurement. I have further argued that probabilities which are functions of the expansion coefficients are not consistent with the Everett interpretation, because these quantities are not then accessible to an observer in the reduced state associated with a branch. An alternative scheme that could be consistent with Everett is one where each branch has the same probability and the probability of a given outcome depends on the number of branches associated with it. However, this also cannot be made consistent with the Born rule and it leads to predictions of chaotic, unpredictable behavior, in contrast to the relatively well-ordered behavior, invariably demonstrated in experiments. I conclude that the Born rule is a vitally important principle in determining quantum behavior, but that it depends on wavefunction collapse, or something very like it, that does not supervene upon the time-dependent Schrödinger equation. It would be possible to retain the many-worlds ontology of the Everett model while allowing information to be transferred through the measurement, but the state evolution would no longer be governed by the Schrödinger equation alone and the economy of postulates would no longer obviously outweigh the metaphysical extravagance associated with the Everett picture.

The debate between the different interpretations of quantum mechanics has often been metaphysical in the sense that they often make the same predictions and cannot therefore be distinguished experimentally. The present paper has argued that this is not so in the case of the Everett interpretation, which predicts results different from those that follow from the Copenhagen interpretation, which in turn are supported by experiment. If this is accepted, the Everett model will have been falsified and the search for a consensual resolution of the quantum measurement problem will have to be focussed elsewhere.

Acknowledgments

I am grateful to Peter Lewis for comments on an earlier draft of this paper and to Simon Saunders and David Wallace for useful discussions. I would also like to thank the referees for comments that have helped me clarify some of the points made in the discussion.

References

Baker, D. (2006). Measurement outcomes and probability in Everettian quantum mechanics. *Philosophy of Science*, archive id2717.

Barnum, H., Caves, C., Finkelstein, J., Fuchs, C. A., & Schack, R. (2000). Quantum probability from decision theory?. *Proceedings of the Royal Society A*, *456*, 1175–1182.

Bohr, N. (1935). Can quantum mechanical description of physical reality be considered complete?. *Physical Review*, *48*, 696–702.

Deutsch, D. (1999). Quantum theory of probability and decisions. *Proceedings of the Royal Society A*, *455*, 3129–3137.

Everett III, H. (1957). "Relative state" formulation of quantum mechanics. *Reviews of Modern Physics*, *29*, 454–462.

Gill, R. D. (2005). *Quantum probability and infinite dimensional analysis: From foundations to applications*. Singapore: World Scientific (pp. 277–292).

Graham, N. (1973). *The measurement of relative frequency*. Princeton: Princeton University Press.

Greaves, H. (2004). Understanding Deutsch's probability in a deterministic universe. *Studies in the History and Philosophy of Modern Physics*, *35*, 423–456.

Hemmo, M., & Pitowski, I. (2007). Quantum probability and many worlds. *Studies in the History and Philosophy of Modern Physics*, *38*, 333–350.

Kent, A. (1990). Against many-worlds interpretations. *International Journal of Modern Physics A*, *5*, 1745–1776.

Lewis, D. (2004). How many lives has Schrödinger's cat?. *Australasian Journal of Philosophy*, *82*, 3–22.

Lewis, P. J. (2005). Probability in Everettian quantum mechanics. *PhilSCI*, eprint.

Lewis, P. J. (2007). Uncertainty and probability for branching selves. *Studies in the History and Philosophy of Modern Physics*, *38*, 1–14.

Lockwood, M. (1989). *Mind, brain and the quantum*. Oxford: Blackwell.

Saunders, S. (2004). Derivation of the Born rule from operational assumptions. *Proceedings of the Royal Society A*, *460*, 1–18.

Squires, E. J. (1990). On an alleged "proof" of the quantum probability law. *Physics Letters A*, *145*, 67–68.

Wallace, D. (2002). Worlds in the Everett interpretation. *Studies in the History and Philosophy of Modern Physics*, *33*, 637–661.

Wallace, D. (2003). Everett and structure. *Studies in the History and Philosophy of Modern Physics*, *34*, 87–105.

Wallace, D. (2003). Everettian rationality: Defending Deutsch's approach to probability in the Everett interpretation. *Studies in the History and Philosophy of Modern Physics*, *34*, 415–442.

Wallace, D. (2007). Quantum probability from subjective likelihood: Improving on Deutsch's proof of the probability rule. *Studies in the History and Philosophy of Modern Physics*, *38*, 311–332.

Zurek, W. H. (2007). Relative states and the environment. quant-ph/0707.2832v1.

Philip C. E. Stamp

Statement

and

Readings

MECHANISMS of DECOHERENCE

Philip Stamp

The topic of decoherence has a long history, beginning in the 1950s by Ludwig, Green and others; by the 1970s simple analyses of real systems had begun (Zeh, Simonius), and the idea that the environment could impose 'selection rules' (what later was called by Zurek the 'pointer basis') had been analysed (see eg., Simonius). A fundamental change in attitude began with the idea of Leggett et al. in the 1980's, that one could give detailed theoretical predictions for the quantum mechanics of large systems, and that SQUIDs could show macroscopic coherence properties (this was finally seen by Chiorescu et al. in 2003). This forced the well-known 'measurement problem' to become a serious topic of investigation in mainstream condensed matter physics; it also stimulated serious experimental tests of the validity of QM on the mesoscopic scale. Perhaps an even more fundamental change in focus was engendered in the discussion, mainly in the last 15 years, of large-scale entanglement, which is required for most kinds of quantum information processing. The quest for quantum communication devices, and for a workable quantum computer, has led to an avalanche of experimental work, in solid-state systems and in quantum-optical systems. As a result of these developments, it has become clear that we need a proper theory of decoherence, which explains not only general features like the connection to quantum measurements, the relationship to dissipative processes, and the possibility that there may be 'intrinsic decoherence' processes in Nature, but which also elucidates the detailed mechanisms involved in decoherence, and which can make quantitative predictions for the dynamics of decoherence in real systems.

This talk will begin by reviewing some of the history, and standard questions that arise, such as the relationship between decoherence and the 'classical' limit of QM, as well as to quantum measurements, dissipation, and so on. I will briefly discuss some recent ideas such as 'intrinsic decoherence' mechanisms, and '3rd party decoherence'. However the main focus of this talk will be on the mechanisms of decoherence arising in Nature, and the ways in which one can try to control or suppress them in the lab. I will emphasize decoherence in condensed matter systems, discussing how one can reduce the description of environmental decoherence to one of two models, in which the environment is described as either an oscillator bath or a spin bath. The implications for important contemporary problems in physics are discussed, with emphasis on solid-state qubits, and on 'quantum critical phenomena'.

Decoherence is often assumed to rule out coherence phenomena at high temperatures. Some of the most interesting examples of large-scale low-temperature quantum coherence will be referred to (involving SQUIDs and magnetic systems) [1]. However one can actually get remarkable examples of room-temperature coherence, even in condensed matter systems: I will make some brief remarks on decoherence in some biological systems, with specific reference to the light-harvesting molecules [2].

[1] *Experimental work on decoherence in solid-state systems is being reviewed in B Barbara's talk.*
[2] *Decoherence in biological systems will be discussed in detail in H Briegel's talk.*

READINGS:

1. PCE Stamp, "The decoherence puzzle", Stud. Hist. Phil. Mod Phys. **37**, 467 (2007) [included in reader]
2. M Simonius, "Spontaneous symmetry-breaking and blocking of metastable states", Phys Rev Lett **40**, 980 (1978)
3. HM Ronnow et al., "Quantum Phase transition of a magnet in a spin bath" Science **308**, 389 (2005) [included in reader—see Bernard Barbara's section]
4. E Collini et al., "Coherently wired light-harvesting in photosynthetic marine algae at ambient temperature" Nature **463**, 644 (2010).

ELSEVIER

Studies in History and Philosophy of Modern Physics 37 (2006) 467–497

Studies in History and Philosophy of Modern Physics

www.elsevier.com/locate/shpsb

The decoherence puzzle

P.C.E. Stamp[a,b,*]

[a]*Department of Physics and Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver BC, Canada V6T 1Z1*

[b]*Pacific Institute for Theoretical Physics, University of British Columbia, 6224 Agricultural Road, Vancouver BC, Canada V6T 1Z1*

Abstract

The understanding of decoherence is critical to philosophical debates on several different topics, including measurements, the 'emergence' of classical mechanics from quantum mechanics, and the arrows of time. This paper first reviews the basic mechanisms of decoherence in Nature, stressing recent discoveries and the crucial importance of 'low-energy' physics. The way in which the interpretation of some recent experiments relates to the problem is also delineated. Finally, some of the more common questions posed by philosophers about decoherence are reformulated, and partial answers are given to these. Throughout the article, the incomplete nature of our understanding is stressed, and the way it depends on several different unresolved questions in both low- and high-energy physics.

Keywords: Decoherence; Quantum environment; Quantum measurement

1. Introduction

'Decoherence' means different things to different people. To most physicists, phase decoherence is a fact of life, important throughout physics (and large parts of chemistry). For those interested in the foundations of quantum mechanics, and for historians and philosophers of physics, decoherence is interesting because of its connection to three main problems, viz., (i) the 'quantum measurement' problem; (ii) the 'emergence' of classical

*Corresponding author at: Department of Physics, Pacific Institute of Theoretical Physics, University of British Colombia, 6224 Agricultural Road, Vancouver BC, Canada. Tel.: +1 604 822 5711; fax: +1 604 822 5324.

E-mail address: stamp@physics.ubc.ca.

1355-2198/$ - see front matter © 2006 Published by Elsevier Ltd.
doi:10.1016/j.shpsb.2006.04.003

from quantum mechanics (and the hinterland between the two); and (iii) the arrows of time. The philosophical literature on decoherence over the last two decades has mostly focussed on what is sometimes called the 'decoherence programme',[1] viz., the effort to explain away problems like the three just mentioned as decoherence phenomena.

Curiously most physicists are not interested in either the decoherence programme or philosophical discussions of it—this in spite of their strong professional interest in decoherence phenomena. One reason for this is probably the empirical bent of most physicists, who quickly lose interest in grand 'scenarios' or 'programmes' when details are not forthcoming,[2] or when the scenario is not experimentally testable (necessarily the case for many discussions of the arrows of time (Halliwell et al., 1994; Savitt, 1995; Schulman, 1997; Zeh, 1989)). However there is also somewhat of a schism, between (a) physicists who feel that decoherence is a fairly trivial process, ubiquitous in physics, about which no interesting general statements can be made, and (b) those who feel it is highly non-trivial, but that meaningful discussion requires models that are both realistic and of some generality. The problem here is that most discussions of decoherence in the context of foundational problems (e.g., quantum measurements) have been based on simple idealised models. There is an obvious need for realistic models of complex macroscopic systems, if we are to address any of the three big questions mentioned above. This problem has occasionally been acknowledged by proponents of the decoherence programme. For example, Omnés (1994, Chapter 7), in his book does recognise some of the limitations of simplified models of decoherence—although this does not stop him from claiming some rather general results for macroscopic systems! For remarks on the validity of such results, see Section 4.

Ironically, in the last 20 years a quiet revolution has taken place in our understanding of the quantum mechanics of large systems, and of decoherence phenomena. The revolution is by no means complete, and we will see that several crucial problems remain to be solved. However we now have at hand many of the details missing from earlier discussions. As often happens, many early general assertions made on the basis of the idealised models can now be seen to be misleading, or just plain wrong. Despite this, the newer advances have had little impact on the philosophical literature. This is surprising and unfortunate, since the results do radically change our perspective on at least the first two questions mentioned above, and possibly also the third.

[1]For an introduction to the philosophical literature on decoherence, see Bacciagaluppi (2005). What is called the 'decoherence programme' by, e.g., Joos et al. (2003) and Zeh (2002), can actually be separated into various strands, depending on whether one is dealing with non-relativistic physics or quantum gravity, and on which question one is interested in (quantum measurements, the interpretation of quantum mechanics, large-scale quantum phenomena, cosmology, etc.). For extensive reviews, see Joos et al. (2003), Omnés (1992, 1994), Zeh (2002) and Zurek (2003), and for reviews of the 'decoherence histories' approach, see Griffiths (1984, 1986), Hartle (1991) and Gell-Mann & Hartle (1993). For discussions of the arrows of time, which touch upon the connections to decoherence, the quantum arrow, etc., see Schulman (1997), Zeh (1989), Halliwell, Pérez-Mercader, & Zurek (1994) and Savitt (1995).

[2]Typically what experimentalists are looking for is testable predictions—or at least something sufficiently precise and realistic that it can be related to some present or future class of experiments. Theorists are also looking for something quantitatively precise, which acquires much greater interest if it is both realistic (i.e., not oversimplified) and of some broad generality. Note that 'theoretical programmes' sometimes have a bad name in physics—an attitude summed up in Pauli's famous letter to Gamow in 1954 (referring to Heisenberg's 'programme' for a unified field theory). Writing "This is to show I can paint like Titian", he drew a simple rectangle, and then wrote "Only technical details are missing".

The purpose of this paper is to

(a) review quickly, for non-specialists,[3] what we now know about the physical mechanisms of decoherence, stressing the recent developments and their broader implications (Section 2). Then, in Section 3, I discuss what this means in the lab—how do experiments bear on the fundamental questions mentioned above? The main interest here for philosophers is to see just how much the interpretation of experiments depends on how one feels about the ultimate validity of quantum mechanics. For those wishing to follow up any of the themes mentioned in these two sections, I have given extensive references;
(b) in the light of the recent developments, to reconsider some of the more general questions mentioned above—concentrating on whether there exist 'intrinsic' sources of decoherence in Nature, how decoherence relates to the emergence of classical physics, to irreversibility and dissipation, to quantum measurements, and to the arrows of time (all in Section 4). The conclusion (Section 5) summarises where we are now.

2. Decoherence and quantum relaxation: models, mechanisms, dynamics

Discussions of decoherence usually begin with the interaction of a physical system $\mathscr{S}$ with an environment $\mathscr{E}$. One imagines that $\mathscr{S}$ starts off in some simple superposition of states, say $\Psi = \sum_j c_j \psi_j$, which upon interaction with $\mathscr{E}$, becomes entangled with it, so that the final state cannot be decomposed into a product state. Averaging over the environmental variables then produces a full or partial mixture, rather than a superposition of states, for $\mathscr{S}$. How this all happens in the real world is part of the 'decoherence problem', and it is interesting to see how views on this have evolved over the years.

Even before decoherence was discussed as such in the literature, mechanisms for it were being discussed in the context of the measurement problem, in the wake of the analyses of the 1930s of quantum measurements (London & Bauer, 1939/1983; Neumann, 1932/1955; Pauli, 1980). Early discussions of decoherence processes emphasised the role of randomisation of phases, and analysed this in terms of simple models of system/environment interactions, leading to irretrievable loss of phase correlations in the environment.[4] Very interesting ideas emerged from these discussions, including the possible role of amplification and relaxation, at least in measurements (Daneri et al., 1962, 1966), and the idea that the structure of interactions in the world might inevitably lead to decoherence in certain 'preferred bases'.[5] Simple models of decoherence were analysed in some of these papers, including Geiger counters, cloud chambers (an analysis going back

[3] It is assumed that the reader is familiar with elementary quantum mechanics. An intuitive understanding like that provided by Feynman (1965) is also useful.

[4] The idea of decoherence goes back at least to Ludwig (Born & Ludwig, 1958; Ludwig, 1953, 1958). Another early paper, concentrating on the role of the environment in the measurement problem, is Green (1958). These papers all argued that environmental dephasing (what we now call decoherence) would destroy large-scale quantum behaviour. This idea was picked up and further developed in Daneri, Loinger, & Prosperi (1962, 1966), Zeh (1970, 1973), Joos & Zeh (1985) and Simonius (1978), amongst others.

[5] The idea of preferred bases and preferred states, selected by decoherence, is described in, e.g., Simonius (1978), where these states are called 'inert states'; and in Zurek (1981, 1982), where they are called 'pointer states'. See also (Zurek, 2003).

to Mott), Stern–Gerlach experiments, sugar molecules, damped oscillators, etc. However, insofar as any macroscopic features of these examples were discussed, this was done in a very crude way, the aim being to demonstrate that decoherence would *always* suppress quantum interference effects except at the atomic or molecular scale (an assertion repeated (Van Kampen, 1988) as late as 1988).

This old orthodoxy was severely upset at the beginning of the 1980s by the now well-known work of Leggett et al. (Caldeira & Leggett, 1983; Leggett, 1984; Leggett et al., 1987), who pointed out that in fact one could expect superconducting SQUIDs to show quantum tunneling and interference properties at the macroscopic scale—and that moreover, one could test quantum theory at the macroscopic scale in this way. Initial scepticism has yielded in most quarters to the weight of experimental evidence—both macroscopic tunneling and coherence have now been seen in superconductors (see Section 3). In related developments, experimentalists have succeeded in the multi-particle entanglement of atoms in traps (Häffner et al., 2005; Leibfried et al., 2005), as well as superpositions of photon states in cavities (Zhao et al., 2004, and refs. therein); and new schemes, involving ideas like 'quantum non-demolition' measurements,[6] have been employed to reduce decoherence and dissipation effects in optical systems and in large Al bars (for gravity wave experiments).

Leggett et al. used an "oscillator bath" representation of the environment—a ploy first described by Feynman and Vernon (1963) and developed much further by Leggett et al. (Caldeira & Leggett, 1983; Leggett, 1984; Leggett et al., 1987). These models clearly lend themselves to problems in particle and string physics, quantum optics, and cosmology,[7] and they are also often used in condensed matter systems at low temperatures (Weiss, 1999). In contrast to the qualitative pre-1980 discussions of decoherence, we have a real theory, quantitatively testable on a large variety of systems. This completely changes the nature of both the scientific and the philosophical debates, as we shall see in the rest of this article.

However, in spite of this remarkable success, there is an important quantitative problem, particularly in solid-state systems—when one comes to compare the decoherence rates predicted by Caldeira–Leggett theory with the measured rates, the experimental rates are typically several orders of magnitude larger than theory predicts (see Section 3). This discrepancy indicates that most of the decoherence is coming from somewhere else, in ways not described by oscillator bath models. Whether this constitutes in some way a real problem of principle, particularly for tests of quantum mechanics at the macroscopic scale, is one of the topics addressed herein.

In Section 2.1 the main features of environmentally induced decoherence are explained, with an emphasis on the physical mechanisms responsible. Since there is a widespread belief that *all* decoherence is caused by direct interaction with an environment, in Section 2.2 I briefly outline another way decoherence can happen. The material of Section 2 is essential if one wishes to address the more philosophical questions associated with decoherence.

[6]For discussions of some novel measurement schemes, including quantum non-demolition schemes, see Braginsky & Khalili (1992) and Caves, Thorne, Drever, Sandberg, & Zimmermann (1980).

[7]For some examples of the use of oscillator bath models in cosmology and string theory see Cornwal & Bruinsma (1988), Callan & Freed (1992) and Callan, Felce, & Freed (1993, and refs. therein).

2.1. Quantum environments

2.1.1. Extended environmental modes; oscillator baths

Research on the dynamics of polarons and related problems led Feynman in the early 1960s to a general discussion of the interaction of a quantum system with its background environment. Feynman and Vernon (1963) considered the case where each environmental mode coupled only weakly to the central system. Arguing that for this weak-coupling case, the effect of *any* environment could be mapped to that of a set of oscillators, they treated a model Hamiltonian in which a central system $\mathscr{S}$, with generalised coordinates P, Q and Hamiltonian $H_0(P, Q)$, interacted with an environment $\mathscr{E}$ of oscillators with generalised coordinates $\{p_q, x_q\}$ and Hamiltonian $H_{\rm env}^{\rm osc}(\{p_q, x_q\})$, via a simple bilinear coupling:

$$H_{\rm eff}^{\rm osc} = H_0 + H_{\rm int} + H_{\rm env}^{\rm osc}, \quad H_{\rm int}(Q, \{x_q\}) = \sum_{q=1}^{N} c_q x_q Q. \tag{1}$$

We assume that the entire Hamiltonian $H_{\rm eff}(\Omega_0)$ is defined with an ultraviolet cutoff energy Ω_0. The important points to bear in mind here are:

(i) the oscillators have bosonic statistics, and typically represent delocalised modes, extending over the whole region of the environment. Typical examples are phonons, magnons, electron–hole pairs, or photons, which are wave-like oscillations of some background field. These are the low-energy modes of the environment—at higher energies the model usually breaks down;
(ii) the couplings $\{c_q\}$ are weak—in fact $c_q \sim O(N^{-1/2})$, where N is the number of low-energy environmental modes (N is thus proportional to the size of the environmental domain). This typically follows because we must normalise the oscillator wave functions (so they are $\sim O(N^{-1/2})$). Typically N is very big, so that mathematical treatments often just adopt the 'thermodynamic limit' $N \to \infty$. Since the effect of each oscillator to second order is $\sim |c_q|^2 \sim O(1/N)$, their total effect is then independent of N, as it should be in this limit. Thus each oscillator is only very weakly affected by the system, but the system may be quite strongly affected by the oscillators.

Curiously, the work of Feynman and Vernon had no impact whatsoever on the discussion of quantum measurements or decoherence for two decades—possibly because it was phrased in the then unfamiliar language of path integrals, and because the community working on the foundations of quantum mechanics was less interested at that time in detailed models.

At the beginning of the 1980s Caldeira and Leggett (1983) introduced a somewhat generalised Feynman–Vernon model, in which the coupling $\sum_q c_q x_q Q$ was replaced by

$$H_{\rm int}^{\rm osc} = \sum_{q=1}^{N} [F_q(Q) x_q + G_q(P) p_q]. \tag{2}$$

The Hamiltonians (1) and (2) are effective ones, which means amongst other things that the couplings c_q, F_q, and G_q, the oscillator frequencies ω_q, and even the system Hamiltonian H_0 depend not only on the UV cutoff Ω_0 but also on the bath temperature T. This may seem strange to some (particularly readers more at home with the models used in particle physics). Recall however that *all* Hamiltonians in physics are effective ones,

written in a quantum system in terms of operators defined over some restricted Hilbert space, depending implicitly or explicitly on energy cutoffs, temperature, and possibly other boundary conditions.[8] It is only when dealing with a very rarified medium that one can ignore these complexities.

Caldeira and Leggett gave arguments for the very general applicability of such effective Hamiltonians to systems at low energy (along with specific application to superconducting SQUIDs). Consider some arbitrary environment, with eigenstates $\phi_\alpha(\mathbf{X})$ and eigenenergies ε_α defined over the environment's full multi-dimensional coordinate space $\mathbf{X}$. Assume the system interacts with this environment via some interaction $V(Q,\mathbf{X})$. Then the arguments go as follows:

(a) Certainly we can recover an oscillator bath model if the coupling between different eigenstates induced by the interaction $V(Q,\mathbf{X})$ is weak, i.e., under the Feynman–Vernon condition that

$$|V_{\alpha\beta}| \ll |(\varepsilon_\alpha - \varepsilon_\beta)| \tag{3}$$

for all relevant environmental states, where $V_{\alpha\beta} = \int d\mathbf{X}\phi^*_\alpha(\mathbf{X})V(Q,\mathbf{X})\phi_\beta(\mathbf{X})$. The oscillator modes then correspond to the transitions between these states, and $\omega_q \equiv (\varepsilon_\alpha - \varepsilon_\beta)$.

(b) However, even if the weak-coupling condition is not obeyed, we can use a Born–Oppenheimer argument to derive a similar criterion. We first define adiabatic environmental eigenstates $\tilde{\phi}_\alpha(\mathbf{X},Q)$ and eigenenergies $\tilde{\varepsilon}_\alpha(Q)$, which depend on the instantaneous system coordinate Q. Now suppose that these states have a *fast* dynamics compared to the slower dynamics of the system coordinate Q (formally, that if Q moves on a frequency scale E_0, then $E_0 \ll \tilde{\varepsilon}_\alpha$). One then defines a fake 'gauge potential' $A_{\alpha\beta}$, describing the effect of the slowly changing Q on the bath modes, given by $iA_{\alpha\beta} = \int d\mathbf{X}\tilde{\phi}^*_\alpha(\mathbf{X})\partial/\partial Q\tilde{\phi}_\beta(\mathbf{X})$; there is no reference to the original interaction between Q and the bath modes, because this has already been incorporated into the renormalised $\tilde{\varepsilon}_\alpha$. Standard manoeuvres then show that we can make a mapping to an oscillator bath provided

$$|A_{\alpha\beta}| \ll |(\tilde{\varepsilon}_\alpha - \tilde{\varepsilon}_\beta)| \tag{4}$$

for all the relevant modes. If (4) is satisfied, then the oscillators now describe transitions between the new adiabatic bath modes, with frequencies $\omega_q \equiv (\tilde{\varepsilon}_\alpha - \tilde{\varepsilon}_\beta)$; and one can also derive the couplings F_q, G_q in terms of the gauge coupling in (4).

(c) Leggett et al. then argued that the low-T, low-energy quantum dynamics of such a system could be related to its higher T dissipative classical dynamics (cf. Fig. 1). From the classical dissipative dynamics one *infers* a low-energy effective Hamiltonian (having the form (1), with the generalised interaction in (2)); in particular, one finds the form of the couplings in (2). This is crucially important—instead of trying to derive the form of H_{eff} from some theory (a move which is always open to criticism given the

[8] The idea of the 'effective Hamiltonian' (or the effective Lagrangian) is rather subtle, and bound up in the recent history of physics with the idea of the renormalisation group (although discussions go back at least to the 19th century). See, e.g., Anderson (1984). For a recent discussion of the effective Hamiltonian, directed to a philosophical readership, see Stamp (to be published).

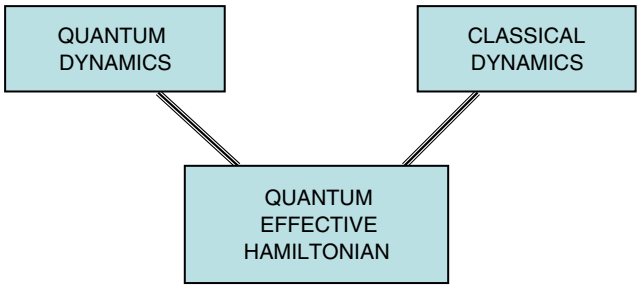


Fig. 1. The epistemological connection between the observable classical (usually high-temperature) dynamics and the low-temperature quantum dynamics (often not so easy to observe), for a system with many degrees of freedom. Both can be derived from the correct quantum effective Hamiltonian. Often (as in the approach of Caldeira and Leggett) one infers the quantum Hamiltonian from experiments on the classical behaviour.

huge complexity of large systems), one instead infers it directly from experiment.[9] One then *derives* the quantum dynamics of the system from this effective Hamiltonian.

At first glance the assumptions behind the oscillator bath model seem restrictive—small oscillations and weak coupling to each mode, use of a Born–Oppenheimer approximation, etc. However appearances are deceptive—oscillator bath models are quite robust in the real world. A large class of effective Hamiltonians (sometimes called a *universality class*), which will describe many physical systems, can be mapped to models of the oscillator bath type (Dubé & Stamp, 2001). Examples include: (i) itinerant fermion baths (e.g., a bath of interacting conduction electrons), in three, two or one dimensions; (ii) systems having weak higher-order 'anharmonic' couplings to extended bath modes—these can be absorbed into modified couplings to a new set of oscillators (the couplings and oscillator frequencies now being very strongly T-dependent); and (iii) systems where bath modes are strongly coupled to the system, provided the condition (4) is not violated (i.e., provided the effective coupling between two environmental states goes to zero fast enough as one reduces the energy difference between them). It is worth remarking here on a point which is crucially important for decoherence. The reduction in the strength of coupling to oscillator bath modes at low energies is a general feature of extended environmental states, whose density of states always goes down with energy, because of decreasing available phase space volume. This means that at as one lowers energies and temperatures towards zero, we can naively expect the decoherence from oscillator baths to also decrease to zero.

We have seen that oscillator bath models of quantum environments are thus much more general than is often assumed in the literature. However they certainly cannot always work, and they clearly fail in many solid-state systems at low temperatures. In order to understand why, we make a little diversion into the real world of low-energy physics.

[9]In Caldeira–Leggett theory, the interaction between system and environment is summarised in a 'spectral function' $J(\omega, T)$, a function of frequency and temperature. If the Caldeira–Leggett effective Hamiltonian applies to some physical system, *and* if one knows $J(\omega, T)$, then the behaviour can be derived theoretically in both classical and quantum regimes. More typically, one *infers* $J(\omega, T)$ from the classical and/or quantum behaviour in experiments.

2.1.2. Interlude: real condensed matter

Condensed matter is all around us—we are directly aware of little else. All measuring systems are made from condensed matter. It is clearly messy, and complex structures and order are evident everywhere (not least in living things). As a result, except for the He liquids (which go superfluid at low T and which can be made in essentially completely pure form) and rarified gases and plasmas, the low-energy effective Hamiltonians of *real* condensed matter systems are extraordinarily subtle (and very far from the descriptions usually given in student textbooks). There is a common misunderstanding that these subtleties have to do with 'dirt' effects (the 'squalid state', in Pauli's famous phrase). In fact they are mostly intrinsic, for the following reasons (footnote 8):

(i) *Topology.* Many-particle wave-functions have topological properties which restrict and sometimes control the dynamics. This often leads to new branches of low-energy 'topological excitations', with their counterpart in the effective Hamiltonian (Thouless, 1998).
(ii) *Lattices + interactions.* In solids, electrons are constrained to move between different atomic orbitals. Strong repulsive interactions between electrons can prevent more than one particle per orbital, imposing a highly non-trivial structure on the Hilbert space of the effective Hamiltonian and even causing the low-energy states to localise.
(iii) *Boundaries or edges.* All systems have boundaries. In conjunction with long-range forces and/or the topological properties of wave-functions, the boundaries and the states localised near them can control the low-energy properties of the whole system.
(iv) *Frustration.* Interactions between two different pairs of particles or spins are often 'incompatible' (i.e., lead to contradictory effects on any one of the particles). The result is typically a large number of almost degenerate low-energy states which hardly communicate.[10] The system can never reach its putative ground state (which then becomes a mere mathematical chimera). Because of frustration, most pure solids, without impurities, are intrinsically disordered. States pile up at low energies—many of these low-energy states are localised (footnote 10).

Clearly none of these effects come from 'junk' or 'dirt'; moreover, because they arise from very general mechanisms, they lead to effects that are ubiquitous in low-temperature experiments. These include peculiar structure in the low-energy density of states, complex and often non-linear long-time relaxation phenomena, including 'glassy' behaviour (the freezing out of dynamics caused by frustration), increasingly subtle kinds of quantum ordering as one lowers the temperature, etc. Over the last four decades a phenomenological description has emerged for these low-energy phenomena, in terms of a set of low-energy discrete modes (i.e., each having a discrete finite set of states, often only two, in the energy range of interest), appropriate to localised states (Anderson, 1994; Binder & Young, 1986; Esquinazi, 1998; Mézard et al., 1987). These states interact both amongst themselves, and with the extended 'oscillator modes'. Thus one ends up with a low energy description in terms of a set of interacting 'two-level systems'; usually the interactions are fairly weak,

[10]The only elementary review of some of the low-energy complexities in real solids seems to be the five short articles by Anderson on 'spin glasses' (Anderson, 1994). More sophisticated reviews are by Binder & Young (1986) and Mézard, Parisi, & Virasoro (1987); this latter book also makes the connection with work in computation and biology.

although they can have important effects. There is certainly no universal agreement about this picture (Yu & Leggett, 1988), but in many cases there is extensive evidence that it gives a good description of the low-energy physics (Anderson, 1994; Binder & Young, 1986; Esquinazi, 1998; Mézard et al., 1987). I emphasise again that these effects are pretty much universal in solids, although their effects are sometimes not obvious until very low temperatures. Their effects on ordinary transport and other dissipative properties can be very small (making them almost invisible at higher temperatures), but we shall see that their contribution to decoherence can be very large.

One is often met by surprise at this situation. How, it is asked, can a simple solid show such 'pathological behaviour', when after all it is made up electrons, protons, etc., which can be described by a simple continuum theory having none of these complexities? The fallacy in this argument is the assumption that the effective Hamiltonian of a composite system will somehow be analysable into that of its constituents.[11] This is not true—the effective theory of the constituents is still an effective theory, applicable only in a certain energy range and assuming a restricted Hilbert space. For this reason neither the vacuum nor the low-energy eigenstates of the high-energy Hamiltonians used in particle physics look anything like a condensed matter system (even though this is physically what a high-energy system becomes if it is cooled!). In many real solids, an infinite hierarchy of effective Hamiltonians, ever more complex, is expected as one lowers the energy scale (footnote 8), and we only have a dim understanding of what their structure might be. In other words, we do not really understand the basic structure of the lowest energy states or Hilbert spaces of most many-body systems. An understanding of this low-energy structure is one of the holy grails of condensed matter physics—in many ways it seems more elusive now than it did 30–40 (or even 100) years ago. One hundred years ago, with the vindication of the atomic hypothesis, but before quantum mechanics, a simple reductionist view of condensed matter looked very reasonable. Thirty to forty years ago, a unification of methods between quantum field theory and condensed matter physics looked imminent—the Ginzburg–Landau–Wilson theory of phase transitions, and the BCS theory of superconductivity, were shaping much of modern particle theory. This unification has happened, but only in the study of 'simple' systems. For a more realistic perspective see Anderson (1994), Binder and Young (1986), Esquinazi (1998) and Mézard et al. (1987).

If some day we ever have a ''complete theory of everything'', with a 'universal Hamiltonian' whose eigenstates (including the ground state) represent the real states of the universe, over all energy scales, then we would presumably find that the low-energy states of this Hamiltonian contain the full complexity of real condensed matter. Right now we have little idea if such a theory would even be meaningful (it is perhaps more likely that the whole Hamiltonian structure will be replaced by something more fundamental). We certainly have not the slightest idea whatsoever what it would look like. Current efforts towards progress range from theory at supra-Planck scale energies, to the exploration of coherence phenomena at temperatures below 10^{-9} K.

[11] It is commonly argued that the 'complexity' of low-energy physics comes only from the large number of constituents (this is certainly the point of view of 'reductionists'). This argument is refuted in a well-known paper by Anderson (1972), which inspired a very large subsequent literature.

2.1.3. Localised modes: spin baths

We return now to the question at hand, which is to understand the sources of decoherence at low energies in real condensed matter systems. The importance of the previous discussion is that we now see we must deal with the large number of low-energy localised states existing in solids, or more generally, low-energy modes having a finite Hilbert space, with discrete excitations. The general nature of these was described above; they include the eigenstates of nuclear spins, of topological defects, and of various more subtle modes associated with frustration, boundaries, and intrinsic disorder. In any real system there will also be 'junk' effects, coming from paramagnetic impurities, 'charge trap' excitations, etc. In many systems we may not know exactly what these discrete modes are, but as noted above, their presence is often very obvious in experiments (Anderson, 1994; Binder & Young, 1986; Esquinazi, 1998; Mézard et al., 1987).

Now one can always map a system having a set of M discrete states to a spin system, with spin σ, such that $2\sigma + 1 = M$. Thus we can in all cases describe an environment of these states as a '*spin bath*' (Prokof'ev & Stamp, 2000). Spin baths have the following general characteristics:

(i) The generic model for a quantum system interacting with a spin bath (corresponding to the generic oscillator bath model defined by Eqs. (1) and (2)) has the effective Hamiltonian:

$$H^{\rm sp}_{\rm eff}(\Omega_0) = H_0 + H^{\rm sp}_{\rm int} + H^{\rm sp}_{\rm env}, \tag{5}$$

where $H_0(P, Q)$ describes the system as before; but now the interaction term is a vector coupling to a set of 'spins' $\{\boldsymbol{\sigma}_k\}$ (which for simplicity we take here to be two-level systems, i.e., spin-$\frac{1}{2}$ systems):

$$H^{\rm sp}_{\rm int} = \sum_{k}^{N_s} \boldsymbol{F}_k(P, Q) \cdot \boldsymbol{\sigma}_k, \tag{6}$$

and the spin bath Hamiltonian itself has the form:

$$H^{\rm sp}_{\rm env} = \sum_{k}^{N_s} \mathbf{h}_k \cdot \boldsymbol{\sigma}_k + \sum_{k,k'}^{N_s} V^{\alpha\beta}_{kk'} \sigma^{\alpha}_k \sigma^{\beta}_{k'}, \tag{7}$$

with a set of external fields $\{\mathbf{h}_k\}$, and interspin interactions $V_{kk'}$. The generalisation of this model to bath modes having $M > 2$ discrete states is straightforward.

(ii) Each bath 'spin' interacts only weakly with its compatriots—formally we require that $\{|\boldsymbol{F}_k|\} \gg |V_{kk'}|$. If the $\{\boldsymbol{\sigma}_k\}$ describe localised modes, this is quite typical. The different bath excitation wave-functions do not overlap and can only communicate via weak long-range interactions $V_{kk'}$, whereas there is nothing limiting the size of the $\{|\boldsymbol{F}_k|\}$ (which are no longer $\sim O(1/N^{1/2})$). The bath dynamics is then under the direct control of the central system (note that inequality (4) is now violated), with its own 'intrinsic dynamics' playing second fiddle. Recall that this is exactly opposite to the oscillator bath system, where the intrinsic dynamics of the oscillator bath is only weakly perturbed by the central system, because the oscillator frequencies $\{\omega_q\}$ are much larger than either the $\{c_q\}$ or the F_q, G_q in (2). This situation is illustrated in Fig. 2.

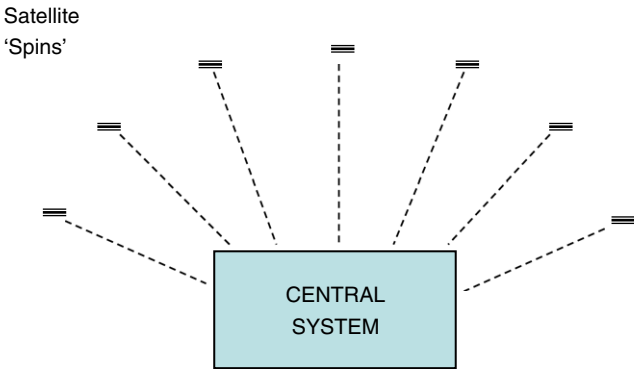


Fig. 2. The 'spin bath' environment—a set of satellite spins couples to the central quantum system of interest. The spins typically represent localised modes (not necessarily spins!) in the environment, each with a finite Hilbert space (often two-dimensional). The coupling between spins is weak compared to the coupling of each to the central system.

Clearly under some circumstances we can map the spin bath onto an oscillator bath. For example, if the interactions $V_{kk'}$ are strong (i.e., if $\{|\boldsymbol{F}_k|\} < |V_{kk'}|$ and if $|\mathbf{h}_k| < |V_{kk'}|$), then the bath spins can couple together to form extended 'spin waves', and $H_{\rm eff}^{\rm sp}(\Omega_0)$ then maps back to a Caldeira–Leggett model. If the central system dynamical energy scale $E_0 \gg |\boldsymbol{F}_k|$, then one goes to an anti-adiabatic (or 'anti-Born–Oppenheimer') limit, in which the system–bath couplings can be treated perturbatively. One can give more complete criteria for the mapping of spin baths to oscillator baths (Prokof'ev & Stamp, 2000), which we see must also involve the static fields $\{\mathbf{h}_k\}$.

In real physical systems the coupling energies $|\boldsymbol{F}_k|$ and static field strengths $\{|\mathbf{h}_k|\}$ are often spread over a very wide range, particularly in systems with frustration, disorder or impurities (note that 'impurities' include nuclear spins, which are almost everywhere; they live in some finite fraction of the nuclei of almost all the elements in solids). We cannot then use either or a Born–Oppenheimer or an anti-Born–Oppenheimer approximation, there are many environmental modes which must be treated directly as localised modes. Because these modes then have characteristic frequencies similar to those of the central system we are interested in, they cause a lot of decoherence.

2.2. Bath-induced decoherence and relaxation

Although the detailed calculation of the dynamics of decoherence is a complicated business, many of the main points can be understood by simple (although qualitative) arguments.

As noted earlier, in the early development of this subject, the idea of decoherence was very much bound up with quantum measurements. Decoherence was, in effect, viewed as a process in which the environment $\mathscr{E}$ 'measured' the state of the system $\mathscr{S}$ being decohered, via a transition

$$\sum_j c_j \psi_j \Phi_0 \;\rightarrow\; \sum_j c_j \chi_j \Phi_j, \tag{8}$$

for the combined $\mathscr{S}\otimes\mathscr{E}$, with the final states $\{\Phi_j\}$ of $\mathscr{E}$ uniquely correlated to the original system states $\{\psi_j\}$. The superposition still exists in the combined state of $\mathscr{S}\otimes\mathscr{E}$, but tracing out the environment gives a reduced density matrix $\rho^S_{jj'}=|c_j|^2\delta_{jj'}$ for $\mathscr{S}$, in which all correlations between the ψ_j (i.e., all off-diagonal matrix elements in $\rho^S_{jj'}$, in this particular Hilbert space basis) have disappeared. Recall that the infamous 'measurement problem' (d'Espagnat, 1976; Wheeler & Zurek, 1983) centres around states like (8).

Any real environment would not align its states so exactly with those of $\mathscr{S}$; such precise correlations can only be engineered by an experimentalist, by deliberate 'state preparation' (Margenau, 1973a,b). Nevertheless over the last 50 years there has been great interest in how the environment might cause $\rho^S_{jj'}$ to diagonalise in certain preferred bases (in the context of measurement theory these states are called the 'inert' or 'pointer' states (footnote 5)). The most popular is the basis set $\{Q\}$ of position eigenstates $\{Q\}$, and one assumes that $\rho^S_{Q,Q'}$ tends over some timescale to a function which is narrowly focussed around $\delta(Q-Q')$ (e.g., a Gaussian function $\rho^S_{Q,Q'}=(1/2\pi\sigma^2)^{1/2}\exp[-(Q-Q')^2/2\sigma]$, with small variance σ^2). This certainly can happen in simple models. The easiest way is to couple some bath to $\mathscr{S}$ with a coupling linear in the system coordinate Q, and the model of a central oscillator coupled bilinearly to a bath of oscillators has been the object of many papers (which usually assume an Ohmic coupling) like Grabert, Schramm, and Ingold (1988).

However these models and their behaviour lack generality, as we will see. Before continuing, it is useful to give some intuition for the dynamics of decoherence, i.e., the time evolution of ρ^S. This is described by a propagator K, which relates the density matrix at some time t_2 to its state at an earlier time t_1. Now let us go to a particular basis, the position basis, which allows us to look at how K evolves in real space, using the highly intuitive path integral formulation of quantum mechanics (Feynman & Hibbs, 1965; Feynman, Leighton, & Sands, 1965). As discussed by Feynman, one can usefully write (see also Feynman & Vernon, 1963):

$$K(Q_2,Q_2';Q_1,Q_1';t,t')=\int_{Q_1}^{Q_2}\mathscr{D}q\int_{Q_1'}^{Q_2'}\mathscr{D}q'\,\mathrm{e}^{-\mathrm{i}/\hbar(S_0[q]-S_0[q'])}\mathscr{F}[q,q'], \tag{9}$$

where $\mathscr{F}[q,q']$ embodies all the effects of the bath on the dynamics of $\mathscr{S}$, after we have averaged over the bath. To interpret (9), suppose first that $\mathscr{F}[q,q']=1$, i.e., that the system $\mathscr{S}$ is completely decoupled from the bath, and propagates freely. Then K propagates along two paths $q(\tau)$ and $q'(\tau)$ between the limiting arguments, and in the usual quantum way, one sums over all possible pairs of paths. Thus $\mathscr{F}[q,q']$ is just a *weighting factor*, defined over these two paths, and it couples them. Moreover, $\mathscr{F}[q,q']$ has a simple form; one can always write $\mathscr{F}[q,q']=\exp[-\mathrm{i}\Phi-\Gamma]$, where the phase $\Phi[q,q']$ and 'damping' $\Gamma[q,q']$ are real.

Now suppose, for example, that $\mathscr{F}[q,q']$ falls off rapidly when the paths q and q' move apart from each other. Then the density matrix will be forced towards an approximate 'pointer basis' in Q-space. Many other behaviours are also possible. The advantage of dealing with $\mathscr{F}[q,q']$ is that it can also be connected in a transparent way with (and calculated from) the effective Hamiltonian, and the behaviour of $\mathscr{F}[q,q']$ is easily visualised.

Decoherence and relaxation in oscillator bath models. The essential properties of $\mathscr{F}[q,q']$ for a system in contact with an oscillator bath were defined by Feynman (Feynman &

Hibbs, 1965; Feynman et al., 1965; Feynman & Vernon, 1963), and as noted above, Caldeira and Leggett were able to relate the quantum and classical dynamics of $\mathscr{S}$ for such environments. This crucial step allows us to understand decoherence for such models in the same terms as we understand ordinary dissipation.

In this way one arrives at the following intuitive picture. The 'fast' environmental modes cause little decoherence or dissipation. They simply adapt their dynamics to the much slower system dynamics, and their main effect is simply to renormalise this slow dynamics (i.e., change its frequency scale somewhat). Dissipation and decoherence both arise from energy exchange between system and bath, either because of thermal or quantum fluctuations in the bath (one can also think of this as a combination of stimulated and spontaneous emission/absorption processes). The connections established for oscillator baths between decoherence and dissipation and between the classical and quantum dynamics, are amongst the most remarkable results derived by Caldeira and Leggett.

The crux of the connection between decoherence and relaxation in oscillator bath models lies in the weak-coupling assumption (i.e., the assumption that each bath mode is only weakly perturbed by the central system; note again that the system itself may be very strongly affected by the combined effect of all the bath modes). Decoherence and dissipation are arising then from the same second-order processes, in which a single bath mode intervenes to exchange energy with the system (foonote 9). By summing over all the bath modes, one rapidly introduces standard results for a situation like this (Leggett, 1984; Caldeira & Leggett, 1983). The 'fluctuation–dissipation' theorem, connecting the fluctuation spectrum of the bath with the dissipation it causes on the system, is an immediate consequence of the weak-coupling assumption (the derivation uses 'linear response' theory). We can thereby connect the decoherence and dissipation in the system directly to the noise spectrum of the environment.

To give a quick intuitive picture of all this, let us pick a really simple central system $\mathscr{S}$, viz., a 'qubit', the elementary component of a quantum computer, which we then couple to an oscillator bath. This model (known as the 'spin–boson' model) has been studied rather exhaustively (Leggett et al., 1987; Weiss, 1999). The qubit itself can be described by a Pauli spin-$\frac{1}{2}$ vector $\boldsymbol{\tau}$, and we choose a Hamiltonian $H_0(\boldsymbol{\tau}) = \Delta_0\tau_x + \varepsilon\tau_z$. Working in the basis of the eigenstates $|\uparrow\rangle$ and $|\downarrow\rangle$ of τ_z, the propagation of the density matrix in time can then be visualised very easily in path integral language (see Fig. 3). The qubit simply 'flips' back and forth between $|\uparrow\rangle$ and $|\downarrow\rangle$. This happens on a very short timescale $\sim 1/\Omega_0$, governed by the high-energy physics of the qubit at energy $\sim\Omega_0$. If we now add the coupling $\sum_q c_q\tau_z x_q$ to the oscillators (the analogue of the Feynman–Vernon coupling $\sum_q c_q Q x_q$ to Q), the effect of the bath is to allow 'second-order' interaction processes between the bath and qubit,[12] with the oscillators shown as wavy lines.

The qubit-bath coupling distinguishes the states $|\uparrow\rangle$ and $|\downarrow\rangle$, so that the bath is in effect 'watching' the qubit. It is not surprising that the general effect of this coupling is to slow down the qubit dynamics and to degrade coherence (i.e., superpositions) between $|\uparrow\rangle$ and $|\downarrow\rangle$. In path integral language, attractive interactions are generated between the jumps, both on the same path and between paths; this causes them to bind together and thereby disappear (thereby making jumps less frequent and also suppressing 'off-diagonal' states

[12]Each wavy line in the figure represents the emission and absorption of a bath excitation, and is thus second order in the interaction. Multiple interaction lines appear in the figure because the influence functional in the path integral (Eq. (9)) is an exponential function of the interaction.

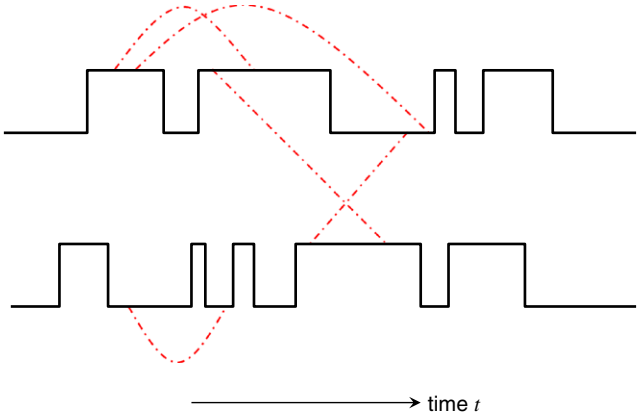


Fig. 3. The behaviour in time t of the density matrix of a simple two-level system (a 'qubit') which is interacting with a bath. The density matrix always involves two paths (the 'forward' path 1 and 'return' path 2); each switches between the two available qubit states as time goes on. The qubit–bath coupling mediates interactions between paths (hatched lines), as well as 'self-energy' interactions between states of the same path at different times.

corresponding to interference). If the qubit-bath interaction is strong enough, this can even cause all transitions to disappear, and the qubit is then frozen by the bath. The dissipative slowing down can be related directly to the decoherence rate, and both can be related to the quantum fluctuations of the bath, as we expect from the remarks made above.

All this is very much what one might expect from the bath. Study of other models, such as a moving particle coupled linearly to the bath according to (1) or (2), and either in a homogeneous medium or in some potential well (or tunneling from it), give similar results.

Finally, we re-emphasise the point already made above that as one lowers the bath temperature and the operating energy scale Δ_0 of the qubit, the decreasing available phase space for transitions in the bath states means that decoherence also falls rapidly (particularly rapidly in insulators) and eventually goes to zero, as T and Δ_0 go to zero. High-energy bath modes cause little decoherence (one has to be careful to distinguish simple renormalisation effects caused by these modes from genuine decoherence (Unruh, 1999)). Under certain circumstances, one can then *engineer* the oscillator bath environment to have very few low-energy states (for example, in a superconductor one has a gap in the low-energy spectrum). In this case we expect very little decoherence from the oscillator bath.

Decoherence from spin baths. Decoherence works in a very different way for a spin bath, and the differences with oscillator bath decoherence are very illuminating (Dubé & Stamp, 2001; Prokof'ev & Stamp, 2000). Consider a particular bath spin $\boldsymbol{\sigma}_k$. Its dynamics is controlled by (i) a *static* field $\mathbf{h}_k$, and (ii) a dynamic field $\boldsymbol{F}_k(P,Q)$ caused by the central system, whose state is evolving in time. The interaction $V_{kk'}$ with other spins is a small perturbation on this.

How $\boldsymbol{\sigma}_k$ actually evolves in time depends on a third energy/frequency scale—the rate $\dot{\boldsymbol{F}}_k(P,Q)$ at which $\boldsymbol{F}_k(P,Q)$ is changing. This rate is controlled by the dynamics of the central system $\mathscr{S}$. If the characteristic frequency scale for changes of $\boldsymbol{F}_k(P,Q)$ is Ω_0, so that $|\dot{\boldsymbol{F}}_k/\boldsymbol{F}_k| \sim \Omega_0$, then we have two limits, viz., (a) if $u_k \ll 1$, where $u_k = |\boldsymbol{F}_k|/\Omega_0 \sim |\boldsymbol{F}_k^2/\dot{\boldsymbol{F}}_k|$, the

system is moving too fast for $\boldsymbol{\sigma}_k$ to follow the field $\boldsymbol{F}_k$ (the 'fast' or 'sudden' limit); whereas (b) if $u_k \gg 1$, the system dynamics is slow, and $\boldsymbol{\sigma}_k$ can track $\boldsymbol{F}_k$ fairly accurately (the adiabatic limit).

To get an idea of how decoherence works here, imagine that $\mathscr{S}$ is a qubit $\boldsymbol{\tau}$, so that its typical paths are those just discussed for the spin–boson model. The coupling between $\boldsymbol{\sigma}_k$ and $\boldsymbol{\tau}$ can be assumed quite generally to take the form $H_{\rm int} = \hat{\tau}_z \boldsymbol{\omega}_k \cdot \boldsymbol{\sigma}_k$, so that the field on $\boldsymbol{\sigma}_k$ from the central qubit, when it is in state $|\Uparrow\rangle$, is $\boldsymbol{\omega}_k$, and this field reverses its direction when the qubit flips. Note that there is also another *static* field $\mathbf{h}_k$ acting on $\boldsymbol{\sigma}_k$ (cf. Eq. (7)); the total field is the sum of the two (we ignore the interaction $\{V_{kk'}\}$ for the moment).

How this then affects the dynamics of our bath spin $\boldsymbol{\sigma}_k$ is shown schematically in Fig. 4, which shows a typical path for the bath spin (*not* for the qubit!). Each time $\boldsymbol{\tau}$ flips, the total field jerks suddenly (on a time scale Ω_0^{-1}) between two orientations. Suppose the bath spin begins in a state oriented along one of these fields. A qubit flip then starts it precessing about the new field; in general when $\boldsymbol{\tau}$ flips back the bath spin will be oriented in some other direction, and it will begin to precess anew around this field. We see immediately that: (i) the bath spin's dynamics is now *entangled* with that of the qubit. In particular it accumulates a 'precessional phase' that depends on the qubit path; and thence (ii) averaging over the bath spins now gives decoherence in the dynamics of the central system $\boldsymbol{\tau}$, this is called 'precessional decoherence'.

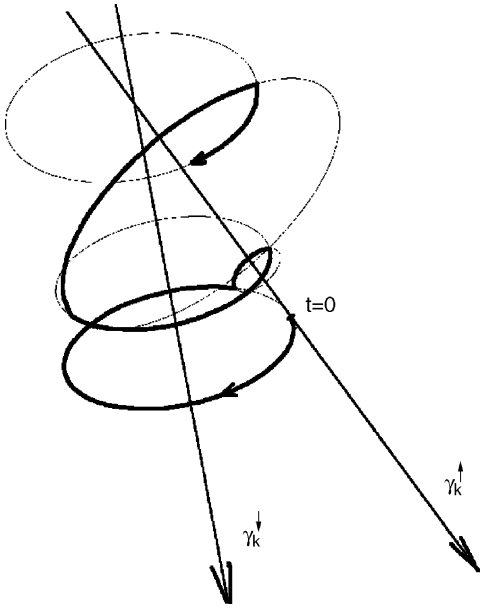


Fig. 4. The dynamics of a bath spin (the 'kth' bath spin) under the influence of a qubit. The qubit exerts fields $\boldsymbol{\gamma}_k^{\uparrow}$ or $\boldsymbol{\gamma}_k^{\downarrow}$, depending on whether the qubit is in state $|\uparrow\rangle$ or $|\downarrow\rangle$. Each time the qubit flips, the bath spin must begin to precess in the new field, causing the bath spin trajectory shown. The dependence of this trajectory on the qubit trajectory means that their quantum dynamics are strongly entangled, even though no energy is exchanged.

Precessional decoherence is almost always the most important mechanism of decoherence coming from a spin bath. However it has very different characteristics from that coming from oscillator baths. In particular

(i) Notice that *no energy transfer* between system and bath is involved in precessional decoherence; no transitions in the bath spin state occur.[13] Nevertheless very strong decoherence can occur—if the total phase accumulated by all the bath spins over a time τ is $\gtrsim 2\pi$, then averaging over this phase will give very strong decoherence in the qubit dynamics over this time.
(ii) The bath spin dynamics is being driven by the qubit. If the qubit dynamics is switched off, only the very weak interspin interactions $\{V_{kk'}\}$ can drive the bath spins. Thus the bath spin dynamics is largely *slaved* to the qubit dynamics and will slow down drastically if the bath spin dynamics is frozen. Again, this is totally different from an oscillator bath, whose internal dynamics is only weakly affected by coupling to the central system. Incidentally this means that the intrinsic noise coming from the spin bath has little connection to the decoherence.

These results underline the fact that decoherence caused by a quantum environment is really about phase exchange between system and environment, and has no necessary connection with either environmental noise or dissipation at all. It can proceed in the complete absence of either. Some physicists, used to the framework of linear response and fluctuation dissipation theorems, are quite surprised by this. It is important to remember the limitations of the fluctuation–dissipation framework—it only works if the bath is weakly perturbed by the central system (or by some probe). In the present case the spin bath will only obey linear response and the fluctuation–dissipation theorem if it is weakly perturbed, which is precisely the point at issue here. Indeed, it should be remarked that all the standard ideas about linear response are more and more difficult to apply as one lowers the temperature, since ever smaller perturbations will take the bath outside the linear response regime.

This naturally leads one to ask how decoherence from spin baths behaves as one goes to the low temperature limit. In contrast to oscillator baths, we can no longer assume that the oscillators will go away in the low-energy limit. Indeed, as remarked above, localised states tend to *pile up* at low energies, in many solids. Thus one can expect very large contributions to decoherence from these states. Thus on purely theoretical grounds one can expect that spin bath decoherence will dominate over oscillator bath decoherence at low temperature. There is no reason to expect it to go to zero, even as $T \to 0$. Thus we see that the intuitive connection between decoherence, dissipation, and environmental noise, all gained from the oscillator bath models, is in no way generic to decoherence.

Nevertheless very surprising features can emerge. For example, a particle hopping quantum mechanically around a lattice of some topology will, if coupled to an oscillator bath, *always* tend at long times to show diffusive dynamics, and this feature is often cited as an example of the inevitable crossover of quantum behaviour to classical stochastic

[13] We omit here discussion of 'topological decoherence' from the spin bath, which is usually much weaker than precessional decoherence (and also causes no dissipation). See, e.g., Prokof'ev & Stamp (1993, 2000). Note that the analogue of this can exist in special oscillator bath models where the longitudinal coupling $c_q x_q \tau_z$ is absent, but a transverse coupling like $c_q^{\perp} x_q \tau_x$ is present—again there will be no dissipation.

behaviour in the presence of decoherence. But one can show (Prokof'ev & Stamp, 2006) that if the particle is instead coupled to a spin bath, this is not so. In fact, in the long time limit, there will always be some part of the particle density matrix which still shows quantum interference behaviour. Since such lattices can be used as very general models for the propagation of quantum information (Kempe, 2003; Kendon, 2003), this result is of some importance.

Finally, I emphasise that none of these results are mere remarks about abstract models. As we will note in Section 3, they are crucial for recent experiments.

2.3. Third-party decoherence

The literature on decoherence deals entirely, as far as I know, with environmentally induced decoherence, in which phase correlations (and possibly also energy, etc.) are transferred from system to environment by some physical coupling between them. The purpose of this section is cautionary, to emphasise that entanglement between system and environment can be set up without such a direct coupling, or even an indirect one. The basic idea discussed briefly here (more details appear elsewhere[14]) is that of 'third-party decoherence', in which decoherence emerges eventually in the dynamics of some system $\mathscr{S}$, not via any direct coupling to the environment $\mathscr{E}$, but through the influence of a third party.

Clearly there is a trivial way in which phase correlations can be set up between a system $\mathscr{S}$ and an environment $\mathscr{E}$, even when they are not directly coupled. One couples $\mathscr{S}$ to a 'third party' $\mathscr{P}_3$ which is in contact with (or is later brought into contact with) $\mathscr{E}$. Thus phase correlations, entanglement, etc., pass through the chain $\mathscr{S} \rightarrow \mathscr{P}_3 \rightarrow \mathscr{E}$. However this is clearly not a fundamentally new situation. Theoretically, we can simply expand our original environment to a new environment $\mathscr{E}' = \mathscr{E} + \mathscr{P}_3$. The details may be non-trivial and important for experiment, and interesting things may happen, since the entanglement held between $\mathscr{S}$ and $\mathscr{P}_3$ may take some time to reach $\mathscr{E}$ (particularly if $\mathscr{P}_3$ is only later allowed to interact with $\mathscr{E}$). One can also extend this chain to include fourth, fifth, etc., parties.

There are however more interesting kinds of third-party decoherence. Consider as an example the famous two-slit experiment, in which particles pass through two slits and an interference pattern is produced on a screen. There will be simple decoherence mechanisms here, in which the particle interacts with a photon bath (the dipolar EM interaction allows photons to track the particle path $\mathbf{Q}(t)$) or even phonons emitted by the particle if it collides inelastically with the slit system on its way through. However a more subtle effect can arise if the particle itself possesses internal degrees of freedom $\{x_l\}$, which themselves do not interact with the particle centre of mass coordinate $\mathbf{Q}(t)$, but which do interact with the slit system. In this way it is possible to entangle the environmental wave-function $\Phi(\{x_l\})$ with the system wave-function $\Psi(\mathbf{Q})$, not through any interaction mediated by the slit, but simply because they interact in similar ways with the slit system.

Without going into details (footnote 14), we can easily see how this works in a 'toy' calculation. If we ignore the internal modes of the particle, we have the usual situation depicted in Fig. 5. Assuming slit states $\psi_j \sim |\psi_j| e^{i\phi_j}$, where $j = A, B$, we then find the

[14]The discussion of third-party decoherence here is simplified and does not include energy relaxation and equilibration, or any dynamics. For a proper analysis, see Stamp (to be published).

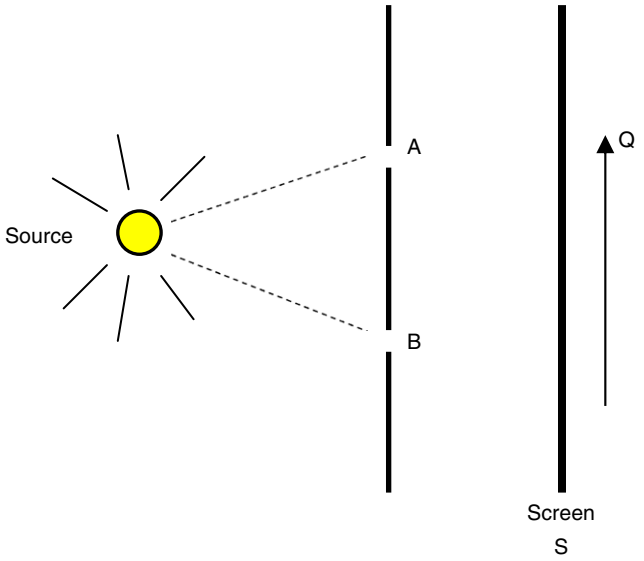


Fig. 5. The two-slit experiment—a particle can pass via either slit A or slit B to reach a point with coordinate Q on the screen S. The probability $P(Q)$ of arrival at Q then shows the standard interference pattern.

probability of arrival of particles at coordinate Q on the screen is given by

$$P(Q) = P_A(Q) + P_B(Q) + 2[P_A(Q)P_B(Q)]^{1/2} \cos\phi_{AB}(Q), \tag{10}$$

where the phase $\phi_{AB}(Q) = \phi_A(Q) - \phi_B(Q)$ in the interference term comes from the difference in phase accumulated by particles traveling through the A or B slits.

Now suppose that when the particle goes through the slit system, internal modes are excited, and these are excited *differently* depending on which slit the particle goes through. In a real experiment on, e.g., buckyballs, which could involve the excitation of phonon modes via the deformation of the buckyball, the deformation will certainly depend on which slit the particle goes through. In this case the internal vibration modes will be excited rather differently. After a passage through slit A, the lth mode will be in some state $\phi_l^A = \sum_{n_l} c_{n_l}^A \chi_{n_l}$, with amplitude $c_{n_l}^A$ to excite this mode into its nth excited state; however passage through the other slit B will give different amplitudes $c_{n_l}^B$, so that the overlap

$$f_l = |\langle\phi_l^A|\phi_l^B\rangle| = \sum_{n_l} |(c_{n_l}^A)^* c_{n_l}^B| < 1. \tag{11}$$

We see that the wave-functions of the internal modes are now entangled with that of the centre of mass motion, even though they have never interacted with them. It is simple to now show that after tracing over these internal modes, we get a suppression of the interference term above by a factor $D = \prod_l f_l$.

The crucial point in the above discussion is that at no point ever do the internal modes of the particle interact with its centre of mass coordinate, either directly or indirectly—instead, they both happen to interact in a similar way with the slit system. In other words, because of the symmetry of the system, a kind of underlying constraint, the two different

systems (centre of mass coordinate, and internal modes) are forced to separately interact with the slit system in such a way that afterwards their states are entangled. One can extend this discussion to other examples, for which there is no space here.

Considering the problem from a more general standpoint, we note that third-party decoherence can affect any system whose behaviour is conditioned by some agency which also happens to condition the behaviour of an environment which we trace over. This means that it can be quite discreet, and not so easy to eliminate from an experiment. As with the spin bath, there is no dissipation in the motion of the central system coordinate Q, yet it still experiences decoherence. The results also mean that the usual discussions of decoherence in terms of interacting system–bath models, described by some effective Hamiltonian, are incomplete. I emphasise that none of the above results are beyond the reach of standard formalism (one can describe them equally well with reduced density matrices or with the decoherence functional formalism, by suitably generalising the averages). The novelty is the necessity for inclusion of the apparently innocuous third party.

3. Decoherence in the lab

The 21st-century lab is the battlefield upon which our ideas on decoherence, confronted by experiment, are going to live or die. Two points are worth emphasising:

(i) The stakes are very high. Questions at issue include: do we really understand what causes decoherence and are there ineluctable or even intrinsic decoherence sources in Nature? Is quantum mechanics valid at large scales? If so, can we use highly entangled multi-particle states (in spite of decoherence)? The experimental answers to these questions will play a major role in the future evolution of physics.
(ii) The relationship between experiment and theory is very complex here. On some fronts, experiment is loath to challenge theory, even where there is striking disagreement. In many cases, the interpretation of the experiments often depends on what theoretical question the experimentalists decide they are probing. Any experiment can be examined through different theoretical lenses.

In the last four decades some landmark experimental tests of quantum mechanics have been formulated and enacted, particularly associated with Bell's theorem and entanglement (Aspect, Dalibard, & Roger, 1982; Clauser & Shimony, 1978). In some of this work, and in offshoots of it, quantum entanglement and superposition have been tested over length scales of many *km* (Marcikic et al., 2004). However none of these tests has involved a large number of particles; rather, they have involved small molecules or a few entangled photons or ions. A number of experimental tests of quantum mechanics at the macroscopic scale, involving very large numbers of particles, were suggested by Leggett et al. (Caldeira & Leggett, 1983; Leggett, 1984, 2002; Leggett & Garg, 1985). These tests all involved the use of superconductors. One set of tests looked at 'macroscopic quantum tunneling' of superconductors—the quantitative theoretical predictions of tunneling rates vs. temperature and applied field (Caldeira & Leggett, 1983; Leggett, 1984) included the dissipative effect on tunneling of the environment (Caldeira–Leggett theory). The later experiments (Clarke, Cleland, Devoret, Esteve, & Martinis, 1988) agreed with this theory over the

whole range of experimental parameters, to within experimental error-a remarkable result. None of these experiments probed decoherence.

Tests of our understanding of decoherence have come both from quantum optics and from solid-state physics. In the former, decoherence in the dynamics of entangled ions is expected to come from interaction with photons (Myatt et al., 2000). Experiment indicates that the mechanisms are understood; there seem to be no hidden sources of decoherence. Tests in the solid state have looked at (i) coherent electron dynamics in mesoscopic conductors, and (ii) coherent tunneling in superconductors and magnetic systems. The main results here are as follows:

(a) Many experiments on mesoscopic conductors measure the time it takes for phase coherence to be lost in the dynamics of the electrons. Some of these experiments indicate that strong decoherence persists down to very low temperature. Since this result conflicts with the standard theory (in which decoherence comes from interactions of electrons between themselves and with phonons and impurities), it has caused much controversy (Aleiner, Altshuler, & Gershenson, 1999; Mohanty, Jariwala, & Webb, 1997). Some more recent experiments (Pierre et al., 2003) indicate that interaction of the electrons with spin impurities may be responsible (i.e., a 'junk' effect, with the junk being the bath of spin impurities).

(b) In a large number of different molecules, the electronic spins lock strongly together to give a 'giant spin', which at low temperatures can quantum tunnel through the energy barrier between two different spin orientations. Many experiments have examined this tunneling (Wernsdorfer, 2001; Tupitsyn & Barbara, 2001), as well as related phenomena in rare earth magnets (Ronnow et al., 2005, and refs. therein). It is now clear what controls the tunneling dynamics of these giant spins. At low temperatures the nuclear spins in the system (coupled strongly to the central giant spin via hyperfine interactions) disrupt the coherent dynamics of the central spin, so that the tunneling is completely incoherent. Present efforts to make spin qubits (for quantum computation) concentrate on suppressing this nuclear spin-mediated decoherence by making the qubit dynamics much faster than the nuclear spin dynamics, bringing in the risk of significant decoherence from phonons (Stamp & Tupitsyn, 2004) (an oscillator bath effect). One can also try to eliminate the nuclear spins by isotopic purification, but this will not be easy.

(c) Tests of the coherent dynamics of a superconducting SQUID between two potential wells are a solid-state realisation of a 'Schrödinger's Cat', in which a macroscopic number of electrons are in a coherent superposition of two different current states (Leggett et al., 1987). In the last few years several experiments have given very strong evidence for Cat states in superconductors (Chiorescu, Nakamura, Harmans, & Mooij, 2003; Nakamura, Pashkin, & Tsai, 1999; Pashkin et al., 2003; Vion et al., 2002). Leggett and Garg (1985) also formulated a criterion of 'macroscopic realism' which can be tested on systems of this kind. The criterion of macrorealism has a clear physical meaning—it distinguishes those properties of a macroscopic system which can be treated as objectively real, in a similar spirit to that discussed by EPR and Bell for microscopic systems. The formal criterion for testing macrorealism involves a set of inequalities pertaining to measurements at different times on a macroscopic variable. These inequalities test quantum mechanics explicitly on the macroscopic scale, but experiments on them have yet to be done. Some of the existing experiments have explicitly measured decoherence rates (Chiorescu et al., 2003; Nakamura et al., 1999; Pashkin et al., 2003; Vion et al., 2002). The experimental decoherence rates in superconductors are always found to be much larger (by up to six orders of magnitude)

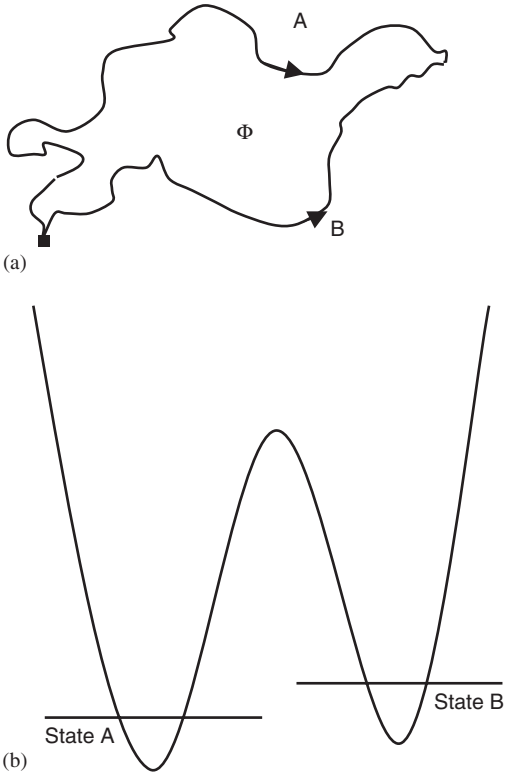


Fig. 6. Tests of large-scale quantum phenomena in condensed matter systems usually involve either interference between propagation along two different paths, which may have some flux Φ enclosed between them (a); or they involve interference between two states quasi-localised in two different potential wells, which communicate weakly by tunneling (b). Experiments on magnetic molecules combine both features (see text).

than those predicted by Caldeira–Leggett theory. Note that this is the same theory which works so well for dissipative tunneling experiments in the same superconductors!

All of the above experiments are of course described using certain theoretical models, and I have summarised the most important ones in Fig. 6. The experiments on decoherence in mesoscopic conductors rely on interference between single electrons following two different paths (Fig. 6(a)); if there is flux enclosed between these paths, we can use the Aharonov–Bohm effect (actually, its suppression) to detect decoherence. In experiments on magnetic molecules, or superconductors (Fig. 6(b)), the relevant mesoscopic or macroscopic coordinate (magnetisation for the molecules, flux for the superconductors) is confined to tunnel between two potential wells. If it moves coherently, one can use it as a qubit, whereas decoherence gradually converts its motion to incoherent tunneling.

We may now summarise the results of the experiments insofar as they concern decoherence. We apparently do know what is causing decoherence in some of the experiments (ions in cavities, possibly mesoscopic conductors, possibly magnetic molecules), but so far theory has not described the decoherence in superconductors. The claims made for intrinsic zero-temperature decoherence in some of these experiments have yet to be properly tested.

Let us now come to a more general discussion of these results, which from a philosophical standpoint is of some interest. The main point I wish to make is that how we interpret the experiments depends mainly on what question we think they are asking. The following questions (amongst others) are thought to be important:

(i) Do we understand the decoherence in these experiments (where it is coming from, how it works)? Are there hidden sources of decoherence? Are there even *intrinsic* mechanisms of decoherence in Nature?
(ii) Is the whole idea of quantum information processing with massively entangled states possible in practice, or even in principle? Can we get rid of decoherence?
(iii) Has macrorealism been tested (and what is the verdict)?

Any response to these questions depends on how one feels about the gap between theory and experiment—in particular, on whether it is felt that standard quantum mechanics, using one mechanism or another, can eventually explain all the decoherence in *all* of the experiments. If so, then one can adopt the view that even very large existing discrepancies are basically just a question of detail. With a lot of work theory and experiment will eventually be brought to agree. If not, then these tests of decoherence mechanisms and rates become of supreme importance in our quest to understand quantum mechanics properly and possibly even to go beyond it. Disagreement between theory and experiment is then very far from being a mere detail.

These are of course two extreme points of view, and there are others lying between them. Nevertheless the point is clear: how the experiments are interpreted depends less on the experiments themselves than on a faith about the validity of the existing theoretical framework.

This point is rather obvious as far as the first two questions are concerned, so I will not belabour it. The question about macrorealism brings the relation between experiment and theory into acute relief. Many (indeed most) physicists, faced with the observations in superconductors of macroscopic coherence, simply remark that the verification of quantum mechanics at the macroscopic scale is not surprising and are then less interested in hearing about tests of macrorealism. The expectation is that quantum mechanics will always prevail. The existence of large amounts of decoherence is then again regarded as a detail, a problem to be solved within the framework of quantum mechanics.

On the other hand there are those who think such tests important, that quantum mechanics does need to be tested at the macroscopic scale and may be found wanting. Apart from the Leggett school of thought (Leggett, 2002; Leggett & Garg, 1985), many papers have discussed non-linear extensions of quantum mechanics, where the non-linearity appears for sufficiently large systems and would be hard to distinguish from decoherence in experiments (Ghirardi, Pearle, & Rimini, 1990; Ghirardi, Rimini, & Weber, 1986; Pearle, 1976, 1989). There are also more exotic ideas, involving intrinsic decoherence sources, coming either from spacetime curvature (intrinsic gravitational decoherence)[15] or from ultra-Planck scale physics ('t Hooft, 1999, 2001), the latter idea having an interesting

[15]For gravitationally induced wave-function collapse, see Diósi (1989), Ghirardi, Grassi, & Rimini (1990) and Penrose (1994, Sections 6.10– 6.12). A possible experimental test was suggested by Marshall, Simon, Penrose, & Bouwmeester (2003).

history.[16] All of these ideas attempt to go beyond existing theory and try to remove some of the paradoxical features of quantum mechanics. Although none of these programmes has actually constructed a comprehensive theory, they do provide possible experimental tests. These test macrorealism in the case of experiments on superconductors, and in the other cases predictions are made for what looks like an intrinsic decoherence rate in Nature, in a way which violates conventional quantum theory. This intrinsic decoherence would show up, in all these cases, in experiments on large-scale quantum phenomena. As noted by 't Hooft (1999, 2001), any such intrinsic decoherence mechanism would put severe limits on quantum information processing (in the holographic approach of 't Hooft and Susskind, it would be impossible for a quantum computer or quantum memory to involve more than roughly 400 entangled qubits).

An important point I wish to make here is that no experiment purporting to test quantum mechanics, according to any of these scenarios, can afford to ignore disagreements between experimental and theoretical decoherence rates; these are no longer a question of detail. One certainly cannot treat any disagreement as a 'dirt' or some other uncontrolled extrinsic effect. This would automatically dismiss any real breakdown of quantum mechanics as a dirt effect and make tests of large-scale quantum mechanics impossible in principle.

The aim of this section has been to give readers a feel for how current experiments bear on some of the really fundamental questions associated with decoherence and possibly on even more fundamental questions about quantum mechanics itself. Perhaps not surprisingly, we see that how the experiments are interpreted depends very much on prevailing views and prejudices, about the expected answers to these questions.[17]

4. Six questions about decoherence and quantum relaxation

With the material in the two previous sections in hand, we may now address directly some of the larger problems mentioned in the introduction. Rather than a lengthy analysis of these, it is simpler to frame the discussion in terms of a set of six questions. Some of these have frequently been posed before, others less so. However in all cases the answers depend in one way or another on what we have been discussing, i.e., on what are the mechanisms of decoherence.

Question 1: What causes decoherence in Nature? Is there a 'generic model' of decoherence (and if so what is it)?

Answer: We have certainly now elucidated some of the decoherence mechanisms operating in Nature, and there is a large variety of them. While the three models discussed in Section 2 (spin bath, oscillator bath, and third-party decoherence) themselves cover

[16]The idea that quantum fluctuations of spacetime at very high energies, up to the Planck scale, might cause decoherence at low energies has been discussed in various contexts. See, e.g., Hawking (1982), Hawking & Laflamme (1988), Coleman (1988) and Ellis, Mohanty, & Nanopoulos (1989).

[17]I emphasise that we are interested here in the theoretical context in which genuine experimental challenges to an established theory (here, quantum mechanics) are mounted and what criteria are used to decide how successful is the challenge. There are currently several controversies raging about the relation between theory and experiment in science, notably over the misuse of experimental data (in, e.g., the debate over evolution vs. 'intelligent design', or in the Schön–Batlogg debacle, where some 20 papers based on fabricated data were published by Nature and Science). This is of course a very different issue, and should not be confused in any way with the present discussion.

many different physical systems, there is no reason to suppose we have found *all* possible mechanisms of decoherence!

At this point one has to insist that the real verdict must come from experiment. Without a quantitative explanation of experimental decoherence rates in terms of known theoretical models, one can always posit undiscovered sources of decoherence 'out there'. As emphasised above, many experimental systems at present show anomalously large experimental decoherence rates (although some discrepancies can probably be explained by spin bath effects). To have a generic model for decoherence would suppose a much better understanding than we presently have of most condensed matter systems. One should beware of general theorems on decoherence rates for large systems, since they usually make very restrictive (and unrealistic) assumptions about the structure of the many-body states.

Thus we are not yet in a position to be talking about a generic model for decoherence.

Question 2: Is decoherence necessarily related to irreversibility and dissipation/relaxation? If so, does decoherence then go to zero with temperature, and can it be eliminated in the real world?

Answer: All these questions have been controversial and are also of fundamental interest. If decoherence were tied to dissipation, then at low energy, with a cold environment, decoherence rates would be very low, going to zero with temperature; and moreover vacuum fluctuations would not cause decoherence at all. Such a conclusion would be of great importance, if true.

As we saw in Section 2, dissipation and decoherence *are* tied together in the oscillator bath models of the environment. However, as we also saw, this result is *not* true for spin baths, where one can have decoherence with no dissipation, even at $T = 0$; and in the case of third-party decoherence there cannot possibly be any environmental dissipation, at any T, since there is no direct coupling to the environment. Thus there is no necessary connection between decoherence and dissipation in the real world, and no necessary reason for it to go to zero at $T = 0$.

This is a problem of real practical interest right now, both for the construction of quantum information processing systems and for the standard physics of solids. There is thus a massive worldwide quest going on for ways to eliminate environmental decoherence. If the three sources of decoherence just mentioned are in fact the only kinds that exist, then one might still entertain hopes of eliminating them; indeed, some very interesting idea for doing this are under present investigation. However, what if there are other decoherence sources? This suggests the next question:

Question 3: Are there 'intrinsic' sources of decoherence in Nature, impossible to eradicate?

Answer: By 'intrinsic' sources, is meant sources which are inevitable in the world as it is, not arising from dissipative processes and perhaps even arising as part of the basic structure of the universe. Such intrinsic sources of decoherence in Nature, operating even at $T = 0$, would not only provide a way of explaining the 'emergence of classical physics' in fields ranging from quantum cosmology to condensed matter physics; they would also place a fundamental limit on the observability of quantum phenomena. This would limit the possibility of seeing macroscopic quantum phenomena, and also place fundamental limits on the superpositions required for quantum computing.

Possibilities for intrinsic decoherence mechanisms have already emerged from both low- and high-energy physics. From low-energy physics there has been a suggestion that zero

point modes of continuous quantum fields (in particular, the photon field) could cause $T = 0$ decoherence. This has, for example, been suggested as an explanation of the decoherence saturation at low T in mesoscopic conductors (Mohanty et al., 1997). This suggestion is controversial and was discussed in Section 3; many feel that the explanation lies instead with magnetic impurities (a spin bath effect). There is also the suggestion of non-linear terms in the dynamics of macroscopic quantum systems (Ghirardi, Pearle et al., 1990; Ghirardi et al., 1986; Pearle, 1976, 1989); this has hardly been tested yet.

On the high-energy side a wide range of possibilities has been canvassed and already noted in Section 3. These include, again, low-energy decoherence from zero point modes—this time from gravitons, or string fields, or from vacuum fluctuations of the spacetime metric (footnote 16) (including the so-called 'baby universe' fluctuations). So far these suggestions have not been met enthusiastically for they fly in the face of conventional ideas about renormalisation, according to which neither very high energy modes nor vacuum fluctuations can enter into any dynamic processes in a low-energy effective Hamiltonian. There are also more exotic ideas, involving modifications of quantum theory. Two recent proposals are an intrinsic decoherence arising from spacetime curvature (intrinsic gravitational decoherence) (footnote 15), and a source arising from ultra-Planck scale physics, suggested by 't Hooft (1999, 2001). Although neither of these programmes has actually constructed a comprehensive theory, they do provide possible experimental tests, in both cases involving an intrinsic decoherence rate, which violates conventional quantum theory. This intrinsic decoherence would show up in both cases in experiments on large-scale quantum phenomena. As noted by 't Hooft, any such intrinsic decoherence mechanism would put severe limits on quantum information processing (in the holographic approach of 't Hooft and Susskind, it would be impossible for a quantum computer or quantum memory to involve more than roughly 400 entangled qubits).

Clearly some pretty crucial experiments are required here. This is one of the very interesting frontiers of physics right now. Such experiments will have to be done with great care, to eliminate, for example, the influence of third-party decoherence processes, not reflected in the effective Hamiltonian of the experimental system but in its previous history. Indeed it is not obvious to the present author how one can eliminate third-party decoherence with certainty.

Question 4: Does decoherence give rise to the 'emergence' of classical physics? If so, then what kind of a theory is quantum mechanics (often held to depend on classical mechanics for its definition in the first place)?

Answer: One interpretation of this question focuses on the more physical question of how classical quasi-deterministic behaviour emerges for large systems, and/or how quasi-classical stochastic behaviour emerges, even for small systems. It should now be completely evident, from Sections 2 and 3, that a proper answer to this question requires understanding the real decoherence mechanisms operating in Nature and that these are not so simple, or necessarily completely understood. Thus we do not yet have a theory which derives classical physics from quantum physics solely using ideas from decoherence, even though we do have some derivations of classical behaviour within certain models. It is important to note that in some other models one can actually find non-classical behaviour emerging in the large-scale dynamics, because of decoherence (this happens, for example, when one is dealing with a spin bath environment (Prokof'ev & Stamp, 2006)). Thus there is nothing inevitable about classical behaviour!

In a second interpretation, it is suggested that decoherence might not only derive classical physics as a limiting case of quantum mechanics but also show how strictly classical concepts such as momentum and position are inevitable in the very formulation of quantum mechanics. The basic argument here is that the structure of interactions in Nature inevitably leads to a preferred 'inert' or 'pointer' basis for the states of macroscopic objects (Simonius, 1978; Zurek, 1981, 1982, 2003). Again, however, this argument has relied on simple models, and a general demonstration would require the use of more general models. It is extremely interesting to ask whether more general models could yield instead 'non-classical' pointer bases. Just as interesting is to ask what physicists will do if, as seems very possible, experiments find that 'macrorealism' (in the Garg–Leggett sense) fails. How then will we formulate quantum mechanics?

Question 5: With the understanding gained into the mechanisms of decoherence, can we now say how quantum measurements work? And does decoherence 'solve' the measurement problem?

Answer: One of the remarkable paradoxes of quantum physics is how difficult to give a theoretical description of most measurement schemes, even though they are being used all the time to do experiments! Detailed accounts, including all steps from the measured degrees of freedom up to the final 'classical' state of the measuring apparatus, are mostly confined to experiments designed for tests of quantum phenomena (often in quantum optics labs) or to sensitive experiments designed to search for very weak effects (e.g., gravity waves (Braginsky & Khalili, 1992; Caves et al., 1980)). Usually in these descriptions assumptions are made about how irreversible amplification processes, accompanied by strong decoherence, lead to definite results, FAPP (For All Practical Purposes). There is no question that if serious tests of quantum mechanics are to be made at the macroscopic scale (e.g., of macrorealism), a more complete analysis will need to be done, carrying the full quantum description right up to the macroscopic scale and including all sources of decoherence at each stage. It hardly matters which verdict the experiments give here. In either case a convincing experimental result will only be attained if all sources of decoherence are understood (including third-party decoherence).

Whether such analyses will 'solve' the measurement problem depends on what the problem is supposed to be. As with question 4, we remark that there is nothing inevitable about a classical behaviour for the measuring system (unless one defines measuring systems so that they must be classical!). On the other hand if tests of quantum mechanics at the macroscopic scale do actually vindicate it, so that macrorealism is falsified, then the measurement problem will surely undergo a radical transformation to a new problem, viz.: how far can we push the 'FAPP barrier' (between the quantum and classical worlds) into what is now considered the classical world? Certainly a new vocabulary will be required by physicists to deal with genuinely macroscopic quantum states.

Question 6: Is decoherence connected to the 'Arrow of Time'? If so, how?

Answer: It is commonly assumed that all arrows, including the thermodynamic arrow, derive from the cosmological arrow. In this view, irreversibility is caused ultimately by the cosmological arrow. If one assumes that decoherence is connected with irreversibility, then the 'quantum arrow' results from the thermodynamic arrow (a commonly adopted point of view) and is also then subservient to the cosmological arrow. In this picture, everything in the universe, even something as basic as classical spacetime, has resulted from special initial conditions.

However, this point of view is by no means universally accepted, and it is possible to write quantum mechanics, including measurements, in a time-symmetric form (Aharonov, Bergmann, & Lebowitz, 1964). We have also seen above that some kinds of decoherence are not at all connected with dissipation or irreversibility and that decoherence does not even necessarily have to lead to classical behaviour. Thus we are driven back to the familiar question about mechanisms—whether or not we should associate decoherence with either the thermodynamic or the cosmological arrow of time depends on what mechanisms are responsible for decoherence. Certainly the results about the mechanisms of decoherence, discussed in Sections 2 and 3, make it clear that there is no necessary or logical connection between decoherence and the thermodynamic arrow. If decoherence is logically independent of the thermodynamic arrow, it is much less obvious that it is connected with the cosmological arrow.

A real handicap in analyses of this question is that many of the current discussions of the arrow of time are framed in terms of theories about the beginning of the universe, or of vague ideas like the 'anthropic principle', which have not been really tested and which change fairly rapidly with time. The most prudent course of action here may be to suspend judgement on any possible connection between the cosmological arrow of time and decoherence until both are understood a little better.

5. Conclusions

In this article I have discussed how general questions about the nature of solids and about the low-energy physics of macroscopic systems have consequences for old questions about quantum measurements, about the relation between classical and quantum mechanics, and about the validity of quantum mechanics itself. If there is a central point here, it is that facts about the physical mechanisms of decoherence are crucial to answering these questions. We now know something about these mechanisms, and what we have found out has radically changed our perspective. Far from asking "how do decoherence and/or dissipation produce classical mechanics at the macroscopic scale?", we are now asking "how can we evade decoherence at the macroscopic scale?". The preparation and use of states with high-level entanglement (i.e., N-entangled states with $N \gg 1$), instead of being treated as a theoretical impossibility, is now a target in many experimental research programmes. Most radical of all, the idea that the investigation of such states could lead to a failure of quantum mechanics itself is being taken seriously by both high- and low-energy theorists, with experiments to test this idea in preparation.

If there is a thread running through all of this, it is that to make progress we need a firm understanding of the physical mechanisms governing decoherence. Decoherence, according to the older ideas, is supposed to explain away the quantum measurement problem and to explain how classical mechanics emerges from quantum mechanics. And yet in the last few years experiments have been gradually bringing decoherence under control, inexorably pushing quantum mechanics to scales that were formerly the preserve of classical physics. Along the way a new picture, a picture of how decoherence operates, has begun to emerge. Far from being associated with ordinary relaxation, the decoherence in most experiments (certainly those in solid-state systems) appears to come from 'sleeper' modes, modes nearly invisible in most experiments because they cause almost no dissipation. Thus decoherence is more subtle, and perhaps more pervasive, than previously thought. There are many

things we still do not understand about decoherence and what causes it, and it should now be clear that this is a very pressing problem.

We have seen that how one views all this depends much on pre-existing prejudices, both about our present understanding of solids and about the validity of quantum mechanics itself. The point of view I have taken is that there are still many things we do not understand about solids, particularly at very low energies, and that the failure of quantum mechanics is a possibility which is certainly worth considering and testing experimentally. It then follows that we cannot dismiss disagreement between theoretical and experimental decoherence rates, which may conceal the very failure we are looking for, whether it comes from the ultra-Planck scale or from very low energies.

It is always remarkable when a combination of theory and experiment has larger philosophical consequences. Perhaps the most dramatic example in recent times has been the impact of Bell's inequalities, where a set of experiments in atomic physics was able to rule out a whole class of possible theories about Nature, and in doing so, consign a widely accepted philosophical view about 'reality' to the dustbin. The fascinating prospect is that future experiments at low temperatures in condensed matter systems looking for 'gravitational decoherence', or non-linear terms in a future quantum mechanics, or something else, may have a similar impact. But this remains to be seen.

Acknowledgements

I was stimulated to write this paper by a PITP workshop which brought together philosophers, historians, and physicists, in December 2004. The paper has benefited from conversations with my present and former colleagues A. Morello, I. S. Tupitsyn, W. G. Unruh, and G. 't Hooft, and also from discussions with M. Devoret, D. Howard, A. J. Leggett, N. V. Prokof'ev, Y. Nakamura, and R. Wald. However none of them are responsible for what is said herein! The research was supported by the National Science and Engineering Council, by the Canadian Institute for Advanced Research, and by the Pacific Institute of Theoretical Physics (PITP).

References

Aharonov, Y., Bergmann, P. G., & Lebowitz, J. L. (1964). Time symmetry in the quantum process of measurement. *Physical Review*, *134*(6B), B1410–B1416.

Aleiner, I. L., Altshuler, B. L., & Gershenson, M. E. (1999). Interaction effects and phase relaxation in disordered systems. *Waves in Random Media*, *9*, 201–239.

Anderson, P. W. (1972). More is different. *Science*, *177*(4047), 393–396.

Anderson, P. W. (1984). *Basic notions of condensed matter physics*. Reading, MA: Benjamin.

Anderson, P. W. (1994). *World Scientific series in 20th century physics*: *Vol.* 7, *A career in theoretical physics* (pp. 525–538). Singapore: World Scientific.

Aspect, A., Dalibard, J., & Roger, G. (1982). Experimental test of Bell's inequalities using time-varying analyzers. *Physical Review Letters*, *49*, 1804–1807.

Bacciagaluppi, G. (2005). The role of decoherence in quantum mechanics. *The Stanford encyclopedia of philosophy*. Retrieved February 18, 2006 from ⟨http://plato.stanford.edu/archives/sum2005/entries/qm-decoherence/⟩.

Binder, K., & Young, A. P. (1986). Spin-glasses—experimental facts, theoretical concepts, and open questions. *Reviews of Modern Physics*, *58*, 801–976.

Born, M., & Ludwig, G. (1958). *Zur Quantenmechanik des Kräftefreien Teilchens* (Quantum mechanics of force free particles). *Zeitschrift für Physik A Hadrons and Nuclei*, *150*(1), 106–117.

Braginsky, V. B., & Khalili, F. Y. (1992). *Quantum measurement*. Cambridge: Cambridge University Press.

Caldeira, A. O., & Leggett, A. J. (1983). Quantum tunnelling in a dissipative system. *Annals of Physics*, *149*, 374–456.

Callan, C. G., Felce, A. G., & Freed, D. (1993). Critical theories of the dissipative Hofstadter model. *Nuclear Physics B*, *392*, 551–592.

Callan, C. G., & Freed, D. (1992). Phase diagram of the dissipative Hofstadter model. *Nuclear Physics B*, *374*, 543–566.

Caves, C. M., Thorne, K. S., Drever, R. W. P., Sandberg, V. D., & Zimmermann, M. (1980). On the measurement of a weak classical force coupled to a quantum-mechanical oscillator. I. Issues of principle. *Review of Modern Physics*, *52*(2), 341–392.

Chiorescu, I., Nakamura, Y., Harmans, C. J. P. M., & Mooij, J. E. (2003). Coherent quantum dynamics of a superconducting flux qubit. *Science*, *299*, 1869–1871.

Clarke, J., Cleland, A. N., Devoret, M. H., Esteve, D., & Martinis, J. (1988). Quantum mechanics of a macroscopic variable—The phase difference of a Josephson junction. *Science*, *239*, 992–997.

Clauser, J. F., & Shimony, A. (1978). Bell's theorem: Experimental tests and implications. *Reports on Progress in Physics*, *41*, 1881–1927.

Coleman, S. (1988). Why there is nothing rather than something: A theory of the cosmological constant. *Nuclear Physics B*, *310*(3 and 4), 643–668.

Cornwal, J. M., & Bruinsma, R. (1988). Quantum evolution of an unstable field in a de Sitter-space thermal bath. *Physical Review D*, *38*, 3146–3157.

Daneri, A., Loinger, A., & Prosperi, G. M. (1962). Quantum theory of measurement and ergodicity conditions. *Nuclear Physics*, *33*, 297.

Daneri, A., Loinger, A., & Prosperi, G. M. (1966). Further remarks on the relation between statistical mechanics and quantum theory of measurement. *Nuovo Cimento*, *44B*, 119.

d'Espagnat, B. (1976). *Conceptual foundations of quantum mechanics*. Reading, MA: Benjamin.

Diósi, L. (1989). Models for universal reduction of macroscopic quantum fluctuations. *Physical Review A*, *40*, 1165–1174.

Dubé, M., & Stamp, P. C. E. (2001). Mechanisms of decoherence at low temperatures. *Chemical Physics*, *268*, 257–272.

Ellis, J., Mohanty, S., & Nanopoulos, D. V. (1989). Quantum gravity and the collapse of the wave function. *Physics Letters B*, *221*, 113–124.

Esquinazi, P. (Ed.). (1998). *Tunneling systems in amorphous and crystalline solids*. Berlin, Heidelberg, New York: Springer.

Feynman, R. P. (1965). *Lecture notes in physics* (Vol. 3). New York: Addison-Wesley.

Feynman, R. P., & Hibbs, A. R. (1965). *Quantum mechanics and path integrals*. New York: McGraw-Hill.

Feynman, R. P., Leighton, R. B., & Sands, M. (1965). *Feynman lectures on physics* (Vol. 3). Reading, MA: Addison-Wesley.

Feynman, R. P., & Vernon, F. L. (1963). The theory of a general quantum system interacting with a linear dissipative system. *Annals of Physics*, *24*, 118–173.

Gell-Mann, M., & Hartle, J. B. (1993). Classical equations for quantum systems. *Physical Review D*, *47*, 3345–3382.

Ghirardi, G. C., Grassi, R., & Rimini, A. (1990). Continuous-spontaneous-reduction model involving gravity. *Physical Review A*, *42*, 1057–1064.

Ghirardi, G. C., Pearle, P., & Rimini, A. (1990). Markov processes in Hilbert space and continuous spontaneous localization of systems of identical particles. *Physical Review A*, *42*, 78–89.

Ghirardi, G. C., Rimini, A., & Weber, T. (1986). Unified dynamics for microscopic and macroscopic systems. *Physical Review D*, *34*, 470–491.

Grabert, H., Schramm, P., & Ingold, G.-L. (1988). Quantum Brownian motion: The functional integral approach. *Physics Reports*, *168*, 115–207.

Green, H. S. (1958). Observation in quantum mechanics. *Nuovo Cimento*, *9*, 880–889.

Griffiths, R. B. (1984). Consistent histories and the interpretation of quantum mechanics. *Journal of Statistical Physics*, *36*(1 and 2), 219–272.

Griffiths, R. B. (1986). Correlations in separated quantum systems: A consistent history analysis of the EPR problem. *American Journal of Physics*, *55*(1), 11–17.

Häffner, H., Hänsel, W., Roos, C. F., Benhelm, J., Chek-al-kar, D., Chwalla, M., et al. (2005). Scalable multiparticle entanglement of trapped ions. *Nature, 438*, 643–646.

Halliwell, J. J., Pérez-Mercader, J., & Zurek, W. H. (Eds.). (1994). *Physical origins of time asymmetry*. Cambridge: Cambridge University Press.

Hartle, J. B. (1991). Spacetime coarse grainings in nonrelativistic quantum mechanics. *Physical Review D*, *44*, 3173–3196.

Hawking, S. W. (1982). The unpredictability of quantum gravity. *Communications in Mathematical Physics*, *87*(3), 395–415.

Hawking, S. W., & Laflamme, R. (1988). Baby universes and the nonrenormalizability of gravity. *Physics Letters B*, *209*, 39–44.

Joos, E., & Zeh, H. D. (1985). The emergence of classical properties through interactions with the environment. *Zeitschrift für Physik*, *B59*, 223.

Joos, E., Zeh, H. D., Kiefer, C., Giulini, D., Kupsch, K., & Stamatescu, I.-O. (2003). *Decoherence and the appearance of a classical world in quantum theory* (2nd ed.). Berlin, Heidelberg: Springer.

Kempe, J. (2003). Quantum random walks: An introductory overview. *Contemporary Physics*, *44*, 307–327 (also published in quant-ph/0303081).

Kendon, V. (2003, August 13). Quantum walks on general graphs. *quant-ph/0306140*. Retrieved February 19, 2006 from ⟨http://arxiv.org/abs/quant-ph/0306140⟩.

Leggett, A. J. (1984). Quantum tunneling in the presence of an arbitrary linear dissipation mechanism. *Physical Review B*, *30*(3), 1208–1218.

Leggett, A. J. (2002). Testing the limits of quantum mechanics: Motivation, state of play, prospects. *Journal of Physics: Condensed Matter*, *14*, R415–R451.

Leggett, A. J., Chakravarty, S., Dorsey, A. T., Fisher, M. P. A., Garg, A., & Zwerger, W. (1987). Dynamics of the dissipative two-state system. *Review of Modern Physics*, *59*, 1–85.

Leggett, A. J., & Garg, A. (1985). Quantum mechanics versus macroscopic realism: Is the flux there when nobody looks? *Physical Review Letters*, *54*, 857–860.

Leibfried, D., Knill, E., Seidelin, S., Britton, J., Blakestad, R. B., Chiaverini, J., et al. (2005). Creation of a six-atom 'schrödinger cat' state. *Nature, 438*, 639–642.

London, F. W., & Bauer, E. (1983). The theory of observation in quantum mechanics. In J. A. Wheeler, & W. H. Zurek (Eds.), *Quantum theory and measurement*. Princeton: Princeton University Press (original work published 1939).

Ludwig, G. (1953). *Der Meßprozeß* (The measurement process). *Zeitschrift für Physik A Hadrons and Nuclei*, *135*(5), 483–511.

Ludwig, G. (1958). *Zum Ergodensatz und zum Begriff der Makroskopischen Observablen* (Ergodic principle and the concept of macroscopic observables). *Zeitschrift für Physik A Hadrons and Nuclei*, *152*(1), 98.

Marcikic, I., de Riedmatten, H., Tittel, W., Zbinden, H., Legré, M., & Gisin, N. (2004). Distribution of time-bin entangled qubits over 50 km of optical fiber. *Physical Review Letters*, *93*, 180502.1–180502.4.

Margenau, H. (1937a). Critical points in modern physical theory. *Philosophy of Science, 4*, 337–370.

Margenau, H. (1937b). Quantum-mechanical description. *Physical Review, 49*, 240–241.

Marshall, W., Simon, C., Penrose, R., & Bouwmeester, D. (2003). Towards quantum superpositions of a mirror. *Physical Review Letters*, *91*, 130401.1–130401.4.

Mézard, M., Parisi, G., & Virasoro, M. A. (1987). *World Scientific lecture notes in physics*: *Vol. 9.*, *spin glass theory and beyond*. Singapore: World Scientific.

Mohanty, P., Jariwala, E. M. Q., & Webb, R. A. (1997). Intrinsic decoherence in mesoscopic systems. *Physical Review Letters*, *78*, 3366–3369.

Myatt, C. J., King, B. E., Turchette, Q. A., Sackett, C. A., Kielpinski, D., Itano, W. M., et al. (2000). Decoherence of quantum superpositions through coupling to engineered reservoirs. *Nature, 403*, 269–273.

Nakamura, Y., Pashkin, Y. A., & Tsai, J. S. (1999). Coherent control of macroscopic quantum states in a single-Cooper-pair box. *Nature*, *398*, 786–788.

Neumann, J. von (1955). *Mathematical foundations of quantum mechanics*. Princeton: Princeton University Press (original work published 1932).

Omnés, R. (1992). Consistent interpretations of quantum mechanics. *Reviews of Modern Physics*, *64*, 339–382.

Omnés, R. (1994). *The interpretation of quantum mechanics*. Princeton: Princeton University Press.

Pashkin, Y. A., Yamamoto, T., Astafiev, O., Nakamura, Y., Averin, D. V., & Tsai, J. S. (2003). Quantum oscillations in two coupled charge qubits. *Nature*, *421*, 823–826.

Pauli, W. (1980). *General principles of quantum mechanics*. Berlin: Springer (original work published 1958).

Pearle, P. (1976). Reduction of the state vector by a nonlinear Schrödinger equation. *Physical Review D*, *13*, 857–868.

Pearle, P. (1989). Combining stochastic dynamical state-vector reduction with spontaneous localization. *Physical Review A*, *39*, 2277–2289.

Penrose, R. (1994). *Shadows of the mind*. Oxford: Oxford University Press.
Pierre, F., Gougam, A. B., Anthore, A., Pothier, H., Esteve, D., & Birge, N. O. (2003). Dephasing of electrons in mesoscopic metal wires. *Physical Review B*, *68*, 085413.1–085413.15.
Prokof'ev, N. V., & Stamp, P. C. E. (1993). Giant spins and topological decoherence: A Hamiltonian approach. *Journal of Physics: Condensed Matter*, *5*, L663–L670.
Prokof'ev, N. V., & Stamp, P. C. E. (2000). Theory of the spin bath. *Reports on Progress in Physics*, *63*, 669–726.
Prokof'ev, N.V., & Stamp, P.C.E. (2006). Decoherence and Quantum walks: anomalous diffusion and ballistic tails. cond-mat/0605097, retrieved May 3rd, 2006 from http://arxiv.org/abs/cond-mat/0605097.
Ronnow, H. M., Parthasarathy, R., Jensen, J., Aeppli, G., Rosenbaum, T. F., & McMorrow, D. F. (2005). Quantum phase transition of a magnet in a spin bath. *Science*, *308*, 389–392.
Savitt, S. F. (1995). *Time's arrow today: Recent physical and philosophical work on the direction of time*. Cambridge: Cambridge University Press.
Schulman, L. S. (1997). *Time's arrow and quantum measurement*. Cambridge: Cambridge University Press.
Simonius, M. (1978). Spontaneous symmetry breaking and blocking of metastable states. *Physical Review Letters*, *40*, 980–983.
Stamp, P. C. E., & Tupitsyn, I. S. (2004). Coherence window in the dynamics of quantum nanomagnets. *Physical Review B*, *69*, 014401.1–014401.5.
't Hooft, G. (1999). Quantum gravity as a dissipative deterministic system. *Classical and Quantum Gravity Class*, *16*(10), 3263–3279.
't Hooft, G. (2001, May 11). Quantum mechanics and determinism. *hep-th/0105105*, Retrieved February 18, 2006 from ⟨http://arxiv.org/abs/hep-th/0105105⟩.
Thouless, D. J. (1998). *Topological quantum numbers in non-relativistic physics*. Singapore: World Scientific.
Tupitsyn, I. S., & Barbara, B. (2001). Quantum tunneling of magnetization in molecular complexes with large spins: Effect of the environment. In M. Drillon, & J. Miller (Eds.), Magnetoscience—From molecules to materials (pp. 109–168). Weinheim: Wiley.
Unruh, W. G. (1999). False loss of coherence. In H. P. Breuer, & F. Petruccione (Eds.), *Relativistic quantum measurement and decoherence* (pp. 125–140). Berlin: Springer.
Van Kampen, N. G. (1988). Ten theorems about quantum mechanical measurements. *Physica A: Statistical and Theoretical Physics*, *153*(1), 97–113.
Vion, D., Aassime, A., Cottet, A., Joyez, P., Pothier, H., Urbina, C., et al. (2002). Manipulating the quantum state of an electrical circuit. *Science, 296*(5569), 886–889.
Weiss, U. (1999). *Quantum dissipative systems*. Singapore: World Scientific.
Wernsdorfer, W. (2001). Classical and quantum magnetization reversal studied in nanometer-sized particles and clusters. *Advances in Chemical Physics*, *118*, 99–190.
Wheeler, J. A., & Zurek, W. H. (Eds.). (1983). *Quantum theory and measurement*. Princeton: Princeton University Press.
Yu, C. C., & Leggett, A. J. (1988). Low-temperature properties of amorphous materials: Through a glass darkly. *Comments on Condensed Matter Physics*, *14*(4), 231–251.
Zeh, H. D. (1970). On the interpretation of measurement in quantum theory. *Foundations of Physics*, *1*(1), 69.
Zeh, H. D. (1973). Toward a quantum theory of observation. *Foundations of Physics*, *3*(1), 109–116.
Zeh, H.-D. (1989). *The physical basis of the direction of time* (4th ed.). Berlin, Heidelberg, New York: Springer.
Zeh, H. D. (2002, June 30). Basic concepts and their interpretation. *quant-ph/9506020*, Retrieved February 18, 2006 from ⟨http://arxiv.org/abs/quant-ph/9506020⟩.
Zhao, Z., Chen, Y.-A., Zhang, T., Yang, T., Briegel, H. J., & Pan, J.-W. (2004). Experimental demonstration of five-photon entanglement and open-destination teleportation. *Nature*, *430*, 54–58.
Zurek, W. H. (1981). Pointer basis of quantum apparatus: Into what mixture does the wave packet collapse? *Physical Review D*, *24*, 1516–1525.
Zurek, W. H. (1982). Environment-induced superselection rules. *Physical Review D*, *26*, 1862–1880.
Zurek, W. H. (2003). Decoherence, einselection, and the quantum origins of the classical. *Reviews of Modern Physics*, *75*, 715–785.

York, 1972), Vol. 2, p. 116.
[10]Aa. S. Sudbø and P. C. Hemmer, Phys. Rev. B 13, 980 (1976); S. Hsu and J. D. Gunton, Phys. Rev. B 15, 2688 (1977).
[11]L. P. Kadanoff, Phys. Rev. Lett. 34, 1005 (1975); L. P. Kadanoff, A. Houghton, and M. C. Yalabik, J. Statist. Phys. 14, 171 (1976). When "shifting bonds" in the Kadanoff scheme, one must remember that the pair correlation function for two spins on opposite sides of the seam is the negative of that for two spins in the same relative spatial positions on the same side of the seam.
[12]The value of $k^{-1}d\gamma^{\sigma}/dT$ at $T=0$ is -0.8723 ± 0.0010 according to J. W. Cahn and R. Kikuchi, J. Phys. Chem. Solids 20, 94 (1961). Our approximation yields -1.14.
[13]That the singularity could be extremely "mild" is suggested by a model calculation of H. van Beijeren, Phys. Rev. Lett. 38, 993 (1977).

Spontaneous Symmetry Breaking and Blocking of Metastable States

Markus Simonius
Laboratorium für Kernphysik, Eidgenossische Technische Hochschule, CH-8093 Zürich, Switzerland
(Received 27 October 1977)

A mechanism for spontaneous symmetry breaking and related phenomena and the corresponding blocking of metastable states is discussed. It is based on the interaction of an object with a background of "probes" like photons or particles, etc., in its natural surrounding. Applications include quasilocalization of macroscopic bodies, spontaneous parity nonconservation of sugar crystals, localization of atoms in molecules (Born-Oppenheimer approximation), stability of metastable compounds, and perhaps also intrinsic symmetries of elementary particles.

The object of this Letter is to discuss a quantum mechanical mechanism of spontaneous symmetry breaking and related phenomena and the corresponding blocking of states which are not eigenstates of the Hamiltonian of the object in question. This mechanism differs from, and is much more powerful than, the one usually discussed in the current literature based on nonsymmetric solutions to symmetric equations.

Consider an object O and a probe P described in Hilbert spaces H^O and H^P, respectively. For simplicity H^O is assumed to be two dimensional; generalizations will be mentioned at the end. Consider further two orthogonal states $\varphi_1{}^O$ and $\varphi_2{}^O$ of the object and an interaction between object and probe leading to the following transitions of the combined system:

$$\varphi_i{}^O\otimes\varphi_0{}^P\rightarrow\varphi_i{}^O\otimes\varphi_i{}^P,\quad i=1,\ 2, \tag{1}$$

where $\varphi_0{}^P$ is the assumed initial state of the probe. By linearity this gives the transition for arbitrary initial state $a\varphi_1{}^O+b\varphi_2{}^O$. If no further observation is performed on the probe, and the object alone is considered after separation of the two, the object has to be described by a density matrix ρ^O on H^O and the trace has to be taken over H^P.

Taking $\varphi_1{}^O,\varphi_2{}^O$ as basis in H^O and assuming $(\varphi_1{}^P,\varphi_2{}^P)=0$, the transition (1) for an initial superposition $a\varphi_1{}^O+b\varphi_2{}^O$ with $|a|^2+|b|^2=1$ leads then for the density matrix ρ^O of the object to

$$\begin{pmatrix}|a|^2 & ab^*\\ a^*b & |b|^2\end{pmatrix}\rightarrow\begin{pmatrix}|a|^2 & ab^*(\varphi_2{}^P,\varphi_1{}^P)\\ a^*b(\varphi_1{}^P\varphi_2{}^P) & |b|^2\end{pmatrix}$$
$$=\begin{pmatrix}|a|^2 & 0\\ 0 & |b|^2\end{pmatrix}. \tag{2}$$

This constitutes a "reduction of the state vector" of a pure initial to a mixed final state of the object.

There are two important aspects in this connection. One is the compatibility of this reduction with the linearity of the law of motion. This is the case by construction. The second is the assessment of the relevance and frequency of occurrence of this phenomenon. This may be judged from its connection to the process of measurement. Indeed Eq. (1) is a (simplified) model of a measurement, in which information is transferred from the object to the probe in such a way that subsequent observation of the probe alone *could* discriminate exactly between the two cases where the object is initially in the state $\varphi_1{}^O$ or $\varphi_2{}^O$. The *possibility* of this discrimination requires $\varphi_1{}^P$ and $\varphi_2{}^P$ in Eq. (1) to be orthogonal and thus leads to the exact depletion of the off-diagonal elements of ρ^O in Eq. (2) if the object alone is considered, i.e., even, and in particular, *in*

the absence of any actual observation of the probe.[1] The following two examples illustrate some important application. (I omit the superscript O for the object where this does not lead to confusion.)

Example 1.—The object is a macroscopic body. φ_1 and φ_2 are two localized spacially well-separated wave functions. The probe consists of one or or several photons in the region where φ_1 or φ_2 is localized. By observation of the photons one could discriminate between the case where the object is in the state φ_1 or φ_2. It follows that the passing, emission, or absorption of such photons destroys any coherence which might have prevailed before. Thus under the usual conditions of the macroscopic world where one can see the objects, i.e., discriminate between different locations by observing photons, it is not possible to preserve the coherence between states with macroscopically different localization.[2] There is no particular (e.g., cosmological) initial condition needed for this. This is responsible also, for instance, for the appearance of bubbles, localized in space, in an overheated liquid.

Example 2.—Consider sugar. It can be in a "right" (φ_1) and "left" (φ_2) state. If the sugar is crystallized, the two may be discriminated by eye (or microscope) by observing the light scattered from it. Thus, again, any coherence between the two states φ_1 and φ_2 is immediately destroyed by the interaction with this light. Now if parity is conserved and the ground state is nondegenerate it is of the form $\varphi = \alpha\varphi_1 + \beta\varphi_2$ with $|\alpha|^2 = |\beta|^2 = \frac{1}{2}$. Symmetry breaking cannot be due to lack of symmetry of the ground state of a symmetric Hamiltonian unless the ground state is degenerate, which is not expected to be the case in general. In both examples the probes could as well be electrons or molecules from the surrounding (in particular for sugar dissolved in a liquid).

An important question which remains to be answered is, what singles out the particular states φ_1 and φ_2 for the reduction (2), i.e., why does the reduction in the examples above not sometimes produce an incoherent mixture between $\varphi_+ = (\sqrt{2})^{-1}(\varphi_1 + \varphi_2)$ and $\varphi_- = (\sqrt{2})^{-1}(\varphi_1 - \varphi_2)$, say, instead of φ_1 and φ_2. But this would imply that if the object is originally in the state φ_1, say, it could be found afterwards with 50% probability in the state φ_2. This means that the interaction with the probe, for which I took photons in my examples, could cause with 50% probability a transition between φ_1 and φ_2 which for large enough distance between the two cannot be the case for a massive body. Thus the inertia of massive objects singles out quasilocalized states for a reduction by interactions with light probes.

I now turn to an important consequence of reduction (2) in case it occurs frequently (in a sense to be discussed), namely the *blocking* of the states φ_1 and φ_2 by stochastic (repeated) reduction.

Consider for concreteness example 2 with a parity-conserving Hamiltonian H. For proper choice of the phases of φ_1 and φ_2 the two eigenstates of H are $\varphi_\pm = (\varphi_1 \pm \varphi_2)/\sqrt{2}$. Their energy difference is $\hbar\omega$. Then, after a reduction the density matrix on the right-hand side of (2) will evolve in a time interval τ according to $\rho \to \exp\{-i\tau H/\hbar\}\rho\exp\{i\tau H/\hbar\}$ to

$$\frac{1}{2}\begin{pmatrix} 1+\delta_0\cos\omega\tau & i\delta_0\sin\omega\tau \\ -i\delta_0\sin\omega\tau & 1-\delta_0\cos\omega\tau \end{pmatrix}, \tag{3}$$

where $\delta_0 = |a|^2 - |b|^2$. For $\delta_0 \neq 0$, in particular for $\delta_0 = \pm 1$, i.e., pure initial state φ_1 or φ_2, this exhibits the expected oscillations with circular frequency ω. Thus the original property (right-handedness of some sugar crystals) disappears after a time $T \sim \omega^{-1}$. The fact that this time is long is usually assumed to be the reason for stability. It will be seen that this is only part of the truth.

If after a time interval τ the object is again hit by a probe leading to a reduction of the density matrix to diagonal form, then

$$\rho \to \frac{1}{2}\begin{pmatrix} 1+\delta & 0 \\ 0 & 1-\delta \end{pmatrix}, \text{ with } \delta = \delta_0\cos\omega t. \tag{4}$$

If this is repeated n times, say, then, after a time $t = \sum_\nu \tau_\nu \approx n\bar{\tau}$, where τ_ν are the time intervals between subsequent reductions and $\bar{\tau}$ their mean value, one obtains $\delta = \delta_t = \delta_0 \prod_\nu \cos\omega\tau_\nu$ in Eq. (4).

I now consider the important cases where[3]

$$\omega\tau_\nu \ll 1, \text{ i.e., } \tau_\nu \ll T = 1/\omega. \tag{5}$$

Then

$$\prod_{\nu=1}^{n} \cos\omega\tau_\nu \approx \prod_{\nu=1}^{n} [1 - \tfrac{1}{2}(\omega\tau_\nu)^2] \approx \exp\left(-\tfrac{1}{2}\omega^2 \sum_{\nu=1}^{n} \tau_\nu^2\right).$$

If all τ_ν were equal this would lead to $\exp\{-n\omega^2\bar{\tau}^2/2\}$. Assuming more realistically a stochastic distribution with distribution function $e^{-\tau/\bar{\tau}}$ corresponding to a Poisson distribution for the counting rate within a given time interval, then $\sum_\nu \tau_\nu^2 = 2n\bar{\tau}^2 = 2\bar{\tau}t$ for large n and one obtains

$$\delta_t = \delta_0 e^{-\lambda t}, \text{ with } \lambda = \omega \times \omega\bar{\tau} = \omega \times \bar{\tau}/T. \tag{6}$$

Thus the relaxation time $1/\lambda = T \times T/\bar{\tau} \gg T$ is much larger than $T = 1/\omega$ if (5) holds. This blocking of the states φ_1 and φ_2 by stochastic "reduction of the state vector" prevents oscillation between φ_1 and φ_2 and is responsible for enhanced stability.

The same analysis may be performed also if the Hamiltonian has not the high symmetry (parity conservation) with respect to φ_1 and φ_2 as assumed above. If its eigenstates are $\varphi = \alpha\varphi_1 + \beta\varphi_2$ and $\varphi' = -\beta^*\varphi_1 + \alpha^*\varphi_2$ with $|\alpha|^2 + |\beta|^2 = 1$, one obtains $\delta_t = \delta_0 \prod_\nu (1 - 8|\alpha\beta|^2 \sin^2\frac{1}{2}\omega\tau_\nu)$. With (5) this leads to the replacement of λ in (6) by $\lambda = 4|\alpha\beta|^2 \times \omega\bar{\tau}/T \leq \omega\bar{\tau}/T$ where one notes that $|\alpha|^2 + |\beta|^2 = 1$ implies $4|\alpha\beta|^2 \leq 1$ with equality holding in the symmetric case $|\alpha|^2 = |\beta|^2 = \frac{1}{2}$.

The model considered so far is rather schematic in several respects. First of all it is clear that, in particular for microscopic systems, the interaction with a probe like a photon does not necessarily lead to orthogonal states φ_1^P and φ_2^P for given incoming state φ_0^P. If it does not, then one obtains only a *partial reduction* of the density matrix ρ:

$$\begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix} \rightarrow \begin{pmatrix} \rho_{11} & \zeta\rho_{12} \\ \zeta^*\rho_{21} & \rho_{22} \end{pmatrix}, \tag{7}$$

where $\zeta = (\varphi_2^P, \varphi_1^P)$. Since the states are normalized to unity, $|\zeta| < 1$ unless $\varphi_1 = \alpha\varphi_2$. After m reduction with ζ_μ the off-diagonal elements are multiplied by

$$\zeta^{(m)} = \prod_{\mu=1}^{m} \zeta_\mu,$$

which approaches zero for large m if $|\zeta_\mu| < 1$. This is in line with the previously discussed relation to the process of measurement since even if the object may not be localized with one photon, in example 1 for instance, this may still be done with many of them, which then may be considered as one probe[1] leading to a (almost) complete reduction, or as many individual probes each leading to a partial reduction. From the physics involved one thus expects the blocking effect to be working similarly where, of course, one has to require now that m partial reductions with $|\zeta^{(m)}| \ll 1$ should take place in a time interval which is short compared to $T = 1/\omega$. This is borne out also by the formal analysis whose result is given here for symmetric (parity-conserving) Hamiltonian as in Eqs. (3)–(6). Under the condition that subsequent reductions are statistically independent one can—for the analysis of the average relaxation—replaced τ_ν by $\bar{\tau}$, τ_ν^2 by $2\bar{\tau}$ (see above), and the ζ_ν by their *average* ζ with $|\zeta| < 1$. The relevant condition replacing (5) turns out to be

$$\omega\bar{\tau}/(1 - |\zeta|) \equiv \omega\tau_0 \ll 1. \tag{8}$$

After a sequence of n evolutions under H as in (3) and subsequent reductions according to (7), an initial density matrix $\rho(0)$ turns into $\rho(t)$, where $t = \sum_\nu \tau_\nu = n\bar{\tau}$,

with[4]

$$\delta_t = e^{-\lambda' t}\left\{\delta_0 - 2\,\mathrm{Im}\left[\frac{\omega\bar{\tau}}{1-\zeta}(1-\zeta^n)\rho_{12}(0)\right] + O_1(\omega^2\tau_0^2)\right\},$$

$$\rho_{12}(t) = e^{-\lambda' t}\left\{\zeta^n\rho_{12}(0) + \frac{i}{2}\frac{\omega\bar{\tau}\zeta}{1-\zeta}(1-\zeta^n)\delta_0 + O_2(\omega^2\tau_0^2)\right\}, \quad \lambda' = \omega^2\bar{\tau}\,\mathrm{Re}\,\frac{1}{1-\zeta} \equiv \omega^2\tau_{\mathrm{eff}}, \tag{9}$$

where $\delta_t = \rho_{11}(t) - \rho_{22}(t)$ as in Eq. (4). Now $|\zeta|^n = \exp(t \ln|\zeta|/\bar{\tau}) < \exp(-t/\tau_0)$, implying that the off-diagonal element ρ_{12} is effectively reduced to zero after a run-in time of a few times τ_0. The condition (8) implies $\tau_0 \ll \omega^{-1} \ll \lambda^{-1}$, and $|\omega\bar{\tau}/(1-\zeta)| \ll 1$ in (9) so that spontaneous symmetry breaking and blocking is recovered with modified relaxation time $1/\lambda'$ as expected. This may, of course, be generalized to nonsymmetric evolution as discussed for complete reduction.

In conclusion it is seen that even if single reductions are very weak, i.e., if $(1 - |\zeta|) \ll 1$, spontaneous symmetry breaking and blocking persist as long as (8) holds. This ensures a wide applicability of this phenomenon.

The second oversimplification of the model analyzed is that H^O is two dimensional and that φ_1^O and φ_2^O are reproduced exactly. Of course, the whole discussion can be generalized to l different mutually orthogonal states $\varphi_1^O \cdots \varphi_l^O$ without any principal change. Of some interest, however, is the generalization to groups of states or subspaces H_i^O of H^O such that starting from given initial states $\varphi_i^O \in H_i^O$ subsequent interactions with probes not necessarily reproduce φ_i^O but still lead to (not necessarily pure) states which are again in H_i^O for all i. This interaction

then leads to the depletion of matrix elements of the density matrix between states from different $H_i{}^O$ in the same way and under the same circumstances as for the case of one-dimensional $H_i{}^O$ analyzed explicitly above. The general features of the blocking effect are expected to prevail.

In conclusion I restate the main ingredients of the mechanism presented here. It is based on the interaction of the object under consideration with a background of probes in its natural surrounding. The relevant background may consist for instance of electromagnetic or corpuscular radiation and perhaps even neutrinos, or molecules within a gas or fluid, and eventually also phonons, etc., in a solid, depending on the kind of object studied. The characteristic behavior of the object under the influence of this background is governed by two criteria:

(i) *The inertness criterion* (a) singles out the *inert states* to which reduction takes place such that (b) the interaction with the background (probes) does not induce transitions between these states. It thus defines the axes of spontaneous symmetry breaking. For macroscopic bodies it obviously singles out macroscopically localized states. If some interactions violate condition (b) independent of the choice of inert states, the state relaxes to unit density matrix, i.e., "total chaos." Of course, small violation of condition (b) can be tolerated as long as the relaxation time obtained from these transitions is long as compared to the other characteristic times of the problem at hand.

(ii) The *frequency criterion* for the blocking effect defines the regime in which the latter dominates over the free evolution under the Hamiltonian describing the object. In the cases calculated here it is given by Eqs. (5) and (8).

From the thermodynamical point of view the frequency criterion implies that the object is in a temperature bath with $kT \gg \hbar\omega$ (T = absolute temperature, k = Boltzmann constant) implying unit equilibrium density matrix. However, under the conditions of the inertness criterion, *the shorter* $\bar{\tau}$ *and the smaller* ζ, *i.e., the more intense the coupling, the longer is the relaxation time* $1/\lambda$ *or* $1/\lambda'$.

The mechanism of spontaneous symmetry breaking and blocking of metastable states so obtained applies to many macroscopic systems and compounds which would not be stable otherwise and thus seems to play a crucial role for the stability and classical property of the macroscopic world.[5] It eliminates the need for special (cosmological) initial conditions in order to obtain macroscopically localized states.[6] It is expected to be important also in molecular physics for the localization of atoms in molecules (and thus the applicability in the widely used Born-Oppenheimer approximation), guaranteeing in particular the stability of steroisomers, chiral molecules (such as sugar), etc. To what extent it operates also for intrinsic degrees of freedom (including P and/or CP) of elementary particles is an open question, which seems, however, certainly worthwhile studying.

This work was supported in part by the Swiss National Science Foundation.

---

[1]The connection with the problem of measurement and the interpretation of quantum mechanics will be discussed in more detail elsewhere. Here I only remark that starting from a purely statistical (ensemble) interpretation of quantum mechanics one obtains in this way under macroscopic conditions (see example 1 below) a description of individual systems.

[2]In *principle* one could recover coherence by including in the observation the final states of *all* such photons.

[3]Of course I implicitly assume also that the collision time between object and probe is short compared to τ_v. This is, however, relevant only for the exact calculation, not for the general feature of the effect discussed. (See the partial reduction discussed below.)

[4]$|O_i(x^2)| \leq C_i\, x^2$ for $x \ll 1$ with C_i independent of n.

[5]For instance, $T = 1/\omega \approx 1$ yr and $\tau \approx 10^{-5}$ sec imply $1/\lambda \approx 3 \times 10^{12}$ yr according to (6). On the other hand, for molecules with $T = \omega^{-1} \lesssim 10^{-2}$ sec it may be possible to study the blocking effect experimentally, though perhaps not in the extreme limit (8), for instance in a molecular beam crossing laser light or going through a gas of appropriate pressure.

[6]For a recent synopsis (with references) of other approaches to establish quantum mechanics as a general theory whose applicability includes the classical domain see J.-M. Lévy-Leblond, in *Quantum Mechanics, A Half Century Later*, edited by J. Leite Lopes and M. Paty (Reidel, Dordrecht, 1977), pp. 187–206. Note, however, that no attempt is made here to prove a reduction of the form (2) for the combined system described on $H^O \otimes H^P$. Unlike other approaches the probe can therefore be a simple microscopic system, which is obviously an important feature. On the other hand, no coherence can prevail between different "pointer" positions (or living and dead cats, etc., so to speak) in the case of a *macroscopic* apparatus under usual conditions, i.e., in the presence of radiation, air, etc. (necessary for living cats).

LETTERS

Coherently wired light-harvesting in photosynthetic marine algae at ambient temperature

Elisabetta Collini[1]*†, Cathy Y. Wong[1]*, Krystyna E. Wilk[2], Paul M. G. Curmi[2], Paul Brumer[1] & Gregory D. Scholes[1]

Photosynthesis makes use of sunlight to convert carbon dioxide into useful biomass and is vital for life on Earth. Crucial components for the photosynthetic process are antenna proteins, which absorb light and transmit the resultant excitation energy between molecules to a reaction centre. The efficiency of these electronic energy transfers has inspired much work on antenna proteins isolated from photosynthetic organisms to uncover the basic mechanisms at play[1–5]. Intriguingly, recent work has documented[6–8] that light-absorbing molecules in some photosynthetic proteins capture and transfer energy according to quantum-mechanical probability laws instead of classical laws[9] at temperatures up to 180 K. This contrasts with the long-held view that long-range quantum coherence between molecules cannot be sustained in complex biological systems, even at low temperatures. Here we present two-dimensional photon echo spectroscopy[10–13] measurements on two evolutionarily related light-harvesting proteins isolated from marine cryptophyte algae, which reveal exceptionally long-lasting excitation oscillations with distinct correlations and anti-correlations even at ambient temperature. These observations provide compelling evidence for quantum-coherent sharing of electronic excitation across the 5-nm-wide proteins under biologically relevant conditions, suggesting that distant molecules within the photosynthetic proteins are 'wired' together by quantum coherence for more efficient light-harvesting in cryptophyte marine algae.

Cryptophytes are eukaryotic algae that live in marine and freshwater environments. They are members of an evolutionary group notable because their photosynthetic apparatus was acquired from red algae by a sequence of endosymbiotic events. As a result, cryptophyte photosynthetic antenna proteins (phycobiliproteins) exhibit exceptional spectral variation between species because they use mainly tunable linear tetrapyroles (bilins) for light-harvesting. Another remarkable feature of cryptophytes is that they can photosynthesize in low-light conditions, which suggests that the absorption of incident sunlight by phycobiliprotein antennae in the intrathylakoid space[14] and the subsequent transfer of that energy among these proteins and eventually to the membrane-bound photosystems is particularly effective[15]. Theory indicates that fast energy transfer is facilitated by small interchromophore separations[2], yet the average nearest-neighbour centre-to-centre separation of chromophores within cryptophyte light-harvesting antenna proteins (Fig. 1) is ~20 Å (ref. 16)—about double that for the major light-harvesting protein in plants. To explore how a light-harvesting antenna can function efficiently with such a counter-intuitive design, we study the antennae of two marine cryptophytes, phycoerythrin PE545 from *Rhodomonas* CS24 and phycocyanin PC645 from *Chroomonas* CCMP270 at ambient temperature (294 K) using two-dimensional photon echo (2DPE) spectroscopy[10–13].

PC645 contains eight light-absorbing bilin molecules covalently bound to a four-subunit protein scaffold[17]. Its structure, determined to 1.4-Å resolution by X-ray crystallography[18] and shown in Fig. 1a, exhibits approximate twofold symmetry. A dihydrobiliverdin (DBV) dimer (green) located in the centre of the protein and two mesobiliverdin (MBV) molecules (blue) located near the protein periphery give rise to the upper half of the complex's absorption spectrum (Fig. 1c), spanned by our laser pulse spectrum. The electronic coupling of ~320 cm^{-1} between the DBV molecules C and D (labelled according to the protein subunit that binds them) leads to delocalization of the excitation and yields the dimer electronic excited states, or so-called molecular excitonic states[19], labelled DBV_+ and DBV_-. Excitation

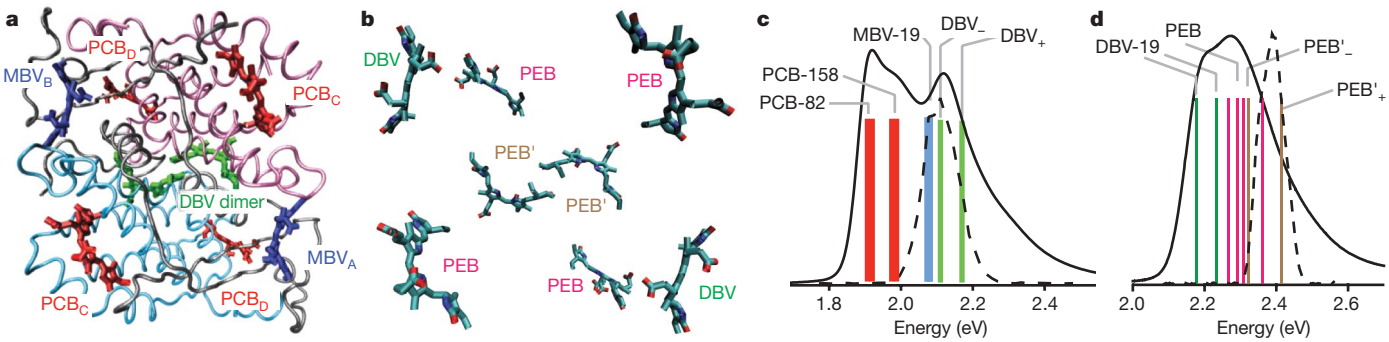


Figure 1 | Structure and spectroscopy of cryptophyte antenna proteins. **a**, Structural model of PC645. The eight light-harvesting bilin molecules are coloured red (PCB), blue (MBV) and green (DBV). **b**, Chromophores from the structural model for PE545 showing the different chromophore incorporation. **c**, Electronic absorption spectrum of isolated PC645 protein in aqueous buffer (294 K). The approximate absorption energies of the bilin molecules are indicated as coloured bars. **d**, Electronic absorption spectrum of isolated PE545 protein in aqueous buffer (294 K) with approximate absorption band positions indicated by the coloured bars. The spectrum of the ultrafast laser pulse is plotted as a dashed line in **c** and **d**.

[1]Department of Chemistry, Institute for Optical Sciences and Centre for Quantum Information and Quantum Control, University of Toronto, 80 St George Street, Toronto, Ontario, M5S 3H6 Canada. [2]School of Physics and Centre for Applied Medical Research, St Vincent's Hospital, The University of New South Wales, Sydney, New South Wales 2052, Australia.
†Present address: Dipartimento di Scienze Chimiche, Università di Padova, via Marzolo 1, 35100, Padova, Italy.
*These authors contributed equally to this work.

energy absorbed by the dimer flows to the MBV molecules. which are each 23 Å from the closest DBV, and ultimately to four phycocyanobilins (PCB, coloured red) that absorb in the lower-energy half of the absorption spectrum.

The structure of PE545 (Fig. 1b) is closely related to that of PC645 except that the bilin types differ[16,20]. The lowest-energy chromophores are DBV bilins. The dimer consists of phycoerythrobilin chromophores PEB′, with the prime indicating they are doubly covalently bound to the protein. The remaining chromophores are singly bound PEBs. The electronic couplings between the chromophores are reported elsewhere[21]. The approximate absorption spectrum and band positions are shown in Fig. 1d.

For the experiments the proteins were isolated from the algae and suspended at low concentration in aqueous buffer at ambient temperature (294 K). The femtosecond laser pulse (25-fs duration) excites a coherent superposition of the antenna protein's electronic-vibrational eigenstates (absorption bands). The initial state of the system is thus prepared in a non-stationary state[22], where electronic excitation is localized to a greater or lesser degree compared to the eigenstates. The time-dependent solution to quantum dynamics for electronically coupled molecules with this initial condition predicts that excitation subsequently oscillates among the molecules under the influence of the system Hamiltonian until the natural eigenstates are restored owing to interactions with the environment. 2DPE provides a means of observing this experimentally, enabling us to explore the significance of quantum coherence.

Representative 2DPE data for PC645 are shown in Fig. 2 with positions on the diagonal assigned to absorption bands. Rich features such as cross-peaks and excited state absorptions are evident. In the 2DPE experiment the two-pulse excitation sequence (sweeping $\tau > 0$) can prepare population density, for example $|\mathrm{DBV}_-\rangle\langle\mathrm{DBV}_-|$, that evolves during the delay time T and can be probed as a bleach signal on the diagonal part of a rephasing 2DPE spectrum. Alternatively, off-diagonal contributions like $|\mathrm{DBV}_+\rangle\langle\mathrm{DBV}_-|$ can be excited when the pump pulse sequence interacts coherently with both absorption bands. The resulting signal will be probed as a cross-peak above the diagonal in rephasing spectra that oscillate as a function of T with frequency $\phi = 2\pi\left(E_{\mathrm{DBV}_+} - E_{\mathrm{DBV}_-}\right)/h$ because it carries a phase $\exp(-i\phi T)$. Similarly the complementary coherence $|\mathrm{DBV}_-\rangle\langle\mathrm{DBV}_+|$ will contribute a cross-peak below the diagonal in rephasing spectra that will carry an opposite phase, $\exp(+i\phi T)$.

These predicted coherent oscillations can be reproducibly seen in our 2DPE spectra by plotting the intensity of rephasing spectra at lower and upper cross-peaks as a function of waiting time T (Fig. 2b and c). The red line indicates the cross-peak above the diagonal, the black line is that below. As mentioned above, the upper and lower cross-peak oscillations should differ by a phase factor determined by the sign of the energy difference between the states in superposition, leading to anti-correlated upper and lower cross-peak beats with a dominant frequency component equal to the eigenvalue energy difference. Such behaviour is indeed clearly seen in the experimental data, with the anti-correlated oscillations providing striking evidence that both DBV dimer and DBV–MBV electronic superposition states persist for more than 400 fs after photo-excitation. It is remarkable that electronic coherence spans from the DBV dimer to the peripheral MBV molecules, over a distance of 25 Å.

The PC645 cross-peak beating is complex; multiple frequencies arise for the same reason they do in a simple mechanical system comprised of a mass connected by a weak spring to a pair of masses coupled by a strong spring. And because these data were recorded at room temperature, line broadening is significant, so that overlapping bands partly obscure oscillating features. Nevertheless, Fourier transforms of these data (Supplementary Fig. 1) suggest the presence of frequencies in these beating patterns that can be related to the frequency differences between absorption bands. A careful global analysis of the data (see Supplementary Information) provides evidence that the oscillations in the 2DPE data can be decomposed into components corresponding to frequency differences between absorption bands and that—most importantly—the cross-peak beats at each frequency are anti-correlated.

For a comparison with the PC645 results, we also undertook experiments on the PE545 antenna protein by exciting the blue side of the absorption. A typical rephasing 2DPE spectrum (that is, scanned so that $\tau > 0$) is shown in Fig. 3a. To show beats across the entire anti-diagonal slice through the PEB/PEB′ cross-peaks we plot the intensity of the 2DPE rephasing spectrum along the anti-diagonal line drawn in Fig. 3a as a function of population time T (Fig. 3b).

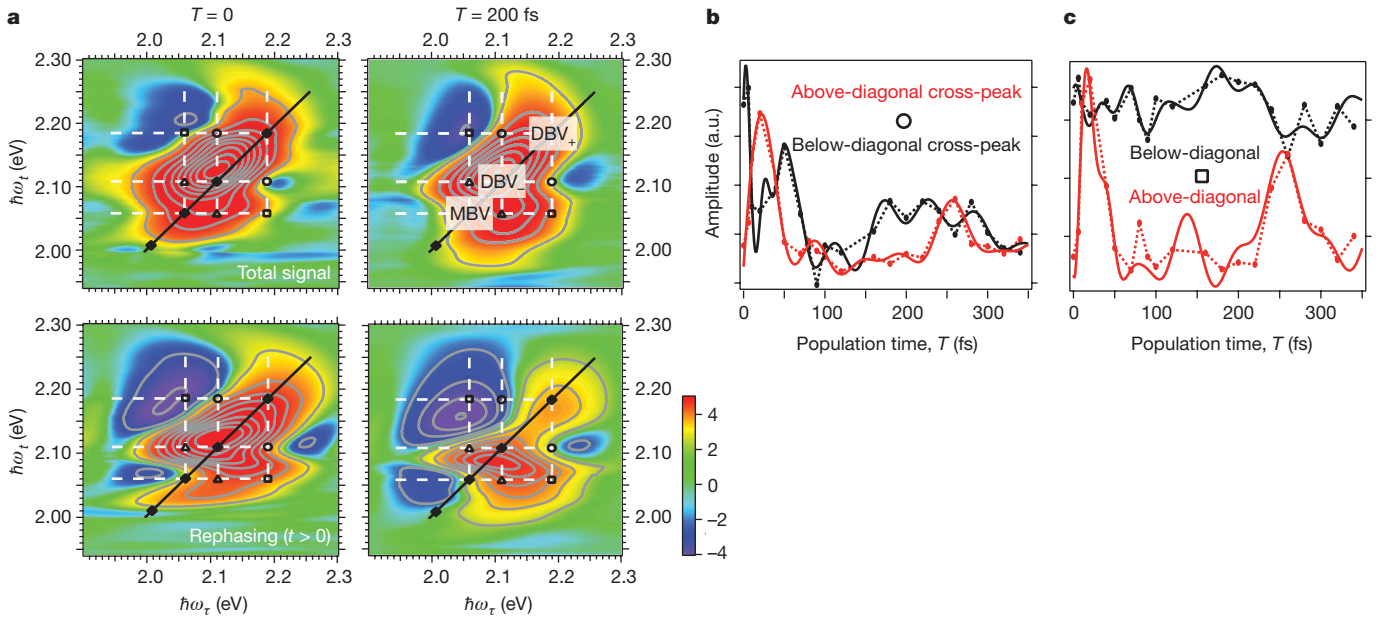


Figure 2 | Two-dimensional photon echo data for PC645. a, The left column shows the total real 2DPE spectrum recorded for PC645 at zero waiting time ($T = 0$), together with the rephasing contribution to this signal. The right column shows the data for $T = 200$ fs. The 2DPE spectra show the signal intensity on an arcsinh scale (colour scale, arbitrary units) plotted as a function of coherence frequency ω_τ and emission frequency ω_t. **b**, Intensity of the DBV dimer cross-peaks (open circle) as a function of time T. **c**, Intensity of the MBV–DBV_+ cross-peaks (open square) as a function of time T. The dashed lines interpolate the data points (solid circles). The solid line is a fit to a sum of damped sine functions (Supplementary Information). a.u., arbitrary units.

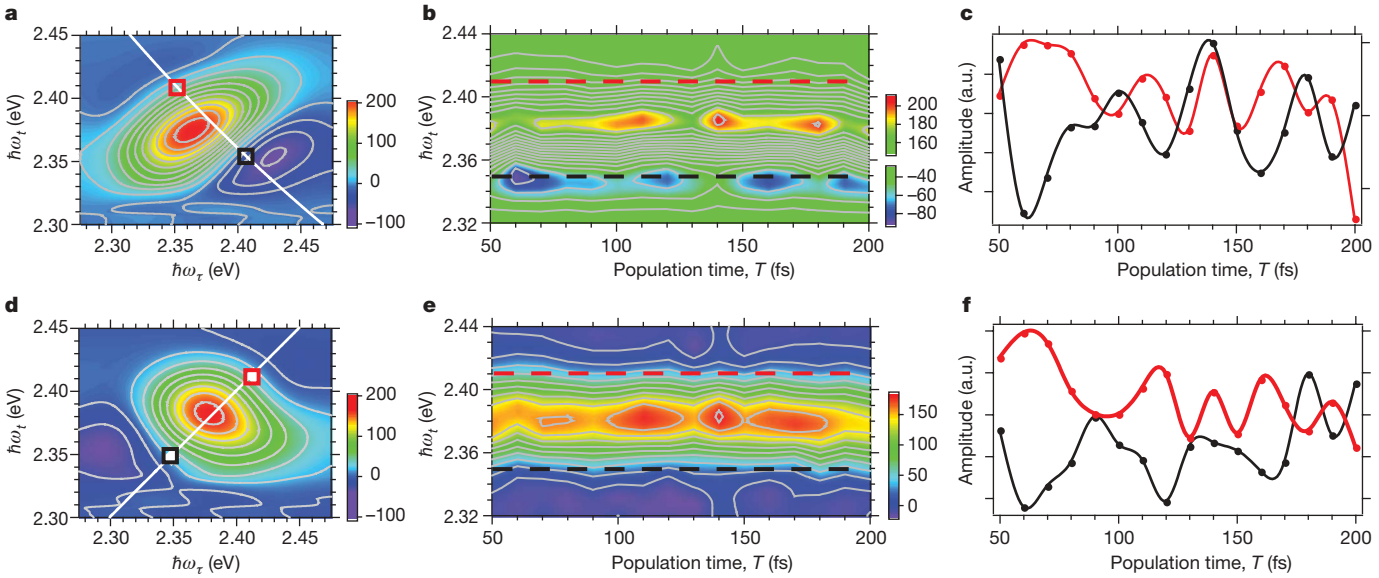


Figure 3 | Two-dimensional photon echo data for PE545. **a**, 2DPE spectrum (rephasing real signal) for PE545 recorded at $T = 100$ fs. **b**, The intensity of the 2DPE rephasing spectrum along an anti-diagonal slice through the cross-peaks versus population time T. Upper and lower cross-peaks are indicated by red and black dashed lines respectively. **c**, Intensity oscillations in the cross-peaks (red and black squares in **a**). **d**, 2DPE spectrum (non-rephasing real signal) for PE545 recorded at $T = 100$ fs. **e**, **f**, As for **b** and **c** but for the 2DPE non-rephasing spectrum along the diagonal slice. The 2DPE spectra are plotted on a linear intensity scale.

Oscillations of the main bleach and excited-state absorption peaks are clearly evident. The beats in the centre of this plot are cross-peaks excited because the PEB and PEB′ absorption bands overlap owing to spectral line broadening. The cross-peaks, indicated by dashed horizontal lines in Fig. 3b (red is the cross-peak above the diagonal, the black line is that below), are more clearly seen in Fig. 3c, in which the beats are well resolved and markedly anti-correlated—a signature of quantum coherence. These oscillations are directly analogous to those observed for PC645 (compare Fig. 2b).

Figure 3d–f shows plots similar to those in Fig. 3a–c, but for non-rephasing spectra ($\tau < 0$) of PE545. The same electronic coherences giving rise to oscillating cross-peaks in rephasing spectra are predicted to cause oscillations at the diagonal positions of non-rephasing 2DPE spectra[23]. Indeed, we observe a clear phase relationship between beats within the rephasing and non-rephasing spectra and also between the data sets, which is compelling evidence for the presence of long-lived quantum coherence. The first ~130 fs of these spectra reproducibly show clear oscillations with a period of ~60 fs ($\nu \approx 500\ \mathrm{cm}^{-1}$). After this time the oscillation pattern becomes more complicated, suggesting that the initial coherence may evolve owing to coupling with other molecules in the protein.

In our experiments the light-harvesting process in both PC645 and PE545 antenna proteins involves quantum coherence at ambient temperature, suggesting that coherence may more generally be used by cryptophyte algae. Quantum coherence occurs in an intermediate regime of energy transfer where there is a complex balance between quantum interference among electronic resonances and coupling to the environment causing decoherence[24]. There still remains the question of precisely how quantum coherence can persist for hundreds of femtoseconds in these biological assemblies. In an isolated molecule, electronic decoherence arises from the decay of the overlap $S(t) = \langle \nu_2(t) \mid \nu_1(t) \rangle$ between the unobserved vibrational wavepackets $|\nu_1(t)\rangle$ and $|\nu_2(t)\rangle$ associated with the lower and upper electronic states respectively[25,26]. In 2DPE experiments, the observable includes both vibrational and electronic components and, as such, decoherence due to decay of $S(t)$ is not manifest in the data. Rather, the slow decay of electronic coherence reflects the interaction of vibronic superposition states with the external environment.

Recent studies have attributed the slow dephasing of electronic coherence to the presence of shared or correlated motions in the surrounding environment[6,7,27,28]. In this context, we note that, unlike most photosynthetic pigments that are non-covalently complexed to their protein environment (chlorophyll via histidine residues, for example), the bilins in PC645 are covalently bound to the protein backbone. Covalent attachment of the chromophores to their protein environment may support or strengthen correlated motions between chromophores and protein and thus be an important factor in slowing down decoherence in cryptophyte antenna proteins at ambient temperature, thereby differentiating them from many other photosynthetic light-harvesting antennae. We also note that the precise manifestation of long-lived quantum-coherence depends on the photo-excitation conditions[22,29,30], and cryptophyte algae are obviously using sunlight that does not arrive in the form of laser pulses as used in our experiments. Nevertheless, the couplings giving rise to the long-lived quantum coherence that we clearly observe at ambient temperature will still be present and strongly suggest that quantum effects facilitate the efficient light-harvesting by cryptophyte algae. That is, long-lived quantum coherence can facilitate energy transfer by 'wiring' together the final energy acceptors (PCB in the case of PC645 and DBV for PE545) across a single protein unit, and thereby help to compensate for the exceptionally large average interchromophore separations in these antenna proteins.

METHODS SUMMARY

Cryptophyte *Chroomonas* sp. (CCMP270 strain, National Culture Collection of Marine Phytoplankton, Bigelow Laboratory for Ocean Sciences, USA) and *Rhodomonas* sp. (CS24) were cultivated and harvested, and the phycobiliproteins were isolated by usual procedures[20]. 2DPE experiments were performed as described in refs 11 and 27. The laser pulse duration and chirp were measured using transient grating frequency resolved optical gating (TG-FROG) experiments on a solvent (typically ethanol). The time-bandwidth product was estimated to be about 0.53, close to the ideal transform-limited condition for a Gaussian pulse. During data collection for PC645, for any given population time T, τ was scanned from –200 to 200 fs with 0.25-fs steps. Each 2D map at a fixed T is the average of at least three separate scans, and each series of 2D scans at different T was further repeated on different days for comparison. For PE545, τ was scanned from –60 to 60 fs with 0.15-fs steps, and each 2D map is the average of two separate scans. Additionally, PE545 was measured at this wavelength numerous times with different T steps, on different days. The samples were moved after each scan and absorption spectra taken before and after each series of scans confirmed that the sample did not degrade during the measurements.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 14 July; accepted 17 December 2009.

1. Green, B. R. & Parson, W. W. (eds) *Light-Harvesting Antennas in Photosynthesis* (Kluwer, Dordrecht, 2003).
2. Scholes, G. D. Long-range resonance energy transfer in molecular systems. *Annu. Rev. Phys. Chem.* **54,** 57–87 (2003).
3. Jang, S., Newton, M. D. & Silbey, R. J. Multichromophoric Förster resonance energy transfer from B800 to B850 in the light harvesting complex 2: evidence for subtle energetic optimization by purple bacteria. *J. Phys. Chem. B* **111,** 6807–6814 (2007).
4. Cheng, Y. C. & Fleming, G. R. Dynamics of light harvesting in photosynthesis. *Annu. Rev. Phys. Chem.* **60,** 241–262 (2009).
5. van Grondelle, R. & Novoderezhkin, V. I. Energy transfer in photosynthesis: experimental insights and quantitative models. *Phys. Chem. Chem. Phys.* **8,** 793–807 (2006).
6. Engel, G. S. *et al.* Evidence for wavelike energy transfer through quantum coherence in photosynthetic systems. *Nature* **446,** 782–786 (2007).
7. Lee, H., Cheng, Y. C. & Fleming, G. R. Coherence dynamics in photosynthesis: protein protection of excitonic coherence. *Science* **316,** 1462–1465 (2007).
8. Mercer, I. P. *et al.* Instantaneous mapping of coherently coupled electronic transitions and energy transfers in a photosynthetic complex using angle-resolved coherent optical wave-mixing. *Phys. Rev. Lett.* **102,** 057402 (2009).
9. Feynman, R. P. Space-time approach to non-relativistic quantum mechanics. *Rev. Mod. Phys.* **20,** 367–387 (1948).
10. Jonas, D. M. Two-dimensional femtosecond spectroscopy. *Annu. Rev. Phys. Chem.* **54,** 425–463 (2003).
11. Brixner, T., Mancal, T., Stiopkin, I. V. & Fleming, G. R. Phase-stabilized two-dimensional electronic spectroscopy. *J. Chem. Phys.* **121,** 4221–4236 (2004).
12. Cho, M. H. Coherent two-dimensional optical spectroscopy. *Chem. Rev.* **108,** 1331–1418 (2008).
13. Abramavicius, D. *et al.* Coherent multidimensional optical spectroscopy of excitons in molecular aggregates; quasiparticle versus supermolecule perspectives. *Chem. Rev.* **109,** 2350–2408 (2009).
14. Spear-Bernstein, L. & Miller, K. R. Unique location of the phycobiliprotein light-harvesting pigment in the cryptophyceae. *J. Phycol.* **25,** 412–419 (1989).
15. van der Weij-De Wit, C. D. *et al.* Phycocyanin sensitizes both photosystem I and photosystem II in cryptophyte *Chroomonas* CCMP270 cells. *Biophys. J.* **94,** 2423–2433 (2008).
16. Wilk, K. E. *et al.* Evolution of a light-harvesting protein by addition of new subunits and rearrangement of conserved elements: crystal structure of a cryptophyte phycoerythrin at 1.63-Ångstrom resolution. *Proc. Natl Acad. Sci. USA* **96,** 8901–8906 (1999).
17. Wedemayer, G. J., Kidd, D. G., Wemmer, D. E. & Glazer, A. N. Phycobilins of cryptophycean algae: occurrence of dihydrobiliverdin and mesobiliverdin in cryptomonad biliproteins. *J. Biol. Chem.* **267,** 7315–7331 (1992).
18. Mirkovic, T. *et al.* Ultrafast light harvesting dynamics in the cryptophyte phycocyanin 645. *Photochem. Photobiol. Sci.* **6,** 964–975 (2007).
19. Scholes, G. D. & Rumbles, G. Excitons in nanoscale systems. *Nature Mater.* **5,** 683–696 (2006).
20. Doust, A. B. *et al.* Developing a structure-function model for the cryptophyte phycoerythrin 545 using ultrahigh resolution crystallography and ultrafast laser spectroscopy. *J. Mol. Biol.* **344,** 135–153 (2004).
21. Scholes, G. D. *et al.* How solvent controls electronic energy transfer and light harvesting. *J. Phys. Chem. B* **111,** 6978–6982 (2007).
22. Rhodes, W. Radiationless transitions in isolated molecules. the effects of molecular size and radiation bandwidth. *J. Chem. Phys.* **50,** 2885–2896 (1969).
23. Cheng, Y. C. & Fleming, G. R. Coherence quantum beats in two-dimensional electronic spectroscopy. *J. Phys. Chem. A* **112,** 4254–4260 (2008).
24. Rackovsky, S. & Silbey, R. Electronic-energy transfer in impure solids. 1. 2 molecules embedded in a lattice. *Mol. Phys.* **25,** 61–72 (1973).
25. Hwang, H. & Rossky, P. J. Electronic decoherence induced by intramolecular vibrational motions in a betaine dye molecule. *J. Phys. Chem. B* **108,** 6723–6732 (2004).
26. Franco, I., Shapiro, M. & Brumer, P. Femtosecond dynamics and laser control of charge transport in trans-polyacetylene. *J. Chem. Phys.* **128,** 244905 (2008).
27. Collini, E. & Scholes, G. D. Coherent intrachain energy migration in a conjugated polymer at room temperature. *Science* **323,** 369–373 (2009).
28. Beljonne, D., Curutchet, C., Scholes, G. D. & Silbey, R. Beyond Förster resonance energy transfer in biological and nanoscale systems. *J. Phys. Chem. B* **113,** 6583–6599 (2009).
29. Langhoff, C. A. & Robinson, G. W. Time decay and untangling of vibronically tangled resonances: naphthalene second singlet. *Chem. Phys.* **6,** 34–53 (1974).
30. Jang, S. Theory of coherent resonance energy transfer for coherent initial condition. *J. Chem. Phys.* **131,** 164101 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by the Natural Sciences and Engineering Research Council of Canada and the Australian Research Council. G.D.S. acknowledges the support of an EWR Steacie Memorial Fellowship.

Author Contributions E.C. performed the experiments on PC645 and analysed those data. C.Y.W. performed the experiments on PE545 and analysed those data. K.E.W. prepared the samples. P.M.G.C. and G.D.S. designed the research. P.B. and G.D.S. examined the interpretation of the results. G.D.S. wrote the paper. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to G.D.S. (gscholes@chem.utoronto.ca).

METHODS

Cryptophyte *Chroomonas* sp. (CCMP270 strain, National Culture Collection of Marine Phytoplankton, Bigelow Laboratory for Ocean Sciences, USA) was cultured at 20 °C under constant low light illumination (12-V white fluorescent tubes, 300 lx at 0.3 m) in a modified 'Fe' medium. The algal cells were harvested and passed twice through a French press cell at a pressure of 1,000 psi. The resultant solution was centrifuged, yielding a supernatant solution containing PC645. PC645 was isolated using gradual ammonium sulphate precipitation from 0 to 80%. The pellets were re-suspended in a minimal volume of 25 mM phosphate buffer at pH 7.1. Further purification was continued using a combination of ion-exchange and size-exclusion chromatography. About 20 ml of pure PC645 was concentrated to approximately 200 µl using a 10-kDa Amicon Centriprep and then frozen using liquid nitrogen before being stored at –80 °C. A buffer solution, prepared from a 25 mM solution of HEPES (4-(2–hydroxyethyl)–1–piperazineethanesulphonic acid) in deionized water, adjusted to pH = 7.5 by the addition of concentrated NaOH solution, was used to prepare dilute PC645 samples for the experiments.

Rhodomonas sp. (CS24) was cultivated and harvested as previously reported[20]. Cell pellets were re-suspended in buffer A (0.05 M Mes (pH 6.5) with 1 mM NaN_3) and homogenized in a Teflon glass homogenizer followed by passage though a French press at a pressure of 1,000 psi. The resultant solution was centrifuged for 30 min at 17,000*g*, producing a pellet of cell debris containing thylakoid membranes and a supernatant containing phycoerythrin. PE545 was purified from the supernatant as described elsewhere[20].

2DPE experiments were performed as described in refs 11 and 27. A Ti:sapphire regeneratively amplified laser system was used to pump a Noncolinear Optical Parametric Amplifier (NOPA) to produce 25-fs duration pulses centred at 590 nm for PC645 experiments or at 520 nm for PE545 for the results reported here, with a spectral bandwidth of about 25 nm and repetition rate of 1 kHz. The laser pulse duration and chirp were measured using TG-FROG experiments on a solvent (typically ethanol). The time-bandwidth product was estimated to be about 0.53, close to the ideal transform-limited condition for a Gaussian pulse. The pulse from the NOPA was split by a 50% beam splitter and the two resulting beams were overlapped in a diffractive optic, producing two pairs of phase-locked beams in a boxcars phase-matched geometry. The delay time T was controlled by a motorized translation stage inserted in one beam path before the diffractive optic, whereas the delay time τ was introduced by means of movable glass wedge pairs, calibrated by spectral interferometry[11].

During data collection for PC645, for any given population time T, τ was scanned from –200 to 200 fs with 0.25-fs steps. Each 2D map at a fixed T is the average of at least three separate scans, and each series of 2D scans at different T was further repeated on different days for comparison. For PE545, τ was scanned from –60 to 60 fs with 0.15-fs steps, and each 2D map is the average of two separate scans. Additionally, PE545 was measured at this wavelength numerous times with different T steps, on different days. The samples were moved after each scan and absorption spectra taken before and after each series of scans confirmed that the sample did not degrade during the measurements.

To ensure that the local oscillator did not influence the response of the system, its intensity was attenuated by about three orders of magnitude relative to the other beams and the time ordering was set so that the local oscillator always preceded the probe by ~500 fs. The resulting local oscillator–signal interference intensity was focused into a 0.63-m spectrograph (25-µm slit) and recorded using a 16-bit, 400 × 1,600 pixel, thermo-electrically cooled charge-coupled device (CCD) detector. Subtraction of unwanted scatter contributions, Fourier windowing, transformation and phase retrieval were performed as reported previously[11].

Bernard Barbara

Statement

and

Readings

Abstract

Bernard Barbara

A short review will be given of experimental aspects of decoherence in solid state qubits, including magnetic and superconducting qubits. Most important decoherence mechanisms will be discussed, for single or ensembles of qubits with or without excitation pulse, with different dimensions and degrees of complexity. More specific subjects, such as the effects of decoherence on magnetic molecules or in quantum phase transitions, will be tackled.

LETTERS

Quantum oscillations in a molecular magnet

S. Bertaina[1]†, S. Gambarelli[2], T. Mitra[3], B. Tsukerblat[4], A. Müller[3] & B. Barbara[1,2]

The term 'molecular magnet' generally refers to a molecular entity containing several magnetic ions whose coupled spins generate a collective spin, S (ref. 1). Such complex multi-spin systems provide attractive targets for the study of quantum effects at the mesoscopic scale. In these molecules, the large energy barriers between collective spin states can be crossed by thermal activation or quantum tunnelling, depending on the temperature or an applied magnetic field[2–4]. There is the hope that these mesoscopic spin states can be harnessed for the realization of quantum bits—'qubits', the basic building blocks of a quantum computer—based on molecular magnets[5–8]. But strong decoherence[9] must be overcome if the envisaged applications are to become practical. Here we report the observation and analysis of Rabi oscillations (quantum oscillations resulting from the coherent absorption and emission of photons driven by an electromagnetic wave[10]) of a molecular magnet in a hybrid system, in which discrete and well-separated magnetic V_{15}^{IV} clusters are embedded in a self-organized non-magnetic environment. Each cluster contains 15 antiferromagnetically coupled $S = 1/2$ spins, leading to an $S = 1/2$ collective ground state[11–13]. When this system is placed into a resonant cavity, the microwave field induces oscillatory transitions between the ground and excited collective spin states, indicative of long-lived quantum coherence. The present observation of quantum oscillations suggests that low-dimension self-organized qubit networks having coherence times of the order of 100 μs (at liquid helium temperatures) are a realistic prospect.

In the context of quantum computing, it was recently discussed how the decoherence of molecular magnet spin quantum bits could be suppressed, with reference to the discrete low spin clusters V_{15} and Cr_7Ni (ref. 7; see also refs 8 and 14). In both systems, their low spin states cause weak environmental coupling[7], making them candidates for the realization of a long-lived quantum memory. Measurement of the spin relaxation time τ_2 in Cr_7Ni was subsequently reported and found to be interestingly large[15,16]; however, the important Rabi quantum oscillations were not observed, probably because electronic and nuclear degrees of freedom were too strongly linked to each other. As these oscillations have until now only been observed in non-molecular spin systems (see, for example, refs 17–20), it has remained an open question whether quantum oscillations could in principle be realized in molecular magnets[7,8]. This question is now answered by our observation of quantum oscillations of the Rabi type in V_{15}. The main reason for this success lies in the fact that the important pairwise decoherence mechanism[7,8] associated with dipolar interactions could be strongly reduced.

Before discussing the observed quantum oscillations, we first briefly describe the magnetic/electronic structure of the V_{15}^{IV} species as determined experimentally. Following the synthesis of the quasi-spherical mesoscopic cluster anion $\left[V_{15}^{IV}As_6^{III}O_{42}(H_2O)\right]^{6-}$ ($\equiv V_{15}$) nearly two decades ago (ref. 11), the properties of this molecule have received considerable attention (see, for example, refs 1, 11, 14, 21–25). The V_{15} cluster with an ~1.3 nm diameter exhibits an unique structure with layers of different magnetizations: a large central V_3^{IV} triangle is sandwiched by two smaller V_6^{IV} hexagons[11] (Fig. 1). The 15 $S = 1/2$ spins are coupled by antiferromagnetic super-exchange and Dzyaloshinsky–Moriya (DM) interaction[13,21–25] (see also refs 26, 27) through different pathways, which results in a collective low spin ground state with $S = 1/2$ (refs 12, 13, 24, 25).

Energy spectrum calculations for the full cluster spin space give two $S = 1/2$ (spin doublet) ground states slightly shifted from each other by DM interactions, and an $S = 3/2$ (spin quartet) excited state; these states are 'isolated' from a quasi-continuum of states lying at energy $E/k_B \approx 250$ K above the $S = 3/2$ excited state. These low-lying energy states can be obtained with a good accuracy using the generally accepted three-spin approximation (valid below 100 K), in which the spins of the inner triangle are coupled by an effective interaction $|J_0| \ll |J'|$ mediated by the spins of the hexagons[12,13,21–25] (Fig. 2 and Methods; J_0 and J' are shown in Fig. 1b).

The spin hamiltonian of V_{15} can be written as:

$$H = -J_0 \sum_{\substack{i,j=1 \\ (i<j)}}^{3} \mathbf{S}_i \mathbf{S}_j + \sum_{ij=12,13,31} \mathbf{D}_{ij}\left(\mathbf{S}_i \times \mathbf{S}_j\right) + A \sum_{i=1}^{3} \mathbf{I}_i \mathbf{S}_j + g\mu_B \mathbf{H} \sum_{i=1}^{3} \mathbf{S}_i \quad (1)$$

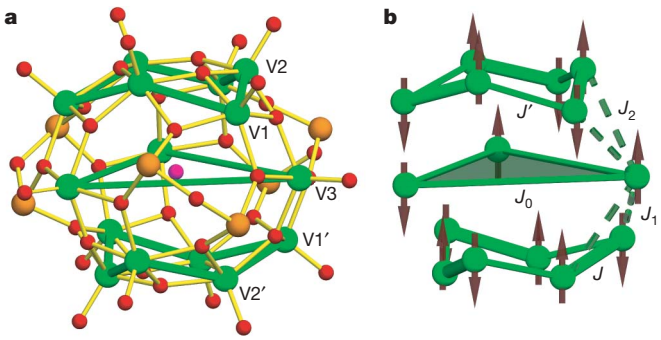


Figure 1 | Structure and exchange interaction pathways of the cluster anion $[V_{15}^{IV}As_6^{III}O_{42}(H_2O)]^{6-}$. a, The cluster is shown in ball-and-stick representation (green, V; orange, As; red, O). The outer V_6 hexagons are highlighted by thick green lines. A weak deviation from trigonal symmetry can be attributed to the water molecule located in the centre of the cavity (O of the encapsulated water molecule in purple) or/and to the presence of water in the lattice between molecules. The different types of V ions, namely V1, V2, V3, V1′ and V2′, are shown for the definition of different exchange pathways. **b**, Sketch showing the spin arrangement at low temperatures (three-spin approximation), emphasizing some of the exchange interaction pathways (J, J_1, J_2, J'). The coupling J_0 between the spins of the inner triangle is not direct but results from different exchange pathways through the hexagons. The magnetic layer system is defined by one V_3 triangle sandwiched by two V_6 hexagons (for further details see text).

[1]Institut Néel, CNRS, 25 Ave. des Martyrs, BP166, 38042 Grenoble Cedex 9, France. [2]Laboratoire de Chimie Inorganique et Biologique (UMR-E3 CEA-UJF), INAC, CEA-Grenoble, 17 Ave. des Martyrs, 38054 Grenoble Cedex 9, France. [3]Fakultät für Chemie, Universität Bielefeld, Postfach 100131, D-33501 Bielefeld, Germany. [4]Department of Chemistry, Ben-Gurion University of the Negev, PO Box 653, 84105 Beer-Sheva, Israel. †Present address: National High Magnetic Field Laboratory, Florida State University, 1800 East Paul Dirac Drive, Tallahassee, Florida 32310, USA.

where $\mathbf{D}_{ij}$ is the antisymmetric vector of the DM interaction associated with the pair ij, and A is the hyperfine coupling constant of the ^{51}V isotope (see below). The six components of $\mathbf{D}_{ij}$ can be expressed in terms of two parameters, namely D_Z (perpendicular to the plane) and D_{XY} (in-plane). The DM interaction removes the degeneracy of the two low-lying doublets and produces a first order zero-field splitting $\Delta_{DM} \approx \sqrt{3} D_Z$ (plus small second order corrections)[22–25]. The excited (quartet) state shows only a second order splitting caused by a small inter-multiplet mixing through the in-plane component of DM coupling, that is, $\Delta'_{DM} = -D_{XY}^2/8J_0$ (refs 24, 25). The energy separation between the doublet states and quartet state is given by $3J_0/2 \approx -3.67$K (refs 13, 21–25). Figure 2 shows the level scheme calculated by diagonalization of the hamiltonian (equation (1)), with only one free parameter $D_Z \approx 43$ mK adjusted to fit the positions of the measured resonances (a value close to that obtained from magnetization data[13,21,24]), and $D_{XY} = 0$, a choice conditioned by the fact that the transverse DM component has a negligible effect on resonance fields below 0.5 T (this is important in the calculation of transition probabilities only). To ensure legibility, hyperfine interactions are not included in Fig. 2 (they simply broaden the levels).

A new hybrid material, based on the use of a cationic surfactant—DODA ($[Me_2N\{(CH_2)_{17}Me\}_2]^+$)—as an embedding material for the anionic clusters, was developed for the present work (see Methods). The related frozen system contains V_{15} clusters integrated into the self-organized environment of the surfactant. The clusters—prepared according to ref. 11—were extracted from aqueous solution into chloroform by the surfactant DODA present in large excess. The surfactants, which wrap up the cluster anions, are amphiphilic cations, with their long hydrophobic tails pointing away from the cluster anions, enabling solubility in chloroform. The procedure ensures that the cluster anions cannot get into direct contact with one another; they are clearly separated by the surfactants (mean distance ~13 nm).

Electron paramagnetic resonance (EPR) experiments were performed on this hybrid material at ~4 K using a Bruker E-580 X-band continuous-wave (CW) and pulsed spectrometer operating at 9.7 GHz. The CW-EPR spectrum, recorded at 16 K on a frozen sample, corresponds precisely to that obtained in the solid state in a previous study[12]. In particular, the resonance field shows the same profile and line-width (~30 mT), compatible with the g-tensor values of a single crystal ($g_{//} = 1.98$ and $g_{\perp} = 1.95$). The measured transition width $W \approx 35$ mT is directly connected with the energy E occurring in the expression of decoherence calculated for a multi-spin molecule[7,8] (see below). Note that this transition width W should be associated with $S = 3/2$, the EPR spectrum being dominated by the excited quartet.

Rabi oscillations were recorded using a nutation pulse of length t, followed (after a delay greater than τ_2) by a $\pi/2-\pi$ sequence. Experimental results showed two different types of Rabi oscillations, corresponding to the resonant transitions 1, 2 and 3 for $S = 3/2$ spins, and 4, 5, 6 and 7 for $S = 1/2$ spins, here called '3/2' and '1/2', respectively (Fig. 3b and a, respectively). Although both types of oscillation are associated with the same collective degrees of freedom of the clusters, they show very different behaviour. In particular, the first type of Rabi frequency compares well with that of a single spin-3/2 system, whereas the Rabi frequency of the second type is much smaller than that of a single spin-1/2. This is a consequence of selection rules: the transition type '3/2' is always allowed, whereas the transitions 5 and 7 of the '1/2' type occur only due to transverse DM interactions or/and breaking of the C_3 symmetry[25] (Methods). Therefore we obtained Rabi oscillations with quite different frequencies, $\Omega_{R\,3/2} \approx 18.5 \pm 0.2$MHz and $\Omega_{R1/2} \approx 4.5 \pm 0.2$MHz, and a small ratio of transition probabilities (or intensities) $R < 6 \times 10^{-2}$ (Fig. 3, Methods). When the transition '1/2' is excited (by a single excitation pulse), a whole spectrum of Rabi oscillations is generated. The frequency of the detected oscillation depends on the characteristics of the detection pulse, such as its length or its amplitude (Fig. 3). This

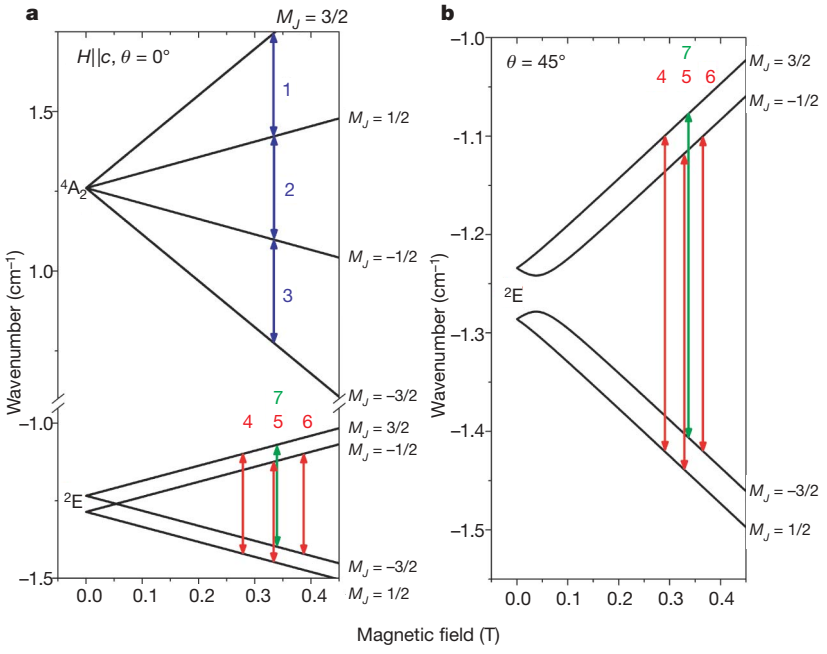


Figure 2 | Low-energy EPR transitions. These calculated diagrams used parameter values $g \approx 1.96$, $J_0 \approx -2.45$ K, $D_Z \approx 43$ mK and $D_{XY} = 0$ (see text for details). **a**, The magnetic field is parallel to the c axis. Whereas the orbital singlet $^4A_2$ ($S = 3/2$) gives the superposition of the three transitions 1, 2, and 3, the orbital doublet ^2E ($S = 1/2$) gives two inter-doublet transitions 4 and 6 which are basically allowed, as well as two intra-doublet transitions 5 and 7 (which are respectively allowed by transverse DM interactions and non-symmetrical exchange interactions due to a small deviation from the trigonal symmetry[24] (Methods)). Second order zero-field splitting of $^4A_2$ and small splitting of the lines 1, 2, and 3 is not shown. The M_J labels correspond to the quantization axis along the DM anisotropy field. **b**, Shown are the transitions 4, 5, 6, and 7 for the angle $\theta = 45^\circ$ between the field and cluster C_3-axis. The M_J labels correspond to the quantization axis along the field in the strong field limit. The boundaries of the measured resonance fields of Fig. 4 correspond to the field distribution given by the positions of labels 4, 5, 6 and 7 above the curves. The blue transitions correspond to $S = 3/2$; the red and the green transitions correspond to $S = 1/2$.

spectrum is due to the presence of an avoided level crossing and the special selection rules; these are caused by the uniaxial anisotropy introduced by the DM interactions in the spin-frustrated (orbitally degenerate) ground state giving the overlapping transitions 4–7 (Fig. 2). The glassy character of the investigated frozen material is also relevant here; this material contains different cluster orientations, leading to a distribution of transverse field components, which gives a scattering of the coefficients of the states entering in the two-level wavefunctions $|\varphi_1\rangle$ and $|\varphi_2\rangle$ and therefore a distribution of the Rabi frequencies $\Omega_{R\,1/2} \propto |\langle\varphi_1|S_+|\varphi_2\rangle|$ (Fig. 2 and Methods). Whereas the splitting of the excited quartet state in a magnetic field is almost isotropic, the distribution function of the associated Rabi frequency is very narrow.

An extension of the experiments shown in Fig. 3 to other values of the applied field showed that Rabi oscillations could be detected for each value of the applied field below 500 mT, while the transitions are inhomogeneously broadened. Figure 4 gives the result of a systematic investigation, consisting of the measurement of the spin-echo intensity at time $t = 0$ in a sweeping magnetic field. Two broad resonance distributions are observed, which correspond to the Rabi oscillations '3/2' and '1/2' of Fig. 3b and a, respectively, which were measured near the maxima $H_{3/2} \approx 357$ mT and $H_{1/2} \approx 335$ mT of the curves of Fig. 4. Whereas the nearly symmetrical type '3/2' distribution shows resonances which are optimally excited by pulse durations and powers similar to those generally used for isolated 3/2 spins, the asymmetrical type '1/2' distribution shows resonances requiring larger power and pulse length, confirming much smaller transition probabilities. The observed inhomogeneous widths ($\sim 50 \pm 10$ mT) result from the existence of different transitions—that is 1 to 3 and 4 to 7 shifted by the longitudinal field components associated with the glassy character of the frozen solution. The width of the resonance of type '1/2' (Fig. 4) fits the transition fields calculated from the hamiltonian (equation (1)) for the resonances 4 to 7 with limiting angles $\vartheta = 0$ and $\pi/2$ (Fig. 2), whereas the width of the resonance of type '3/2' is simply given by the unique resonance field of transitions 1 to 3 (Fig. 2 a). In both cases, the ^{51}V hyperfine interactions contribute equally to the resonance widths.

To conclude, it was possible to entangle the 15 spins of a molecular magnet—a complex system which, formally speaking, entails a Hilbert space of dimension $D_H = 2^{15}$ (Methods)—with photons by performing pulse EPR experiments on a frozen solution of randomly oriented and well separated clusters. Despite the complexity of the system[11–14,21–25] (involving in a formal consideration dozens of cluster electrons and nuclear spins of ^{51}V, ^{75}As and ^{1}H), long-lived Rabi oscillations[10] were generated and selectively detected. An analysis, based on the widely used three-spin approximation of V_{15} (refs 12, 13, 21–25; the related interactions are mediated by the 12 other spins) gives a global interpretation of the results.

The observed coherence on the microsecond timescale seems to be mainly limited by the bath of nuclear spins. Each V_{15} cluster is correspondingly weakly coupled to 36 first-neighbour protons of the six DODA methyl groups distributed around the cluster, and to two water protons at the cluster centre. According to the charge (6−) of V_{15}, six cationic DODA surfactants are relevant, with their positively charged parts (six dimethyl groups) attached to the O atoms of the cluster surface (see also ref. 28); the corresponding neutral hybrid just leads to the solubility in the organic solvent. The distance from the H atoms of a methyl group to a V^{IV} is ~0.45 nm. For this typical spin–proton distance, the half-width of the gaussian distribution of the coupling energy of a cluster/surfactant unit is $E \approx 3.5$ mK, giving, for the level separation $\Delta \approx 0.4$ K (Fig. 2), the coherence time[7,8]

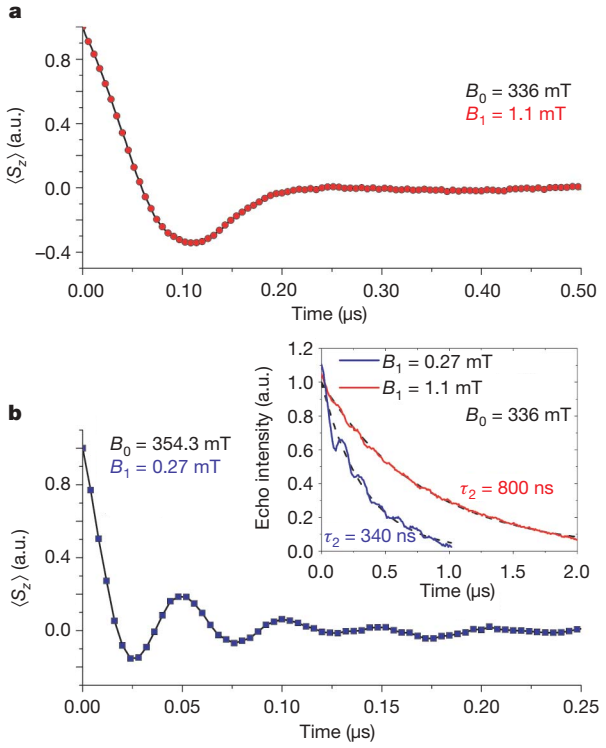


Figure 3 | Generation and detection of Rabi oscillations. **a**, Time evolution of the average spin $\langle S_z\rangle$ after a spin-echo sequence. The '1/2' type transition observed near the maximum of the corresponding resonance of Fig. 4 (B_0, 336.0 mT) requires unusually large excitation power $B_1 = 1.1$ mT and pulse length $T_{\pi/2} = 64$ ns. It corresponds to the transitions 4–7 of Fig. 2. The Rabi frequency $\Omega_R = 4.5$ MHz was selected by a detection pulse with characteristics $B_1 = 0.3$ mT and $T_{\pi/2} = 200$ ns. **b**, The '3/2' type transition with Rabi frequency $\Omega_R = 18.5$ MHz was excited near the maximum of the corresponding resonance of Fig. 4 ($B_0 = 354.3$ mT). It requires excitation and detection pulses similar to those usually used for a single spin of 3/2 ($B_1 = 0.27$ mT, $T_{\pi/2} = 16$ ns) and corresponds to the transitions 1–3 of Fig. 2. Inset, spin-echo intensity measured versus time for both oscillations. The coherence times τ_2 obtained from exponential fits are inverse functions of the spin values: 800 ns for $S = 1/2$ (red) and 340 ns for $S = 3/2$ (blue). Superimposed oscillations, mainly observed on the '3/2' type curve come from the precession of proton spins[19]. These oscillations correspond to only a weak perturbation of the Rabi coherence. Temperature, 4 K for all results shown.

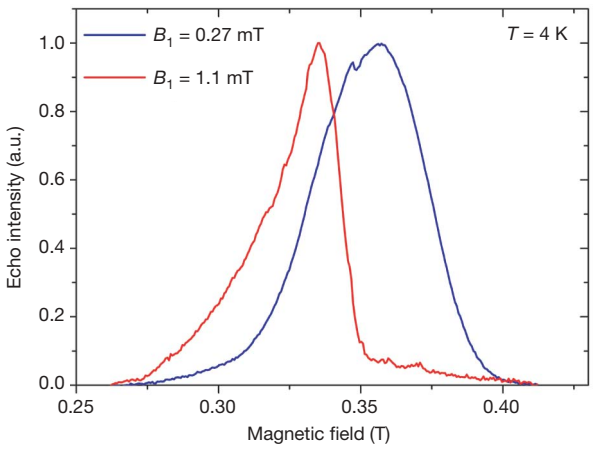


Figure 4 | Distribution of spin-echo intensities. The measurements were done in field sweep experiments for two excitation pulse configurations. The blue curve, corresponding to '3/2' type transitions (obtained with the excitation pulse $B_1 = 0.27$ mT and $T_{\pi/2} = 16$ ns), is nearly symmetrical and has a high transition probability. The red curve, corresponding to '1/2' type transitions (obtained with the excitation pulse $B_1 = 1.1$ mT and $T_{\pi/2} = 64$ ns), is asymmetrical and has a low probability (involving collective orbital degrees of freedom). The resonance fields form a 'band' due to random cluster orientations, while the corresponding distribution widths can be well explained by the dispersions of the resonance transitions 1–3 and 4–7.

$\tau_2^H = 4\pi\Delta/E^2 \approx 18\,\mu s$. The contribution of more distant neighbouring protons should reduce this value to a few microseconds. Regarding the decoherence effect from ^{51}V, the transition width $W \approx 35\,mK$ gives $E = W/2 \approx 17\,mK$ and $\tau_2^V = 0.75\,\mu s$, suggesting that the observed decoherence of the $S = 3/2$ resonances is almost entirely caused by the ^{51}V nuclear spins. The observed larger coherence time of the $S = 1/2$ transitions is presumably due to their smaller hyperfine coupling. In spite of the relatively high temperature of the measurement, the phonons' decoherence[7,8] $\tau_2^{ph} \propto S^{-4}$ is strongly lowered due to the low spin and anisotropy values involved in the electron–phonon[29,30] coupling $\propto |\langle i|S_yS_z + S_zS_y|f\rangle|^2$, giving $\tau_2^{ph} \approx 100\,\mu s$, that is, $\tau_2^{ph} \gg \tau_2^H > \tau_2^V$. Finally, the pairwise decoherence mechanism originating from electronic dipolar interaction[7], which is usually considered as the most destructive, is nearly negligible, owing to the strong dilution of the clusters that results from the surfactant environment. This allows weak dipolar interactions only ($\sim$0.5 μK) and very large coherence times ($\tau_2^{pw} \approx 100\,\mu s$). A comparison of the different decoherence mechanisms suggests that coherence times greater than 100 μs should be obtained in molecular magnets at liquid-helium temperatures if nuclear-spin-free molecules and deuterated surfactants are used.

The control of complex coherent spin states of molecular magnets—in which exchange interactions can be tuned by well defined chemical changes of the metal cluster ligand spheres—could finally lead to a way to avoid the 'roadblock' of decoherence. This would be particularly important in the case of self-organized one- or two-dimensional supramolecular networks, where well separated magnetic species could be addressed selectively, following different schemes already proposed for the molecular magnet option.

METHODS SUMMARY

When we refer to the three-spin approximation of V_{15} (refs 12, 13, 21–25), we consider the three spins located on each corner of the inner triangle (Fig. 1b). However these spins do not interact directly but via the other spins of the cluster. Strictly speaking, each hexagon contains three pairs of spins strongly coupled with $J \approx -800\,K$ ('dimers') and each spin of the inner triangle is coupled to two of those pairs, one belonging to the upper hexagon and one belonging to the lower hexagon ($J_1 \approx -150\,K$ and $J_2 \approx -300\,K$). This gives three groups of five spins with resultant spin $S = 1/2$ (superposition of 'entangled' states, coupled through inter-dimer hexagon superexchange $J' \approx -150\,K$), showing that, in fact, the three-spin approximation involves all of the 15 spins of the cluster and therefore the Hilbert space has the dimension $D_H = 2^{15}$ (D_H for the three-spin system is 2^3). This approximation simplifies the evaluation of the low-lying energy levels of the 15 'entangled' states of the V_{15} cluster. For $D_Z \neq 0$ the $S = 1/2$ orbital doublet ^2E, whose basis functions can be labelled by the quantum number of the total pseudo-angular momentum $M_J = M_L + M_S$, is associated with the pseudo-orbital momentum $M_L = +1$ or $M_L = -1$ (refs 24, 25). The allowed EPR transitions satisfy the subsequent selection rules: $\Delta M_L = 0$, $\Delta M_S = \pm 1$, that is $\Delta M_J = \pm 1$ for the inter-doublet transitions 4 and 6, and $\Delta M_L = \mp 1$, $\Delta M_S = \pm 2$, that is $\Delta M_J = \pm 1$ for the weak intra-doublet transition 5 whose transition probability is caused by a small intermultiplet mixing through the in-plane component of the DM coupling. The intensity of this transition is significantly increased when transition 7 becomes allowed due to a weak deviation from the C_3 symmetry (Fig. 1). This also leads to an increased zero-field gap $\left[\Delta_{DM}^2 + \delta^2\right]^{1/2}$ where δ is the parameter in the exchange shift $\delta \mathbf{S}_1\mathbf{S}_2$.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 28 January; accepted 1 April 2008.

1. Gatteschi, D., Sessoli, R. & Villain, J. *Molecular Nanomagnets* (Oxford Univ. Press, Oxford, UK, 2006).
2. Barbara, B. *et al.* Mesoscopic quantum tunneling of the magnetization. *J. Magn. Magn. Mater.* **140–144,** 1825–1828 (1995).
3. Thomas, L. *et al.* Macroscopic quantum tunneling of magnetization in a single crystal of nanomagnets. *Nature* **383,** 145–147 (1996).
4. Friedman, J. R. *et al.* Macroscopic measurements of resonant magnetization tunneling in high spin molecules. *Phys. Rev. Lett.* **76,** 3830–3833 (1996).
5. Leuenberger, M. N. & Loss, D. Quantum computing in molecular magnets. *Nature* **410,** 789–793 (2001).
6. Aharonov, D., Kitaev, A. & Preskill, J. Fault-tolerant quantum computation with long-range correlated noise. *Phys. Rev. Lett.* **96,** 050504 (2006).
7. Stamp, P. C. E. & Tupitsyn, I. S. Coherence window in the dynamics of quantum nanomagnets. *Phys. Rev. B* **69,** 014401 (2004).
8. Morello, A., Stamp, P. C. E. & Tupitsyn, I. S. Pairwise decoherence in coupled spin qubit networks. *Phys. Rev. Lett.* **97,** 207206 (2006).
9. Prokof'ev, N. V. & Stamp, P. C. E. Theory of the spin bath. *Rep. Prog. Phys.* **63,** 669–726 (2000).
10. Rabi, I. I. Space quantization in a gyrating magnetic field. *Phys. Rev.* **51,** 652–654 (1937).
11. Müller, A. & Döring, J. A novel heterocluster with D_3-symmetry containing twenty-one core atoms: $[As^{III}_6V^{IV}_{15}O_{42}(H_2O)]^{6-}$. *Angew. Chem. Int. Edn Engl.* **27,** 1721 (1988).
12. Gatteschi, D., Pardi, L., Barra, A. L., Müller, A. & Döring, J. Layered magnetic structure of a metal cluster ion. *Nature* **354,** 463–465 (1991).
13. Barbara, B. On the richness of supra-molecular chemistry and its openings in physics. *J. Mol. Struct.* **656,** 135–140 (2003).
14. Wernsdorfer, W., Müller, A., Mailly, D. & Barbara, B. Resonant photon absorption in the low spin molecule V_{15}. *Europhys. Lett.* **66,** 861–867 (2004).
15. Ardavan, A. *et al.* Will spin-relaxation times in molecular magnets permit quantum information processing? *Phys. Rev. Lett.* **98,** 057201 (2007).
16. Wernsdorfer, W. A long-lasting phase. *Nature Mater.* **6,** 174–176 (2007).
17. Mehring, M., Mende, J. & Scherer, W. Entanglement between an electron and a nuclear spin ½. *Phys. Rev. Lett.* **90,** 153001 (2003).
18. Morton, J. J. L. *et al.* Bang–bang control of fullerene qubits using ultrafast phase gates. *Nature Phys.* **2,** 40–43 (2006).
19. Bertaina, S. *et al.* Rare earth solid state qubits. *Nature Nanotechnol.* **2,** 39–42 (2007).
20. Nellutla, S. *et al.* Coherent manipulation of electron spins up to ambient temperatures in Cr^{5+} (S=1/2) doped K_3NbO_8. *Phys. Rev. Lett.* **99,** 137601 (2007).
21. Chiorescu, I., Wernsdorfer, W., Müller, A., Bögge, H. & Barbara, B. Butterfly hysteresis loop and dissipative spin reversal in the $S = 1/2$, V_{15} molecular complex. *Phys. Rev. Lett.* **84,** 3454–3457 (2000).
22. De Raedt, H. D., Miyashita, S., Michielsen, K. & Machida, M. Dzyaloshinskii-Moriya interactions and adiabatic magnetization dynamics in molecular magnets. *Phys. Rev. B* **70,** 064401 (2004).
23. Chaboussant, G. *et al.* Mechanism of ground-state selection in the frustrated molecular spin cluster V_{15}. *Europhys. Lett.* **66,** 423–429 (2004).
24. Tarantul, A., Tsukerblat, B. & Müller, A. Static magnetization of V_{15} cluster at ultra-low temperatures: Precise estimation of antisymmetric exchange. *Inorg. Chem.* **46,** 161–169 (2007).
25. Tsukerblat, B., Tarantul, A. & Müller, A. Low temperature EPR spectra of the mesoscopic cluster V_{15}: The role of antisymmetric exchange. *J. Chem. Phys.* **125,** 054714 (2006).
26. Dzyaloshinsky, I. A thermodynamic theory of "weak" ferromagnetism of antiferromagnetics. *J. Phys. Chem. Solids* **4,** 241–255 (1958).
27. Moriya, T. Anisotropic superexchange interactions and weak ferromagnetism. *Phys. Rev.* **120,** 91–98 (1960).
28. Volkmer, D. *et al.* Towards nanodevices: Synthesis and characterization of the nanoporous surfactant-encapsulated keplerate $(DODA)_{40}(NH_4)_2[(H_2O)_n \subset Mo_{132}O_{372}(CH_3COO)_{30}(H_2O)_{72}]$. *J. Am. Chem. Soc.* **122,** 1995–1998 (2000).
29. Prokof'ev, N. V. & Stamp, P. C. E. Quantum relaxation of magnetisation in magnetic particles. *J. Low Temp. Phys.* **104,** 143–210 (1996).
30. Hartmann-Boutron, F., Politi, P. & Villain, J. Tunneling and magnetic relaxation in mesoscopic molecules. *Int. J. Mod. Phys. B* **10,** 2577–2637 (1996).

Acknowledgements We acknowledge I. Chiorescu from NHMFL-FSU, Tallahassee, USA, for discussions. We thank M.-N. Collomb for help in processing samples for EPR measurements, and G. Desfonds for technical support. B.B. and A.M. thank the European Research Council for support through network projects MAGMANet, MolNanoMag, QueMolNa and INTAS; A.M. thanks the Deutsche Forschungsgemeinschaft and the Fonds der Chemischen Industrie for support; and B.T. and A.M. thank the German–Israeli Foundation for Scientific Research and Development for support.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to B.B. (bernard.barbara@grenoble.cnrs.fr) or A.M. (a.mueller@uni-bielefeld.de).

METHODS

Sample synthesis. 0.04 g (0.0175 mmol) of freshly prepared brown $K_6\left[V^{IV}_{15}As^{III}_6O_{42}(H_2O)\right]\cdot 8H_2O$ obtained as reported[10] was dissolved in 20 ml of degassed water. After addition of 25 ml of a (degassed) trichloromethane solution of [DODA]Br (1.10 g/1.75 mmol) the reaction medium was stirred under inert atmosphere. The stirring was continued until the olive-brown coloured aqueous layer turned colourless and the corresponding colour appeared in the organic phase. The organic layer was then quickly separated, put into an EPR tube and frozen to liquid nitrogen temperature. All operations were done in an inert atmosphere.

Comparing Rabi frequencies. The frequency of the Rabi oscillations between two states 1 and 2 is given by[6–8,19]:

$$\Omega_R = \Omega_{R0}\,|\langle\varphi_1|S_+|\varphi_2\rangle| \tag{2}$$

Here $\Omega_{R0} = 2g\mu_B B_1/h_{Planck} = 55.96B_1$ (MHz, mT) is the Rabi frequency of a spin 1/2, B_1 is the amplitude of the a.c. microwave fields, $g \approx 2$ the Landé factor, S_+ the ladder operator and $|\varphi_1\rangle$, $|\varphi_2\rangle$ the wavefunctions associated with these states. The probability of a transition, defined as $P = |\langle\varphi_1|S_+|\varphi_2\rangle|^2$, is directly connected with its Rabi frequency:

$$P = (\Omega_R/\Omega_{R0})^2 \tag{3}$$

This allows one to evaluate the ratio (R) of the probabilities associated with two transitions (here the '3/2' and '1/2' types) from the measurement of their Rabi frequencies without the knowledge of their wavefunctions:

$$R = P_{3/2}/P_{1/2} = \left(\Omega_{R\,3/2}/\Omega_{R\,1/2}\right)^2 \tag{4}$$

Using the values of the Rabi frequencies given in Fig. 3, one gets $R \approx (4.5/18.5)^2 \approx 5.9 \times 10^{-2}$. The time $T_{\pi/2}$, during which the excitation pulse is applied to induce a $\pi/2$ rotation, is by definition equal to $1/4\Omega_R$ (refs 6, 19), showing that equation (4) is equivalent to:

$$R = P_{3/2}/P_{1/2} = \left(T_{\pi/2,1/2}/T_{\pi/2,3/2}\right)^2 \tag{5}$$

This gives another way to determine R. Using the $T_{\pi/2}$ values given in Fig. 4 legend, one gets $R \approx (16/64)^2 \approx 6.2 \times 10^{-2}$, which is very close to the first one and shows that the probability associated with the '1/2' type transition is much smaller than the one associated with '3/2'.

Rare-earth solid-state qubits

S. BERTAINA[1,2], S. GAMBARELLI[3], A. TKACHUK[4], I. N. KURKIN[5], B. MALKIN[5], A. STEPANOV[6] AND B. BARBARA[1]*

[1]Institut Néel, Département Nanosciences, CNRS, 25 Ave. des Martyrs, BP166, 38042 Grenoble Cedex 9, France
[2]Ecole Nationale Supérieure de Physique, INP de Grenoble, Minatec 3 parvis Louis Néel, BP 257, 38016 Grenoble Cedex 9, France
[3]Laboratoire de Chimie Inorganique et Biologie (UMR-E 3 CEA-UJF), DRFMC, CEA-Grenoble, 17 rue des Martyrs 38054 Grenoble Cedex 9, France
[4]S.I. Vavilov State Optical Institute, St Petersburg 199034, Russian Federation
[5]Kazan State University, Kazan 420008, Russian Federation
[6]Laboratoire de Matériaux et Microélectronique de Provence, Faculté St Jérôme, C142, 13397, Marseille Cedex 20, France
*e-mail: bernard.barbara@grenoble.cnrs.fr

Published online: 3 January 2007; doi:10.1038/nnano.2006.174

Quantum bits (qubits) are the basic building blocks of any quantum computer. Superconducting qubits have been created with a top-down approach that integrates superconducting devices into macroscopic electrical circuits[1–3], and electron-spin qubits have been demonstrated in quantum dots[4–6]. The phase coherence time (τ_2) and the single qubit figure of merit (Q_M) of superconducting and electron-spin qubits are similar — at $\tau_2 \sim \mu$s and $Q_M \sim 10$–1,000 below 100 mK — and it should be possible to scale up these systems, which is essential for the development of any useful quantum computer. Bottom-up approaches based on dilute ensembles of spins have achieved much larger values of τ_2 (up to tens of milliseconds; refs 7,8), but these systems cannot be scaled up, although some proposals for qubits based on two-dimensional nanostructures should be scalable[9–11]. Here we report that a new family of spin qubits based on rare-earth ions demonstrates values of τ_2 (∼50 μs) and Q_M (∼1,400) at 2.5 K, which suggests that rare-earth qubits may, in principle, be suitable for scalable quantum information processing at ^{4}He temperatures.

In general, a spin qubit state is a linear superposition of the two spin states of an electron $|\uparrow>$ and $|\downarrow>$. This means that the qubit can be represented as $|\psi_s> = \alpha|\uparrow> + \beta|\downarrow>$, where α and β are probability amplitudes, and $|\alpha|^2 + |\beta|^2 = 1$. When measuring this qubit, the probability of outcome $|\uparrow>$ (or $|\downarrow>$) is $|\alpha|^2$ (or $|\beta|^2$). In rare earth (RE) systems, the total spin, **S**, is no longer a good quantum number, because the spin-orbit coupling between **S** and the total orbital angular momentum, **L**, is larger than the coupling of **L** with the electric field gradient of environmental ionic charges (crystal field). The good quantum number is the total angular momentum, $\mathbf{J} = \mathbf{L} + \mathbf{S}$, which is coupled with the crystal field through **L**. The RE qubit states are therefore crystal-field states. In addition, RE elements often have isotopes with a nuclear spin, **I**, that has large hyperfine interactions with **J**, leading to electro-nuclear crystal-field states with wavefunctions $|\Psi_{en}>$ (see Methods).

Qubits based on these electro-nuclear states differ from typical spin qubits in several ways: (1) the crystal field strongly affects the Rabi frequencies that depend on the direction and the strength of applied magnetic fields and electric field gradients, and this could open up new possibilities for scaling; (2) the hyperfine interactions produce up to $3(2I + 1) - 2$ qubits per RE, all with slightly different resonance frequencies, which means that it should be quite easy to selectively address them with superimposed (low) field pulses; (3) owing to their large magnetic moment (∼10 μ_B), it should be simple to manipulate RE qubits; and (4) the single qubit figure of merit, Q_M, should be large enough to allow quantum information processing at ^{4}He temperatures (Q_M is the number of coherent single-qubit operations, defined as $\Omega_R\tau_2/\pi$, where Ω_R is the Rabi frequency; equivalently, it is the coherence time divided by half the Rabi period).

This work is an extension of previous research that explored the quantum tunnelling of the magnetization in Mn_{12}-ac and Ho:$YLiF_4$ (refs 12–14). Owing to the strong hyperfine interactions in the latter system, **J** tunnels simultaneously with **I** (electro-nuclear tunnelling). The system chosen to illustrate the concept of RE qubits consists of Er^{3+} ions ($J = 15/2$ and $g_J = 6/5$) diluted in a single crystalline matrix of $CaWO_4$, which is isomorphic with $YLiF_4$. The main reason for replacing $YLiF_4$ with $CaWO_4$ is to reduce the proportion of nuclear spins, which are an important source of decoherence[15] (the phenomenon by which a quantum system seems to be classical as a result of interactions with its environment).

Continuous-wave electron paramagnetic resonance (CW-EPR) measurements were first performed in Er^{3+}:$CaWO_4$. The transitions for the isotopes with $I = 0$ and $I = 7/2$ were observed at ^{4}He temperatures using a Bruker X-band spectrometer at 9.7 GHz. These transitions occur either between pure crystal-field levels ($I = 0$) or between electro-nuclear crystal-field sublevels ($I = 7/2$) (see Methods). In both cases, the observed line-width is small enough for the lifetime of the levels to be much larger than calculated periods of Rabi oscillations (weak decoherence). In order to observe these oscillations, a series of experiments was performed in pulsed-wave EPR (PW-EPR) mode. Eight transitions were observed (Fig. 1).

An example of the measured Rabi oscillations[16] is given in Fig. 2, for $I = 0$, where the z component of the magnetization, M_z, is plotted against time. It is possible to fit the data to

$$\langle M_z \rangle = M_{z(t=0)} e^{-t/\tau_R} \sin(\Omega_R t) \qquad (1)$$

using a single exponential damping parameter $\tau_R \sim 0.2$ μs (Ω_R having been previously obtained from a Fourier transform of the

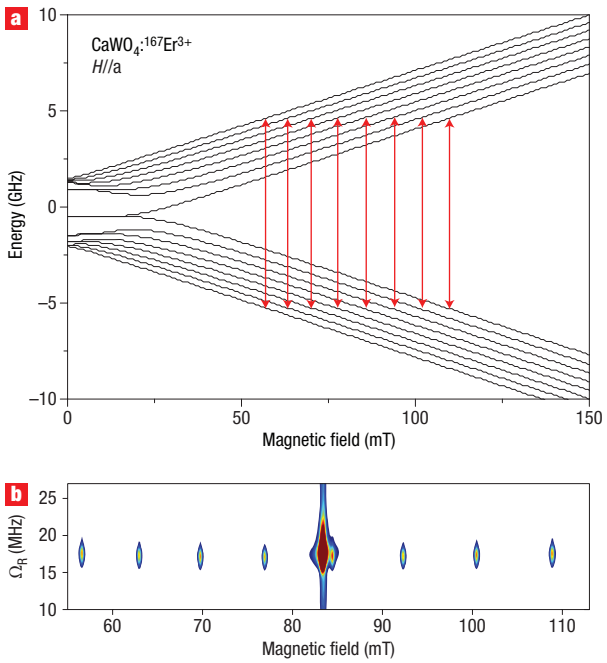


Figure 1 Energy levels and Rabi frequencies for the erbium-doped RE system $^{167}Er^{3+}$:$CaWO_4$. a, Energy spectrum calculated for a magnetic field perpendicular to the c-axis. In the zero field, the spectrum contains 16 electro-nuclear states ($(2S+1)(2I+1)$ with $S=1/2$ and $I=7/2$) consisting of a singlet, 7 doublets and another singlet (nine sublevels). The fourth doublet, near the centre of the figure, is well separated from the other levels. When the Zeeman splitting caused by the magnetic field becomes larger than the hyperfine splitting, which sets the energy scale at zero field, the levels vary linearly with the magnetic field, which gives 8 states with effective spin $1/2$ and 8 states with effective spin $-1/2$. Each of these states is labelled by the nuclear spin projection, m_I, which increases from $-7/2$ for the two states at the centre of the figure to $+7/2$ for the lower- and upper-most states. EPR transitions between spins $\pm 1/2$ and $\Delta m_I = 0$ are represented by the vertical arrows. **b**, Rabi frequencies, measured versus static field $H//a$ and ac-field $\mu_0 h = 0.12$ mT//b, on a single crystal of Er^{3+}:$CaWO_4$ ($2 \times 2.5 \times 3$ mm^3, 10^{-5} atomic % Er). They show an intense central peak (for the isotopes $I=0$) and 8 smaller peaks separated by $\Delta H \approx 6$–8 mT (for the isotope $I=7/2$, $^{167}Er^{3+}$). Exact diagonalization of equation (4) (see Methods) permits accurate calculation of these frequencies (using the crystal-field and hyperfine constants only[30,32]); one finds $\Omega_R/2\pi = 17.546$, 17.302, 17.166, 17.115, 17.137, 17.238, 17.394 and 17.605 MHz. The colour scale shows the proportion of ions with Rabi frequency Ω_R at a given magnetic field (white <80, blue $=80$, red >800 arbitrary units).

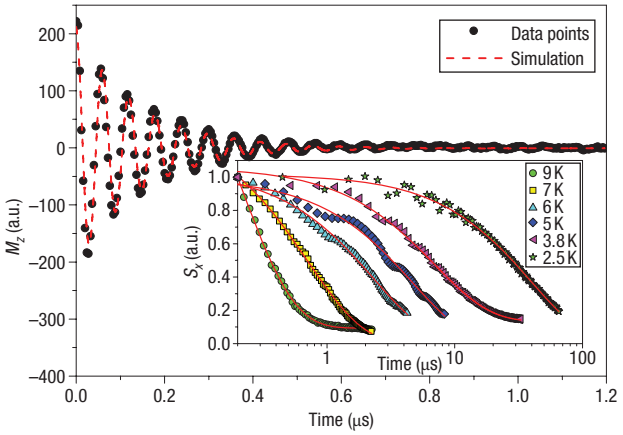


Figure 2 Rabi oscillations and coherence times. Rabi oscillations measured on $I=0$ isotopes of the same single crystal for $\mu_0 H = 0.522$ T//c, $\mu_0 h = 0.15$ mT//b and $T = 3.5$ K. These oscillations are obtained by the application of a nutation pulse of length t followed, after a delay greater than τ_2 (permitting the transverses spin components to relax), by a $\pi/2 - \pi$ sequence. The resulting echo intensity is averaged over $\sim 10^3$ measurements, giving the z-component of the nutating magnetization at time t (M_z). The dashed line is a fit to equation (1) (see text) giving an exponential decay time $\tau_R = 0.2\ \mu s \ll \tau_2 \approx 7\ \mu s$ (see Fig. 3). The inset shows the decay of the transverse spin component, S_x, obtained by a conventional spin-echo method at different temperatures, showing that the coherence time τ_2 reaches the 100 μs scale at ^{4}He temperatures. Weak superimposed oscillations come from the ESEEM effect (Electron Spin Echo Envelope Modulation)[34] produced by the super-hyperfine coupling with second neighbour W nuclear spins. One can verify that the oscillation frequency perfectly matches the W nucleus spin Larmor frequency in the applied field (small super hyperfine limit).

data). Other experiments performed at different microwave powers show that τ_R increases as the power decreases, and the number of Rabi oscillations, $N(c)$ (where c is the concentration of Er), remains nearly unchanged; that is, $N(c) \approx \tau_R(c)\Omega_R$, with $N(c) \sim 20$ in the example of Fig. 2. This increase of τ_R is always limited by τ_2 (Fig. 3). All of this suggests the phenomenological expression

$$1/\tau_R(c) \approx \Omega_R/N(c) + 1/\tau_2(c) \qquad (2)$$

where $\tau_R(c)$, $N(c)$ and $\tau_2(c)$ are concentration-dependent. Rabi oscillations are lost for $t \gg \tau_2$ in the low power limit where $\Omega_R \to 0$, and for $t \gg N(c)/\Omega_R$ in the large power limit where $\Omega_R \gg N(c)/\tau_2$. In the first case, τ_2 should be limited by RE spin-diffusion because of long-range dipolar interactions, as in nuclear magnetic resonance. In the second case, the observed behaviour is characteristic of inhomogeneous nutation frequency. In fact, a weak random crystal field, responsible for the CW line-width[17,18], feeds into some distribution of the $|J,m_J,I,m_I>$ coefficients, resulting in destructive interference of Rabi oscillations ($\Omega_R \propto \langle \phi_1, m_I | J_+ | \phi_2, m_I' \rangle$, see Methods), which go out of phase after a certain number of periods. However, the number of oscillations $N(c)$ depends on concentration, indicating that dipolar interactions must also be taken into account.

Recently, a model relying on the assumption that each spin experiences a stochastic field of mean-square amplitude β, oscillating at the resonance frequency ω, led to the expression

$$1/\tau_R = \beta\Omega_R + 1/2\tau_2 \qquad (3)$$

very similar to equation (2) (ref. 19). This linear dependence on Ω_R was tested on pure $S=1/2$ spins in amorphous-SiO_2 containing E′ centres where a concentration effect has also been obtained[20]. In the frame of the present study, the origin of the stochastic field should be related to both crystal-field distribution and dipolar interactions[21]. In order to check equations (2) and (3) more carefully, $1/\tau_R$ versus Ω_R is plotted for two different directions of the microwave field h (Fig. 3, inset; see also Fig. 4). The obtained curve is continuous, showing that the damping rate scales with the Rabi frequency (and not with the microwave field h when the dipole matrix elements are different) according to an S-shaped curve of, for example, the type $1/\tau_R = 1/\tau_2 f(\Omega_R \tau_2)$, with a progressive saturation at τ_2 when $\Omega_R \to 0$. The dependence

of Ω_R with the direction of the microwave field, h, is demonstrated in Fig. 4, where a simple rotation from $h//b$ to $h//c$ reduces the Rabi frequency by the factor $\Omega_{R//b}/\Omega_{R//c} \sim 6$. This ratio is slightly smaller than the one derived from the proportionality $\Omega_{R//b}/\Omega_{R//c} \approx g_{effb}/g_{effc} \approx 6.7$, because $\Omega_{R//b}$ drops in a few degrees from its maximum value $\Omega_{R//b} \sim 6.7\Omega_{R//c}$ to its minimum value $\sim\Omega_{R//c}$. A better agreement would simply require better angular accuracy in the crystal orientation.

Finally, RE qubits have large Q_M at ^{4}He temperatures and, in principle, they should be scalable. Indeed τ_2 increases with dilution and cooling (Fig. 2b); an extrapolation down to 1.5 K for a concentration of 10^{-6} atomic Er:$CaWO_4$ gives $Q_M \sim 10^4$, which is enough for quantum information processing. Moreover, RE qubits could in principle be selectively addressed and their couplings manipulated, according to variants of existing proposals and realizations[4–6,10,11]. As a matter of fact, they could be inserted in all kinds of matrices structured by lithography, including films, quantum dots or nanowires of semiconducting Si (ref. 22) or GaN (ref. 23), and coupled by controlled carrier injection through the gate voltage[24]. They could be addressed selectively by application of (1) local field pulses of amplitude <25 mT adding algebraically to the static field (this is limited to $n \leq 3(2I+1)-2$ qubits; Fig. 1), and (2) continuous electric field gradients for $n > 3(2I+1)-2$. A gradient of 10 mV $(nm)^{-2}$ is enough to modify the crystal-field parameters by $\sim$10% in most matrices and therefore the resonance frequency. Interestingly, the $3(2I+1)-2$ Rabi oscillations of each ^{167}Er (Fig. 1) may also be used to implement Grover's algorithm[25] on single RE ions (this is a general property of electro-nuclear RE qubits with $I \neq 0$). Spin-state detection could follow schemes like those in refs 4 and 6, but alternative ways using the fast photoluminescent properties of RE (refs 22,23,26) might ultimately be better. Finally, instead of dots one might also use single molecules containing a RE ion[27].

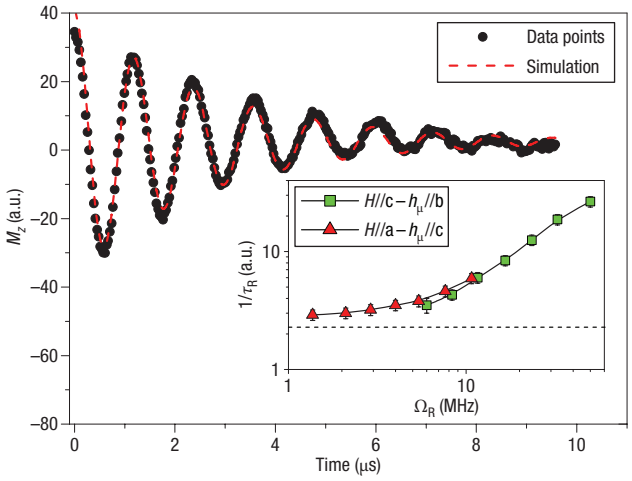


Figure 3 Changing the damping time with the microwave power. When the experiment in Fig. 2 is repeated with the microwave field reduced by a factor of 20, the period of the Rabi oscillations becomes longer (by the same factor of 20), but the number of periods remains of the order of 20 (up to 20 μs). The same fit as in Fig. 2 gives $\tau_R \sim 3$ μs, which is comparable with the $\tau_2 \sim 7$ μs obtained in spin-echo measurements under the same experimental conditions. The inset shows the damping rate of the Rabi oscillations, $1/\tau_R$, plotted against the Rabi frequency, Ω_R, for two directions of the microwave field. The continuity of the curve proves that $1/\tau_R$ depends on Ω_R only and tends to $1/\tau_2$ at low microwave power (dashed line).

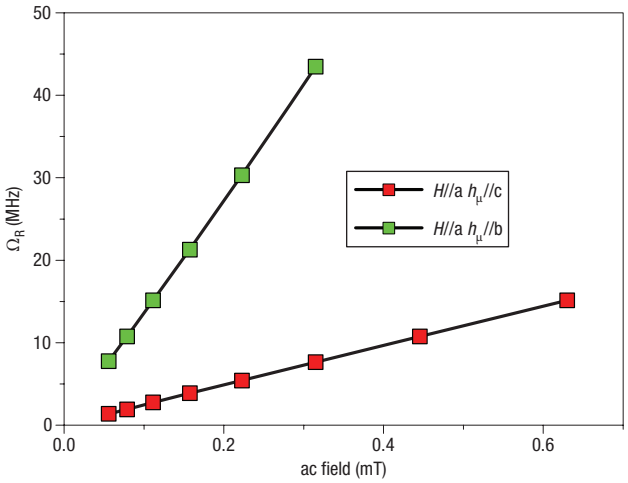


Figure 4 Maximum and minimum coupling of the microwave field to Er effective spins and direction-dependent Rabi frequencies. Rabi frequency, Ω_R, measured for two directions of the microwave field, h_μ, on a single crystal with atomic Er concentration of 5×10^{-4}. The ac field was calibrated by comparison with a coal sample. Owing to the 'easy' plane anisotropy (see Methods), the coupling between Er effective spins and the microwave field is maximum when the latter is in the easy plane (giving large Ω_R) and minimum when it is perpendicular to it (giving small Ω_R).

In conclusion, Rabi oscillations of the angular moment $J = 15/2$ of Er:$CaWO_4$ have been observed for the first time and analysed, evincing a new type of anisotropic electro-nuclear spin qubits. Isotopes with $I = 0$ give a single purely electronic Rabi frequency (single qubit, $\Delta M_J = \pm 1$), and the isotope $I = 7/2$ (^{167}Er) gives a set of eight electro-nuclear frequencies (eight qubits, $\Delta M_J = \pm 1$ and $\Delta M_I = 0$), which are addressed independently. Because the spin-orbit coupling, the magnetic moments and the hyperfine interactions are all large, it should be possible to couple and address selectively a large number of RE qubits using weak electric and magnetic fields. Furthermore, each RE ion could be used to implement Grover's algorithm. All this, together with large Q_M factors ($\sim 10^3$–10^4 between 2.5 and 1.5 K), suggests that RE qubits are good candidates for implementation of quantum computation at ^{4}He temperatures.

METHODS

CRYSTAL-FIELD BACKGROUND

The hamiltonian The single-ion hamiltonian for Er^{3+}:$CaWO_4$ (tetragonal space group $I4_1/a$ and S_4 point symmetry[28]) contains crystal-field, hyperfine and Zeeman terms:

$$H_{CF} = \alpha_J B_2^0 O_2^0 + \beta_J (B_4^0 O_4^0 + B_4^4 O_4^4) + \gamma_J (B_6^0 O_6^0 + B_6^4 O_6^4 + B_6^{-4} O_6^{-4}) + A_J \mathbf{I}\cdot\mathbf{J} + g_J \mu_B \mu_0 \mathbf{JH} \quad (4)$$

The O_l^m are the Stevens' equivalent operators with the reduced matrix elements α_J, β_J, γ_J (ref. 29), and the B_l^m are the crystal-field parameters determined by high-resolution optical spectroscopy ($B_2^0 = 231\ cm^{-1}$, $B_4^0 = -90\ cm^{-1}$, $B_4^4 = \pm 852\ cm^{-1}$, $B_6^0 = -0.6\ cm^{-1}$, $B_6^4 = \pm 396\ cm^{-1}$ and $B_6^{-4} = \pm 75\ cm^{-1}$; ref. 30).

Energy spectra and wavefunctions Exact diagonalization of the 16×16 matrix of equation (4) with $\mathbf{I} = \mathbf{H} = 0$ reveals an easy plane perpendicular to the c-axis with a doublet ground state of wavefunctions $|\phi_1>$ and $|\phi_2>$. This doublet, with effective spin 1/2 and anisotropic g_{eff} tensor ($g_{//} = 1.247$, $g_\perp = 8.38$; ref. 31), permits a single EPR transition ($\Delta m_J = \pm 1$), which can be

observed on I = 0 isotopes (~77%). The Rabi frequency is given by $\Omega_R = 2g_J\mu_B\langle\phi_1|J_\mu h_\mu|\phi_2\rangle/h \propto g_{eff}$, where h is Planck's constant. Natural Er also contains ^{167}Er with I = 7/2 (~23%) and $A_J = -4.16 \times 10^{-3}$ cm^{-1} (−125 MHz) (ref. 32). In this case the 128 × 128 matrix leads to the energy spectrum of Fig. 1a. The degeneracy is completely removed by **H** and the new set of wavefunctions $|\Psi_{en}\rangle = \Sigma b_i|J,m_J,I,m_I\rangle$ on the space product $|L,S,J,m_J\rangle \otimes |I,m_I\rangle$ differs from $|\phi_1\rangle$ and $|\phi_2\rangle$ owing to the nuclear degrees of freedom. Figure 1 also shows that $3(2I + 1) - 2$ EPR transitions are allowed, giving, for I = 7/2, eight transitions with conservation of I ($\Delta m_J = \pm 1$ and $\Delta m_I = 0$) and 14 transitions without ($\Delta I_J = \pm 1$ and $\Delta m_I = \pm 1$).

Received 23 October 2006; accepted 27 November 2006; published 3 January 2007.

References

1. Leggett, A. J. Superconducting qubits — a major roadblock dissolved? *Science* **296,** 861–862 (2002).
2. Chiorescu, I. *et al.* Coherent dynamics of a flux qubit coupled to a harmonic oscillator. *Nature* **431,** 159–163 (2004).
3. Pashkin, A., Astafiev, O., Nakamura, Y. & Tsai, J. S. Demonstration of conditional gate operation using superconducting charge qubits. *Nature* **425,** 941–944 (2003).
4. Koppens, F. L. H. *et al.* Driven coherent oscillations at a single electron spin in a quantum dot. *Nature* **442,** 766–771 (2006).
5. Oosterkamp, T. H. *et al.* Microwave spectroscopy of a quantum-dot molecule. *Nature* **395,** 873–876 (1998).
6. Petta, J. R. *et al.* Coherent manipulation of coupled electron spins in semiconductor quantum dots. *Science* **309,** 2180–2184 (2005).
7. Tyryshkin, A. M., Lyon, S. A., Astashkin, A. V. & Raitsimring, A. M. Electron spin relaxation times in phosphorous donors in silicon. *Phys. Rev. B* **68,** 193207 (2003).
8. Mehring, M., Scherer, W. & Weidinger, A. Pseudo-entanglement of spin states in the multilevel $^{15}N@C_{60}$ system. *Phys. Rev. Lett.* **93,** 206603 (2004).
9. Vandersypen, L. M. K. *et al.* Experimental realization of Shor's quantum factoring algorithm using nuclear magnetic resonance. *Nature* **414,** 883–887 (2001).
10. Loss, D. & DiVincenzo, D. P. Quantum computation with quantum dots. *Phys. Rev. A* **57,** 120–126 (1998).
11. Kane, B. E. A silicon based nuclear spin quantum computer. *Nature* **393,** 133–137 (1998).
12. Thomas, L. *et al.* Macroscopic quantum tunneling of the magnetization in a single crystal of nanomagnets. *Nature* **383,** 145–148 (1996)
13. Giraud, R., Wernsdorfer, W., Tkatchuk, A., Mailly, D. & Barbara, B. Nuclear spin driven quantum relaxation in $LiY_{0.998}Ho_{0.002}F_4$. *Phys. Rev. Lett.* **87,** 057203 (2001).
14. Giraud, R., Tkachuk, A. M. & Barbara, B. Quantum dynamics of atomic magnets: co-tunneling and dipolar-biased-tunneling. *Phys. Rev. Lett.* **91,** 257204 (2003).
15. Stamp, P. C. E. & Tupitsyn, I. S. Coherence window in the dynamics of quantum nanomagnets. *Phys. Rev. B* **69,** 014401 (2004).
16. Rabi, I. I. Space quantization in a gyrating magnetic field. *Phys. Rev.* **51,** 652–655 (1937).
17. Shakurov, G. S. *et al.* Direct measurement of anti-crossings of the electron-nuclear energy levels in $LiYF_4$: Ho with submillimeter EPR spectroscopy. *Appl. Magn. Reson.* **28,** 251–265 (2005).
18. Kurkin, I. N. & Shekun, L. Ya. Paramagnetic resonance linewidths for impurity ions in scheelite single crystals. *Fiz. Tverd. Tela* (*Leningrad*) **9,** 444–448 (1967).
19. Shakhmuratov, R. N., Gelardi, F. M. & Cannas, M. Non-Bloch transients in solids: free induction decay and transient nutations. *Phys. Rev. Lett.* **79,** 2963–2966 (1997).
20. Agnello, S., Boscaino, R., Cannas, M., Gelardi, F. M. & Shakhmuratov, R. N. Transient nutation decay: the effect of field-modified dipolar interactions. *Phys. Rev. A* **59,** 4087–4090 (1999).
21. Prokof'ev, N. V. & Stamp, P. C. E. Theory of the spin-bath. *Rep. Prog. Phys.* **63,** 669–726 (2000).
22. Gallis, S. *et al.* Photoluminescence in erbium doped amorphous silicon oxycarbide thin films. *Appl. Phys. Lett.* **87,** 091901 (2005).
23. Hori, Y. *et al.* GaN quantum dots doped with Tb. *Appl. Phys. Lett.* **88,** 53102 (2006).
24. Ohno, H. *et al.* Electric-field control of ferromagnetism. *Nature* **408,** 944–946 (2000).
25. Leuenberger, M. N. & Loss, D. Grover algorithm for large nuclear spins in semiconductors. *Phys. Rev. B* **68,** 165317 (2003).
26. Tkachuk, A. M., Razumova, I. K., Malyshev, A. V. & Gapontsev, V. P. Population of lasing erbium in YLT : Er^{3+} crystals under upconversion cw LD pumping. *J. Luminescence* **94–95,** 317–320 (2001).
27. Ishikawa, N., Sugita, M., Ishikawa, T., Koshihara, S. & Kaisu, Y. Lanthanide double-decker complexes functioning as magnets at single-molecular level. *J. Am. Chem. Soc.* **125,** 8694–8695 (2003).
28. Zhang, Y., Holzwarth, N. A. W. & Williams, R. T. Electronic band structures of the scheelite materials $CaMoO_4$, $CaWO_4$, $PbMoO_4$, and $PbWO_4$. *Phys. Rev. B* **57,** 12738 (1998).
29. Stevens, K. W. H. The theory of paramagnetic relaxation. *Proc. Phys. Soc. London* **A65,** 209–217 (1952).
30. Bernal, E. G. Optical spectrum and magnetic properties of Er^{3+} in $CaWO_4$. *J. Chem. Phys.* **55,** 2538–2549 (1971).
31. Antipin, A. A. *et al.* Paramagnetic resonance and spin-lattice relaxation of Er^{3+} and Tb^{3+} ions in $CaWO_4$ crystal lattice. *Sov. Phys. Solid State* **10,** 468–474 (1968).
32. Abragam, A. & Bleaney, B. *Electron Paramagnetic Resonance of Transition Ions* (Clarendon Press, Oxford, 1970).
33. Rowan, L. G., Hahn, E. L. & Mims, W. B. Electronic spin-echo envelope modulation. *Phys. Rev. A* **137,** A61–A71 (1969).

Acknowledgements

The authors acknowledge the support of INTAS contract no. 2003/03-51-4943. B.M. and I.K. acknowledge the Ministry of Education and Science of the Russian Federation (project RNP 2.1.1.7348) and B.B. the interdisciplinary European Network of Excellence 'MAGMANet' for support during the first year of the research.

Author contributions

A.T. provided the samples. S.B. and S.G. performed the experiments, and analysed and discussed them with A.S., I.K., B.M. and B.B. B.B. proposed this study and wrote the manuscript, which was commented on by all the authors.

Competing financial interests

The authors declare that they have no competing financial interests.

pole, then α would be the co-latitude of the struts and the longitudinal angle between the struts would be 120°.
16. The presence of the factor of c^2 confirms that this is a relativistic effect. There is no swimming effect in the analogous Newtonian problem.
17. The fact that the swimming displacement per stroke is so small means that, strictly speaking, one should consider the swimming effect relative to the ordinary nonswimming geodesic motion of the swimmer. However, the calculation that is presented is enough to show the existence of the effect that is surely also present in more complicated situations. Perhaps the most interesting case to consider would be a swimmer in a circular orbit, where the swimming effect could be used to gradually increase the radius of the orbit.
18. I thank J. Touma for infecting me with his interest in geometric phase and for bringing the articles of A. Shapere and F. Wilczek to my attention. I thank H. Abelson, E. Bertschinger, D. Finkelstein, R. Hermann, P. Kumar, G. J. Sussman, J. Touma, and F. Wilczek for helpful and pleasant conversations.

11 December 2002; accepted 12 February 2003
Published online 27 February 2003;
10.1126/science.1081406
Include this information when citing this paper.

REPORTS

Coherent Quantum Dynamics of a Superconducting Flux Qubit

I. Chiorescu,[1]* Y. Nakamura,[1,2] C. J. P. M. Harmans,[1] J. E. Mooij[1]

We have observed coherent time evolution between two quantum states of a superconducting flux qubit comprising three Josephson junctions in a loop. The superposition of the two states carrying opposite macroscopic persistent currents is manipulated by resonant microwave pulses. Readout by means of switching-event measurement with an attached superconducting quantum interference device revealed quantum-state oscillations with high fidelity. Under strong microwave driving, it was possible to induce hundreds of coherent oscillations. Pulsed operations on this first sample yielded a relaxation time of 900 nanoseconds and a free-induction dephasing time of 20 nanoseconds. These results are promising for future solid-state quantum computing.

It is becoming clear that artificially fabricated solid-state devices of macroscopic size may, under certain conditions, behave as single quantum particles. We report on the controlled time-dependent quantum dynamics between two states of a micron-size superconducting ring containing billions of Cooper pairs (*1*). From a ground state in which all the Cooper pairs circulate in one direction, application of resonant microwave pulses can excite the system to a state where all pairs move oppositely, and make it oscillate coherently between these two states. Moreover, multiple pulses can be used to create quantum operation sequences. This is of strong fundamental interest because it allows experimental studies on decoherence mechanisms of the quantum behavior of a macroscopic-sized object. In addition, it is of great importance in the context of quantum computing (*2*) because these fabricated structures are attractive for a design that can be scaled up to large numbers of quantum bits or qubits (*3*).

Superconducting circuits with mesoscopic Josephson junctions are expected to behave according to the laws of quantum mechanics if they are separated sufficiently from external degrees of freedom, thereby reducing the decoherence. Quantum oscillations of a superconducting two-level system have been observed in the Cooper pair box qubit using the charge degree of freedom (*4*). An improved version of the Cooper pair box qubit showed that quantum oscillations with a high quality factor could be achieved (*5*). In addition, a qubit based on the phase degree of freedom in a Josephson junction was presented, consisting of a single, relatively large Josephson junction current-biased close to its critical current (*6, 7*).

Our flux qubit consists of three Josephson junctions arranged in a superconducting loop threaded by an externally applied magnetic flux near half a superconducting flux quantum $\Phi_0 = h/2e$ [(*8*); a one-junction flux qubit is described in (*9*)]. Varying the flux bias controls the energy level separation of this effectively two-level system. At half a flux quantum, the two lowest states are symmetric and antisymmetric superpositions of two classical states with clockwise and anticlockwise circulating currents. As shown by previous microwave spectroscopy studies, the qubit can be engineered such that the two lowest eigenstates are energetically well separated from the higher ones (*10*). Because the qubit is primarily biased by magnetic flux, it is relatively insensitive to the charge noise that is abundantly present in circuits of this kind.

The central part of the circuit, fabricated by electron beam lithography and shadow evaporation of Al, shows the three in-line Josephson junctions together with the small loop defining the qubit in which the persistent current can flow in two directions, as shown by arrows (Fig. 1A). The area of the middle junction of the qubit is $\alpha = 0.8$ times the area of the two outer ones. This ratio, together with the charging energy $E_C = e^2/2C$ and the Josephson energy $E_J = hI_C/4\pi e$ of the outer junctions (where I_C and C are their critical current and capacitance, respectively), determines the qubit energy levels (Fig. 2A) as a function of the superconductor phase γ_q across the junctions (Fig. 1B). Close to $\gamma_q =$

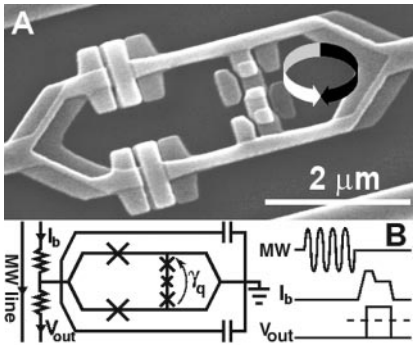


Fig. 1. (**A**) Scanning electron micrograph of a flux qubit (small loop with three Josephson junctions of critical current ~0.5 μA) and the attached SQUID (large loop with two big Josephson junctions of critical current ~2.2 μA). Evaporating Al from two different angles with an oxidation process between them gives the small overlapping regions (the Josephson junctions). The middle junction of the qubit is 0.8 times the area of the other two, and the ratio of qubit/SQUID areas is about 1:3. Arrows indicate the two directions of the persistent current in the qubit. The mutual qubit/SQUID inductance is $M \approx 9$ pH. (**B**) Schematic of the on-chip circuit; crosses represent the Josephson junctions. The SQUID is shunted by two capacitors (~5 pF each) to reduce the SQUID plasma frequency and biased through a resistor (~150 ohms) to avoid parasitic resonances in the leads. Symmetry of the circuit is introduced to suppress excitation of the SQUID from the qubit-control pulses. The MW line provides microwave current bursts inducing oscillating magnetic fields in the qubit loop. The current line provides the measuring pulse I_b and the voltage line allows the readout of the switching pulse V_{out}. The V_{out} signal is amplified, and a threshold discriminator (dashed line) detects the switching event at room temperature.

[1]Quantum Transport Group, Department of NanoScience, Delft University of Technology and Delft Institute for Micro Electronics and Submicron Technology (DIMES), Lorentzweg 1, 2628 CJ Delft, Netherlands. [2]NEC Fundamental Research Laboratories, 34 Miyukigaoka, Tsukuba, Ibaraki 305-8501, Japan.

*To whom correspondence should be addressed. E-mail: chiorescu@qt.tn.tudelft.nl

π, the loop behaves as a two-level system with an energy separation $E_{10} = E_1 - E_0$ of the eigenstates $|0\rangle$ and $|1\rangle$ described by the effective Hamiltonian $H = -\varepsilon\sigma_z/2 - \Delta\sigma_x/2$, where $\sigma_{z,x}$ are the Pauli spin matrices, Δ is the level repulsion, and $\varepsilon \approx I_p\Phi_0(\gamma_q - \pi)/\pi$ (where $I_p \approx 2\pi\alpha E_J/\Phi_0$ is the qubit maximum persistent current) (*11*).

The sample is enclosed in a gold-plated copper shielding box kept at cryogenic temperatures $T = 25$ mK ($k_BT \ll \Delta$). The qubit is initialized to the ground state simply by allowing it to relax. Coherent control of the qubit state is achieved by applying resonant microwave excitations on the microwave (MW) line (Fig. 1B), thereby inducing an oscillating magnetic field through the qubit loop. The qubit state evolves driven by a time-dependent term $(-\frac{1}{2})\varepsilon_{mw}\cos(2\pi Ft)\sigma_z$ in the Hamiltonian where F is the microwave frequency and ε_{mw} is the energy-modulation amplitude proportional to the microwave amplitude. This dynamic evolution is similar to that of spins in magnetic resonance. When the MW frequency equals the energy difference of the qubit, the qubit oscillates between the ground state and the excited state. This phenomenon is known as Rabi oscillation. The Rabi frequency depends linearly on the MW amplitude (*12*–*14*).

Readout is performed with an underdamped superconducting quantum interference device (SQUID) with a hysteretic current-voltage characteristic in direct contact with the qubit loop (Fig. 1A). The mutual coupling M is relatively large because of the shared kinetic and geometric inductances of the joint part enhancing the qubit signal. After performing the qubit operation, a bias current pulse I_b is applied to the SQUID (*15*). The I_b pulse consists of a short current pulse of length ~50 ns followed by a trailing plateau of ~500 ns (Fig. 1B). During the current pulse, the SQUID either switches to the gap voltage or stays at zero voltage. The pulse height and length are set to optimize the distinction of the switching probability between the two qubit states, which couple to the SQUID through the associated circulating currents. Because the readout electronics has a limited bandwidth of ~100 kHz, a voltage pulse of 50 ns is too short to be detected. For that reason the trailing plateau is added, with a current just above the retrapping current of the SQUID. The whole shape is adjusted for maximum readout fidelity. The switching probability is obtained by repeating the whole sequence of reequilibration, microwave control pulses, and readout typically 5000 times.

When the SQUID bias current is switched on, the circulating current in the SQUID changes. This circulating current, coupled to the qubit through the mutual inductance, changes the phase bias of the qubit by an estimated amount 0.01π. Consequently, the phase bias at which the quantum operations are performed is different from the phase bias at readout. This can be very useful because at the phase bias near π, where the qubit is least sensitive to flux noise, the expectation values for the qubit circulating current are extremely small. The automatic phase bias shift can be used to operate near π and to perform readout at a bias with a good qubit signal (*11*). Care must be taken that the fast shift remains adiabatic and that the whole sequence is completed within the relaxation time.

The average SQUID switching current I_{sw} versus applied flux shows the change of the qubit ground-state circulating current (Fig. 2B). Here, the I_b pulse amplitude is adjusted such that the averaged switching probability is maintained at 50%. A step corresponding to the change of qubit circulating current was observed (around the dashed line). The relative variation of 2.5% of I_{sw} is in agreement with the estimation based on the qubit current I_p and the qubit-SQUID mutual inductance M.

The relevant two energy levels of the qubit were first examined by spectroscopic means. Before each readout, a long microwave pulse (1 μs) at a series of frequencies was applied to observe resonant absorption peaks/dips each time the qubit energy separation E_{10}—adjusted by changing the external flux—coincides with the MW frequency F (*10*). The dots in Fig. 2C are measured peak/dip positions, obtained by varying F, whereas the continuous line is a numerical fit produced by exact diagonalization (compare Fig. 2A) giving an energy gap $\Delta \approx 3.4$ GHz. The curves in Fig. 2, B and C, are plotted against the change $\Delta\Phi_{ext}$ in external flux from the symmetry position indicated by the dotted line. In agreement with our numerical simulations, the step (Fig. 2B) is shifted away from the symmetry position of the energy spectrum (Fig. 2C) by a phase bias shift $\Delta\gamma_q \approx 2\pi(\Delta\Phi_{ext}/\Phi_0) \approx 0.008\pi$. The step reflects the external-flux dependence of the qubit circulating current at $I_b \approx I_{sw}$ (after the shift), whereas the spectrum reflects E_{01} at $I_b = 0$ (before the shift) (*16*).

Next, we used different MW pulse sequences to induce coherent quantum dynamics of the qubit in the time domain. For a given level separation E_{10}, a short resonant MW pulse of variable length with frequency $F = E_{10}$ was applied. Together with the MW amplitude, the pulse length defines the relative occupancy of the ground state and the excited state. The corresponding switching probability was measured with a fixed-bias current pulse amplitude. We obtained coherent Rabi oscillations of the qubit circulating current for a frequency $F = 6.6$ GHz and three different values of the MW power A (Fig. 3). The variation in switching probability is around 60%, indicating that the fidelity in a single readout is of that order. By varying A, we verified the linear dependence of the Rabi frequency on the MW amplitude, a key signature of the Rabi process (Fig. 3B). The oscillation pattern can be fitted to a damped sinusoid. For relatively strong driving (Rabi period below 10 ns), decay times τ_{Rabi} up to ~150 ns are obtained. This large decay time resulted in hundreds of coherent oscillations at large microwave power.

The Rabi scheme also allows the study of the state occupancy relaxation. This can be done by applying a coherent π pulse for full rotation of the qubit into the excited state and varying the delay time before readout. Experiments performed at $F = 5.71$ GHz gave an exponential decay with relaxation time $\tau_{relax} \approx 900$ ns.

As a next step we measured the undriven, free-evolution dephasing time τ_φ by performing a Ramsey interference experiment (*17*) as follows. Two $\pi/2$ pulses, whose length is determined from the Rabi precession presented above, are applied to the qubit. The first pulse creates a superposition of the $|0\rangle$ and $|1\rangle$ states. If the microwave frequency is detuned by $\delta F = E_{10} - F$ away from resonance, the superposition phase increases with a rate

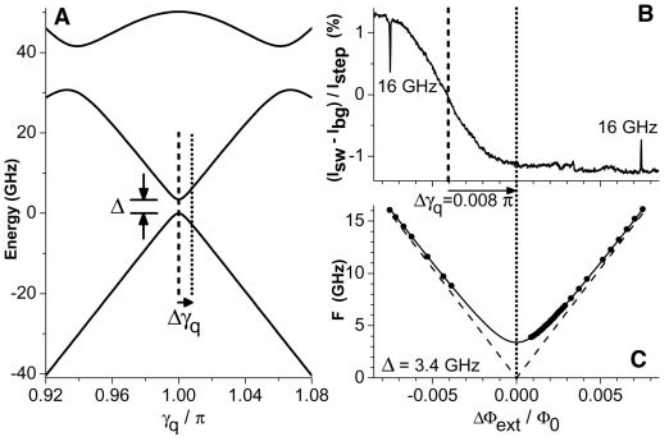


Fig. 2. (**A**) Calculated energy diagram for the three-junction qubit, for $E_J/E_C = 35$, $E_C = 7.4$ GHz, and $\alpha = 0.8$ (*11*). $\Delta\gamma_q$ indicates the phase shift induced by the SQUID bias current. (**B**) Ground-state transition step: The sinusoidal background modulation of the SQUID (I_{bg}) is subtracted from the I_b pulse amplitude corresponding to 50% switching probability (I_{sw}) and then normalized to I_{step}, the middle value (at the dashed line). A sharp peak and dip are induced by a long (1 μs) MW radiation burst at 16 GHz, allowing the symmetry point to be found (midpoint of the peak/dip positions, dotted line). Data show I_{sw} versus $\Delta\Phi_{ext}$, the deviation in external flux from this point. The transition step is displaced from this point by $\Delta\gamma_q/2\pi$. (**C**) Frequency of the resonant peaks/dips (dots) versus $\Delta\Phi_{ext}$; the continuous line is a numerical fit with the same parameters as in (A) leading to a value of $\Delta = 3.4$ GHz, whereas the dashed line depicts the case $\Delta = 0$.

$2\pi\delta F$, in the frame rotating with the MW frequency F. After a varying delay time, we apply another $\pi/2$ pulse to measure the final $|0\rangle$ and $|1\rangle$ state occupancy via the switching probability. The readout shows Ramsey fringes with a period $1/\delta F$, as in Fig. 4A, where $E_{10} = 5.71$ GHz and $\delta F = 220$ MHz. The dots represent experimental data, whereas the continuous line is an exponentially damped sinusoidal fitting curve, yielding a free-evolution dephasing time $\tau_\varphi \approx 20$ ns. Note that the oscillation period of 4.5 ns agrees well with $1/\delta F$.

Additional information on the spectral properties of the decohering fluctuations can be obtained with a modified Ramsey experiment. By inserting a π pulse between the two $\pi/2$ pulses (Fig. 4B), we obtain a spin-echo pulse configuration. The role of the π pulse is to reverse the noise-driven diffusion of the qubit phase at the midpoint in time of the free evolution. Dephasing due to fluctuations of lower frequencies should be cancelled by their opposite influence before and after the π pulse (*18*). Spin-echo oscillations (Fig. 4B) are taken under the same conditions as the Ramsey fringes, but are here recorded as a function of the π pulse position. The period ($\sim$2.3 ns) is half that of the Ramsey interference. We measured the decay of the maximum spin-echo signal (i.e., with the π pulse in the center) versus the delay time between the two $\pi/2$ pulses. The data can be fitted to a half-Gaussian (not shown) with a decay time $\tau_{echo} \approx 30$ ns.

We conclude that with the present device and setup, the dephasing time $\tau_\varphi \approx 20$ ns, as measured with the Ramsey pulses, is much shorter than the relaxation time $\tau_{relax} \approx 900$ ns. Dephasing is probably caused by a variation in time of the qubit energy splitting, attributable to external or internal noise. A likely source is external flux noise, which can be reduced in the future. The present qubit could not be operated at the symmetry point $\gamma_q = \pi$ where the influence of flux noise is minimal (*5*), presumably as the result of an accidentally close SQUID resonance (*19*). Other possible noise sources are thermal, charge, critical current, and spin fluctuations. From estimations of the Johnson noise in the bias circuit (*20, 21*), we find a contribution that is several orders of magnitude weaker.

For strong driving, Rabi oscillations persisted for times much longer than τ_φ. This constitutes no inconsistency. The dependence of the Rabi period on the detuning, due to fluctuations of the qubit energy E_{10}, is weak when the Rabi period is short. The fact that coherence is only marginally improved by the π pulse in the spin-echo experiment seems to indicate the presence of noise at frequencies beyond 10 MHz. Further analysis and additional measurements are needed.

These first results on the coherent time evolution of a flux qubit are very promising. The already high fidelity of qubit excitation and readout can no doubt be improved. Quite likely it is also possible to reduce the dephasing rate. Taken together, these results establish the superconducting flux qubit as an attractive candidate for solid-state quantum computing.

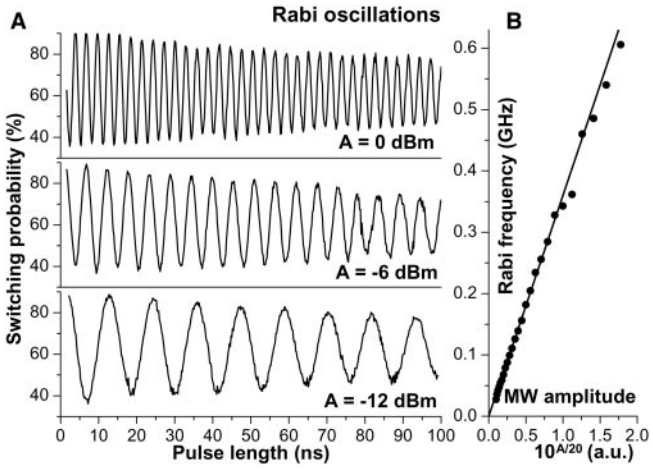


Fig. 3. (**A**) Rabi oscillations for a resonant frequency $F = E_{10} = 6.6$ GHz and three different microwave powers $A = 0$, -6, and -12 dBm, where A is the nominal microwave power applied at room temperature. The data are well fitted by exponentially damped sinusoidal oscillations. The resulting decay time is $\sim$150 ns for all powers. (**B**) Linear dependence of the Rabi frequency on the microwave amplitude, expressed as $10^{A/20}$. The slope is in agreement with estimations based on sample design.

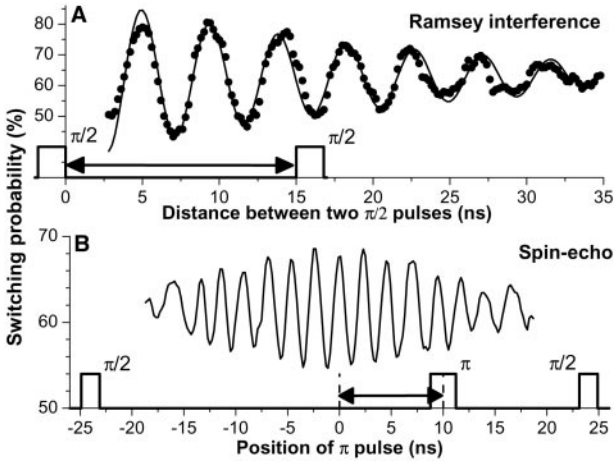


Fig. 4. (**A**) Ramsey interference: The measured switching probability (dots) is plotted against the time between the two $\pi/2$ pulses. The continuous line is a fit by exponentially damped oscillations with a decay time of 20 ns. The Ramsey interference period of 4.5 ns agrees with the inverse of the detuning from resonance, 220 MHz. The resonant frequency is 5.71 GHz and microwave power A = 0 dBm. (**B**) Spin-echo experiment: switching probability versus position of the π pulse between two $\pi/2$ pulses. The period of $\sim$2.3 ns corresponds well to half the inverse of the detuning. The width and timing of microwave pulses in the MW line are shown in each graph. The readout pulse in the bias line immediately follows the last $\pi/2$ pulse (see Fig. 1B).

References and Notes

1. A. J. Leggett, A. Garg, *Phys. Rev. Lett.* **54**, 857 (1985).
2. M. A. Nielsen, I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge Univ. Press, Cambridge, 2000).
3. Y. Makhlin *et al.*, *Rev. Mod. Phys.* **73**, 357 (2001).
4. Y. Nakamura *et al.*, *Nature* **398**, 786 (1999).
5. D. Vion *et al.*, *Science* **296**, 886 (2002).
6. Y. Yu *et al.*, *Science* **296**, 889 (2002).
7. J. M. Martinis *et al.*, *Phys. Rev. Lett.* **89**, 117901 (2002).
8. J. E. Mooij *et al.*, *Science* **285**, 1036 (1999).
9. J. R. Friedman *et al.*, *Nature* **406**, 43 (2000).
10. C. H. van der Wal *et al.*, *Science* **290**, 773 (2000).
11. The two opposite persistent current states, depicted by arrows in Fig. 1A, describe the basis $\{|\uparrow\rangle, |\downarrow\rangle\}$ of Pauli spin matrices. Using the notation $\tan 2\theta = \Delta/\varepsilon$, the qubit eigenstates can be written as $|0\rangle = \cos\theta|\uparrow\rangle + \sin\theta|\downarrow\rangle$ and $|1\rangle = -\sin\theta|\uparrow\rangle + \cos\theta|\downarrow\rangle$ and the expectation values of the corresponding circulating currents as $I_{q0,1} = \pm I_p \cos 2\theta$.
12. I. I. Rabi, *Phys. Rev.* **51**, 652 (1937).
13. M. Grifoni, P. Hänggi, *Phys. Rep.* **304**, 229 (1998).
14. In the frame rotating at the MW frequency $F = E_{10}$, the Rabi precession is around the x axis with a frequency $(\varepsilon_{mw} \sin 2\theta)/h$ [with θ as in (*11*)].
15. During the qubit initialization and control, the SQUID bias current is set to zero and, as a result of the SQUID symmetry, the qubit is decoupled from the external current noise to first order. At $I_b = 0$, small external noise current flows equally in the two branches of the SQUID even in the presence of the circulating current in the SQUID. See also (*21*).
16. A part of the energy spectrum is missing, because the readout is not efficient around the step and thus the spectroscopy signal is weak.
17. N. F. Ramsey, *Phys. Rev.* **78**, 695 (1950).
18. E. L. Hahn, *Phys. Rev.* **80**, 580 (1950).
19. In the present device, $\Delta \approx 3.4$ GHz was rather close to the SQUID plasma frequency designed to be $\sim$2 GHz (at $I_b \approx I_{sw}$). This could be a possible explanation for the absence of coherent oscillations for $F = \Delta$.
20. Lin Tian, S. Lloyd, T. P. Orlando, *Phys. Rev. B* **65**, 144516 (2002).
21. C. H. van der Wal, F. K. Wilhelm, C. J. P. M. Harmans, J. E. Mooij, *Eur. Phys. J. B* **31**, 111 (2003).

22. We thank R. N. Schouten, J. B. Majer, A. Lupascu, and K. Semba for experimental help and discussion; D. Esteve and C. Urbina for valuable input; and D. Vion, A. Aassime, C. H. van der Wal, A. C. J. ter Haar, T. Orlando, S. Lloyd, M. Grifoni, F. K. Wilhelm, L. Vandersypen, and P. C. E. Stamp for fruitful discussions. Supported by the Dutch Foundation for Fundamental Research on Matter (FOM), the European Union SQUBIT project, and the U.S. Army Research Office (grant DAAD 19-00-1-0548).

2 December 2002; accepted 4 February 2003
Published online 13 February 2003;
10.1126/science.1081045
Include this information when citing this paper.

ARTICLES

Driven coherent oscillations of a single electron spin in a quantum dot

F. H. L. Koppens[1], C. Buizert[1], K. J. Tielrooij[1], I. T. Vink[1], K. C. Nowack[1], T. Meunier[1], L. P. Kouwenhoven[1] & L. M. K. Vandersypen[1]

The ability to control the quantum state of a single electron spin in a quantum dot is at the heart of recent developments towards a scalable spin-based quantum computer. In combination with the recently demonstrated controlled exchange gate between two neighbouring spins, driven coherent single spin rotations would permit universal quantum operations. Here, we report the experimental realization of single electron spin rotations in a double quantum dot. First, we apply a continuous-wave oscillating magnetic field, generated on-chip, and observe electron spin resonance in spin-dependent transport measurements through the two dots. Next, we coherently control the quantum state of the electron spin by applying short bursts of the oscillating magnetic field and observe about eight oscillations of the spin state (so-called Rabi oscillations) during a microsecond burst. These results demonstrate the feasibility of operating single-electron spins in a quantum dot as quantum bits.

The use of quantum mechanical superposition states and entanglement in a computer can theoretically solve important mathematical and physical problems much faster than classical computers[1,2]. However, the realization of such a quantum computer represents a formidable challenge, because it requires fast and precise control of fragile quantum states. The prospects for accurate quantum control in a scalable system are thus being explored in a rich variety of physical systems, ranging from nuclear magnetic resonance and ion traps to superconducting devices[3].

Electron spin states were identified early on as an attractive realization of a quantum bit[4], because they are relatively robust against decoherence (uncontrolled interactions with the environment). Advances in the field of semiconductor quantum dots have made this system very fruitful as a host for the electron spin. Since Loss and DiVincenzo's proposal[5] on electron spin qubits in quantum dots in 1998, many of the elements necessary for quantum computation have been realized experimentally. It is now routine to isolate with certainty a single electron in each of two coupled quantum dots[6–9]. The spin of this electron can be reliably initialized to the ground state, spin-up, via optical pumping[10] or by thermal equilibration at sufficiently low temperatures and strong static magnetic fields (for example, $T = 100$ mK and $B_{ext} = 1$ T). The spin states are also very long-lived, with relaxation times of the order of milliseconds[11–13]. Furthermore, a lower bound on the spin coherence time exceeding 1 μs was established, using spin-echo techniques on a two-electron system[14]. These long relaxation and coherence times are possible in part because the magnetic moment of a single electron spin is so weak. On the other hand, this property makes read-out and manipulation of single spins particularly challenging. By combining spin-to-charge conversion with real-time single-charge detection[15–17], it has nevertheless been possible to accomplish single-shot read-out of spin states in a quantum dot[13,18].

The next major achievement was the observation of the coherent exchange of two electron spins in a double dot system, controlled by fast electrical switching of the tunnel coupling between the two quantum dots[14]. Finally, free evolution of a single electron spin about a static magnetic field (Larmor precession) has been observed, via optical pump–probe experiments[19,20]. The only missing ingredient for universal quantum computation with spins in dots remained the demonstration of driven coherent spin rotations (Rabi oscillations) of a single electron spin.

The most commonly used technique for inducing spin flips is electron spin resonance (ESR)[21]. ESR is the physical process whereby electron spins are rotated by an oscillating magnetic field B_{ac} (with frequency f_{ac}) that is resonant with the spin precession frequency in an external magnetic field B_{ext}, oriented perpendicularly to B_{ac} ($hf_{ac} = g\mu_B B_{ext}$, where μ_B is the Bohr magneton and g the electron spin g-factor). Magnetic resonance of a single electron spin in a solid has been reported in a few specific cases[22–24], but has never been realized in semiconductor quantum dots. Detecting ESR in a single quantum dot is conceptually simple[25], but experimentally difficult to realize, as it requires a strong, high-frequency magnetic field at low temperature, while accompanying alternating electric fields must be minimized. Alternative schemes for driven rotations of a spin in a dot have been proposed, based on optical excitation[26] or electrical control[27–29], but this is perhaps even more challenging and has not been accomplished either.

Here, we demonstrate the ability to control the spin state of a single electron confined in a double quantum dot via ESR. In a double dot system, spin-flips can be detected through the transition of an electron from one dot to the other[30,31] rather than between a dot and a reservoir, as would be the case for a single dot. This has the advantage that there is no need for the electron spin Zeeman splitting (used in a single dot for spin-selective tunnelling) to exceed the temperature of the electron reservoirs (~100 mK; the phonon temperature was ~40 mK). The experiment can thus be performed at a smaller static magnetic field, and consequently with lower, technically less demanding, excitation frequencies. Furthermore, by applying a large bias voltage across the double dot, the spin detection can be made much less sensitive to electric fields than is possible in the single-dot case (electric fields can cause photon-assisted tunnelling; see Supplementary Discussion). Finally, in a double dot, single-spin

[1]Kavli Institute of NanoScience, Delft University of Technology, PO Box 5046, 2600 GA, Delft, The Netherlands.

operations can in future experiments be combined with two-qubit operations to realize universal quantum gates[5], and with spin read-out to demonstrate entanglement[32,33].

Device and ESR detection concept

Two coupled semiconductor quantum dots are defined by surface gates (Fig. 1a) on top of a two-dimensional electron gas. By applying the appropriate negative voltages to the gates the dots can be tuned to the few-electron regime[8]. The oscillating magnetic field that drives the spin transitions is generated by applying a radio-frequency (RF) signal to an on-chip coplanar stripline (CPS) which is terminated in a narrow wire, positioned near the dots and separated from the surface gates by a 100-nm-thick dielectric (Fig. 1b). The current through the wire generates an oscillating magnetic field B_{ac} at the dots, perpendicular to the static external field B_{ext} and slightly stronger in the left dot than in the right dot (see Supplementary Fig. S1).

To detect the ESR-induced spin rotations, we use electrical transport measurements through the two dots in series in the spin blockade regime where current flow depends on the relative spin state of the electrons in the two dots[30,34]. In brief, the device is operated so that current is blocked owing to spin blockade, but this blockade is lifted if the ESR condition ($hf_{ac} = g\mu_B B_{ext}$) is satisfied.

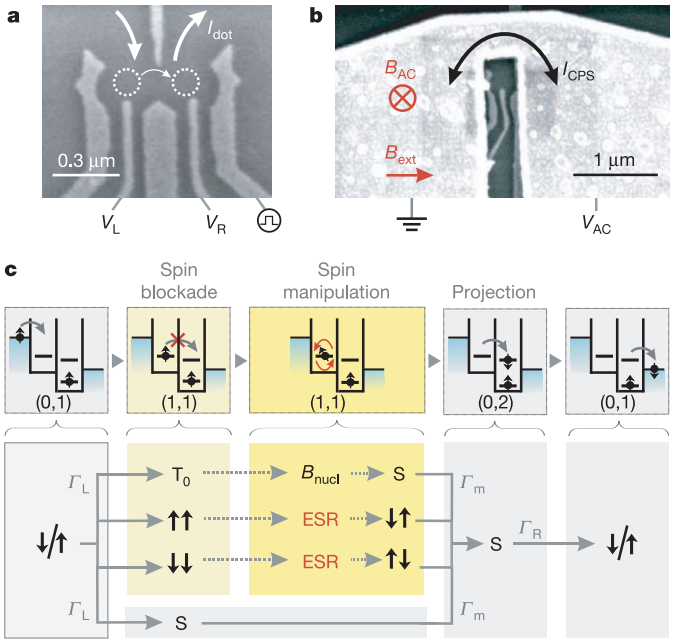


Figure 1 | Device and ESR detection scheme. a, Scanning electron microscope (SEM) image of a device with the same gate pattern as used in the experiment. The Ti/Au gates are deposited on top of a GaAs/AlGaAs heterostructure containing a two-dimensional electron gas 90 nm below the surface. White arrows indicate current flow through the two coupled dots (dotted circles). The right side gate is fitted with a homemade bias-tee (rise time 150 ps) to allow fast pulsing of the dot levels. **b**, SEM image of a device similar to the one used in the experiment. The termination of the coplanar stripline is visible on top of the gates. The gold stripline has a thickness of 400 nm and is designed to have a 50 Ω characteristic impedance, Z_0, up to the shorted termination. It is separated from the gate electrodes by a 100-nm-thick dielectric (Calixerene)[50]. **c**, Diagrams illustrating the transport cycle in the spin blockade regime. This cycle can be described via the occupations (m,n) of the left and right dots as (0,1) → (1,1) → (0,2) → (0,1). When an electron enters the left dot (with rate Γ_L) starting from (0,1), the two-electron system that is formed can be either a singlet S(1,1) or a triplet T(1,1). From S(1,1), further current flow is possible via a transition to S(0,2) (with rate Γ_m). When the system is in T(1,1), current is blocked unless this state is coupled to S(1,1). For T_0, this coupling is provided by the inhomogeneous nuclear field ΔB_N. For T_+ or T_-, ESR causes a transition to ↑↓ or ↓↑, which contains a S(1,1) component and a T_0 component (which is in turn coupled to S(1,1) by the nuclear field).

This spin blockade regime is accessed by tuning the gate voltages such that one electron always resides in the right dot, and a second electron can tunnel from the left reservoir to the left dot (Fig. 1c and Supplementary Fig. S2). If this electron forms a double-dot singlet state with the electron in the right dot ($S = \uparrow\downarrow - \downarrow\uparrow$; normalization omitted for brevity), it is possible for the left electron to move to the right dot, and then to the right lead (leaving behind an electron in the right dot with spin ↑ or spin ↓), since the right dot singlet state is energetically accessible. If, however, the two electrons form a double-dot triplet state, the left electron cannot move to the right dot because the right dot's triplet state is much higher in energy. The electron also cannot move back to the lead and therefore further current flow is blocked as soon as any of the (double-dot) triplet states is formed.

Role of the nuclear spin bath for ESR detection

In fact, the situation is more complex, because each of the two spins experiences a randomly oriented and fluctuating effective nuclear field of ~1–3 mT (refs 35, 36). This nuclear field, B_N, arises from the hyperfine interaction of the electron spins with the Ga and As nuclear spins in the host material, and is in general different in the two dots, with a difference of ΔB_N. At zero external field and for sufficiently small double dot singlet–triplet splitting (see Supplementary Fig. S2d), the inhomogeneous component of the nuclear field causes all three triplet states (T_0, T_+ and T_-) to be admixed with the singlet S (for example, $T_0 = \uparrow\downarrow + \downarrow\uparrow$ evolves into $S = \uparrow\downarrow - \downarrow\uparrow$ due to $\Delta B_{N,z}$, and $T_+ = \uparrow\uparrow$ and $T_- = \downarrow\downarrow$ evolve into S owing to $\Delta B_{N,x}$). As a result, spin blockade is lifted. For $B_{ext} \gg \sqrt{\langle B_N^2 \rangle}$, however, the T_+ and T_- states split off in energy, which makes hyperfine-induced admixing between $T_\pm$ and S ineffective (T_0 and S remain admixed; see Fig. 2a). Here spin blockade does occur, whenever a state with parallel spins (↑↑ or ↓↓) becomes occupied.

ESR is then detected as follows (see Fig. 1c). An oscillating magnetic field resonant with the Zeeman splitting can flip the spin in the left or the right dot. Starting from ↑↑ or ↓↓, the spin state then changes to ↑↓ (or ↓↑). If both spins are flipped, transitions occur between ↑↑ and ↓↓ via the intermediate state $\frac{\uparrow\downarrow+\downarrow\uparrow}{\sqrt{2}}\frac{\uparrow\downarrow+\downarrow\uparrow}{\sqrt{2}}$. In both cases, states with anti-parallel spins ($S_z = 0$) are created owing to ESR. Expressed in the singlet-triplet measurement basis, ↑↓ or ↓↑ is a superposition of the T_0 and S state ($\uparrow\downarrow = T_0 + S$). For the singlet component of this state, the left electron can transition immediately to the right dot and from there to the right lead. The T_0 component first evolves into a singlet due to the nuclear field and then the left electron can move to the right dot as well. Thus whenever the spins are anti-parallel, one electron charge moves through the dots. If such transitions from parallel to anti-parallel spins are induced repeatedly at a sufficiently high rate, a measurable current flows through the two dots.

ESR spectroscopy

The resonant ESR response is clearly observed in the transport measurements as a function of magnetic field (Fig. 2a, b), where satellite peaks develop at the resonant field $B_{ext} = \pm hf_{ac}/g\mu_B$ when the RF source is turned on (the zero-field peak arises from the inhomogeneous nuclear field, which admixes all the triplets with the singlet[36,37]). The key signature of ESR is the linear dependence of the satellite peak location on the RF frequency, which is clearly seen in the data of Fig. 2c, where the RF frequency is varied from 10 to 750 MHz. From a linear fit through the top of the peaks we obtain a g-factor with modulus 0.35 ± 0.01, which lies within the range of reported values for confined electron spins in GaAs quantum dots[11,38–40]. We also verified explicitly that the resonance we observe is magnetic in origin and not caused by the electric field that the CPS generates as well; negligible response was observed when RF power is applied to the right side gate, generating mostly a RF electric field (see Supplementary Fig. S3).

The amplitude of the peaks in Fig. 2b increases linearly with RF power (~B_{ac}^2) before saturation occurs, as predicted[25] (Fig. 2b, inset). The ESR satellite peak is expected to be broadened by either the

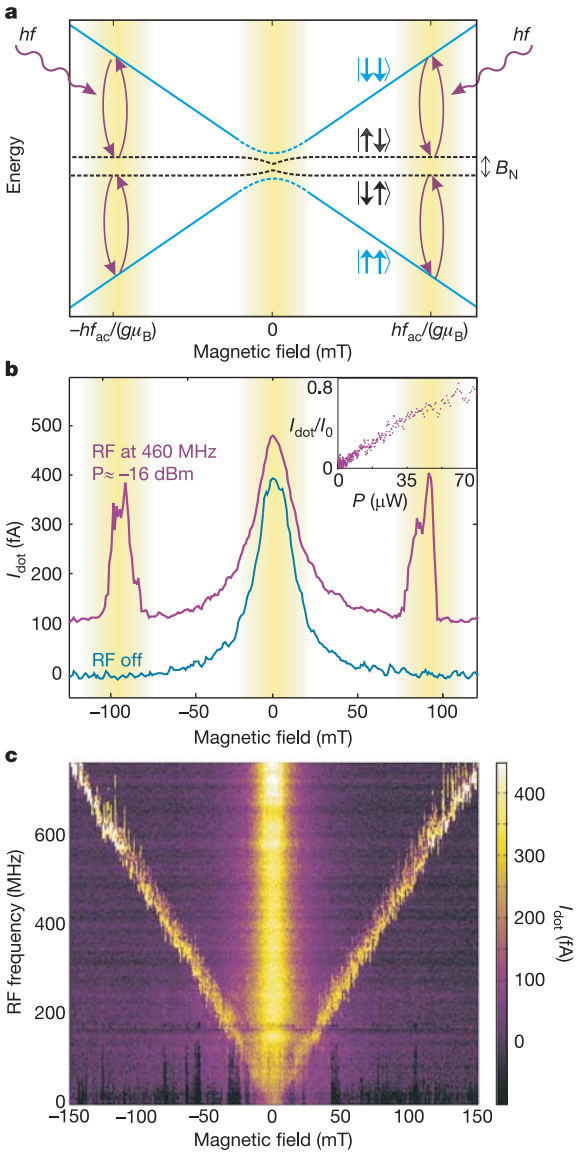


Figure 2 | ESR spin state spectroscopy. a, Energy diagram showing the relevant eigenstates of two electron spins in a double-dot, subject to an external magnetic field and nuclear fields. Because the nuclear field is generally inhomogeneous, the Zeeman energy is different in the two dots and results therefore in a different energy for $\uparrow\downarrow$ and $\downarrow\uparrow$. ESR turns the spin states $\uparrow\uparrow$ and $\downarrow\downarrow$ into $\uparrow\downarrow$ or $\downarrow\uparrow$, depending on the nuclear fields in the two dots. The yellow bands denote the ranges in B_{ext} where spin blockade is lifted (by the nuclear field or ESR) and current will flow through the dots. **b**, Current measured through the double-dot in the spin blockade regime, with (red trace, offset by 100 fA for clarity) and without (blue trace) a RF magnetic field. Satellite peaks appear as the external magnetic field is swept through the spin resonance condition. Each measurement point is averaged for one second, and is therefore expected to represent an average response over many nuclear configurations. The RF power P applied to the CPS is estimated from the power applied to the coax line and the attenuation in the lines. Inset, satellite peak height versus RF power ($f = 408$ MHz, $B_{ext} = 70$ mT, taken at slightly different gate voltage settings). The current is normalized to the current at $B_{ext} = 0$ ($= I_0$). Unwanted electric field effects are reduced by applying a compensating signal to the right side gate with opposite phase as the signal on the stripline (see Supplementary Fig. S4). This allowed us to obtain this curve up to relatively high RF powers. **c**, Current through the dots when sweeping the RF frequency and stepping the magnetic field. The ESR satellite peak is already visible at a small magnetic field of 20 mT and RF excitation of 100 MHz, and its location evolves linearly in field when increasing the frequency. For higher frequencies the satellite peak is broadened asymmetrically for certain sweeps, visible as vertical stripes. This broadening is time dependent, hysteretic in sweep direction, and changes with the dot level alignment. The horizontal line at 180 MHz is due to a resonance in the transmission line inside the dilution refrigerator.

excitation amplitude B_{ac} or incoherent processes, like cotunnelling, inelastic transitions (to the S(0,2) state) or the statistical fluctuations in the nuclear field, whichever of the four has the largest contribution. No dependence of the width on RF power was found within the experimentally accessible range ($B_{ac} < 2$ mT). Furthermore, we suspect that the broadening is not dominated by cotunnelling or inelastic transitions because the corresponding rates are smaller than the observed broadening (see Supplementary Figs S4b and S2d). The observed ESR peaks are steeper on the flanks and broader than expected from the nuclear field fluctuations. In many cases, the peak width and position are even hysteretic in the sweep direction, suggesting that the resonance condition is shifted during the field sweep. We speculate that dynamic nuclear polarization due to feedback of the electron transport on the nuclear spins plays a central part here[37].

Coherent Rabi oscillations

Following the observation of magnetically induced spin flips, we next test whether we can also coherently rotate the spin by applying RF bursts with variable length. In contrast to the continuous-wave experiment, where detection and spin rotation occur at the same time, we pulse the system into Coulomb blockade during the spin manipulation. This eliminates decoherence induced by tunnel events from the left to the right dot during the spin rotations. The experiment consists of three stages (Fig. 3): initialization through spin blockade in a statistical mixture of $\uparrow\uparrow$ and $\downarrow\downarrow$, manipulation by a RF burst in Coulomb blockade, and detection by pulsing back for projection (onto S(0,2)) and tunnelling to the lead. When one of the electrons is rotated over $(2n + 1)\pi$ (with integer n), the two-electron state evolves to $\uparrow\downarrow$ (or $\downarrow\uparrow$), giving a maximum contribution to the current (as before, when the two spins are anti-parallel, one electron charge moves through the dots). However, no electron flow is expected after rotations of $2\pi n$, where one would find two parallel spins in the two dots after the RF burst.

We observe that the dot current oscillates periodically with the RF burst length (Fig. 4). This oscillation indicates that we performed driven, coherent electron spin rotations, or Rabi oscillations. A key characteristic of the Rabi process is a linear dependence of the Rabi frequency on the RF burst amplitude, B_{ac} ($f_{Rabi} = g\mu_B B_1/h$ with $B_1 = B_{ac}/2$ due to the rotating wave approximation). We verify this by extracting the Rabi frequency from a fit of the current oscillations of Fig. 4b with a sinusoid, which gives the expected linear behaviour

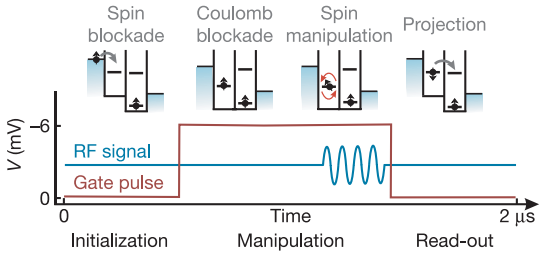


Figure 3 | The control cycle for coherent manipulation of the electron spin. During the 'initialization' stage the double-dot is tuned in the spin blockade regime. Electrons will move from left to right until the system is blocked with two parallel spins (either $\uparrow\uparrow$ or $\downarrow\downarrow$; in the figure only the $\uparrow\uparrow$ case is shown). For the 'manipulation' stage, the right dot potential is pulsed up so none of the levels in the right dot are accessible (Coulomb blockade), and a RF burst with a variable duration is applied. 'Read-out' of the spin state at the end of the manipulation stage is done by pulsing the right dot potential back; electron tunnelling to the right lead will then take place only if the spins were anti-parallel. The duration of the read-out and initialization stages combined was 1 μs, long enough ($1\ \mu s >> 1/\Gamma_L, 1/\Gamma_M, 1/\Gamma_R$) to have parallel spins in the dots at the end of the initialization stage with near certainty (this is checked by signal saturation when the pulse duration is prolonged). The duration of the manipulation stage is also held fixed at 1 μs to keep the number of pulses per second constant. The RF burst is applied just before the read-out stage starts.

(Fig. 4b, inset). From the fit we obtain $B_{ac} = 0.59$ mT for a stripline current I_{CPS} of ~1 mA, which agrees well with predictions from numerical finite element simulations (see Supplementary Fig. S1). The maximum B_1 we could reach in the experiment before electric field effects hindered the measurement was 1.9 mT, corresponding to $\pi/2$ rotations of only 27 ns (that is, a Rabi period of 108 ns, see Fig. 4b). If the accompanying electric fields from the stripline excitation could be reduced in future experiments (for example, by improving the impedance matching from coax to CPS), considerably faster Rabi flopping should be attainable.

The oscillations in Fig. 4b remain visible throughout the entire measurement range, up to 1 μs. This is striking, because the Rabi period of ~100 ns is much longer than the time-averaged coherence time T_2^* of 10–20 ns (refs 14, 19, 35, 36) caused by the nuclear field fluctuations. The slow damping of the oscillations is only possible because the nuclear field fluctuates very slowly compared to the timescale of spin rotations and because other mechanisms, such as the spin-orbit interaction, disturb the electron spin coherence only on even longer timescales[13,41,42]. We also note that the decay is not exponential (grey line in Fig. 4a), which is related to the fact that the nuclear bath is non-markovian (it has a long memory)[43].

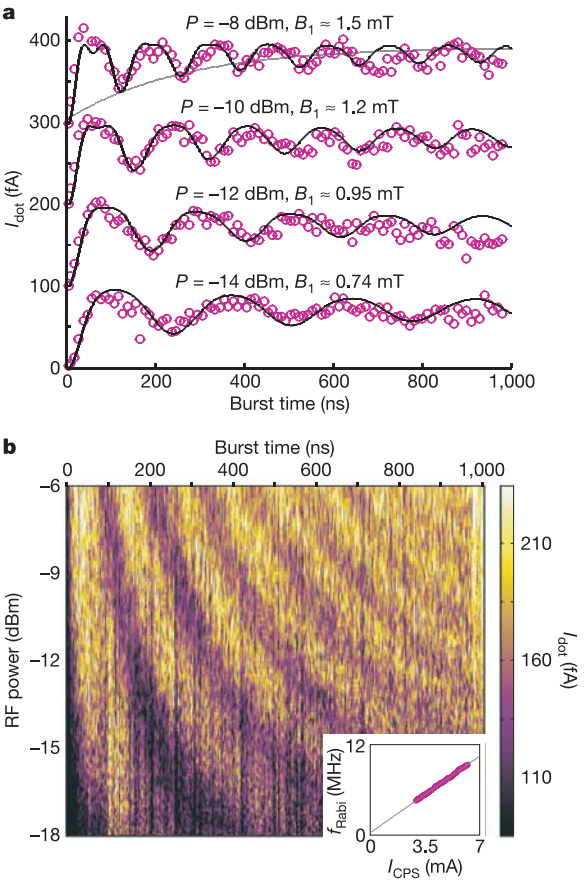


Figure 4 | Coherent spin rotations. a, The dot current—reflecting the spin state at the end of the RF burst—oscillates as a function of RF burst length (curves offset by 100 fA for clarity). The frequency of B_{ac} is set at the spin resonance frequency of 200 MHz ($B_{ext} = 41$ mT). The period of the oscillation increases and is more strongly damped for decreasing RF power. The RF power P applied to the CPS is estimated from the power applied to the coax line and the attenuation in the lines and RF switch. From P, the stripline current is calculated via the relation $P = \frac{1}{2}\left(\frac{I_{CPS}}{2}\right)^2 Z_0$ assuming perfect reflection of the RF wave at the short. Each measurement point is averaged over 15 s. We correct for a current offset which is measured with the RF frequency off-resonance (280 MHz). The solid lines are obtained from numerical computation of the time evolution, as discussed in the text. The grey line corresponds to an exponentially damped envelope. **b**, The oscillating dot current (represented in colourscale) is displayed over a wide range of RF powers (the sweep axis) and burst durations. The dependence of the Rabi frequency f_{Rabi} on RF power is shown in the inset. f_{Rabi} is extracted from a sinusoidal fit with the current oscillations from 10 to 500 ns for RF powers ranging from −12.5 dBm up to −6 dBm.

Theoretical model

To understand better the amplitudes and decay times of the oscillations, we model the time evolution of the spins throughout the burst duration. The model uses a hamiltonian that includes the Zeeman splitting for the two spins and the RF field, which we take to be of equal amplitude in both dots ($\mathbf{S}_L$ and $\mathbf{S}_R$ refer to the electron spins in the left and right dot respectively):

$$H = g\mu_B(\mathbf{B}_{ext} + \mathbf{B}_{L,N})\mathbf{S}_L + g\mu_B(\mathbf{B}_{ext} + \mathbf{B}_{R,N})\mathbf{S}_R$$
$$+ g\mu_B \cos(\omega t)\mathbf{B}_{ac}(\mathbf{S}_L + \mathbf{S}_R)$$

where $\mathbf{B}_{L,N}$ and $\mathbf{B}_{R,N}$ correspond to a single frozen configuration of the nuclear field in the left and right dot. This is justified because the electron spin dynamics is much faster than the dynamics of the nuclear system. From the resulting time evolution operator and assuming that the initial state is a statistical mixture of $\uparrow\uparrow$ and $\downarrow\downarrow$, we can numerically obtain the probability for having anti-parallel spins after the RF burst. This is also the probability that the left electron tunnels to the right dot during the read-out stage.

In the current measurements of Fig. 4a, each data point is averaged over 15 s, which presumably represents an average over many nuclear configurations. We include this averaging over different nuclear configurations in the model by taking 2,000 samples from a gaussian distribution of nuclear fields (with standard deviation $\sigma = \sqrt{\langle B_N^2 \rangle}$), and computing the probability that an electron tunnels out after the RF burst. When the electron tunnels, one or more additional electrons, say m, may subsequently tunnel through before $\uparrow\uparrow$ or $\downarrow\downarrow$ is formed and the current is blocked again. Taking m and σ as fitting parameters, we find good agreement with the data for m=1.5 and $\sigma = 2.2$ mT (solid black lines in Fig. 4a). This value for σ is comparable to that found in refs 35 and 36. The value found for m is different from what we would expect from a simple picture where all four spin states are formed with equal probability during the initialization stage, which would give $m = 1$. We do not understand this discrepancy, but it could be due to different tunnel rates for $\uparrow$ and $\downarrow$ or more subtle details in the transport cycle that we have neglected in the model.

Time evolution of the spin states during RF bursts

We now discuss in more detail the time evolution of the two spins during a RF burst. The resonance condition in each dot depends on the effective nuclear field, which needs to be added vectorially to B_{ext}. Through their continuous reorientation, the nuclear spins will bring the respective electron spins in the two dots on and off resonance as time progresses.

When a RF burst is applied to two spins initially in $\uparrow\uparrow$, and is on-resonance with the right spin only, the spins evolve as:

$$|\uparrow\rangle|\uparrow\rangle \rightarrow |\uparrow\rangle\frac{|\uparrow\rangle+|\downarrow\rangle}{\sqrt{2}} \rightarrow |\uparrow\rangle|\downarrow\rangle \rightarrow$$
$$|\uparrow\rangle\frac{|\uparrow\rangle-|\downarrow\rangle}{\sqrt{2}} \rightarrow |\uparrow\rangle|\uparrow\rangle$$

When the RF burst is on-resonance with both spins, the time evolution is:

$$|\uparrow\rangle|\uparrow\rangle \rightarrow \frac{|\uparrow\rangle+|\downarrow\rangle}{\sqrt{2}}\frac{|\uparrow\rangle+|\downarrow\rangle}{\sqrt{2}} \rightarrow |\downarrow\rangle|\downarrow\rangle \rightarrow$$
$$\frac{|\uparrow\rangle-|\downarrow\rangle}{\sqrt{2}}\frac{|\uparrow\rangle-|\downarrow\rangle}{\sqrt{2}} \rightarrow |\uparrow\rangle|\uparrow\rangle$$

In both cases, the RF causes transitions between the ↑ and ↓ states of single spin-half particles. When the RF is on-resonance with both spins, such single-spin rotations take place for both spins simultaneously. Because the current through the dots is proportional to the $S_z = 0$ probability (↑↓ or ↓↑), we see that when both spins are excited simultaneously, the current through the dots will oscillate twice as fast as when only one spin is excited, but with only half the amplitude.

In the experiment, the excitation is on-resonance with only one spin at a time for most of the frozen nuclear configurations (Fig. 5). Only at the highest powers ($B_1/\sqrt{\langle B_{N,z}^2\rangle} > 1$), both spins may be excited simultaneously (but independently) and a small double Rabi frequency contribution is expected, although it could not be observed, owing to the measurement noise.

Quantum gate fidelity

We can estimate the angle over which the electron spins are rotated in the Bloch sphere based on our knowledge of B_1 and the nuclear field fluctuations in the z-direction, again using the hamiltonian H. For the maximum ratio of $B_1/\sqrt{\langle B_{N,z}^2\rangle} = B_1/(\sigma/\sqrt{3}) = 1.5$ reached in the present experiment, we achieve an average tip angle of 131° for an intended 180° rotation, corresponding to a fidelity of 73% (Fig. 5). Apart from using a stronger B_1, the tip angle can be increased considerably by taking advantage of the long timescale of the nuclear field fluctuations. First, application of composite pulses, widely used in nuclear magnetic resonance to compensate for resonance off-sets[44], can greatly improve the quality of the rotations. A second solution comprises a measurement of the nuclear field (nuclear state narrowing[45–47]), so that the uncertainty in the nuclear field is reduced, and accurate rotations can be realized for as long as the nuclear field remains constant.

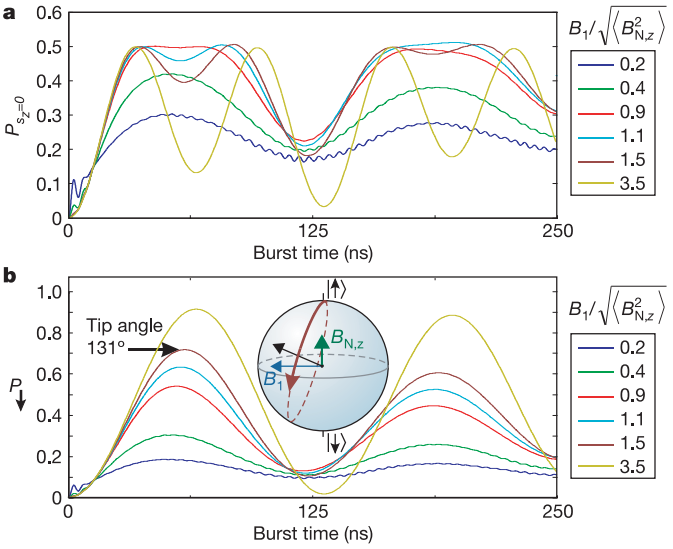


Figure 5 | Time evolution of the spin states. a, Probability for the two spins to be in ↑↓ or ↓↑ ($S_z = 0$) at the end of a RF burst, with initial state ↑↑, computed using the hamiltonian H presented in the main text, for six different values of $\sigma_{N,z} = \langle B_{N,z}^2\rangle^{1/2}$ (fixed $B_1 = 1.5$ mT, $B_{ext} = 40$ mT, each of the traces is averaged over 2,000 static nuclear configurations). As expected, the oscillation contains a single frequency for B_1 small compared to $\sigma_{N,z}$, corresponding to the Rabi oscillation of a single spin. The oscillation develops a second frequency component, twice as fast as the first, when $B_1/\sigma_{N,z} > 1$. For $B_1/\sigma_{N,z} > 4$ the double frequency component is dominant, reflecting the simultaneous Rabi oscillation of the two spins. **b**, Probability for one of the spins to be ↓ at the end of a RF burst. The spin state evolution is computed as in **a**. This oscillation represents the Rabi oscillation of one spin by itself. For increasing B_1, the maximum angle over which the spin is rotated in the Bloch sphere increases as well. In the experiment, this angle could not be measured directly, because the current measurement constitutes a two-spin measurement, not a single-spin measurement. We can, however, extract the tip angle from $P_\downarrow$.

In future experiments, controllable addressing of the spins in the two dots separately can be achieved through a gradient in either the static or the oscillating magnetic field. Such gradient fields can be created relatively easily using a ferromagnet or an asymmetric stripline. Alternatively, the resonance frequency of the spins can be selectively shifted using local g-factor engineering[48,49]. The single spin rotations reported here, in combination with single-shot spin read-out[13,18] and the tunable exchange coupling in double dots[14], offers many new opportunities, such as measuring the violation of Bell's inequalities or the implementation of simple quantum algorithms.

Received 26 April; accepted 6 July 2006.

1. Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge Univ. Press, Cambridge, 2000).
2. Shor, P. W. in *Proc. 35th Annu. Symp. on the Foundations of Computer Science* (ed. Goldwasser, S.) 124–134 (IEEE Computer Society Press, Los Alamitos, California, 1994).
3. Zoller, P. *et al.* Quantum information processing and communication, Strategic report on current status, visions and goals for research in Europe. *Eur. Phys. J. D* **36,** 203–228 (2005).
4. DiVincenzo, D. P. Quantum computation. *Science* **270,** 255–261 (1995).
5. Loss, D. & DiVincenzo, D. P. Quantum computation with quantum dots. *Phys. Rev. A* **57,** 120–126 (1998).
6. Austing, D. G., Honda, T., Muraki, K., Tokura, Y. & Tarucha, S. Quantum dot molecules. *Phys. B Cond. Matter* **249–251,** 206–209 (1998).
7. Ciorga, M. *et al.* Addition spectrum of a lateral dot from Coulomb and spin-blockade spectroscopy. *Phys. Rev. B* **61,** R16315 (2000).
8. Elzerman, J. M. *et al.* Few-electron quantum dot circuit with integrated charge read out. *Phys. Rev. B* **67,** 161308 (2003).
9. Bayer, M. *et al.* Coupling and entangling of quantum states in quantum dot molecules. *Science* **291,** 451–453 (2001).
10. Atature, M. *et al.* Quantum-dot spin-state preparation with near-unity fidelity. *Science* **312,** 551–553 (2006).
11. Hanson, R. *et al.* Zeeman energy and spin relaxation in a one-electron quantum dot. *Phys. Rev. Lett.* **91,** 196802 (2003).
12. Fujisawa, T., Austing, D. G., Tokura, Y., Hirayama, Y. & Tarucha, S. Allowed and forbidden transitions in artificial hydrogen and helium atoms. *Nature* **419,** 278–281 (2002).
13. Elzerman, J. M. *et al.* Single-shot read-out of an individual electron spin in a quantum dot. *Nature* **430,** 431–435 (2004).
14. Petta, J. R. *et al.* Coherent manipulation of coupled electron spins in semiconductor quantum dots. *Science* **309,** 2180–2184 (2005).
15. Schleser, R. *et al.* Time-resolved detection of individual electrons in a quantum dot. *Appl. Phys. Lett.* **85,** 2005–2007 (2004).
16. Vandersypen, L. M. K. *et al.* Real-time detection of single-electron tunneling using a quantum point contact. *Appl. Phys. Lett.* **85,** 4394–4396 (2004).
17. Lu, W., Ji, Z. Q., Pfeiffer, L., West, K. W. & Rimberg, A. J. Real-time detection of electron tunnelling in a quantum dot. *Nature* **423,** 422–425 (2003).
18. Hanson, R. *et al.* Single-shot readout of electron spin states in a quantum dot using spin-dependent tunnel rates. *Phys. Rev. Lett.* **94,** 196802 (2005).
19. Dutt, M. V. G. *et al.* Stimulated and spontaneous optical generation of electron spin coherence in charged GaAs quantum dots. *Phys. Rev. Lett.* **94,** 227403 (2005).
20. Greilich, A. *et al.* Optical control of spin coherence in singly charged (In,Ga)As/GaAs quantum dots. *Phys. Rev. Lett.* **96,** 227401 (2006).
21. Poole, C. P. *Electron Spin Resonance* 2nd edn (Wiley, New York, 1983).
22. Xiao, M., Martin, I., Yablonovitch, E. & Jiang, H. W. Electrical detection of the spin resonance of a single electron in a silicon field-effect transistor. *Nature* **430,** 435–439 (2004).
23. Jelezko, F., Gaebel, T., Popa, I., Gruber, A. & Wrachtrup, J. Observation of coherent oscillations in a single electron spin. *Phys. Rev. Lett.* **92,** 076401 (2004).
24. Rugar, D., Budakian, R., Mamin, H. J. & Chui, B. W. Single spin detection by magnetic resonance force microscopy. *Nature* **430,** 329–332 (2004).
25. Engel, H. A. & Loss, D. Detection of single spin decoherence in a quantum dot via charge currents. *Phys. Rev. Lett.* **86,** 4648–4651 (2001).
26. Imamoglu, A. *et al.* Quantum information processing using quantum dot spins and cavity QED. *Phys. Rev. Lett.* **83,** 4204–4207 (1999).
27. Kato, Y., Myers, R. C., Gossard, A. C. & Awschalom, D. D. Coherent spin manipulation without magnetic fields in strained semiconductors. *Nature* **427,** 50–53 (2003).
28. Golovach, V. N., Borhani, M. & Loss, D. Electric dipole induced spin resonance in quantum dots. Preprint at ⟨www.arXiv.org/cond-mat/0601674⟩ (2006).
29. Tokura, Y., Van der Wiel, W. G., Obata, T. & Tarucha, S. Coherent single electron spin control in a slanting Zeeman field. *Phys. Rev. Lett.* **96,** 047202 (2006).
30. Ono, K., Austing, D. G., Tokura, Y. & Tarucha, S. Current rectification by Pauli exclusion in a weakly coupled double quantum dot system. *Science* **297,** 1313–1317 (2002).

31. Engel, H. A. *et al.* Measurement efficiency and n-shot readout of spin qubits. *Phys. Rev. Lett.* **93,** 106804 (2004).
32. Blaauboer, M. & DiVincenzo, D. P. Detecting entanglement using a double-quantum-dot turnstile. *Phys. Rev. Lett.* **95,** 160402 (2005).
33. Engel, H. A. & Loss, D. Fermionic bell-state analyzer for spin qubits. *Science* **309,** 586–588 (2005).
34. Johnson, A. C., Petta, J. R., Marcus, C. M., Hanson, M. P. & Gossard, A. C. Singlet-triplet spin blockade and charge sensing in a few-electron double quantum dot. *Phys. Rev. B* **72,** 165308 (2005).
35. Johnson, A. C. *et al.* Triplet-singlet spin relaxation via nuclei in a double quantum dot. *Nature* **435,** 925–928 (2005).
36. Koppens, F. H. L. *et al.* Control and detection of singlet-triplet mixing in a random nuclear field. *Science* **309,** 1346–1350 (2005).
37. Jouravlev, O. N. & Nazarov, Y. V. Electron transport in a double quantum dot governed by a nuclear magnetic field. *Phys. Rev. Lett.* **96,** 176804 (2006).
38. Potok, R. M. *et al.* Spin and polarized current from Coulomb blockaded quantum dots. *Phys. Rev. Lett.* **91,** 016802 (2003).
39. Willems van Beveren, L. H. W. *et al.* Spin filling of a quantum dot derived from excited-state spectroscopy. *New J. Phys.* **7,** 182 (2005).
40. Kogan, A. *et al.* Measurements of Kondo and spin splitting in single-electron transistors. *Phys. Rev. Lett.* **93,** 166602 (2004).
41. Kroutvar, M. *et al.* Optically programmable electron spin memory using semiconductor quantum dots. *Nature* **432,** 81–84 (2004).
42. Golovach, V. N., Khaetskii, A. & Loss, D. Phonon-induced decay of the electron spin in quantum dots. *Phys. Rev. Lett.* **93,** 016601 (2004).
43. Coish, W. A. & Loss, D. Hyperfine interaction in a quantum dot: Non-Markovian electron spin dynamics. *Phys. Rev. B* **70,** 195340 (2004).
44. Vandersypen, L. M. K. & Chuang, I. L. NMR techniques for quantum control and computation. *Rev. Mod. Phys.* **76,** 1037–1069 (2004).
45. Klauser, D., Coish, W. A. & Loss, D. Nuclear spin state narrowing via gate-controlled Rabi oscillations in a double quantum dot. *Phys. Rev. Lett.* **96,** 176804 (2006).
46. Giedke, G., Taylor, J. M., D'Alessandro, D., Lukin, D. & Imamoglu, A. Quantum measurement of the nuclear spin polarization in quantum dots. Preprint at ⟨www.arXiv.org/quant-ph/0508144⟩ (2005).
47. Stepanenko, D., Burkard, G., Giedke, G. & Imamoglu, A. Enhancement of electron spin coherence by optical preparation of nuclear spins. *Phys. Rev. Lett.* **96,** 136401 (2006).
48. Salis, G. *et al.* Electrical control of spin coherence in semiconductor nanostructures. *Nature* **414,** 619–622 (2001).
49. Jiang, H. W. & Yablonovitch, E. Gate-controlled electron spin resonance in $GaAs/Al_xGa_{1-x}As$ heterostructures. *Phys. Rev. B* **64,** 041307 (2001).
50. Holleitner, A. W., Blick, R. H. & Eberl, K. Fabrication of coupled quantum dots for multiport access. *Appl. Phys. Lett.* **82,** 1887–1889 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank W. Coish, J. Elzerman, D. Klauser, A. Lupascu, D. Loss and in particular J. Folk for discussions; R. Schouten, B. van der Enden and W. den Braver for technical assistance; The International Research Centre for Telecommunication and Radar at the Delft University of Technology for assistance with the stripline simulations. Supported by the Dutch Organization for Fundamental Research on Matter (FOM), the Netherlands Organization for Scientific Research (NWO) and the Defense Advanced Research Projects Agency Quantum Information Science and Technology programme.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to L.M.K.V. (l.m.k.vandersypen@tudelft.nl) and F.H.L.K. (f.h.l.koppens@tudelft.nl).

tinctive features are very similar to those observed in the ~2500-Ma Mt. McRae Shale, and their age is supported by more thorough analytical protocols (*24*). The discovery and careful analysis of biomarkers in rocks of still greater age and of different Archean environments will potentially offer new insights into early microbial life and its evolution.

References and Notes

1. J. W. Schopf, *Science* **260**, 640 (1993).
2. M. R. Walter, in *Earth's Earliest Biosphere*, J. W. Schopf, Ed. (Princeton Univ. Press, Princeton, NJ, 1983), pp. 187–213.
3. S. J. Mojzsis *et al.*, *Nature* **384**, 55 (1996).
4. J. M. Hayes, I. R. Kaplan, K. W. Wedeking, in (*2*), pp. 93–134.
5. F. D. Mango, *Nature* **352**, 146 (1991).
6. A. Dutkiewicz, B. Rasmussen, R. Buick, *ibid.* **395**, 885 (1998).
7. K. E. Peters and J. M. Moldowan, *The Biomarker Guide* (Prentice-Hall, Englewood Cliffs, NJ, 1993).
8. R. C. Morris, *Precambrian Res.* **60**, 243 (1993).
9. A. F. Trendall, D. R. Nelson, J. R. de Laeter, S. W. Hassler, *Aust. J. Earth Sci.* **45**, 137 (1998).
10. N. T. Arndt, D. R. Nelson, W. Compston, A. F. Trendall, A. M. Thorne, *ibid.* **38**, 261 (1991).
11. R. E. Smith, J. L. Perdrix, T. C. Parks, *J. Petrol.* **23**, 75 (1982).
12. J. M. Gressier, thesis, University of Sydney, Sydney, Australia (1996).
13. T. C. Hoering and V. Navale, *Precambrian Res.* **34**, 247 (1987).
14. D. R. Nelson, A. F. Trendall, J. R. de Laeter, N. J. Grobler, I. R. Fletcher, *ibid.* **54**, 231 (1992).
15. R. E. Summons, T. G. Powell, C. J. Boreham, *Geochim. Cosmochim. Acta* **52**, 1747 (1988).
16. R. E. Summons and M. R. Walter, *Am. J. Sci.* **290A**, 212 (1990).
17. T. C. Hoering, *Carnegie Inst. Wash. Yearb.* **64**, 215 (1965); *ibid.* **65**, 365, (1966).
18. G. A. Logan, J. M. Hayes, G. B. Heishima, R. E. Summons, *Nature* **376**, 53 (1995); G. A. Logan, R. E. Summons, J. M. Hayes, *Geochim. Cosmochim. Acta* **61**, 5391 (1997).
19. G. A. Logan *et al.*, *Geochim. Cosmochim. Acta*, **63**, 1345 (1999).
20. S. J. Rowland, *Org. Geochem.* **15**, 9 (1990).
21. J. M. Hayes, in *Early Life on Earth*, Nobel Symposium No. 84, S. Bengtson, Ed. (Columbia Univ. Press, New York, 1994), pp. 220–236.
22. J. W. Schopf and B. M. Packer, *Science* **237**, 70 (1987).
23. R. Buick, *ibid.* **255**, 74 (1992).
24. R. E. Summons, L. L. Jahnke, J. M. Hope, G. A. Logan, *Nature*, in press.
25. H. D. Holland and N. J. Beukes, *Am. J. Sci.* **290A**, 1 (1990); A. H. Knoll and H. D. Holland, in *Effects of Past Global Change on Life*, S. M. Stanley, Ed. (National Academy Press, Washington, DC, 1995), pp. 21–33.
26. P. S. Braterman, A. G. Cairns-Smith, R. W. Sloper, *Nature* **303**, 163 (1983).
27. F. Widdel *et al.*, *ibid.* **362**, 834 (1993).
28. P. Cloud, *Science* **160**, 729 (1968); *Econ. Geol.* **68**, 1135 (1973).
29. G. Ourisson, M. Rohmer, K. Poralla, *Annu. Rev. Microbiol.* **41**, 301 (1987).
30. W. Kohl, A. Gloe, H. Reichenbach, *J. Gen. Microbiol.* **129**, 1629 (1983).
31. T.-M. Han and B. Runnegar, *Science* **257**, 232 (1992); A. H. Knoll, *ibid.* **256**, 622 (1992).

32. Supported by the Studienstiftung des Deutschen Volkes (J.J.B.) and American Chemical Society Petroleum Research Fund (R.B.). We thank J. Gressier and RioTinto Exploration for samples, the AGSO Isotope & Organic Geochemistry staff for technical assistance, J. Kamprad for x-ray diffraction analyses, T. Blake, D. Des Marais, L. L. Jahnke, C. J. Boreham, D. S. Edwards, T. G. Powell, D. E. Canfield, and M. R. Walter for advice, and J. M. Hayes, A. Knoll, and an anonymous reviewer for their thoughtful comments. G.A.L. and R.E.S. publish with the permission of the Executive Director of AGSO.

19 May 1999; accepted 13 July 1999

REPORTS

Josephson Persistent-Current Qubit

J. E. Mooij,[1,2]* T. P. Orlando,[2] L. Levitov,[3] Lin Tian,[3] Caspar H. van der Wal,[1] Seth Lloyd[4]

A qubit was designed that can be fabricated with conventional electron beam lithography and is suited for integration into a large quantum computer. The qubit consists of a micrometer-sized loop with three or four Josephson junctions; the two qubit states have persistent currents of opposite direction. Quantum superpositions of these states are obtained by pulsed microwave modulation of the enclosed magnetic flux by currents in control lines. A superconducting flux transporter allows for controlled transfer between qubits of the flux that is generated by the persistent currents, leading to entanglement of qubit information.

In a quantum computer, information is stored on quantum variables such as spins, photons, or atoms (*1*–*3*). The elementary unit is a two-state quantum system called a qubit. Computations are performed by the creation of quantum superposition states of the qubits and by controlled entanglement of the information on the qubits. Quantum coherence must be conserved to a high degree during these operations. For a quantum computer to be of practical value, the number of qubits must be at least 10^4. Qubits have been implemented in cavity quantum electrodynamics systems (*4*), ion traps (*5*), and nuclear spins of large numbers of identical molecules (*6*). Quantum coherence is high in these systems, but it seems difficult or impossible to realize the desired high number of interacting qubits. Solid state circuits lend themselves to large-scale integration, but the multitude of quantum degrees of freedom leads in general to short decoherence times. Proposals have been put forward for future implementation of qubits with spins of individual donor atoms in silicon (*7*), with spin states in quantum dots (*8*), and with d-wave superconductors (*9*); the technology for practical realization still needs to be developed.

In superconductors, all electrons are condensed in the same macrosopic quantum state, separated by a gap from the many quasi-particle states. This gap is a measure for the strength of the superconducting effects. Superconductors can be weakly coupled with Josephson tunnel junctions (regions where only a thin oxide separates them). The coupling energy is given by $E_J(1 - \cos \gamma)$, where the Josephson energy E_J is proportional to the gap of the superconductors divided by the normal-state tunnel resistance of the junction and γ is the gauge-invariant phase difference of the order parameters. The current through a Josephson junction is equal to $I_o \sin \gamma$, with $I_o = (2e/\hbar) E_J$, where e is the electron charge and $\hbar$ is Planck's constant divided by 2π. In a Josephson junction circuit with small electrical capacitance, the numbers of excess Cooper pairs on islands n_i, n_j and the phase differences γ_i, γ_j are related as noncommuting conjugate quantum variables (*10*). The Heisenberg uncertainty between phase and charge and the occurrence of quantum superpositions of charges as well as phase excitations (vortexlike fluxoids) have been demonstrated in experiments (*11*). Coherent charge oscillations in a superconducting quantum box have recently been observed (*12*). Qubits for quantum computing based on charge states have been suggested (*13*, *14*). However, in actual practice, fabricated Josephson circuits exhibit a high level of static and dynamic charge noise due to charged impurities. In contrast, the magnetic background is clean and stable. Here, we present the design of a qubit with persistent currents of opposite sign as its basic states. The qubits

[1]Department of Applied Physics and Delft Institute for Microelectronics and Submicron Technologies, Delft University of Technology, Post Office Box 5046, 2600 GA Delft, Netherlands. [2]Department of Electrical Engineering and Computer Science, [3]Department of Physics and Center for Materials Science and Engineering, [4]Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

*To whom correspondence should be addressed. E-mail: mooij@qt.tn.tudelft.nl

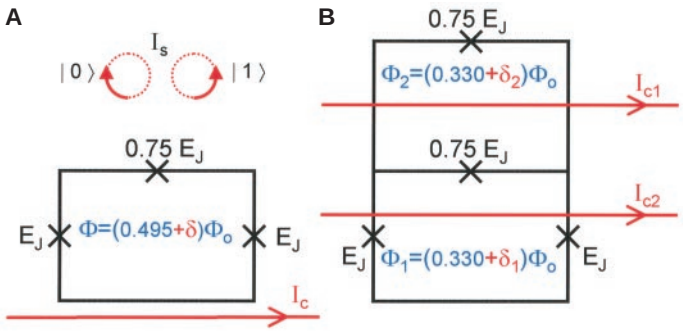


Fig. 1. Persistent current qubit. (**A**) Three-junction qubit. A superconducting loop with three Josephson junctions (indicated with crosses) encloses a flux that is supplied by an external magnet. The flux is $f\Phi_o$, where Φ_o is the superconducting flux quantum and f is 0.495. Two junctions have a Josephson coupling energy E_J, and the third junction has αE_J, where $\alpha = 0.75$. This system has two (meta)stable states I0> and I1> with opposite circulating persistent current. The level splitting is determined by the offset from $\Phi_o/2$ of the flux. The barrier between the states depends on the value of α. The qubit is operated by resonant microwave modulation of the enclosed magnetic flux by a superconducting control line (indicated in red). (**B**) Four-junction qubit. The top junction of (A) is replaced by a parallel junction (SQUID) circuit. There are two loops with equal areas; a magnet supplies a static flux $0.330\Phi_o$ to both. Qubit operations are performed with currents in superconducting control lines (indicated in red) on top of the qubit, separated by a thin insulator. The microwave current I_{c1} couples only to the bottom loop and performs qubit operations as in (A). I_{c2} couples to both loops; it is used for qubit operations with suppressed σ_z action and for an adiabatic increase of the tunnel barrier between qubit states to facilitate the measurement.

can be driven individually by magnetic microwave pulses; measurements can be made with superconducting magnetometers [superconducting quantum interference devices (SQUIDs)]. They are decoupled from charges and electrical signals, and the known sources of decoherence allow for a decoherence time of more than 1 ms. Switching is possible at a rate of 100 MHz. Entanglement is achieved by coupling the flux, which is generated by the persistent current, to a second qubit. The qubits are small (of order 1 μm), can be individually addressed, and can be integrated into large circuits.

Our qubit in principle consists of a loop with three small-capacitance Josephson junctions in series (Fig. 1A) that encloses an applied magnetic flux $f\Phi_o$ (Φ_o is the superconducting flux quantum $h/2e$, where h is Planck's constant); f is slightly smaller than 0.5. Two of the junctions have equal Josephson coupling energy E_J; the coupling in the third junction is αE_J, with $0.5 < \alpha < 1$. Useful values are $f = 0.495$ and $\alpha = 0.75$ (as chosen in Fig. 1A). This system has two stable classical states with persistent circulating currents of opposite sign. For $f = 0.5$, the energies of the two states are the same; the offset from 0.5 determines the level splitting. The barrier for quantum tunneling between the states depends strongly on the value of α. The four-junction version (Fig. 1B) allows modulation of this barrier in situ. Here, the third junction has been converted into a parallel circuit of two junctions, each with a coupling energy αE_J. The four-junction qubit behaves as the three-junction circuit of Fig. 1A, with an enclosed flux $(f_1 + f_2/2)\Phi_o$ and a third-junction (SQUID) strength $2\alpha E_J \cos(f_2\pi)$. The constant fluxes $f\Phi_o$, $f_1\Phi_o$, and $f_2\Phi_o$ are supplied by an external, static, homogeneous magnetic field. Control lines on a separate fabrication level couple inductively to individual qubit loops. All operations on qubits are performed with currents in the control lines.

When γ_1 and γ_2 are the gauge-invariant phase differences across the left and right junctions, the Josephson energy of the four-junction qubit U_J is

$$U_J/E_J = 2 + 2\alpha - \cos\gamma_1 - \cos\gamma_2 - 2\alpha\cos(f_2\pi)\cos(2f_1\pi + f_2\pi + \gamma_1 - \gamma_2) \quad (1)$$

In this expression, the self-generated flux has been neglected. Although this flux will be used for coupling of qubits, it is much smaller than the flux quantum and only slightly changes the picture here. U_J is 2π periodic in γ_1 and γ_2 (Fig. 2A) for the parameter values $\alpha = 0.75$ and $f_1 = f_2 = 0.330$. Each unit cell has two minima L_{ij} and R_{ij} with left- and right-handed circulating currents of about $0.75I_o$ at approximate γ_1,γ_2 values of $\pm 0.27\pi$. The minima would have been symmetric for $2f_1 + f_2 = 1$, which corresponds to a three-junction loop enclosing half a flux quantum. The set of all L minima yields one qubit state and the set of R minima the other. In γ_1,γ_2 space, there are saddle-point connections between L and R minima as indicated with red (intracell, in) and blue lines (intercell, out). Along such trajectories, the system can tunnel between its macroscopic quantum states. The Josephson energy along the trajectories is plotted in Fig. 2B. The saddle-point energies U_{in} and U_{out} depend on α and f_2; lower SQUID coupling gives lower U_{in} but higher U_{out}. For $2\alpha\cos(f_2\pi) < 0.5$, the barrier for intracell tunneling has disappeared, and there is only one minimum with zero circulating current.

Motion of the system in γ_1,γ_2 space can be discussed in analogy with motion of a mass-carrying particle in a landscape with periodic potential energy. Motion in phase space leads to voltages across junctions. The kinetic energy is the associated Coulomb charg-

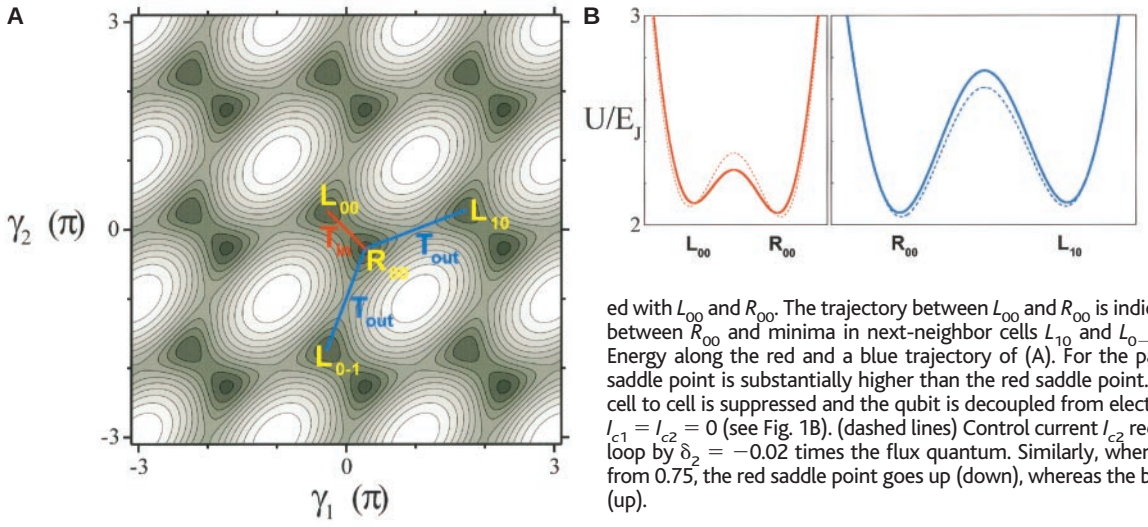


Fig. 2. Josephson energy of qubit in phase space. (**A**) Energy plotted as a function of the gauge-invariant phase differences γ_1 and γ_2 across the left and right junctions of Fig. 1A. The energy is periodic with period 2π. There are two minima in each unit cell, for the center cell indicated with L_{00} and R_{00}. The trajectory between L_{00} and R_{00} is indicated in red; the trajectories between R_{00} and minima in next-neighbor cells L_{10} and L_{0-1} are indicated in blue. (**B**) Energy along the red and a blue trajectory of (A). For the parameters chosen, the blue saddle point is substantially higher than the red saddle point. As a result, tunneling from cell to cell is suppressed and the qubit is decoupled from electrical potentials. (solid lines) $I_{c1} = I_{c2} = 0$ (see Fig. 1B). (dashed lines) Control current I_{c2} reduces the flux in the SQUID loop by $\delta_2 = -0.02$ times the flux quantum. Similarly, when α is increased (decreased) from 0.75, the red saddle point goes up (down), whereas the blue saddle point goes down (up).

ing energy of the junction capacitances. The mass is proportional to the junction capacitance C because other capacitance elements are small. The effective mass tensor has principal values M_a and M_b in the $\gamma_1 - \gamma_2 = 0$ and $\gamma_1 + \gamma_2 = 0$ directions. For the chosen values of the circuit parameters, these principal values are $M_a = \hbar^2/(4E_C)$ and $M_b = \hbar^2/(E_C)$, where the charging energy is defined as $E_C = e^2/2C$. The system will perform plasma oscillations in the potential well with frequencies $\hbar\omega_b \approx 1.3(E_C E_J)^{1/2}$ and $\hbar\omega_a \approx 2.3(E_C E_J)^{1/2}$. The tunneling matrix elements can be estimated by calculation of the action in the Wentzel-Kramers-Brillouin approximation. For tunneling within the unit cell between the minima L and R, the matrix element is $T_{in} \approx \hbar\omega_b \exp[-0.64(E_J/E_C)^{1/2}]$; for tunneling from cell to cell, the matrix element is $T_{out} \approx 1.6\hbar\omega_b \exp[-1.5(E_J/E_C)^{1/2}]$. For the qubit, a subtle balance has to be struck: The plasma frequency must be small enough relative to the barrier height to have well-defined states with a measurable circulating current but large enough (small enough mass) to have substantial tunneling. The preceding qualitative discussion has been confirmed by detailed quantitative calculations in phase space and in charge space (*15*). From these calculations, the best parameters for qubits can be determined. In practice, it is possible to controllably fabricate aluminum tunnel junctions with chosen E_J and E_C values in a useful range.

It is strongly desirable to suppress the intercell tunneling T_{out}. This suppression leads to independence from electrical potentials, even if the charges on the islands are conjugate quantum variables to the phases. The qubit system in phase space is then comparable to a crystal in real space with nonoverlapping atomic wave functions. In such a crystal, the electronic wave functions are independent of momentum; similarly, charge has no influence in our qubit.

Mesoscopic aluminum junctions can be reliably fabricated by shadow evaporation with critical current densities up to 500 A/cm^2. In practice, a junction of 100 nm^2 by 100 nm^2 has E_J around 25 GHz and E_C around 20 GHz. A higher E_J/E_C ratio can be obtained by increasing the area to which E_J is proportional and E_C is inversely proportional. A practical qubit would, for example, have junctions with an area of 200 nm^2 by 400 nm^2, $E_J \sim 200$ GHz, $E_J/E_c \sim 80$, level splitting $\Delta E \sim 10$ GHz, barrier height around 35 GHz, plasma frequency around 25 GHz, and tunneling matrix element $T_{in} \sim 1$ GHz. The matrix element for undesired tunneling T_{out} is smaller than 1 MHz. The qubit size would be of order 1 μm; with an estimated inductance of 5 pH, the flux generated by the persistent currents is about $10^{-3}\Phi_o$.

To calculate the dependence of the level splitting on f_1 and f_2, we apply a linearized approximation in the vicinity of $f_1 = f_2 = 1/3$, defining F as the change of U_J away from the minimum of $U_J(\gamma_1,\gamma_2)$. This yields $F/E_J = 1.2[2(f_1 - 1/3)+(f_2 - 1/3)]$. The level splitting without tunneling would be $2F$. With tunneling, symmetric and antisymmetric combinations are created; the level splitting is now $\Delta E = 2(F^2 + T_{in}^2)^{1/2}$. As long as $F \gg T_{in}$, the newly formed eigenstates are localized in the minima of $U_J(\gamma_1,\gamma_2)$.

We discuss qubit operations for the four-junction qubit. They are driven by the currents I_{ca} and I_{cb} in the two control lines (Fig. 1B). The fluxes induced in the two loops, normalized to the flux quantum, are $\delta_1 = (L_{a1}I_{ca} + L_{b1}I_{cb})/\Phi_o$ and $\delta_2 = (L_{a2}I_{ca} + L_{b2}I_{cb})/\Phi_o$. The control line positions are chosen such that $L_{a2} = 0$ and $L_{b2} = -2L_{b1}$. When the two loops have equal areas, $f_1 = f_2$ for zero control current. We assume that the qubit states are defined with zero control current and that δ_1 and δ_2 act as perturbations to this system. The effective Hamiltonian operator (H_{op}) in terms of Pauli spin matrices σ_x and σ_z for the chosen parameters is about

$$H_{op}/\Delta E \approx (80\delta_1 + 42\delta_2)\sigma_z - (9.2\delta_1 + 8.3\delta_2)\sigma_x \quad (2)$$

The numerical prefactors follow from the variational analysis of the influence of δ_1 and δ_2 on the tunnel barrier and the level splitting.

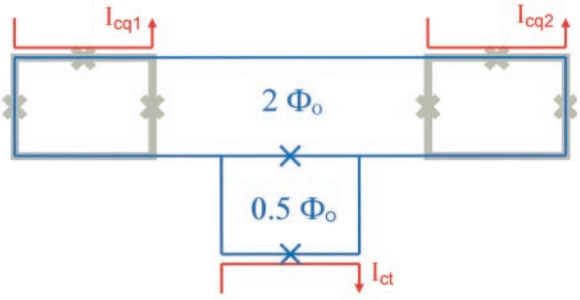


Fig. 3. Switchable qubit coupler. A superconducting flux transporter (blue) is placed on top of two qubits, separated by a thin insulator. The transporter is a closed loop that contains two Josephson junctions in parallel (SQUID) with high critical current. In the off state, the two loops of the transporter contain an integer number of flux quanta (main loop) and half a flux quantum (SQUID loop), supplied by a permanent magnet. The current response to a flux change is very small. In the on state, the flux in the SQUID loop is made integer by means of a control current I_{ct} (red). As the transporter attempts to keep the flux in its loop constant, a flux change induced by qubit 1 is transmitted to qubit 2. As shown here, the two three-junction qubits experience $\sigma_z \otimes \sigma_z$-type coupling. The flux values have to be adjusted for the influence of circulating currents.

The terms that contain σ_x can be used to induce Rabi oscillations between the two states, applying microwave pulses of frequency $\Delta E/h$. There are two main options, connected to one of the two control lines. Control current I_{ca} changes δ_1, which leads to a Rabi oscillation (σ_x term) as well as a strong modulation of the Larmor precession (σ_z term). As long as the Rabi frequency is far enough below the Larmor frequency, this is no problem. For $\delta_1 = 0.001$, the Rabi frequency is 100 MHz. This mode is the only one available for the three-junction qubit and is most effective near the symmetry point $f = 0.5$ or $f_1 = f_2 = 1/3$. Control current I_{cb} is used to modulate the tunnel barrier. Here, the σ_z action is suppressed by means of the choice $L_{b2}/L_{b1} = \delta_2/\delta_1 = -2$. However, a detailed analysis shows that with δ_2 modulation, it is easy to excite the plasma oscillation with frequency ω_b. One has to restrict δ_2 to remain within the two-level system. Values of 0.001 for δ_1 or δ_2 correspond to about 50-pW microwave power at 10 GHz in the control line. These numbers are well within practical range.

Two or more qubits can be coupled by means of the flux that the circulating persistent current generates. The current is about 0.3 μA, the self-inductance of the loop is about 5 pH, and the generated flux is about $10^{-3}\Phi_o$. When a superconducting closed loop (a flux transporter) with high critical current is placed on top of both qubits, the total enclosed flux is constant. A flux change $\Delta\Phi$ that is induced by a reversal of the current in one qubit leads to a change of about $\Delta\Phi/2$ in the flux that is enclosed by the other qubit. One can choose to couple the flux, generated in the main loop of qubit 1, to the main loop of qubit 2 ($\sigma_z \otimes \sigma_z$ coupling) or to the SQUID loop of qubit 2 ($\sigma_z \otimes \sigma_x$ coupling). A two-qubit gate operation is about as efficient as a single qubit operation driven with $\delta_1 = 0.001$. An example of a possible controlled-NOT operation with fixed coupling runs as follows: The level splitting of qubit 2 depends on the state of qubit 1, the values are ΔE_{20} and ΔE_{21}. When Rabi microwave pulses, resonant with ΔE_{21}, are applied to qubit 2, it will only react if qubit 1 is in its I1> state. In principle, qubits can be coupled at larger distances. An array scheme as proposed by Lloyd (*1*, *3*), where only nearest neighbor qubits are coupled, is also very feasible. It is possible to create a flux transporter that has to be switched on by a control current (Fig. 3).

The typical switching times for our qubit are 10 to 100 ns. To yield a practical quantum computer, the decoherence time should be at least 100 μs. We can estimate the influence of known sources of decoherence for our system, but it is impossible to determine the real decoherence time with certainty, except by measurement. We discuss

some decohering influences here. All quasiparticle states in the superconductor have to remain unoccupied. In equilibrium, the number is far below 1 at temperatures below 30 mK. Extreme care must be taken to shield the sample from photons. Even 4 K blackbody photons have enough energy to break a Cooper pair. Adequate shielding is possible on the time scale of our computer. Inductive coupling to bodies of normal metal has to be avoided. By decoupling the qubit from electrical potentials, we have eliminated coupling to charged defects in substrate or tunnel barriers. The aluminum nuclei have a spin that is not polarized by the small magnetic fields at our temperature of 25 mK. Statistical fluctuations will occur, but their time constant is very long because of the absence of electronic quasi-particles. The net effect will be a small static offset of the level splitting, within the scale of the variations due to fabrication. The dephasing time that results from unintended dipole-dipole coupling of qubits is longer than 1 ms if the qubits are farther apart than 1 μm. Emission of photons is negligible for the small loop. Overall, the sources of decoherence that we know allow for a decoherence time above 1 ms.

Requirements for a quantum computer are that the qubits can be prepared in well-defined states before the start of the computation and that their states can be measured at the end. Initialization will proceed by cooling the computer to below 50 mK and having the qubits settle in the ground state. For the measurement, a generated flux of $10^{-3}\Phi_o$ in an individual qubit can be detected with a SQUID if enough measuring time is available. A good SQUID has a sensitivity of $10^{-5}\Phi_o/\mathrm{Hz}^{1/2}$, so that a time of 100 μs is required. Usual SQUIDs have junctions that are shunted with normal metal. The shunt introduces severe decoherence in a qubit when the SQUID is in place, even if no measurement is performed. We are developing a nonshunted SQUID that detects its critical current by discontinuous switching. For a measurement at the end of a quantum computation scheme, the qubit can be frozen by an adiabatic increase of the tunnel barrier between the two qubit states. As Fig. 2 indicates, we can increase the barrier by a change of control current. A similar procedure, as suggested by Shnirman and Schön (*14*), for charge qubits can be followed.

The proposed qubit should be of considerable interest for fundamental studies of macroscopic quantum coherence, apart from its quantum computing potential. Compared with the radio frequency SQUID systems that have been used in attempts to observe such effects (*16*) and also have been suggested as possible qubits for quantum computation (*17*), the much smaller size of the qubit decouples it substantially better from the environment.

References and Notes

1. S. Lloyd, *Science* **261**, 1569 (1993).
2. C. H. Bennett, *Phys. Today* **48** (no. 10), 24 (1995); D. P. DiVincenzo, *Science* **270**, 255 (1995); T. P. Spiller, *Proc. IEEE* **84**, 1719 (1996).
3. S. Lloyd, *Sci. Am.* **273** (no. 4), 140 (1995).
4. Q. A. Turchette, C. J. Hood, W. Lange, H. Mabuchi, H. J. Kimble, *Phys. Rev. Lett.* **75**, 4710 (1995).
5. C. Monroe, D. M. Meekhof, B. E. King, W. M. Itano, D. J. Wineland, *ibid.*, p. 4714.
6. N. A. Gershenfeld and I. L. Chuang, *Science* **275**, 350 (1997).
7. B. Kane, *Nature* **393**, 133 (1998).
8. D. Loss and D. DiVincenzo, *Phys. Rev. A* **57**, 120 (1998).
9. L. B. Ioffe, V. B. Geshkenbein, M. V. Feigel'man, A. L. Fauchère, G. Blatter, *Nature* **398**, 679 (1999).
10. D. V. Averin and K. K. Likharev, in *Mesoscopic Phenomena in Solids*, B. L. Altshuler, P. A. Lee, R. A. Webb, Eds. (North Holland, Amsterdam, 1991), pp. 173–271.
11. W. J. Elion, M. Matters, U. Geigenmuller, J. E. Mooij, *Nature* **371**, 594 (1994); L. S. Kuzmin and D. B. Haviland, *Phys. Rev. Lett.* **67**, 2890 (1991); P. Joyez, D. Esteve, M. H. Devoret, *ibid.* **80**, 1956 (1998).
12. Y. Nakamura, Yu. A. Pashkin, J. S. Tsai, *Nature* **398**, 786 (1999).
13. A. Shnirman, G. Schön, Z. Hermon, *Phys. Rev. Lett.* **79**, 2371 (1997); D. V. Averin, *Solid State Commun.* **105**, 659 (1998); Yu. Makhlin, G. Schön, A. Shnirman, *Nature* **398**, 305 (1999).
14. A. Shnirman and G. Schön, *Phys. Rev. B* **57**, 15400 (1998).
15. T. P. Orlando *et al.*, in preparation.
16. C. D. Tesche, *Phys. Rev. Lett.* **64**, 2358 (1990); R. Rouse, S. Han, J. E. Lukens, *ibid.* **75**, 1614 (1995).
17. M. F. Bocko, A. M. Herr, M. J. Feldman, *IEEE Trans. Appl. Supercond.* **7**, 3638 (1997).

18. We thank J. J. Mazo, C. J. P. M. Harmans, A. C. Wallast, and H. Tanaka for important discussions. This work is partially supported by Army Research Office grant DAAG55-98-1-0369, Stichting voor Fundamenteel Onderzoek der Materie, NSF Award 67436000IRG, and the New Energy and Industrial Technology Development Organization.

22 April 1999; accepted 7 July 1999

Energetic Iron(VI) Chemistry: The Super-Iron Battery

Stuart Licht,* Baohui Wang, Susanta Ghosh

Higher capacity batteries based on an unusual stabilized iron(VI) chemistry are presented. The storage capacities of alkaline and metal hydride batteries are largely cathode limited, and both use a potassium hydroxide electrolyte. The new batteries are compatible with the alkaline and metal hydride battery anodes but have higher cathode capacity and are based on available, benign materials. Iron(VI/III) cathodes can use low-solubility K_2FeO_4 and $BaFeO_4$ salts with respective capacities of 406 and 313 milliampere-hours per gram. Super-iron batteries have a 50 percent energy advantage compared to conventional alkaline batteries. A cell with an iron(VI) cathode and a metal hydride anode is significantly (75 percent) rechargeable.

Improved batteries are needed for various applications such as consumer electronics, communications devices, medical implants, and transportation needs. The search for higher capacity electrochemical storage has focused on a wide range of materials, such as carbonaceous materials (*1*), tin oxide (*2*), grouped electrocatalysts (*3*), or macroporous minerals (*4*). Of growing importance are rechargeable (secondary) batteries such as metal hydride (MH) batteries (*5*), which this year have increased the commercial electric car range to 250 km per charge. In consumer electronics, primary, rather than secondary, batteries dominate. Capacity, power, cost, and safety factors have led to the annual global use of approximately 6×10^{10} alkaline or dry batteries, which use electrochemical storage based on a Zn anode, an aqueous electrolyte, and a MnO_2 cathode, and which constitute the vast majority of consumer batteries. Despite the need for safe, inexpensive, higher capacity electrical storage, the aqueous MnO_2/Zn battery has been a dominant primary battery chemistry for over a century. Contemporary alkaline and MH batteries have two common features: Their storage capacity is largely cathode limited and both use a KOH electrolyte.

We report a new class of batteries, referred to as super-iron batteries, which contain a cathode that uses a common material (Fe) but in an unusual (greater than 3) valence state. Although they contain the same Zn anode and electrolyte as conventional alkaline batteries, the super-iron batteries provide >50% more energy capacity. In addition, the Fe(VI) chemistry is rechargeable, is based on abundant starting materials, has a relatively environmentally benign discharge product, and appears to be compatible with the anode of either the primary alkaline or secondary MH batteries.

The fundamentals of MnO_2 chemistry continue to be of widespread interest (*6*). The storage capacity of the aqueous MnO_2/Zn

Department of Chemistry and Institute of Catalysis Science, Technion—Israel Institute of Technology, Haifa 32000, Israel.

*To whom correspondence should be addressed. E-mail: chrlicht@techunix.technion.ac.il

18. The International HapMap Consortium, *Nature* **426**, 789 (2003).
19. A. R. Templeton, E. Boerwinkle, C. F. Sing, *Genetics* **117**, 343 (1987).
20. D. W. Schultz *et al.*, *Hum. Mol. Genet.* **12**, 3315 (2003).
21. M. Hayashi *et al.*, *Ophthalmic Genet.* **25**, 111 (2004).
22. G. J. McKay *et al.*, *Mol. Vis.* **10**, 682 (2004).
23. S. Rodríguez de Cordoba, J. Esparza-Gordillo, E. Goicoechea de Jorge, M. Lopez-Trascasa, P. Sanchez-Corral, *Mol. Immunol.* **41**, 355 (2004).
24. L. V. Johnson, W. P. Leitner, M. K. Staples, D. H. Anderson, *Exp. Eye Res.* **73**, 887 (2001).
25. R. F. Mullins, S. R. Russell, D. H. Anderson, G. S. Hageman, *FASEB J.* **14**, 835 (2000).
26. J. Ambati *et al.*, *Nat. Med.* **9**, 1390 (2003).
27. G. S. Hageman *et al.*, *Prog. Retinal Eye Res.* **20**, 705 (2001).
28. J. Esparza-Gordillo *et al.*, *Immunogenetics* **56**, 77 (2004).
29. G. Wistow *et al.*, *Mol. Vis.* **8**, 205 (2002).
30. R. F. Mullins, N. Aptsiauri, G. S. Hageman, *Eye* **15**, 390 (2001).
31. A. M. Blom, L. Kask, B. Ramesh, A. Hillarp, *Arch. Biochem. Biophys.* **418**, 108 (2003).
32. J. M. Seddon, G. Gensler, R. C. Milton, M. L. Klein, N. Rifai, *JAMA* **291**, 704 (2004).
33. The Raymond and Beverly Sackler Fund for Arts and Sciences' generous support made this project possible. We thank Raymond Sackler, J. Sackler, and E. Vosburg for their input and encouragement. We also thank AREDS participants and investigators; G. Gensler, T. Clemons, and A. Lindblad for work on the AREDS Genetic Repository; S. Westman and A. Evan for assistance with the microarrays; R. Fariss for the human retinal sections and advice on confocal microscopy; E. Johnson for assistance with immunostaining; and J. Majewski for constructive comments on the manuscript. Partially funded by NIH-K25HG000060 and NIH-R01EY015771 (J.H.), Macula Vision Research Foundation and the David Woods Kemper Memorial Foundation (C.B.), NIH-R01MH44292 (J.O.), and NIH-K01RR16090 and Yale Pepper Center for Study of Diseases in Aging (C.Z.). This work also benefited from the International HapMap Consortium making their data available prior to publication.

Supporting Online Material
www.sciencemag.org/cgi/content/full/1109557/DC1
Materials and Methods
Fig. S1
Tables S1 to S5
References

10 January 2005; accepted 22 February 2005
Published online 10 March 2005;
10.1126/science.1109557
Include this information when citing this paper.

REPORTS

Quantum Phase Transition of a Magnet in a Spin Bath

H. M. Rønnow,[1,2,3]* R. Parthasarathy,[2] J. Jensen,[4] G. Aeppli,[5] T. F. Rosenbaum,[2] D. F. McMorrow[3,4,6]

The excitation spectrum of a model magnetic system, $LiHoF_4$, was studied with the use of neutron spectroscopy as the system was tuned to its quantum critical point by an applied magnetic field. The electronic mode softening expected for a quantum phase transition was forestalled by hyperfine coupling to the nuclear spins. We found that interactions with the nuclear spin bath controlled the length scale over which the excitations could be entangled. This generic result places a limit on our ability to observe intrinsic electronic quantum criticality.

The preparation and preservation of entangled quantum states is particularly relevant for the development of quantum computers, where interacting quantum bits (qubits) must produce states sufficiently long lived for meaningful manipulation. The state lifetime, typically referred to as decoherence time, is derived from coupling to the background environment. For solid-state quantum computing schemes, the qubits are typically electron spins, and they couple to two generic background environments (*1*). The oscillator bath—that is, delocalized environmental modes (*2*) such as thermal vibrations coupled via magnetoelastic terms to the spins—can be escaped by lowering the temperature to a point where the lattice is essentially frozen. Coupling to local degrees of freedom, such as nuclear magnetic moments that form a spin bath, may prove more difficult to avoid, because all spin-based candidate materials for quantum computation have at least one naturally occurring isotope that carries nuclear spin.

Experimental work in this area has been largely restricted to the relaxation of single, weakly interacting magnetic moments such as those on large molecules (*3*); much less is known about spins as they might interact in a real quantum computer. In this regard, the insight that quantum phase transitions (QPTs) (*4*) are a good arena for looking at fundamental quantum properties of strongly interacting spins turns out to be valuable, as it has already been for explorations of entanglement. In particular, we show that coupling to a nuclear spin bath limits the distance over which quantum mechanical mixing affects the electron spin dynamics.

QPTs are transitions between different ground states driven not by thermal fluctuations but by quantum fluctuations controlled by a parameter such as doping, pressure, or magnetic field (*5*, *6*). Much of the interest in QPTs stems from their importance for understanding materials with unconventional properties, such as heavy fermion systems and high-temperature superconductors. However, these materials are rather complex and do not easily lend themselves to a universal understanding of QPTs. To this end, it is desirable to identify quantum critical systems with a well-defined and solvable Hamiltonian and with a precisely controllable tuning parameter. One very simple model displaying a QPT is the Ising ferromagnet in a transverse magnetic field (*5*, *7*–*9*) with the Hamiltonian

$$\mathcal{H} = -\sum_{ij} \mathcal{J}_{ij}\, \sigma_i^z \cdot \sigma_j^z - \Gamma \sum_i \sigma_i^x \qquad (1)$$

where $\mathcal{J}_{ij}$ is the coupling between the spins on sites i and j represented by the Pauli matrices σ^z with eigenvalues ±1. In the absence of a magnetic field, the system orders ferromagnetically below a critical temperature T_c. The transverse-field Γ mixes the two states and leads to destruction of long-range order in a QPT at a critical field Γ_c, even at zero temperature. In the ferromagnetic state at zero field and temperature, the excitation spectrum is momentum independent and is centered at the energy $4\sum_j J_{ij}$ associated with single-spin reversal. Upon application of a magnetic field, however, the excitations acquire a dispersion, softening to zero at the zone center $q = 0$ when the QPT is reached.

We investigated the excitation spectrum around the QPT in $LiHoF_4$, which is an excellent physical realization of the transverse-field Ising model, with an added term accounting for the hyperfine coupling between electronic and nuclear moments (*10*–*12*). The dilution series $LiHo_xY_{1-x}F_4$ is the host for a wide variety of collective quantum effects, ranging from tunneling of single moments and domain walls to quantum annealing, entanglement, and Rabi oscillations (*13*–*17*). These intriguing properties rely largely on the ability of a transverse field, whether applied externally or generated internally by the off-diagonal part of the magnetic dipolar interaction, to mix two degenerate crystal field states of each Ho ion.

[1]Laboratory for Neutron Scattering, ETH-Zürich and Paul Scherrer Institut, 5232 Villigen, Switzerland. [2]James Franck Institute and Department of Physics, University of Chicago, Chicago, IL 60637, USA. [3]Risø National Laboratory, DK-4000 Roskilde, Denmark. [4]Ørsted Laboratory, Niels Bohr Institute fAPG, Universitetsparken 5, 2100 Copenhagen, Denmark. [5]London Centre for Nanotechnology and Department of Physics and Astronomy, University College London, London WC1E 6BT, UK. [6]ISIS, Rutherford Appleton Laboratory, Chilton, Didcot OX11 0QX, UK.

*To whom correspondence should be addressed. E-mail: henrik.ronnow@psi.ch

The Ho ions in $LiHoF_4$ are placed on a tetragonal Scheelite lattice with parameters $a = 5.175$ Å and $c = 10.75$ Å. The crystal-field ground state is a $\Gamma_{3,4}$ doublet with only a c component to the angular momentum and hence can be represented by the $\sigma^z = \pm 1$ Ising states. A transverse field in the a-b plane mixes the higher lying states with the ground state; this produces a splitting of the doublet, equivalent to an effective Ising model field. The phase diagram of $LiHoF_4$ (Fig. 1A) was determined earlier by susceptibility measurements (*10*) and displays a zero-field T_c of 1.53 K and a critical field of $H_c = 49.5$ kOe in the zero temperature limit. The same measurements confirmed the strong Ising anisotropy, with longitudinal and transverse g factors differing by a factor of 18 (*10*). The sudden increase in H_c below 400 mK was explained by alignment of the Ho nuclear moments through the hyperfine coupling. Corrections to phase diagrams as a result of hyperfine couplings have a long history (*18*) and were noted for the Li*RE*F_4 (*RE* = rare earth) series, of which $LiHoF_4$ is a member, more than 20 years ago (*19*). What is new here is that the application of a transverse field and the use of high-resolution neutron scattering spectroscopy allow us to carefully study the dynamics as we tune through the quantum critical point (QCP).

We measured the magnetic excitation spectrum of $LiHoF_4$ with the use of the TAS7 neutron spectrometer at Risø National Laboratory, with an energy resolution (full width at half maximum) of 0.06 to 0.18 meV (*20*). The transverse field was aligned to better than 0.35°, and the sample was cooled in a dilution refrigerator. At the base temperature of 0.31 K, giving a critical field of 42.4 kOe, the excitation spectrum was mapped out below, at, and above the critical field (Fig. 2). For all fields, a single excitation branch disperses upward from a minimum gap at (2,0,0) toward (1,0,0). From (1,0,0) to (1,0,1), the mode shows little dispersion but appears to broaden. The discontinuity on approaching $(1,0,1-\varepsilon)$ and $(1+\varepsilon,0,1)$ as $\varepsilon \rightarrow 0$ reflects the anisotropy and long-range nature of the magnetic dipole coupling. However, the most important observation is that the (2,0,0) energy, which is always lower than the calculated single-ion energy (~0.39 meV at 42.4 kOe), shrinks upon increasing the field from 36 to 42.4 kOe and then hardens again at 60 kOe. At this qualitative level, what we see agrees with the mode softening predicted for the simple Ising model in a transverse field. However, it appears that the mode softening is incomplete. At the critical field of 42.4 kOe, the mode retains a finite energy of 0.24 ± 0.01 meV. This result is apparent in Fig. 1B, which shows the gap energy as a function of the external field.

To obtain a quantitative understanding of our experiments, we consider the full rare-earth Hamiltonian, which closely resembles that of

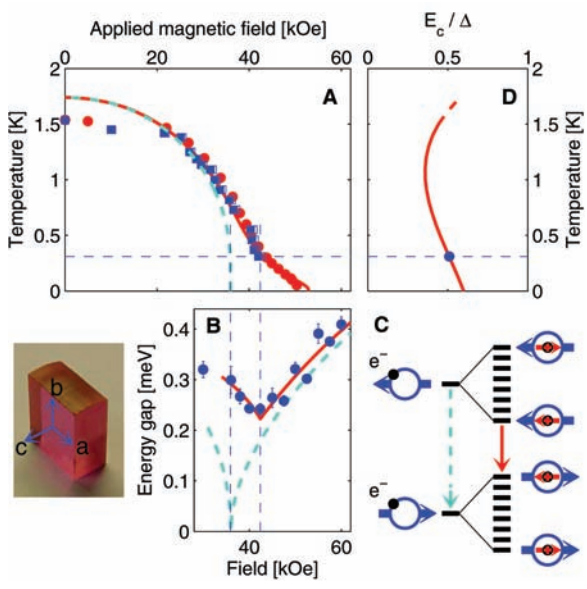


Fig. 1. (**A**) Phase diagram of $LiHoF_4$ as a function of transverse magnetic field and temperature from susceptibility (*10*) (circles) and neutron scattering (squares) measurements. Lines are 1/z calculations with (solid) and without (dashed) hyperfine interaction. Horizontal dashed guide marks the temperature 0.31 K at which inelastic neutron measurements were performed. (**B**) Field dependence of the lowest excitation energy in $LiHoF_4$ measured at $Q = (1+\varepsilon,0,1)$. Lines are calculated energies scaled by $Z = 1.15$ with (solid) and without (dashed) hyperfine coupling. The dashed vertical guides show how in either case the minimum energy occurs at the field of the transition [compare with (A)]. (**C**) Schematic of electronic (blue) and nuclear (red) levels as the transverse field is lowered toward the QCP. Neglecting the nuclear spins, the electronic transition (light blue arrow) would soften all the way to zero energy. Hyperfine coupling creates a nondegenerate multiplet around each electronic state. The QCP now occurs when the excited-state multiplet through level repulsion squeezes the collective mode of the ground-state multiplet to zero energy, hence forestalling complete softening of the electronic mode. Of course, the true ground and excited states are collective modes of many Ho ions and should be classified in momentum space. (**D**) Calculated ratio of the minimum excitation energy E_c to the single-ion splitting Δ at the critical field as a function of temperature. This measures how far the electronic system is from the coherent limit, for which $E_c/\Delta = 0$.

HoF_3 (*21*, *22*). Each Ho ion is subject to the crystal field, the Zeeman coupling, and the hyperfine coupling. The interaction between moments is dominated by the long-range dipole coupling, with a small nearest neighbor exchange interaction J_{12}:

$$\mathcal{H} = \sum_i [\mathcal{H}_{CF}(\mathbf{J}_i) + A\mathbf{J}_i \cdot \mathbf{I}_i - g\mu_B \mathbf{J}_i \cdot \mathbf{H}] - \frac{1}{2}\sum_{ij}\sum_{\alpha\beta} \mathcal{J}_D D_{\alpha\beta}(ij) J_{i\alpha} J_{j\beta} - \frac{1}{2}\sum_{ij}^{n.n.} \mathcal{J}_{12}\, \mathbf{J}_i \cdot \mathbf{J}_j \qquad (2)$$

where $\mathbf{J}$ and $\mathbf{I}$ are the electronic and nuclear moments, respectively, and for $^{165}Ho^{3+}$ $J = 8$ and $I = 7/2$. Hyperfine resonance (*23*) and heat capacity measurements (*24*) show the hyperfine coupling parameter $A = 3.36$ μeV as for the isolated ion, with negligible nuclear-quadrupole coupling. The Zeeman term is reduced by the demagnetization field. The normalized dipole tensor $D_{\alpha\beta}(ij)$ is directly calculable, and the dipole coupling strength J_D is simply fixed by lattice constants and the magnetic moments of the ions at $J_D = (g\mu_B)^2 N = 1.1654$ μeV, where μ_B is the Bohr magneton. This leaves as free parameters various numbers appearing in the crystal-field Hamiltonian H_{CF} and the exchange constant J_{12}. The former are determined (*25*) largely from electron spin resonance for dilute Ho atoms substituted for Y in $LiYF_4$, whereas the latter is constrained by the phase diagram determined earlier (*10*) (Fig. 1A). We have used an effective medium theory (*9*) previously applied to HoF_3 (*26*) to fit the phase diagram, and we conclude that a good overall description—except for a modest (14%) overestimate of the zero-field transition temperature—is obtained for $J_{12} = -0.1$ μeV. On the basis of quantum Monte Carlo simulation data, others (*27*) have also concluded that J_{12} is substantially smaller than J_D.

Having established a good parameterization of the Hamiltonian, we model the dynamics, where expansion to order 1/z (where z is the number of nearest neighbors of an ion in the lattice) leads to an energy-dependent renormalization $[1 + \Sigma(\omega)]^{-1}$ (on the order of 10%) of the dynamic susceptibility calculated in the random phase approximation, with the self energy $\Sigma(\omega)$ evaluated as described in (*26*). For the three fields investigated in detail, the dispersion measured by neutron scattering is closely reproduced throughout the Brillouin zone. As indicated by the solid lines in Fig. 2, the agreement becomes excellent if the calculated excitation energies are multiplied by a renormalization factor $Z = 1.15$. The point is not that the calculation is imperfect but rather that it matches the data as closely as it does. Indeed, it also predicts a weak mode splitting of about 0.08 meV at $(1,0,1-\varepsilon)$, consistent with the increased width in the measurements. The agreement for the discontinuous jump between $(1,0,1-\varepsilon)$ and $(1+\varepsilon,0,1)$ as a result of the long-range nature of the dipole coupling shows that this is indeed the dominant coupling.

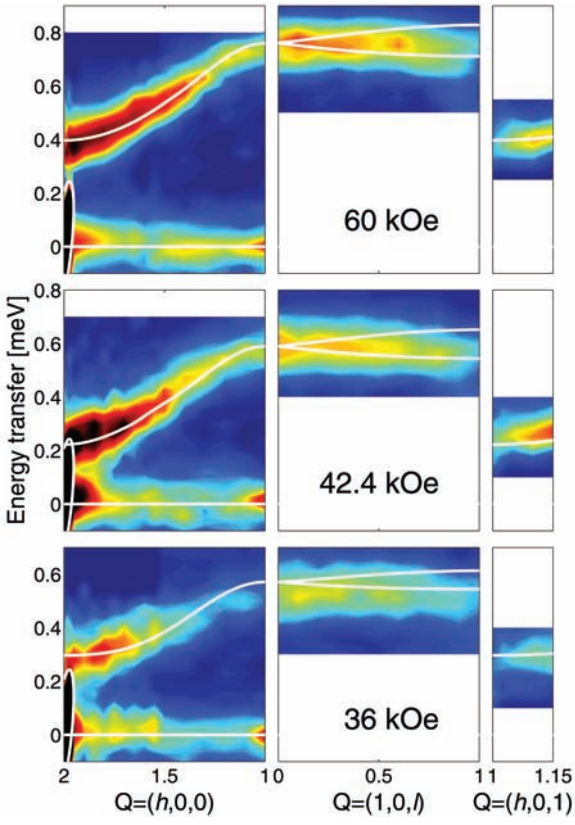


Fig. 2. Pseudocolor representation of the inelastic neutron scattering intensity for $LiHoF_4$ at T = 0.31 K observed along the reciprocal space trace (2,0,0) → (1,0,0) → (1,0,1) → (1.15,0,1). White lines show the 1/z calculation for the excitation energies as described in the text. White ellipses around the (2,0,0) Bragg peak indicate 5 times the resolution tail (full width at half maximum).

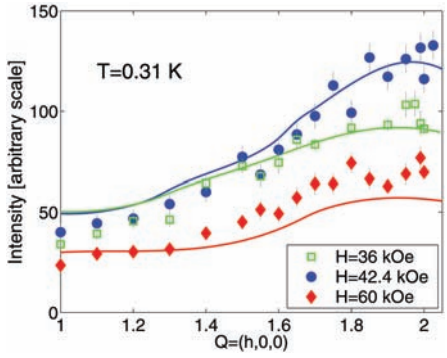


Fig. 3. Measured intensities of the excitations along Q = (h,0,0) at the same values of the field as in Fig. 2. Lines are calculated with geometric and resolution corrections applied to allow comparison to the neutron data.

The simple origin of the incomplete softening and enhanced critical field (Fig. 1, B and C) is easiest to understand if we start from the polarized paramagnetic state above H_c, where the experiment, the purely electronic calculation, and the theory including the hyperfine coupling all coincide. At high fields, the only effect of the hyperfine term is to split both the ground state and the electronic excitation modes into multiplets that are simply the direct products of the electronic and nuclear levels, with a total span of $2A\langle J\rangle I \approx 0.1$ meV (Fig. 1C). Upon lowering the field, the electronic mode softens and would reach zero energy at H_c^0 = 36 kOe in the absence of hyperfine coupling. The hyperfine coupling, however, already mixes the original ground and excited (soft mode) states above H_c. As this happens, the formation of a composite spin from mixed nuclear and electronic contributions immediately stabilizes ordering along the c axis of the crystal. In other words, the hyperfine coupling shunts the electronic mode, raising the critical field to the observed H_c = 42.4 kOe, where the mode reaches a nonzero minimum. This process is accompanied by transfer of intensity from the magnetic excitation of electronic origin to soft modes of much lower energy (in the 10-μeV range) that have an entangled nuclear/electronic character. Cooling to very low temperatures would reveal these modes as propagating and softening to zero at the QCP, but at the temperatures reachable in our measurements there is thermalization, dephasing the composite modes to yield the strong quasi-elastic scattering appearing around Q = (2,0,0) and zero energy at the critical field, as in Fig. 2.

The intensities of the excitations are simply proportional to the matrix elements $|\langle f|\sum_j \exp(iQ \cdot R_j)J_j^+|0\rangle|^2$, and therefore provide a direct measure of the wave functions via the interference effects implicit in the spatial Fourier transform of J_j. Figure 3 shows intensities recorded along (h,0,0) for the three fields 36, 42.4, and 60 kOe. They follow a momentum dependence characterized by a broad peak near (2,0,0), which is well described by our theory. In the absence of hyperfine interactions, the intensity at H_c^0 would diverge as q approaches (2,0,0), reflecting that the real-space dynamical coherence length ξ_c of the excited state grows to infinity. The finite width of the peak observed at H_c corresponds in real space to a distance on the order of the interholmium spacing; because the hyperfine interactions forestall the softening of the electronic mode, the implication is that these interactions also limit the distance over which the electronic wave functions can be entangled (*4*). Thus, Fig. 3 is a direct demonstration of the limitation of quantum coherence in space via coupling to a nuclear spin bath. ξ_c is obtained from a sum over matrix elements connecting the ground state to a particular set of excited states, whereas the thermodynamic correlation length ξ_t is derived from the equal time correlation function $S(r)$, which is the sum over all final states. ξ_t diverges at second-order transitions such as those in $LiHoF_4$, where the quasielastic component seen in our data dominates the long-distance behavior of $S(r)$ at $T_c(H)$. It is the electronic mode, and hence ξ_c, that dictates to what extent $LiHoF_4$ can be characterized and potentially exploited as a realization of the ideal transverse-field Ising model.

Beyond providing a quantitative understanding of the excitations near the QCP of a model experimental system, we obtain new insight by bringing together the older knowledge from rare-earth magnetism and the contemporary ideas of entanglement, qubits, and decoherence. Although the notion of the spin bath was developed to address decoherence in localized magnetic clusters and molecules (*1*), our work discloses its importance for QPTs. In particular, we establish that the spin bath is a generic feature that will limit our ability to observe intrinsic electronic quantum criticality. This may not matter much for transition metal oxides with very large exchange constants, but it could matter for rare earth and actinide intermetallic compounds, which show currently unexplained crossovers to novel behaviors at low (<1 K) temperatures [see, e.g., (*28*)].

For magnetic clusters, decoherence can be minimized in a window between the oscillator bath–dominated high-temperature regions and the spin bath–dominated low-temperature regions (*29*). Our calculations suggest that the dense quantum critical magnet shows analogous behavior. Here the interacting electron spins themselves constitute the oscillator bath, and the extent to which the magnetic excitation softens at $T_c(H)$, as measured by the ratio of the zone center energy E_c to the field-induced single-ion splitting Δ (Fig. 1D), gauges the electronic decoherence. E_c/Δ achieves its minimum not at T = 0 but rather at an intermediate temperature $T \approx 1$ K, exactly where the phase boundary in Fig. 1A begins to be affected by the nuclear hyperfine interactions.

References and Notes

1. N. V. Prokof'ev, P. C. E. Stamp, *Rep. Prog. Phys.* **63**, 669 (2000).
2. R. P. Feynman, F. L. Vernon, *Ann. Phys.* **24**, 118 (1963).
3. W. Wernsdorfer, S. Bhaduri, R. Tiron, D. N. Hendrickson, G. Christou, *Phys. Rev. Lett.* **89**, 197201 (2002).
4. A. Osterloh, L. Amico, G. Falci, R. Fazio, *Nature* **416**, 608 (2002).
5. S. Sachdev, *Phys. World* **12**, 33 (1999).
6. S. Sachdev, *Quantum Phase Transitions* (Cambridge Univ. Press, Cambridge, 1999).
7. P. G. de Gennes, *Solid State Commun.* **1**, 132 (1963).
8. R. J. Elliott, P. Pfeuty, C. Wood, *Phys. Rev. Lett.* **25**, 443 (1970).
9. R. B. Stinchcombe, *J. Phys. C* **6**, 2459 and 2484 (1973).
10. D. Bitko, T. F. Rosenbaum, G. Aeppli, *Phys. Rev. Lett.* **77**, 940 (1997).
11. T. F. Rosenbaum *et al.*, *J. Appl. Phys.* **70**, 5946 (1991).
12. D. Bitko, thesis, University of Chicago (1997).
13. R. Giraud *et al.*, *Phys. Rev. Lett.* **87**, 057203 (2001).
14. J. Brooke, D. Bitko, T. F. Rosenbaum, G. Aeppli, *Science* **284**, 779 (1999).
15. J. Brooke, T. F. Rosenbaum, G. Aeppli, *Nature* **413**, 610 (2001).
16. S. Ghosh, R. Parthasarathy, T. F. Rosenbaum, G. Aeppli, *Science* **296**, 2195 (2002).
17. S. Ghosh *et al.*, *Nature* **425**, 48 (2003).
18. K. Andres, *Phys. Rev. B* **7**, 4295 (1973).
19. R. W. Youngblood, G. Aeppli, J. D. Axe, J. A. Griffin, *Phys. Rev. Lett.* **49**, 1724 (1982).
20. H. M. Rønnow, thesis, Risø National Laboratory, Denmark (2000).
21. M. J. M. Leask *et al.*, *J. Phys. C* **6**, 505 (1994).
22. A. P. Ramirez, J. Jensen, *J. Phys. C* **6**, L215 (1994).
23. J. Magariño, J. Tuchendler, P. Beauvillain, I. Laursen, *Phys. Rev. B* **21**, 18 (1980).
24. G. Mennenga, L. J. de Jongh, W. J. Huiskamp, *J. Magn. Magn. Mater.* **44**, 59 (1984).
25. H. M. Rønnow *et al.*, in preparation.
26. J. Jensen, *Phys. Rev. B* **49**, 11833 (1994).
27. P. B. Chakraborty, P. Henelius, H. Kjønsberg, A. W. Sandvik, S. M. Girvin, *Phys. Rev. B* **70**, 144411 (2004).
28. P. Gegenwart *et al.*, *Phys. Rev. Lett.* **89**, 056402 (2002).
29. P. C. E. Stamp, I. S. Tupitsyn, *Phys. Rev. B* **69**, 014401 (2004).

30. We thank G. McIntyre for his expert assistance during complementary measurements on the D10 diffractometer at the Institut Laue Langevin, Grenoble, France. Work at the University of Chicago was supported by NSF Materials Research Science and Engineering Centers grant DMR-0213745. Work in London was supported by the Wolfson–Royal Society Research Merit Award Program and the Basic Technologies program of the UK Research Councils.

6 December 2004; accepted 23 February 2005
10.1126/science.1108317

Atomic-Scale Visualization of Inertial Dynamics

A. M. Lindenberg,[1] J. Larsson,[2] K. Sokolowski-Tinten,[3] K. J. Gaffney,[1] C. Blome,[4] O. Synnergren,[2] J. Sheppard,[5] C. Caleman,[6] A. G. MacPhee,[7] D. Weinstein,[7] D. P. Lowney,[7] T. K. Allison,[7] T. Matthews,[7] R. W. Falcone,[7] A. L. Cavalieri,[8] D. M. Fritz,[8] S. H. Lee,[8] P. H. Bucksbaum,[8] D. A. Reis,[8] J. Rudati,[9] P. H. Fuoss,[10] C. C. Kao,[11] D. P. Siddons,[11] R. Pahl,[12] J. Als-Nielsen,[13] S. Duesterer,[4] R. Ischebeck,[4] H. Schlarb,[4] H. Schulte-Schrepping,[4] Th. Tschentscher,[4] J. Schneider,[4] D. von der Linde,[14] O. Hignette,[15] F. Sette,[15] H. N. Chapman,[16] R. W. Lee,[16] T. N. Hansen,[2] S. Techert,[17] J. S. Wark,[5] M. Bergh,[6] G. Huldt,[6] D. van der Spoel,[6] N. Timneanu,[6] J. Hajdu,[6] R. A. Akre,[18] E. Bong,[18] P. Krejcik,[18] J. Arthur,[1] S. Brennan,[1] K. Luening,[1] J. B. Hastings[1]

The motion of atoms on interatomic potential energy surfaces is fundamental to the dynamics of liquids and solids. An accelerator-based source of femtosecond x-ray pulses allowed us to follow directly atomic displacements on an optically modified energy landscape, leading eventually to the transition from crystalline solid to disordered liquid. We show that, to first order in time, the dynamics are inertial, and we place constraints on the shape and curvature of the transition-state potential energy surface. Our measurements point toward analogies between this nonequilibrium phase transition and the short-time dynamics intrinsic to equilibrium liquids.

In a crystal at room temperature, vibrational excitations, or phonons, only slightly perturb the crystalline order. In contrast, liquids explore a wide range of configurations set by the topology of a complex and time-dependent potential energy surface (*1*, *2*). By using light to trigger changes in this energy landscape, well-defined initial and final states can be generated to which a full range of time-resolved techniques may be applied. In particular, light-induced structural transitions between the crystalline and liquid states of matter may act as simple models for dynamics intrinsic to the liquid state or to transition states in general (*3*).

In this context, a new class of nonthermal processes governing the ultrafast solid-liquid melting transition has recently emerged, supported by time-resolved optical (*4*–*7*) and x-ray (*8*–*10*) experiments and with technological applications ranging from micromachining to eye surgery (*11*). Intense femtosecond excitation of semiconductor materials results in the excitation of a dense electron-hole plasma, with accompanying dramatic changes in the interatomic potential (*12*–*14*). At sufficiently high levels of excitation, it is thought that this process leads to disordering of the crystalline lattice on time scales faster than the time scale for thermal equilibration [often known as the electron-phonon coupling time, on the order of a few picoseconds (*15*)]. In a pioneering study, Rousse *et al.* (*9*) determined that the structure of indium antimonide (InSb) changes on sub-picosecond time scales, but the mechanism by which this occurs and the microscopic pathways the atoms follow have remained elusive, in part because of uncertainties in the pulse duration of laser-plasma sources and signal-to-noise limitations.

Research and development efforts leading toward the Linac Coherent Light Source (LCLS) free-electron laser have facilitated the construction of a new accelerator-based x-ray source, the Sub-Picosecond Pulse Source (SPPS), which uses the same linac-based acceleration and electron bunch compression schemes to be used at future free-electron lasers (*16*, *17*). In order to produce femtosecond x-ray bursts, electron bunches at the Stanford Linear Accelerator Center (SLAC) are chirped and then sent through a series of energy-dispersive magnetic chicanes to create 80-fs electron pulses. These pulses are then transported through an undulator to create sub-100-femtosecond x-ray pulses (*18*). In order to overcome the intrinsic jitter between x-rays and a Ti:sapphire-based femtosecond laser

[1]Stanford Synchrotron Radiation Laboratory/Stanford Linear Accelerator Center (SLAC), Menlo Park, CA 94025, USA. [2]Department of Physics, Lund Institute of Technology, Post Office Box 118, S-22100, Lund, Sweden. [3]Institut für Optik und Quantenelektronik, Friedrich-Schiller Universität Jena, Max-Wien-Platz 1, 07743 Jena, Germany. [4]Deutsches Elektronen-Synchrotron DESY, Notkestrasse 85, 22607 Hamburg, Germany. [5]Department of Physics, Clarendon Laboratory, Parks Road, University of Oxford, Oxford OX1 3PU, UK. [6]Department of Cell and Molecular Biology, Biomedical Centre, Uppsala University, SE-75124 Uppsala, Sweden. [7]Department of Physics, University of California, Berkeley, CA 94720, USA. [8]FOCUS (Frontiers in Optical Coherent and Ultrafast Science) Center, Department of Physics and Applied Physics Program, University of Michigan, Ann Arbor, MI 48109, USA. [9]Advanced Photon Source, [10]Materials Science Division, Argonne National Laboratory, Argonne, IL 60439, USA. [11]National Synchrotron Light Source, Brookhaven National Laboratory, Upton, NY 11973, USA. [12]Consortium for Advanced Radiation Sources, University of Chicago, Chicago, IL 60637, USA. [13]Niels Bohr Institute, Copenhagen University, 2100 Copenhagen Ø, Denmark. [14]Institut für Experimentelle Physik, Universität Duisburg-Essen, D-45117 Essen, Germany. [15]European Synchrotron Radiation Facility, 38043 Grenoble Cedex 9, France. [16]Physics Department, Lawrence Livermore National Laboratory, Livermore, CA 94550, USA. [17]Max Plank Institute for Biophysical Chemistry, Am Faßberg 11, 37077 Göttingen, Germany. [18]SLAC, Menlo Park, CA 94025, USA.

A silicon-based nuclear spin quantum computer

B. E. Kane

Semiconductor Nanofabrication Facility, School of Physics, University of New South Wales, Sydney 2052, Australia

Quantum computers promise to exceed the computational efficiency of ordinary classical machines because quantum algorithms allow the execution of certain tasks in fewer steps. But practical implementation of these machines poses a formidable challenge. Here I present a scheme for implementing a quantum-mechanical computer. Information is encoded onto the nuclear spins of donor atoms in doped silicon electronic devices. Logical operations on individual spins are performed using externally applied electric fields, and spin measurements are made using currents of spin-polarized electrons. The realization of such a computer is dependent on future refinements of conventional silicon electronics.

Although the concept of information underlying all modern computer technology is essentially classical, phsyicists know that nature obeys the laws of quantum mechanics. The idea of a quantum computer has been developed theoretically over several decades to elucidate fundamental questions concerning the capabilities and limitations of machines in which information is treated quantum mechanically[1,2]. Specifically, in quantum computers the ones and zeros of classical digital computers are replaced by the quantum state of a two-level system (a qubit). Logical operations carried out on the qubits and their measurement to determine the result of the computation must obey quantum-mechanical laws. Quantum computation can in principle only occur in systems that are almost completely isolated from their environment and which consequently must dissipate no energy during the process of computation, conditions that are extraordinarily difficult to fulfil in practice.

Interest in quantum computation has increased dramatically in the past four years because of two important insights: first, quantum algorithms (most notably for prime factorization[3,4] and for exhaustive search[5]) have been developed that outperform the best known algorithms doing the same tasks on a classical computer. These algorithms require that the internal state of the quantum computer be controlled with extraordinary precision, so that the coherent quantum state upon which the quantum algorithms rely is not destroyed. Because completely preventing decoherence (uncontrolled interaction of a quantum system with its surrounding environment) is impossible, the existence of quantum algorithms does not prove that they can ever be implemented in a real machine.

The second critical insight has been the discovery of quantum error-correcting codes that enable quantum computers to operate despite some degree of decoherence and which may make quantum computers experimentally realizable[6,7]. The tasks that lie ahead to create an actual quantum computer are formidable: Preskill[8] has estimated that a quantum computer operating on 10^6 qubits with a 10^{-6} probability of error in each operation would exceed the capabilities of contemporary conventional computers on the prime factorization problem. To make use of error-correcting codes, logical operations and measurement must be able to proceed in parallel on qubits throughout the computer.

The states of spin 1/2 particles are two-level systems that can potentially be used for quantum computation. Nuclear spins have been incorporated into several quantum computer proposals[9–12] because they are extremely well isolated from their environment and so operations on nuclear spin qubits could have low error rates. The primary challenge in using nuclear spins in quantum computers lies in measuring the spins. The bulk spin resonance approach to quantum computation[11,12] circumvents the single-spin detection problem essentially by performing quantum calculations in parallel in a large number of molecules and determining the result from macroscopic magnetization measurements. The measurable signal decreases with the number of qubits, however, and scaling this approach above about ten qubits will be technically demanding[37].

To attain the goal of a 10^6 qubit quantum computer, it has been suggested that a 'solid state' approach[13] might eventually replicate the enormous success of modern electronics fabrication technology. An attractive alternative approach to nuclear spin quantum computation is to incorporate nuclear spins into an electronic device and to detect the spins and control their interactions electronically[14]. Electron and nuclear spins are coupled by the hyperfine interaction[15]. Under appropriate circumstances, polarization is transferred between the two spin systems and nuclear spin polarization is detectable by its effect on the electronic properties of a sample[16,17]. Electronic devices for both generating and detecting nuclear spin polarization, implemented at low temperatures in $GaAs/Al_xGa_{1-x}As$ heterostructures, have been developed[18], and similar devices have been incorporated into nanostructures[19,20]. Although the number of spins probed in the nanostructure experiments is still large ($\sim 10^{11}$; ref. 19), sensitivity will improve in optimized devices and in systems with larger hyperfine interactions.

Here I present a scheme for implementing a quantum computer on an array of nuclear spins located on donors in silicon, the semiconductor used in most conventional computer electronics. Logical operations and measurements can in principle be performed independently and in parallel on each spin in the array. I describe specific electronic devices for the manipulation and measurement of nuclear spins, fabrication of which will require significant advances in the rapidly moving field of nanotechnology. Although it is likely that scaling the devices proposed here into a computer of the size envisaged by Preskill[8] will be an extraordinary challenge, a silicon-based quantum computer is in a unique position to benefit from the resources and ingenuity being directed towards making conventional electronics of ever smaller size and greater complexity.

Quantum computation with a ^{31}P array in silicon

The strength of the hyperfine interaction is proportional to the probability density of the electron wavefunction at the nucleus. In semiconductors, the electron wavefunction extends over large distances through the crystal lattice. Two nuclear spins can consequently interact with the same electron, leading to electron-mediated or indirect nuclear spin coupling[15]. Because the electron is sensitive to externally applied electric fields, the hyperfine inter-

action and electron-mediated nuclear spin interaction can be controlled by voltages applied to metallic gates in a semiconductor device, enabling the external manipulation of nuclear spin dynamics that is necessary for quantum computation.

The conditions required for electron-coupled nuclear spin computation and single nuclear spin detection can arise if the nuclear spin is located on a positively charged donor in a semiconductor host. The electron wavefunction is then concentrated at the donor nucleus (for *s* orbitals and energy bands composed primarily of them), yielding a large hyperfine interaction energy. For shallow-level donors, however, the electron wavefunction extends tens or hundreds of ångströms away from the donor nucleus, allowing electron-mediated nuclear spin coupling to occur over comparable distances. The quantum computer proposed here comprises an array of such donors positioned beneath the surface of a semiconductor host (Fig. 1). A quantum mechanical calculation proceeds by the precise control of three external parameters: (1) gates above the donors control the strength of the hyperfine interactions and hence the resonance frequency of the nuclear spins beneath them; (2) gates between the donors turn on and off electron-mediated coupling between the nuclear spins[13]; (3) a globally applied a.c. magnetic field B_{ac} flips nuclear spins at resonance. Custom adjustment of the coupling of each spin to its neighbours and to B_{ac} enables different operations to be performed on each of the spins simultaneously. Finally, measurements are performed by transferring nuclear spin polarization to the electrons and determining the electron spin state by its effect on the orbital wavefunction of the electrons, which can be probed using capacitance measurements between adjacent gates.

An important requirement for a quantum computer is to isolate the qubits from any degrees of freedom that may lead to decoherence. If the qubits are spins on a donor in a semiconductor, nuclear spins in the host are a large reservoir with which the donor spins can interact. Consequently, the host should contain only nuclei with spin $I = 0$. This simple requirement unfortunately eliminates all III–V semiconductors as host candidates, because none of their constituent elements possesses stable $I = 0$ isotopes[21]. Group IV semiconductors are composed primarily $I = 0$ isotopes and can in principle be purified to contain only $I = 0$ isotopes. Because of the advanced state of Si materials technology and the tremendous effort currently underway in Si nanofabrication, Si is the obvious choice for the semiconductor host.

The only $I = 1/2$ shallow (group V) donor in Si is ^{31}P. The Si:^{31}P system was exhaustively studied 40 years ago in the first electron–nuclear double-resonance experiments[22,23]. At sufficiently low ^{31}P concentrations at temperature $T = 1.5$ K, the electron spin relaxation time is thousands of seconds and the ^{31}P nuclear spin relaxation time exceeds 10 hours. It is likely that at millikelvin temperatures the phonon limited ^{31}P relaxation time is of the order of 10^{18} seconds (ref. 24), making this system ideal for quantum computation.

The purpose of the electrons in the computer is to mediate nuclear spin interactions and to facilitate measurement of the nuclear spins. Irreversible interactions between electron and nuclear spins must not occur as the computation proceeds: the electrons must be in a non-degenerate ground state throughout the computation. At sufficiently low temperatures, electrons only occupy the lowest energy-bound state at the donor, whose twofold spin degeneracy is broken by an applied magnetic field B. (The valley degeneracy of the Si conduction band is broken in the vicinity of the donor[25]. The lowest donor excited state is approximately 15 meV above the ground state[23].) The electrons will only occupy the lowest energy spin level when $2\mu_B B \gg kT$, where μ_B is the Bohr magneton. (In Si, the Landé g-factor is very close to +2, so $g = 2$ is used throughout this discussion.) The electrons will be completely spin-polarized ($n_\uparrow/n_\downarrow < 10^{-6}$) when $T \leqslant 100$ mK and $B \geqslant 2$ tesla. A quantum-mechanical computer is non-dissipative and can consequently operate at low temperatures. Dissipation will arise external to the computer from gate biasing and from eddy currents caused by B_{ac}, and during polarization and measurement of the nuclear spins. These effects will determine the minimum operable temperature of the computer. For this discussion, I will assume $T = 100$ mK and $B = 2$ T. Note that these conditions do not fully polarize the nuclear spins, which are instead aligned by interactions with the polarized electrons.

Magnitude of spin interactions in Si:^{31}P

The size of the interactions between spins determines both the time

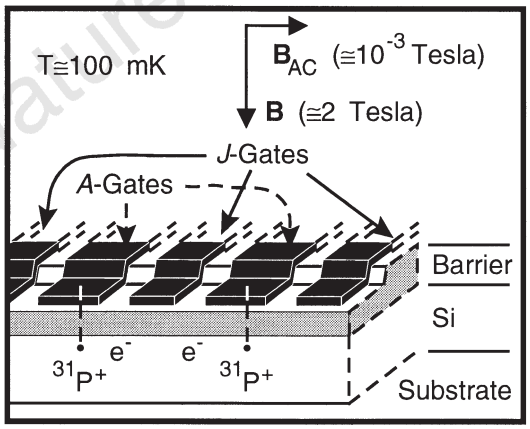


Figure 1 Illustration of two cells in a one-dimensional array containing ^{31}P donors and electrons in a Si host, separated by a barrier from metal gates on the surface. '*A* gates' control the resonance frequency of the nuclear spin qubits; '*J* gates' control the electron-mediated coupling between adjacent nuclear spins. The ledge over which the gates cross localizes the gate electric field in the vicinity of the donors.

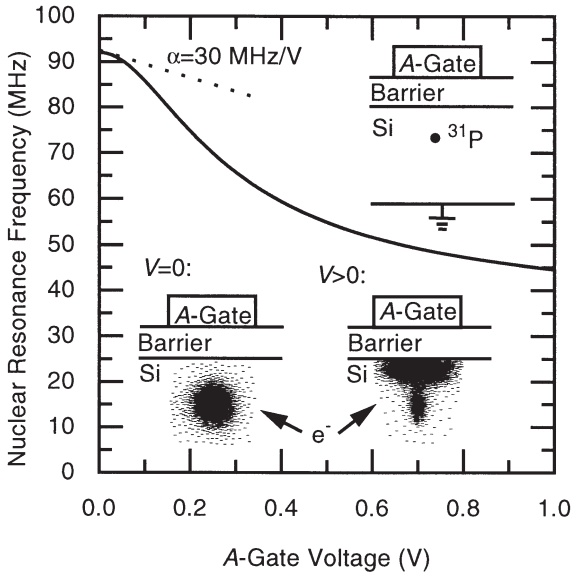


Figure 2 An electric field applied to an *A* gate pulls the electron wavefunction away from the donor and towards the barrier, reducing the hyperfine interaction and the resonance frequency of the nucleus. The donor nucleus–electron system is a voltage-controlled oscillator with a tuning parameter α of the order of 30 MHz V^{-1}.

required to do elementary operations on the qubits and the separation necessary between donors in the array. The hamiltonian for a nuclear spin–electron system in Si, applicable for an $I = 1/2$ donor nucleus and with $B \| z$ is $H_{en} = \mu_B B\sigma_z^e - g_n\mu_n B\sigma_z^n + A\sigma^e\cdot\sigma^n$, where σ are the Pauli spin matrices (with eigenvalues ± 1), μ_n is the nuclear magneton, g_n is the nuclear g-factor (1.13 for ^{31}P; ref. 21), and $A = \frac{8}{3}\pi\mu_B g_n\mu_n|\Psi(0)|^2$ is the contact hyperfine interaction energy, with $|\Psi(0)|^2$, the probability density of the electron wavefunction, evaluated at the nucleus. If the electron is in its ground state, the frequency separation of the nuclear levels is, to second order

$$h\nu_A = 2g_n\mu_n B + 2A + \frac{2A^2}{\mu_B B} \quad (1)$$

In Si:^{31}P, $2A/h = 58$ MHz, and the second term in equation (1) exceeds the first term for $B < 3.5$ T.

An electric field applied to the electron-donor system shifts the electron wavefunction envelope away from the nucleus and reduces the hyperfine interaction. The size of this shift, following estimates of Kohn[25] of shallow donor Stark shifts in Si, is shown in Fig. 2 for a donor 200 Å beneath a gate. A donor nuclear spin–electron system close to an '*A* gate' functions as a voltage-controlled oscillator: the precession frequency of the nuclear spin is controllable externally, and spins can be selectively brought into resonance with B_{ac}, allowing arbitrary rotations to be performed on each nuclear spin.

Quantum mechanical computation requires, in addition to single spin rotations, the two-qubit 'controlled rotation' operation, which rotates the spin of a target qubit through a prescribed angle if, and only if, the control qubit is oriented in a specified direction, and leaves the orientation of the control qubit unchanged[26,27]. Performing the controlled rotation operation requires nuclear-spin exchange between two donor nucleus-electron spin systems[13], which will arise from electron-mediated interactions when the donors are sufficiently close to each other. The hamiltonian of two coupled donor nucleus–electron systems, valid at energy scales small compared to the donor-electron binding energy, is $H = H(B) + A_1\sigma^{1n}\cdot\sigma^{2e} + A_2\sigma^{2n}\cdot\sigma^{2e} + J\sigma^{1e}\cdot\sigma^{2e}$, where $H(B)$ are the magnetic field interaction terms for the spins. A_1 and A_2 are the hyperfine interaction energies of the respective nucleus–electron systems. $4J$, the exchange energy, depends on the overlap of the electron wavefunctions. For well separated donors[28]

$$4J(r) \cong 1.6\frac{e^2}{\epsilon a_B}\left(\frac{r}{a_B}\right)^{\frac{5}{2}}\exp\left(\frac{-2r}{a_B}\right) \quad (2)$$

where r is the distance between donors, ϵ is the dielectric constant of the semiconductor, and a_B is the semiconductor Bohr radius. This function, with values appropriate for Si, is plotted in Fig. 3. Equation (2), originally derived for H atoms, is complicated in Si by its valley degenerate anisotropic band structure[29]. Exchange coupling terms from each valley interfere, leading to oscillatory behaviour of $J(r)$. In this discussion, the complications introduced by Si band structure will be neglected. In determining $J(r)$ in Fig. 3, the transverse mass for Si ($\cong 0.2m_e$) has been used, and $a_B = 30$ Å. Because J is proportional to the electron wave function overlap, it can be varied by an electrostatic potential imposed by a '*J*-gate' positioned between the donors[13]. As shall be seen below, significant coupling between nuclei will occur when $4J \approx \mu_B B$, and this condition approximates the necessary separation between donors of 100–200 Å. Whereas actual separations may be considerably larger than this value because the *J* gate can be biased positively to reduce the barrier between donors, the gate sizes required for the quantum computer are near the limit of current electronics fabrication technology.

For two-electron systems, the exchange interaction lowers the electron singlet ($|\uparrow\downarrow - \downarrow\uparrow\rangle$) energy with respect to the triplets[30]. (The $|\uparrow\downarrow\rangle$ notation is used here to represent the electron spin state, and the $|01\rangle$ notation the nuclear state; in the $|\downarrow\downarrow 11\rangle$ state, all spins point in the same direction. For simplicity, normalization constants are omitted.) In a magnetic field, however, $|\downarrow\downarrow\rangle$ will be the electron ground state if $J < \mu_B B/2$ (Fig. 4a). In the $|\downarrow\downarrow\rangle$ state, the energies of the nuclear states can be calculated to second order in A using perturbation theory. When $A_1 = A_2 = A$, the $|10 - 01\rangle$ state is lowered in energy with respect to $|10 + 01\rangle$ by:

$$h\nu_J = 2A^2\left(\frac{1}{\mu_B B - 2J} - \frac{1}{\mu_B B}\right) \quad (3)$$

The $|11\rangle$ state is above the $|10 + 01\rangle$ state and the $|00\rangle$ state below the $|10 - 01\rangle$ state by an energy $h\nu_A$, given in equation (1). For the Si:^{31}P system at $B = 2$ T and for $4J/h = 30$ GHz, equation (3) yields $\nu_J = 75$ kHz. This nuclear spin exchange frequency approximates the rate at which binary operations can be performed on the computer (ν_J can be increased by increasing J, but at the expense of also increasing the relaxation rate of the coupled nuclear–electron spin excitations). The speed of single spin operations is determined by the size of B_{ac} and is comparable to 75 kHz when $B_{ac} = 10^{-3}$ T.

Spin measurements

Measurement of nuclear spins in the proposed quantum computer is accomplished in a two-step process: distinct nuclear spin states are adiabatically converted into states with different electron polarization, and the electron spin is determined by its effect on the symmetry of the orbital wavefunction of an exchange-coupled two-electron system. A procedure for accomplishing this conversion is shown in Fig. 4. While computation is done when $J < \mu_B B/2$ and the electrons are fully polarized, measurements are made when $J > \mu_B B/2$, and $|\uparrow\downarrow - \downarrow\uparrow\rangle$ states have the lowest energy (Fig. 4a). As the electron levels cross, the $|\downarrow\downarrow\rangle$ and $|\uparrow\downarrow - \downarrow\uparrow\rangle$ states are coupled by hyperfine interactions with the nuclei. During an adiabatic increase in J, the two lower-energy nuclear spin states at $J = 0$ evolve into $|\uparrow\downarrow - \downarrow\uparrow\rangle$ states when $J > \mu_B B/2$, whereas the two higher-energy nuclear states remain $|\downarrow\downarrow\rangle$. If, at $J = 0$, $A_1 > A_2$, the orientation of nuclear spin 1 alone will determine whether the system evolves into the $|\uparrow\downarrow - \downarrow\uparrow\rangle$ or the $|\downarrow\downarrow\rangle$ state during an adiabatic increase in J.

A method to detect the electron spin state by using electronic

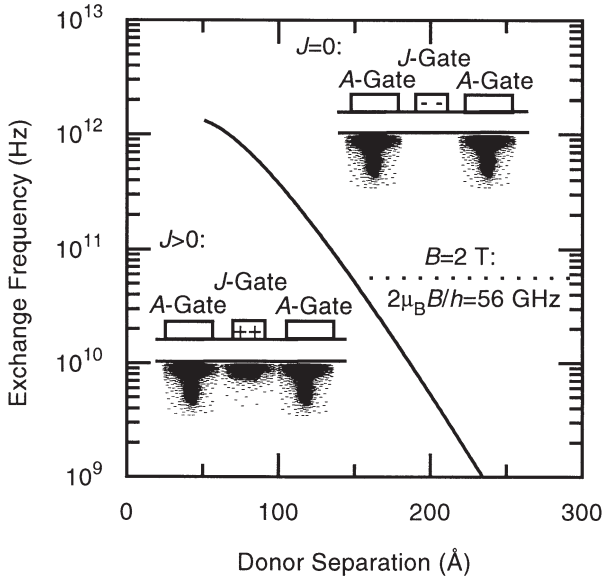


Figure 3 *J* gates vary the electrostatic potential barrier *V* between donors to enhance or reduce exchange coupling, proportional to the electron wavefunction overlap. The exchange frequency ($4J/h$) when $V = 0$ is plotted for Si.

means is shown in Fig. 4b. Both electrons can become bound to the same donor (a D^- state) if the A gates above the donors are biased appropriately. In Si:P, the D^- state is always a singlet with a second electron binding energy of 1.7 meV (refs 31, 32). Consequently, a differential voltage applied to the A gates can result in charge motion between the donors that only occurs if the electrons are in a singlet state. This charge motion is measurable using sensitive single-electron capacitance techniques[33]. This approach to spin measurement produces a signal that persists until the electron spin relaxes, a time that, as noted above, can be thousands of seconds in Si:P.

The spin measurement process can also be used to prepare nuclear spins in a prescribed state by first determining the state of a spin and flipping it if necessary so that it ends up in the desired spin state. As with the spin computation procedures already discussed, spin measurement and preparation can in principle be performed in parallel throughout the computer.

Initializing the computer

Before any computation, the computer must be initialized by calibrating the A gates and the J gates. Fluctuations from cell to cell in the gate biases necessary to perform logical operations are an inevitable consequence of variations in the positions of the donors and in the sizes of the gates. The parameters of each cell, however, can be determined individually using the measurement capabilities of the computer, because the measurement technique discussed here does not require precise knowledge of the J and A couplings. The A-gate voltage at which the underlying nuclear spin is resonant with an applied B_{ac} can be determined using the technique of adiabatic fast passage[34]: when $B_{ac} = 0$, the nuclear spin is measured and the A gate is biased at a voltage known to be off resonance. B_{ac} is then switched on, and the A gate bias is swept through a prescribed voltage interval. B_{ac} is then switched off and the nuclear spin is measured again. The spin will have flipped if, and only if, resonance occurred within the prescribed A-gate voltage range. Testing for spin flips in increasingly small voltage ranges leads to the determination of the resonance voltage. Once adjacent A gates have been calibrated, the J gates can be calibrated in a similar manner by sweeping J-gate biases across resonances of two coupled cells.

This calibration procedure can be performed in parallel on many cells, so calibration is not a fundamental impediment to scaling the computer to large sizes. Calibration voltages can be stored on capacitors located on the Si chip adjacent to the quantum computer. External controlling circuitry would thus need to control only the timing of gate biases, and not their magnitudes.

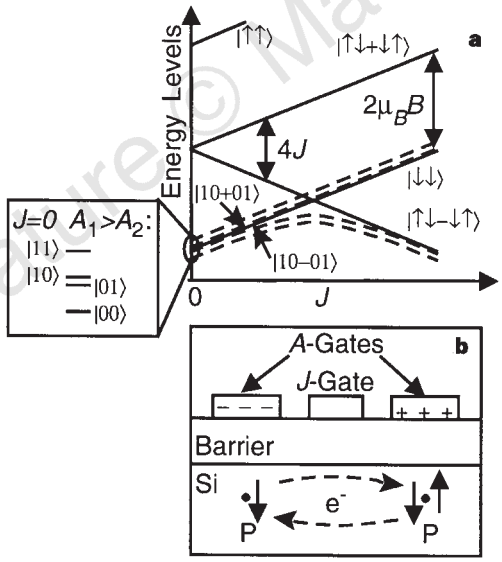


Figure 4 Two qubit quantum logic and spin measurement. **a**, Electron (solid lines) and lowest energy-coupled electron-nuclear (dashed lines) energy levels as a function of J. When $J < \mu_B B/2$, two qubit computations are performed by controlling the $|10-01\rangle - |10+01\rangle$ level splitting with a J gate. Above $J = \mu_B B/2$, the states of the coupled system evolve into states of differing electron polarization. The state of the nucleus at $J = 0$ with the larger energy splitting (controllable by the A gate bias) determines the final electron spin state after an adiabatic increase in J. **b**, Only $|\uparrow\downarrow - \downarrow\uparrow\rangle$ electrons can make transitions into states in which electrons are bound to the same donor (D^- states). Electron current during these transitions is measurable using capacitive techniques, enabling the underlying spin states of the electrons and nuclei to be determined.

Spin decoherence introduced by gates

In the quantum computer architecture outlined above, biasing of A gates and J gates enables custom control of the qubits and their mutual interactions. The presence of the gates, however, will lead to decoherence of the spins if the gate biases fluctuate away from their desired values. These effects need to be considered to evaluate the performance of any gate-controlled quantum computer. During the computation, the largest source of decoherence is likely to arise from voltage fluctuations on the A gates. (When $J < \mu_B B/2$, modulation of the state energies by the J gates is much smaller than by the A gates. J exceeds $\mu_B B/2$ only during the measurement process, when decoherence will inevitably occur.) The precession frequencies of two spins in phase at $t = 0$ depends on the potentials on their respective A gates. Differential fluctuations of the potentials produce differences in the precession frequency. At some later time $t = t_\phi$, the spins will be 180° out of phase; t_ϕ can be estimated by determining the transition rate between $|10+01\rangle$ (spins in phase) and $|10-01\rangle$ (spins 180° out of phase) of a two-spin system. The hamiltonian that couples these states is $H_\phi = \frac{1}{4}h\Delta(\sigma_z^{1n} - \sigma_z^{2n})$, where Δ is the fluctuating differential precession frequency of the spins. Standard treatment of fluctuating hamiltonians[34] predicts: $t_\phi^{-1} = \pi^2 S_\Delta(\nu_{st})$, where S_Δ is the spectral density of the frequency fluctuations, and ν_{st} is the frequency difference between the $|10-01\rangle$ and $|10+01\rangle$ states. At a particular bias voltage, the A gates have a frequency tuning parameter $\alpha = d\Delta/dV$. Thus:

$$t_\phi^{-1} = \pi^2 \alpha^2(V) S_V(\nu_{st}) \qquad (4)$$

where S_V is the spectral density of the gate voltage fluctuations.

S_V for good room temperature electronics is of order 10^{-18} V^2/Hz, comparable to the room temperature Johnson noise of a 50-Ω resistor. The value of α, estimated from Fig. 2, is 10–100 MHz V^{-1}, yielding t_ϕ = 10–1,000 s; α is determined by the size of the donor array cells and cannot readily be reduced (to increase t_ϕ) without reducing the exchange interaction between cells. Because α is a function of the gate bias (Fig. 2), t_ϕ can be increased by minimizing the voltage applied to the A gates.

Although equation (4) is valid for white noise, at low frequencies it is likely that materials-dependent fluctuations ($1/f$ noise) will be the dominant cause of spin dephasing. Consequently, it is difficult to give hard estimates of t_ϕ for the computer. Charge fluctuations within the computer (arising from fluctuating occupancies of traps and surface states, for example) are likely to be particularly important, and minimizing them will place great demands on computer fabrication.

Although materials-dependent fluctuations are difficult to estimate, the low-temperature operations of the computer and the dissipationless nature of quantum computing mean that, in principle, fluctuations can be kept extremely small: using low-temperature electronics to bias the gates (for instance, by using on chip capacitors as discussed above) could produce $t_\phi \approx 10^6$ s. Electronically controlled nuclear spin quantum computers thus have the theoretical capability to perform at least 10^5 to perhaps 10^{10}

logical operations during t_ϕ, and can probably meet Preskill's criterion[8] for an error probability of 10^{-6} per qubit operation.

Constructing the computer

Building the computer presented here will obviously be an extraordinary challenge: the materials must be almost completely free of spin ($I \neq 0$ isotopes) and charge impurities to prevent dephasing fluctuations from arising within the computer. Donors must be introduced into the material in an ordered array hundreds of Å beneath the surface. Finally, gates with lateral dimensions and separations $\sim$100 Å must be patterned on the surface, registered to the donors beneath them. Although it is possible that the computer can use SiO_2 as the barrier material (the standard MOS technology used in most current conventional electronics), the need to reduce disorder and fluctuations to a minimum means that heteroepitaxial materials, such as Si/SiGe, may ultimately be preferable to Si/SiO_2.

The most obvious obstacle to building to the quantum computer presented above is the incorporation of the donor array into the Si layer beneath the barrier layer. Currently, semiconductor structures are deposited layer by layer. The δ-doping technique produces donors lying on a plane in the material, with the donors randomly distributed within the plane. The quantum computer envisaged here requires that the donors be placed into an ordered one- or two-dimensional array; furthermore, precisely one donor must be placed into each array cell, making it extremely difficult to create the array by using lithography and ion implantation or by focused deposition. Methods currently under development to place single atoms on surfaces using ultra-high-vacuum scanning tunnelling microscopy[35] or atom optics techniques[36] are likely candidates to be used to position the donor array. A challenge will be to grow high-quality Si layers on the surface subsequent to placement of the donors.

Fabricating large arrays of donors may prove to be difficult, but two-spin devices, which can be used to test the logical operations and measurement techniques presented here, can be made using random doping techniques. Although only a small fraction of such devices will work properly, adjacent conventional Si electronic multiplexing circuitry can be used to examine many devices separately. The relative ease of fabricating such 'hybrid' (quantum-conventional) circuits is a particularly attractive feature of Si-based quantum computation.

In a Si-based nuclear spin quantum computer, the highly coherent quantum states necessary for quantum computation are incorporated into a material in which the ability to implement complex computer architectures is well established. The substantial challenges facing the realization of the computer, particularly in fabricating 100-Å-scale gated devices, are similar to those facing the next generation of conventional electronics; consequently, new manufacturing technologies being developed for conventional electronics will bear directly on efforts to develop a quantum computer in Si. Quantum computers sufficiently complex that they can achieve their theoretical potential may thus one day be built using the same technology that is used to produce conventional computers. □

Received 10 November 1997; accepted 24 February 1998.

1. Steane, A. Quantum computing. *Rep. Prog. Phys.* **61,** 117–173 (1998).
2. Bennett, C. H. Quantum information and computation. *Physics Today* 24–30 (Oct. 1995).
3. Shor, P. W. in *Proc. 35th Annu. Symp. Foundations of Computer Science* (ed. Goldwasser, S.) 124–134 (IEEE Computer Society, Los Alamitos, CA, 1994).
4. Ekert, A. & Jozsa, R. Quantum computation and Shor's factoring algorithm. *Rev. Mod. Phys.* **68,** 733–753 (1996).
5. Grover, L. K. Quantum mechanics helps in searching for a needle in a haystack. *Phys. Rev. Lett.* **79,** 325–328 (1997).
6. Calderbank, A. R. & Shor, P. W. Good quantum error correcting codes exist. *Phys. Rev. A* **54,** 1098–1105 (1996).
7. Steane, A. M. Error correcting codes in quantum theory. *Phys. Rev. Lett.* **77,** 793–797 (1996).
8. Preskill, J. Reliable quantum computers. *Proc. R. Soc. Lond. A* **454,** 385–410 (1998).
9. Lloyd, S. A potentially realizable quantum computer. *Science* **261,** 1569–1571 (1993).
10. DiVincenzo, D. P. Quantum computation. *Science* **270,** 255–261 (1995).
11. Gershenfeld, N. A. & Chuang, I. L. Bulk spin-resonance quantum computation. *Science* **275,** 350–356 (1997).
12. Cory, D. G., Fahmy, A. F. & Havel, T. F. Ensemble quantum computing by NMR spectroscopy. *Proc. Natl Acad. Sci. USA* **94,** 1634–1639 (1997).
13. Loss, D. & DiVincenzo, D. P. Quantum computation with quantum dots. *Phys. Rev. A* **57,** 120–126 (1998).
14. Privman, V., Vagner, I. D. & Kventsel, G. Quantum computation in quantum Hall systems. *Phys. Lett. A* **239,** 141–146 (1998)..
15. Slichter, C. P. *Principles of Magnetic Resonance* 3rd edn, Ch 4 (Springer, Berlin, 1990).
16. Dobers, M., Klitzing, K. v., Schneider, J., Weimann, G. & Ploog, K. Electrical detection of nuclear magnetic resonance in GaAs-Al_xGa_{1-x}As heterostructures. *Phys. Rev. Lett.* **61,** 1650–1653 (1988).
17. Stich, B., Greulich-Weber, S. & Spaeth, J.-M. Electrical detection of electron nuclear double resonance in silicon. *Appl. Phys. Lett.* **68,** 1102–1104 (1996).
18. Kane, B. E., Pfeiffer, L. N. & West, K. W. Evidence for an electric-field-induced phase transition in a spin-polarized two-dimensional electron gas. *Phys. Rev. B* **46,** 7264–7267 (1992).
19. Wald, K. W., Kouwenhoven, L. P., McEuen, P. L., van der Vaart, N. C. & Foxon, C. T. Local dynamic nuclear polarization using quantum point contacts. *Phys. Rev. Lett.* **73,** 1011–1014 (1994).
20. Dixon, D. C., Wald, K. R., McEuen, P. L. & Melloch, M. R. Dynamic polarization at the edge of a two-dimensional electron gas. *Phys. Rev. B* **56,** 4743–4750 (1997).
21. *CRC Handbook of Chemistry and Physics* 77th edn 11–38 (CRC Press, Boca Raton, Florida, 1996).
22. Feher, G. Electron spin resonance on donors in silicon. I. Electronic structure of donors by the electron nuclear double resonance technique. *Phys. Rev.* **114,** 1219–1244 (1959).
23. Wilson, D. K. & Feher, G. Electron spin resonance experiments on donors in silicon. III. Investigation of excited states by the application of uniaxial stress and their importance in relaxation processes. *Phys. Rev.* **124,** 1068–1083 (1961).
24. Waugh, J. S. & Slichter, C. P. Mechanism of nuclear spin-lattice relaxation in insulators at very low temperatures. *Phys. Rev. B* **37,** 4337–4339 (1988).
25. Kohn, W. *Solid State Physics* Vol. 5 (eds Seitz, F. & Turnbull, D.) 257–320 (Academic, New York, 1957).
26. DiVincenzo, D. P. Two-bit gates are universal for quantum computation. *Phys. Rev. A* **51,** 1015–1021 (1995).
27. Lloyd, S. Almost any quantum logic gate is universal. *Phys. Rev. Lett.* **75,** 346–349 (1995).
28. Herring, C. & Flicker, M. Asymptotic exchange coupling of two hydrogen atoms. *Phys. Rev.* **134,** A362–A366 (1964).
29. Andres, K., Bhatt, R. N., Goalwin, P., Rice, T. M. & Walstedt, R. E. Low-temperature magnetic susceptibility of Si:P in the nonmetallic region. *Phys. Rev. B* **24,** 244–260 (1981).
30. Ashcroft, N. W. & Mermin, N. D. in *Solid State Physics* Ch. 32 (Saunders College, Philadelphia, 1976).
31. Larsen, D. M. Stress dependence of the binding energy of D^- centers in Si. *Phys. Rev. B* **23,** 5521–5526 (1981).
32. Larsen, D. M. & McCann, S. Y. Variational studies of two- and three-dimensional D^- centers in magnetic fields. *Phys. Rev. B* **46,** 3966–3970 (1992).
33. Ashoori, R. C. Electrons in artificial atoms. *Nature* **379,** 413–419 (1996).
34. Abragam, A. *Principles of Nuclear Magnetism* (Oxford Univ. Press, London, 1961).
35. Lyding, J. W. UHV STM nanofabrication: progress, technology spin-offs, and challenges. *Proc. IEEE* **85,** 589–600 (1997).
36. Adams, C. S., Sigel, J. & Mlynek, J. Atom optics. *Phys. Rep.* **240,** 143–210 (1994).
37. Warren, W. S. The usefulness of NMR quantum computing. *Science* **277,** 1688–1690 (1997).

Acknowledgements. This work has been supported by the Australian Research Council. I thank R. G. Clark for encouragement and E. Hellman for suggesting that the work in ref. 18 could be relevant to quantum computation.

Correspondence should be addressed to the author (e-mail: kane@newt.phys.unsw.edu.au).

Robert Raussendorf

Statement

and

Readings

Decoherence in quantum computation - foe or friend?

Robert Raussendorf, University of British Columbia

Abstract: Decoherence is detrimental to quantum computation because it makes the computation "noisy". Or is it? Upon closer inspection, it turns out that decoherence can both compromize and help realize quantum computation. Which of the two applies does very much depend on the decoherence model considered.
I will start out by proving the expected, namely that decoherence, for a certain (justifiable) class of decoherence models, does indeed compromise quantum computation. In this regard, I will review a result of Bravyi and Kitaev [1b]/ van Dam and Howard [1a] demonstrating an *upper* bound to the error threshold for fault-tolerant quantum computation. The significance of this upper bound is that no method of error correction, however clever, can put the quantum computation back on track if the decoherence level per elementary gate operation is above the threshold value.
In the second part of my introduction, I will discuss two computational models [2], [3] that *use* decoherent dynamics to realize quantum computation. In the case of [2], the computation is driven by local projective measurements on a highly entangled quantum state. Therein, the entanglement of the initial quantum state is progressively destroyed as the computation proceeds. Thus, entanglement is a resource for this computational model. In the second case, Ref. [3], universal quantum computation is implemented in a dissipative quantum system whose evolution is governed by time-independent and local couplings to the environment. Due to the purely dissipative nature of the process, this way of doing quantum computation exhibits some inherent robustness and defies some of the DiVincenzo criteria for quantum computation.

Suggested Reading:

[1a] Sergey Bravyi and Alexei Kitaev, Phys. Rev. A **71**, 022316 (2005).

[1b] Wim van Dam and Mark Howard, Phys. Rev. Lett. **103**, 170504 (2009).

[2] R. Raussendorf and H.J Briegel, Phys. Rev. Lett. **86**, 5188 (2001).

[3] F. Verstraete, M. Wolf and J.I. Cirac, Nature Physics **5**, 633 - 636 (2009).

Remark: The results of refs. [1a] and [1b] are very closely related. For background reading, I recommend Ref. [1b] over [1a] because it is shorter. In my introduction, I will discuss [1a] (first part only), however, because the result therein is better suited for graphical display.

Universal quantum computation with ideal Clifford gates and noisy ancillas

Sergey Bravyi* and Alexei Kitaev†
Institute for Quantum Information, California Institute of Technology, Pasadena, 91125 California, USA
(Received 6 May 2004; published 22 February 2005)

We consider a model of quantum computation in which the set of elementary operations is limited to Clifford unitaries, the creation of the state $|0\rangle$, and qubit measurement in the computational basis. In addition, we allow the creation of a one-qubit ancilla in a mixed state ρ, which should be regarded as a parameter of the model. Our goal is to determine for which ρ universal quantum computation (UQC) can be efficiently simulated. To answer this question, we construct purification protocols that consume several copies of ρ and produce a single output qubit with higher polarization. The protocols allow one to increase the polarization only along certain "magic" directions. If the polarization of ρ along a magic direction exceeds a threshold value (about 65%), the purification asymptotically yields a pure state, which we call a magic state. We show that the Clifford group operations combined with magic states preparation are sufficient for UQC. The connection of our results with the Gottesman-Knill theorem is discussed.

DOI: 10.1103/PhysRevA.71.022316 PACS number(s): 03.67.Lx, 03.67.Pp

I. INTRODUCTION AND SUMMARY

The theory of fault-tolerant quantum computation defines an important number called the error threshold. If the physical error rate is less than the threshold value δ, it is possible to stabilize computation by transforming the quantum circuit into a fault-tolerant form where errors can be detected and eliminated. However, if the error rate is above the threshold, then errors begin to accumulate, which results in rapid decoherence and renders the output of the computation useless. The actual value of δ depends on the error correction scheme and the error model. Unfortunately, this number seems to be rather small for all known schemes. Estimates vary from 10^{-6} (see Ref. [1]) to 10^{-4} (see Refs. [2–4]), which is hardly achievable with the present technology.

In principle, one can envision a situation in which qubits do not decohere, and a subset of the elementary gates is realized *exactly* due to special properties of the physical system. This scenario could be realized experimentally using spin, electron, or other many-body systems with topologically ordered ground states. Excitations in two-dimensional topologically ordered systems are anyons—quasiparticles with unusual statistics described by nontrivial representations of the braid group. If we have sufficient control of anyons, i.e., are able to move them around each other, fuse them, and distinguish between different particle types, then we can realize some set of unitary operators and measurements exactly. This set may or may not be computationally universal. While the universality can be achieved with sufficiently nontrivial types of anyons [5–8], more realistic systems offer only decoherence protection and an incomplete set of topological gates. (See Refs. [9,10] about non-Abelian anyons in quantum Hall systems and Refs. [11,12] about topological orders in Josephson junction arrays.) Nevertheless, universal computation is possible if we introduce some additional operations (e.g., measurements by Aharonov-Bohm interference [13] or some gates that are not related to topology at all). Of course, these nontopological operations cannot be implemented exactly and thus are prone to errors.

In this situation, the threshold error rate δ may become significantly larger than the values given above because we need to correct only errors of certain special type and we introduce a smaller amount of error in the correction stage. The main purpose of the present paper is to illustrate this statement by a particular computational model.

The model is built upon the *Clifford group*—the group of unitary operators that map the group of Pauli operators to itself under conjugation. The set of elementary operations is divided into two parts: $\mathcal{O}=\mathcal{O}_{\text{ideal}}\cup\mathcal{O}_{\text{faulty}}$. Operations from $\mathcal{O}_{\text{ideal}}$ are assumed to be perfect. We list these operations below:

(i) prepare a qubit in the state $|0\rangle$;

(ii) apply unitary operators from the Clifford group;

(iii) measure an eigenvalue of a Pauli operator (σ^x, σ^y, or σ^z) on any qubit.

Here we mean nondestructive projective measurement. We also assume that no errors occur between the operations.

It is well known that these operations are not sufficient for universal quantum computation (UQC) (unless a quantum computer can be efficiently simulated on a classical computer). More specifically, the Gottesman-Knill theorem states that by operations from $\mathcal{O}_{\text{ideal}}$ one can only obtain quantum states of a very special form called *stabilizer states*. Such a state can be specified as an intersection of eigenspaces of pairwise commuting Pauli operators, which are referred to as *stabilizers*. Using the stabilizer formalism, one can easily simulate the evolution of the state and the statistics of measurements on a classical probabilistic computer (see Ref. [14] or a textbook [15] for more details).

The set $\mathcal{O}_{\text{faulty}}$ describes faulty operations. In our model, it consists of just one operation: prepare an ancillary qubit in a mixed state ρ. The state ρ should be regarded as a parameter of the model. From the physical point of view, ρ is mixed due to imperfections of the preparation procedure (entanglement of the ancilla with the environment, thermal fluctua-

*Email address: serg@cs.caltech.edu
†Email address: kitaev@iqi.caltech.edu

tions, etc.). An essential requirement is that by preparing n qubits we obtain the state $\rho^{\otimes n}$, i.e., all ancillary qubits are independent. The independence assumption is similar to the uncorrelated errors model in the standard fault-tolerant computation theory.

Our motivation for including all Clifford group gates into $\mathcal{O}_{\text{ideal}}$ relies mostly on the recent progress in the fault-tolerant implementation of such gates. For instance, using a concatenated stabilizer code with good error correcting properties to encode each qubit and applying gates transversally (so that errors do not propagate inside code blocks) one can implement Clifford gates with an arbitrary high precision, see Ref. [16]. However, these nearly perfect gates act on *encoded* qubits. To establish a correspondence with our model, one needs to prepare an *encoded* ancilla in the state ρ. It can be done using the schemes for fault-tolerant encoding of an arbitrary *known* one-qubit state described by Knill in Ref. [17]. In the more recent paper [18] Knill constructed a scheme of fault-tolerant quantum computation which combines (i) the teleported computing and error correction technique by Gottesman and Chuang [19]; (ii) the method of purification of CSS states by Dür and Briegel [20]; and (iii) the magic states distillation algorithms described in the present paper. As was argued in Ref. [18], this scheme is likely to yield a much higher value for the threshold δ (it may be up to 1%).

Unfortunately, ideal implementation of the Clifford group cannot be currently achieved in any realistic physical system with a topological order. What universality classes of anyons allow one to implement all Clifford group gates (but do not allow one to simulate UQC) is an interesting open problem.

To fully utilize the potential of our model, we allow *adaptive* computation. It means that a description of an operation to be performed at step t may be a function of all measurement outcomes at steps $1,\ldots,t-1$. (For even greater generality, the dependence may be probabilistic. This assumption does not actually strengthen the model since tossing a fair coin can be simulated using $\mathcal{O}_{\text{ideal}}$) At this point, we need to be careful because the proper choice of operations should not only be defined mathematically—it should be computed by some *efficient algorithm*. In all protocols described below, the algorithms will actually be very simple. (Let us point out that dropping the computational complexity restriction still leaves a nontrivial problem: can we prepare an arbitrary multiqubit pure state with any given fidelity using only operations from the basis $\mathcal{O}$?)

The main question that we address in this paper is as follows: For which density matrices ρ can one efficiently simulate universal quantum computation by adaptive computation in the basis $\mathcal{O}$?

It will be convenient to use the Bloch sphere representation of one-qubit states:

$$\rho = \tfrac{1}{2}(I + \rho_x\sigma^x + \rho_y\sigma^y + \rho_z\sigma^z).$$

The vector (ρ_x,ρ_y,ρ_z) will be referred to as the *polarization vector* of ρ. Let us first consider the subset of states satisfying

$$|\rho_x| + |\rho_y| + |\rho_z| \leqslant 1.$$

This inequality says that the vector (ρ_x,ρ_y,ρ_z) lies inside the octahedron O with vertices (±1, 0, 0), (0, ±1, 0), (0, 0, ±1),

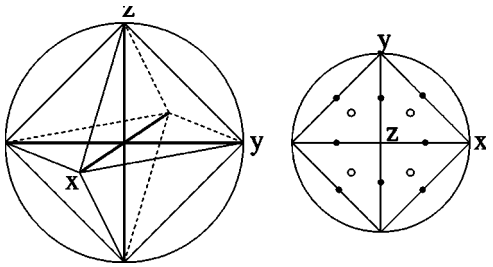


FIG. 1. Left: the Bloch sphere and the octahedron O. Right: the octahedron O projected on the $x-y$ plane. The magic states correspond to the intersections of the symmetry axes of O with the Bloch sphere. The empty and filled circles represent T-type and H-type magic states, respectively.

see Fig. 1. The six vertices of O represent the six eigenstates of the Pauli operators σ^x,σ^y, and σ^z. We can prepare these states by operations from O_{ideal} only. Since ρ is a convex linear combination (probabilistic mixture) of these states, we can prepare ρ by operations from $\mathcal{O}_{\text{ideal}}$ and by tossing a coin with suitable weights. Thus we can rephrase the Gottesman-Knill theorem in the following way.

Theorem 1. Suppose the polarization vector (ρ_x,ρ_y,ρ_z) of the state ρ belongs to the convex hull of (±1, 0, 0), (0, ±1, 0), (0, 0, ±1). Then any adaptive computation in the basis $\mathcal{O}$ can be efficiently simulated on a classical probabilistic computer.

This observation leads naturally to the following question: is it true that UQC can be efficiently simulated whenever ρ lies in the exterior of the octahedron O? In an attempt to provide at least a partial answer, we prove the universality for a large set of states. Specifically, we construct two particular schemes of UQC simulation based on a method which we call *magic states distillation*. Let us start by defining the magic states.

Definition 1. Consider pure states $|H\rangle,|T\rangle \in \mathbb{C}^2$ such that

$$|T\rangle\langle T| = \frac{1}{2}\left[I + \frac{1}{\sqrt{3}}(\sigma^x + \sigma^y + \sigma^z)\right],$$

and

$$|H\rangle\langle H| = \frac{1}{2}\left[I + \frac{1}{\sqrt{2}}(\sigma^x + \sigma^z)\right].$$

The images of $|T\rangle$ and $|H\rangle$ under the action of one-qubit Clifford operators are called magic states of T type and H type, respectively.

[This notation is chosen since $|H\rangle$ and $|T\rangle$ are eigenvectors of certain Clifford group operators: the Hadamard gate H and the operator usually denoted T, see Eq. (7).] Denote the one-qubit Clifford group by $\mathcal{C}_1$. Overall, there are 8 magic states of T type, $\{U|T\rangle, U \in \mathcal{C}_1\}$ (up to a phase) and 12 states of H type, $\{U|H\rangle, U \in \mathcal{C}_1\}$, see Fig. 1. Clearly, the polarization vectors of magic states are in one-to-one correspondence with rotational symmetry axes of the octahedron O (H-type states correspond to 180° rotations and T-type states correspond to 120° rotations). The role of magic states in our construction is twofold. First, adaptive computation in the basis $\mathcal{O}_{\text{ideal}}$ together with the preparation of magic states (of either type) allows one to simulate UQC (see Sec. III). Second, by adap-

tive computation in the basis $\mathcal{O}_{\text{ideal}}$ one can "purify" imperfect magic states. It is a rather surprising coincidence that one and the same state can comprise both of these properties, and that is the reason why we call them magic states.

More exactly, a magic state distillation procedure yields one copy of a magic state (with any desired fidelity) from several copies of the state ρ, provided that the initial fidelity between ρ and the magic state to be distilled is large enough. In the course of distillation, we use only operations from the set $\mathcal{O}_{\text{ideal}}$. By constructing two particular distillation schemes, for T-type and H-type magic states, respectively, we prove the following theorems.

Theorem 2. Let $F_T(\rho)$ be the maximum fidelity between ρ and a T-type magic state, i.e.,

$$F_T(\rho) = \max_{U \in \mathcal{C}_1} \sqrt{\langle T|U^\dagger \rho U|T\rangle}.$$

Adaptive computation in the basis $\mathcal{O}=\mathcal{O}_{\text{ideal}} \cup \{\rho\}$ allows one to simulate universal quantum computation whenever

$$F_T(\rho) > F_T = \left[\frac{1}{2}\left(1+\sqrt{\frac{3}{7}}\right)\right]^{1/2} \approx 0.910.$$

Theorem 3. Let $F_H(\rho)$ be the maximum fidelity between ρ and an H-type magic state,

$$F_H(\rho) = \max_{U \in \mathcal{C}_1} \sqrt{\langle H|U^\dagger \rho U|H\rangle}.$$

Adaptive computation in the basis $\mathcal{O}=\mathcal{O}_{\text{ideal}} \cup \{\rho\}$ allows one to simulate universal quantum computation whenever

$$F_H(\rho) > F_H \approx 0.927.$$

The quantities F_T and F_H have the meaning of threshold fidelity since our distillation schemes increase the polarization of ρ, converging to a magic state as long as the inequalities $F_T(\rho) > F_T$ or $F_H(\rho) > F_H$ are fulfilled. If they are not fulfilled, the process converges to the maximally mixed state. The conditions stated in the theorems can also be understood in terms of the polarization vector (ρ_x, ρ_y, ρ_z). Indeed, let us associate a "magic direction" with each of the magic states. Then Theorems 2 and 3 say that the distillation is possible if there is a T direction such that the projection of the vector (ρ_x, ρ_y, ρ_z) onto that T direction exceeds the threshold value of $2F_T^2 - 1 \approx 0.655$, or if the projection on some of the H directions is greater than $2F_H^2 - 1 \approx 0.718$.

Let us remark that, although the proposed distillation schemes are probably not optimal, the threshold fidelities F_T and F_H cannot be improved significantly. Indeed, it is easy to check that the octahedron O corresponding to probabilistic mixtures of stabilizer states can be defined as

$$\mathcal{O} = \{\rho : F_T(\rho) \leqslant F_T^*\},$$

where

$$F_T^* = \left[\frac{1}{2}\left(1+\sqrt{\frac{1}{3}}\right)\right]^{1/2} \approx 0.888.$$

It means that F_T^* is a lower bound on the threshold fidelity F_T for any protocol distilling T-type magic states. Thus any potential improvement to Theorem 2 may only decrease F_T from 0.910 down to $F_T^*=0.888$. From a practical perspective, the difference between these two numbers is not important.

On the other hand, such an improvement would be of great theoretical interest. Indeed, if Theorem 2 with F_T replaced by F_T^* is true, it would imply that the Gottesman-Knill theorem provides necessary and sufficient conditions for the classical simulation, and that a transition from classical to universal quantum behavior occurs at the boundary of the octahedron O. This kind of transition has been discussed in context of a general error model [21]. Our model is simpler, which gives hope for sharper results.

By the same argument, one can show that the quantity

$$F_H^* \overset{\text{def}}{=} \max_{\rho \in O} \sqrt{\langle H|\rho|H\rangle} = \left[\frac{1}{2}\left(1+\sqrt{\frac{1}{2}}\right)\right]^{1/2} \approx 0.924$$

is a lower bound on the threshold fidelity F_H for any protocol distilling H-type magic states.

A similar approach to UQC simulation was suggested in Ref. [22], where Clifford group operations were used to distill the entangled three-qubit state $|000\rangle+|001\rangle+|010\rangle+|100\rangle$, which is necessary for the realization of the Toffoli gate.

The rest of the paper is organized as follows. Section II contains some well-known facts about the Clifford group and stabilizer formalism, which will be used throughout the paper. In Sec. III we prove that magic states together with operations from $\mathcal{O}_{\text{ideal}}$ are sufficient for UQC. In Sec. IV ideal magic are substituted by faulty ones and the error rate that our simulation algorithm can tolerate is estimated. In Sec. V we describe a distillation protocol for T-type magic states. This protocol is based on the well-known five-qubit quantum code. In Sec. VI a distillation protocol for H-type magic states is constructed. It is based on a certain CSS stabilizer code that encodes one qubit into 15 and admits a nontrivial automorphism [23]. Specifically, the bitwise application of a certain *non-Clifford* unitary operator preserves the code subspace and effects the same operator on the encoded qubit. We conclude with a brief summary and a discussion of open problems.

II. CLIFFORD GROUP, STABILIZERS, AND SYNDROME MEASUREMENTS

Let $\mathcal{C}_n$ denote the n-qubit *Clifford group*. Recall that it is a finite subgroup of $\mathrm{U}(2^n)$ generated by the Hadamard gate H (applied to any qubit), the phase-shift gate K (applied to any qubit), and the controlled-not gate $\Lambda(\sigma^x)$ (which may be applied to any pair qubits),

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad K = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}, \quad \Lambda(\sigma^x) = \begin{pmatrix} I & 0 \\ 0 & \sigma^x \end{pmatrix}. \tag{1}$$

The Pauli operators $\sigma^x, \sigma^y, \sigma^z$ belong to $\mathcal{C}_1$, for instance, $\sigma^z = K^2$ and $\sigma^x = HK^2H$. The *Pauli group* $P(n) \subset \mathcal{C}_n$ is generated by the Pauli operators acting on n qubits. It is known [24] that the Clifford group $\mathcal{C}_n$ augmented by scalar unitary operators $e^{i\varphi}I$ coincides with the normalizer of $P(n)$ in the uni-

tary group $U(2^n)$. Hermitian elements of the Pauli group are of particular importance for quantum error correction theory; they are referred to as *stabilizers*. These are operators of the form

$$\pm\sigma^{\alpha_1}\otimes\cdots\otimes\sigma^{\alpha_n},\quad \alpha_j\in\{0,x,y,z\},$$

where $\sigma^0=I$. Let us denote by $S(n)$ the set of all n-qubit stabilizers:

$$S(n)=\{S\in P(n)\ :\ S^\dagger=S\}.$$

For any two stabilizers S_1,S_2 we have $S_1S_2=\pm S_2S_1$ and $S_1^2=S_2^2=I$. It is known that for any set of pairwise commuting stabilizers $S_1,\ldots,S_k\in S(n)$ there exists a unitary operator $V\in\mathcal{C}_n$ such that

$$VS_jV^\dagger=\sigma^z[j],\quad j=1,\ldots,k,$$

where $\sigma^z[j]$ denotes the operator σ^z applied to the jth qubit, e.g., $\sigma^z[1]=\sigma^z\otimes I\otimes\cdots\otimes I$.

These properties of the Clifford group allow us to introduce a very useful computational procedure which can be realized by operations from $\mathcal{O}_{\text{ideal}}$. Specifically, we can perform a joint nondestructive eigenvalue measurement for any set of pairwise commuting stabilizers $S_1,\ldots,S_k\in S(n)$. The outcome of such a measurement is a sequence of eigenvalues $\lambda=(\lambda_1,\ldots,\lambda_k)$, $\lambda_j=\pm1$, which is usually called a *syndrome*. For any given outcome, the quantum state is acted upon by the projector

$$\Pi_\lambda=\prod_{j=1}^{k}\frac{1}{2}(I+\lambda_jS_j).$$

Now, let us consider a computation that begins with an arbitrary state and consists of operations from $\mathcal{O}_{\text{ideal}}$. It is clear that we can defer all Clifford operations until the very end if we replace the Pauli measurements by general syndrome measurements. Thus the most general transformation that can be realized by $\mathcal{O}_{\text{ideal}}$ is an *adaptive syndrome measurement*, meaning that the choice of the stabilizer S_j to be measured next depends on the previously measured values of $\lambda_1,\ldots,\lambda_{j-1}$. In general, this dependence may involve coin tossing. Without loss of generality one can assume that S_j commutes with all previously measured stabilizers $S_1,\ldots,S_{j-1}$ (for all possible values of $\lambda_1,\ldots,\lambda_{j-1}$ and coin tossing outcomes). Adaptive syndrome measurement has been used in Ref. [25] to distill entangled states of a bipartite system by local operations.

III. UNIVERSAL QUANTUM COMPUTATION WITH MAGIC STATES

In this section, we show that operations from $\mathcal{O}_{\text{ideal}}$ are sufficient for universal quantum computation if a supply of *ideal* magic states is also available. First, consider a one-qubit state

$$|A_\theta\rangle=2^{-1/2}(|0\rangle+e^{i\theta}|1\rangle) \tag{2}$$

and suppose that θ is not a multiple of $\pi/2$. We now describe a procedure that implements the phase shift gate

$$\Lambda(e^{i\theta})=\begin{pmatrix}1&0\\0&e^{i\theta}\end{pmatrix}$$

by consuming several copies of $|A_\theta\rangle$ and using only operations from $\mathcal{O}_{\text{ideal}}$.

Let $|\psi\rangle=a|0\rangle+b|1\rangle$ be the unknown initial state which should be acted on by $\Lambda(e^{i\theta})$. Prepare the state $|\Psi_0\rangle=|\psi\rangle\otimes|A_\theta\rangle$ and measure the stabilizer $S_1=\sigma^z\otimes\sigma^z$. Note that both outcomes of this measurement appear with probability 1/2. If the outcome is "+1", we are left with the state

$$|\Psi_1^+\rangle=(a|0,0\rangle+be^{i\theta}|1,1\rangle).$$

In the case of "−1" outcome, the resulting state is

$$|\Psi_1^-\rangle=(ae^{i\theta}|0,1\rangle+b|1,0\rangle).$$

Let us apply the gate $\Lambda(\sigma^x)[1,2]$ (the first qubit is the control one). The above two states are mapped to

$$|\Psi_2^+\rangle=\Lambda(\sigma^x)[1,2]|\Psi_1^+\rangle=(a|0\rangle+be^{i\theta}|1\rangle)\otimes|0\rangle,$$

$$|\Psi_2^-\rangle=\Lambda(\sigma^x)[1,2]|\Psi_1^-\rangle=(ae^{i\theta}|0\rangle+b|1\rangle)\otimes|1\rangle.$$

Now the second qubit can be discarded, and we are left with the state $a|0\rangle+be^{\pm i\theta}|1\rangle$, depending upon the measured eigenvalue. Thus the net effect of this circuit is the application of a unitary operator that is chosen randomly between $\Lambda(e^{i\theta})$ and $\Lambda(e^{-i\theta})$ (and we know which of the two possibilities has occurred).

Applying the circuit repeatedly, we effect the transformations $\Lambda(e^{ip_1\theta})$, $\Lambda(e^{ip_2\theta})$,... for some integers $p_1,p_2,\ldots$ which obey the random-walk statistics. It is well known that such a random walk visits each integer with the probability 1. It means that sooner or later we will get $p_k=1$ and thus realize the desired operator $\Lambda(e^{i\theta})$. The probability that we will need more than N steps to succeed can be estimated as $cN^{-1/2}$ for some constant $c>0$. Note also that if θ is a rational multiple of 2π, we actually have a random walk on a cyclic group $\mathbb{Z}_q$. In this case, the probability that we will need more than N steps decreases exponentially with N.

The magic state $|H\rangle$ can be explicitly written in the standard basis as

$$|H\rangle=\cos\left(\frac{\pi}{8}\right)|0\rangle+\sin\left(\frac{\pi}{8}\right)|1\rangle. \tag{3}$$

Note that $HK|H\rangle=e^{i\pi/8}|A_{-\pi/4}\rangle$. So if we are able to prepare the state $|H\rangle$, we can realize the operator $\Lambda(e^{-i\pi/4})$. It does not belong to the Clifford group. Moreover, the subgroup of U(2) generated by $\Lambda(e^{-i\pi/4})$ and $\mathcal{C}_1$ is dense in U(2). [1] Thus the operators from $\mathcal{C}_1$ and $\mathcal{C}_2$ together with $\Lambda(e^{-i\pi/4})$ constitute a universal basis for quantum computation.

The magic state $|T\rangle$ can be explicitly written in the standard basis:

[1] Recall that the action of the Clifford group $\mathcal{C}_1$ on the set of operators $\pm\sigma^x$, $\pm\sigma^y$, $\pm\sigma^z$ coincides with the action of rotational symmetry group of a cube on the set of unit vectors $\pm e_x$, $\pm e_y$, $\pm e_z$, respectively.

$$|T\rangle = \cos\beta|0\rangle + e^{i(\pi/4)}\sin\beta|1\rangle, \quad \cos(2\beta) = \frac{1}{\sqrt{3}}. \qquad (4)$$

Let us prepare an initial state $|\Psi_0\rangle = |T\rangle \otimes |T\rangle$ and measure the stabilizer $S_1 = \sigma^z \otimes \sigma^z$. The outcome +1 appears with probability $p_+ = \cos^4\beta + \sin^4\beta = 2/3$. If the outcome is −1, we discard the reduced state and try again, using a fresh pair of magic states. (On average, we need three copies of the $|T\rangle$ state to get the outcome +1.) The reduced state corresponding to the outcome +1 is

$$|\Psi_1\rangle = \cos\gamma|0,0\rangle + i\sin\gamma|1,1\rangle, \quad \gamma = \frac{\pi}{12}.$$

Let us apply the gate $\Lambda(\sigma^x)[1,2]$ and discard the second qubit. We arrive at the state

$$|\Psi_2\rangle = \cos\gamma|0\rangle + i\sin\gamma|1\rangle.$$

Next apply the Hadamard gate H:

$$|\Psi_3\rangle = H|\Psi_2\rangle = 2^{-1/2}e^{i\gamma}(|0\rangle + e^{-2i\gamma}|1\rangle) = |A_{-\pi/6}\rangle.$$

We can use this state as described above to realize the operator $\Lambda(e^{-i\pi/6})$. It is easy to check that Clifford operators together with $\Lambda(e^{-i\pi/6})$ constitute a universal set of unitary gates.

Thus we have proved that the sets of operations $\mathcal{O}_{\text{ideal}} \cup \{|H\rangle\}$ and $\mathcal{O}_{\text{ideal}} \cup \{|T\rangle\}$ are sufficient for universal quantum computation.

IV. ERROR ANALYSIS

To establish a connection between the simulation algorithms described in Sec. III and the universality theorems stated in the introduction we have to substitute *ideal* magic states by *faulty* ones. Before doing that let us discuss the ideal case in more detail. Suppose that a quantum circuit to be simulated uses a gate basis in which the only non-Clifford gate is the phase shift $\Lambda(e^{-i\pi/4})$ or $\Lambda(e^{-i\pi/6})$. One can apply the algorithm of Sec. III to simulate each non-Clifford gate independently. To avoid fluctuations in the number of magic states consumed at each round, let us set a limit of K magic states per round, where K is a parameter to be chosen later. As was pointed out in Sec. III, the probability for some particular simulation round to "run out of budget" scales as $\exp(-\alpha K)$ for some constant $\alpha > 0$. If at least one simulation round runs out of budget, we declare a failure and the whole simulation must be aborted. Denote the total number of non-Clifford gates in the circuit by L. The probability p_a for the whole simulation to be aborted can be estimated as

$$p_a \sim 1 - [1 - \exp(-\alpha K)]^L \sim L\exp(-\alpha K) \ll 1,$$

provided that $L\exp(-\alpha K) \ll 1$. We will assume

$$K \gtrsim \alpha^{-1}\ln L,$$

so the abort probability can be neglected.

Each time the algorithm requests an ideal magic state, it actually receives a slightly nonideal one. Such nearly perfect magic states must be prepared using the distillation methods described in Secs. V and VI. Let us estimate an affordable error rate ϵ_{out} for *distilled* magic states. Since there are L non-Clifford gates in the circuit, one can tolerate an error rate of the order $1/L$ in implementation of these gates.[2] Each non-Clifford gate requires $K \sim \ln L$ magic states. Thus the whole simulation is reliable enough if one chooses

$$\epsilon_{\text{out}} \sim 1/(L\ln L). \qquad (5)$$

What are the resources needed to distill one copy of a magic state with the error rate ϵ_{out}? To be more specific, let us talk about H-type states. It will be shown in Sec. VI that the number n of raw (undistilled) ancillas needed to distill one copy of the $|H\rangle$ magic state with an error rate not exceeding ϵ_{out} scales as

$$n \sim [\ln(1/\epsilon_{\text{out}})]^\gamma, \quad \gamma = \log_3 15 \approx 2.5,$$

see Eq. (39). Taking ϵ_{out} from Eq. (5), one gets

$$n \sim (\ln L)^\gamma.$$

Since the whole simulation requires $KL \sim L\ln L$ copies of the distilled $|H\rangle$ state, we need

$$N \sim L(\ln L)^{\gamma+1}$$

raw ancillas overall.

Summarizing, the simulation theorems stated in the introduction follow from the following results (the last one will be proved later):

(i) the circuits described in Sec. III allow one to simulate UQC with the sets of operations $\mathcal{O}_{\text{ideal}} \cup \{|H\rangle\}$ and $\mathcal{O}_{\text{ideal}} \cup \{|T\rangle\}$;

(ii) these circuits work reliably enough if the states $|H\rangle$ and $|T\rangle$ are slightly noisy, provided that the error rate does not exceed $\epsilon_{\text{out}} \sim 1/(L\ln L)$;

(iii) a magic state having an error rate ϵ_{out} can be prepared from copies of the raw ancillary state ρ using the distillation schemes provided that $F_T(\rho) > F_T$ or $F_H(\rho) > F_H$. The distillation requires resources that are polynomial in $\ln L$.

V. DISTILLATION OF *T*-TYPE MAGIC STATES

Suppose we are given n copies of a state ρ, and our goal is to distill one copy of the magic state $|T\rangle$. The polarization vector of ρ can be brought into the positive octant of the Bloch space by a Clifford group operator, so we can assume that

$$\rho_x, \rho_y, \rho_z \geqslant 0.$$

In this case, the fidelity between ρ and $|T\rangle$ is the largest one among all T-type magic states, i.e.,

$$F_T(\rho) = \sqrt{\langle T|\rho|T\rangle}.$$

A related quantity,

[2] This fault tolerance does not require any redundancy in the implementation of the circuit (e.g., the use of concatenated codes). It is achived automatically because in the worst case the error probability accumulates linearly in the number of gates. In our model only non-Clifford gates are faulty.

$$\epsilon = 1 - \langle T|\rho|T\rangle = \frac{1}{2}\left[1 - \frac{1}{\sqrt{3}}(\rho_x + \rho_y + \rho_z)\right],$$

will be called the *initial error probability*. By definition, $0 \leqslant \epsilon \leqslant 1/2$.

The output of the distillation algorithm will be some one-qubit mixed state ρ_{out}. To quantify the proximity between ρ_{out} and $|T\rangle$, let us define a *final error probability*:

$$\epsilon_{\text{out}} = 1 - \langle T|\rho_{\text{out}}|T\rangle.$$

It will be certain function of n and ϵ. The asymptotic behavior of this function for $n \to \infty$ reveals the existence of a *threshold error probability*,

$$\epsilon_0 = \frac{1}{2}\left(1 - \sqrt{\frac{3}{7}}\right) \approx 0.173,$$

such that for $\epsilon < \epsilon_0$ the function $\epsilon_{\text{out}}(n,\epsilon)$ converges to zero. We will see that for small ϵ,

$$\epsilon_{\text{out}}(n,\epsilon) \sim (5\epsilon)^{n^\xi}, \quad \xi = 1/\log_2 30 \approx 0.2. \tag{6}$$

On the other hand, if $\epsilon > \epsilon_0$, the output state converges to the maximally mixed state, i.e., $\lim_{n\to\infty}\epsilon_{\text{out}}(n,\epsilon) = 1/2$.

Before coming to a detailed description of the distillation algorithm, let us outline the basic ideas involved in its construction. The algorithm recursively iterates an elementary distillation subroutine that transforms five copies of an imperfect magic state into one copy having a smaller error probability. This elementary subroutine involves a syndrome measurement for certain commuting stabilizers $S_1, S_2, S_3, S_4 \in S(5)$. If the measured syndrome $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ is nontrivial ($\lambda_j = -1$ for some j), the distillation attempt fails and the reduced state is discarded. If the measured syndrome is trivial ($\lambda_j = 1$ for all j), the distillation attempt is successful. Applying a decoding transformation (a certain Clifford operator) to the reduced state, we transform it to a single-qubit state. This qubit is the output of the subroutine.

Our construction is similar to concatenated codes used in many fault-tolerant quantum computation techniques, but it differs from them in two respects. First, we do not need to *correct* errors—it suffices only to *detect* them. Once an error has been detected, we simply discard the reduced state, since it does not contain any valuable information. This allows us to achieve higher threshold error probability. Second, we do not use quantum codes in the way for which they were originally designed: in our scheme, the syndrome is measured on a product state.

The state $|T\rangle$ is an eigenstate for the unitary operator

$$T = e^{i\pi/4}KH = \frac{e^{i\pi/4}}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \in \mathcal{C}_1. \tag{7}$$

Note that T acts on the Pauli operators as follows:[3]

$$T\sigma^x T^\dagger = \sigma^z, \quad T\sigma^z T^\dagger = \sigma^y, \quad T\sigma^y T^\dagger = \sigma^x. \tag{8}$$

We will denote its eigenstates by $|T_0\rangle$ and $|T_1\rangle$, so that

$$T|T_0\rangle = e^{+i\pi/3}|T_0\rangle, \quad T|T_1\rangle = e^{-i\pi/3}|T_1\rangle,$$

$$|T_{0,1}\rangle\langle T_{0,1}| = \frac{1}{2}\left[I \pm \frac{1}{\sqrt{3}}(\sigma^x + \sigma^y + \sigma^z)\right].$$

Note that $|T_0\rangle \overset{\text{def}}{=} |T\rangle$ and $|T_1\rangle = \sigma^y H|T_0\rangle$ are T-type magic states.

Let us apply a dephasing transformation,

$$D(\eta) = \frac{1}{3}(\eta + T\eta T^\dagger + T^\dagger \eta T), \tag{9}$$

to each copy of the state ρ. The transformation D can be realized by applying one of the operators I, T, T^{-1} chosen with probability $1/3$ each. Since

$$D(|T_0\rangle\langle T_1|) = D(|T_1\rangle\langle T_0|) = 0,$$

we have

$$D(\rho) = (1-\epsilon)|T_0\rangle\langle T_0| + \epsilon|T_1\rangle\langle T_1|. \tag{10}$$

We will assume that the dephasing transformation is applied at the very first step of the distillation, so ρ has the form (10). Thus the initial state for the elementary distillation subroutine is

$$\rho_{\text{in}} = \rho^{\otimes 5} = \sum_{x\in\{0,1\}^5} \epsilon^{|x|}(1-\epsilon)^{5-|x|}|T_x\rangle\langle T_x|, \tag{11}$$

where $x = (x_1, \ldots, x_5)$ is a binary string, $|x|$ is the number of 1's in x, and

$$|T_x\rangle \overset{\text{def}}{=} |T_{x_1}\rangle \otimes \cdots \otimes |T_{x_5}\rangle.$$

The stabilizers $S_1, \ldots, S_4$ to be measured on the state ρ_{in} correspond to the famous five-qubit code, see Refs. [26,27]. They are defined as follows:

$$S_1 = \sigma^x \otimes \sigma^z \otimes \sigma^z \otimes \sigma^x \otimes I,$$

$$S_2 = I \otimes \sigma^x \otimes \sigma^z \otimes \sigma^z \otimes \sigma^x,$$

$$S_3 = \sigma^x \otimes I \otimes \sigma^x \otimes \sigma^z \otimes \sigma^z,$$

$$S_4 = \sigma^z \otimes \sigma^x \otimes I \otimes \sigma^x \otimes \sigma^z. \tag{12}$$

This code has a cyclic symmetry, which becomes explicit if we introduce an auxiliary stabilizer, $S_5 = S_1S_2S_3S_4 = \sigma^z \otimes \sigma^z \otimes \sigma^x \otimes I \otimes \sigma^x$. Let $\mathcal{L}$ be the two-dimensional code subspace specified by the conditions $S_j|\Psi\rangle = |\Psi\rangle$, $j = 1, \ldots, 4$, and Π be the orthogonal projector onto $\mathcal{L}$:

$$\Pi = \frac{1}{16}\prod_{j=1}^{4}(I + S_j). \tag{13}$$

It was pointed out in Ref. [16] that the operators

$$\hat{X} = (\sigma^x)^{\otimes 5}, \quad \hat{Y} = (\sigma^y)^{\otimes 5}, \quad \hat{Z} = (\sigma^z)^{\otimes 5},$$

and

[3]The operator denoted by T in Ref. [16] does not coincide with our T. They are related by the substitution $T \to e^{-i\pi/4}T^\dagger$ though.

$$\hat{T} = (T)^{\otimes 5} \tag{14}$$

commute with Π, thus preserving the code subspace. Moreover, $\hat{X},\hat{Y},\hat{Z}$ obey the same algebraic relations as one-qubit Pauli operators, e.g., $\hat{X}\hat{Y}=i\hat{Z}$. Let us choose a basis in $\mathcal{L}$ such that $\hat{X},\hat{Y}$, and $\hat{Z}$ become logical Pauli operators σ^x,σ^y, and σ^z, respectively. How does the operator $\hat{T}$ act in this basis? From Eq. (8) we immediately get

$$\hat{T}\hat{X}\hat{T}^\dagger = \hat{Z}, \quad \hat{T}\hat{Z}\hat{T}^\dagger = \hat{Y}, \quad \hat{T}\hat{Y}\hat{T}^\dagger = \hat{X}.$$

Therefore $\hat{T}$ coincides with the logical operator T up to an overall phase factor. This factor is fixed by the condition that the logical T has eigenvalues $e^{\pm i(\pi/3)}$.

Let us find the eigenvectors of $\hat{T}$ that belong to $\mathcal{L}$. Consider two particular states from $\mathcal{L}$, namely

$$|T_1^L\rangle = \sqrt{6}\Pi|T_{00000}\rangle, \quad \text{and } |T_0^L\rangle = \sqrt{6}\Pi|T_{11111}\rangle.$$

In the Appendix we show that

$$\langle T_{00000}|\Pi|T_{00000}\rangle = \langle T_{11111}|\Pi|T_{11111}\rangle = \frac{1}{6}, \tag{15}$$

so that the states $|T_0^L\rangle$ and $|T_1^L\rangle$ are normalized. Taking into account that $[\hat{T},\Pi]=0$ and that

$$\hat{T}|T_x\rangle = e^{i(\pi/3)(5-2|x|)}|T_x\rangle \text{ for all } x \in \{0,1\}^5, \tag{16}$$

we get

$$\hat{T}|T_1^L\rangle = \sqrt{6}\hat{T}\Pi|T_{00000}\rangle = \sqrt{6}\Pi\hat{T}|T_{00000}\rangle = e^{-i\pi/3}|T_1^L\rangle.$$

Analogously, one can check that

$$\hat{T}|T_0^L\rangle = e^{+i\pi/3}|T_0^L\rangle.$$

It follows that $\hat{T}$ is exactly the logical operator T, including the overall phase, and $|T_0^L\rangle$ and $|T_1^L\rangle$ are the logical states $|T_0\rangle$ and $|T_1\rangle$ (up to some phase factors, which are not important for us). Therefore we have

$$|T_{0,1}^L\rangle\langle T_{0,1}^L| = \Pi\frac{1}{2}\left[I \pm \frac{1}{\sqrt{3}}(\hat{X}+\hat{Y}+\hat{Z})\right]. \tag{17}$$

Now we are in a position to describe the syndrome measurement performed on the state ρ_{in}. The unnormalized reduced state corresponding to the trivial syndrome is as follows:

$$\rho_s = \Pi\rho_{\text{in}}\Pi = \sum_{x\in\{0,1\}^5} \epsilon^{|x|}(1-\epsilon)^{5-|x|}\Pi|T_x\rangle\langle T_x|\Pi, \tag{18}$$

see Eq. (11). The probability for the trivial syndrome to be observed is

$$p_s = \text{Tr}\,\rho_s.$$

Note that the state $\Pi|T_x\rangle$ is an eigenvector of $\hat{T}$ for any $x \in \{0,1\}^5$. But we know that the restriction of $\hat{T}$ on $\mathcal{L}$ has eigenvalues $e^{\pm i\pi/3}$. At the same time, Eq. (16) implies that

$$\hat{T}\Pi|T_x\rangle = -\Pi|T_x\rangle$$

whenever $|x|=1$ or $|x|=4$. This eigenvalue equation is not a contradiction only if

$$\Pi|T_x\rangle = 0 \text{ for } |x| = 1,4.$$

This equality can be interpreted as an error correction property. Indeed, the initial state ρ_{in} is a mixture of the desired state $|T_{00000}\rangle$ and unwanted states $|T_x\rangle$ with $|x|>0$. We can interpret the number of "1" components in x as a number of errors. Once the trivial syndrome has been measured, we can be sure that either no errors or at least two errors have occurred. Such error correction, however, is not directly related to the minimal distance of the code.

It follows from Eq. (16) that for $|x|=2$, 3 one has $\hat{T}\Pi|T_x\rangle = e^{\pm i\pi/3}\Pi|T_x\rangle$, so that $\Pi|T_x\rangle$ must be proportional to one of the states $|T_0^L\rangle$, $|T_1^L\rangle$. Our observations can be summarized as follows:

$$\Pi|T_x\rangle = \begin{cases} 6^{-1/2}|T_1^L\rangle, & \text{if } |x|=0, \\ 0, & \text{if } |x|=1, \\ a_x|T_0^L\rangle, & \text{if } |x|=2, \\ b_x|T_1^L\rangle, & \text{if } |x|=3, \\ 0, & \text{if } |x|=4, \\ 6^{-1/2}|T_0^L\rangle, & \text{if } |x|=5. \end{cases} \tag{19}$$

Here the coefficients a_x, b_x depend upon x in some way. The output state (18) can now be written as

$$\rho_s = \left[\frac{1}{6}\epsilon^5 + \epsilon^2(1-\epsilon)^3\sum_{x:|x|=2}|a_x|^2\right]|T_0^L\rangle\langle T_0^L| + \left[\frac{1}{6}(1-\epsilon)^5 + \epsilon^3(1-\epsilon)^2\sum_{x:|x|=3}|b_x|^2\right]|T_1^L\rangle\langle T_1^L|. \tag{20}$$

To exclude the unknown coefficients a_x and b_x, we can use the identity

$$|T_0^L\rangle\langle T_0^L| + |T_1^L\rangle\langle T_1^L| = \Pi = \sum_{x\in\{0,1\}^5}\Pi|T_x\rangle\langle T_x|\Pi.$$

Substituting Eq. (19) into this identity, we get

$$\sum_{x:|x|=2}|a_x|^2 = \sum_{x:|x|=3}|b_x|^2 = \frac{5}{6}.$$

So the final expression for the output state ρ_s is as follows:

$$\rho_s = \left[\frac{\epsilon^5 + 5\epsilon^2(1-\epsilon)^3}{6}\right]|T_0^L\rangle\langle T_0^L| + \left[\frac{(1-\epsilon)^5 + 5\epsilon^3(1-\epsilon)^2}{6}\right] \times |T_1^L\rangle\langle T_1^L|. \tag{21}$$

Accordingly, the probability to observe the trivial syndrome is

$$p_s = \frac{\epsilon^5 + 5\epsilon^2(1-\epsilon)^3 + 5\epsilon^3(1-\epsilon)^2 + (1-\epsilon)^5}{6}. \tag{22}$$

A decoding transformaion for the five-qubit code is a unitary operator $V \in \mathcal{C}_5$ such that

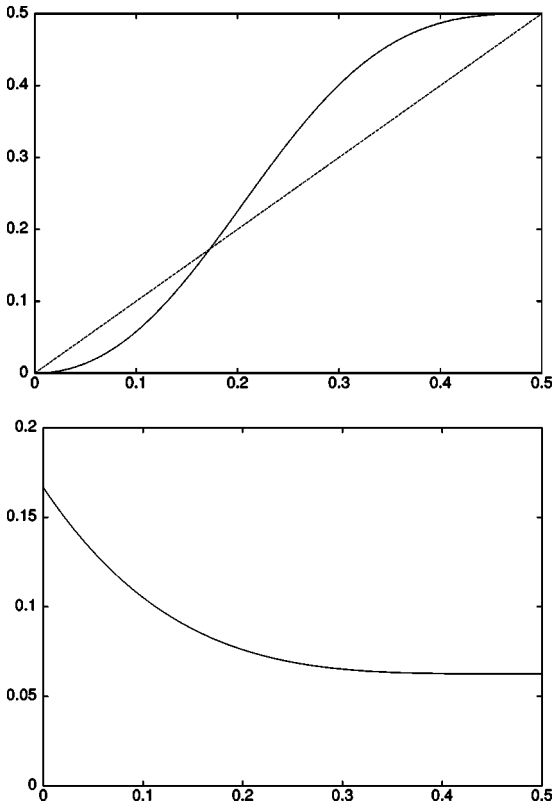


FIG. 2. The final error probability ϵ_{out} and the probability p_s to measure the trivial syndrome as functions of the initial error probability ϵ for the T-type states distillation.

$$V\mathcal{L} = \mathbb{C}^2 \otimes |0,0,0,0\rangle.$$

In other words, V maps the stabilizers S_j, $j=2$, 3, 4, 5 to $\sigma^z[j]$. The logical operators $\hat{X},\hat{Y},\hat{Z}$ are mapped to the Pauli operators $\sigma^x,\sigma^y,\sigma^z$ acting on the first qubit. From Eq. (17) we infer that

$$V|T_{0,1}^L\rangle = |T_{0,1}\rangle \otimes |0,0,0,0\rangle$$

(maybe up to some phase). The decoding should be followed by an additional operator $A=\sigma^y H \in \mathcal{C}_1$, which swaps the states $|T_0\rangle$ and $|T_1\rangle$ (note that for small ϵ the state ρ_s is close to $|T_1^L\rangle$, while our goal is to distill $|T_0\rangle$). After that we get a normalized output state

$$\rho_{\text{out}} = (1-\epsilon_{\text{out}})|T_0\rangle\langle T_0| + \epsilon_{\text{out}}|T_1\rangle\langle T_1|,$$

where

$$\epsilon_{\text{out}} = \frac{t^5+5t^2}{1+5t^2+5t^3+t^5}, \quad t=\frac{\epsilon}{1-\epsilon}. \tag{23}$$

The plot of the function $\epsilon_{\text{out}}(\epsilon)$ is shown on Fig. 2. It indicates that the equation $\epsilon_{\text{out}}(\epsilon)=\epsilon$ has only one nontrivial solution, $\epsilon=\epsilon_0\approx 0.173$. The exact value is

$$\epsilon_0 = \frac{1}{2}\left(1-\sqrt{\frac{3}{7}}\right).$$

If $\epsilon<\epsilon_0$, we can recursively iterate the elementary distillation subroutine to produce as good an approximation to the state $|T_0\rangle$ as we wish. On the other hand, if $\epsilon>\epsilon_0$, the distillation subroutine increases the error probability and iterations converge to the maximally mixed state. Thus ϵ_0 is a threshold error probability for our scheme. The corresponding threshold polarization is $1-2\epsilon_0=\sqrt{3/7}\approx 0.655$. For a sufficiently small ϵ, one can use the approximation $\epsilon_{\text{out}}(\epsilon) \approx 5\epsilon^2$.

The probability $p_s=p_s(\epsilon)$ to measure the trivial syndrome decreases monotonically from $1/6$ for $\epsilon=0$ to $1/16$ for $\epsilon=1/2$, see Fig. 2. In the asymptotic regime where ϵ is small, we can use the approximation $p_s\approx p_s(0)=1/6$.

Now the construction of the whole distillation scheme is straightforward. We start from $n\gg 1$ copies of the state $\rho=(1-\epsilon)|T_0\rangle\langle T_0|+\epsilon|T_1\rangle\langle T_1|$. Let us split these states into groups containing five states each and apply the elementary distillation subroutine described above to each group independently. In some of these groups the distillation attempt fails, and the outputs of such groups must be discarded. The average number of "successful" groups is obviously $p_s(\epsilon)\times(n/5)\approx n/30$ if ϵ is small. Neglecting the fluctuations of this quantity, we can say that our scheme provides a constant *yield* $r=1/30$ of output states that are characterized by the error probability $\epsilon_{\text{out}}(\epsilon)\approx 5\epsilon^2$. Therefore we can obtain $r^2 n$ states with $\epsilon_{\text{out}}\approx 5^3\epsilon^4$, $r^3 n$ states with $\epsilon_{\text{out}}\approx 5^7\epsilon^8$, and so on. We have created a hierarchy of states with n states on the first level and four or fewer states on the last level. Let k be the number of levels in this hierarchy and ϵ_{out} the error probability characterizing the states on the last level. Up to small fluctuations, the numbers $n,k,\epsilon_{\text{out}}$, and ϵ are related by the following obvious equations:

$$\epsilon_{\text{out}} \approx \tfrac{1}{5}(5\epsilon)^{2^k}, \quad r^k n \approx 1. \tag{24}$$

Their solution yields Eq. (6).

VI. DISTILLATION OF *H*-TYPE MAGIC STATES

A distillation scheme for H-type magic states also works by recursive iteration of a certain elementary distillation subroutine based on a syndrome measurement for a suitable stabilizer code. Let us start with introducing some relevant coding theory constructions, which reveal an unusual symmetry of this code and explain why it is particularly useful for H-type magic states distillation.

Let $\mathbb{F}_2^n$ be the n-dimensional binary linear space and A be a one-qubit operator such that $A^2=I$. With any binary vector $u=(u_1,\ldots,u_n)\in\mathbb{F}_2^n$ we associate the n-qubit operator

$$A(u) = A^{u_1}\otimes A^{u_2}\otimes\cdots\otimes A^{u_n}.$$

Let $(u,v)=\Sigma_{i=1}^n u_i v_i \bmod 2$ denote the standard binary inner product. If $\mathcal{L}\subseteq\mathbb{F}_2^n$ is a linear subspace, we denote by $\mathcal{L}^\perp$ the set of vectors which are orthogonal to $\mathcal{L}$. The Hamming weight of a binary vector u is denoted by $|u|$. Finally, $u\cdot v\in\mathbb{F}_2^n$ designates the bitwise product of u and v, i.e., $(u\cdot v)_i=u_i v_i$.

A systematic way of constructing stabilizer codes was suggested by Calderbank, Shor, and Steane, see Refs. [28,29]. Codes that can be described in this way will be referred to as *standard CSS codes*. In addition, we consider

their images under an arbitrary unitary transformation $V \in \mathrm{U}(2)$ applied to every qubit. Such "rotated" codes will be called *CSS codes*.

Definition 2. Consider a pair of one-qubit Hermitian operators A, B such that

$$A^2 = B^2 = I, \quad AB = -BA,$$

and a pair of binary vector spaces $\mathcal{L}_A, \mathcal{L}_B \subseteq \mathbb{F}_2^n$, such that

$$(u,v) = 0 \text{ for all } u \in \mathcal{L}_A, v \in \mathcal{L}_B.$$

A quantum code $\mathrm{CSS}(A, \mathcal{L}_A; B, \mathcal{L}_B)$ is a decomposition

$$(\mathbb{C}^2)^{\otimes n} = \bigoplus_{\mu \in \mathcal{L}_A^*} \bigoplus_{\eta \in \mathcal{L}_B^*} \mathcal{H}(\mu, \eta), \tag{25}$$

where the subspace $\mathcal{H}(\mu, \eta)$ is defined by the conditions

$$A(u)|\Psi\rangle = (-1)^{\mu(u)}|\Psi\rangle, \quad B(v)|\Psi\rangle = (-1)^{\eta(v)}|\Psi\rangle$$

for all $u \in \mathcal{L}_A$ and $v \in \mathcal{L}_B$. The linear functionals μ and η are referred to as A syndrome and B syndrome, respectively. The subspace $\mathcal{H}(0,0)$ corresponding to the trivial syndromes $\mu = \eta = 0$ is called the code subspace.

The subspaces $\mathcal{H}(\mu, \eta)$ are well defined since the operators $A(u)$ and $B(v)$ commute for any $u \in \mathcal{L}_A$ and $v \in \mathcal{L}_B$:

$$A(u)B(v) = (-1)^{(u,v)}B(v)A(u) = B(v)A(u).$$

The number of logical qubits in a CSS code is

$$k = \log_2[\dim \mathcal{H}(0,0)] = n - \dim \mathcal{L}_A - \dim \mathcal{L}_B.$$

Logical operators preserving the subspaces $\mathcal{H}(\mu, \eta)$ can be chosen as

$$\{A(u) \,:\, u \in \mathcal{L}_B^{\perp}/\mathcal{L}_A\} \text{ and } \{B(v) \,:\, v \in \mathcal{L}_A^{\perp}/\mathcal{L}_B\}.$$

(By definition, $\mathcal{L}_A \subseteq \mathcal{L}_B^{\perp}$ and $\mathcal{L}_B \subseteq \mathcal{L}_A^{\perp}$, so the factor spaces are well defined.) In the case where A and B are Pauli operators, we get a standard CSS code. Generally, $A = V\sigma^z V^{\dagger}$ and $B = V\sigma^x V^{\dagger}$ for some unitary operator $V \in \mathrm{SU}(2)$, so an arbitrary CSS code can be mapped to a standard one by a suitable bitwise rotation. By a syndrome measurement for a CSS code we mean a projective measurement associated with the decomposition (25).

Consider a CSS code such that some of the operators $A(u), B(v)$ do not belong to the Pauli group $P(n)$. Let us pose this question: can one perform a syndrome measurement for this code by operations from $\mathcal{O}_{\mathrm{ideal}}$ only? It may seem that the answer is no, because by definition of $\mathcal{O}_{\mathrm{ideal}}$ one cannot measure an eigenvalue of an operator unless it belongs to the Pauli group. Surprisingly, this naive answer is wrong. Indeed, imagine that we have measured part of the operators $A(u)$, $B(v)$ (namely, those that belong to the Pauli group). Now we may restrict the remaining operators to the subspace corresponding to the obtained measurement outcomes. It may happen that the restriction of some unmeasured operator $A(u)$, which does not belong to the Pauli group, coincides with the restriction of some other operator $\widetilde{A}(\widetilde{u}) \in P(n)$. If this is the case, we can safely measure $\widetilde{A}(\widetilde{u})$ instead of $A(u)$. The 15-qubit code that we use for the distillation is actually the simplest (to our knowledge) CSS code exhibiting this strange behavior. We now come to an explicit description of this code.

Consider a function f of four Boolean variables. Denote by $[f] \in \mathbb{F}_2^{15}$ the table of all values of f except $f(0000)$. The table is considered as a binary vector, i.e.,

$$[f] = (f(0001), f(0010), f(0011), \ldots, f(1111)).$$

Let $\mathcal{L}_1$ be the set of all vectors $[f]$, where f is a linear function satisfying $f(0) = 0$. In other words, $\mathcal{L}_1$ is the linear subspace spanned by the four vectors $[x_j]$, $j = 1, 2, 3, 4$ (where x_j is an indicator function for the jth input bit):

$$\mathcal{L}_1 = \text{linear span}([x_1], [x_2], [x_3], [x_4]).$$

Let also $\mathcal{L}_2$ be the set of all vectors $[f]$, where f is a polynomial of degree at most 2 satisfying $f(0) = 0$. In other words, $\mathcal{L}_2$ is the linear subspace spanned by the four vectors $[x_j]$ and the six vectors $[x_i x_j]$:

$$\mathcal{L}_2 = \text{linear span}([x_1], [x_2], [x_3], [x_4], [x_1x_2], [x_1x_3], [x_1x_4], [x_2x_3], [x_2x_4], [x_3x_4]). \tag{26}$$

The definition of $\mathcal{L}_1$ and $\mathcal{L}_2$ resembles the definition of punctured Reed-Muller codes of order 1 and 2, respectively, see Ref. [30]. Note also that $\mathcal{L}_1$ is the dual space for the 15-bit Hamming code. The relevant properties of the subspaces $\mathcal{L}_j$ are stated in the following lemma.

Lemma 1.

(1) For any $u \in \mathcal{L}_1$ one has $|u| \equiv 0 (\mathrm{mod}\ 8)$.

(2) For any $v \in \mathcal{L}_2$ one has $|v| \equiv 0 (\mathrm{mod}\ 2)$.

(3) Let $[1]$ be the unit vector (1, 1,…, 1, 1). Then $\mathcal{L}_1^{\perp} = \mathcal{L}_2 \oplus [1]$ and $\mathcal{L}_2^{\perp} = \mathcal{L}_1 \oplus [1]$.

(4) For any vectors $u, v \in \mathcal{L}_1$ one has $|u \cdot v| \equiv 0 (\mathrm{mod}\ 4)$.

(5) For any vectors $u \in \mathcal{L}_1$ and $v \in \mathcal{L}_2^{\perp}$ one has $|u \cdot v| \equiv 0 (\mathrm{mod}\ 4)$.

Proof.

(1) Any linear function f on $\mathbb{F}_2^4$ satisfying $f(0) = 0$ takes value 1 exactly eight times (if $f \neq 0$) or zero times (if $f = 0$).

(2) All basis vectors of $\mathcal{L}_2$ have weight equal to 8 (the vectors $[x_i]$) or 4 (the vectors $[x_i x_j]$). By linearity, all elements of $\mathcal{L}_2$ have even weight.

(3) One can easily check that all basis vectors of $\mathcal{L}_1$ are orthogonal to all basis vectors of $\mathcal{L}_2$, therefore $\mathcal{L}_1 \subseteq \mathcal{L}_2^{\perp}$, $\mathcal{L}_2 \subseteq \mathcal{L}_1^{\perp}$. Besides, we have already proved that $[1] \in \mathcal{L}_1^{\perp}$ and $[1] \in \mathcal{L}_2^{\perp}$. Now the statement follows from dimension counting, since $\dim \mathcal{L}_1 = 4$ and $\dim \mathcal{L}_2 = 10$.

(4) Without loss of generality we may assume that $u \neq 0$ and $v \neq 0$. If $u = v$, the statement has been already proved, see property 1. If $u \neq v$, then $u = [f]$, $v = [g]$ for some linearly independent linear functions f and g. We can introduce new coordinates (y_1, y_2, y_3, y_4) on $\mathbb{F}_2^4$ such that $y_1 = f(x)$ and $y_2 = g(x)$. Now $|u \cdot v| = |[y_1 y_2]| = 4$.

(5) Let $u \in \mathcal{L}_1$ and $v \in \mathcal{L}_2^{\perp}$. Since $\mathcal{L}_2^{\perp} = \mathcal{L}_1 \oplus [1]$, there are two possibilities: $v \in \mathcal{L}_1$ and $v = [1] + w$ for some $w \in \mathcal{L}_1$. The first case has been already considered. In the second case we have

$$|u\cdot v| = \sum_{j=1}^{15} u_j(1-w_j) = |u| - |u\cdot w|.$$

It follows from properties 1 and 4 that $|u\cdot v| \equiv 0 (\mathrm{mod}\, 4)$. □

Now consider the one-qubit Hermitian operator

$$A = \frac{1}{\sqrt{2}}(\sigma^x + \sigma^y) = \begin{pmatrix} 0 & e^{-i(\pi/4)} \\ e^{+i(\pi/4)} & 0 \end{pmatrix} = e^{-1(\pi/4)} K \sigma^x,$$

where K is the phase shift gate, see Eq. (1). By definition, A belongs to the Clifford group $\mathcal{C}_1$. One can easily check that $A^2 = I$ and $A\sigma^z = -\sigma^z A$, so the code $\mathrm{CSS}(\sigma^z, \mathcal{L}_2; A, \mathcal{L}_1)$ is well defined. We claim that its code subspace coincides with the code subspace of a certain stabilizer code.

Lemma 2. Consider the decomposition

$$(\mathbb{C}^2)^{\otimes 15} = \bigoplus_{\mu\in\mathcal{L}_2^*} \bigoplus_{\eta\in\mathcal{L}_1^*} \mathcal{H}(\mu,\eta),$$

associated with the code $\mathrm{CSS}(\sigma^z, \mathcal{L}_2; A, \mathcal{L}_1)$ and the decomposition

$$(\mathbb{C}^2)^{\otimes 15} = \bigoplus_{\mu\in\mathcal{L}_2^*} \bigoplus_{\eta\in\mathcal{L}_1^*} \mathcal{G}(\mu,\eta),$$

associated with the stabilizer code $\mathrm{CSS}(\sigma^z, \mathcal{L}_2; \sigma^x, \mathcal{L}_1)$. For any syndrome $\eta \in \mathcal{L}_1^*$ one has

$$\mathcal{H}(0,\eta) = \mathcal{G}(0,\eta).$$

Moreover, for any $\mu \in \mathcal{L}_2^*$ there exists some $w \in \mathbb{F}_2^{15}$ such that for any $\eta \in \mathcal{L}_1^*$

$$\mathcal{H}(\mu,\eta) = A(w)\mathcal{G}(0,\eta). \tag{27}$$

This Lemma provides a strategy to measure a syndrome of the code $\mathrm{CSS}(\sigma^z, \mathcal{L}_2; A, \mathcal{L}_1)$ by operations from $\mathcal{O}_{\mathrm{ideal}}$. Specifically, we measure μ (i.e., the σ^z part of the syndrome) first, compute $w = w(\mu)$, apply $A(w)^\dagger$, measure η using the stabilizers $\sigma^x([x_j])$, and apply $A(w)$.

Proof of the lemma. Consider an auxiliary subspace,

$$\mathcal{H} = \bigoplus_{\eta\in\mathcal{L}_1^*} \mathcal{H}(0,\eta) = \bigoplus_{\eta\in\mathcal{L}_1^*} \mathcal{G}(0,\eta),$$

corresponding to the trivial σ^z syndrome for both CSS codes. Each state $|\Psi\rangle \in \mathcal{H}(0)$ can be represented as

$$|\Psi\rangle = \sum_{v\in\mathcal{L}_2^\perp} c_v |v\rangle,$$

where c_v are some complex amplitudes and $|v\rangle = |v_1,\ldots,v_{15}\rangle$ are vectors of the standard basis. Let us show that

$$A(u)|\Psi\rangle = \sigma^x(u)|\Psi\rangle \text{ for any } |\Psi\rangle \in \mathcal{H}, \quad u \in \mathcal{L}_1.$$

To this end, we represent A as $\sigma^x e^{i\pi/4} K^\dagger$. For any $u \in \mathcal{L}_1$ and $v \in \mathcal{L}_2^\perp$ we have

$$A(u)|v\rangle = \sigma^x(u) e^{i(\pi/4)|u| - i(\pi/2)|u\cdot v|}|v\rangle = \sigma^x(u)|v\rangle,$$

because $|u| \equiv 0 (\mathrm{mod}\, 8)$ and $|u\cdot v| \equiv 0 (\mathrm{mod}\, 4)$ (see Lemma 1, parts 1 and 5).

Since for any $u \in \mathcal{L}_1$ the operators $A(u)$ and $\sigma^x(u)$ act on $\mathcal{H}$ in the same way, their eigenspaces must coincide, i.e., $\mathcal{H}(0,\eta) = \mathcal{G}(0,\eta)$ for any $\eta \in \mathcal{L}_1^*$.

Let us now consider the subspace $\mathcal{H}(\mu,\eta)$ for arbitrary $\mu \in \mathcal{L}_2^*$, $\eta \in \mathcal{L}_1^*$. By definition, μ is a linear functional on $\mathcal{L}_2 \subseteq \mathbb{F}_2^{15}$; we can extend it to a linear functional on $\mathbb{F}_2^{15}$, i.e., represent it in the form $\mu(v) = (w,v)$ for some $w \in \mathbb{F}_2^{15}$. Then for any $|\Psi\rangle \in \mathcal{H}(\mu,\eta)$, $v \in \mathcal{L}_2$, and $u \in \mathcal{L}_1$ we have

$$\sigma^z(v)A(w)^\dagger|\Psi\rangle = (-1)^{(w,v)} A(w)^\dagger \sigma^z(v)|\Psi\rangle = A(w)^\dagger|\Psi\rangle,$$

$$A(u)A(w)^\dagger|\Psi\rangle = A(w)^\dagger A(u)|\Psi\rangle = (-1)^{\eta(v)} A(w)^\dagger|\Psi\rangle$$

(as σ^z and A anticommute), hence $A(w)^\dagger|\Psi\rangle \in \mathcal{H}(0,\eta)$. Thus

$$\mathcal{H}(\mu,\eta) = A(w)\mathcal{H}(0,\eta) = A(w)\mathcal{G}(0,\eta).$$

□

Lemma 2 is closely related to an interesting property of the stabilizer code $\mathrm{CSS}(\sigma^z, \mathcal{L}_2; \sigma^x, \mathcal{L}_1)$, namely the existence of a non-Clifford automorphism [23]. Consider a one-qubit unitary operator W such that

$$W\sigma^z W^\dagger = \sigma^z \text{ and } W\sigma^x W^\dagger = A.$$

It is defined up to an overall phase and obviously does not belong to the Clifford group $\mathcal{C}_1$. However, the bitwise application of W, i.e., the operator $W^{\otimes 15}$, preserves the code subspace $\mathcal{G}(0,0)$. Indeed, $W^{\otimes 15}\mathcal{G}(0,0)$ corresponds to the trivial syndrome of the code

$$\mathrm{CSS}(W\sigma^z W^\dagger, \mathcal{L}_2; W\sigma^x W^\dagger, \mathcal{L}_1) = \mathrm{CSS}(\sigma^z, \mathcal{L}_2; A, \mathcal{L}_1).$$

Thus $W^{\otimes 15}\mathcal{G}(0,0) = \mathcal{H}(0,0)$. But $\mathcal{H}(0,0) = \mathcal{G}(0,0)$ due to the lemma.

Now we are in a position to describe the distillation scheme and to estimate its threshold and yield. Suppose we are given 15 copies of the state ρ, and our goal is to distill one copy of an H-type magic state. We will actually distill the state,

$$|A_0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{i\frac{\pi}{4}}|1\rangle) = e^{i\frac{\pi}{8}} HK^\dagger |H\rangle.$$

Note that $|A_0\rangle$ is an eigenstate of the operator A; specifically, $A|A_0\rangle = |A_0\rangle$. Let us also introduce the state

$$|A_1\rangle = \sigma^z|A_0\rangle,$$

which satisfies $A|A_1\rangle = -|A_1\rangle$. Since the Clifford group $\mathcal{C}_1$ acts transitively on the set of H-type magic states, we can assume that the fidelity between ρ and $|A_0\rangle$ is the maximum one among all H-type magic states, so that

$$F_H(\rho) = \sqrt{\langle A_0|\rho|A_0\rangle}.$$

As in Sec. V we define the initial error probability

$$\epsilon = 1 - [F_H(\rho)]^2 = \langle A_1|\rho|A_1\rangle.$$

Applying the dephasing transformation

$$D(\eta) = \frac{1}{2}(\eta + A\,\eta A^\dagger)$$

to each copy of ρ, we can guarantee that ρ is diagonal in the $\{A_0, A_1\}$ basis, i.e.,

$$\rho = D(\rho) = (1-\epsilon)|A_0\rangle\langle A_0| + \epsilon|A_1\rangle\langle A_1|.$$

Since $A \in \mathcal{C}_1$, the dephasing transformation can be realized by operations from $\mathcal{O}_{\text{ideal}}$. Thus our initial state is

$$\rho_{\text{in}} = \rho^{\otimes 15} = \sum_{u \in \mathbb{F}_2^{15}} \epsilon^{|u|}(1-\epsilon)^{15-|u|}|A_u\rangle\langle A_u|, \tag{28}$$

where $|A_u\rangle = |A_{u_0}\rangle \otimes \cdots \otimes |A_{u_{15}}\rangle$.

According to the remark following the formulation of Lemma 2, we can measure the syndrome (μ, η) of the code $\text{CSS}(\sigma^z, \mathcal{L}_2; A, \mathcal{L}_1)$ by operations from $\mathcal{O}_{\text{ideal}}$ only. Let us follow this scheme, omitting the very last step. So, we begin with the state ρ_{in}, measure μ, compute $w = w(\mu)$, apply $A(w)^\dagger$, and measure η. We consider the distillation attempt successful if $\eta = 0$. The measured value of μ is not important at this stage. In fact, for any $\mu \in \mathcal{L}_2^*$ the unnormalized post-measurement state is

$$\rho_s = \Pi A(w)^\dagger \rho_{\text{in}} A(w) \Pi = \Pi \rho_{\text{in}} \Pi.$$

In this equation Π is the projector onto the code subspace $\mathcal{H}(0,0) = \mathcal{G}(0,0)$, i.e., $\Pi = \Pi_z \Pi_A$ for

$$\Pi_z = \frac{1}{|\mathcal{L}_2|}\sum_{v \in \mathcal{L}_2} \sigma^z(v), \quad \Pi_A = \frac{1}{|\mathcal{L}_1|}\sum_{u \in \mathcal{L}_1} A(u). \tag{29}$$

Let us compute the state $\rho_s = \Pi \rho_{\text{in}} \Pi$. Since

$$A(u)|A_w\rangle = (-1)^{(u,w)}|A_w\rangle, \quad \sigma^z(v)|A_w\rangle = |A_{w+v}\rangle,$$

one can easily see that $\Pi_A|A_w\rangle = |A_w\rangle$ if $w \in \mathcal{L}_1^\perp$, otherwise $\Pi_A|A_w\rangle = 0$. On the other hand, $\Pi_z|A_w\rangle$ does not vanish and depends only on the coset of $\mathcal{L}_2$ that contains w. There are only two such cosets in $\mathcal{L}_1^\perp$ (because $\mathcal{L}_1^\perp = \mathcal{L}_2 \oplus [1]$, see Lemma 1), and the corresponding projected states are

$$|A_0^L\rangle = \sqrt{|\mathcal{L}_2|}\Pi_z|A_{0\ldots0}\rangle = \frac{1}{\sqrt{|\mathcal{L}_2|}}\sum_{v \in \mathcal{L}_2}|A_v\rangle,$$

$$|A_1^L\rangle = \sqrt{|\mathcal{L}_2|}\Pi_z|A_{1\ldots1}\rangle = \frac{1}{\sqrt{|\mathcal{L}_2|}}\sum_{v \in \mathcal{L}_2}|A_{v+[1]}\rangle. \tag{30}$$

The states $|A_{0,1}^L\rangle$ form an orthonormal basis of the code subspace. The projections of $|A_w\rangle$ for $w \in \mathcal{L}_1^\perp$ onto the code subspace are given by these formulas:

$$\Pi|A_w\rangle = \frac{1}{\sqrt{|\mathcal{L}_2|}}|A_0^L\rangle \text{ if } w \in \mathcal{L}_2,$$

$$\Pi|A_w\rangle = \frac{1}{\sqrt{|\mathcal{L}_2|}}|A_1^L\rangle \text{ if } w \in \mathcal{L}_2 + [1].$$

Now the unnormalized final state $\rho_s = \Pi\rho_{\text{in}}\Pi$ can be expanded as

$$\rho_s \frac{1}{|\mathcal{L}_2|}\sum_{v \in \mathcal{L}_2}(1-\epsilon)^{15-|v|}\epsilon^{|v|}|A_0^L\rangle\langle A_0^L|$$

$$\times + \frac{1}{|\mathcal{L}_2|}\sum_{v \in \mathcal{L}_2}\epsilon^{15-|v|}(1-\epsilon)^{|v|}|A_1^L\rangle\langle A_1^L|.$$

The distillation succeeds with probability

$$p_s = |\mathcal{L}_2|\text{Tr}\,\rho_s = \sum_{v \in \mathcal{L}_1^\perp}\epsilon^{15-|v|}(1-\epsilon)^{|v|}.$$

(The factor $|\mathcal{L}_2|$ reflects the number of possible values of μ, which all give rise to the same state ρ_s.)

To complete the distillation procedure, we need to apply a decoding transformation that would map the two-dimensional subspace $\mathcal{H}(0,0) \subset (\mathbb{C}^2)^{\otimes 15}$ onto the Hilbert space of one qubit. Recall that $\mathcal{H}(0,0) = \mathcal{G}(0,0)$ is the code subspace of the stabilizer code $\text{CSS}(\sigma^z, \mathcal{L}_2; \sigma^x, \mathcal{L}_1)$. Its logical Pauli operators can be chosen as

$$\hat{X} = (\sigma^x)^{\otimes 15}, \quad \hat{Y} = (\sigma^y)^{\otimes 15}, \quad \hat{Z} = -(\sigma^z)^{\otimes 15}.$$

It is easy to see that $\hat{X}, \hat{Y}, \hat{Z}$ obey the correct algebraic relations and preserve the code subspace. The decoding can be realized as a Clifford operator $V \in \mathcal{C}_{15}$ that maps $\hat{X}, \hat{Y}, \hat{Z}$ to the Pauli operators $\sigma^x, \sigma^y, \sigma^z$ acting on the first qubit. (The remaining 14 qubits become unentangled with the first one, so we can safely disregard them.) Let us show that the logical state $|A_0^L\rangle$ is transformed into $|A_0\rangle$ (up to some phase). For this, it suffices to check that $\langle A_0^L|\hat{X}|A_0^L\rangle = \langle A_0|\sigma^x|A_0\rangle$, $\langle A_0^L|\hat{Y}|A_0^L\rangle = \langle A_0|\sigma^y|A_0\rangle$, and $\langle A_0^L|\hat{Z}|A_0^L\rangle = \langle A_0|\sigma^z|A_0\rangle$. Verifying these identities becomes a straightforward task if we represent $|A_0^L\rangle$ in the standard basis:

$$|A_0^L\rangle = |\mathcal{L}_2|^{1/2}2^{-15/2}\sum_{u \in \mathcal{L}_2^\perp} e^{i(\pi/4)|u|}|u\rangle$$

$$= 2^{-5/2}\sum_{u \in \mathcal{L}_1}(|u\rangle + e^{-i(\pi/4)}|u + [1]\rangle).$$

To summarize, the distillation subroutine consists of the following steps.

(1) Measure eigenvalues of the Pauli operators $\sigma^z([x_j])$, $\sigma^z([x_j x_k])$ (for $j,k = 1,2,3,4$). The outcomes determine the σ^z syndrome, $\mu \in \mathcal{L}_2^*$.

(2) Find $w = w(\mu) \in \mathbb{F}_2^{15}$ such that $(w,v) = \mu(v)$ for any $v \in \mathcal{L}_2$.

(3) Apply the correcting operator $A(w)^\dagger$.

(4) Measure eigenvalues of the operators $\sigma^x([x_j])$. The outcomes determine the A syndrome, $\eta \in \mathcal{L}_1^*$.

(5) Declare failure if $\eta \neq 0$, otherwise proceed to the next step.

(6) Apply the decoding transformation, which takes the

code subspace to the Hilbert space of one qubit.

The subroutine succeeds with probability

$$p_s = \sum_{v \in \mathcal{L}_1^\perp} \epsilon^{15-|v|}(1-\epsilon)^{|v|}. \tag{31}$$

In the case of success, it produces the normalized output state

$$\rho_{\text{out}} = (1-\epsilon_{\text{out}})|A_0\rangle\langle A_0| + \epsilon_{\text{out}}|A_1\rangle\langle A_1| \tag{32}$$

characterized by the error probability

$$\epsilon_{\text{out}} = p_s^{-1} \sum_{v \in \mathcal{L}_2} \epsilon^{15-|v|}(1-\epsilon)^{|v|}. \tag{33}$$

The sums in Eqs. (31) and (33) are special forms of so-called weight enumerators. The *weight enumerator* of a subspace $\mathcal{L} \subseteq \mathbb{F}_2^n$ is a homogeneous polynomial of degree n in two variables, namely

$$W_{\mathcal{L}}(x,y) = \sum_{u \in \mathcal{L}} x^{n-|u|}y^{|u|}.$$

In this notation,

$$p_s = W_{\mathcal{L}_1^\perp}(\epsilon, 1-\epsilon), \quad \epsilon_{\text{out}} = \frac{W_{\mathcal{L}_2}(\epsilon,1-\epsilon)}{W_{\mathcal{L}_1^\perp}(\epsilon,1-\epsilon)}.$$

The MacWilliams identity [30] relates the weight enumerator of $\mathcal{L}$ to that of $\mathcal{L}^\perp$:

$$W_{\mathcal{L}}(x,y) = \frac{1}{|\mathcal{L}^\perp|} W_{\mathcal{L}^\perp}(x+y, x-y).$$

Applying this identity and taking into account that $\mathcal{L}_2^\perp = \mathcal{L}_1 \oplus [1]$ and that $|u| \equiv 0 (\text{mod } 2)$ for any $u \in \mathcal{L}_1$ (see Lemma 1), we get

$$p_s = \frac{1}{16} W_{\mathcal{L}_1}(1, 1-2\epsilon), \quad \epsilon_{\text{out}} = \frac{1}{2}\left(1 - \frac{W_{\mathcal{L}_1}(1-2\epsilon,1)}{W_{\mathcal{L}_1}(1,1-2\epsilon)}\right). \tag{34}$$

The weight enumerator of the subspace $\mathcal{L}_1$ is particularly simple:

$$W_{\mathcal{L}_1}(x,y) = x^{15} + 15x^7y^8.$$

Substituting this expression into Eq. (34), we arrive at the following formulas:

$$p_s = \frac{1 + 15(1-2\epsilon)^8}{16}, \tag{35}$$

$$\epsilon_{\text{out}} = \frac{1 - 15(1-2\epsilon)^7 + 15(1-2\epsilon)^8 - (1-2\epsilon)^{15}}{2[1 + 15(1-2\epsilon)^8]}. \tag{36}$$

The function $\epsilon_{\text{out}}(\epsilon)$ is plotted in Fig. 3. Solving the equation $\epsilon_{\text{out}}(\epsilon) = \epsilon$ numerically, we find the threshold error probability:

$$\epsilon_0 \approx 0.141. \tag{37}$$

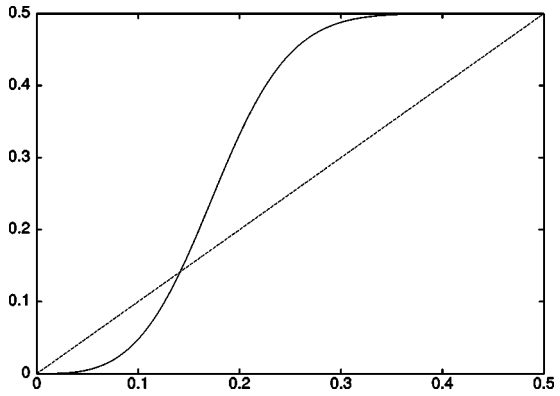


FIG. 3. The final error probability $\epsilon_{\text{out}}(\epsilon)$ for the H-type states distillation.

Let us examine the asymptotic properties of this scheme. For small ϵ the distillation subroutine succeeds with probability close to 1, therefore the yield is close to $1/15$. The output error probability is

$$\epsilon_{\text{out}} \approx 35\epsilon^3. \tag{38}$$

Now suppose that the subroutine is applied recursively. From n copies of the state ρ with a given ϵ, we distill one copy of the magic state $|A_0\rangle$ with the final error probability

$$\epsilon_{\text{out}}(n,\epsilon) \approx \frac{1}{\sqrt{35}}(\sqrt{35}\epsilon)^{3^k}, \quad 15^k \approx n,$$

where k is the number of recursion levels (here we neglect the fluctuations in the number of successful distillation attempts). Solving these equation, we obtain the relation

$$\epsilon_{\text{out}}(n,\epsilon) \sim (\sqrt{35}\epsilon)^{n^\xi}, \quad \xi = 1/\log_3 15 \approx 0.4. \tag{39}$$

It characterizes the efficiency of the distillation scheme.

VII. CONCLUSION AND SOME OPEN PROBLEMS

We have studied a simplified model of fault-tolerant quantum computation in which operations from the Clifford group are realized exactly, whereas decoherence occurs only during the preparation of nontrivial ancillary states. The model is fully characterized by a one-qubit density matrix ρ describing these states. It is shown that a good strategy for simulating universal quantum computation in this model is "magic states distillation." By constructing two particular distillation schemes we find a threshold polarization of ρ above which the simulation is possible.

The most exciting open problem is to understand the computational power of the model in the region of parameters $1 < |\rho_x| + |\rho_y| + |\rho_z| \leqslant 3/\sqrt{7}$ (which corresponds to $F_T^* < F_T(\rho) \leqslant F_T$, see Sec. I). In this region, the distillation scheme based on the five-quit code does not work, while the Gottesman-Knill theorem does not yet allow the classical simulation. One possibility is that a transition from classical to universal quantum behavior occurs on the octahedron boundary, $|\rho_x| + |\rho_y| + |\rho_z| = 1$.

To prove the existence of such a transition, one it suffices to construct a T-type states distillation scheme having the threshold fidelity F_T^*. A systematic way of constructing such schemes is to replace the five-qubit by a $GF(4)$-linear stabilizer code. A nice property of these codes is that the bitwise application of the operator T preserves the code subspace and acts on the encoded qubit as T, see Ref. [31] for more details. One can check that the error-correcting effect described in Sec. V takes place for an arbitrary GF(4)-linear stabilizer code, provided that the number of qubits is $n=6k-1$ for any integer k. Unfortunately, numerical simulations we performed for some codes with $n=11$ and $n=17$ indicate that the threshold fidelity increases as the number of qubits increases. So it may well be the case that the five-qubit code is the best GF(4)-linear code as far as the distillation is concerned.

From the experimental point of view, an exciting open problem is to design a physical system in which reliable storage of quantum information and its processing by Clifford group operations is possible. Since our simulation scheme tolerates strong decoherence on the ancilla preparation stage, such a system would be a good candidate for a practical quantum computer.

ACKNOWLEDGMENTS

We thank Mikhail Vyalyi for bringing to our attention many useful facts about the Clifford group. This work has been supported in part by the National Science Foundation under Grant No. EIA-0086038.

APPENDIX

The purpose of this section is to prove Eq. (15). Let us introduce this notation:

$$|\hat{T}_0\rangle = |T_{00000}\rangle \text{ and } |\hat{T}_1\rangle = |T_{11111}\rangle.$$

Consider the set $S_+(5)\subset S(5)$ consisting of all possible tensor products of the Pauli operators $\sigma^x,\sigma^y,\sigma^z$ on five qubits (clearly, $|S_+(5)|=4^5=|S(5)|/2$ since elements of $S(5)$ may have a plus or minus sign). For each $g\in S_+(5)$ let $|g|\in[0,5]$ be the number of qubits on which g acts nontrivially (e.g., $|\sigma^x\otimes\sigma^x\otimes\sigma^y\otimes I\otimes I|=3$). We have

$$|\hat{T}_0\rangle\langle\hat{T}_0| = \frac{1}{2^5}\sum_{g\in S_+(5)}\left(\frac{1}{\sqrt{3}}\right)^{|g|} g.$$

Now let us expand the formula (13) for the projector Π. Denote by $G\subset P(5)$ the Abelian group generated by the stabilizers S_1,S_2,S_3,S_4. It consists of 16 elements. Repeatedly conjugating the stabilizer S_1 by the operator $\hat{T}=T^{\otimes 5}$, we get three elements of G:

$$S_1 = \sigma^x\otimes\sigma^z\otimes\sigma^z\otimes\sigma^x\otimes I,$$

$$S_1S_3S_4 = \sigma^z\otimes\sigma^y\otimes\sigma^y\otimes\sigma^z\otimes I,$$

$$S_3S_4 = \sigma^y\otimes\sigma^x\otimes\sigma^x\otimes\sigma^y\otimes I.$$

Due to the cyclic symmetry mentioned in Sec. V, the 15 cyclic permutations of these elements also belong to G; together with the identity operator they exhaust the group G. Thus $G\subset S_+(5)$, and we have

$$\Pi = \frac{1}{16}\sum_{h\in G} h.$$

Taking into account that $\mathrm{Tr}(gh)=2^5\delta_{g,h}$ for any $g,h\in S_+(5)$, we get

$$\langle\hat{T}_0|\Pi|\hat{T}_0\rangle = \frac{1}{2^9}\sum_{h\in G}\sum_{g\in S_+(5)} 3^{-|g|/2}\mathrm{Tr}(gh) = \frac{1}{16}\sum_{g\in G} 3^{-|g|/2} = \frac{1}{6}.$$

Similar calculations show that $\langle\hat{T}_1|\Pi|\hat{T}_1\rangle=\frac{1}{6}$.

[1] E. Knill, R. Laflamme, and W. Zurek, Science **279**, 342 (1998).
[2] C. Zalka, e-print quant-ph/9612028.
[3] A. Steane, Phys. Rev. Lett. **78**, 2252 (1997).
[4] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, J. Math. Phys. **43**, 4452 (2002).
[5] A. Kitaev, Ann. Phys. (N.Y.) **303**, 2 (2003).
[6] M. Freedman, M. Larsen, and Z. Wang, e-print quant-ph/0001108.
[7] M. Freedman, A. Kitaev, M. Larsen, and Z. Wang, Bull., New Ser., Am. Math. Soc. **40**, 31 (2002).
[8] C. Mochon, Phys. Rev. A **69**, 032306 (2004).
[9] G. Moore and N. Read, Nucl. Phys. B **360**, 362 (1991).
[10] C. Nayak and F. Wilczek, Nucl. Phys. B **479**, 529 (1996).
[11] B. Doucot and J. Vidal, Phys. Rev. Lett. **88**, 227005 (2001).
[12] M. Feigel'man and L. Ioffe, Phys. Rev. B **66**, 224503 (2002).
[13] J. Preskill, e-print quant-ph/9712048.
[14] D. Gottesman, Ph.D. thesis, Caltech, Pasadena, 1997, URL http://arxiv.org/abs/quant-ph/9705052.
[15] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).
[16] D. Gottesman, Phys. Rev. A **57**, 127 (1998).
[17] E. Knill, e-print quant-ph/0402171.
[18] E. Knill, e-print quant-ph/0404104.
[19] D. Gottesman and I. Chuang, Nature (London) **402**, 390 (1999).
[20] W. Dur and H. Briegel, Phys. Rev. Lett. **90**, 067901 (2003).
[21] D. Aharonov, e-print quant-ph/9602019.
[22] E. Dennis, Phys. Rev. A **63**, 052314 (2001).
[23] E. Knill, R. Laflamme, and W. Zurek, e-print quant-ph/9610011.

[24] A. Calderbank, E. Rains, P. Shor, and N. Sloane, Phys. Rev. Lett. **78**, 405 (1997).
[25] A. Ambainis and D. Gottesman, e-print quant-ph/0310097.
[26] C. Bennett, D. DiVincenzo, J. Smolin, and W. Wootters, Phys. Rev. A **54**, 3824 (1996).
[27] R. Laflamme, C. Miquel, J. Paz, and W. Zurek, Phys. Rev. Lett. **77**, 198 (1996).
[28] A. Calderbank and P. Shor, Phys. Rev. A **54**, 1098 (1996).
[29] A. Steane, Proc. R. Soc. London, Ser. A **452**, 2551 (1996).
[30] F. MacWilliams and N. Sloane, *The Theory of Error-Correcting Codes* (North-Holland, Amsterdam, 1981).
[31] A. Calderbank, E. Rains, P. Shor, and N. Sloane, e-print quant-ph/9608006.

Tight Noise Thresholds for Quantum Computation with Perfect Stabilizer Operations

Wim van Dam*
Department of Computer Science, University of California, Santa Barbara, California 93106, USA and Department of Physics, University of California, Santa Barbara, California 93106, USA

Mark Howard†
Department of Physics, University of California, Santa Barbara, California 93106, USA
(Received 21 July 2009; published 23 October 2009)

We study how much noise can be tolerated by a universal gate set before it loses its quantum-computational power. Specifically we look at circuits with perfect stabilizer operations in addition to imperfect nonstabilizer gates. We prove that for all unitary single-qubit gates there exists a tight depolarizing noise threshold that determines whether the gate enables universal quantum computation or if the gate can be simulated by a mixture of Clifford gates. This exact threshold is determined by the Clifford polytope spanned by the 24 single-qubit Clifford gates. The result is in contrast to the situation wherein nonstabilizer qubit states are used; the thresholds in that case are not currently known to be tight.

DOI: 10.1103/PhysRevLett.103.170504 PACS numbers: 03.67.Lx, 03.67.Pp

Introduction.—A way to study the resources needed for universal quantum computation (UQC) is to analyze the transition from a system that can provide UQC to one that is classically efficiently simulable. A particularly useful example of a classically simulable system is given by the stabilizer operations, which are made by a combination of preparation of $|0\rangle$ states, unitary Clifford gates, measurements in the $\{|0\rangle, |1\rangle\}$ basis, and classical control determined by the measurement outcomes. The Gottesman-Knill theorem tells us that stabilizer operations can be efficiently simulated classically (see, for example, [1], Theorem 10.7), while it also known that the addition of any other one-qubit gate outside the Clifford group will enable the system to perform UQC. This fact provides us with a framework for testing tolerance to noise—one can examine how noisy this additional non-Clifford gate can be before it becomes classically simulable itself. If the non-Clifford operation has become a probabilistic combination of Clifford gates due to the noise, then we know that we are unequivocally in the classical computational regime. The noise rate where the extra gate becomes simulable (where it enters the ''Clifford polytope'' [2]) is thus an upper bound for fault tolerance. If the converse is true—i.e., if any operation outside the Clifford polytope enables UQC, then the threshold is tight. In this Letter we show that for single-qubit gates undergoing depolarizing such a tight noise threshold does indeed apply. We will do so by proving that any depolarized gate that lies outside the Clifford polytope of single-qubit operations, in combination with noiseless stabilizing operations, allows for UQC. This result should be contrasted to the situation for nonstabilizer qubit *states* where the thresholds in that case are not currently known to be tight. In fact, a recent result by Campbell and Browne [3] states that achieving tight thresholds for all nonstabilizer qubit states is impossible if the number of copies of the resource state must be finite.

We will consistently assume that Clifford gates can be implemented perfectly, motivated by the fact that these gates can be implemented fault tolerantly by applying them transversally and to encoded states [4–7]. The fault-tolerant implementation of Clifford gates naturally carries with it a threshold of its own, independent of the kind we discuss in this Letter. The current model is particularly relevant to the so-called Pfaffian quantum Hall state in topological quantum computation [8,9], the two-qubit Clifford group (but only the Clifford group) can be implemented using braiding making these operations naturally fault tolerant. The additional resource required to perform UQC will likely be highly noisy, and so we can see the parallels with our model.

We will begin by listing a couple of previously known results in this area. Next, we will discuss the connection between the geometry of the Clifford polytope and stabilizer measurements, and show that tightness of a magic-state distillation procedure for single-qubit states automatically ensures tight thresholds for non-Clifford gates undergoing any kind of unital noise. Finally, we show that currently known magic-state distillation techniques are sufficient to prove tight thresholds for a non-Clifford gate undergoing depolarizing noise.

Previously known results.—The idea of using perfect stabilizer operations in conjunction with imperfect nonstabilizer states to perform UQC originates with Knill [7]. Shortly after, Bravyi and Kitaev [10] showed that most nonstabilizer qubit states (when sufficiently many copies are available) can be purified (''distilled''), using only stabilizer operations, towards a pure nonstabilizer state (a ''magic state''). Since a universal gate set can be created from perfect stabilizer operations and a supply of magic states [10], we see that allowing access to a supply of appropriate (possibly impure) nonstabilizer qubit ancillas promotes the power of stabilizer operations from classi-

cally simulable to UQC. The conditions on the ancillary qubits to enable UQC is that they are sufficiently close to being one of the 20 so-called magic states that lie on the surface of the Bloch sphere. The two classes of magic state (see Fig. 1) are the $|H\rangle$ type and $|T\rangle$ type, where all $|H\rangle$ type states can be derived by applying a Clifford operation to some canonical representative $|H\rangle = (|0\rangle + e^{i\pi/4}|1\rangle)/\sqrt{2}$, and similarly for $|T\rangle$-type states like $|T\rangle = \cos(\vartheta)|0\rangle + e^{i\pi/4}\sin(\vartheta)|1\rangle$, where $\cos(2\vartheta) = 1/\sqrt{3}$. The routines used in [10] were unable to distill qubit states just outside the edges and faces of the octahedron of Fig. 1. Reichardt [11] subsequently proposed an improved routine that closed the gap in the $|H\rangle$ direction (along the edges of the octahedron). Virmani *et al.* [12] suggested using the convex hull of Clifford operations in order to find gates' robustness to various types of noise. In particular, they considered gates that are diagonal in the computational basis. Plenio and Virmani [13] subsequently extended this idea by analyzing cases where noise was allowed to affect the stabilizer operations too. Buhrman *et al.* [2] used a similar idea (that noise causes non-Clifford gates to eventually become able to be implemented via Clifford gates only) to find the non-Clifford gate that is most resistant to depolarizing noise—a $\pi/8$ rotation about the Z axis (or the same gate modulo some Clifford operation). Reichardt [14] showed that this particular gate enabled UQC right up to its threshold noise rate (about 45%), as well as considering in detail the process of reducing multiqubit states to single-qubit states using postselected stabilizer operations. Our current result here generalizes this tightness result to all possible single-qubit gates.

Preliminaries and notation.—Let us parameterize an arbitrary single-qubit SU(2) gate as follows

$$U(\theta, \gamma, \delta) = \begin{pmatrix} e^{i\gamma}\cos(\theta) & -e^{i\delta}\sin(\theta) \\ e^{-i\delta}\sin(\theta) & e^{-i\gamma}\cos(\theta) \end{pmatrix}. \quad (1)$$

The representation of this rotation in SO(3) is denoted by $R(\theta, \gamma, \delta)$. Implementing a rotation R while suffering depolarizing noise (with noise rate p), means that this noisy operation is represented by the rescaling $M = (1 - p)R$, a fact that we will need later.

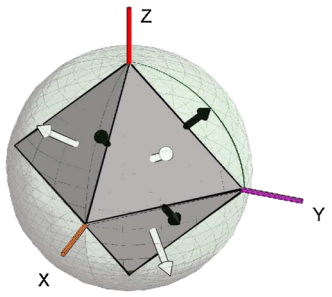


FIG. 1 (color online). Magic states and the octahedron: Some of the single-qubit magic states: $|H\rangle$ type states are designated with black arrows, $|T\rangle$ type states with white arrows. The octahedron defined by $|x| + |y| + |z| \leq 1$ depicts the single-qubit states that can be created by stabilizer operations. Reichardt [11] showed that distillation techniques work right up to the edges of the octahedron (i.e., tight in the $|H\rangle$ direction). Current distillation techniques are unable to distill states just outside the faces of the octahedron (i.e., not tight in the $|T\rangle$ direction).

Often we will apply the unitary $U(\theta, \gamma, \delta)$ to one half of an entangled Bell pair, $|\Phi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, yielding

$$\rho = (I \otimes U)|\Phi\rangle\langle\Phi|(I \otimes U)^\dagger. \quad (2)$$

If we use the two-qubit Pauli operators as a basis for the density matrix ρ then we can find the 16 real coefficients $c_{ij} = \mathrm{Tr}(\rho(\sigma_i \otimes \sigma_j))$ so that

$$\rho = \frac{1}{4}\sum c_{ij}(\sigma_i \otimes \sigma_j), \qquad i, j \in \{I, X, Y, Z\}. \quad (3)$$

Since we have applied a local unitary to a maximally entangled state, the coefficients (c_{IX}, c_{IY}, c_{IZ}, c_{XI}, c_{YI}, c_{ZI}) are always zero. Comparing the 9 coefficients $\{c_{XX}, c_{XY}, \ldots, c_{ZZ}\}$ one can see that these are the same as the entries of the SO(3) matrix $R(\theta, \gamma, \delta)$. More precisely,

$$R(\theta, \gamma, \delta) = \begin{pmatrix} c_{XX} & -c_{YX} & c_{ZX} \\ c_{XY} & -c_{YY} & c_{ZY} \\ c_{XZ} & -c_{YZ} & c_{ZZ} \end{pmatrix}, \quad (4)$$

where the c_{ij} are obviously also functions of (θ, γ, δ).

If we represent the 24 single-qubit Clifford operations as SO(3) matrices, then they are simply signed permutation matrices with unit determinant (they are a matrix representation of the elements of the chiral octahedral symmetry group or, equivalently, the symmetry group S_4). We label these operations C_i and so the convex hull of the C_i (the so-called Clifford polytope) is given by

$$\mathcal{P} = \left\{ \sum_{i=1}^{24} p_i C_i \;\middle|\; \text{with } p_i \geq 0 \text{ and } \sum_{i=1}^{24} p_i = 1 \right\}. \quad (5)$$

Geometrically, the Clifford polytope is a closed polyhedron in $\mathbb{R}^9$ that has 24 vertices (each vertex representing one of the C_i). This polytope can also be defined by the bounding inequalities of its 120 facets. The concise description of these facets used by Buhrman *et al.* [2] is given by the set

$$\mathcal{F} = \{C_i F C_j | i, j \in \{1, \ldots, 24\}, F \in \{A, A^T, B\}\}, \quad (6)$$

where

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 1 & 0 & 1 \end{pmatrix}. \quad (7)$$

At times we will have reason to refer to different subsets of the set of facets $\mathcal{F}$ so we use the obvious notation $\mathcal{F} = \mathcal{F}_A \cup \mathcal{F}_{A^T} \cup \mathcal{F}_B$. It is useful to note that all the facets derived from A comprise a single column with ± 1 entries and zeros elsewhere, and similarly for the row facets derived from A^T; hence, $|\mathcal{F}_A| = |\mathcal{F}_{A^T}| = 3 \times 2^3 = 24$. There are $|\mathcal{F}_B| = 72$ "B-type" facets, which can be constructed as follows: (i) Pick one position in a 3×3 matrix F, e.g., row i and column j and put ± 1 there ($9 \times 2 = 18$

choices), (ii) Fill the remaining entries not in row i or column j with ± 1 such that $\det(F) = -2$ (4 choices).

To determine whether or not an operation M is inside the Clifford polytope $\mathcal{P}$ we take the elementwise inner product (or Frobenius inner product) between M and the facets $F \in \mathcal{F}$ of the polytope

$$M \cdot F = \sum_{i,j=1}^{3} M_{i,j} F_{i,j} = \mathrm{Tr}(M^T F). \tag{8}$$

Using the above notation, a 3×3 matrix M is inside the polytope $\mathcal{P}$ if and only if for all $F \in \mathcal{F}$ we have $M \cdot F \leq 1$.

Interpreting the facets of the Clifford polytope.—Our proof will involve applying some non-Clifford gate to one half of a Bell Pair [as in Eq. (2)] and then postselecting on the outcomes of various stabilizer measurements. After some further stabilizer operations, this measurement ultimately has the effect of taking our two-qubit state ρ to a single-qubit state ρ' (times some stabilizer state that we do not care about), which we then distill using magic-state distillation (see [14] for a more general discussion of these kinds of techniques). For example, performing a YX measurement on ρ and postselecting on a "$+1$" outcome (i.e., projecting with $\Pi = \frac{1}{2}(II + YX)$) leads to a single-qubit state ρ' with a Bloch vector given by

$$\vec{r}(\rho') = \left(0, \frac{c_{XZ} - c_{ZY}}{c_{II} + c_{YX}}, -\frac{c_{XY} + c_{ZZ}}{c_{II} + c_{YX}}\right). \tag{9}$$

The form of this vector means that it lies in the YZ plane (see Fig. 1), where we know that distillation techniques work right up to the edge $|y| + |z| = 1$ of the octahedron. We can check if $\vec{r}$ is outside the octahedron by simply comparing the L^1 norm of $\vec{r}$ with 1. Rearranged, the condition $\|\vec{r}\|_1 > 1$ for being outside the octahedron is

$$|c_{XZ} - c_{ZY}| + |-(c_{XY} + c_{ZZ})| > |c_{II} + c_{YX}|. \tag{10}$$

Given the correspondence between the coefficients c_{ij} and the elements of R [see Eq. (4)] we can rewrite the above condition (dropping the absolute value operators) as a facet inequality

$$R \cdot F > 1 \quad \text{where} \quad F = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & -1 \\ 1 & 0 & -1 \end{pmatrix}. \tag{11}$$

This facet is a legitimate "B-type" facet and a little thought shows that, had we applied the single-qubit Pauli operations [as SO(3) rotations] X, Y or Z to $\vec{r}(\rho')$ above, we would arrive at three other "B-type" facets

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 1 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 1 \\ -1 & 0 & -1 \end{pmatrix}, \tag{12}$$

respectively. Note that all four facet inequalities combined could be simplified to the form Eq. (10) above. We omit the details, but it is straightforward to show that all 72 "B-type" facets correspond to postselecting on some (weight two) Pauli operator, and possibly performing a single-qubit Pauli rotation on the resulting ρ'.

It is somewhat more straightforward to see the geometrical interpretation of the "$A^{(T)}$-type" facets. For example, the canonical A given in Eq. (7), merely returns the sum of the elements of the Bloch vector $\vec{r}$, arising from a rotation applied to the X "$+1$" eigenstate.

$$R \cdot A = \sum_{i=1}^{3} r_i \quad \text{where} \quad \vec{r} = R \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \tag{13}$$

In general, an operation M having an inner product greater than one with some "A-type" facet simply means that M, applied to some initial vector corresponding to a stabilizer state, brings that vector to a final position outside the octahedron.

The preceding discussion shows us that if magic-state distillation were possible everywhere outside the octahedron, then every unital operation outside the Clifford polytope would be distillable—either straightforwardly or by using postselection, depending on which facet it violated. Using current (not tight) distillation routines however, we would be unable to deal with some operations violating an "A-type" facet by a fairly small amount. In the next section we show that, for depolarizing noise, any noisy rotation violating an "A-type" facet also violates a "B-type" facet. Since "B-type" facets correspond to $|H\rangle$ state distillation, the results we obtain are tight.

Tight threshold for depolarizing noise.—The claim we shall prove is that, anytime a matrix $M = (1 - p)R$, representing a depolarized rotation, is outside some "A-type" facet then there exists a "B-type" facet that M also lies outside. In fact, we will prove the slightly stronger statement that for all $R \in \mathrm{SO}(3)$

$$\forall A \in \mathcal{F}_A \cup \mathcal{F}_{A^T}, \quad \exists B \in \mathcal{F}_B \ \text{ such that } R \cdot (B - A) \geq 0. \tag{14}$$

To simplify the proof we will repeatedly make use of the symmetries of the problem [see Eq. (6)]. We will pick a canonical "A-type" facet and assume that this gives the largest inner product with R of all the $F \in \mathcal{F}_A$. If there was a larger inner product with some $F \in \mathcal{F}_{A^T}$, then we could just relabel R^T as R. We can assume that the facet with ones in the first column [the A in Eq. (7)] gives the biggest inner product since $\mathrm{Tr}[R^T(C_i A C_j)] = \mathrm{Tr}(C_j R^T C_i A) = \mathrm{Tr}[(C_k R C_l)^T A]$, which shows that the inner product of R with a different $F \in \mathcal{F}_A$ is the same as the inner product between a different (but related via Clifford operations) rotation and the canonical "A-type" facet.

The proof will hinge on an entry of R, outside of the first column, being larger than the rest of the elements outside the first column. As such, let us define 12 matrices closely related to A and call them A'_i

$$A'_1 = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad A'_2 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \cdots \quad A'_{12} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix} \tag{15}$$

such that the index i of the largest inner product $R \cdot A_i'$ tells us the sign and location of the largest magnitude element outside the first column. Once again, symmetry allows us to assume that A_1' yields the largest inner product because the rest of the A_i' can be derived from A_1' via Clifford rotations

$$\{A_i'\} = \left\{ \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}^j A_1' \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}^k \;\middle|\; \begin{array}{l} j \in \{1,2,3\} \\ k \in \{1,2,3,4\} \end{array} \right\}. \tag{16}$$

For a matrix R to be an SO(3) rotation there are constraints on the signs of the elements $R_{i,j}$; i.e., there are eight choices for the first column, 6 choices for the second column and 2 for the third. Given that A is the maximum facet for R, we have fixed the signs positively in the first column, reducing the number of types of rotation to $6 \times 2 = 12$. Since A_1' gives the maximum inner product with R of all A_i' we have that $R_{1,2} < 0$, which reduces the number of rotation types further to $3 \times 2 = 6$. It can be shown that $R_{1,2}$ having larger magnitude than the rest of the elements $R_{i,j} (i \in \{1, 2, 3\}, j \in \{2, 3\})$ restricts the type of rotation further to one the following four types

$$R \in \left\{ \begin{pmatrix} + & - & + \\ + & + & - \\ + & + & + \end{pmatrix}, \begin{pmatrix} + & - & - \\ + & + & - \\ + & + & + \end{pmatrix}, \begin{pmatrix} + & - & - \\ + & + & - \\ + & - & + \end{pmatrix}, \begin{pmatrix} + & - & + \\ + & - & - \\ + & + & + \end{pmatrix} \right\}. \tag{17}$$

This should not be surprising if one considers that $R_{1,2} = -(R_{2,1}R_{3,3} - R_{2,3}R_{3,1})$ because of the structure of SO(3) matrices, and the sign patterns listed above ensure $|R_{1,2}|$ is as large as possible.

We claim that the $B \in \mathcal{F}_B$ of Eq. (9) will suffice to prove the desired inequality $R \cdot (B - A) \geq 0$, which reads in matrix form

$$\begin{pmatrix} + & - & \cdot \\ + & \cdot & - \\ + & \cdot & + \end{pmatrix} \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{pmatrix} \geq 0. \tag{18}$$

Using the relevant entries of R we define a pair of 2 vectors $\vec{u}$ and $\vec{v}$ as $\vec{u} = (R_{1,1}, R_{1,2})$, $\vec{v} = (R_{2,3}, R_{3,3})$ so that we can rewrite the above inequality Eq. (17) as

$$\|\vec{v}\|_1 - \|\vec{u}\|_1 \geq 0. \tag{19}$$

The L^2 normalization of all the rows and columns of the rotation matrix R means that $\vec{u}$ and $\vec{v}$ have the same L^2 norm. With reference to Fig. 2, it should be clear that because $\vec{u}$ has an L^∞ norm at least as big as that of $\vec{v}$ (because $|R_{1,2}| \geq |R_{2,3}|, |R_{3,3}|$), it holds that the L^1 norm of $\vec{v}$ is automatically at least as large as the L^1 norm of $\vec{u}$, as desired.

Summary.—We showed that for any unitary one-qubit gate undergoing depolarizing noise with rate p it holds that if its SO(3) representation $M = (1 - p)R$ lies outside the Clifford polytope $\mathcal{P}$, then it must be the case that there is a facet $B \in \mathcal{F}_B$ such that $M \cdot B > 1$. In turn, this means that if this noisy gate is applied to a Bell pair $|\Phi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ and an appropriate stabilizer measurement is performed, then, conditionally on the outcome of the measurement, one obtains a state that can be transformed using Clifford gates into a single-qubit state with $|y| + |z| > 1$ in the Bloch ball representation. By the result of Reichardt [11] such states enable stabilizing operations to perform universal quantum computation.

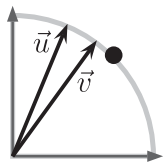


FIG. 2. Proof of Eq. (19): For any pair of two vectors $\vec{u}$ and $\vec{v}$ with the same L^2 norm, the vector with greater L^∞ norm has smaller L^1 norm. A vector pointing towards the black dot has simultaneously minimal L^∞ norm and maximal L^1 norm.

This material is based upon work supported by the National Science Foundation under Grant No. 0917244.

*vandam@cs.ucsb.edu
†mhoward@physics.ucsb.edu

[1] Michael A. Nielsen and Isaac L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).
[2] H. Buhrman, R. Cleve, M. Laurent, N. Linden, A. Schrijver, and F. Unger, *Annual IEEE Symposium on Foundations of Computer Science* (2006), p 411.
[3] E. T. Campbell and D. E. Browne, arXiv:0908.0836v2.
[4] A. R. Calderbank and P. W. Shor, Phys. Rev. A **54**, 1098 (1996).
[5] A. Steane, Proc. R. Soc. A **452**, 2551 (1996).
[6] D. Gottesman, Phys. Rev. A **57**, 127 (1998).
[7] E. H. Knill, arXiv:quant-ph/0402171v1.
[8] M. Freedman, C. Nayak, and K. Walker, Phys. Rev. B **73**, 245307 (2006).
[9] L. S. Georgiev, Phys. Rev. B **74**, 235112 (2006).
[10] S. Bravyi and A. Kitaev, Phys. Rev. A **71**, 022316 (2005).
[11] Ben W. Reichardt, Quant. Info. Proc. **4**, 251 (2005).
[12] S. Virmani, S. F. Huelga, and M. B. Plenio, Phys. Rev. A **71**, 042328 (2005).
[13] M. B. Plenio and S. Virmani, arXiv:0810.4340.
[14] B. W. Reichardt, arXiv:quant-ph/0608085v1.

A One-Way Quantum Computer

Robert Raussendorf and Hans J. Briegel
Theoretische Physik, Ludwig-Maximilians-Universität München, Germany
(Received 25 October 2000)

We present a scheme of quantum computation that consists entirely of one-qubit measurements on a particular class of entangled states, the cluster states. The measurements are used to imprint a quantum logic circuit on the state, thereby destroying its entanglement at the same time. Cluster states are thus one-way quantum computers and the measurements form the program.

DOI: 10.1103/PhysRevLett.86.5188 PACS numbers: 03.67.Lx, 03.65.Ud

A quantum computer promises efficient processing of certain computational tasks that are intractable with classical computer technology [1]. While basic principles of a quantum computer have been demonstrated in the laboratory [2], scalability of these systems to a large number of qubits [3], essential for practical applications such as the Shor algorithm, represents a formidable challenge. Most of the current experiments are designed to implement sequences of highly controlled interactions between selected particles (qubits), thereby following models of a quantum computer as a (sequential) network of quantum logic gates [4,5].

Here we propose a different model of a scalable quantum computer. In our model, the entire resource for the quantum computation is provided initially in the form of a specific entangled state (a so-called cluster state [6]) of a large number of qubits. Information is then written onto the cluster, processed, and read out from the cluster by one-particle measurements only. The entangled state of the cluster thereby serves as a universal "substrate" for any quantum computation. Cluster states can be created efficiently in any system with a quantum Ising-type interaction (at very low temperatures) between two-state particles in a lattice configuration.

We consider two- and three-dimensional arrays of qubits that interact via an Ising-type next-neighbor interaction [6] described by a Hamiltonian $H_{\mathrm{int}} = g(t) \times \sum_{\langle a,a'\rangle} \frac{1+\sigma_z^{(a)}}{2} \frac{1-\sigma_z^{(a')}}{2} \cong -\frac{1}{4} g(t) \sum_{\langle a,a'\rangle} \sigma_z^{(a)} \sigma_z^{(a')}$ [7] whose strength $g(t)$ can be controlled externally. A possible realization of such a system is discussed below. A qubit at site a can be in two states $|0\rangle_a \equiv |0\rangle_{z,a}$ or $|1\rangle_a \equiv |1\rangle_{z,a}$, the eigenstates of the Pauli phase flip operator $\sigma_z^{(a)}$ $[\sigma_z^{(a)}|i\rangle_a = (-1)^i |i\rangle_a]$. These two states form the computational basis. Each qubit can equally be in an arbitrary superposition state $\alpha|0\rangle + \beta|1\rangle$, $|\alpha|^2 + |\beta|^2 = 1$. For our purpose, we initially prepare all qubits in the superposition $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, an eigenstate of the Pauli spin flip operator σ_x $[\sigma_x|\pm\rangle = \pm|\pm\rangle]$. H_{int} is then switched on for an appropriately chosen finite time interval T, where $\int_0^T dt\, g(t) = \pi$, by which a unitary transformation S is realized. Since H_{int} acts uniformly on the lattice, entire clusters of neighboring particles become entangled in one single step. The quantum state $|\Phi\rangle_C$, the state of a cluster (C) of neighboring qubits, which is thereby created provides in advance all entanglement that is involved in the subsequent quantum computation. It has been shown [6] that the cluster state $|\Phi\rangle_C$ is characterized by a set of eigenvalue equations

$$\sigma_x^{(a)} \bigotimes_{a' \in ngbh(a)} \sigma_z^{(a')} |\Phi\rangle_C = \pm |\Phi\rangle_C \,, \qquad (1)$$

where $ngbh(a)$ specifies the sites of all qubits that interact with the qubit at site $a \in C$. The eigenvalues are determined by the distribution of the qubits on the lattice. The equations (1) are central for the proposed computation scheme. As an example, a measurement on an individual qubit of a cluster has a random outcome. On the other hand, Eqs. (1) imply that any two qubits at sites $a, a' \in C$ can be projected into a Bell state by measuring a subset of the other qubits in the cluster. This property will be used to define quantum channels that allow us to propagate quantum information through a cluster.

We show that a cluster state $|\Phi\rangle_C$ can be used as a substrate on which any quantum circuit can be imprinted by one-qubit measurements. In Fig. 1 this scheme is illustrated. For simplicity, we assume that in a certain region of the lattice each site is occupied by a qubit. This requirement is not essential as will be explained below [see (d)]. In the first step of the computation, a subset of qubits is measured in the basis of σ_z which effectively removes them. In Fig. 1 these qubits are denoted by " ⊙."

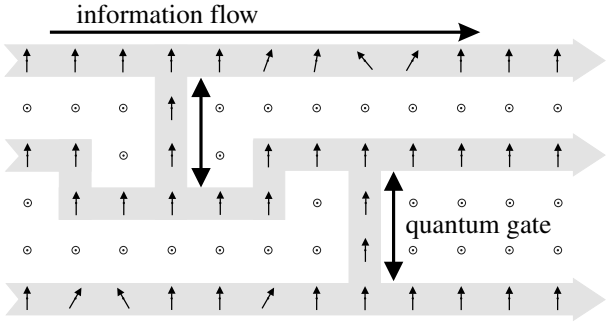


FIG. 1. Quantum computation by measuring two-state particles on a lattice. Before the measurements the qubits are in the cluster state $|\Phi\rangle_C$ of (1). Circles ⊙ symbolize measurements of σ_z, vertical arrows are measurements of σ_x, while tilted arrows refer to measurements in the x-y plane.

 0031-9007/01/86(22)/5188(4)$15.00

The state $|\Phi\rangle_C$ is thereby projected into a tensor product $|\mu\rangle_{C\setminus\mathcal{N}} \otimes |\tilde{\Phi}\rangle_{\mathcal{N}}$ consisting of the state $|\mu\rangle_{C\setminus\mathcal{N}}$ of all measured particles (subset $C\setminus\mathcal{N}$) on one side and an entangled state $|\tilde{\Phi}\rangle_{\mathcal{N}}$ of yet unmeasured particles (subset $\mathcal{N} \subset C$), on the other side. These unmeasured particles define a "network" $\mathcal{N}$ corresponding to the shaded structure in Fig. 1. The state $|\tilde{\Phi}\rangle_{\mathcal{N}}$ of the network is related to a cluster state $|\Phi\rangle_{\mathcal{N}}$ on $\mathcal{N}$ by a local unitary transformation which depends on the set of measurement results μ. More specifically, $|\tilde{\Phi}\rangle_{\mathcal{N}}$ satisfies Eqs. (1)—with C replaced by the subcluster $\mathcal{N}$—except for a possible difference in the sign factors, which are determined by the measurement results μ.

To process quantum information with this network, it suffices to measure its particles in a certain order and in a certain basis. Quantum information is thereby propagated horizontally through the cluster by measuring the qubits on the wire while qubits on vertical connections are used to realize two-bit quantum gates. The basis in which a certain qubit is measured depends in general on the results of preceding measurements. The processing is finished once all qubits except the last one on each wire have been measured. At this point, the results of previous measurements determine in which basis these "output" qubits need to be measured for the final readout. We note that, in the entire process, only one-qubit measurements are required. The amount of entanglement therefore decreases with every measurement [8] and all entanglement involved in the process is provided by the initial resource, the cluster state. This is different from the scheme of Ref. [11], which uses Bell measurements (capable of producing entanglement) to realize quantum gates.

In the following, we show that any quantum logic circuit can be implemented on a cluster state. The purpose of this is twofold. First, it serves as an illustration of how to implement a particular quantum circuit in practice. Second, in showing that any quantum circuit can be implemented on a sufficiently large cluster we demonstrate the universality of the proposed scheme. For pedagogical reasons we first explain a scheme with one essential modification with respect to the proposed scheme: before the entanglement operation S, certain qubits are selected as input qubits and the input information is written onto them, while the remaining qubits are prepared in $|+\rangle$. This step weakens the scheme since it affects the character of the cluster state as a genuine resource. It can, however, be avoided [see (e)]. Points (a) to (c) are concerned with the basic elements of a quantum circuit, quantum gates, and wires, point (d) with the composition of gates to circuits.

(a) Information propagation in a wire for qubits. A qubit can be teleported from one site of a cluster to any other site. In particular, consider a chain of an odd number of qubits 1 to n prepared in the state $|\psi_{\text{in}}\rangle_1 \otimes |+\rangle_2 \otimes \cdots \otimes |+\rangle_n$ and subsequently entangled by S. The state that was originally encoded in qubit 1, $|\psi_{\text{in}}\rangle$, is now delocalized and can be transferred to site n by performing σ_x measurements (basis $\{|+\rangle_j = |0\rangle_{x,j}, |-\rangle_j = |1\rangle_{x,j}\}$) at qubit sites $j = 1, \ldots, n-1$ with measurement outcomes $s_j \in \{0, 1\}$. The resulting state is $|s_1\rangle_{x,1} \otimes \cdots \otimes |s_{n-1}\rangle_{x,n-1} \otimes |\psi_{\text{out}}\rangle_n$. The output state $|\psi_{\text{out}}\rangle$ is related to the input state $|\psi_{\text{in}}\rangle$ by a unitary transformation $U_\Sigma \in \{1, \sigma_x, \sigma_z, \sigma_x\sigma_z\}$ which depends on the outcomes of the σ_x measurements at sites 1 to $n-1$. A similar argument can be given for an even number of qubits. The effect of U_Σ can be accounted for at the end of a computation as shown below [see (d)]. It is noteworthy that not all classical information gained by the σ_x measurements needs to be stored to identify the transformation U_Σ. Instead, U_Σ is determined by the values of only two classical bits which are updated with every measurement.

(b) An arbitrary rotation $U_R \in \text{SU}(2)$ can be achieved in a chain of five qubits. Consider a rotation in its Euler representation $U_R(\xi, \eta, \zeta) = U_x(\zeta)U_z(\eta)U_x(\xi)$, where $U_x(\alpha) = \exp(-i\alpha\frac{\sigma_x}{2}), U_z(\alpha) = \exp(-i\alpha\frac{\sigma_z}{2})$. Initially, the first qubit is in some state $|\psi_{\text{in}}\rangle$, which is to be rotated, and the other qubits are in $|+\rangle$; i.e., their common state reads $|\Psi\rangle_{1,\ldots,5} = |\psi_{\text{in}}\rangle_1 \otimes |+\rangle_2 \otimes |+\rangle_3 \otimes |+\rangle_4 \otimes |+\rangle_5$. After the five qubits are entangled by S they are in the state $S|\Psi\rangle_{1,\ldots,5} = 1/2|\psi_{\text{in}}\rangle_1|0\rangle_2|-\rangle_3|0\rangle_4|-\rangle_5 - 1/2|\psi_{\text{in}}\rangle_1|0\rangle_2|+\rangle_3|1\rangle_4|+\rangle_5 - 1/2|\psi^*_{\text{in}}\rangle_1|1\rangle_2|+\rangle_3|0\rangle_4|-\rangle_5 + 1/2|\psi^*_{\text{in}}\rangle_1|1\rangle_2|-\rangle_3|1\rangle_4|+\rangle_5$, where $|\psi^*_{\text{in}}\rangle = \sigma_z|\psi_{\text{in}}\rangle$. Now, the state $|\psi_{\text{in}}\rangle$ can be rotated by measuring qubits 1 to 4, while it is teleported to site 5 at the same time. The qubits $1, \ldots, 4$ are measured in appropriately chosen bases $\mathcal{B}_j(\alpha_j) = \{\frac{|0\rangle_j + e^{i\alpha_j}|1\rangle_j}{\sqrt{2}}, \frac{|0\rangle_j - e^{i\alpha_j}|1\rangle_j}{\sqrt{2}}\}$ whereby the measurement outcomes $s_j \in \{0, 1\}$ for $j = 1, \ldots, 4$ are obtained. Here, $s_j = 0$ means that qubit j is projected into the first state of $\mathcal{B}_j(\alpha_j)$. The resulting state is $|s_1\rangle_{\alpha_1,1} \otimes |s_2\rangle_{\alpha_2,2} \otimes |s_3\rangle_{\alpha_3,3} \otimes |s_4\rangle_{\alpha_4,4} \otimes |\psi_{\text{out}}\rangle_5$ with $|\psi_{\text{out}}\rangle = U|\psi_{\text{in}}\rangle$. For the choice $\alpha_1 = 0$ (measuring σ_x of qubit 1) the rotation U has the form $U = \sigma_x^{s_2+s_4}\sigma_z^{s_1+s_3}U_R[(-1)^{s_1+1}\alpha_2, (-1)^{s_2}\alpha_3, (-1)^{s_1+s_3}\alpha_4]$. In summary, the procedure to implement an arbitrary rotation $U_R(\xi, \eta, \zeta)$, specified by its Euler angles ξ, η, ζ is (i) measure qubit 1 in $\mathcal{B}_1(0)$; (ii) measure qubit 2 in $\mathcal{B}_2((-1)^{s_1+1}\xi)$; (iii) measure qubit 3 in $\mathcal{B}_3((-1)^{s_2}\eta)$; (iv) measure qubit 4 in $\mathcal{B}_4((-1)^{s_1+s_3}\zeta)$. In this way the rotation U'_R is realized: $U'_R(\xi, \eta, \zeta) = \sigma_x^{s_2+s_4}\sigma_z^{s_1+s_3}U_R(\xi, \eta, \zeta)$. The extra rotation $U_\Sigma = \sigma_x^{s_2+s_4}\sigma_z^{s_1+s_3}$ can be accounted for at the end of the computation, as is described below in (d).

(c) To perform the gate $\text{CNOT}(c, t_{\text{in}} \to t_{\text{out}}) = |0\rangle_{cc}\langle 0| \otimes 1^{(t_{\text{in}}\to t_{\text{out}})} + |1\rangle_{cc}\langle 1| \otimes \sigma_x^{(t_{\text{in}}\to t_{\text{out}})}$ between a control qubit c and a target qubit t, four qubits, arranged as depicted Fig. 2a, are required. During the action of the gate, the target qubit t is transferred from t_{in} to t_{out}. The following procedure has to be implemented. Let qubit 4 be the control qubit. First, the state $|i_1\rangle_{z,1} \otimes |i_4\rangle_{z,4} \otimes |+\rangle_2 \otimes |+\rangle_3$ is prepared and then the entanglement operation S is performed. Second, σ_x of qubits 1 and 2 is measured. The measurement results

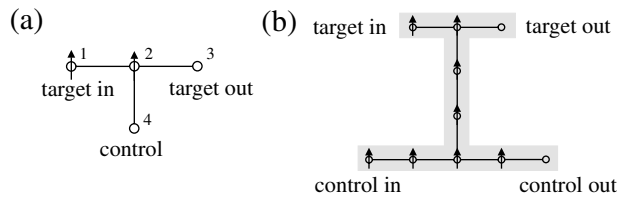


FIG. 2. Realization of a CNOT gate by one-particle measurements. See text.

$s_j \in \{0, 1\}$ correspond to projections of the qubits j into $|s_j\rangle_{x,j}$, $j = 1, 2$. The quantum state created by this procedure is $|s_1\rangle_{x,1} \otimes |s_2\rangle_{x,2} \otimes U_{\Sigma}^{(34)} |i_4\rangle_{z,4} \otimes |i_1 + i_4 \bmod 2\rangle_{z,3}$, where $U_{\Sigma}^{(34)} = \sigma_z^{(3)^{s_1+1}} \sigma_x^{(3)^{s_2}} \sigma_z^{(4)^{s_1}}$. The input state is thus acted upon by the CNOT and successive σ_x and σ_z rotations $U_{\Sigma}^{(34)}$, depending on the measurement results s_1, s_2. These unwanted extra rotations can again be accounted for as described in (d). For practical purposes it is more convenient if the control qubit is, as the target qubit, transferred to another site during the action of the gate. When a CNOT is combined with other gates to form a quantum circuit it will be used in the form shown in Fig. 2b.

To explain the working principle of the CNOT gate we, for simplicity, refer to the minimal implementation with four qubits. The minimal CNOT can be viewed as a wire from qubit 1 to qubit 3 with an additional qubit, No. 4, attached. From the eigenvalue equations (1) it can now be derived that, if qubit 4 is in an eigenstate $|i_4\rangle_{z,4}$ of σ_z, then the value of $i_4 \in \{0, 1\}$ determines whether a unit wire or a spin flip σ_x (modulo the same correction $U_{\Sigma}^{(3)}$ for both values of i_4) is being implemented. In other words, once σ_x of qubits 1 and 2 have been measured, the value i_4 of qubit 4 controls whether the target qubit is flipped or not.

(d) Quantum circuits. The gates described—the CNOT and arbitrary one-qubit rotations—form a universal set [5]. In the implementation of a quantum circuit on a cluster state the site of every output qubit of a gate overlaps with the site of an input qubit of a subsequent gate. Because of this, the entire entanglement operation can be performed at the beginning. To see this, compare the following two strategies. Given a quantum circuit implemented on a network $\mathcal{N}$ of qubits which is divided into two consecutive circuits, circuit 1 is implemented on network $\mathcal{N}_1$ and circuit 2 is implemented on network $\mathcal{N}_2$, and $\mathcal{N} = \mathcal{N}_1 \cup \mathcal{N}_2$. There is an overlap $\mathcal{O} = \mathcal{N}_1 \cap \mathcal{N}_2$ which contains the sites of the output qubits of circuit 1 (these are identical to the sites of the input qubits of circuit 2). The sites of the readout qubits form a set $\mathcal{R} \subset \mathcal{N}_2$. Strategy (i) consists of the following steps: (1) write input and entangle all qubits on $\mathcal{N}$; (2) measure qubits $\in \mathcal{N} \backslash \mathcal{R}$ to implement the circuit. Strategy (ii) consists of (1) write input and entangle the qubits on $\mathcal{N}_1$, (2) measure the qubits in $\mathcal{N}_1 \backslash \mathcal{O}$. This implements the circuit on $\mathcal{N}_1$ and writes the intermediate output to $\mathcal{O}$; (3) entangle the qubits on $\mathcal{N}_2$; (4) measure all qubits in $\mathcal{N}_2 \backslash \mathcal{R}$. Steps 3 and 4 implement the circuit 2 on $\mathcal{N}_2$. The measurements on $\mathcal{N}_1 \backslash \mathcal{O}$ commute with the entanglement operation restricted to $\mathcal{N}_2$, since they act on different subsets of particles. Therefore the two strategies are mathematically equivalent and yield the same results. It is therefore consistent to entangle in a single step at the beginning and perform all measurements afterwards.

Two further points should be addressed in connection with circuits. First, the randomness of the measurement results does not jeopardize the function of the circuit. Depending on the measurement results, extra rotations σ_x and σ_z act on the output qubit of every implemented gate. By use of the relations $U_R(\xi, \eta, \zeta) \sigma_z^s \sigma_x^{s'} = \sigma_z^s \sigma_x^{s'} U_R((-1)^s \xi, (-1)^{s'} \eta, (-1)^s \zeta)$, and $\mathrm{CNOT}(c, t) \sigma_z^{(t)^{s_t}} \sigma_z^{(c)^{s_c}} \sigma_x^{(t)^{s'_t}} \sigma_x^{(c)^{s'_c}} = \sigma_z^{(t)^{s_t}} \sigma_z^{(c)^{s_c + s_t}} \sigma_x^{(t)^{s'_c + s'_t}} \sigma_x^{(c)^{s'_c}} \mathrm{CNOT}(c, t)$, these extra rotations can be pulled through the network to act upon the output state. There they can be accounted for by adjusting the measurement basis for the final readout. The above relations imply that for a rotation $U_R(\xi, \eta, \zeta)$—different from the CNOT gate—the accumulated extra rotations U_Σ at the input side of U_R need to be determined before the measurement bases that realize U_R can be specified. This introduces a partial temporal ordering of the measurements on the whole cluster. Second, quantum circuits can also be implemented on irregular clusters. In that case, qubits may be missing which are required for the standard implementation of the circuit. This can be compensated by a large flexibility in shape of the gates and wires. The components can be bent and stretched to fit to the cluster structure as long as the topology of the circuit implementation does not change. Irregular clusters are found in lattices with a finite site occupation probability $0 < p < 1$. In such a situation, the possibility of *universal* quantum computation is closely linked to the phenomenon of percolation. For p above a certain critical value p_c, which depends on the dimension of the lattice, an infinitely extended cluster exists that may be used as the carrier of the quantum circuit. In two dimensions, for example, exactly one such cluster C exists. Suppose this cluster is divided into two subclusters C_1 and C_2 by a one-dimensional cut $\mathcal{O} = C_1 \cap C_2$. It can be shown, e.g., by using Russo's formula [12] from percolation theory that, for any cut $\mathcal{O}$, $|\mathcal{O}| = \infty$. Therefore there is no upper bound, in principle, to the "capacity" of the cluster, i.e., to the number of qubits that can be processed across such a cut.

(e) Full scheme. It is important to note that the step of writing the input information onto the qubits before the cluster is entangled was introduced only for pedagogical reasons. For illustration of this point consider a chain of five qubits in the state $S|+\rangle_1 \otimes |+\rangle_2 \otimes \cdots \otimes |+\rangle_5$. Clearly, there is no local information on any of the qubits. However, by measuring qubits 1 to 4 along suitable directions, qubit 5 can be projected into any desired state (modulo U_Σ). What is used here is the knowledge that the

resource has been prepared with qubit 1 in the state $|+\rangle_1$ before the entanglement operation. By the four measurements, this qubit is then rotated as described in (b). In order to use qubit 5 for further processing, the five-qubit chain considered here should, of course, be part of a larger cluster such that particle 5 is still entangled with the remaining network, after particles 1 to 4 have been measured. The method of preparing the input state remains the same, in this case, as explained in (d). In a similar manner any desired input state can be prepared if the rotations are replaced by a circuit preceding the proper circuit for computation. In summary, no input information needs to be written to the qubits before they are entangled. Cluster states are thus a genuine resource for quantum computation via measurements only.

For a cluster of a given *finite* size, the number of computational steps may be too large to fit on the cluster. In this case, the computation can be split into consecutive parts, for each of which there is sufficient space on the cluster. The modified procedure consists then of repeatedly (re)entangling the cluster and imprinting the actual part of the circuit — by measuring all of the lattice qubits except the ones carrying the intermediate quantum output — until the whole calculation is performed. This procedure has also the virtue that qubits involved in the later part of a calculation need not be protected from decoherence for a long time while the calculation is still being performed at a remote place of the cluster. Standard error-correction techniques [13,14] may then be used on each part of the circuit to stabilize the computation against decoherence.

A possible implementation of such a quantum computer uses neutral atoms stored in periodic micropotentials [15–18] where Ising-type interactions can be realized by controlled collisions between atoms in neighboring potential wells [16,18]. This system combines small decoherence rates with a high scalability. The question of scalability is linked to the percolation phenomenon, as mentioned earlier. For a site occupation probability above the percolation threshold, there exists a cluster which is bounded in size only by the trap dimensions. For optical lattices in three dimensions, single-atom site occupation with a filling factor of 0.44 has been reported [19] which is significantly above the percolation threshold of 0.31 [20]. As in other proposed implementations for quantum computing, the addressability of single qubits in the lattice is, however, still a problem. (For recent progress, see Ref. [21]). Recently, it has also been shown that implementations based on arrays of capacitively coupled quantum dots may be used to realize an Ising-type interaction [22].

In conclusion, we have described a new scheme of quantum computation that consists entirely of one-qubit measurements on a particular class of entangled states, the cluster states. The measurements are used to imprint a quantum circuit on the state, thereby destroying its entanglement at the same time. Cluster states are thus one-way quantum computers and the measurements form the program.

We thank D. E. Browne, D. P. DiVincenzo, A. Schenzle, and H. Wagner for helpful discussions. This work was supported by the Deutsche Forschungsgemeinschaft.

[1] C. H. Bennett and D. P. DiVincenzo, Nature (London) **404**, 247 (2000).
[2] See Ref. [1] for a recent review.
[3] J. I. Cirac and P. Zoller, Nature (London) **404**, 579 (2000).
[4] D. Deutsch, Proc. R. Soc. London **425**, 73 (1989).
[5] A. Barenco *et al.,* Phys. Rev. A **52**, 3457 (1995).
[6] H.-J. Briegel and R. Raussendorf, Phys. Rev. Lett. **86**, 910 (2001).
[7] The second Hamiltonian is of the standard Ising form. The symbol "≅" means that the states generated from a given initial state, under the action of these Hamiltonians, are identical up to a local rotation on certain qubits. We use the first Hamiltonian to make the computational scheme more transparent. The conclusions drawn in the paper are, however, the same for both Hamiltonians.
[8] By the "amount of entanglement" contained in the resource, we mean any measure that satisfies the criteria of an entanglement monotone [9]. For cluster states, the entanglement can be calculated, e.g., in terms of the Schmidt measure of Ref. [10].
[9] G. Vidal, J. Mod. Opt. **47**, 355 (2000).
[10] J. Eisert and H.-J. Briegel, quant-ph/0007081.
[11] D. Gottesman and I. L. Chuang, Nature (London) **402**, 390 (1999).
[12] See, e.g., G. Grimmett, *Percolation* (Springer-Verlag, New York, 1989).
[13] A. Calderbank and P. W. Shor, Phys. Rev. A **54**, 1098 (1996).
[14] A. Steane, Phys. Rev. Lett. **77**, 793 (1996).
[15] G. K. Brennen *et al.,* Phys. Rev. Lett. **82**, 1060 (1999).
[16] D. Jaksch *et al.,* Phys. Rev. Lett. **82**, 1975 (1999).
[17] T. Calarco *et al.,* Phys. Rev. A **61**, 022304 (2000).
[18] H.-J. Briegel *et al.,* J. Mod. Opt. **47**, 415 (2000).
[19] M. T. DePue *et al.,* Phys. Rev. Lett. **82**, 2262 (1999).
[20] J. M. Ziman, *Models of Disorder* (Cambridge University Press, Cambridge, United Kingdom, 1979).
[21] R. Scheunemann *et al.,* Phys. Rev. A **62**, 051801(R) (2000).
[22] T. Tanamoto, quant-ph/0009030.

nature physics

LETTERS

PUBLISHED ONLINE: 20 JULY 2009 | DOI: 10.1038/NPHYS1342

Quantum computation and quantum-state engineering driven by dissipation

Frank Verstraete[1]*, Michael M. Wolf[2] and J. Ignacio Cirac[3]*

The strongest adversary in quantum information science is decoherence, which arises owing to the coupling of a system with its environment[1]. The induced dissipation tends to destroy and wash out the interesting quantum effects that give rise to the power of quantum computation[2], cryptography[2] and simulation[3]. Whereas such a statement is true for many forms of dissipation, we show here that dissipation can also have exactly the opposite effect: it can be a fully fledged resource for universal quantum computation without any coherent dynamics needed to complement it. The coupling to the environment drives the system to a steady state where the outcome of the computation is encoded. In a similar vein, we show that dissipation can be used to engineer a large variety of strongly correlated states in steady state, including all stabilizer codes, matrix product states[4], and their generalization to higher dimensions[5].

The situation we have in mind is shown in Fig. 1. A quantum system composed of N particles (such as qubits) is organized in space according to a particular geometry (in the figure, a one-dimensional lattice). Neighbouring systems are coupled to some local environments, which are dissipative in nature and tend to drive the system to a steady state. Our idea is to engineer those couplings, so that the environments drive the system to a desired final state. The coupling to the environment will be static, so that the desired state is obtained after some time without having to actively control the system. Note that the role of the environments is to dissipate (or, more precisely, evacuate) the entropy of the system, and by choosing the couplings appropriately we can use this effect to drive our system.

We will show first how to design the interactions with the environment to implement universal quantum computation. This new method, which we refer to as dissipative quantum computation (DQC), defies some of the standard criteria for quantum computation because it requires neither state preparation, nor unitary dynamics[6]. However, it is nevertheless as powerful as standard quantum computation. Then we will show that dissipation can be engineered[7] to prepare ground states of frustration-free Hamiltonians. Those include matrix product states[4,8,9] (MPSs) and projected entangled pair states[5,9] (PEPSs), such as graph states[10] and Kitaev[11] and Levin–Wen[12] topological codes. Both DQC and dissipative state engineering (DSE) are robust in the sense that, given the dissipative nature of the process, the system is driven towards its steady state independent of the initial state and hence of eventual perturbations along the way.

Here, we will concentrate first on DQC, showing how given any quantum circuit one can construct a locally acting master equation for which the steady state is unique, encodes the outcome of the circuit and is reached in polynomial time (with respect to the one corresponding to the circuit). Then we will show how to construct dissipative processes that drive the system to the ground state of any frustration-free Hamiltonian. In the Methods section, we will prove that MPS (ref. 9) and certain kinds of PEPS (ref. 9) can be efficiently prepared using this method, and in Supplementary Information we will give details of the proofs. In this letter we will not consider specific physical set-ups where our ideas can be implemented. Nevertheless, the Methods section will provide a universal way of engineering the master equations required for DQC and DSE, which can be easily adapted to current experiments[13] based on, for example, atoms in optical lattices[14] or trapped ions[15]. Thus, we expect that our predictions may be experimentally tested in the near future.

Let us start with DQC by considering N qubits in a line and a quantum circuit specified by a sequence of nearest-neighbour qubit operations $\{U_t\}_{t=1}^{T}$. We define $|\psi_t\rangle := U_t U_{t-1} \ldots U_1 |0\rangle_1 \otimes \ldots |0\rangle_N$, so that $|\psi_T\rangle$ is the final state after the computation. Our goal is to find a master equation $\dot{\rho} = \mathcal{L}(\rho)$ with a Liouvillian in Lindblad form[16]

$$\mathcal{L}(\rho) = \sum_k L_k \rho L_k^\dagger - \frac{1}{2}\left\{L_k^\dagger L_k, \rho\right\}_+ \qquad (1)$$

where the L_k acts locally and has a steady state, ρ_0: (1) that is unique; (2) that can be reached in a time poly(T); (3) such that ψ_T can be extracted from it in a time poly(T). As in Feynman's construction of a quantum simulator[3], we consider another auxiliary register with states $\{|t\rangle\}_{t=0}^{T}$, which will represent the time. We choose the Lindblad operators

$$L_i = |0\rangle_i\langle 1| \otimes |0\rangle_t\langle 0|$$

$$L_t = U_t \otimes |t+1\rangle\langle t| + U_t^\dagger \otimes |t\rangle\langle t+1|$$

where $i = 1, \ldots, N$ and $t = 0, \ldots, T$. It is clear that the L terms act locally except for the interaction with the extra register, which can be made local as well. Furthermore,

$$\rho_0 = \frac{1}{T+1}\sum_t |\psi_t\rangle\langle\psi_t| \otimes |t\rangle\langle t|$$

is a steady state, that is, $\mathcal{L}(\rho_0) = 0$. Given such a state, the result of the actual quantum computation can be read out with probability $1/T$ by measuring the time register. In Supplementary Information, we show that ρ_0 is the unique steady state and that the Liouvillian has a spectral gap $\Delta = \pi^2/(2T+3)^2$. This means indeed that the steady state will be reached in polynomial time in T. Note that this gap is independent of N as well as of the actual quantum computation that is carried out (that is, independent of the U_t). It is also shown that the same gap is retained if the clock register is encoded in the unary

[1]Fakultät für Physik, Universität Wien, 1090 Wien, Austria, [2]Niels Bohr Institute, 2100 Copenhagen, Denmark, [3]Max-Planck-Institut für Quantenoptik, 85748 Garching, Germany. *e-mail: fverstraete@gmail.com; ignacio.cirac@mpq.mpg.de.

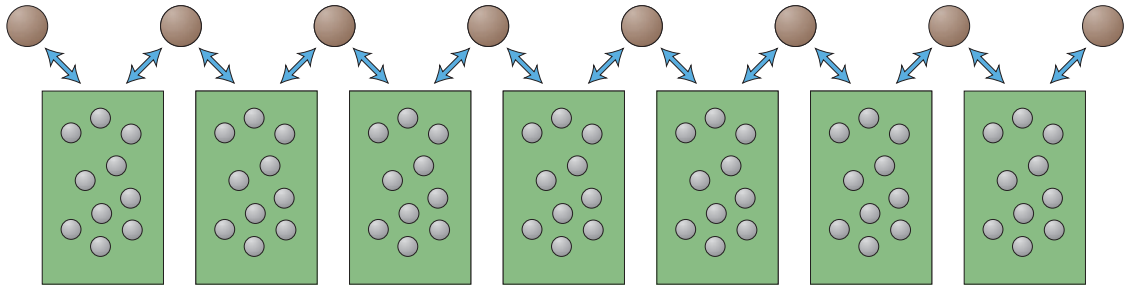


Figure 1 | Schematic representation of the set-up. We consider a collection of N quantum particles, locally coupled to a set of environments. The couplings are engineered in such a way that the system reaches the desired state in the long-time limit.

way proposed by Kitaev and co-workers[17], making the Lindblad operators strictly local. A sketch of the proof is as follows. First, we do a similarity transformation on $\mathcal{L}$ that replaces all gates U_i with the identity gates, showing that its spectrum is independent of the actual quantum computation. Second, another similarity transformation is done that makes $\mathcal{L}$ Hermitian and block-diagonal. Each block can then be diagonalized exactly leading to the claimed gap.

In some sense, the present formalism can be seen as a robust way of doing adiabatic quantum computation[18] (errors do not accumulate and the path does not have to be engineered carefully) and implementing quantum random walks[19], and it might therefore be easier to tackle interesting open questions, such as the quantum probabilistically-checkable-proofs theorem, in this setting[20]. In addition, it seems that the dissipative way of preparing ground states is more natural than to use adiabatic time evolution, as nature itself prepares them by cooling.

Let us now turn to DSE and consider again a quantum system with N particles on a lattice in any dimension. We are interested in ground states Ψ, of Hamiltonians

$$H = \sum_{\lambda} H_{\lambda}$$

that are frustration-free, meaning that Ψ minimizes the energy of each H_λ individually, and local in the sense that H_λ acts non-trivially only on a small set $\lambda \subset \{1,\ldots,N\}$ of sites (for example, nearest neighbours). We can assume the terms H_λ to be projectors and we will denote the orthogonal projectors by $P_\lambda = \mathbf{1} - H_\lambda$. States Ψ of the considered form are, for example, all PEPS (including MPS and stabilizer states[21]).

We will consider discrete time evolution generated by a trace-preserving completely positive map instead of a master equation. These two approaches are basically equivalent[22] as every local completely positive map $\mathcal{T}$ can be associated with a local Liouvillian through $\mathcal{L}(\rho) = N[\mathcal{T}(\rho) - \rho]$, which leads to the same fixed points and spectrum. We choose completely positive maps of the form

$$\mathcal{T}(\rho) = \sum_{\lambda} p_{\lambda} \left[P_{\lambda}\rho P_{\lambda} + \frac{1}{m}\sum_{i=1}^{m} U_{\lambda,i} H_{\lambda}\rho H_{\lambda} U_{\lambda,i}^{\dagger} \right] \qquad (2)$$

where the p_λ terms are probabilities and $U_{\lambda,1},\ldots,U_{\lambda,m}$ is a set of unitaries acting non-trivially only within region λ. They effectively rotate part of the high-energy space (with support of H_λ) to the zero-energy space, so that $\mathrm{tr}[\mathcal{T}(\rho)\Psi] \geq \mathrm{tr}[\rho\Psi]$ increases. As for Liouvillians (1), we could similarly take $L_{\lambda,i} = U_i H_\lambda$, or the ones associated with the completely positive map.

We show now that for every frustration-free Hamiltonian, the completely positive map in equation (2) converges to the ground-state space if we choose the unitaries $U_{\lambda,i}$ to be completely depolarizing, that is, $\mathcal{T}(\rho) \propto \sum_\lambda P_\lambda \rho P_\lambda + \mathbf{1}_\lambda \otimes \mathrm{tr}_\lambda[H_\lambda\rho]/\mathrm{tr}[\mathbf{1}_\lambda]$. For ease of notation, we will explain the proof for the case of a one-dimensional ring with nearest-neighbour interactions labelled by the first site $\lambda = 1,\ldots,N$. Assume ρ is such that its expectation value with respect to the projector Ψ onto the ground-state space of H is non-increasing under applications of $\mathcal{T}$, that is, in particular $\mathrm{tr}[\rho\Psi] = \mathrm{tr}[\mathcal{T}^N(\rho)\Psi]$. Expressing this in the Heisenberg picture in which $\mathcal{T}^*(\Psi) = \Psi + \sum_\lambda H_\lambda \mathrm{tr}_\lambda(\Psi)/(d^2 N)$, we get

$$\begin{aligned} \mathrm{tr}[\rho\Psi] &\geq \mathrm{tr}[\rho\Psi] + \frac{1}{(d^2N)^N}\mathrm{tr}\left[\rho \sum_{\mu=1}^{N}\prod_{\lambda=1}^{N}\left(H_{\lambda+\mu}\mathrm{tr}_{\lambda+\mu}\right)(\Psi)\right] \\ &\geq \mathrm{tr}[\rho\Psi] + \frac{\nu^N}{(d^2N)^N}\mathrm{tr}[\rho H] \end{aligned}$$

where the first inequality comes from discarding (positive) terms in the sum and the second one is due to bounding all partial traces of H_λ from below by the respective smallest eigenvalue ν. Note that the latter is strictly positive unless H has a product state as the ground state (in which case the statement becomes trivial). Hence, we must have $\mathrm{tr}[\rho H] = 0$; that is, ρ is a ground state of H. It is easily seen that the same argument applies for more general interactions on arbitrary lattices.

Once we have shown that the steady state after the application of the completely positive map lies within the desired subspace (the ground-state space of the frustration-free Hamiltonian), the next question to be addressed is how efficient the process is. This depends on the spectral gap, δ, of the completely positive map (or, equivalently, of the corresponding Liouvillian), as the time to reach the steady state, $\tau = \mathcal{O}(1/\delta)$. Thus, the above procedure will be efficient as long as the gap vanishes only polynomially with the number of systems, N. Similarly to what occurs with many-body Hamiltonians, the determination of such a gap is, in general, very complicated. For a wide range of interesting models, however, it can be proved that this gap scales favourably. This is the case for all MPS as well as for a rich subfamily of PEPS that includes all stabilizer states (such as Kitaev's toric code[11] and the Levin–Wen states[12]). In the Methods section, we characterize such a subfamily of states, and in Supplementary Information we give the technical proofs of our statements. Here, we will qualitatively explain how our method works efficiently for some families of states. For that we note that the action of the completely positive map (2) can be interpreted as randomly choosing a region λ (according to p_λ, which we may set equal to $1/N$), then measuring P_λ and applying a correction according to the unitaries if the outcome was negative. We denote by R_n the set of regions λ where φ satisfies the condition $H_\lambda|\varphi\rangle = 0$. If we measure now in one of those regions, we will obviously obtain a positive result, and thus R_n will remain the same. If we measure in another region, we may have a positive or negative result, something that may change the set R_n. By imposing certain conditions on the operators H_λ and $U_{\lambda,i}$, we can make sure that in each step R_n cannot be reduced and that the probability of

being enlarged is non-vanishing. This automatically ensures that the τ scales only polynomially with the number of systems. In one dimension, however, one can get rid of all those restrictions and show that any MPS can be prepared in a time that also scales favourably with N. The fact that all MPS states can be prepared with our method, together with the results reported in refs 23, 24, automatically implies the existence of phase transitions driven by dissipation in the following sense. By changing the parameters of the operators H_λ appearing in the completely positive map (2), we change the steady state of that map. It is possible to choose models for which that state changes abruptly at some particular value of that parameter in such a way that the correlation length diverges and an order parameter appears (an example can be found in the Supplementary Information).

We have investigated the computational power of purely dissipative processes, and proved that it is equivalent to that of the quantum circuit model of quantum computation. We have also shown that dissipative dynamics can be used to create ground states (such as MPS or PEPS) of frustration-free Hamiltonians of strongly correlated quantum spin systems. We believe that these new methods can be experimentally tested using atoms or ions with current set-ups (see the Methods section).

Let us stress that we have been concerned here with a proof-of-principle demonstration that dissipation provides us with an alternative way of carrying out quantum computations or state engineering. We believe, however, that much more efficient and practical schemes can be developed and adapted to specific implementations. We also think that these results open up some interesting questions that deserve further investigation: for example, how the use of fault-tolerant computations can make our scheme more robust, or how one can design translationally invariant completely positive maps that prepare MPS more efficiently, or the importance and generality of the set of commuting Hamiltonians (see the Methods section), which is intimately connected to the fixed points of the renormalization group transformations on PEPS (as it happens with MPS; ref. 25). Furthermore, the model of DQC might well lead to the construction of new quantum algorithms, as, for example, quantum random walks can more easily be formulated within this context. Finally, other ideas related to this work can be easily addressed using the methods introduced; for example, thermal states of commuting Hamiltonians can be engineered using DSE because the Metropolis way of sampling over classical spin configurations can be adopted to the case of commuting operators. Similar techniques could be applied to free fermionic and bosonic systems, and, more generally, it should be possible to devise DSE schemes converging to the ground or thermal states of frustrated Hamiltonians by combining unitary and dissipative dynamics.

Note added. Concurrently with the submission of this paper, refs 26 and 27 appeared in which a similar quantum-reservoir engineering was used to prepare many-body states and non-equilibrium quantum phases.

Methods

Engineering dissipation. Here we show how to engineer the local dissipation that gives rise to the master equations (1) and completely positive maps (2). They are composed of local terms, involving few particles (typically two), so that we just have to show how to implement those. To simplify the exposition, we will treat those particles as a single one and assume that one has full control over its dynamics (for example, one can apply arbitrary gates).

Let us start with the completely positive maps. It is clear that by applying a quantum gate to the particle and a 'fresh' ancilla and then tracing the ancilla one can generate any physical action (that is, completely positive map) on the system. Furthermore, by repeating the same process with short time intervals one can subject the system to an arbitrary time-independent master equation. This last process may not be efficient. An alternative way works as follows. Let us assume that the ancilla is a qubit interacting with a reservoir such that it fulfils a master equation with Liouville operator $L_a=\sqrt{\Gamma}\sigma_-$, where $\sigma_-=|0\rangle\langle 1|$. Now, we couple the ancilla to the system with a Hamiltonian $H=\Omega(\sigma_- L^\dagger+\sigma_-^\dagger L)$. In the limit $\Gamma\gg\Omega$, one can adiabatically eliminate the level $|1\rangle$ of the ancilla[28] by applying second-order perturbation theory to the Liouvillian (albeit for non-Hermitian operators). In this way we obtain an effective master equation for ρ describing the system alone, with Liouville operator $\Omega/\sqrt{\Gamma}L$. By using several ancillas with Hamiltonians $H=\Omega(\sigma_- L_i+\sigma_-^\dagger L_i^\dagger)$ and following the same procedure we obtain the desired master equation. Although we have not specified here a physical system, one could use atoms. In that case, the ancilla could be an atom itself with $|0\rangle$ and $|1\rangle$ an electronic ground and excited level, respectively, so that spontaneous emission gives rise to the dissipation. The coupling to the system (other atoms) could be achieved using standard ideas used in the implementation of quantum computation using those systems[13].

Efficient state preparation. We have shown that it is possible to engineer dissipative processes that prepare ground states of frustration-free Hamiltonians in steady state. In the proof, the time for this preparation scales as N^N, which may be an issue for experiments with large number of particles. Here we give much more efficient methods for certain classes of frustration-free Hamiltonians.

We consider first frustration-free Hamiltonians for which $[H_\lambda,H_\mu]=0$ and show that, under certain conditions, the corresponding ground states can be prepared in a time that scales only polynomially with the number of particles. The corresponding set of ground states contains important families, such as stabilizer states (for example, cluster states and topological codes), or certain kinds of PEPS, namely, those that have (commuting) parent Hamiltonians with the injectivity condition (as defined in refs 8, 29). Note that there was no known way of efficient preparation for the latter.

Loosely speaking, we will consider two classes of Hamiltonians. (1) Hamiltonians for which all excitations can be locally annihilated. In this case the time of convergence scales as $\tau=\mathcal{O}(\log N)$. (2) Interactions where excitations have to be moved along the lattice before they can annihilate and $\tau=\mathcal{O}(N\log N)$.

To see how the first case can occur notice that, when iterating $\mathcal{T}$, the correction on λ does not change the outcome of previous measurements on neighbouring regions because

$$\forall\lambda\neq\lambda'\colon [U_{\lambda,i},H_{\lambda'}]=0 \tag{3}$$

In fact, this can always be achieved by regrouping the regions into larger ones having an interior $I(\lambda)\subset\lambda$ on which only H_λ acts non-trivially and letting the $U_{\lambda,i}$ solely act on $I(\lambda)$. Denote by q the largest probability for obtaining twice a negative measurement outcome on the same region λ. The energy $\mathrm{tr}[H\mathcal{T}^M(\rho)]$ after M applications of $\mathcal{T}$ decreases then as $N(1-(1-q)/N)^M$ such that it takes $\mathcal{O}((N\log N)/(1-q))$ steps to converge to a ground state. The relaxation time of the corresponding Liouvillian is thus $\tau=\mathcal{O}(\log N^{1/1-q})$. Clearly, this is a reasonable bound only if $q<1$, a condition possibly incompatible with equation (3).

Note that for all stabilizer states we can achieve $q=0$, because there exists always a local unitary (acting on a single qubit) so that $H_\lambda U_\lambda H_\lambda=0$. A class of stabilizer states where this is compatible with equation (3) are the so-called graph states[10]. In this case, λ labels (with some abuse of notation) a vertex of a graph and $H_\lambda=(\mathbf{1}-\sigma_x^{(\lambda)}\prod_{(\lambda,\mu)\in\mathcal{E}}\sigma_z^{(\mu)})/2$, where $\sigma^{(\lambda)}$ is a Pauli operator acting on site λ and $\mathcal{E}$ is the set of edges of the graph. Obviously, $U_\lambda=\sigma_z^{(\lambda)}$ does the job. In this special case, we can get even faster convergence when using the Liouvillian

$$\mathcal{L}(\rho)=\left(\sum_\lambda U_\lambda H_\lambda\rho H_\lambda U_\lambda^\dagger\right)-\frac{1}{2}\left\{H,\rho\right\}_+$$

The corresponding relaxation time can be determined exactly by realizing that the spectrum of $\mathcal{L}$ equals that of $-(H\otimes\mathbf{1}+\mathbf{1}\otimes H)/2$ so that $\tau=1$ (see Supplementary Information).

For the second type of commuting Hamiltonians, equation (3) and $q<1$ are incompatible. However, we can still prove fast convergence by relaxing equation (3) such that within each region λ the U_λ acts on a site closest to a predetermined site (say the origin) on the lattice and thus commutes with all terms H_λ that are further away (see Supplementary Information for details). In this way excitations are moved over the lattice before they can annihilate. As this requires extra time proportional to the system's size, we get $\tau=\mathcal{O}(N\log N)$.

We turn now to another family of ground states of frustration-free Hamiltonians, namely MPS (ref. 9). For the sake of clearness, we will consider here translationally invariant Hamiltonians, although the analysis can be straightforwardly extended to systems without that symmetry. We will specify a completely positive map to prepare states of the form

$$|\Psi\rangle=\sum_{i=1}^{d}\mathrm{tr}(A_{i_1}\ldots A_{i_N})|i_1\ldots i_N\rangle$$

where the A terms are $D\times D$ matrices. We assume the injectivity property[29], which implies that Ψ is the unique ground state of a nearest-neighbour frustration-free

'parent' Hamiltonian that has a gap. Denoting by ρ the reduced density operator corresponding to particles k and $k+1$, H_k and $P_k = 1 - H_k$ will denote the projectors onto its kernel and range, respectively. Note that $\mathrm{tr}(P_k) = D^2$. We take $N = 2^n$ for simplicity, but this is clearly not necessary. We construct the channel $\mathcal{T}$ in several steps. We first define a channel acting on two neighbouring particles $k, k+1$, as follows

$$\mathcal{R}_{r,c}(X) := P_k X P_k + \frac{P_k}{D^2}\mathrm{tr}(H_k X)$$

Here, $k = 2^{r-1}(2c-1)$, where $r = 1,\ldots,n$ and $c = 1,\ldots,2^{n-r}$. The action of these maps has a tree structure, where the index r indicates the row in the tree, whereas c does it for the column. Now we define recursively,

$$\mathcal{S}_{r,c} := \frac{(1-\epsilon_r)}{2}(\mathcal{S}_{r-1,2c} + \mathcal{S}_{r-1,2c+1}) + \epsilon_r \mathcal{R}_{r,c}$$

Here, $r = 2,\ldots,n$, $c = 1,\ldots,2^{n-r}$, $\mathcal{S}_{1,c} := \mathcal{R}_{1,c}$ and $\epsilon_{r+1} = 1/M^r$, where $M = CN^2$ and $C \gg 1$ (see Supplementary Information). Note that $\mathcal{S}_{r,1}$ acts on the first 2^r particles, $\mathcal{S}_{r,2}$ on the next 2^r and so on. We finally define

$$\mathcal{T} := (1-\epsilon_{n+1})\mathcal{S}_{n,1} + \epsilon_{n+1}\mathcal{R}_{n,2} \tag{4}$$

In the Supplementary Information, we show that this map achieves the fixed point (up to an exponentially small error in C) in a time $\mathcal{O}(N^{\log_2(N)})$. The intuition behind the completely positive map (4) is that the channels $\mathcal{S}_{1,c}$, which are the ones that most often applied, project the state of every second nearest neighbour onto the right subspace. Then $\mathcal{S}_{2,c}$ do the same with half of the pairs that have not been projected. Then $\mathcal{S}_{3,c}$ does the same on half of the rest, and so on.

Received 11 March 2008; accepted 18 June 2009; published online 20 July 2009

References

1. Aliferis, P., Gottesman, D. & Preskill, J. Quantum accuracy threshold for concatenated distance-3 codes. *Quant. Inf. Comput.* **6,** 97–165 (2006).
2. Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge Univ. Press, 2000).
3. Feynman, R. P. Simulating physics with computers. *Int. J. Theor. Phys.* **21,** 467–488 (1982).
4. Fannes, M., Nachtergaele, B. & Werner, R. F. Finitely correlated states on quantum spin chains. *Commun. Math. Phys.* **144,** 443–490 (1992).
5. Verstraete, F. & Cirac, J. I. Renormalization algorithms for quantum-many body systems in two and higher dimensions. Preprint at <http://arxiv.org/abs/cond-mat/0407066> (2004).
6. DiVincenzo, D. P. The physical implementation of quantum computation. *Fortschr. Phys.* **48,** 771–783 (2000).
7. Poyatos, J. F., Cirac, J. I. & Zoller, P. Quantum Reservoir Engineering with laser cooled trapped ions. *Phys. Rev. Lett.* **77,** 4728–4731 (1996).
8. Perez-Garcia, D., Verstraete, F., Wolf, M. M. & Cirac, J. I. Matrix product state representations. *Quant. Inf. Comput.* **7,** 401–430 (2007).
9. Verstraete, F., Murg, V. & Cirac, J. I. Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems. *Adv. Phys.* **57,** 143–224 (2008).
10. Briegel, H. J. & Raussendorf, R. Persistent entanglement in arrays of interacting qubits. *Phys. Rev. Lett.* **86,** 910–913 (2001).
11. Kitaev, A. Y. Fault-tolerant quantum computation by anyons. *Ann. Phys.* **303,** 2–30 (2003).
12. Levin, M. A. & Wen, X. G. String-net condensation: A physical mechanism for topological phases. *Phys. Rev. B* **71,** 045110 (2005).
13. Cirac, J. I. & Zoller, P. New frontiers in quantum information with atoms and ions. *Phys. Today* **57,** 38–44 (2004).
14. Bloch, I., Dalibard, J. & Zwerger, W. Many-body physics with ultracold gases. *Rev. Mod. Phys.* **80,** 885–964 (2008).
15. Leibfried, D., Blatt, R., Monroe, C. & Wineland, D. Quantum dynamics of single trapped ions. *Rev. Mod. Phys.* **75,** 281–324 (2003).
16. Lindblad, G. On the generators of quantum dynamical semigroups. *Commun. Math. Phys.* **48,** 119–130 (1976).
17. Kempe, J., Kitaev, A. Y. & Regev, O. The complexity of the local Hamiltonian problem. *SIAM J. Comput.* **35,** 1070–1097 (2004).
18. Aharonov, D. *et al.* Adiabatic quantum computation is equivalent to standard quantum computation. *SIAM J. Comput.* **37,** 166–194 (2007).
19. Kempe, J. Quantum random walks—an introductory overview. *Contemp. Phys.* **44,** 307–327 (2003).
20. Arora, S. & Safra, S. Probabilistic checking of proofs: A new characterization of NP. *J. ACM* **45,** 70–122 (1998).
21. Gottesman, D. A theory of fault-tolerant quantum computation. *Phys. Rev. A* **57,** 127–137 (1998).
22. Wolf, M. M. & Cirac, J. I. Dividing quantum channels. *Commun. Math. Phys.* **279,** 147–168 (2008).
23. Wolf, M. M., Ortiz, G., Verstraete, F. & Cirac, J. I. Quantum phase transitions in matrix product systems. *Phys. Rev. Lett.* **97,** 110403 (2006).
24. Verstraete, F., Wolf, M. M., Perez-Garcia, D. & Cirac, J. I. Criticality, the area law, and the computational power of PEPS. *Phys. Rev. Lett.* **96,** 220601 (2006).
25. Verstraete, F., Cirac, J. I., Latorre, J. I., Rico, E. & Wolf, M. M. Renormalization-group transformations on quantum states. *Phys. Rev. Lett.* **94,** 140601 (2005).
26. Diehl, S. *et al.* Quantum states and phases in driven open quantum systems with cold atoms. *Nature Phys.* **4,** 878–883 (2008).
27. Kraus, B. *et al.* Preparation of entangled states by quantum Markov processes. *Phys. Rev. A* **78,** 042307 (2008).
28. Cohen-Tannoudji, C., Dupont-Roc, J. & Grynberg, G. *Atom-Photon Interactions* (Wiley, 1992).
29. Perez-Garcia, D., Verstraete, F., Cirac, J. I. & Wolf, M. M. PEPS as unique ground states of local Hamiltonians. *Quant. Inf. Comput.* **8,** 0650–0663 (2008).

Acknowledgements

We thank D. Perez-Garcia for discussions and acknowledge financial support by the EU projects QUEVADIS, SCALA, the FWF, QUANTOP, FNU, SFB FoQuS, the DFG, Forschungsgruppe 635, the Munich Center for Advanced Photonics (MAP) and Caixa Manresa.

Author contributions

All authors have contributed equally to this paper.

Additional information

Supplementary information accompanies this paper on www.nature.com/naturephysics. Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions. Correspondence and requests for materials should be addressed to F.V. or J.I.C.

Charles Bennett

Statement

and

Readings

Entanglement, Decoherence, and the fate of Jimmy Hoffa

Charles Bennett

Quantum information theory provides a coherent picture of the emergence and obliteration of correlations, even in macroscopic systems exhibiting few traditional quantum hallmarks. It helps explain why the future is more uncertain than the past, and how decoherence causes information to become classical by becoming redundantly replicated throughout a system's environment, a process Zurek nicknamed "quantum Darwinism". The most private information, exemplified by which path a particle takes through an interferometer, evades this replication, and so is evanescent in the sense that after the experiment is over even God does not remember what "happened". Less private kinds of information include classical secrets, facts known only to a few, or information—like the lost literature of antiquity—that once was public but has been forgotten over time. Finally there is information that has been replicated and propagated so widely as to be infeasible to conceal and unlikely to be forgotten. Modern information technology has caused a proliferation of such information, eroding personal privacy while at the same time deterring crime and tyranny. At a fundamental level, one might hope that whenever information is amplified to the point of becoming macroscopic and classical, it becomes permanent and ineradicable. However, by comparing entropy flows into and out of the Earth with estimates of the planet's storage capacity, we conclude that most macroscopic information about the past—for example the pattern of drops in last week's rainfall or rice grains in last night's dinner—is impermanent, soon becoming nearly as ambiguous, from a terrestrial perspective, as the which-path information of an interferometer. Depending on the diligence and forgetfulness of their enemies, the fate of mysteriously disappeared persons such as US labor leader Jimmy Hoffa, thought to have been murdered in 1977, may by now have acquired this ambiguous epistemological status. Finally we discuss prerequisites for a system to accumulate and maintain in its present state, as our world does, a complex and redundant record of at least some features of its past. Not all Hamiltonians and initial conditions lead to this behavior, and in those that do, the behavior itself tends to be temporary, with the system losing its memory as it relaxes to thermal equilibrium.

Hans J. Briegel

Statement

and

Readings

Entanglement and coherence in biological systems

Hans J. Briegel

We discuss the possibility of the existence of entanglement in biological systems. Our arguments center on the fact that biological systems are thermodynamic open driven systems far from equilibrium. In such systems error correction can occur which may maintain entanglement despite high levels of decoherence. We also discuss the possibility of cooling (classical or quantum) at the molecular level.

Joint work with Sandu Popescu and Jianming Cai.

Entanglement and intra-molecular cooling in biological systems? – A quantum thermodynamic perspective

Hans J. Briegel[1,2] and Sandu Popescu[3,4]

[1] *Institute for Theoretical Physics, University of Innsbruck, Technikerstraße 25, Innsbruck, Austria*
[2] *Institute for Quantum Optics and Quantum Information of the Austrian Academy of Sciences, Technikerstraße 21a, A-6020 Innsbruck, Austria*
[3] *H.H. Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, U.K.*
[4] *Hewlett-Packard Laboratories, Stoke Gifford, Bristol BS12 6QZ, U.K.*

Abstract

We discuss the possibility of existence of entanglement in biological systems. Our arguments centre on the fact that biological systems are thermodynamic open driven systems far from equilibrium. In such systems error correction can occur which may maintain entanglement despite high levels of de-coherence. We also discuss the possibility of cooling (classical or quantum) at molecular level.

1 Introduction

The notion of entanglement plays a central role in quantum physics. It is a key concept in quantum information theory, with applications in quantum computation and communication [1, 2], and considerable effort has been devoted to identifying physical systems in which entanglement can be created and exploited for information processing purposes [3]. From a broader perspective, entanglement and the concepts of quantum physics play a fundamental role for our understanding of Nature: The occurrence of entanglement in any system indicates that we are on a terrain of Nature where classical concepts are likely to be quite insufficient for its proper understanding.

Experience shows that entanglement - other than the one that is there because it is a property of the ground state of a system and protected by brute force via an energy gap - is extremely fragile and easily destroyed by

noise. Laboratories work very hard to produce such entanglement, e.g. by isolating single ions in traps and by cooling them close to their motional ground state or, more generally, by complete control and manipulation of matter on the atomic level. Hence most researchers believe that entanglement will be hard to find in natural (i.e. uncontrolled) systems outside the laboratories, not to mention biological systems, which operate at room temperature and which are complex, noisy and "wet" (for a recent review see e.g. [4]). Here however, we would like to make the case that contrary to the standard view, entanglement might actually exist and play a role also in biological systems; one should be aware of this and look for signatures of entanglement in such systems.

In fact, possible evidence for quantum coherence (though not entanglement) in photosynthetic systems has been recently reported in [5, 6] and studied theoretically in a number of papers [7, 8, 9, 10].

We want to emphasize from the outset that this is not a research paper in the sense that we do not have any concrete results to prove that persistent, controlled entanglement exists in biological systems. We will however try and present circumstantial evidence for it.

Our main point is that the intuitions about the fragility of entanglement that are used for dismissing the possibility of entanglement in biological systems are misleading because they generally ignore a fundamental fact, namely that biological systems are open driven systems. This opens many possibilities that, as far as we know, have not yet been carefully considered.

Since the issue is so important, we feel that it would be a great mistake to dismiss the possibility of biological entanglement without a much more careful investigation. The scope of our paper is to call for such a vigorous program of research.

Related to the problem of biological entanglement is possibility of intra-molecular cooling, that is, the possibility that parts of a molecule are actively cooled relative to the environmental temperature. This process is in fact much more general than the generation of entanglement. In fact intra-molecular cooling might occur also in instances when there is no quantum coherence whatsoever. As such, the probability that intra-molecular cooling actually takes place in biological systems is far larger than the probability of existence of controlled quantum coherent phenomena and entanglement that form the main subject of the present paper.

2 Live or dead entanglement?

To start with, we should discuss more carefully what kind of entanglement are we looking for.

That entanglement, and, more generally, coherent quantum effects always exist - at some level - in all systems (including biological ones) is quite clear. After all the laws of quantum physics enter on the level of quantum chemistry, determining the structure and energy spectra of the molecules and their interactions. Coming to biology, most scientists would probably share the view that quantum mechanical effects play a role, but only an indirect one: Quantum mechanics would thus be responsible for the molecular basis or substrate, while the biological functionality of the molecules can be explained by classical statistical physics, combined with the principles of molecular Darwinism [11].

This makes it clear that, even before asking whether entanglement exists or not, we need to better define what we are actually talking about, that is: what kind of entanglement? In the following, we would like to distinguish three different kinds of entanglement that we expect to play a role in biology; the same classification holds also for any quantum coherent processes that may occur in biological systems.

One has to say from the beginning that the boundaries between theses different types of entanglement are fuzzy. The classification of some phenomena is clear-cut, while for others one may argue whether it is of type 1 or type 2, etc.; nevertheless, we believe this classification to be essential when proceeding to study the possibility of biological quantum effects.

The three types of entanglement are:

- entanglement of basic constituents
- dead entanglement
- live entanglement.

Entanglement of basic constituents. Think of any atom or molecule; these elementary systems contain a lot of entanglement. Indeed, all their electrons are entangled, the protons and the neutrons in the nucleus are entangled, and so on. The existence of this sort of entanglement is obvious and in most cases trivial.

As always the boundary cases may be more interesting. For example, while the existence of a delocalized electron state in a benzene molecule is

also quite trivial, finding delocalized states that extend over much larger molecules may have important consequences.

A similar situation occurs in condensed matter systems, in which the ground state is typically highly entangled, too. The role of this entanglement e.g. in quantum magnets or super-conducting materials has attracted a lot of attention recently [12]. The phenomenon of entanglement is thus not restricted to the microscopic domain of very few particles, but it occurs also in macroscopic bodies and even at finite temperatures.

The entanglement in these systems (which could be called "static entanglement") has many interesting facets. In any case, this is not the type of entanglement we are concerned with in the present paper.

Dead entanglement. Dead entanglement occurs in molecules that have biological origin or occur in biological cells. However the occurrence of this kind of entanglement does not require metabolic processes to function. As such, these molecules can be taken out of the cell, and they will continue to work.

Here we are talking about systems that are generally in thermal equilibrium. When some appropriate external perturbation comes, it takes them out of equilibrium, some coherent phenomenon takes place and it quickly dies out. A paradigmatic example of this, outside the biological context, would be a piece of an optical fiber. When a photon comes it propagates through the optical fiber and gets out at the other end. However, during the rest of the time, the piece of optical fiber just stays there, say on an optical table, and nothing happens to it. A similar example, taken from biology, would be the more complicated molecule of chlorophyll, which absorbs energy (in form of a photon) that then propagates (in form of an exciton) from one centre/part of the molecule to another part, until it reaches the reaction center. Again, in the absence of the photon, the molecule is at equilibrium and nothing interesting happens to it.

There are some key words and properties that we generally expect to be associated with this type of entanglement: *Incidental*, *Side effect*, *Short time*, *May have biological functionality*, and *May be evolutionary selected*.

Incidental: The main characteristic of the process may not require entanglement, or coherence, but in a particular implementation of the process they may just occur. For example, when a photon propagates through an optical fiber, maintaining polarisation coherence is not necessary. It so happens however that present day optical fibers are so good that polarisation coherence is maintained. But as far as functionality is concerned, which is transmitting the light, this is not important.

Side effect: Entanglement may be just some sort of side effect of the process. That is, it may always accompany a given process, but not play any role.

Short time: Generally these phenomena are short time because this external perturbation produces some modification but then the environment immediately brings it down to equilibrium. These are phenomena that may typically take pico- or femto-seconds. As a matter of fact, if you work on a very short time scale, there is always some quantum coherence, because it requires some time to die out, to de-cohere.

On the other hand, one must also be aware that although the absolute time scales involved in these phenomena are very short, this doesn't necessarily mean that entanglement/coherence does not play a significant role. Indeed, a relevant time scale is that of the duration of the process itself. If entanglement/coherence is present during the whole process, then it may plays a significant role; otherwise its role is most probably irrelevant.

Biological functionality and evolution: In some instances, the entanglement *may* have some biological functionality, and it *may* have been that this type of quantum coherence was evolutionary selected. It is also possible that the although the entanglement is just a side effect without biological functionality, the primary effect that leads to it was biologically selected. That is, other, important, things were selected and evolved, and with them the entanglement, but just as an accompanying effect. As an example, we can give here photosynthesis again.

The key word here is "may". That is, in the case of this type of entanglement/quantum coherence, while evolution may occur, it is not a sine-qua-non condition for its very existence.

By no means are "dead" entanglement/quantum coherence non-interesting phenomena. On the opposite. They are extremely interesting and very complicated: Although conceptually they are the same, there is an enormous difference between say, propagation of a photon through an optical fiber and propagation of an exciton through chlorophyll. Showing that even such "simple" quantum effects actually take place in biological systems is a great challenge. In fact at present all the experimental work on biological quantum effects is focussed exactly on this type of processes.

Live entanglement. The defining property of this type of biological entanglement is that it exists only while metabolic processes take place. In other words, it exists only as long as the system is actively maintained far from thermal equilibrium, i.e. in **open, driven systems far from equilibrium**. When the metabolism stops and the system reaches equilibrium,

this type of entanglement disappears.

The key properties that we expect this class of phenomena to posses are: *It is persistent*, *It is dynamically controllable*, *It has biological functionality*, *It is evolutionary selected*.

Persistency: By their very nature, these processes are such that as long as they are active, entanglement/coherence is maintained. This is in fact the purpose of the entire process. From this point of view they are fundamentally different from dead-entanglement phenomena, in which the entanglement appears only as a transitory phenomenon.

Dynamical control: In the case of dead entanglement, such as, if confirmed, propagation of an exciton through a chlorophyll molecule, the details of the process are governed by the structure of the molecule itself. This is a structural and therefore relatively static parameter. On the other hand, in processes that are driven far from equilibrium, changes in the way in which the process is driven can immediately alter the characteristics of the process, and hence of the entanglement.

Biological functionality and evolutionary selection: These are, of course, much more complicated processes than the ones that give rise to dead entanglement. Unless the live entanglement has biological functionality, evolution most probably could not have arrived at it.

Again, all the above are properties that we expect this class to posses. This is not to say, of course, that there could be no exceptions. For example it is not impossible that some process leads to persistent entanglement as a side effect with no biological functionality. All we say is that this seems to us highly improbable.

To conclude this section, we want to emphasize once more that we do not have any (experimental) evidence, at present, that this type of entanglement (or similarly: quantum coherence) exists. The very scope of this paper is to investigate the possibility that such entanglement actually exists.

We would also like to mention that we are by no means suggesting the possibility of entanglement at very large scale - such as super-positions of brain states leading possibly to quantum computation in the brain, etc.. This seems to us virtually impossible and here we fully agree with the sceptical view expressed in Ref. [4] (see also [13]). What we are interested in is persistent and controllable entanglement with presumably biological function, at the level of bio-chemical processes.

3 Why should we look for entanglement in biological systems?

At first sight, biological systems, which are warm and wet, seem very unlikely places to look for entanglement. It is our claim however that if entanglement (of the type we described above) is to be found anywhere in Nature outside physics labs, it is inside living beings.

To stabilize entanglement is generally difficult and requires complex setups - obtaining and maintaining entanglement in a laboratory is not an easy task. Of course, some instances of naturally occurring entanglement could exist in nature, such as a piece of nonlinear crystal that in sunlight may produce entangled photons by down-conversion. But such occurrences are purely accidental and probably very rare. On the other hand, biology itself could be a driving force: if entanglement turns out to be useful, biological evolution could select for this. In fact it is quite probable that entanglement offers advantages. The point is that the number of possible entangled states is so much larger than that of non-entangled ones that it is, in fact, inconceivable not to find entangled states that will offer advantages. Indeed, the possibilities offered by entanglement are much richer than those offered by non-entangled mixtures: If the efficiency of a certain biological process depends on the state that occurs in the system, the state space over which evolution can optimize that process is much larger if one allows entangled states. This is essentially a probability argument. To conclude, from this perspective, if entanglement is to be found anywhere in Nature, then in biological systems. Incidentally, the only place where we have non-trivial entanglement, i.e. in some laboratories, it is of ultimate biological origin! It is only at the end of a long evolution that the required complexity for producing controlled entanglement has been achieved.

On the other hand, there is a caveat: Evolution takes a long time and it is quite possible that even though entanglement is useful, nature has not succeeded yet in exploiting it.

4 Open driven systems and entanglement

Living organisms depend on permanent consumption of energy in the form of food (or photons, as in plants). Different from e.g. some solid-state material, a living cell cannot be described as an isolated system. It continuously exchanges particles with its environment, and with them energy and entropy. In the language of thermodynamics, biological systems are open

driven (quantum) systems, whose steady state is far away from thermodynamic equilibrium. As we mentioned in the introduction, this fact has major implications for the issue of the presence of entanglement.

To begin with, the fact that in open driven systems entanglement can exist at room temperature is absolutely clear, once one realizes that every quantum physics laboratory is such a system. So, having established that, as a matter of principle, controllable entanglement can exist in room temperature systems, the question is only one of scale and complexity. Do we need sophisticated lasers and large fridges (which work at room temperature but cool a subpart of them) or can they be present in the small scale of a living cell? After all, there are numerous studies that show that analogues of large scale, man-made machines (ratchets, rotors, etc.) do actually exist on a bio-molecular level (for a review see e.g. [14]). In what follows, we will give a number of specific examples that seem suitable to be scaled down.

Before going to specific examples however, it is important to understand why open driven systems make a difference. The reason is that such systems can perform error correction. Decoherence introduces noise into systems and increases their entropy. On the other hand, open driven systems have, by definition, access to a source of free energy and can use it to get rid of the errors. In fact for the issue we consider, namely merely producing and stabilizing a particular entangled state, one doesn't even need full quantum error correction, i.e. an error correction protocol that can stabilize any arbitrary superposition of states (in a given Hilbert subspace) [15, 16]. The error correction we require here is, computationally speaking, trivial - the stabilization of a single state. Nevertheless, entropically the task is similar, i.e. the entropy continuously produced by noise has to be removed from the system.

Coming back to biology, we note that probably the most striking characteristic of biological systems is that they are error correcting systems - a dead animal starts decomposing in a matter of hours so it must be that while living there are continuous error correction processes going on which maintain the body. So once we realize that error correction takes place in any living organism, whether or not it is enough to stabilize entangled states becomes a matter of scale not one of principle.

4.1 Toy models I

As a first example, consider the famous procedure used to produce entanglement in present day optical labs, namely parametric down conversion. A typical parametric down-conversion experiment uses a laser to produce a stream

of photons that are directed towards a nonlinear crystal. Upon impinging onto the crystal, some of these photons generate pairs of polarization-entangled photons.

The entanglement in the experiment described above is totally dependent on the fact that we are dealing with an open driven system: The laser is powered by a power source; when the source is turned off the entanglement disappears.

The parametric down-conversion experiment discussed above requires complex equipment including a laser and a power source. Is such complicated equipment necessary? The answer is no - there is no fundamental principle of nature that requires complex equipment. The following simple device (see Fig. 1) could do the same thing.

Suppose we have a thermo-electric element - two simple wires, each made of a different metal - that is connected to a simple light-emitting-diode LED. The LED itself is a very simple device as well - two semiconducting crystals, each with a different impurity, joined together (a so called "n-p junction"). If one of the joints of the thermo-electrical element is heated relative to the other, e.g. by a simple flame, the LED will produce a light-beam. When the output of the LED is directed to a non-linear crystal entangled photons are produced by parametric down-conversion.

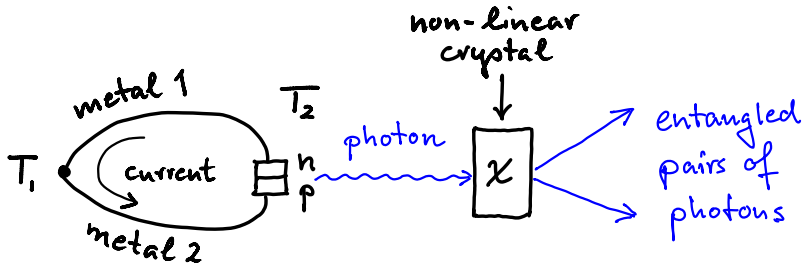


Figure 1: Simple entanglement-generating device (see text).

Note that there is something quite remarkable about this simple device. (a) It is quite robust and has no moving parts. (b) The entire system is entirely driven by a temperature difference, thus, the source of free energy is not coherent, and neither is the transmission of current in the wires a coherent process (see also [17]). Furthermore, (c), the junction and the crystal are both at room temperature. This simple device is able to create something highly quantum mechanical, such as entanglement! How can that

be? Of course, it has to do with the (coherent) interaction that goes on inside the crystal, when a photon is absorbed and converted into a pair of photons with lower energy. The main point to be observed is that this process is embedded into a hot and rather noisy environment, and it is only due to the continuous pumping of the system that the entanglement generating interaction in the crystal can be exploited.

We may further imagine that the photons in the entangled pair are further directed to two atoms that absorb them. Upon absorbing the photons, the atoms become entangled. If the atoms are subjected to noise, the entanglement could be subsequently destroyed; whether any entanglement can survive for a longer time is a question of how quick is the decoherence versus the rate of pumping new entanglement into the atoms by subsequent entangled photon pairs.

A molecular analogue (see Fig. 2) would be a chemical process where a large molecule (such as a protein) serves as a catalyzer for an exothermic chemical reaction. We assume that the chemical reaction takes place at a certain functional centre ("docking site") of the catalyzer molecule. The chemical process may lead to a transfer of free energy along the molecule; the free energy transfer may be conveyed by various channels, for example by excitons, phonons, or electric current (electron displacement). This energy transport over the molecule need not be coherent and the molecule may be at room temperature. What matters is that, at some other site of the catalyzer molecule, the energy may be converted by some nonlinear process, similar to the one that occurs in the non-linear crystal of the parametric down-conversion, into two entangled modes that pump entanglement into, say, some receiver atoms.

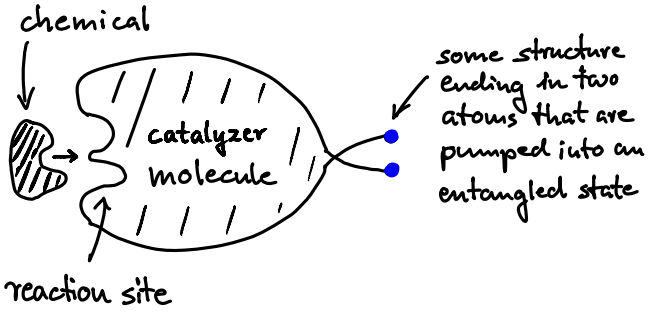


Figure 2: Hypothetical molecular device that is capable of creating entanglement by a molecular pumping process.

The important point is here that the two atoms, depicted in blue in Fig. 2, will not be entangled if we stop driving the system, i.e. if we stop the supply of the reactant chemical. This emphasizes the difference between passive (static) entanglement, as between electrons in an atom, and the dynamic equilibrium entanglement.

4.2 Toy models II. Conformational changes and time-dependent Hamiltonians.

In the previous section we emphasized that set-ups that produce entanglement need not be very complex, As a matter of fact however biological processes are actually very complex and sophisticated. A very important process that occurs in many instances is that of controlled conformational changes in proteins [18]. These processes could lead to controlled time-dependent Hamiltonians, with obvious implications for possible coherent quantum processes and existence of entanglement.

A conceivable process is the following.

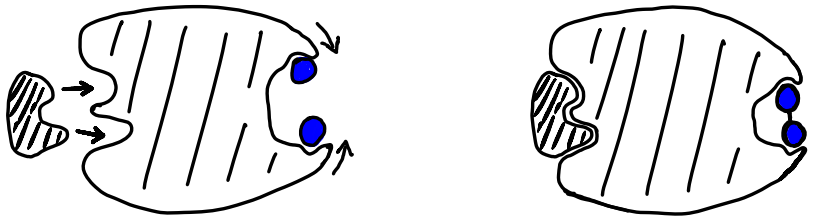


Figure 3: Entanglement of two atoms (blue) in a molecule, induced by a stream of reactant chemicals which dock to the catalyzing molecule, leading to a conformation change (see text).

A chemical process that occurs at one binding site of a protein can lead to a configuration change of another site of the same protein [28]. This conformation change may then lead to an interaction e.g. between two atoms that would otherwise be in a shielded site of the molecule (see Fig. 3). The two atoms depicted in blue in Fig. 3 are not interacting in the "rest" state of the protein; they are at thermal equilibrium and non-entangled. When the conformation change occurs they are brought together and start interacting. The interaction takes the atoms out of their equilibrium state and may entangle them. The entanglement survives until de-coherence kills it and the atoms reach a new thermal equilibrium state. Then protein

reverts to the rest conformation and the atoms revert to the rest mode and then the process starts again. The point here is that although both thermal equilibrium states may be non-entangled, the state of the atoms during the transition time (which may be relatively long) may well be entangled. Again, to complete the cycle the protein needs to be supplied with free energy; if we stop driving the system, i.e. if we stop the supply of reactant chemicals then the atoms simply reach thermal equilibrium and remain there and entanglement is lost. As in the previous section, this emphasizes the difference between passive (static) entanglement, as between electrons in an atom, and dynamic entanglement that exists due to the (non-equilibrium) process of conformational changes.

In the above example we described a very simple cycle. It is conceivable however that more complicated cycles could exist, with many different conformational changes occurring sequentially. Such a process may then implement a more intricate time-dependent Hamiltonian corresponding to what in quantum information is called a "sequence of gates".

4.3 Toy models III. Intra-molecular cooling.

Thinking of different ways of obtaining controllable entanglement in the laboratory, with an eye to possible biological implementation has led us to the idea of intra-molecular cooling. This process is far more general than the generation of entanglement and needs not be associated with entanglement at all. In fact intra-molecular cooling might occur also in instances when there is no quantum coherence whatsoever. As such, the probability that intra-molecular cooling actually takes place in biological systems is far larger than the probability of existence of controlled quantum coherent phenomena and entanglement that form the main subject of the present paper.

Cooling, if it could actually be achieved at molecular level, would have obvious benefits. One example is when there are many possible reaction channels and, when the site is cooled, a preferred channel is selected. Another example is increased efficiency of catalysis: Many proteins act as catalyzers with very high specificity. They have active sites in the shape of cavities, which bind only molecules have that fit precisely into the cavity like a hand in a glove. At high temperature the protein vibrates stronger and the shape of the cavity is deformed which leads to a decline in the efficiency of catalysis. Obviously, if the active site could be cooled this would be a benefit.

Regarding the potential role of cooling in biological systems, it should be realized that many organisms have indeed the ability of cooling parts of

their body. For example our body temperature is kept largely stable at a temperature around 37 ℃, even when we live in hot climate. This is only possible since our body has a built-in cooling system. The only question is therefore not whether or not cooling exists in biological organisms but at which scale. Does it exist only at large scale - at the scale of the whole organism, or at scale of internal organs - or is it present all the way down to molecular level? Given the potential advantages for intra-molecular cooling, it is not inconceivable that such mechanisms were evolutionary selected. In fact, it is even possible that cooling at molecular level is much stronger than at the scale of the whole body. The reason for this is two fold. First in smaller systems, the coupling of parts of the system with the environment may be much weaker than of other parts so some parts, the "inside of the fridge" could be rather well isolated from the environment. Second, there may be a lot of free energy available (from chemical reactions), which will be channeled preferentially within the molecule.

Set-ups that allow cooling can be surprisingly simple. An explicit example is the so called "algorithmic cooling" [19, 20, 21] which could be realized even on a molecular level. In algorithmic cooling, we have in mind that some atoms at a given site, presumably an active but protected site, might be cooled. A very simple example of algorithmic cooling [20] involves only three qubits (i.e. two-level systems). In their original example Brassard *et al.* [20] considered the qubits to be nuclear magnetic moments, but any other physical system (such as vibrational levels of atoms) could be used. In this example one qubit can be cooled while the two other qubits get warmer - very much like the back spiral of an ordinary refrigerator. From these two warmer qubits heat is dissipated into the environment and the process continues, achieving persistent cooling of the first qubit.

In the algorithmic cooling process, the interactions between the qubits are time dependent, driven by an external mechanism. Essentially, this can be realized by a (short) sequence of pair wise couplings between the qubits. In practice, this could be realized by, say, changing the relative position of atoms according to a predetermined sequence, something that could be realized by a sequence of conformational changes in a protein. Consider for example a protein molecule with two active sites, similar to the one in Fig. 3. The active site on the left is where we supply free-energy (by some catalytic reaction) which is required for driving the cooling process. The active site on the right (where the blue atoms in Fig. 3 are) is the "main active site" that we want to cool. The chemical reaction at the active site on the left drives the cooling process at the active site on the right.

Of course, it is impossible to measure directly the temperature of an

active site of a protein. However one possible signature for the presence of a cooling process would be the following: As mentioned above the typical rate of reaction for biological catalytic processes is the following. At low temperature the reaction is slow because the reactants move slower. Then the rate increases with the temperature. However, after a certain temperature the reaction rate starts decreasing because the reaction site becomes deformed due to vibrations and the reactants don't fit in it anymore.

Consider now the activity of the site subjected to cooling, described in Fig. 4. The temperature axis of the graph represents the temperature of the environment of the molecule (the temperature of the cell, or of the liquid that surrounds the cell. The two curves describe the rate of this reaction in two cases - when the protein is supplied with free-energy and when not. Recall that we supply free-energy in the form of chemical reactants that react on the active site at left. When there are no reactants to drive the cooling process, the main active site is at the same temperature as the environment. On the other hand, when there is a supply of reactants the main active site is cooled relative to the environment and the rate of the reaction it catalyzes remains high at higher environmental temperatures. Note however that there should be no difference between the two cases at low temperatures because then the rate is simply limited by the slow movement of the molecules of interest, i.e. they enter less frequently in the main active site.

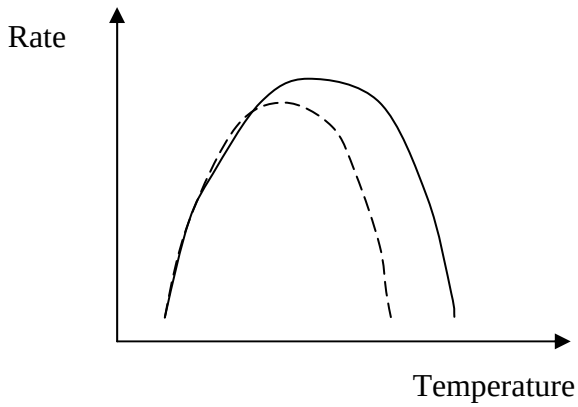


Figure 4: Potential control of enzyme activity through intra-molecular cooling (see text).

Of course, there could be many reasons why the rate of reaction at the

main active site increases when a reaction occurs at the other site of the molecule - the molecule's configuration could simply be optimized for the case when both reactions occur simultaneously. However, the very specific temperature dependence described above could be a good indicator.

4.4 Toy models IV. Reset mechanisms.

Our last example is not derived from a macroscopic device, but from the study of another instance of an open driven quantum system away from thermal equilibrium. This is a gas-type system - specifically a spin gas - that was studied in [22]. This particular system is relatively simple and it is accessible to quantitative analysis. The reason why it is instructive to study this example has to do with the role of de-coherence in such a system, which couples to the individual gas particles and thus quickly destroys any transient entanglement. It is thus clear that no static entanglement can occur in such a system. Let us consider a simple gas cell, where gas particles move on classical trajectories, but have some internal structure which is described by quantum mechanics (see Fig. 5). The interaction between these particles is capable of entangling their internal degree of freedom. As a concrete realization, we may e.g. imagine ultra-cold atoms with two internal hyperfine states $|0\rangle$ and $|1\rangle$, or molecules carrying a nuclear spin with two possible states. For the sake of the argument, suppose that we have full knowledge about the entire interaction history of the gas particles. As was found in a series of studies, the quantum states generated under such a simple dynamics, can already display a number of interesting and highly non-trivial entanglement properties. These include states with a high persistency of entanglement and states that are universal resources for quantum computation. This illustrates that even a system as simple as this (toy-model of a) gas can exhibit highly non-trivial and complex quantum effects.

Consider now that, as in any real scenario, these states will be exposed to de-coherence, which will quickly destroy the entanglement. A possible mechanism for de-coherence could be collisions of the gas particles with the environment consisting e.g. of other species of particles - a "buffer gas" - within the gas cell (see Fig. 5). The effect of such additional interactions, even if they are weak, will be devastating and in general no entanglement will survive in steady state.

However, as was shown in [23], there is a simple mechanism that can in principle sustain the entanglement in the gas in the presence of de-coherence, without introducing entanglement by itself! Imagine that, when particles enter a certain region in the cell, indicated by a red spot in Fig. 6, they are

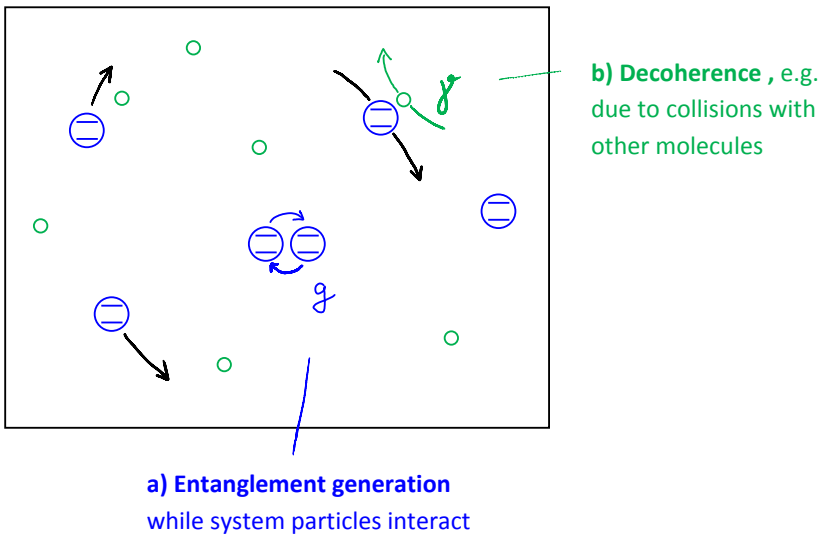


Figure 5: a) Illustration of a spin gas (blue): Particles move on classical trajectories, but each particle carries an internal quantum mechanical degree of freedom, such as a spin. Upon collision, the spin degrees of freedom get entangled. b) Collisions of the gas particles with other particles (buffer gas, green) lead to de-coherence.

reset in some pure state $|\chi\rangle = \alpha|0\rangle + \beta|1\rangle$. This could be realized by various mechanisms, for example by an interaction with some local structure in the cell.

Such a reset mechanism will have two effects. It destroys existing entanglement between this particle (i.e. its spin) and the other particles in the gas, but it also destroys this particle's entanglement with the environment! On first sight, this does not appear to be very constructive, but it has the effect that the particles that leave the light spot are in a pure state, and, if two of such particles collide later on (e.g. in the vicinity of the spot), they are capable of creating fresh entanglement! A detailed analysis shows indeed that, for a specific choice of parameters this (toy-model of a) gas can have a steady state, where entanglement persists in the cell. Depending on the mean free path of the particles, the regions of persisting entanglement may be confined to the vicinity of the reset region or they may extend over the entire gas cell.

While the details of this process are not of interest here, it should be men-

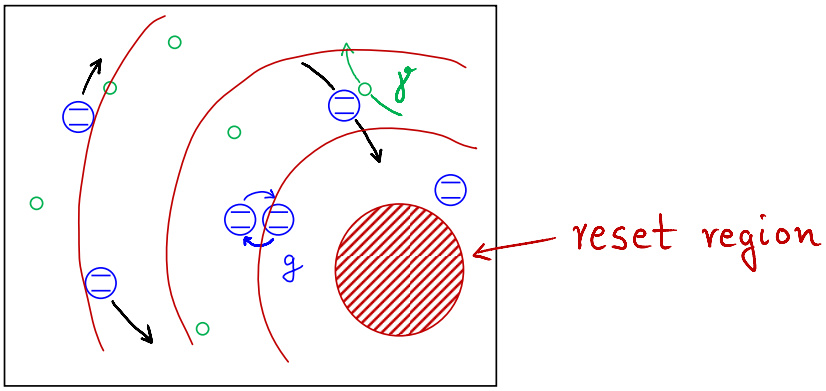


Figure 6: Particles that pass through the reset region indicated by a red spot will be reset in some standard internal state of low entropy. Particles that cross the spot will thus leave the spot in a pure state $|\chi\rangle = \alpha|0\rangle + \beta|1\rangle$. As a consequence, entanglement will build up around the spot and persist – in dynamic equilibrium – despite of de-coherence.

tioned that there are two regimes, one of vanishing and one of non-vanishing entanglement, and there is a sharp transition between them. Entanglement can be sustained if both the coherent interaction strength and the reset rate are sufficiently large. Furthermore, the particles need not be reset to a pure state; it suffices if they leave the region in a mixed state of sufficiently low entropy/high purity.

The main purpose of the preceding discussion of the spin gas was to demonstrate that the possibility of entanglement is not confined to highly controlled systems at very low temperatures: Simple reset processes allow entanglement to persist also in a hot and noisy environment! Even though this is a simple model, it is conceivable that similar processes could play a role also in biological systems, where various analogues of reset processes exist.

5 Conclusions

Biological systems are of extraordinary complexity and diversity. As such, at the moment we don't know where to start searching for entanglement and/or molecular cooling, be it an experimental search or a theoretical one. Furthermore, the specific toy models presented here are almost surely with

very little direct relevance in biology. However, our goal here is far more limited - it is to argue that the presence of controlled entanglement with biological functionality cannot be discounted automatically, without a careful study. Indeed, although our specific toy models may well have very little direct relevance, we are confident that the processes we described (entanglement pumping, resetting, etc.) are to be found in a way or another; the same applies to the idea of molecular cooling. Ultimately, the power of biological evolution coupled with the fact that biological organisms are open, driven systems, may open the door for many unexpected quantum phenomena. Similarly, they also open the door to highly non-trivial thermodynamic phenomena.

Note added: Since the first version of this paper was put on ArXiv, a number of preprints have appeared which investigate the role of quantum entanglement in specific biological scenarios, including e.g. photosynthesis [24] and the chemical compass mechanism for magnetoreception [25, 26]. The possibility of dynamic entanglement generated by conformational changes of molecules, as described in Sec. 4.2, has in the meantime been studied quantitatively in [27].

References

[1] C. H. Bennett and D.P. DiVincenzo, Quantum Information and Computation, *Nature* **404**, 247-255 (2000).

[2] A. Steane, Quantum Computing, In *Rep. Prog. Phys.*, **61**, 117-173 (1998).

[3] P. Zoller *et al.*, Quantum information processing and communication, *Eur. Phys. J. D* **36**, 203 (2005).

[4] H. M. Wiseman and J. Eisert, Nontrivial quantum effects in biology: A skeptical physicists' view, Abstract: Invited contribution to "Quantum Aspects of Life", D. Abbott Ed. (World Scientific, Singapore, 2007). Available also as preprint http://xxx.lanl.gov/abs/0705.1232.

[5] G. S. Engel *et al.*, Evidence for wavelike energy transfer through quantum coherence in photosynthetic systems, *Nature* **446**, 782-784 (2007).

[6] H. Lee, Y.-C. Cheng, and G. R. Fleming, Coherence dynamics in photosynthesis: protein protection of excitonic Coherence. *Science* **316**, 1462(2007).

[7] M. Mohseni, P. Rebentrost, S. Lloyd, and A. Aspuru-Guzik, Environment-assisted quantum walks in photosynthetic energy transfer. *J. Chem. Phys* **129**, 174106 (2008).

[8] M. B. Plenio and S. F. Huelga, Dephasing assisted transport: quantum networks and biomolecules. *New J. Phys* **10**, 113019 (2008).

[9] A. Olaya-Castro, C. F. Lee, F. F. Olsen, and N. F. Johnson, Efficiency of energy transfer in light-harvesting system under quantum coherence. *Phys. Rev. B* **78**, 08115 (2008).

[10] P. Rebentrost, M. Mohseni, and A. Aspuru-Guzik, Role of quantum coherence in chromophoric energy transport. *J. Phys. Chem. B* **113**, 9942 (2009). Preprint version arXiv:0806.4627

[11] A. L. Hughes, Adaptive Evolution of Genes and Genomes, Oxford University Press, 1999.

[12] L. Amico, R. Fazio, A. Osterloh, V. Vedral, Entanglement in many-body systems, *Rev. Mod. Phys.* **80**, 517 (2008).

[13] C. Koch and K. Hepp, Quantum mechanics in the brain, *Nature* **440**, 611 (2006).

[14] W. R. Browne and B. L. Feringa, Making molecular machines work, *Nature Nanotechnology* **1**, 25–35 (2006).

[15] A. M. Steane, Error Correcting Codes in Quantum Theory, *Phys. Rev. Lett.* **77**, 793 (1996).

[16] A. R. Calderbank, P. W. Shor, Good quantum error-correcting codes exist, *Phys. Rev. A.* **54**, 1098 (1996).

[17] M. Plenio, S. Huelga, Entangled light from white noise, *Phys. Rev. Lett.* **88**, 197901 (2002).

[18] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, P. Walter, Molecular Biology of the Cell, Garland Science, Taylor & Francis, New York, 5 th Edition, 2008.

[19] J. Baugh, O. Moussa, C. A. Ryan, A. Nayak and R. Laflamme, Experimental implementation of heat-bath algorithmic cooling using solid-state nuclear magnetic resonance, *Nature* **438**, 470–473 (2002).

[20] G. Brassard, Y. Elias, J. M. Fernandez, H. Gilboa, J. A. Jones, T. Mor, Y. Weinstein, L. Xiao, Experimental Heat-Bath Cooling of Spins, Preprint available at http://xxx.lanl.gov/abs/quant-ph/0511156.

[21] L. J. Schulman, U. V. Vazirani, Molecular scale heat engines and scalable quantum computation, in Proc. 31st Annual ACM Symp. on Theory of Computing. (eds Vitter, J. S., Larmore, L. & Leighton, T.) 322–329 (ACM Press, New York, 1999).

[22] J. Calsamiglia, L. Hartmann, W. Dür, H. J. Briegel, Spin Gases. Quantum Entanglement Driven by Classical Kinematics, *Phys. Rev. Lett.* **95**, 180502 (2005).

[23] L. Hartmann, W. Dür, H. J. Briegel, Steady state entanglement in open and noisy quantum systems at high temperature, *Physical Review A* **74**, 052304 (2006).

[24] M. Sarovar M., A. Ishizaki, G. R. Fleming and K. B. Whaley, Quantum entanglement in photosynthetic light harvesting complexes. arXiv: 0905.3787.

[25] J.-M. Cai, J. J. Guerreschi, and H. J. Briegel, Quantum control and entanglement in a chemical compass. arXiv:0906.2383.

[26] E. Rieper, E. Gauger, J. J. L. Morton, S. C. Benjamin, and V. Vedral, Quantum coherence and entanglement in the avian compass. arXiv:0906.3725.

[27] J.-M. Cai, S. Popescu, and H. J. Briegel, Dynamical entanglement in oscillating molecules and potential biological implications. arXiv:0809.4906.

[28] This effect is known as allostery. Allosteric transitions, which are conformational changes of proteins conditioned to the binding of a ligand molecule at a specific site, play a key role in various bio-molecular processes, e.g. in the regulation of enzyme activity, in motor-proteins, and in ion-transport through membranes (see [18], Chapter 3, pp. 171-188). A related effect is the isomeric transformation of retinal in rhodopsin which plays an important role in vision; here the conformational change is not brought about by the adsorption of a ligand molecule, but by the absorption of a photon.

Dirk Bouwmeester

Statement & Readings

Towards quantum superpositions of a mirror

Dirk Bouwmeester
Department of Physics, Center for Spintronics and Quantum Computation,
University of California, Santa Barbara, CA 93106, USA
&
Huygens Laboratory, Leiden University, P.O. Box 9504
2300 RA Leiden, The Netherlands

Abstract

Theoretical and experimental aspects of testing the quantum superposition principle for relatively massive objects will be presented. In collaboration with Prof. R. Penrose a specific experiment for achieving this was proposed [1]. The motivation for Prof. R. Penrose to consider such an experiment stems from his prediction that gravitational effects play a role in the reduction of a quantum superposition to a single "branch" of the superposition. There are two levels of reasoning that led Penrose to his prediction [2]. The highest level is related to the evolution of the universe; starting from the Big Bang the universe appears to evolve into a collection of black holes that evaporate into zero-mass fields. In order to reconcile this observation with the notion that entropy increases, Penrose concludes that there must be a physical process by which state-space reduces (in order to counterbalance the increase of entropy). He then identifies this with quantum measurements in which a state-space reduction (collapse of the wave-function) seems to take place. The second level of reasoning is based on the observation that a quantum wave function $\Psi(x, y, z, t)$ describing a spatial superposition of a massive object is simply inconsistent with the theory of relativity; a spatial superposition of a massive object implies two different space time structures (space-time curvature is determined by the position of the mass) and can therefore not be described by a single set of coordinates *x, y, z, t*. Penrose has estimated the energy associated with a superposition of two space-time structures resulting from a spatial superposition of a massive object, and argues that this amount of excess energy is allowed for a certain amount of time, invoking the Heisenberg uncertainty relation for time and energy. Based on this argument one can estimate that objects observed in quantum superpositions to date (such as atoms, BECs, superconducting currents and C60 molecules, and even the mechanical resonator recently investigated by A, Cleland [Nature 2010]) had much too small mass to observe the reduction of the wave function within the time scale of the performed experiments. It will, according to Prof. R. Penrose, require objects of about 10^{-12} kg (approximately corresponding to the objects considered below) to witness the collapse of the wave function on a time scale of the order of a second. The designed experiment [1] should bring a tiny mirror in a quantum superposition for several microseconds, thus approaching the regime of interest to test Penrose's predictions.

The experiment envisioned in 2003 has received a lot of attention and the challenge to design, fabricate and test such opto-mechanical systems that can display quantum-mechanical behavior is currently a strongly emerging field of science and technology. The reason for this huge interest is partly the fundamental importance, partly the application of optical cooling (an active cooling scheme will be presented [3]), partly the interest in quantum information processing where quantum mechanical resonators can couple to other quantum systems, and partly the potential for metrology applications in weak force and momentum detection. A good overview of state-of-the-art research in this field can be found in [4].

References

[1] W. Marshall, C. Simon, R. Penrose, D. Bouwmeester, *"Towards quantum superpositions of a tiny mirror"*, Phys. Rev. Lett. **91,** 130401-1 (2003).

[2] R. Penrose, *"Quantum Wave-function Collapse as a Real Gravitational Effect"*, Mathematical Physics 2000, Ed by A. Fokas, A. Grigoryan, T. Kibble and B. Zegarlinski (Imperial College, London, 2000**)** *et al.*, p. 266-282.

[3] D. Kleckner and D. Bouwmeester, "*Sub Kelvin optical cooling of a micro-mechanical resonator*", Nature. **444**, 75 (2006).

[4] T.J. Kippenberg and K.J. Vahala, Science **321**, 1172 (2008).

Towards Quantum Superpositions of a Mirror

William Marshall,[1,2] Christoph Simon,[1] Roger Penrose,[3,4] and Dik Bouwmeester[1,2]
[1]*Department of Physics, University of Oxford, Oxford OX1 3PU, United Kingdom*
[2]*Department of Physics, University of California, Santa Barbara, California 93106, USA*
[3]*Center for Gravitational Physics and Geometry, The Pennsylvania State University, University Park, Pennsylvania 16802, USA*
[4]*Department of Mathematics, University of Oxford, Oxford OX1 3LB, United Kingdom*
(Received 30 September 2002; published 23 September 2003; publisher error corrected 25 September 2003)

We propose an experiment for creating quantum superposition states involving of the order of 10^{14} atoms via the interaction of a single photon with a tiny mirror. This mirror, mounted on a high-quality mechanical oscillator, is part of a high-finesse optical cavity which forms one arm of a Michelson interferometer. By observing the interference of the photon only, one can study the creation and decoherence of superpositions involving the mirror. A detailed analysis of the requirements shows that the experiment is within reach using a combination of state-of-the-art technologies.

DOI: 10.1103/PhysRevLett.91.130401 PACS numbers: 03.65.Ta, 03.65.Yz, 42.50.Ct

Introduction.—In 1935 Schrödinger pointed out that according to quantum mechanics even macroscopic systems can be in superposition states [1]. The associated quantum interference effects are expected to be hard to detect due to environment induced decoherence [2]. Nevertheless, there have been proposals on how to create and observe macroscopic superpositions in various systems [3–7], as well as experiments demonstrating superposition states of superconducting devices [8] and large molecules [9]. One long-term motivation for this kind of experiment is the search for unconventional decoherence processes [5,10].

In several of the above proposals a small quantum system (e.g., a photon [4–6] or a superconducting island [7]) is reversibly coupled to a large system (e.g., a moveable mirror [4–6] or a cantilever [7]) in order to create a macroscopic superposition. The existence of the quantum superposition of the large system is verified by observing the disappearance and reappearance of interference for the small system, as the large system is driven into a superposition and then returns to its initial state. The challenge is to find a feasible implementation of this idea.

Our proposal develops on the ideas in Refs. [4,5]. We also use results from Ref. [6], which relies on coupling between atoms and photons in a microcavity to create and detect superposition states of a moveable mirror. In particular, the formalism used in Ref. [6], based on Refs. [11,12], is applicable to our case. The main purpose here is to show that our purely optical proposal has the potential to be performed with current technology.

Principle.—The proposed setup, shown in Fig. 1, consists of a Michelson interferometer which has a high-finesse cavity in each arm. The cavity in arm (A) contains a tiny mirror attached to a micromechanical oscillator, similar to the cantilevers in atomic force microscopes. The cavity is used to enhance the radiation pressure of the photon on the mirror. The initial superposition of the photon being in either arm causes the system to evolve into a superposition of states corresponding to two distinct locations of the mirror. The observed interference of the photon allows one to study the creation of coherent superposition states of the mirror.

The system can be described by a Hamiltonian [6,11]

$$H = \hbar\omega_c a^\dagger a + \hbar\omega_m b^\dagger b - \hbar G a^\dagger a(b + b^\dagger), \quad (1)$$

where ω_c and a are the frequency and creation operator for the photon in the cavity, ω_m and b are the frequency and phonon creation operator for the center of mass motion of the mirror, and $G = (\omega_c/L)\sqrt{(\hbar/2M\omega_m)}$ is the coupling constant, where L is the cavity length and M is the mass of the mirror.

Let us suppose that initially the photon is in a superposition of being in either arm A or B, and the mirror is in

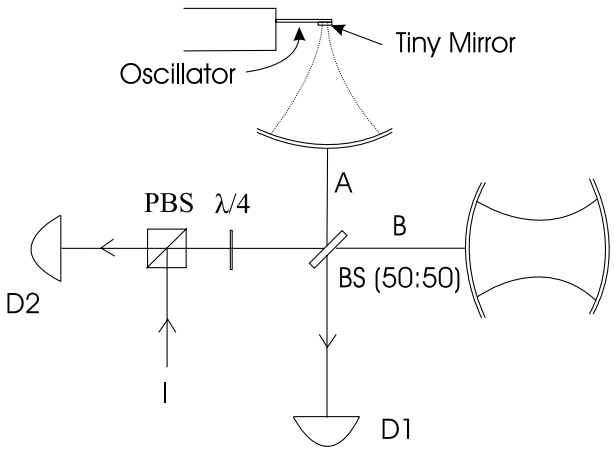


FIG. 1. The proposed setup: a Michelson interferometer for a single photon, where in each arm there is a high-finesse cavity. The cavity in arm A has a very small end mirror mounted on a micromechanical oscillator. The single photon comes in through I. If the photon is in arm A, the motion of the small mirror is affected by its radiation pressure. The photon later leaks out of either cavity and is detected at D1 or D2.

its ground state $|0\rangle_m$. Then the initial state is $|\psi(0)\rangle = (1/\sqrt{2})(|0\rangle_A|1\rangle_B + |1\rangle_A|0\rangle_B)|0\rangle_m$. After a time t the state of the system will be given by [6,12]

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}} e^{-i\omega_c t}[|0\rangle_A|1\rangle_B|0\rangle_m + e^{i\kappa^2(\omega_m t - \sin\omega_m t)}|1\rangle_A|0\rangle_B \times |\kappa(1 - e^{-i\omega_m t})\rangle_m], \quad (2)$$

where $\kappa = G/\omega_m$, and $|\kappa(1 - e^{-i\omega_m t})\rangle_m$ denotes a coherent state with amplitude $\kappa(1 - e^{-i\omega_m t})$. In the second term on the right-hand side the mirror moves under the influence of the radiation pressure of the photon in cavity A. The mirror oscillates around a new equilibrium position determined by the driving force. The parameter κ quantifies the displacement of the mirror in units of the size of the ground state wave packet.

The maximum interference visibility for the photon is given by twice the modulus of the off-diagonal element of the photon's reduced density matrix. By tracing over the mirror one finds from Eq. (2) that the off-diagonal element has the form $\frac{1}{2} e^{-\kappa^2(1-\cos\omega_m t)} e^{i\kappa^2(\omega_m t - \sin\omega_m t)}$. The first factor is the modulus, reaching a minimum after half a period at $t = \pi/\omega_m$, when the mirror is at its maximum displacement. The second factor gives the phase, which is identical to that obtained classically due to the varying length of the cavity.

In the absence of decoherence, after a full period, the system is in the state $(1/\sqrt{2})(|0\rangle_A|1\rangle_B + e^{i\kappa^2 2\pi}|1\rangle_A|0\rangle_B) \times |0\rangle_m$, such that the mirror is again disentangled from the photon. Full interference can be observed if the photon is detected at this time, provided that the phase factor $e^{i\kappa^2 2\pi}$ is taken into account. This revival, shown in Fig. 2, demonstrates the coherence of the superposition state that exists at intermediate times. For $\kappa^2 \gtrsim 1$ the superposition involves two distinct mirror positions. If the environment of the mirror "remembers" that the mirror has moved, then, even after a full period, the photon will still be entangled with the mirror's environment, and thus the revival will not be complete. Therefore the setup can be used to measure the decoherence of the mirror.

Here we have assumed that the mirror starts out in its ground state. We will argue below that optical cooling close to the ground state should be possible. However, in Ref. [6] it was shown that this is not necessary for observing the revival, although for a thermal mirror state with an average phonon number $\bar{n} = 1/(e^{\hbar\omega_m/kT} - 1)$ the revival peak is narrowed by a factor of $\sqrt{\bar{n}}$, leading to stricter requirements on the stability; see Fig. 2 and the discussion below. We now discuss the experimental requirements for achieving a superposition of distinct mirror positions and for observing the revival at $t = 2\pi/\omega_m$.

Conditions for displacement by ground state size.—We require $\kappa^2 \gtrsim 1$, which implies the momentum imparted by the photon has to be larger than the initial quantum uncertainty of the mirror's momentum. Let N denote the number of round-trips of the photon in the cavity during one period of the mirror's motion, such that $2NL/c = 2\pi/\omega_m$. The condition $\kappa^2 \gtrsim 1$ can be written

$$\frac{2\hbar N^3 L}{\pi c M \lambda^2} \gtrsim 1, \quad (3)$$

where λ is the wavelength of the light. The factors entering Eq. (3) are not all independent. The achievable N, determined by the quality of the mirrors, and the minimum mirror size (and hence M) both depend on λ. The mirror's lateral dimension should be an order of magnitude larger than λ to limit diffraction losses. The thickness required in order to achieve sufficiently high reflectivity depends on λ as well.

Equation (3) allows one to compare the viability of different wavelength ranges. While the highest values for N are achievable for microwaves using superconducting mirrors (up to 10^{10}), this is counteracted by their longer wavelengths. On the other hand, there are no good mirrors for highly energetic photons. The optical regime is optimal, given current mirror technology. We propose an experiment with λ around 630 nm.

The cavity mode needs to have a sharp focus on the tiny mirror, which requires the other cavity end mirror to be large due to beam divergence. The maximum cavity length is therefore limited by the difficulty of making large high-quality mirrors. We propose a cavity length of 5 cm, and a small mirror size of $10 \times 10 \times 10$ μm, leading to a mass of order 5×10^{-12} kg.

Such a mirror on a mechanical oscillator can be fabricated by coating a silicon cantilever with alternating layers of SiO_2 and a metal oxide. The best current optical mirrors are made in this way. A larger silicon oscillator has been coated with SiO_2/Ta_2O_5 and used as part of a high-finesse cavity in Ref. [13].

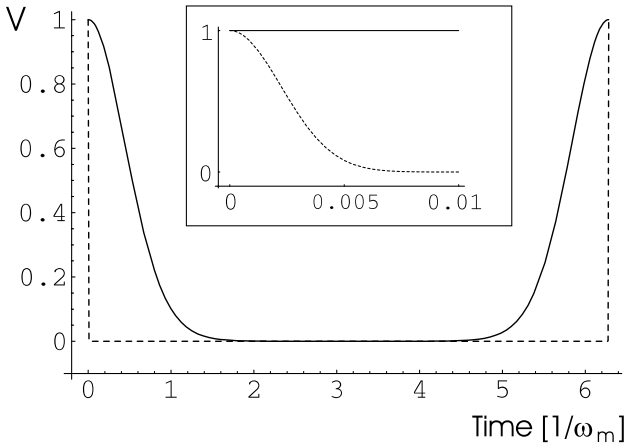


FIG. 2. Time evolution of the interference visibility V of the photon over one period of the mirror's motion for the case where the mirror has been optically cooled close to its ground state ($\bar{n} = 2$, solid line) and for $T = 2$ mK, which corresponds to $\bar{n} = 100\,000$ (dashed line—see also inset). The visibility decays after $t = 0$, but in the absence of decoherence there is a revival of the visibility after a full period. The width of the revival peak scales like $1/\sqrt{\bar{n}}$.

For the above dimensions the condition Eq. (3) is satisfied for $N = 5.6 \times 10^6$. Correspondingly, photon loss per reflection must be smaller than 3×10^{-7}, about a factor of 4 below reported values for such mirrors [14] and for a transmission of 10^{-7}, consistent with a 10 μm mirror thickness. For these values, about 1% of the photons are still left in the cavity after a full period of the mirror. For the above values of N and L one obtains a frequency $\omega_m = 2\pi \times 500$ Hz. This corresponds to a spread of the mirror's ground state wave function of order 10^{-13} m.

The fact that a relatively large L is needed to satisfy Eq. (3) implies that the creation of superpositions following the microcavity based proposal of Ref. [6] imposes requirements beyond current technology. A large L is helpful because, for a given N, it allows a lower frequency ω_m, and thus a more weakly bound mirror that is easier to displace by the photon.

Decoherence.—The requirement of observing the revival puts a bound on the acceptable environmental decoherence. To estimate the expected decoherence we model the mirror's environment by an (Ohmic) bath of harmonic oscillators. The effect of this can approximately be described by a decoherence rate $\gamma_D = \gamma_m k T_E M(\Delta x)^2/\hbar^2$ governing the decay of off-diagonal elements between different mirror positions [2]. Here γ_m is the damping rate for the mechanical oscillator, T_E is the temperature of the environment, which is constituted mainly by the internal degrees of freedom of the mirror and cantilever, and Δx is the separation of two coherent states that are originally in a superposition. This approximation is strictly valid only for times much longer than $2\pi/\omega_m$ and for Δx large compared to the width of the individual wave packets. Here we assume that the order of magnitude of the decoherence is well captured by γ_D. If the experiment achieves $\kappa^2 \gtrsim 1$, i.e., a separation by the size of a coherent state wave packet, $\Delta x \sim \sqrt{(\hbar/M\omega_m)}$, the condition $\gamma_D \lesssim \omega_m$ can be cast in the form

$$Q \gtrsim \frac{kT_E}{\hbar\omega_m}, \tag{4}$$

where $Q = \omega_m/\gamma_m$ is the quality factor of the mechanical oscillator. For $Q \gtrsim 10^5$, which has been achieved [15] for silicon cantilevers of approximately the right dimensions and frequency, this implies that the temperature of the environment has to be of the order of 2 mK, which is achievable with state-of-the-art dilution refrigerators.

Optical cooling.—Cooling the mirror's center of mass motion significantly eases the stability requirements for the proposed experiment. A method for optical cooling of a mirror via feedback was first proposed in Ref. [16]. By observing the phase of the output field of a cavity, its length can be measured with high precision. This can be used to implement a feedback mechanism that cools the center of mass motion of the mirror far below the temperature of its environment. A variation of the original scheme was experimentally implemented in Ref. [17], where a vibrational mode of a macroscopic mirror was cooled using a feedback force proportional to the natural damping force, but larger by a gain factor g. The size of g determines the achievable final temperature for a given T_E. For a tiny mirror, large gain values are realistic using the radiation pressure of a second laser beam to implement the feedback force. To analyze cooling to the quantum regime, one has to take into account the fact that measurement and feedback introduce noise, Ref. [18].

For our proposed experiment the constant component of the feedback laser has to balance the force from the measurement field, since otherwise the mirror would start to oscillate when the light is turned off. Adapting Ref. [19], the final energy of the cooled mirror is given by

$$E_c = \frac{\hbar\omega_m}{2}\frac{1}{2(1+g)}\left[\frac{4k_B T_E}{\hbar\omega_m} + 2\zeta + \frac{g^2}{\eta\zeta}\right], \tag{5}$$

where T_E is the temperature of the mirror's environment, $\zeta = (64\pi c P/M\gamma_m\omega_m\lambda\gamma_c^2 L^2)$, with P the light intensity incident on the measurement cavity and γ_c the cavity decay rate, and η the detection efficiency. The first term in Eq. (5) comes from the original thermal fluctuations, which are suppressed by the feedback. The second term is the back action noise from the measurement and feedback light. It differs from the formula of Ref. [19] by a factor of 2 to include the noise from the feedback laser. The third term is the noise due to imperfect measurement. Increasing the light intensity in the cavity improves the measurement precision, but also increases the back action noise.

The energy of the mirror can be made very close to its ground state energy choosing realistic parameter values; $E_c = \hbar\omega_m$ can be achieved with $g = 6 \times 10^5$, $T_E = 2$ mK, $P = 10^{-8}$ W, $\gamma_c = 3 \times 10^7$ s^{-1}, $\lambda = 800$ nm, $\eta = 0.8$, $\gamma_m = 0.03$ s^{-1}, and M, ω_m, L as before. The necessary feedback force for such a high value of g can be achieved with a feedback laser intensity modulation of $\Delta P_{fb} = 10^{-6}$ W. To balance the measurement field, the constant component of the feedback laser should be $\bar{P}_{fb} = 4 \times 10^{-6}$ W. The relatively large value of γ_c can be achieved in the cavity used in the superposition experiment by working at a wavelength away from where the mirrors are optimal.

Once the mirror has been cooled close to its ground state, which is reached in a time of order $1/(\gamma_m g)$ [20], the measurement and feedback laser fields should be turned off simultaneously. Then the experiment proceeds as described above. Reheating of the mirror happens at a time scale of $1/\gamma_m$ [20] and thus is not a problem for a high-Q oscillator. After every run of the experiment, the mirror has to be reset to its initial state by the optical cooling procedure.

Stability.—The distance between the large cavity end mirror and the equilibrium position of the small mirror has to be stable to of order $\lambda/20N = 0.6 \times 10^{-14}$ m over

the whole measurement time, which is determined as follows. A single run of the experiment starts by sending a weak pulse into the interferometer, such that on average 0.1 photons go into either cavity. This probabilistically prepares a single-photon state as required to a good approximation. The two-photon contribution has to be kept low because it causes noise in the interferometer. Considering the required low value of ω_m and the fact that approximately 1% of the photons remain after a full period for the assumed loss, this implies a detection rate of approximately 10 photons per minute in the revival interval. Thus we demand stability to of order 10^{-14} m over a few minutes. Stability of order 10^{-13} m/min for an STM at 8 K was achieved with a rather simple suspension [21]. Gravitational wave observatories using interferometers also require very high stability in order to have a length sensitivity of 10^{-19} m over time scales of a ms or greater, for arm lengths of order 1 km [22]. If the mirror is in a thermal state, the revival peak is narrowed by a factor $\sqrt{\bar{n}}$ [6], leading to lower count rates in the revival interval and thus making the stability requirements stricter by the same factor, cf. Fig. 2.

The experiment also requires ultrahigh vacuum conditions in order to ensure that events where an atom hits the cantilever are sufficiently rare not to cause significant errors, which is at the level of about 5/s. Background gas particle densities of order $100/\mathrm{cm}^3$ have been achieved [23] and are sufficient for our purposes.

Outlook and conclusions.—In principle the proposed setup has the potential to test wave function reduction models, in particular, the one of Ref. [5]. We estimate that the ratio Q/T needs to be improved by about 6 orders of magnitude from the values discussed in this Letter ($Q = 10^5$ and $T = 2$ mK) to make the predicted wave function decoherence rate comparable to the environmental decoherence rate. However, temperatures as low as 60 μK have been achieved with adiabatic demagnetization [24], while Q is known to increase with decreasing temperature [15] and through annealing [25].

We have performed a detailed study of the experimental requirements for the creation and observation of quantum superposition states of a mirror consisting of 10^{14} atoms, approximately 9 orders of magnitude more massive than any superposition observed to date. Our analysis shows that, while very demanding, this goal appears to be within reach of current technology.

This work was supported by the E.U. (IST-1999-10033). W. M. is supported by EPSRC (Contract No. 00309297). C. S. is supported by a Marie Curie Fund of the E.U. (No. HPMF-CT-2001-01205) and thanks the University of California Santa Barbara for its hospitality. R. P. thanks the NSF for support under Contract No. 00-90091 and the Leverhulme Foundation. We would like to thank S. Bose, M. Davies, M. de Dood, T. Knuuttila, R. Lalezari, A. Lamas-Linares, and J. Pethica for useful discussions.

[1] E. Schrödinger, Naturwissenschaften **23**, 807 (1935).
[2] E. Joos *et al.*, *Decoherence and the Appearance of a Classical World in Quantum Theory* (Springer, Berlin, 1996); W. H. Zurek, Phys. Today **44**, 36 (1991).
[3] J. Ruostekoski, M. J. Collett, R. Graham, and D. F. Walls, Phys. Rev. A **57**, 511 (1998); J. I. Cirac, M. Lewenstein, K. Molmer, and P. Zoller, Phys. Rev. A **57**, 1208 (1998).
[4] D. Bouwmeester, J. Schmiedmayer, H. Weinfurter, and A. Zeilinger, in *Gravitation and Relativity: At the Turn of the Millennium*, edited by N. Dadhich and J. Narlikar (IUCAA, Pune, 1998). This paper was based on discussions between J. Schmiedmayer, R. Penrose, D. Bouwmeester, J. Dapprich, H. Weinfurter, and A. Zeilinger (1997).
[5] R. Penrose, in *Mathematical Physics 2000*, edited by A. Fokas *et al.* (Imperial College, London, 2000).
[6] S. Bose, K. Jacobs, and P. L. Knight, Phys. Rev. A **59**, 3204 (1999).
[7] A. D. Armour, M. P. Blencowe, and K. C. Schwab, Phys. Rev. Lett. **88**, 148301 (2002).
[8] C. H. van der Wal *et al.*, Science **290**, 773 (2000); J. R. Friedman, V. Patel, W. Chen, S. K. Tolpygo, and J. E. Lukens, Nature (London) **406**, 43 (2000).
[9] M. Arndt *et al.*, Nature (London) **401**, 680 (1999); W. Schöllkopf and J. P. Toennies, Science **266**, 1345 (1994).
[10] G. C. Ghirardi, A. Rimini, and T. Weber, Phys. Rev. D **34**, 470 (1986); G. C. Ghirardi, P. Pearle, and A. Rimini, Phys. Rev. A **42**, 78 (1990); I. C. Percival, Proc. R. Soc. London A **447**, 189 (1994); D. I. Fivel, Phys. Rev. A **56**, 146 (1997); L. Diósi, Phys. Rev. A **40**, 1165 (1989).
[11] C. K. Law, Phys. Rev. A **51**, 2537 (1994); C. K. Law, Phys. Rev. A **49**, 433 (1993).
[12] S. Mancini, V. I. Man'ko, and P. Tombesi, Phys. Rev. A **55**, 3042 (1997).
[13] I. Tittonen *et al.*, Phys. Rev. A **59**, 1038 (1999).
[14] G. Rempe, R. J. Thompson, H. J. Kimble, and R. Lalezari, Opt. Lett. **17**, 363 (1992); C. J. Hood, H. J. Kimble, and J. Ye, Phys. Rev. A **64**, 033804 (2001).
[15] H. J. Mamin and D. Rugar, Appl. Phys. Lett. **79**, 3358 (2001).
[16] S. Mancini, D. Vitali, and P. Tombesi, Phys. Rev. Lett. **80**, 688 (1998).
[17] P. F. Cohadon, A. Heidmann, and M. Pinard, Phys. Rev. Lett. **83**, 3174 (1999).
[18] J.-M. Courty, A. Heidmann, and M. Pinard, Eur. Phys. J. D **17**, 399 (2001).
[19] D. Vitali, S. Mancini, L. Ribichini, and P. Tombesi, quant-ph/0211102.
[20] M. Pinard, P. F. Cohadon, T. Briant, and A. Heidmann, Phys. Rev. A **63**, 013808 (2000).
[21] B. C. Stipe, M. A. Rezaei, and W. Ho, Rev. Sci. Instrum. **70**, 137 (1999).
[22] S. Rowan and J. Hough, Living Rev. Relativity **3**, 2000 (2000).
[23] G. Gabrielse *et al.*, Phys. Rev. Lett. **65**, 1317 (1990).
[24] W. Yao *et al.*, J. Low Temp. Phys. **120**, 121 (2000).
[25] J. Yang, T. Ono, and M. Esashi, Appl. Phys. Lett. **77**, 3860 (2000).

LETTERS

Sub-kelvin optical cooling of a micromechanical resonator

Dustin Kleckner[1] & Dirk Bouwmeester[1]

Micromechanical resonators, when cooled down to near their ground state, can be used to explore quantum effects such as superposition and entanglement at a macroscopic scale[1–3]. Previously, it has been proposed to use electronic feedback to cool a high frequency (10 MHz) resonator to near its ground state[4]. In other work, a low frequency resonator was cooled from room temperature to 18 K by passive optical feedback[5]. Additionally, active optical feedback of atomic force microscope cantilevers has been used to modify their response characteristics[6], and cooling to approximately 2 K has been measured[7]. Here we demonstrate active optical feedback cooling to 135 ± 15 mK of a micromechanical resonator integrated with a high-quality optical resonator. Additionally, we show that the scheme should be applicable at cryogenic base temperatures, allowing cooling to near the ground state that is required for quantum experiments—near 100 nK for a kHz oscillator.

Using a laser tuned to the resonance fringe of a high finesse optical cavity, it is possible to observe very small fluctuations in the length of the cavity due to brownian motion of one or both of the end mirrors. We have developed an optical cavity with one rigid large mirror, 6 mm in diameter and with a 25 mm radius of curvature, and one tiny plane mirror, 30 μm in diameter, attached to a commercial atomic force microscope cantilever of dimensions $450 \times 50 \times 2$ μm with a fundamental resonance of 12.5 kHz (Fig. 1b). An optical finesse of 2,100 and a mechanical quality factor of 137,000 have been achieved with the system[8]. The motion of the tiny mirror/cantilever is monitored by measuring the transmission of the cavity at a frequency on the side of an optical resonance peak. To do this, we use about 1 mW from a 780 nm tunable diode laser which is locked to the resonance fringe using the integrated signal from a photo-multiplier tube which monitors the light transmitted through the cavity (Fig. 1a). The time derivative of this signal is proportional to the velocity of the cantilever tip and is used to modulate the amplitude of a second, 980 nm, diode laser focused on the cantilever less than 100 μm away from the tiny mirror. The radiation pressure exerted by this feedback laser counteracts the motion of the mirror and effectively provides cooling of the fundamental mode.

The effective feedback gain can be varied over several orders of magnitude by sending the feedback laser through a variable neutral density filter. The average power in the feedback beam when it reaches the cantilever is of the order of 1 mW at the highest gain settings and proportionally lower otherwise. The mean modulation depth of the feedback beam varies from nearly 100% to less than 5% as gain is increased. The vibration spectrum of the cantilever as a function of gain is shown in Fig. 2. The r.m.s. thermal amplitude of the cantilever without feedback is 1.2 ± 0.1 Å. From this value, one can calculate that the spring constant of the cantilever is $0.15 \pm 0.01\ \mathrm{N\,m^{-1}}$, in agreement with the manufacturer-specified range, and the effective mass of the cantilever fundamental mode is $(2.4 \pm 0.2) \times 10^{-11}$ kg.

To determine the effective gain of the feedback loop and the temperature of the fundamental mode, we fit a lorentzian plus a constant background to the vibration spectrum of the cantilever for each value of feedback gain. The temperature is determined from the area under the lorentzian without the background, while the gain is determined by the width of the resonance. The linewidth provides a good measure of gain because it is directly determined by the damping rate whereas the cantilever amplitude may be affected by other sources of noise in the feedback loop. Cooling is observed over more than three orders of magnitude. The lowest temperature we are able to measure is 135 ± 15 mK, or a cantilever r.m.s. amplitude of 0.023 ± 0.002 Å, with a gain (the ratio of feedback to mechanical damping) of $g = 2{,}490 \pm 90$ (Fig. 2b). The lowest trace in Fig. 2b, indicating an even lower temperature, cannot be reliably fitted owing to the laser noise floor. Since the optical finesse is not the current limiting factor, we operate the opto-mechanical system at a finesse of only 200,

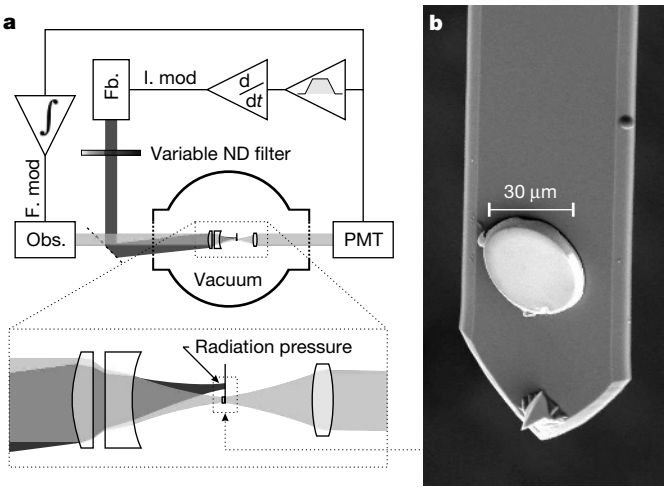


Figure 1 | The experimental system. a, Diagram of the feedback mechanism: a 780 nm observation laser (Obs.) is frequency locked to the optical cavity (shown magnified at bottom) with an integrating circuit (via the laser frequency modulation input, f. mod), using the signal from a photomultiplier tube (PMT). This signal is also sent through a 1.25 kHz bandpass filter at 12.5 kHz and a derivative circuit (d/d*t*) to provide an intensity-modulating signal (I. mod.) for the 980 nm feedback laser (Fb.). The feedback laser is attenuated with a variable neutral density (ND) filter to adjust the gain of the feedback. The feedback force is exerted on the cantilever via this laser's radiation pressure. **b**, Scanning electron microscope image of the tip of the cantilever with attached mirror.

[1]Department of Physics, University of California, Santa Barbara, California 93106, USA.

produced by slight cavity misalignment, which makes the system less sensitive to transient vibrations.

The amplitude of the mirror motion can be calculated in the presence of feedback by assuming that the Langevin force—the effective thermal force that maintains brownian motion—remains constant while the mechanical susceptibility of the mirror is reduced by the dissipation due to the radiation feedback pressure. It suffices to consider only the fundamental mode of the mirror motion, represented by a damped harmonic oscillator. In this approximation, the power spectrum of the mirror's motion in the presence of feedback becomes[9]:

$$S_x^{\mathrm{fb}}[\Omega] = \frac{2\Gamma_0 k_B T_0}{M} \frac{1}{(\omega^2-\Omega^2)^2+(1+g)^2\Gamma_0^2\Omega^2} \quad (1)$$

where Ω is the observation frequency, ω is the resonator frequency, Γ_0 is the mechanical damping factor, M is the effective mass of the resonator mode, k_B is Boltzmann's constant, T_0 is the bulk temperature of the resonator and g is the gain. $g=0$ corresponds to the vibration spectrum in the absence of feedback. The motion of the oscillator in the presence of feedback is the same as that of an oscillator with lower temperature and a higher damping constant:

$$T_{\mathrm{fb}} = (1+g)^{-1} T_0 \quad (2)$$

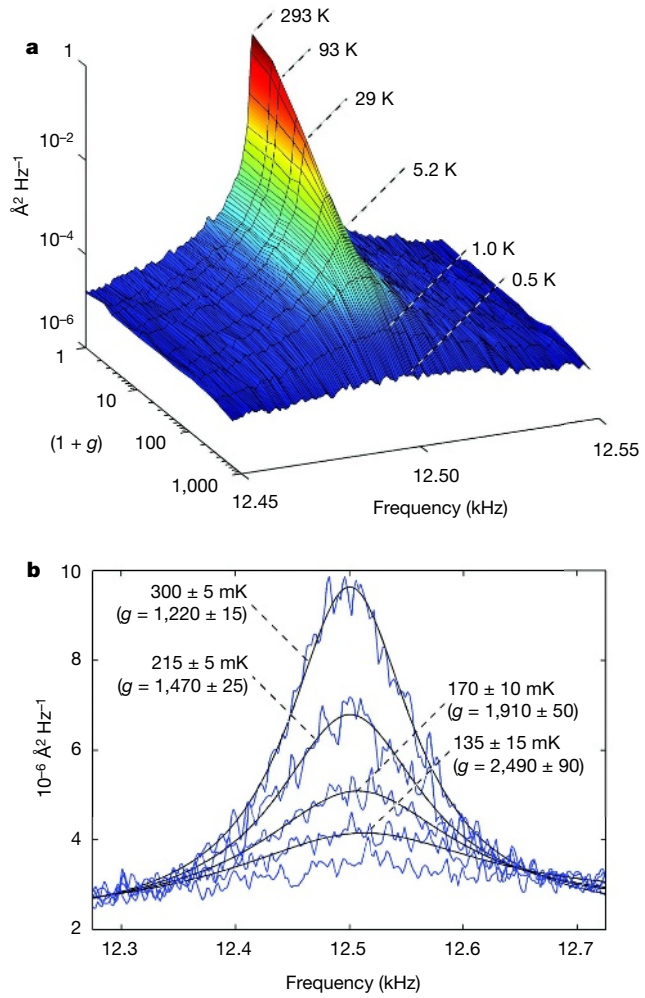


Figure 2 | Single-sided thermal vibration spectrum of the cantilever as it is cooled. g is the dimensionless gain factor, which is the ratio of feedback to mechanical damping. **a**, Spectrum at low to moderate gains. **b**, Spectrum near the background noise level for large gains. The blue curves correspond to experimental data, and the black curves to fits of a gaussian function plus a background. The lowest trace cannot be reliably fitted.

$$\Gamma_{\mathrm{fb}} = (1+g)\Gamma_0 \quad (3)$$

The optical feedback scheme, when analysed in terms of noiseless classical light fields, can be seen as a virtual viscous force, which unlike a real viscous force creates dissipation without introducing fluctuations. As discussed below, the cooling temperature as demonstrated here is limited by laser frequency fluctuations. Ultimately, optical cooling should be limited by the balance of residual heating and quantum noise in the observation and feedback laser signals.

For a signal-to-noise ratio of one in spectral density at the peak of the mechanical resonance, the temperature of the cantilever would be (as can be derived from equation (1)):

$$T_{\mathrm{min}} \cong \sqrt{\frac{T_0 M \omega^3 S_{\mathrm{noise}}}{2k_B Q}} \quad (4)$$

where S_{noise} is the equivalent position noise in the interferometer measurement and $Q=\omega/\Gamma_0$ is the mechanical quality factor. For higher values of gain, the feedback signal is mostly noise and lower temperatures can not be conclusively demonstrated. For our experiment, the equivalent noise level is $\sqrt{S_{\mathrm{noise}}} \approx 10^{-3}\,\text{Å}\,\mathrm{Hz}^{-1/2}$. This corresponds to the expected noise due to the frequency fluctuations of a free running tunable laser diode, which are of order $10^3\,\mathrm{Hz}\,\mathrm{Hz}^{-1/2}$ at the resonance frequency of 12.5 kHz (ref. 10). With the system in vacuum at pressures of 10^{-6} mbar, so as to maximize the mechanical quality factor of the cantilever, this noise level corresponds to a minimum temperature of the order of 100 mK, in good agreement with the experimental data.

An alternative approach to study the cooling is to analyse the temporal response of the system by gating the signal to the feedback laser. The characteristic time constant for the system to reach equilibrium after the cooling is turned on is given by:

$$\tau_{\mathrm{fb}} = \Gamma_{\mathrm{fb}}^{-1} = (1+g)^{-1}\Gamma_0^{-1} \quad (5)$$

To observe this behaviour, we monitor the cantilever over many 10 s periods during each of which the cooling is on for 3 s. Data for cooling to 1.8 ± 0.2, 4.0 ± 0.2 and 6.4 ± 0.1 K and returning to thermal equilibrium are shown in Fig. 3. The cooling times are measured to be 9.0 ± 0.5, 19 ± 1 and 27 ± 1 ms, respectively. The reheating time is found to be indistinguishable for all three gains with an average of $\tau_0 = 1.30 \pm 0.05$ s. This is in agreement with the linewidth of the cantilever measured without feedback, $\Gamma_0 = 680 \pm 50$ mHz. In

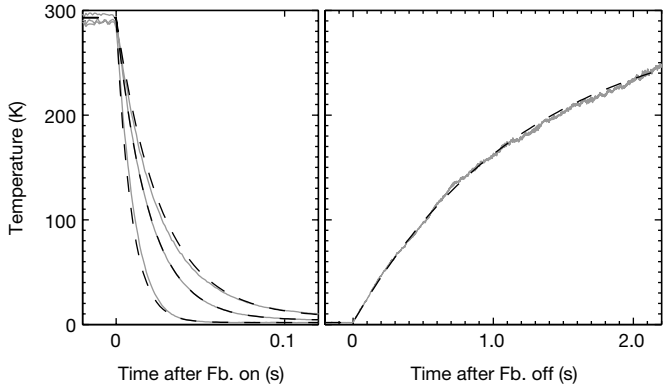


Figure 3 | Temporal response of the cantilever to cooling pulses. The temperature is determined by calculating the total vibrational amplitude of the cantilever between 12 and 13 kHz in 1 ms bins and subtracting the background. Each data set is the average of 1,000 samples. The three sets in the left panel correspond to cooling to 6.4, 4.0 and 1.8 K (solid lines, top to bottom). Heating is shown (right panel) for only one data set (1.8 K), as all three are nearly coincident. The dashed lines are fits to exponential decays, used to determine the cooled temperature and the cooling and reheating times. Fb. refers to the feedback system.

accordance with theory, the ratio of the reheating to the cooling times, τ_0/τ_{fb}, and the corresponding ratio of the spectral linewidths from the earlier measurements, Γ_{fb}/Γ_0, are found to be the same as the cooling factor, T_0/T_{fb}, within statistical uncertainties.

In experiments where optical feedback is used on cantilevers with non-uniform composition, radiation pressure is typically overwhelmed by the photothermal force, which is an effective force due to thermally induced bending[5,6]. Although this is not the case for single-crystal silicon cantilevers, the addition of a tiny mirror on the tip of our cantilever should produce a weak photothermal force. This force can be distinguished from radiation pressure by its dependence on the intensity modulation frequency of the feedback laser. Whereas radiation pressure is independent of modulation frequency, the photothermal force is not, because it has a characteristic response time, τ, related to the thermal relaxation time of the cantilever. A simple model for the frequency dependence of the photothermal force, $F_{pt}(\Omega)$, gives:

$$F_{pt}(\Omega) \cong \int_0^\infty \frac{F_{pt}(0)}{\tau} e^{-\frac{t}{\tau}} e^{-i\Omega t} dt = \frac{F_{pt}(0)}{1+i\Omega\tau} \quad (6)$$

where $e^{-i\Omega t}$ corresponds to the input power modulation, and $e^{-\frac{t}{\tau}}$ is due to the thermal relaxation. This is consistent with the frequency dependence of the photothermal force as described in previous work[6]. To test for the presence of photothermal force in our resonator, the feedback laser was modulated at a range of frequencies from 100 Hz to 20 kHz and the mechanical response of the cantilever was measured as before (Fig. 4). The power in the feedback laser reflected from the cantilever was determined to have a mean of 2.7 ± 0.5 mW and a modulation amplitude of 1.0 ± 0.2 mW, independent of the modulation frequency. This results in a radiation pressure force of $F_{rad} = 2P_{mod}/c = 6.7 \pm 1.3$ pN(where P_{mod} is the amplitude of the power modulation and c is the speed of light) at the modulation frequency.

If the driving frequency is sufficiently far from the cantilever resonance, the mechanical damping constant can be ignored and the amplitude of the cantilever's motion should be of the form:

$$A(\Omega) = \left| \frac{\frac{A_{pt}}{1+i\Omega\tau} + A_{rad}}{1-\left(\Omega/\omega\right)^2} \right| \quad (7)$$

where Ω is the driving frequency, ω is the resonance frequency, τ is the photothermal characteristic time, and A_{rad} and A_{pt} are the magnitudes of the motion due to the radiation pressure and photothermal force alone, at zero frequency. The term in the denominator is due to mechanical amplification by the cantilever resonance. This equation fits well to the measured response (Fig. 4), resulting in $A_{rad} = 0.470 \pm 0.005$ Å, $A_{pt} = -6.3 \pm 0.2$ Å and $\tau = 30 \pm 2$ ms. At frequencies greater than 5 kHz, radiation pressure is observed to be the dominant force mechanism, whereas the photothermal force is relevant only at lower frequencies.

Assuming the constant force background described by A_{rad} is entirely due to radiation pressure, one can calculate the spring constant of the cantilever at the position where the feedback laser is focused to be $k = F_{rad}/A_{rad} = 0.14 \pm 0.03$ N m^{-1}, in agreement with the value for the spring constant obtained earlier. Near the fundamental resonance of the cantilever, the radiation pressure is calculated to be almost 5 times larger than the photothermal force. Additionally, the two forces should be nearly 90° out of phase at this frequency, given that the time constant of the photothermal force is found to be 30 ± 2 ms. Thus the radiation pressure is responsible for almost all of the demonstrated feedback cooling; in the absence of photothermal force, the total feedback force would be reduced by less than 3%.

When optical cooling is active, the cantilever's motion is strongly damped, making it undesirable for many types of measurements. In some cases this problem can be overcome with a stroboscopic cooling scheme, where measurements are only made in the periods when the cooling is off. In addition to being of direct importance for the aforementioned massive superposition experiment, this scheme has already been theoretically shown to be useful for high sensitivity measurements of position and weak impulse forces[11]. Because the cooling is faster than the heating by a factor $(1 + g)$, a low temperature can be maintained even when the cooling is off the majority of the time. However, maintaining low temperatures requires that the measurement window be short; if it is, for example, one oscillation period long, the temperature of the oscillator will have increased by $\Delta T \approx 2\pi T_0/Q$ by the end of each measurement window, meaning that cooling past this point results in marginal improvement.

We now evaluate the potential for reaching even lower temperatures for the purpose of studying quantum effects in similar systems. Reference 3 proposes an experiment: putting a mechanical oscillator in a quantum superposition of vibrating and not-vibrating by interaction with the light pressure of a single photon in an optical cavity of which one end mirror is attached to the oscillator. Appropriate for such a scheme would be a 250-μm-long silicon cantilever with a 20-μm-diameter dielectric mirror on the tip and a resonance frequency of 1 kHz. Because of the constraints of environmentally induced decoherence[12,13], the bulk temperature must be less than $T_{EID} = Q\hbar\omega/k_B = 8$ mK for the cantilever to remain coherent over one period, given $Q \approx 150{,}000$. This temperature is achievable by conventional means; nuclear adiabatic demagnetization of $PrNi_5$ (ref. 14) could be employed as the final, vibration-free, cooling stage, as it is able to be started from temperatures previously demonstrated

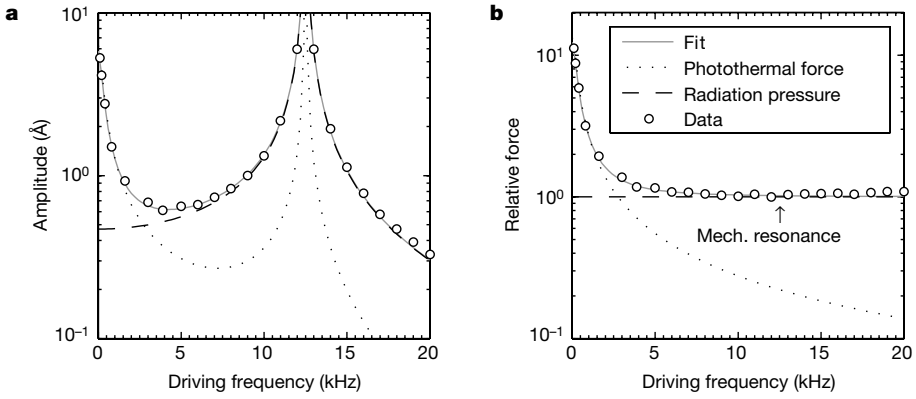


Figure 4 | Response of the cantilever to an external intensity-modulated laser. a, The amplitude of the cantilever's motion at the driving frequency. **b**, The force on the cantilever, calculated by dividing the amplitude by the mechanical amplification of the cantilever. In both graphs the magnitude of the contributions (ignoring phase differences) of the photothermal force and radiation pressure are shown as dashed and dotted lines, respectively. The slight deviation of the fit from the data at higher frequencies is due to higher-order flexural modes.

for vibration-isolated cold stages ($\sim$100 mK)[15,16]. The observation period for a massive superposition experiment is one oscillation long, thus the maximum useful cooling factor is $Q/2\pi \approx 25{,}000$ as discussed above. This corresponds with a temperature of 300 nK or a mean oscillator quantum number of only 2π.

It has been shown theoretically that optical feedback still works in the quantum regime, allowing cooling to the ground state[17]. Experimentally, cooling to the quantum regime requires the capability to accurately monitor the position of the cantilever without introducing significant heating. With an optical finesse $F = 5 \times 10^5$, which should be technologically achievable[8], an observation beam power of 1 aW, or about 5,000 photons per second, is enough to reduce shot noise to the appropriate level. Assuming the thermal conductivity of the cantilever is reduced to the one-dimensional quantum limit, the cantilever's thermal resistivity will be roughly 30 mK aW^{-1} (ref. 18). The heating from the feedback laser can be reduced by use of a sufficiently long wavelength laser so that absorption is negligible; this is not possible for the readout beam, which must be resonant with the optical cavity. Thus as long as the observation laser has relatively low absorption in the cantilever/mirror, it should not significantly affect the bulk temperature. This implies that cooling a kHz oscillator to near its ground state should be possible, drastically simplifying the experimental requirements to observe quantum phenomena in this system.

We have demonstrated active laser feedback cooling of a micromechanical oscillator, using only radiation pressure, from room temperature to 135 mK. Furthermore, we have shown that this cooling method could be used in addition to traditional cryogenics to reach much lower temperatures, even near the ground state of a kHz oscillator. This in turn would significantly aid the realization of proposals to create and investigate massive quantum superpositions.

Received 5 June; accepted 24 August 2006.

1. Bose, S., Jacobs, K. & Knight, P. L. Scheme to probe the decoherence of a macroscopic object. *Phys. Rev. A* **59**, 3204–3210 (1999).
2. Mancini, S., Vitali, D., Giovannetti, V. & Tombesi, P. Stationary entanglement between macroscopic mechanical oscillators. *Eur. Phys. J. D* **22**, 417–422 (2003).
3. Marshall, W., Simon, C., Penrose, R. & Bouwmeester, D. Towards quantum superpositions of a mirror. *Phys. Rev. Lett.* **91**, 130401 (2003).
4. Hopkins, A., Jacobs, K., Habib, S. & Schwab, K. Feedback cooling of a nanomechanical resonator. *Phys. Rev. B* **68**, 235238 (2003).
5. Metzger, C. H. & Karrai, K. Cavity cooling of a microlever. *Nature* **432**, 1002–1005 (2004).
6. Mertz, J., Marti, O. & Mlynek, K. Regulation of a microcantilever response by force feedback. *Appl. Phys. Lett.* **62**, 2344–2346 (1993).
7. Bruland, K. J., Garbini, J. L., Dougherty, W. M. & Sidles, J. A. Optimal control of force microscope cantilevers. II. Magnetic coupling implementation. *J. Appl. Phys.* **80**, 1959–1964 (1996).
8. Kleckner, D. *et al.* High finesse opto-mechanical cavity with a movable thirty-micron-size mirror. *Phys. Rev. Lett.* **96**, 173901 (2006).
9. Cohadon, P. F., Heidmann, A. & Pinard, M. Cooling of a mirror by radiation pressure. *Phys. Rev. Lett.* **83**, 3174–3177 (1999).
10. Turner, L. D., Weber, W. P., Hawthorn, C. J. & Scholten, R. E. Frequency noise characterisation of narrow linewidth diode lasers. *Opt. Commun.* **201**, 391–397 (2002).
11. Vitali, B. D., Mancini, S., Ribichini, L. & Tombesi, P. Mirror quiescence and high-sensitivity position measurements with feedback. *Phys. Rev. A* **65**, 063803 (2002).
12. Joos, E. *et al. Decoherence and the Appearance of a Classical World in Quantum Theory* 2nd edn (Springer, New York, 2003).
13. Zurek, W. H. Decoherence and the transition from quantum to classical. *Phys. Today* **44**, 36–44 (1991).
14. Parpia, J. M. *et al.* Optimization procedure for the cooling of liquid ^{3}He by adiabatic demagnetization of praseodymium nickel. *Rev. Sci. Instrum.* **56**, 437–443 (1985).
15. Moussy, N., Courtois, A. & Pannetier, B. A very low temperature scanning tunnelling microscope for the local spectroscopy of mesoscopic structures. *Rev. Sci. Instrum.* **71**, 128–131 (2001).
16. Mamin, H. J. & Rugar, D. Sub-attonewton force detection at millikelvin temperatures. *Appl. Phys. Lett.* **79**, 3358–3360 (2001).
17. Courty, J. M., Heidman, A. & Pinard, M. Quantum limits of cold damping with optomechanical coupling. *Eur. Phys. J. D* **17**, 399–408 (2001).
18. Schwab, K., Henriksen, E. A., Worlock, J. M. & Roukes, M. L. Measurement of the quantum of thermal conductance. *Nature* **404**, 974–977 (2000).

Acknowledgements This work was supported by the National Science Foundation. We thank M. de Dood, H. Eisenberg, S. Hastings, W. Irvine, A. Kahl, G. Khoury, W. Marshall and C. Simon for their contributions at earlier stages of this work.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to D.K. (dkleckner@physics.ucsb.edu).

Cavity Optomechanics: Back-Action at the Mesoscale

T. J. Kippenberg, *et al.*

Science **321**, 1172 (2008);
DOI: 10.1126/science.1156032

Cavity Optomechanics: Back-Action at the Mesoscale

T. J. Kippenberg[1]*† and K. J. Vahala[2]*

The coupling of optical and mechanical degrees of freedom is the underlying principle of many techniques to measure mechanical displacement, from macroscale gravitational wave detectors to microscale cantilevers used in scanning probe microscopy. Recent experiments have reached a regime where the back-action of photons caused by radiation pressure can influence the optomechanical dynamics, giving rise to a host of long-anticipated phenomena. Here we review these developments and discuss the opportunities for innovative technology as well as for fundamental science.

The reflection of a photon entails momentum transfer, generally referred to as "radiation pressure," with the resulting force called the scattering force. Besides this scattering force, the spatial variation of an intensity distribution can give rise to a gradient or dipole force. Interest in radiation pressure was first generated by the trapping of dielectric particles using laser radiation (*1*). This technique is widely adapted today in the biological and biophysical sciences and is known as the "optical tweezer." In atomic physics, the ability to cool atoms with the use of radiation pressure (*2*, *3*) has enabled many advances (*4*), including the realization of exotic quantum states such as Bose-Einstein condensates.

Radiation pressure can also have an effect on macroscale mechanical masses (such as on an optical interferometer's mirror) and has been considered theoretically for decades (*5*, *6*). The mutual coupling of optical and mechanical degrees of freedom in an optical resonator (or optical cavity) has been explored in laser-based gravitational wave interferometers, in which radiation pressure imposes limits on continuous position detection. Beyond setting detection limits, radiation pressure can also influence the dynamics of a harmonically bound mirror. A discernable effect on mirror motion was first demonstrated in the optical bistability resulting from the static elongation of cavity length caused by radiation pressure (*7*), and later, in work demonstrating the optical spring effect (a radiation-pressure–induced change in stiffness of the "mirror spring") (*8*). These phenomena, however, do not rely on the cavity delay; rather, each results from an adiabatic response of the cavity field to mechanical motion. Phenomena of a purely dynamical nature were predicted (*5*, *9*) to arise when the decay time of the photons inside the cavity is comparable to or longer than the mechanical oscillator period. Creating such delays through an electro-optic hybrid system was later proposed and demonstrated to induce radiation-pressure "feedback cooling" of a cavity mirror (*10*, *11*), also known as cold damping. Whereas in subsequent attempts dynamic radiation-pressure phenomena were masked by thermal effects (*12*), recent advances in micro- and nanofabrication made it possible to access the regime where the effects of cavity-enhanced radiation pressure alone dominate the mechanical dynamics. Demonstrations of mechanical amplification (*13*, *14*) and cooling (*14*–*16*) via dynamical back-action signal that a paradigm shift (*17*) in the ability to manipulate mechanical degrees of freedom is now under way, which has long been anticipated (*18*, *19*). Central to all current work is the role of back-action in setting dynamical control and performance limits. This review is intended to provide context for these recent accomplishments and also to present an overview of possible and anticipated future research directions.

Dynamical Back-Action Versus Quantum Back-Action

Photons at optical frequencies are uniquely suited to measure mechanical displacement for several reasons. First, because of the high energy of optical photons (~1 eV), thermal occupation is negligible at room temperature. Moreover, present-day laser sources are available that offer noise performance that is limited only by quantum noise. To measure displacement, a commonly used experimental apparatus is a Fabry-Perot interferometer,

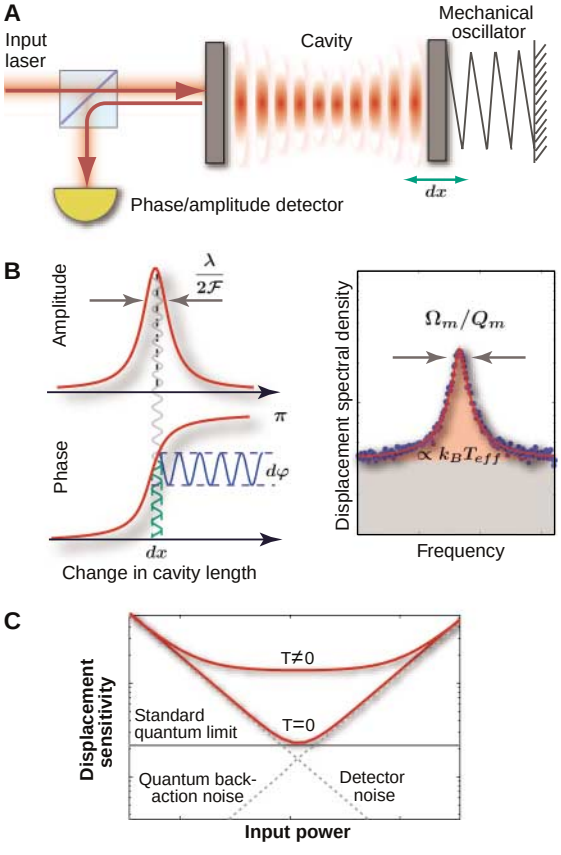


Fig. 1. (**A**) Schematic of the cavity optomechanical interaction of a cavity field (red) and a moveable mirror. (**B**) Transduction mechanism for the laser resonantly probing the cavity. The mechanical motion (green) causes the reflected field to be phase modulated around its steady-state value. This occurs because the mirror motion changes the total cavity length and thereby changes the resonance frequency of the cavity by $\omega_0 \frac{dx}{L}$, where L is the separation between the two mirrors and dx is the mirror displacement. Owing to the high Finesse of the cavity ($\mathcal{F}_\pi$, which describes the number of reflections a photon undergoes on average before escaping the cavity), the conversion of mechanical amplitude to the phase of the field is enhanced (i.e., $d\varphi \approx \frac{\mathcal{F}}{\lambda} \cdot dx$, where $d\varphi$ is the change in the phase of the reflected laser field and λ is the incident wavelength of the laser), allowing minute mirror displacements to be detected. The reflected amplitude is left unchanged. (Right) Fourier analysis of the reflected phase reveals the mechanical spectrum of the mirror motion. Mechanical resonance frequency (Ω_m), quality factor (Q_m), and temperature (T_{eff}) can be determined using this spectrum. (**C**) Sensitivity of the interferometer measurement process for the case of a zero-temperature mechanical oscillator mirror and for finite temperature T. For low-input laser power, detector noise due to the quantum shot noise of the laser field dominates, whereas at higher laser power the quantum fluctuations of the light field cause the mirror to undergo random fluctuations (quantum back-action). At the optimum power, the two sources of fluctuation contribute equally to the measurement imprecision, constituting the SQL. At finite temperature, the mechanical zero-point motion is masked by the presence of thermal noise.

[1]Max Planck Institute für Quantenoptik, 85748 Garching, Germany. [2]Department of Applied Physics, California Institute of Technology, Pasadena, CA 91125, USA.

*To whom correspondence should be addressed. E-mail: tjk@mpq.mpg.de (T.J.K.); vahala@caltech.edu (K.J.V.)
†Present address: Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland.

whose purpose is to determine differential changes in distance between the two end mirrors (Fig. 1A). To account for the mirror suspension or the internal mechanical modes of a mirror, it is assumed that the end mirror is free to oscillate. This harmonic confinement can be either intentional or intrinsic, as we will discuss later. The high-reflectivity end mirrors enhance the number of roundtrips photons undergo (by a factor $\mathcal{F}/\pi$, where $\mathcal{F}$ is the cavity Finesse) and enable very sensitive measurement of the end mirror position (Fig. 1B). For a laser resonant with the cavity, small changes in cavity length shift the cavity resonance frequency and, enhanced by the cavity Finesse, imprint large changes in the reflected phase of the laser field. To date, the best displacement sensitivities attained with optical interferometers [such as those at the Laser Interferometer Gravitational Wave Observatory (LIGO) or Fabry-Perot cavities (*20*)] are already exceeding 10^{-19} m/$\sqrt{\text{Hz}}$, which implies that a displacement equivalent to 1/1000 of the radius of a proton can be measured in 1 s.

This extremely high sensitivity, however, also requires that the disturbances of the measurement process itself must be taken into account. The ultimate sensitivity of an interferometer depends on the back-action that photons exert onto the mechanically compliant mirror, caused by radiation pressure. In terms of mirror-displacement measurement, two fundamental sources of imprecision exist (Fig. 1C). First, there is the detector noise that, for an ideal laser source (emitting a coherent state) and an ideal detector, is given by the random arrival of photons at the detector; i.e., shot noise. The detector signal-to-noise ratio increases with laser power, thereby improving the measurement precision. Increasing power, however, comes at the expense of increased intracavity optical power, causing a back-action onto the mirror. This leads to a second source of imprecision: The resulting random momentum kicks of reflected photons create a mirror-displacement noise. This random force causes the mechanical oscillator to be driven and thus effectively heated. Although this noise can also contain a contribution due to classical sources of noise (excess phase or amplitude noise), it is ultimately, under ideal circumstances, bound by the quantum nature of light and is termed quantum back-action (*21*, *22*). Taking into account both contributions, the optimum sensitivity of an interferometer is achieved at the standard quantum limit (SQL). At the SQL, detector noise and quantum back-action noise contribute each a position uncertainty equal to half of the zero-point motion of the mirror, where the latter is given by $x_0 = \sqrt{\hbar/2m\Omega_m}$ [$\hbar$ is Planck's constant divided by 2π, m is the effective mass (*23*) of the mirror, and Ω_m is the mirror's harmonic frequency]. Much research in the past decade has also focused on ways of circumventing this limit. For example, the use of squeezed light (*24*) can enable surpassing this limit. So far, however, experiments with mechanical mirrors have not observed the radiation-pressure quantum back-action because it is masked by the random, thermal motion of the mirror (Fig. 1C). Fluctuations of the radiation-pressure force have been observed in the field of atomic laser cooling (*25*), where they are responsible for a temperature limit (the Doppler limit).

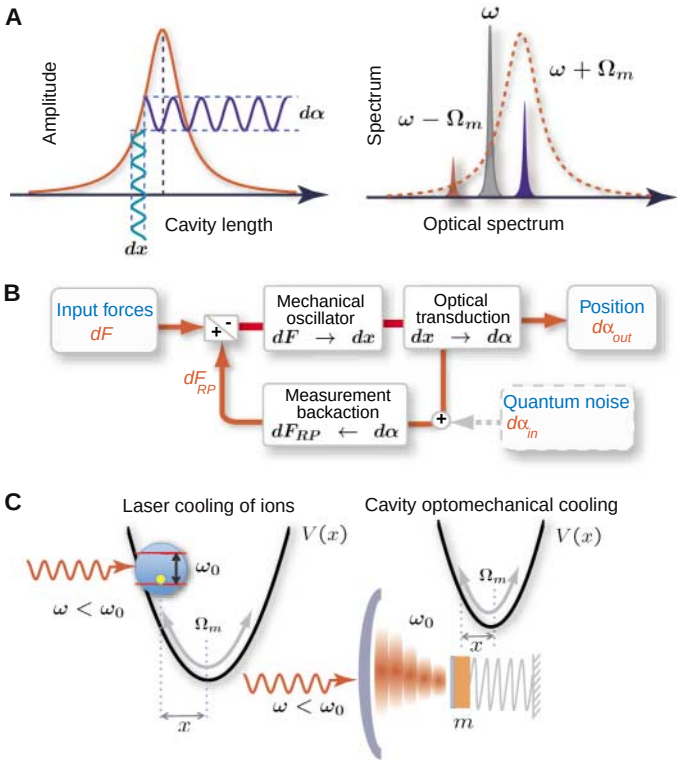


Fig. 2. (**A**) Dynamic back-action results from the coupling of the mechanical motion to the fluctuations of the intracavity field amplitude ($d\alpha$), which occurs when the cavity is excited in a detuned manner. In the frequency domain, the amplitude modulation at Ω_m can be interpreted as sidebands around the optical laser frequency (shown at right). The sideband amplitudes are asymmetric because of the density of states of the cavity. This photon imbalance results in work on the mechanical oscillator (either amplification or cooling), as is further detailed in Fig. 4. (**B**) Basic elements of a feedback loop describing the measurement process and its back-action on the mirror. The mechanical oscillator is subject to a force dF (e.g., because of the thermal force or an externally applied signal force) that induces a mechanical response (dx). The latter causes a change in the optical field (either in amplitude $d\alpha$ or in phase, depending on the detuning), allowing measurement of mechanical position. This transduction is not instantaneous on account of the finite cavity lifetime. For a detuned laser, the amplitude change caused by this measurement process feeds back to the mechanical oscillator through the radiation-pressure force, closing the feedback loop. The sign of the feedback depends on the cavity detuning and can produce either damping (red-detuned pump) or amplification (blue-detuned pump). In a quantum description, this feedback branch is not noiseless but is subjected to quantum noise of the optical field ($d\alpha_{in}$), which yields a random force due to the quantum fluctuations of the field (i.e., the quantum back-action). Although dynamic back-action can be prevented by probing the cavity on resonance (causing $d\alpha = 0$ and thereby preventing feedback), the quantum back-action nevertheless feeds into the mechanical oscillators' input (and thereby reinforces the SQL. $d\alpha_{out}$ are the amplitude fluctuations of the reflected laser field; dF_{RP} are the fluctuations in the radiation pressure force. (**C**) Analogy of dynamical back-action cooling to the laser cooling of harmonically bound ions. In both the case of a harmonically trapped ion and a harmonically oscillating end mirror of a cavity, a dissipative force arises because of the Doppler effect. $V(x)$ denotes the trapping potential of the mirror and ion.

The optical cavity mode not only measures the position of the mechanical mode, but the dynamics of these two modes can also be mutually coupled. This coupling arises when the mechanical motion changes the intracavity field amplitude, which thereby changes the radiation-pressure force experienced by the mirror. For small displacements, this occurs when the laser is detuned with respect to the cavity resonance (Fig. 2A). This mutual coupling of optical and mechanical degrees of freedom can produce an effect called dynamic back-action that arises from the finite cavity delay. This delay leads to a component of the radiation-pressure force that is in quadrature (out of phase) with respect to the mechanical motion. The component is substantial when the cavity photon lifetime is comparable to, or larger than, the mechanical oscillator period and creates an effective mechanical damping of electromagnetic origin. This is the essence of dynamic back-action (*5*), which, like quantum back-action, modifies the motion of the object being measured (the mirror). Unlike quantum back-action, which effectively sets a measurement precision (by causing the mirror to be subjected to a stochastic force resulting from quantum fluctuations of the field), the effect of dynamic back-action is to modify the dynamical behavior of the mirror in a predictable manner. Two consequences of this form of back-action in the context of gravitational wave detection have been identified. With a laser field blue-detuned relative to the optical cavity mode, the mirror motion can be destabilized (*5*) as a result of mechanical amplification (*13*). Similar to the operation of a laser, the onset of this instability occurs when the mechanical gain equals the mechanical loss rate and could thus create an effective limit to boosting detection sen-

sitivity by increasing optical power in interferometers. On the other hand, a red-detuned pump wave can create a radiation component of mechanical damping that leads to cooling of the mechanical mode; i.e., a reduction of the mechanical mode's Brownian motion (*9*, *26*).

One description of this process is given in Fig. 2B, wherein a feedback loop that is inherent to the cavity optomechanical system is described. The elements of this loop include the mechanical and optical oscillators coupled through two distinct paths. Along the upper path, a force acting on the mechanical oscillator (for instance, the thermal Langevin force or a signal force) causes a mechanical displacement, which (for a detuned laser) changes the cavity field due to the optomechanical coupling (the interferometric measurement process). However, the amplitude fluctuations, which contain information on the mirror position, are also coupled back to the mechanical oscillator via radiation pressure (lower path), resulting in a back-action. A blue-detuned pump wave sets up positive feedback (the instability), whereas red detuning introduces negative feedback. Resonant optical probing (where the excitation frequency equals the cavity resonance frequency, $\omega = \omega_0$) interrupts the feedback loop because changes in position only change the phase, not the amplitude, of the field. As described below, this feedback circuit also clarifies the relation between "feedback cooling" and cooling by dynamic back-action.

Experimental Systems

Systems that exhibit radiation-pressure dynamic back-action must address a range of design considerations, including physical size as well as dissipation. Dynamic back-action relies on optical retardation; i.e., is most prominent for photon lifetimes comparable to or exceeding the mechanical oscillation period. Very low optical dissipation also means that photons are recycled many times, thereby enhancing the weak photon pressure on the mirror. On the other hand, the mechanical dissipation rate governs the rate of heating of the mechanical mirror mode by the environment, limiting the effectiveness of optomechanical cooling. It also sets the required amplification level necessary to induce regenerative oscillations. These considerations illustrate the importance of high optical Finesse and mechanical Q in system design.

It is only in the past 3 years that a series of innovative geometries (shown in Fig. 3) has reached a regime where the observation of radiation-pressure dynamic back-action could be observed. These advances have relied on the availability and improvements in high-Finesse mirror coatings (as used in gravity wave detectors) and also on micro- and nanofabrication techniques [which are the underlying enabling technology for nano- and microelectromechanical systems (*27*)]. A commonly used hybrid system consists of a conventional-input mirror made with a high-reflectivity coating and an end mirror whose dimensions are mesoscopic and which is harmonically suspended. This end mirror has been realized in multiple ways, such as from an etched, high-reflectivity mirror substrate (*14*, *15*), a miniaturized and harmonically suspended gram-scale mirror (*28*), or an atomic force cantilever on which a high-reflectivity and micron-sized mirror coating has been transferred (*29*). A natural optomechanical coupling can occur in optical microcavities, such as microtoroidal cavities (*13*) or microspheres, which contain coexisting high-Q, optical whispering gallery modes, and radio-frequency mechanical modes. This coupling can also be optimized for high optical and mechanical Q (*30*). In the case of hybrid systems, yet another approach has separated optical and mechanical degrees of freedom by using a miniature high-Finesse optical cavity and a separate nanometric membrane (*31*). Whereas the aforementioned embodiments have been in the optical domain, devices in the micro- and radiowave domain have also been fabricated (*22*, *32*), such as a nanomechanical resonator coupled to a superconducting microwave resonator (*33*).

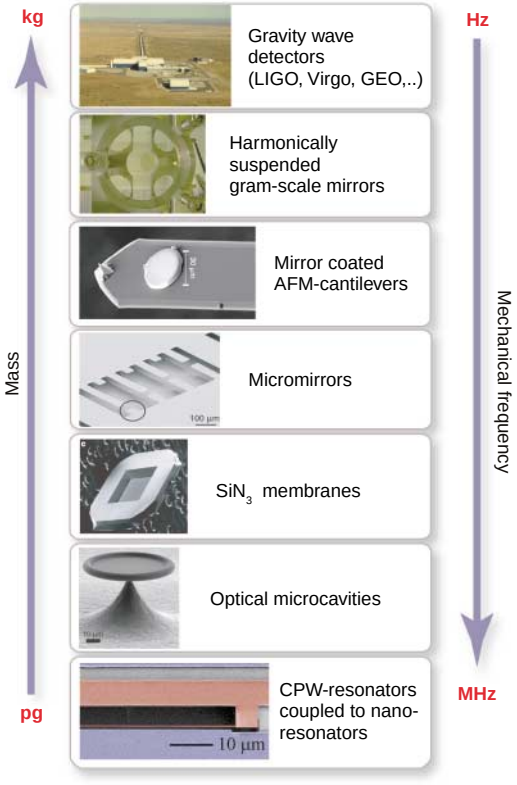


Fig. 3. Experimental cavity optomechanical systems. (**Top** to **Bottom**) Gravitational wave detectors [photo credit LIGO Laboratory], harmonically suspended gram-scale mirrors (*28*), coated atomic force microscopy cantilevers (*29*), coated micromirrors (*14*, *15*), SiN_3 membranes dispersively coupled to an optical cavity (*31*), optical microcavities (*13*, *16*), and superconducting microwave resonators coupled to a nanomechanical beam (*33*). The masses range from kilograms to picograms, whereas frequencies range from tens of megahertz down to the hertz level. CPW, coplanar waveguide.

Many more structures exist that should also realize an optomechanical interaction in an efficient manner. In particular, nanophotonic devices such as photonic crystal membrane cavities or silicon ring resonators might be ideal candidates owing to their small mode volume, high-Finesse, and finite rigidity. Owing to their small length scale, these devices exhibit fundamental flexural frequencies well into the gigahertz regime, but their mechanical quality factors have so far not been studied, nor has optomechanical coupling been observed. As described in the next section, such high frequencies are interesting in the context of regenerative oscillation and ground state cooling.

Cooling and Amplification Using Dynamical Back-Action

The cooling of atoms or ions using radiation pressure has received substantial attention and has been a successful tool in atomic and molecular physics. Dynamical back-action allows laser cooling of mechanical oscillators in a similar manner. The resemblance between atomic laser cooling and the cooling of a mechanical oscillator coupled to an optical (or electronic) resonator is a rigorous one (*34*). In both cases, the motion (of the ion, atom, or mirror) induces a change in the resonance frequency, thereby coupling the motion to the optical (or cavity) resonance (Fig. 2C). Indeed, early work has exploited this coupling to sense the atomic trajectories of single atoms in Fabry-Perot cavities (*35*, *36*) and, more recently, in the context of collective atomic motion (*37*, *38*). This coupling is not only restricted to atoms or cavities but also has been predicted for a variety of other systems. For example, the cooling of a mechanical oscillator can be achieved using coupling to a quantum dot (*39*), a trapped ion (*40*), a Cooper pair box (*41*), an LC circuit (*5*, *32*), or a microwave stripline cavity (*33*). Although the feedback loop of Fig. 2B explains how damping and instability can be introduced into the cavity optomechanical system, the origins of cooling and mechanical amplification are better understood with the use of a motional sideband approach, as described in Fig. 4 (*13*).

Cooling has been first demonstrated for micromechanical oscillators coupled to optical cavities (*14*–*16*) and, using an electromechanical analog, for a Cooper pair box coupled to a nanomechanical beam (*41*). Because the mechanical modes in experiments are high Q (and are thus very well isolated from the reservoir), they are easily resolved in the spectra of detected probe light reflected from the optical cavity (Fig. 1B). Furthermore, their effective temperature can be inferred from the thermal energy $k_B T$ (where k_B is

the Boltzman constant), which is directly proportional to the area of detected mechanical spectral peak (Fig. 1B). In the first back-action cooling experiments, a temperature of ~10 K was achieved for a single mechanical mode. The bath and all other modes in these experiments were at room temperature, owing to the highly targeted nature of cooling (Fig. 4). Since the completion of this work, cooling of a wide variety of experimental embodiments ranging from nanomembranes (*31*) and gram-scale mirrors (*28*) to the modes of kilogram-scale gravitational bar detectors (such as Aurega) has been demonstrated. At this stage, temperatures are rapidly approaching a regime of low phonon number, where quantum effects of the mechanical oscillator become important. To this end, cooling with the use of a combination of conventional cryogenic technology with dynamical back-action cooling is being investigated. Technical hurdles include collateral reheating of the mechanical mode, exacerbated by the very high mechanical Q, which leads to relatively long equilibration times.

Quantum back-action sets a fundamental limit of radiation-pressure cooling (*34*, *42*) that is equivalent to the Doppler temperature in atomic laser cooling (*25*). It may also be viewed as a consequence of the Heisenberg uncertainty relation in that a photon decaying from the resonator has an uncertainty in energy given by $\Delta E = \hbar\kappa$ (where κ is the cavity decay rate), implying that the mechanical oscillator cannot be cooled to a temperature lower than this limit. It has been theoretically shown (*34*, *42*) that ground state cooling is nevertheless possible in the resolved sideband regime (also called the good-cavity limit), in analogy to atomic laser cooling, where this technique has led to ground state cooling of ions (*43*). This regime is characterized by mechanical sidebands that fall well outside the cavity bandwidth and has recently been demonstrated experimentally (*44*). Detection of the ground state could probably prove to be as challenging as its preparation. Proposals to measure the occupancy are diverse, but one method is to measure the weights of the motional sidebands generated by the mechanical motion (*34*).

It is important to note that cooling of mechanical oscillators is also possible using electronic (active) feedback (*10*, *11*, *29*, *45*). This scheme is similar to "stochastic cooling" (*46*) of ions in storage rings and uses a "pick-up" (in the form of an optical cavity interferometer) to measure the mechanical motion and a "kicker" (a radiation-pressure force exerted by a laser on the mirror) to provide a viscous (feedback) force. The idea can also be understood in terms of the feedback loop in Fig. 2B, wherein the lower right optical-feedback branch is replaced by an electrical path driving a second pump laser, which acts as a force actuator on the mirror.

Finally, although originally conceived as a potential limitation in gravitational wave detection, the parametric instability (blue detuned operation of the pump wave) can also be understood as the result of amplification (negative damping) of the mechanical motion (*13*, *17*, *47*). In this sense, the instability is simply the threshold condition in which intrinsic mechanical loss is compensated by amplification. This threshold phenomenon and the subsequent regenerative mechanical oscillation have been studied as a new type of optomechanical oscillator (*48*). Above threshold, the oscillator is regenerative, and oscillation at microwave rates (*49*) has been demonstrated. Additionally, the phase noise of the oscillator has been characterized and observed to obey an inverse power dependence, characteristic of fundamental, Brownian noise (*48*). Quantum back-action is also predicted to set a fundamental low-temperature limit to this linewidth (*50*). The ability to amplify mechanical motion is potentially useful as a means to boost displacements and forces sensitivity (*51*). Finally, returning to the analogy with atomic physics, it is interesting to note that regenerative oscillation (i.e., amplification of mechanical motion) would be expected to occur for trapped ions under blue-detuned excitation.

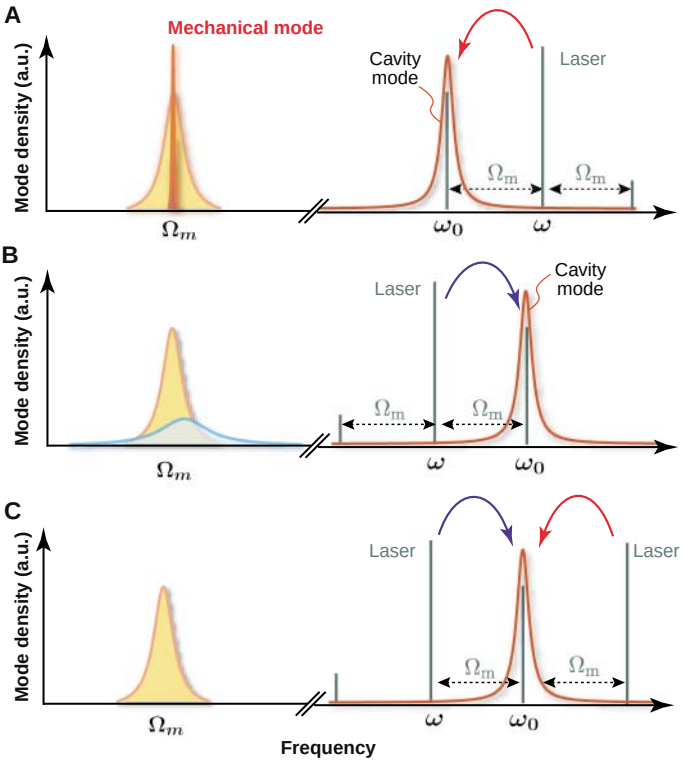


Fig. 4. Frequency domain interpretation of optomechanical interactions in terms of motional sidebands. These sidebands are created on the optical probe wave as photons are Doppler shifted from the mirror surface (which undergoes harmonic motion driven by its thermal energy). Doppler scattering rates into the red (Stokes) and blue (anti-Stokes) sidebands are imbalanced when the probe wave resides to one side of the optical resonance, which can be viewed as a consequence of the asymmetric density of electromagnetic states (Fig. 2A). This imbalance favors the Stokes sideband for a blue-detuned pump and the anti-Stokes sideband for the red-detuned pump, thereby creating a net imbalance in electromagnetic power upon scattering. This imbalance is the origin of mechanical amplification (blue detuning) and cooling (red detuning). (Cooling in this fashion is similar to cavity cooling of atoms.) Only mechanical modes that produce appreciable sideband asymmetry will experience significant gain or cooling. Moreover, the degree of asymmetry can be controlled in an experiment so that a particular mechanical mode can be selected for amplification or cooling. (**A**) Dynamic back-action amplification of mechanical motion via a blue-detuned laser field. The laser scatters pump photons into the cavity, thus creating phonons and leading to amplification. (**B**) Dynamic back-action cooling via a red-detuned laser. Pump photons are scattered into the cavity resonance, thereby removing thermal mechanical quanta from the mechanical oscillator. (**C**) Two-transducer scheme. By symmetrically pumping the cavity on both upper and lower sideband, only one of the quadratures of the mechanical motion is measured with a precision that can exceed the standard limit, thus providing a route to preparing a mechanical oscillator in a squeezed state of mechanical motion via measurement-induced squeezing. a.u., arbitrary units.

Cavity Quantum Optomechanics

A mechanical oscillator has a set of quantum states with energies $E_N = (N + \frac{1}{2})\hbar\Omega_m$, where N is the number of mechanical quanta, and $N = 0$ denotes the quantum ground state. For a mechanical oscillator in the ground state, the ground state energy, $E_0 = \hbar\Omega_m/2$, gives rise to the zero-point motion, characterized by the length scale $x_0 = \sqrt{\hbar/2m\Omega_m}$. As noted earlier, this length scale sets the SQL of mirror position uncertainty in an interferometer such as in Fig. 1. The zero-point motion for structures shown in Fig. 3 ranges from $\sim 10^{-17}$ m for a macroscopic mirror to $\sim 10^{-12}$ m for the nanomechanical beam. Such small motions are masked by the thermal motion of the mechanical oscillator, and to enter the regime where quantum fluctuations become dominant and observable requires that the mechanical mode's temperature satisfy $k_B T << \hbar\Omega_m$, equivalently a thermal occupation less than unity. Over the past decade, cryogenically cooled nanomechanical oscillators coupled to an electronic readout have been steadily approaching the quantum regime (*19*, *52*, *53*). Cavity optomechanical systems exhibit high readout sensitivity, in principle already sufficient to detect the minute zero-point motion of a mesoscopic system. The main challenge toward

observing quantum phenomena in cavity optomechanical systems lies in reducing the mechanical mode thermal occupation. Using conventional cryogenic cooling, the latter is challenging (1 MHz, corresponding to a temperature of only 50 μK). However, in principle, cooling to these temperatures and even lower is possible with the use of optomechanical back-action cooling.

If a sufficiently low occupancy of the mechanical oscillator is reached (using, for instance, a combination of cryogenic precooling and back-action laser cooling), quantum phenomena of a mesoscopic mechanical object may arise. For example, the quantum back-action by photons could become observable (*54*) or signatures of the quantum ground state. Moreover, the interaction of cold mechanics and a light field can give rise to squeezing of the optical field (*55*). This can be understood by noting that the mechanical oscillator couples the amplitude and phase quadrature of the photons. Moreover, the optomechanical coupling Hamiltonian has been predicted to allow quantum nondemolition measurement of the intracavity photon number (*56*, *57*). The coupling afforded by radiation pressure might even allow the production of squeezed states of mechanical motion. These highly nonintuitive quantum states have been produced for electromagnetic fields over the past decades, and producing them in the mechanical realm would be a notable achievement. Such highly nonclassical states may be possible to generate using measurement-induced squeezing. In this method (*22*), one quadrature component of the mechanical oscillator motion is measured (and no information of the complementary variable is gained) so as to project the mechanical oscillator into a squeezed state of motion. This method (Fig. 4C) involves two incident waves and moreover requires that the mechanical frequency exceeds the cavity decay rate (the resolved sideband regime). A great deal of theoretical work has also been devoted to the question of entangling mechanical motion with an electromagnetic field, or even entangling two mechanical modes. Examples include proposals to achieve quantum super-positions of a single photon and a mirror via a "which path" experiment (*58*) or entangling two mirrors via radiation pressure (*59*).

Emerging Cavity Optomechanical Technologies

Cavity optomechanics may also enable advances in several other areas. First, the ability to provide targeted cooling of nano- and micromechanical oscillators (which are otherwise part of devices at room temperature) bodes well for practical applications because, in principle, conventional cryogenics are unnecessary. Beyond providing a better understanding of fluctuation and dissipative mechanisms, the fact that high displacement sensing is an important element of cavity optomechanics will have collateral benefits in other areas of physics and technology, ranging from scanning probe techniques (*60*) to gravitational-wave detection. Moreover, the ability to create all-optical photonic oscillators on a chip with narrow linewidth and at microwave oscillation frequencies may have applications in radio frequency–photonics. Equally important, cavity optomechanical systems already exhibit strong nonlinearity at small driving amplitudes, which offer new functions related to optical mixing (*61*). Finally, although all current interest is focused on radiation-pressure coupling, cavity optomechanical systems based on gradient forces are also possible. Although aimed at a separate set of applications, there has been substantial progress directed toward gradient-force control of mechanical structures using cavity optomechanical effects (*62*–*64*).

Summary

The interaction of mechanical and optical degrees of freedom by radiation pressure is experiencing a paradigm shift in control and measurement of mechanical motion. Radiation-pressure coupling has opened an extremely broad scope of possibilities, both applied and fundamental in nature. With the continued trends toward miniaturization and dissipation reduction, radiation pressure can become an increasingly important phenomenon that will probably allow advances, both in terms of technology as well as in fundamental science. It may well provide a way to probe the quantum regime of mechanical systems and give rise to entirely new ways of controlling mechanics, light, or both. It also seems likely that beyond precision measurement, there will be new technologies that leverage cooling and amplification.

References and Notes

1. A. Ashkin, *Phys. Rev. Lett.* **24**, 156 (1970).
2. T. W. Hänsch, A. L. Schawlow, *Opt. Commun.* **13**, 68 (1975).
3. D. Wineland, H. Dehmelt, *Bull. Am. Phys. Soc.* **20**, 637 (1975).
4. A. Ashkin, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 4853 (1997).
5. V. B. Braginsky, *Measurement of Weak Forces in Physics Experiments* (Univ. of Chicago Press, Chicago, 1977).
6. C. M. Caves, *Phys. Rev. D Part. Fields* **23**, 1693 (1981).
7. A. Dorsel, J. D. McCullen, P. Meystre, E. Vignes, H. Walther, *Phys. Rev. Lett.* **51**, 1550 (1983).
8. B. S. Sheard, M. B. Gray, C. M. Mow-Lowry, D. E. McClelland, S. E. Whitcomb, *Phys. Rev. A* **69**, 051801 (2004).
9. M. I. Dykman, *Sov. Phys. Solid State* **20**, 1306 (1978).
10. S. Mancini, D. Vitali, P. Tombesi, *Phys. Rev. Lett.* **80**, 688 (1998).
11. P. F. Cohadon, A. Heidmann, M. Pinard, *Phys. Rev. Lett.* **83**, 3174 (1999).
12. C. H. Metzger, K. Karrai, *Nature* **432**, 1002 (2004).
13. T. J. Kippenberg, H. Rokhsari, T. Carmon, A. Scherer, K. J. Vahala, *Phys. Rev. Lett.* **95**, 033901 (2005).
14. O. Arcizet, P. F. Cohadon, T. Briant, M. Pinard, A. Heidmann, *Nature* **444**, 71 (2006).
15. S. Gigan *et al.*, *Nature* **444**, 67 (2006).
16. A. Schliesser, P. Del'Haye, N. Nooshi, K. J. Vahala, T. J. Kippenberg, *Phys. Rev. Lett.* **97**, 243905 (2006).
17. T. J. Kippenberg, K. J. Vahala, *Opt. Express* **15**, 17172 (2007).
18. K. Jacobs, I. Tittonen, H. M. Wiseman, S. Schiller, *Phys. Rev. A* **60**, 538 (1999).
19. K. C. Schwab, M. L. Roukes, *Phys. Today* **58**, 36 (2005).
20. O. Arcizet *et al.*, *Phys. Rev. Lett.* **97**, 133601 (2006).
21. C. M. Caves, K. S. Thorne, R. W. P. Drever, V. D. Sandberg, M. Zimmermann, *Rev. Mod. Phys.* **52**, 341 (1980).
22. V. B. Braginsky, F. Y. Khalili, *Quantum Measurement* (Cambridge Univ. Press, Cambridge, 1992).
23. M. Pinard, Y. Hadjar, A. Heidmann, *Eur. Phys. J. D* **7**, 107 (1999).
24. H. J. Kimble, Y. Levin, A. B. Matsko, K. S. Thorne, S. P. Vyatchanin, *Phys. Rev. D Part. Fields* **65**, 022002 (2002).
25. S. Stenholm, *Rev. Mod. Phys.* **58**, 699 (1986).
26. V. B. Braginsky, S. P. Vyatchanin, *Phys. Lett. A* **293**, 228 (2002).
27. H. G. Craighead, *Science* **290**, 1532 (2000).
28. T. Corbitt *et al.*, *Phys. Rev. Lett.* **98**, 150802 (2007).
29. D. Kleckner, D. Bouwmeester, *Nature* **444**, 75 (2006).
30. G. Anetsberger, R. Riviere, A. Schliesser, O. Arcizet, T. J. Kippenberg, *Nat. Photonics*, in press; preprint available at http://arxiv.org/abs/0802.4384 (2008).
31. J. D. Thompson *et al.*, *Nature* **452**, 72 (2008).
32. K. R. Brown *et al.*, *Phys. Rev. Lett.* **99**, 137205 (2007).
33. C. A. Regal, J. D. Teufel, K. W. Lehnert, preprint available at http://arxiv.org/abs/0801.1827 (2008).
34. I. Wilson-Rae, N. Nooshi, W. Zwerger, T. J. Kippenberg, *Phys. Rev. Lett.* **99**, 093902 (2007).
35. C. J. Hood, T. W. Lynn, A. C. Doherty, A. S. Parkins, H. J. Kimble, *Science* **287**, 1447 (2000).
36. P. Maunz *et al.*, *Nature* **428**, 50 (2004).
37. K. W. Murch, K. L. Moore, S. Gupta, D. M. Stamper-Kurn, *Nat. Phys.* **4**, 561 (2008).
38. F. Brennecke, S. Ritter, T. Donner, T. Esslinger, preprint available at http://arxiv.org/abs/0807.2347 (2008).
39. I. Wilson-Rae, P. Zoller, A. Imamoglu, *Phys. Rev. Lett.* **92**, 075507 (2004).
40. L. Tian, P. Zoller, *Phys. Rev. Lett.* **93**, 266403 (2004).
41. A. Naik *et al.*, *Nature* **443**, 193 (2006).
42. F. Marquardt, J. P. Chen, A. A. Clerk, S. M. Girvin, *Phys. Rev. Lett.* **99**, 093902 (2007).
43. D. Leibfried, R. Blatt, C. Monroe, D. Wineland, *Rev. Mod. Phys.* **75**, 281 (2003).
44. A. Schliesser, R. Riviere, G. Anetsberger, O. Arcizet, T. J. Kippenberg, *Nat. Phys.* **4**, 415 (2008).
45. M. Poggio, C. L. Degen, H. J. Mamin, D. Rugar, *Phys. Rev. Lett.* **99**, 017201 (2007).
46. S. Vandermeer, *Rev. Mod. Phys.* **57**, 689 (1985).
47. H. Rokhsari, T. J. Kippenberg, T. Carmon, K. J. Vahala, *Opt. Express* **13**, 5293 (2005).
48. M. Hossein-Zadeh, H. Rokhsari, A. Hajimiri, K. J. Vahala, *Phys. Rev. A* **74**, 023813 (2006).
49. T. Carmon, K. J. Vahala, *Phys. Rev. Lett.* **98**, 123901 (2007).
50. K. J. Vahala, *Phys. Rev. A* **78**, 2 (2008).
51. O. Arcizet, T. Briant, A. Heidmann, M. Pinard, *Phys. Rev. A* **73**, 033819 (2006).
52. R. G. Knobel, A. N. Cleland, *Nature* **424**, 291 (2003).
53. M. D. LaHaye, O. Buu, B. Camarota, K. C. Schwab, *Science* **304**, 74 (2004).
54. I. Tittonen *et al.*, *Phys. Rev. A* **59**, 1038 (1999).
55. C. Fabre *et al.*, *Phys. Rev. A* **49**, 1337 (1994).
56. P. Alsing, G. J. Milburn, D. F. Walls, *Phys. Rev. A* **37**, 2970 (1988).
57. A. Heidmann, Y. Hadjar, M. Pinard, *Appl. Phys. B* **64**, 173 (1997).
58. W. Marshall, C. Simon, R. Penrose, D. Bouwmeester, *Phys. Rev. Lett.* **91**, 173901 (2003).
59. S. Mancini, V. Giovannetti, D. Vitali, P. Tombesi, *Phys. Rev. Lett.* **88**, 120401 (2002).
60. D. Rugar, R. Budakian, H. J. Mamin, B. W. Chui, *Nature* **430**, 329 (2004).
61. M. Hossein-Zadeh, K. J. Vahala, *Photonics Technol. Lett.* **20**, 4 (2007).
62. M. L. Povinelli *et al.*, *Opt. Express* **13**, 8286 (2005).
63. P. T. Rakich, M. A. Popovic, M. Soljacic, E. P. Ippen, *Nat. Photonics* **1**, 658 (2007).
64. M. Eichenfield, C. P. Michael, R. Perahia, O. Painter, *Nat. Photonics* **1**, 416 (2007).

65. K.J.V. acknowledges support from the Alexander von Humboldt Foundation. T.J.K. gratefully acknowledges support via a Max Planck Independent Junior Research Group, a Marie Curie Excellence Grant, the Deutsche Forschungsgemeinschaft (DFG project Ground State Cooling), and Nanosystems Initiative Munich. The authors thank T. W. Hänsch for discussions.

10.1126/science.1156032

Claus Kiefer

Statement

and

Readings

DECOHERENCE IN COSMOLOGY

Claus Kiefer

Assuming the universal validity of quantum theory, the quantum-to-classical transition is also of crucial importance in cosmology. Firstly, any linear theory of quantum gravity predicts superpositions of different metrics even at the macroscopic level. Secondly, primordial fluctuations in the early Universe, out of which galaxies and clusters of galaxies are expected to develop, are of a genuine quantum nature. In my talk, I shall discuss both cases and show how and to which extent classical behaviour emerges through decoherence. The emphasis is on the main conceptual aspects rather than on technical issues.

4 Decoherence in Quantum Field Theory and Quantum Gravity

C. Kiefer

...

4.2 Decoherence and the gravitational field[1]

4.2.1 Emergence of classical spacetime

According to the Copenhagen interpretation of quantum theory, the existence of a classical world is needed from the outset in order to interpret quantum theory. Appropriate classical apparata are assumed to *define* the occurrence of quantum phenomena. The presence of such classical measurement agencies seems to be possible only if spacetime exists as a classical entity.

The discussion of the previous chapters has, however, convincingly demonstrated that quantum theory has a much wider range of applicability than the pioneers had imagined. Classical properties are not intrinsic to objects but emerge through the irreversible interaction with the environment. The experiments discussed in Chap. 3 are an impressive confirmation of this idea.

What about the structure of spacetime itself? Before the advent of the general theory of relativity, spacetime was considered to be a given, non-dynamical background structure. This is also the case in quantum field theories such as QED (Sect. 4.1). In general relativity, however, the geometry of spacetime is associated with the gravitational field and thereby becomes *dynamical*. If the gravitational field is fundamentally described by quantum theory, then spacetime cannot be a classical entity.

But has gravity to be described by quantum theory? Quite generally, it does not seem possible to find a fundamental hybrid description that couples a quantum system to a classical system in a consistent way (Kiefer 2003). This does of course not mean that there exists no *effective* theory which couples quantum to classical systems. For example, one can develop a formalism in which a decohered ("classical") system is coupled to a quantum system that does not exhibit decoherence (Halliwell 1998).

As has already been mentioned, it was important already during the early discussions between Einstein and Bohr to apply the uncertainty relations to macroscopic objects (screens, photographic plates etc.) in order to save them

[1] Extract from Chapter 4 of *Decoherence and the Appearance of a Classical World in Quantum Theory*, by E. Joos, H. D. Zeh, C. Kiefer, D. Giulini, J. Kupsch, and I.-O. Stamatescu (Springer, Berlin, 2003).

for microscopic systems. This is reasonable because macroscopic objects are composed of atoms. Such *consistency arguments* are at the heart of these discussions. At the Solvay conference in 1930, Bohr and Einstein had a debate concerning the time-energy uncertainty relation, $\Delta E \Delta t \geq \hbar/2$. In the discussion, Bohr had to invoke general relativity to counter Einstein's objections. But only very little structure from general relativity does in fact enter the argument; it is only the equivalence principle and therefore the curved nature of spacetime, from which the redshift of light follows as a consequence. The redshift may be derived by just applying the energy law to the expression $\hbar\omega$ for the energy of a photon. One could thus phrase Bohr's argument in the way that a violation of the uncertainty relation would entail a violation of energy conservation.

In fact, the possible violation of conservation laws often plays an important role in such consistency arguments. Eppley and Hannah (1977), for example, consider the interaction of classical gravitational waves with quantum systems. They find, as a consequence, a violation of either momentum conservation or the uncertainty relations for the quantum system, or the occurrence of signals faster than light. Since not many peculiarities of the gravitational field enter their discussion, these results hold also for other systems such as the electromagnetic field. This type of arguments is certainly enforced for the gravitational field due to its coupling to *all* other degrees of freedom. Taking then the quantum nature of the gravitational field for granted, one would expect that efficient decoherence results from this universal coupling for both the gravitational field and other variables.

In a heuristic example, where quantum theory is applied to Newtonian gravity, one finds that the gravitational field is decohered by its action with quantum matter (Joos 1986b). Suppose that a (homogeneous) gravitational field within a box of side length L is in a quantum superposition of different strengths, i.e.

$$|\psi\rangle = c_1|g\rangle + c_2|g'\rangle, \; g \neq g'. \tag{4.1}$$

A particle with mass m in a state $|\chi\rangle$, which moves through this volume, "measures" the value of g, since its trajectory depends on the metric, yielding the total state

$$|g\rangle|\chi_g(t)\rangle \; . \tag{4.2}$$

This correlation destroys the coherence between g and g', and the reduced density matrix can be estimated to assume the following form after many such interactions are taken into account:

$$\rho(g, g', t) = \rho(g, g', 0) \exp\left(-\Gamma(g - g')^2 t\right) \; , \tag{4.3}$$

where

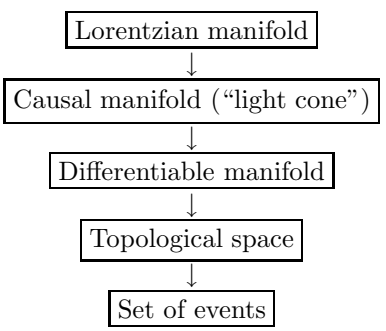
$$\Gamma = nL^4 \left(\frac{\pi m}{2k_B T}\right)^{3/2} \; ,$$

for a gas with particle density n and temperature T. For example, air under ordinary conditions, and $L = 1$ cm, $t = 1$ s yields a remaining coherence width of $\Delta g/g \approx 10^{-6}$.

One can give quite general arguments that the gravitational field is fundamentally of quantum nature (Kiefer 2000, 2003):

- *Singularity theorems of general relativity*: Under very general conditions, the occurrence of a singularity, and therefore the breakdown of the unquantised theory, is unavoidable. A more fundamental theory is therefore needed to overcome this breakdown, and the natural expectation is that this fundamental theory is a quantum theory of gravity. This is similar to ordinary quantum theory preventing the singularity that classical electromagnetism would predict for atoms.
- *Initial conditions in cosmology*: This is related to the singularity theorems which predict the existence of a "big bang" where the known laws of physics break down. To fully understand the evolution of our Universe, its initial state must be amenable to a physical description.
- *Unification*: Apart from general relativity, all known fundamental theories are *quantum* theories. It would thus seem awkward if gravity, which couples to all other fields, should remain the only classical entity in a fundamental description.
- *Gravity as a regulator*: Many models indicate that the consistent inclusion of gravity in a quantum framework would automatically eliminate the divergences that plague ordinary quantum field theory.
- *Problem of time*: In ordinary quantum theory, the presence of an external time parameter t is crucial for the interpretation of the theory: "Measurements" take place at a certain time, matrix elements are evaluated at fixed times, and the norm of the wave function is conserved *in* time. Since in general relativity, on the other hand, time as part of spacetime is a dynamical quantity (as defined by the metric), both concepts of time must be modified at a fundamental level.

But what does the "quantisation" of spacetime mean? In other words, to which classical structures does one have to apply the superposition principle, while the rest remains classical? Isham (1994) presents the following hierarchy of structures where this decision can be made at each level:



A straightforward quantisation of general relativity, for example, would dissolve spacetime as a fundamental classical entity, but would retain a fixed three-dimensional manifold in the formalism. This *canonical approach* is briefly described in the next subsection and will be the basis for the calculations presented below. Path integration, for example, would entail a superposition of different manifolds. This should also be true in a "theory of everything" (for which superstring theory is a candidate) which encompasses all interactions of Nature in a single quantum framework. In such a fundamental theory it is probably only very little structure, if any, that remains classical, although this is not yet clear, cf. Seiberg and Witten (1999).

Quantum effects of gravity are expected to become relevant at the Planck scale. This is the scale where, for an elementary particle, Schwarzschild radius and Compton wavelength coincide. The Planck mass is given by

$$m_P = \sqrt{\frac{\hbar c}{G}} \approx 10^{-5}\ \mathrm{g}\ , \tag{4.4}$$

while Planck length and Planck time are given by the following expressions, respectively,

$$l_P = \sqrt{\frac{\hbar G}{c^3}} \approx 10^{-33}\ \mathrm{cm}\ , \quad t_P = \sqrt{\frac{\hbar G}{c^5}} \approx 10^{-44}\ \mathrm{s}\ . \tag{4.5}$$

As we discuss at length in this volume, quantum effects are not a priori restricted to a particular scale. In Chap. 3 we have demonstrated that it is not the large mass by itself that provokes classical behaviour for a quantum object, but its interaction (whose strength of course depends on the mass) with the environment. Analogously, it is not the smallness of the Planck length by itself that a priori prevents quantum-gravity effects to occur at larger scales. The classical *appearance* of spacetime at larger scales should again be due to the unvoidable interaction with other degrees of freedom. It is for this reason that we can restrict ourselves in the following discussion to canonical quantum gravity, since this should be valid as an *effective* theory for scales $l \gg l_P$, independent of whether this theory is also valid at the Planck scale itself or not (in the latter case a unified theory such as string theory must be invoked).

We mention that gravity is assigned a fundamental role also in approaches which *modify* the formalism of quantum theory, see e.g. Károlyházy *et al.* (1986), Penrose (1986), as well as Chap. 8, but this will not be considered in this chapter.

4.2.2 The formalism of quantum cosmology

The basic intention in the canonical approach to quantum gravity is to derive equations for wave functionals on an appropriate configuration space, analogously to the Schrödinger picture in quantum mechanics. Technically, this is

achieved by foliating, in the classical theory, the classical spacetime into spatial hypersurfaces and choosing the *spatial* metric as a canonical variable (the "q"). In the spacetime which is classically constructed by dynamically developing the initial data on a particular hypersurface, the canonical momentum is linearly related to the embedding of the hypersurfaces into spacetime. (In the case of a Friedmann universe, the radius, a, is the configuration variable, while the canonical momentum corresponds to the Hubble parameter.) The postulate of nontrivial commutation relations between these quantities in quantum gravity then means that spacetime is no longer a fundamental concept, since one cannot specify both the spatial metric and the embedding. The role of spacetime is taken over by the space of all three-dimensional geometries, which is called *superspace* and which serves as the configuration space for the theory. For a detailed physical introduction into these concepts we refer to Zeh (2001); the details of the canonical formalism are presented, for example, by Wald (1984). The central kinematical quantity is thus a wave functional defined on superspace and on matter field degrees of freedom. It is often labeled $\Psi[{}^3\mathcal{G}, \Phi]$, where ${}^3\mathcal{G}$ stands for "three-dimensional geometry" (to express the fact that this wave functional is independent of particular coordinates on the three-dimensional space, as being guaranteed by the three "momentum constraints" of general relativity), and Φ symbolically denotes all non-gravitational fields. The invariance of general relativity (called invariance under coordinate transformations or under diffeomorphisms) leads to the presence of constraints: the total Hamiltonian must vanish.[2] In the quantum theory, the constraints are implemented à la Dirac as restrictions on physically allowed wave functionals. The wave functional then obeys the Wheeler-DeWitt equation (DeWitt 1967, Wheeler 1968),

$$H\Psi = 0, \tag{4.6}$$

where H denotes the full Hamiltonian for gravity and other fields. In classical general relativity, spacetimes can be parametrised by some arbitrary time coordinates (which have lost their absolute status). Since due to the uncertainty relations no spacetimes exist anymore on the level of quantum gravity (only a wave function for spatial metrics), there is no time parameter available to parametrise them – the Wheeler-DeWitt equation is "timeless". This gives rise to the *problem of time* in quantum gravity which is extensively discussed in the literature, see e.g. Barbour (1994a,b), Isham (1992), Kuchař (1992), Zeh (1986, 2001), Kiefer and Zeh (1995), and Kiefer (2000, 2003).

We have to emphasise that this approach at present exists only on a formal level, since the explicit treatment of (4.6) is unclear.[3] In this respect the discussion in the present section is different from the rest of the book and

[2] We consider only the case of spatially closed hypersurfaces. In the asymptotically flat case, the total Hamiltonian can be written as a surface integral.

[3] It is known that (4.6) does not give rise to a unitary evolution in a Fock space built over three-dimensional slices.

should be considered of heuristic value only. However, from general arguments like reparametrisation invariance one would expect the fundamental equation to be of the constraint form (4.6), although the exact form of H may be different. Therefore, the main *interpretational* part of the discussion in this section would remain unaffected, and only the details of the calculations would have to be changed.

The main features of the canonical approach can already be recognised in a simple two-dimensional model – a closed Friedmann universe characterised by its scale factor a, containing a homogeneous massive scalar field φ as a representation of matter, cf. Kiefer (1988) and Halliwell (1991). Taking the units $2G = 3\pi$, the classical action for this model is the sum of the gravitational part and the matter part,

$$\begin{aligned} S &= \int dt\ L(a, \dot{a}, \varphi, \dot{\varphi}, N) \\ &\equiv \frac{1}{2} \int dt\ N a^3 \left(-\frac{\dot{a}^2}{N^2 a^2} + \frac{\dot{\varphi}^2}{N^2} + \frac{1}{a^2} - m^2 \varphi^2 \right) . \end{aligned} \tag{4.7}$$

This action is invariant with respect to arbitrary reparametrisations of the time variable t, a fact which is encoded in the presence of the non-dynamical *lapse function* N which appears undifferentiated in the action. A characteristic feature of the gravitational field is the occurrence of an indefinite kinetic term in the action.

The standard canonical formalism proceeds with the definition of the canonical momenta,

$$p_N = \frac{\partial L}{\partial \dot{N}} = 0,\ p_a = \frac{\partial L}{\partial \dot{a}} = -\frac{a\dot{a}}{N},\ p_\varphi = \frac{\partial L}{\partial \dot{\varphi}} = \frac{a^3 \dot{\varphi}}{N} . \tag{4.8}$$

The canonical Hamiltonian is then given by

$$\begin{aligned} H &= p_N \dot{N} + p_a \dot{a} + p_\varphi \dot{\varphi} - L \\ &= \frac{N}{2} \left(-\frac{p_a^2}{a} + \frac{p_\varphi^2}{a^3} - a + m^2 \varphi^2 a^3 \right) \\ &\equiv \frac{N}{2} G^{AB} p_A p_B + V(a, \varphi) . \end{aligned} \tag{4.9}$$

The important point is that $p_N = 0$ is a constraint that should hold at all times. Therefore, from Hamiton's equations of motion one gets $\partial H / \partial N = 0$ which gives the constraint

$$H = 0 \Leftrightarrow \dot{a}^2 = -1 + a^2 (\dot{\varphi}^2 + m^2 \varphi^2) . \tag{4.10}$$

This is nothing but the classical Friedmann equation which is well known from cosmology. Variation of (4.7) with respect to a and φ give the classical equations of motion. The equation for φ, in particular, reads

$$\ddot{\varphi} + \frac{3\dot{a}}{a} \dot{\varphi} + m^2 \varphi = 0 . \tag{4.11}$$

This is the Klein-Gordon equation for a homogeneous field in an evolving universe, whose effect on φ is the second ("friction") term.

Following Dirac's procedure, the classical constraint (4.10) is then turned into the Wheeler-DeWitt equation (4.6). Using a particular factor ordering,[4] the explicit form of this equation in the present model reads

$$H\psi \equiv \left(\hbar^2 a \frac{\partial}{\partial a} \left(a \frac{\partial}{\partial a} \right) - \hbar^2 \frac{\partial^2}{\partial \varphi^2} + m^2 \varphi^2 a^6 - a^4 \right) \psi(a, \varphi) = 0 \ . \qquad (4.12)$$

Note that the indefiniteness of the kinetic term has led to a *hyperbolic* equation for ψ – in contrast to the Schrödinger equation. In the next subsection, a more complicated model is used in which the variables a and φ play the role of the background, supplemented by additional degrees of freedom ("higher multipoles") $\{f_n\}$. (In the following we set again $\hbar = 1$.)

The Wheeler-DeWitt equation (4.6), (4.12) does not contain a classical time parameter. This is not surprising, since the classical metric is known to determine time. An *approximate* concept of time-dependence of a wave function can be recovered in a Born-Oppenheimer type of approximation scheme in which part of the degrees of freedom are semiclassical (given by WKB wave functions), while the rest is fully quantum. This limit is obtained, for example, if the full wave functional in (4.6) is of the form

$$\Psi[{}^3\mathcal{G}, \Phi] \approx \sum_{(n)} C_{(n)}[{}^3\mathcal{G}] e^{iS_0^{(n)}[{}^3\mathcal{G}]/G\hbar} \psi^{(n)}[{}^3\mathcal{G}, \Phi] \equiv \sum_{(n)} \psi_0^{(n)} \psi^{(n)}, \qquad (4.13)$$

where the $S_0^{(n)}$ are solutions to the gravitational Hamilton–Jacobi equations which are fully equivalent to Einstein's field equations. The gravitational part of the total state is thus treated semiclassically. The semiclassical part may also comprise part of the matter degrees of freedom. In fact, in the discussion of decoherence in Sect. 4.2.3, the scalar field φ will belong to this part. Note the analogy to Equation (**??**) discussed in the last section.

The sum in (4.13) runs over a whole set of indices (n) (which may also be continuous). It turns out that the matter states $\psi^{(n)}$ obey the following approximate equation in each component,

$$i\nabla S_0^{(n)} \cdot \nabla \psi^{(n)} \approx H_m^{(n)} \psi^{(n)}, \qquad (4.14)$$

where $H_m^{(n)}$ denotes the Hamiltonian for the non-gravitational fields (which of course depends on the particular solution $S_0^{(n)}$ chosen for the gravitational field). Note the analogy of $H_m^{(n)}$ to H_ϕ discussed in the last section. The expression $\nabla S_0^{(n)} \cdot \nabla \equiv \partial / \partial t^{(n)}$ is a directional derivative in the gravitational part of the full configuration space, which parametrises the family of

[4] The chosen factor ordering is given by the Laplace-Beltrami operator in the configuration space spanned by a and φ.

classical spacetimes described by $S_0^{(n)}$. The parameters $t^{(n)}$ are often called *WKB times* – they control the "dynamical evolution" of the states $\chi^{(n)}$ along the WKB trajectories. Equation (4.14) is thus nothing but the Schrödinger equation, while t represents our phenomenological time. Details of the semiclassical approximation to quantum gravity are described in Kiefer (1994), see also Kiefer (1993), Giulini and Kiefer (1995).

We note that due to the central input of the Born-Oppenheimer expansion the situation here is analogous to that of Sect. 4.1.2 only ("measurement" of fields by charges), since the reverse effect (which would here correspond to "measurement" of matter by the gravitational field) is too weak to become important.

4.2.3 Decoherence in quantum cosmology

In quantum cosmology, all variables are fundamentally quantum and there is no classical spacetime. How does a classical spacetime emerge? It has been suggested that global degrees of freedom such as the volume of the Universe appear classical after the interaction with other degrees of freedom is taken into account (Zeh 1986). The role of such additional variables may be played by density fluctuations and gravitational waves. All these degrees of freedom are of course within the Universe, but they are "environmental" to the volume-degree of freedom in configuration space. From the viewpoint of a "local" observer who can measure the size of the Universe but has no access to small fluctuations, these other degrees of freedom have to be traced over. In this sense they are able to produce decoherence for the volume degree of freedom. We have emphasised before that the issue of classicality only arises *after* a quantum system has been chosen, for which the straightforward application of the superposition principle would lead to a macroscopically entangled state. In a sense, a classical spacetime thus arises by a "self-measurement" of the Universe.

Calculations for decoherence in quantum cosmology can be done with the help of the Wheeler-DeWitt equation (4.6), see Kiefer (1987). As a necessary prerequisite, the semiclassical approximation to quantum gravity is employed, in which an approximate Schrödinger equation is recovered for the cosmological fluctuations (see Sect. 4.2.2). The time parameter corresponding to this equation is defined by the semiclassical degrees of freedom (Halliwell and Hawking 1985). In Kiefer (1987) the relevant system was taken to be the scale factor ("radius") a of the Universe together with a homogeneous scalar field φ, cf. the model discussed in Sect. 4.2.2. The field φ plays a crucial role in modern cosmological theories where an exponential, "inflationary", expansion is assumed to have happened in an early phase of the Universe, starting about 10^{-33} s after the big bang. It is in fact the "inflaton field" φ itself that causes inflation. The inhomogeneous modes of the gravitational field and the scalar field (gravitational waves and density fluctuations) can then be shown to *decohere* the global variables a and φ.

An open problem in these early papers was the issue of regularisation; the number of fluctuations is infinite and would cause divergences, which is why a cutoff was suggested. The issue was again addressed in Barvinsky *et al.* (1999a) where a physically motivated regularisation scheme was introduced. In the following we shall briefly review this approach.

As a (semi)classical solution for a and φ one may use

$$\varphi(t) \approx \varphi, \tag{4.15}$$

$$a(t) \approx \frac{1}{H(\varphi)} \cosh H(\varphi) t \ , \tag{4.16}$$

where $H(\varphi) = 8\pi V(\varphi)/3m_P^2$ is the Hubble parameter generated by the inflaton potential $V(\varphi)$. It is approximately constant during the inflationary phase in which φ slowly "rolls down" the potential. We take into account fluctuations of a field $f(t, \mathbf{x})$ which can be a field of any spin. Space is assumed to be a closed three-sphere, so $f(t, \mathbf{x})$ can be expanded into a discrete series of spatial orthonormal harmonics $Q_n(\mathbf{x})$,

$$f(t, \mathbf{x}) = \sum_n f_n(t) Q_n(\mathbf{x}) \ . \tag{4.17}$$

One can thus represent the fluctuations by the degrees of freedom f_n.

Our intention now is to solve the Wheeler-DeWitt equation (4.6) in the semiclassical approximation. This leads to the following solution:

$$\Psi(t|\varphi, f) = \frac{1}{\sqrt{v_\varphi^*(t)}} \, e^{-I(\varphi)/2 + iS_{\rm cl}(t,\varphi)} \prod_n \psi_n(t, \varphi|f_n) \ . \tag{4.18}$$

The time t that appears here is the semiclassical ("WKB") time and is defined by the background-degrees of freedom a and φ through the "eikonal" $S_{\rm cl}$ which is a solution of the Hamilton-Jacobi equation; t is formally identical with the time that appears in the classical equations (4.15) and (4.16). Since φ is thus determined by a, only one variable (a *or* φ) occurs in the argument of Ψ. The wave functions ψ_n for the fluctuations f_n obey each an approximate Schrödinger equation (4.14) with respect to t, and their Hamiltonian H_n has the form of a ("time-dependent") harmonic-oscillator Hamiltonian. The first exponent contains the euclidean action $I(\varphi)$ from the classically forbidden region (the "De Sitter instanton") and is independent of t. Its form depends on the *boundary conditions* imposed. In the present case the so-called Hartle-Hawking condition is chosen, see e.g. Halliwell (1991), which amounts to $I(\varphi) \approx -3m_P^4/8V(\varphi)$. The detailed form is, however, not necessary for the discussion below. The function $v_\varphi(t)$ is the so-called basis function for φ and is a solution of the classical equation of motion. In the following we shall choose units such that $G = c = \hbar = 1$.

For the ψ_n we shall take – in analogy to (**??**) – Gaussian states that correspond to the so-called De Sitter-invariant vacuum state (Starobinsky 1979,

Allen 1985). This is the maximally symmetric state which possesses properties very similar to the standard vacuum state in Minkowski space. (In the massless case, this state is invariant only under a subgroup of the De Sitter group.) It is given by

$$\psi_n(t,\varphi|f_n) = \frac{1}{\sqrt{v_n^*(t)}} \exp\left(-\frac{1}{2}\Omega_n(t) f_n^2\right), \tag{4.19}$$

$$\Omega_n(t) = -ia^3(t)\frac{\dot{v}_n^*(t)}{v_n^*(t)} \ . \tag{4.20}$$

The functions v_n are the basis functions of the De Sitter-invariant vacuum state; they satisfy the classical equation of motion

$$F_n\left(\frac{d}{dt}\right) v_n \equiv \left(\frac{d}{dt}a^3\frac{d}{dt} + a^3m^2 + a(n^2-1)\right) v_n = 0 \tag{4.21}$$

with the boundary condition that they should correspond to a standard Minkowski positive-frequency function for constant a. In the simple special case of a spatially flat section of De Sitter space one would have

$$av_n = \frac{e^{-in\eta}}{\sqrt{2n}}\left(1 - \frac{i}{n\eta}\right) \ , \tag{4.22}$$

where η is the *conformal time* defined by $ad\eta = dt$. We note that the corresponding *negative*-frequency function enters the exponent of the Gaussian, see (4.20).

An important property of these vacuum states is that their norm is conserved *along any semiclassical solution* (4.15), (4.16),

$$\left\langle \psi_n | \psi_n \right\rangle \equiv \int df_n |\psi_n(f_n)|^2 = \sqrt{2\pi}[\Delta_n(\varphi)]^{-1/2}, \tag{4.23}$$

$$\Delta_n(\varphi) \equiv ia^3(v_n^*\dot{v}_n - \dot{v}_n^* v_n) = \text{constant} \ . \tag{4.24}$$

Note that $\Delta_n(\varphi)$ is just the (constant) Wronskian corresponding to (4.21). (The corresponding Wronskian for the homogeneous mode φ is $\Delta_\varphi \equiv ia^3(v_\varphi^*\dot{v_\varphi} - \dot{v_\varphi^*}v_\varphi)$.) We must emphasise that Δ_n is a nontrivial function of the background variable φ, since it is defined on full configuration space and not only along semiclassical trajectories (it gives the weights in the "Everett branches".) It is therefore *not* possible to normalise the ψ_n artificially to one, since this would be inconsistent with respect to the full Wheeler-DeWheeler equation (Barvinsky *et al.* 1999a).

The solution (4.18) forms the basis for our discussion of decoherence. Since the $\{f_n\}$ are interpreted as the environmental degrees of freedom, they have to be integrated out to get the reduced density matrix for φ or a (a and φ can be used interchangeably, since they are connected by t). The reduced density matrix thus reads

$$\rho(t|\varphi,\varphi') = \int df \Psi(t|\varphi,f)\Psi^*(t|\varphi',f) \ , \tag{4.25}$$

where Ψ is given by (4.18), and it is understood that $df = \prod_n df_n$. After the integration one finds

$$\rho(t|\varphi,\varphi') = C\frac{1}{\sqrt{v^*_\varphi(t)v'_\varphi(t)}}\exp\Big[-\frac{1}{2}I-\frac{1}{2}I'+i(S-S')\Big] \times\prod_n\Big[v^*_n v'_n(\Omega_n+\Omega'^*_n)\Big]^{-1/2}, \tag{4.26}$$

where C is a numerical constant. The diagonal elements $\rho(t|\varphi,\varphi)$ describe the probabilities for certain values of the inflaton field to occur. In an appropriate model, one can find that these probabilities are peaked at the onset of inflation around values of φ that lead to phenomenologically satisfying results (for example, with respect to structure formation) without having to invoke the anthropic principle, see Barvinsky *et al.* (1999b) and the references therein.

It is convenient to rewrite the expression for the density matrix (4.26) in the form

$$\rho(t|\varphi,\varphi') = C\frac{\Delta_\varphi^{1/4}\Delta_\varphi'^{1/4}}{\sqrt{v^*_\varphi(t)v'_\varphi(t)}}\exp\left(-\frac{1}{2}\boldsymbol{\Gamma}-\frac{1}{2}\boldsymbol{\Gamma}'+i(S-S')\right) \times\boldsymbol{D}(t|\varphi,\varphi'), \tag{4.27}$$

where

$$\boldsymbol{\Gamma} = I(\varphi)+\boldsymbol{\Gamma}_{1-\mathrm{loop}}(\varphi) \tag{4.28}$$

is the full Euclidean effective action including the classical part and the one-loop part. The latter comes from the next-order WKB approximation and is important for the normalisability of the wave function with respect to φ. The last factor in (4.27) is the *decoherence factor*

$$\boldsymbol{D}(t|\varphi,\varphi') = \prod_n\left(\frac{4\mathrm{Re}\,\Omega_n\,\mathrm{Re}\,\Omega'^*_n}{(\Omega_n+\Omega'^*_n)^2}\right)^{1/4}\left(\frac{v_n}{v^*_n}\frac{v'^*_n}{v'_n}\right)^{1/4}. \tag{4.29}$$

It is equal to one for coinciding arguments. While the decoherence factor is time-dependent, the one-loop contribution to (4.27) does not depend on time and may play only a role at the onset of inflation. In a particular model with non-minimal coupling (Barvinsky *et al.* 1997), the size of the non-diagonal elements is at the onset of inflation approximately equal to those of the diagonal elements. The Universe would thus be essentially quantum at this stage, i.e. in a non-classical state.

The amplitude of the decoherence factor can be rewritten in the form

$$|\boldsymbol{D}(t|\varphi,\varphi')| = \exp\frac{1}{4}\sum_n\ln\frac{4\mathrm{Re}\,\Omega_n\,\mathrm{Re}\,\Omega'^*_n}{|\Omega_n+\Omega'^*_n|^2}. \tag{4.30}$$

The convergence of this series is far from being guaranteed. Moreover, the divergences might not be renormalisable by local counterterms in the bare quantised action. We shall now analyse this question in more detail.

We start with a minimally coupled massive scalar field. Equation (4.21) for the basis functions reads

$$\frac{d}{dt}\left(a^3\frac{dv_n}{dt}\right)+a^3\left(\frac{n^2-1}{a^2}+m^2\right)v_n=0\ . \tag{4.31}$$

The appropriate solution to this equation is (Barvinsky *et al.* 1992)

$$v_n(t)=(\cosh Ht)^{-1}P^{-n}_{-\frac{1}{2}+i\sqrt{m^2/H^2-9/4}}(i\sinh Ht)\ , \tag{4.32}$$

where P denotes an associated Legendre function of the first kind. The expansion of (4.32) for large masses was derived in Barvinsky *et al.* (1992). The corresponding expression for (4.20) is given by

$$\Omega_n=a^2\left[\sqrt{n^2+m^2a^2}+i\sinh Ht\left(1+\frac{1}{2}\frac{m^2a^2}{n^2+m^2a^2}\right)\right]+O\left(\frac{1}{m}\right) \tag{4.33}$$

The leading contribution to the amplitude of the decoherence factor is therefore

$$\ln|\boldsymbol{D}(t|\varphi,\varphi')|\simeq\frac{1}{4}\sum_{n=0}^{\infty}n^2\ln\frac{4a^2a'^2\sqrt{n^2+m^2a^2}\,\sqrt{n^2+m^2a'^2}}{\left(a^2\sqrt{n^2+m^2a^2}+a'^2\sqrt{n^2+m^2a'^2}\right)^2} \tag{4.34}$$

The first term, n^2, in the sum comes from the degeneracy of the eigenfunctions. This expression has divergences which *cannot* be represented as additive functions of a and a'. This means that no one-argument counterterm to $\boldsymbol{\Gamma}$ and $\boldsymbol{\Gamma}'$ in (4.27) can cancel these divergences of the amplitude (Paz and Sinha 1992). One might try to apply standard regularisation schemes from quantum field theory, such as dimensional regularisation. The corresponding calculations have been performed in Barvinsky *et al.* (1999a) and will not be given here. The important result is that, although they render the sum (4.34) convergent, they lead to a *positive* value of this expression. This means that the decoherence factor would diverge for $(\varphi-\varphi')\to\infty$ and thus spoil one of the crucial properties of a density matrix – the boundedness of $\mathrm{tr}\,\hat{\rho}^2$. The dominant term in the decoherence factor would read

$$\ln|\boldsymbol{D}|=\frac{\pi}{24}(ma)^3+O(m^2),\quad a\gg a' \tag{4.35}$$

and would thus be unacceptable for a density matrix. Reduced density matrices are usually not considered in quantum field theory, so this problem has not been encountered before. A behaviour such as in (4.35) is even obtained in the case of massless conformally invariant fields, for which one would expect a decoherence factor equal to one, since they decouple from the gravitational

background. How, then, does one have to proceed in order to obtain a sensible regularisation?

The crucial point is to perform a *redefinition* of environmental fields and to invoke a physical principle to fix this redefinition. The situation is somewhat analogous to the treatment of the S-matrix in quantum field theory: off-shell S-matrix and effective action depend on the parametrisation of the quantum fields (Vilkovisky 1984), in analogy to the non-diagonal elements of the reduced density matrix. In Laflamme and Louko (1991) and Kiefer (1992) it has been proposed within special models to rescale the environmental fields by a power of the scale factor. It was therefore suggested in Barvinsky *et al.* (1999a) to redefine the environmental fields by a power of the scale factor that corresponds to the conformal weight of the field (which is defined by the invariance of the conformally invariant wave equation). For a scalar field in four spacetime dimensions this amounts to a multiplication by a:

$$v_n(t) \rightarrow \tilde{v}_n(t) = a\, v_n(t)\ , \tag{4.36}$$

$$\tilde{\Omega}_n \quad = -ia\frac{d}{dt}\ln \tilde{v}_n^*\ . \tag{4.37}$$

An immediate test of this proposal is to see whether the decoherence factor is equal to one for a massless conformally invariant field. In this case, the basis functions and frequency functions read, respectively,

$$\tilde{v}_n^*(t) = \left(\frac{1+i\sinh Ht}{1-i\sinh Ht}\right)^{\frac{n}{2}}, \tag{4.38}$$

$$\tilde{\Omega}_n = -ia\frac{d}{dt}\ln \tilde{v}_n^*(t) = n\ . \tag{4.39}$$

Hence, $\tilde{\boldsymbol{D}}(t|\varphi,\varphi') \equiv 1$. The same holds also for the electromagnetic field (which in four spacetime dimensions is conformally invariant). It is interesting to note that the degree of decoherence caused by a certain field depends on the spacetime dimension, since its conformal properties are dimension-dependent.

For a massive minimally coupled field the new frequency function reads

$$\tilde{\Omega}_n = \left[\sqrt{n^2+m^2a^2} + i\sinh Ht\left(\frac{1}{2}\frac{m^2a^2}{n^2+m^2a^2}\right)\right] + O(1/m)\ . \tag{4.40}$$

Note that, in contrast to (4.33), there is no factor of a in front of this expression. Since (4.40) is valid in the large-mass limit, it corresponds to modes which evolve adiabatically on the gravitational background, the imaginary part in (4.40) describing particle creation.

It turns out that the imaginary part of the decoherence factor has at most logarithmic divergences and, therefore, affects only the phase of the density matrix. Moreover, these divergences decompose into an *additive* sum of one-argument functions and can thus be cancelled by adding counterterms to the classical action S (and S') in (4.27) (Paz and Sinha 1992). The real part

is simply convergent and gives a finite decoherence amplitude. This result is formally similar to the result for the decoherence factor in QED (Kiefer 1992).

For $a \gg a'$ (far off-diagonal terms) one gets the expression

$$|\tilde{\boldsymbol{D}}(t|\varphi,\varphi')| \simeq \exp\left[-\frac{(ma)^3}{24}\left(\pi-\frac{8}{3}\right)+O(m^2)\right] . \tag{4.41}$$

Compared with the naively regularised (and inconsistent) expression (4.35), π has effectively been replaced by $8/3-\pi$. In the vicinity of the diagonal, one obtains

$$\ln|\tilde{\boldsymbol{D}}(t|\varphi,\varphi')| = -\frac{m^3\pi a(a-a')^2}{64} , \tag{4.42}$$

a behaviour similar to (4.41).

An interesting case is also provided by minimally coupled massless scalar fields and by gravitons. They share the basis- and frequency functions in their respective conformal parametrisations:

$$\tilde{v}_n^*(t) = \left(\frac{1+i\sinh Ht}{1-i\sinh Ht}\right)^{\frac{n}{2}}\left(\frac{n-i\sinh Ht}{n+1}\right), \tag{4.43}$$

$$\tilde{\Omega}_n = \frac{n(n^2-1)}{n^2-1+H^2a^2} - i\frac{H^2a^2\sqrt{H^2a^2-1}}{n^2-1+H^2a^2}. \tag{4.44}$$

They differ only by the range of the quantum number n ($2 \leq n$ for inhomogeneous scalar modes and $3 \leq n$ for gravitons) and by the degeneracies of the n-th eigenvalue of the Laplacian,

$$\dim(n)_{\rm scal} = n^2 , \tag{4.45}$$

$$\dim(n)_{\rm grav} = 2(n^2-4). \tag{4.46}$$

For far off-diagonal elements one obtains the decoherence factor

$$|\tilde{\boldsymbol{D}}(t|\varphi,\varphi')| \sim e^{-C(Ha)^3}, \;\; a \gg a', \;\; C>0 , \tag{4.47}$$

while in the vicinity of the diagonal one finds

$$|\tilde{\boldsymbol{D}}(t|\varphi,\varphi')| \sim \exp\left(-\frac{\pi^2}{32}(H-H')^2t^2e^{4Ht}\right) , \tag{4.48}$$

$$\sim \exp\left(-\frac{\pi^2H^4a^2}{8}(a-a')^2\right) , \;\; Ht \gg 1 . \tag{4.49}$$

These expressions exhibit a rapid disappearance of non-diagonal elements during the inflationary evolution.

It is interesting that the behaviour of fermions concerning decoherence is different from the behaviour of bosons (Barvinsky *et al.* 1999c). Since their conformal weight is $-3/2$ in four dimensions, the environmental fermionic fields are reparametrised by a factor $a^{-3/2}$. For $m=0$ this does, as in the

bosonic case, render the decoherence factor finite and, due to conformal invariance, makes it equal to one. The situation for $m \neq 0$ is, however, different. In spite of the conformal reparametrisation, the decoherence factor is divergent. Moreover, dimensional regularisation would again spoil crucial properties of the density matrix and make it inconsistent. There remains, however, a freedom of reparametrisation in the fermionic case (Barvinsky *et al.* 1999c): this is a Bogoliubov transformation that is analogous to a Foldy-Wouthuysen transformation in Minkowski space (the decoupling of spinor components in the nonrelativistic limit). Since it is explicitly n-dependent, it corresponds to a *nonlocal* field redefinition. Instead of m one has now an effective n-dependent mass $\tilde{m}$ depending on the transformation. How can one fix this field redefinition? In Barvinsky *et al.* (1999c) the principle was put forward that decoherence should be *minimal* in the absence of particle creation. This is already implemented in the massless case. In the massive case, it means that decoherence is absent for a stationary spacetime which exhibits no particle creation. This leads to a decoherence factor

$$|\tilde{\boldsymbol{D}}(t|\varphi,\varphi')| \sim \exp\left(-C'm^2H^2a^2(a-a')^2\right), \quad C' > 0\ . \tag{4.50}$$

While decoherence is thus absent in the absence of particle creation, for bosons it is minimal in the sense that it is absent in the conformally-coupled case, but still present in the massive case – the expressions (4.41) and (4.42) do not depend on H. Formally, this is due to the fact that in the fermionic case one has m^2 instead of m^3 in the exponent; since one would expect to find factors of a in the nominator of the exponent (as is suggested by the coupling in the action), they have to be accompanied by corresponding factors of H for dimensional reason. Comparing (4.50) with (4.41) and (4.42) (which are valid for $m \gg H$), one recognises that fermions are *less efficient* in producing decoherence. In the massless case, there influence is totally absent. The point that decoherence is linked with particle creation has been made before (Calzetta and Hu 1994, Hu and Matacz 1995). Using the influence-functional approach to decoherence, see Chap. 5, one can derive an explicit formula connecting the decoherence factor with the Bogoliubov coefficients describing particle creation (Hu and Matacz 1995).[5] Given a special initial state (a "vacuum"), this encodes the irreversible aspect of decoherence. In the massless bosonic case, (4.47) and (4.49), the effect may be interpreted as arising from a cutoff at a mode number $n \approx aH$, i.e., a cutoff of modes whose wavelength a/n is smaller than the Hubble scale H^{-1} (Halliwell 1989). As we shall see in the next subsection, these are exactly the modes that experience particle creation.

[5] The decoherence factor in the massive bosonic case, (4.41) and (4.42), comes from the adiabatic part of $\tilde{\Omega}_n$ and is not directly related to particle creation. This is not in conflict with Hu and Matacz (1995), since there the assumption is being made that the state separates between system and environment in the past, which is not the case here.

The above analysis of decoherence was based on the state (4.18). One might, however, start with a quantum state that is a superposition of many semiclassical components, i.e. many components of the form $\exp(iS^k_{\rm cl})$, where each $S^k_{\rm cl}$ is a solution of the Hamilton-Jacobi equation for a and φ. Decoherence between different such semiclassical *branches* has also been the subject of intense investigation (Halliwell 1989, Kiefer 1992). The important point is that decoherence between different branches is usually weaker than the above discussed decoherence within one branch. Moreover, it usually follows from the presence of decoherence within one branch. In the special case of a superposition of (4.18) with its complex conjugate, one can immediately recognise that decoherence between the semiclassical components is smaller than within one component: in the expression (4.29) for the decoherence factor, the term $\Omega_n + \Omega_n'^*$ in the denominator is replaced by $\Omega_n + \Omega_n'$. Therefore, the imaginary parts of the frequency functions add up instead of partially cancelling each other and (4.29) becomes smaller. One also finds that the decoherence factor is equal to one for vanishing expansion of the semiclassical universe (Kiefer 1992).

We note that the decoherence between the $\exp(iS_{\rm cl})$ and $\exp(-iS_{\rm cl})$ components can be interpreted as a *symmetry breaking* analogously to the case of sugar molecules, see Sect. 3.2.4 and Chap. 9. There, the Hamiltonian is invariant under space reflections, but the state of the sugar molecules exhibits chirality. Here, the Hamiltonian in (4.6) is invariant under complex conjugation,[6] while the "actual states" (i.e., one decohering WKB component in the total superposition) are of the form $\exp(iS_{\rm cl})$ and are thus intrinsically complex. It is therefore not surprising that the recovery of the classical world follows only for complex states, in spite of the real nature of the Wheeler-DeWitt equation (see in this context Barbour 1993). Since this is a prerequisite for the derivation of the Schrödinger equation, one might even say that *time* (the WKB time parameter in the Schrödinger equation) arises from symmetry breaking.

The above considerations thus lead to the following picture. The Universe was essentially "quantum" at the onset of inflation. Mainly due to bosonic fields, decoherence set in and led to the emergence of many "quasi-classical branches" which are dynamically independent of each other. Strictly speaking, the very concept of time makes only sense after decoherence has occurred. In addition to the horizon problem etc., inflation also solves the "classicality problem". It remains of course unclear why inflation happened in the first place (if it really did). Looking back from our Universe (our semiclassical branch) to the past, one would notice that at the time of the onset of inflation our component would interfere with other components to form a timeless quantum-gravitational state. The Universe would thus cease to be transparent to earlier times (because there was no time). This demonstrates in an

[6] We ignore here alternative approaches which use a complex Hamiltonian from the very beginning (Kiefer 1993).

impressive way that quantum-gravitational effects are not restricted to the Planck scale.

It is interesting that a similar kind of constructive interference would occur near the turning point of a classically recollapsing universe (Kiefer and Zeh 1995). This is a direct consequence of the consistent way in which boundary conditions have to be imposed in this case. Again, this demonstrates that quantum effects are not restricted a priori to a particular scale and that it is a quantitative question referring to the dynamics when and to which extent classical properties emerge.

Our analysis has been restricted to the case where the "system" is taken to be a Friedmann universe containing a homogeneous scalar field. This is justified from phenomenological grounds, since our Universe appears isotropic and homogeneous on largest scales. Again, this may be traced back to the presumed occurrence of an inflationary phase and the validity of the cosmic no-hair conjecture. In spite of this, one can discuss decoherence in the context of anisotropic models, too (Gangui *et al.* 1991, Camacho and Camacho-Galván 1999), and find classical properties for the corresponding scale factors.

We want finally to stress the importance of decoherence for the *origin of irreversibility* in our Universe (Zeh 2001; Kiefer and Zeh 1995). Since the entropy of the present Universe (defined by its "relevant" degrees of freedom) is still extremely small compared to its maximal possible value (which would be achieved if the whole mass of the Universe were present in the form a black hole), the evolution of the Universe must have been started with a state of almost zero entropy (Penrose 1981). A possible explanation of this fact must necessarily invoke the fundamental quantum theory of gravity. It has been argued in the above references that a simple boundary condition at $a \to 0$ for the wave function of the Universe may be sufficient to explain the observed arrow of time, and may even lead to macroscopic quantum effects near the turning point of a classically recollapsing universe as well as for black holes. Such a boundary condition was proposed, for example, in Conradi and Zeh (1991). It roughly states that the wave function for small a depends only on a itself, but not on further degrees of freedom. This is consistent with the special form of the potential in the Wheeler-DeWitt equation. The wave function is thus independent, in this limit, of the "higher multipoles" introduced in this section. For increasing size of the Universe, the total state becomes entangled with these further degrees of freedom, and the decoherence for the "relevant subsystem" can be recognised after the "irrelevant" part is integrated out. The local entropy connected with the scale factor and other "relevant" variables, as calculated from the reduced density matrix in the standard way, $S = -k_{\mathrm{B}}\mathrm{tr}(\rho \ln \rho)$, thus *increases* and gives rise to the observed arrow of time in the Universe. An interesting consequence is the occurrence of recoherence in the case of a classically recollapsing universe (Kiefer and Zeh 1995).

...

References Chapter 4

Albrecht, A., Ferreira, P., Joyce, M., and Prokopec, T. (1994): "Inflation and squeezed quantum states." *Phys. Rev.* **D50**, 4807–20.

Allen, B. (1985): "Vacuum states in de Sitter space." *Phys. Rev.* **D32**, 3136–49.

Allen, B. (1997): "The stochastic gravity-wave background: sources and detection." In: *Relativistic gravitation and gravitational radiation*, ed. by J.-A. Marck and J.-P. Lasota (Cambridge University Press, Cambridge).

Amann, A. (1991b): "Molecules coupled to their environment.". In: *Large Scale Molecular Systems*, ed. by W. Gans, A. Blumen, and A. Amann (Plenum Press, New York), p. 3–22.

Amann, A. (1993): "The Gestalt problem in quantum theory: Generation of molecular shape by the environment." *Synthese* **97**, 125–156.

Amati, D. and Russo, J. G. (1999): "Fundamental strings as black bodies." *Phys. Lett. B* **454**, 207–12 (1999).

Anglin, J.R. and Zurek, W.H. (1996a): "Decoherence of Quantum Fields: Pointer States and Predictability." Phys. Rev. **D53**, 7327–35.

Anglin, J.R. and Zurek, W.H. (1996b): "A precision test of decoherence." In: *Dark matter in cosmology, quantum measurements, experimental gravitation*, ed. by R. Ansari, Y. Giraud-Heraud, and J. Tran Thanh Van (Editions Frontières, Gif-sur-Yvette), p. 263–70.

Audretsch, J., Mensky, M., and Müller, R. (1995): "Continuous measurement and localization in the Unruh effect." *Phys. Rev.* **D51**, 1716–27.

Barbour, J.B. (1993): "Time and complex numbers in canonical quantum gravity." *Phys. Rev.* **D47**, 5422–5429.

Barbour, J.B. (1994a): "The emergence of time and its arrow from timelessness." In: *Physical Origins of Time Asymmetry*, ed. by J.J. Halliwell, J.P. Pérez-Mercader, and W.H. Zurek (Cambridge University Press), p. 405–414.

Barbour, J.B. (1994b): "The timelessness of quantum gravity: II. The appearance of dynamics in static configurations." *Class. Quantum Grav.* **11**, 2875–2897.

Barvinsky, A.O., Kamenshchik, A.Yu, and Karmazin, I.P. (1992): "One loop quantum cosmology: Zeta function technique for the Hartle-Hawking wave function of the universe." *Ann. Phys. (N.Y.)* **219**, 201–42.

Barvinsky, A.O., Kamenshchik, A.Yu, and Mishakov, I.V. (1997): "Quantum origin of the early inflationary universe." *Nucl. Phys.* **B491**, 387–426.

Barvinsky, A.O., Kamenshchik, A.Yu., Kiefer, C., and Mishakov, I.V. (1999a): "Decoherence in quantum cosmology at the onset of inflation." *Nucl. Phys.* **B551**, 374–96.

Barvinsky, A.O., Kamenshchik, A.Yu., and Kiefer, C (1999b): "Origin of the inflationary Universe." *Mod. Phys. Lett.* **A14**, 1083–88.

Barvinsky, A.O., Kamenshchik, A.Yu., and Kiefer, C (1999c): "Effective action and decoherence by fermions in quantum cosmology." *Nucl. Phys.* **B552**, 420–44.

Benatti, F. and Floreanini, R. (2002): "Planck's scale dissipative effects in atom interferometry." `quant-ph/0208164`.

Bohr, N. and Rosenfeld, L. (1933): "Zur Frage der Messbarkeit der elektromagnetischen Feldgrössen." *Det Kgl. Danske Videnskabernes Selskab. Mathematisk - fysiske Meddelelser* **XII**, 8, 3–65; English translation in Wheeler and Zurek (1983).

Börner, G. and Gottlöber, S. (eds.) (1997): *The evolution of the Universe* (Wiley, Chichester).

Brandenberger, R., Laflamme, R., and Mijić, M. (1990): "Classical perturbations from decoherence of quantum fluctuations in the inflationary universe." *Mod. Phys. Lett.* **A5**, 2311–2317.

Breuer, H.-P. and Petruccione, F. (2001): "Destruction of quantum coherence through emission of bremsstrahlung." *Phys. Rev.* **A63**, 032102.

Caldeira, A.O. and Leggett, A.J. (1983a): "Path integral approach to quantum Brownian motion." *Physica* **121A**, 587–616.

Caldeira, A.O. and Leggett, A.J. (1983b): "Quantum tunnelling in a dissipative system". *Ann. Phys. (N.Y.)* **149**, 374–456.

Callan, C.G., Giddings, S.B., Harvey, J.A., and Strominger, A. (1992): "Evanescent black holes." *Phys. Rev.* **D45**, R1005–09.

Calzetta, E. and Hu, B.L. (1994): "Noise and fluctuations in semiclassical gravity." *Phys. Rev.* **D51**, 6636–55.

Calzetta, E. and Hu, B.L. (1995): "Quantum fluctuations, decoherence of the mean field, and structure formation in the early Universe." *Phys. Rev.* **D52**, 6770–88.

Camacho, A. and Camacho-Galván, A. (1999): "Time emergence by self-measurement in a quantum isotropic universe." *Nuovo Cim.* **B114**, 923–38.

Chou, K., Su, S., Hao, B., and Yu, L. (1985): "Equilibrium and nonequilibrium formalism made unified." *Phys. Rep.* **118**, 1–131.

Cohen-Tannoudji, C. (1992): *Atom-Photon Interactions* (Wiley, New York).

Coleman, S. (1988a): "Black holes as red herrings: Topological fluctuations and the loss of quantum coherence." *Nucl. Phys.* **B307**, 867–82.

Coleman, S. (1988b): "Why there is nothing rather than something: A theory of the cosmological constant." *Nucl. Phys.* **B310**, 643–68.

Conradi, H.D. and Zeh, H.D. (1991): "Quantum cosmology as an initial value problem." *Phys. Lett.* **A154**, 321–326.

Davidovich, L., Brune, M., Raimond, J.M., and Haroche, S. (1996): "Mesoscopic quantum coherences in cavity QED: preparation and decoherence monitoring schemes." *Phys. Rev.* **A53**, 1295–1309.

Demers, J.-G. and Kiefer, C. (1996): "Decoherence of black holes by Hawking radiation." Phys. Rev. **D53**, 7050–7061.

DeWitt, B.S. (1967): "Quantum theory of gravity I. The canonical theory." *Phys. Rev.* **160**, 1113–1148.

Diósi, L. (1990a): "Landau's density matrix in quantum electrodynamics." *Found. Phys.* **20**, 63–70.

Diósi, L. (1990b): "Relativistic theory for continuous measurement of quantum fields." *Phys. Rev.* **A42**,

Diósi, L. (1995a): "Comments on 'Objectification of classical properties induced by quantum vacuum fluctuations'." *Phys. Lett.* **A97**, 183–184.

Dürr, D. and Spohn, H. (2000): "Decoherence through coupling to the radiation field." In: *Decoherence: Theoretical, experimental, and conceptual problems*, ed. by Ph. Blanchard, D. Giulini, E. Joos, C. Kiefer, and I.-O. Stamatescu (Springer, Berlin), p. 77–86.

Ellis, J., Hagelin, J., Nanopoulos, D.V., and Srednicki, M. (1984): "Search For Violations Of Quantum Mechanics." *Nucl. Phys.* **B241**, 381–405.

Ellis, J., Mohanty, S., and Nanopoulos, D.V. (1990): "Wormholes violate quantum mechanics in SQUIDs." *Phys. Lett.* **B235**, 305–12.

Ellis, J., Mavromatos, N.E., and Nanopoulos, D.V. (1996): "CPT and superstring." Report hep-ph/9607434 (unpublished).

Ellis, J., Kanti, P., Mavromatos, N.E., Nanopoulos, D.V., and Winstanley, E. (1998): "Decoherent scattering of light particles in a D-brane background." *Mod. Phys. Lett.* **A13**, 303–20.

Eppley, K. and Hannah, E. (1977): "The necessity of quantizing the gravitational field." *Found. Phys.* **7**, 51–68.

Ford, L.H. (1993): "Electromagnetic vacuum fluctuations and electron coherence." *Phys. Rev.* **D47**, 5571–5580.

Gangui, A., Mazzitelli, F.D., and Castagnino, M.A. (1991): "Loss of coherence in multidimensional minisuperspaces." *Phys. Rev.* **D43**, 1853–1858.

Giddings, S.B. and Strominger, A. (1988a): "Axion induced topology change in quantum gravity and string theory." *Nucl. Phys.* **B306**, 890–907.

Giddings, S.B. and Strominger, A. (1988b): "Loss of incoherence and determination of coupling constants in quantum gravity." *Nucl. Phys.* **B307**, 854–66.

Giulini, D. and Kiefer, C. (1995): "Consistency of semiclassical gravity." *Class. Quantum Grav.* **12**, 403–411.

Giulini, D., Kiefer, C., and Zeh, H.D. (1995): "Symmetries, superselection rules, and decoherence." *Phys. Lett.* **A199**, 291–298.

Grishchuk, L.P. and Sidorov, Yu.V. (1989): "On the quantum state of relic gravitons." *Class. Quantum Grav.* **6**, L161–65.

Habib, S., Kluger, Y., Mottola, E., and Paz, J.P. (1996): "Dissipation and decoherence in mean field theory." *Phys. Rev. Lett.* **76**, 4660-4663.

Halliwell, J.J. (1989): "Decoherence in quantum cosmology." *Phys. Rev.* **D39**, 2912–2923.

Halliwell, J.J. (1991): "Introductory lectures on quantum cosmology." In: *Quantum cosmology and baby universes*, ed. by S. Coleman, J.B. Hartle, T. Piran, and S. Weinberg (World Scientific, Singapore), p. 159–243.

Halliwell, J.J. (1998): "Effective theories of coupled classical and quantum variables from decoherent histories: A new approach to the backreaction problem." *Phys. Rev.* **D57**, 2337–48.

Halliwell, J.J. and Hawking, S.W. (1985): "Origin of structure in the Universe." *Phys. Rev.* **D31**, 1777–1791.

Hasselbach, F., Kiesel, H., and Sonnentag, P. (2000): "Exploration of the fundamentals of quantum mechanics by charged particle interferometry." In: *Decoherence: Theoretical, experimental, and conceptual problems*, ed. by Ph. Blanchard, D. Giulini, E. Joos, C. Kiefer, and I.-O. Stamatescu (Springer, Berlin), p. 201–12.

Hawking, S.W. and Ross, S. (1997): "Loss of quantum coherence through scattering off virtual black holes." *Phys. Rev.* **D56**, 6403–15.

Horowitz, G.T. (1997): "Quantum states of black holes." In: *Black holes and relativistic stars*, ed. by R.M. Wald (The University of Chicago Press, Chicago), p. 241–66.

Hu, B.L. and Matacz, A. (1995): "Back reaction in semiclassical gravity: The Einstein-Langevin equation." *Phys. Rev.* **D51**, 1577–86.

Isham, C.J. (1992): "Canonical Quantum Gravity and the Problem of Time." In: *Integrable Systems, Quantum Groups, and Quantum Field Theories*, ed. by L.A. Ibart and M.A. Rodrigues (Kluwer, Dordrecht), p. 157–287.

Isham, C.J. (1994): "Prima facie questions in quantum gravity." In: *Canonical gravity: From classical to quantum*, ed. by J. Ehlers and H. Friedrich (Springer, Berlin), p. 1–21.

Jackiw, R. (1988): "Analysis on infinite-dimensional manifolds - Schrödinger representation for quantized fields." In *Field Theory and Particle Physics*, edited by O. Eboli, M. Gomes, and A. Santano (World Scientific, Singapore).

Joos, E. (1986b): "Why do we observe a classical spacetime?" *Phys. Lett.* **A116**, 6–8.

Joos, E. and Zeh, H.D. (1985): "The Emergence of Classical Properties Through Interaction with the Environment." *Z. Phys.* **B59**, 223–243.

Károlyházy, F., Frenkel, A., and Lukács, B. (1986): "On the possible role of gravity in the reduction of the wave function." In: *Quantum Concepts in Space and Time*, ed. by C.J. Isham and R. Penrose (Clarendon Press), p. 109–128.

Kiefer, C. (1987): "Continuous measurement of mini-superspace variables by higher multipoles." *Class. Quantum Grav.* **4**, 1369–1382.

Kiefer, C. (1988): "Wave packets in minisuperspace." *Phys. Rev.* **D38**, 1761–72.

Kiefer, C. (1992): "Decoherence in quantum electrodynamics and quantum cosmology." *Phys. Rev.* **D46**, 1658–1670.

Kiefer, C. (1993): "Topology, decoherence, and semiclassical gravity." *Phys. Rev.* **D47**, 5414–5421.

Kiefer, C. (1994): "The semiclassical approximation to quantum gravity." In *Canonical gravity: From classical to quantum*, ed. by J. Ehlers and H. Friedrich (Springer, Berlin), p. 170–212.

Kiefer, C. (1998a): "Interference of two independent sources." *Am. J. Phys.* **66**, 661–62.

Kiefer, C. (1998b): "Towards a full quantum theory of black holes." In: *Black holes: Theory and observation*, ed. by F.W. Hehl, C. Kiefer, and R. Metzler (Springer, Berlin), p. 416–50.

Kiefer, C. (1999): "Thermodynamics of black holes and Hawking radiation." In: *Classical and quantum black holes*, Studies in High Energy Physics, Cosmology and Gravitation, ed. by P. Fré, V. Gorini, G. Magli, and U. Moschella (IOP Publishing, Bristol), p. 17–74.

Kiefer, C. (2000): "Conceptual issues in quantum cosmology." In: *Towards quantum gravity*, ed. by J. Kowalski-Glikman (Springer, Berlin), p. 158–87.

Kiefer, C. (2001): "Hawking radiation from decoherence." *Class. Quantum Grav.* **18**, L151–54.

Kiefer, C. (2003): *Quantum gravity* (Oxford University Press, forthcoming).

Kiefer, C., Lesgourgues, J., Polarski, D., and Starobinsky, A.A. (1998a): "The coherence of primordial fluctuations produced during inflation." *Class. Quantum Grav.* **15**, L67–72.

Kiefer, C., Padmanabhan, T., and Singh, T.P. (1991): "A comparison between semiclassical gravity and semiclassical electrodynamics." *Class. Quantum Grav.* **8**, L185–L192.

Kiefer, C. and Polarski, D. (1998): "Emergence of classicality for primordial fluctuations: Concepts and analogies." *Ann. Phys. (Leipzig)* **7**, 137–58.

Kiefer, C., Polarski, D., and Starobinsky, A.A. (1998b): "Quantum-to-classical transition for fluctuations in the early universe." *Int. J. Mod. Phys.* **D7**, 455-62.

Kiefer, C., Polarski, D., and Starobinsky, A.A. (2000): "Entropy of gravitons produced in the early Universe." *Phys. Rev.* **D62**, 043518.

Kiefer, C. and Zeh, H.D. (1995): "Arrow of time in a recollapsing quantum universe." *Phys. Rev.* **D51**, 4145–4153.

Kübler, O. and Zeh, H.D. (1973): "Dynamics of Quantum Correlations." *Ann. Phys. (N.Y.)* **76**, 405–418.

Kuchař, K.V. (1992): "Time and Interpretations of Quantum Gravity." In: *Proceedings of the 4th Canadian Conference on General Relativity and Relativistic Astrophysics*, ed. by G. Kunstatter, D. Vincent, and J. Williams (World Scientific, Singapore), p. 211–314.

Laflamme, R. and Louko, J. (1991): "Reduced density matrices and decoherence in quantum cosmology." *Phys. Rev.* **D43**, 3317–3331.

Landau, L. (1927): "Das Dämpfungsproblem in der Wellenmechanik." *Z. Phys.* **45**, 430–441.

Liddle, A.R. and Lyth, D.H. (2000): *Cosmological inflation and large-scale structure* (Cambridge University Press, Cambridge).

Meekhof, C., Monroe, C., King, B.E., Itano, W.M., and Wineland, D.J. (1996): "Generation of nonclassical motional states of a trapped atom." *Phys. Rev. Lett.* **76**, 1796–99.

Meyers, R. (1997): "Pure states don't wear black." *Gen. Rel. Grav.* **29**, 1217–22.

Mijič, M. (1998): "Particle production and classical condensates in de Sitter space." *Phys. Rev.* **D57**, 2138–46.

Netterfield, C. B., et al. (2002). "A measurement by BOOMERANG of multiple peaks in the angular power spectrum of the cosmic microwave background." *Astrophys. J.* **571**, 604–14.

Nicklaus, M. and Hasselbach, F. (1993): "Wien filter: A wave-packet-shifting device for restoring longitudinal coherence in charged-matter-wave interferometers." *Phys. Rev.* **A48**, 152–61.

Paz, J.P. and Sinha, S. (1992): "Decoherence and back reaction in quantum cosmology: Multidimensional minisuperspace examples." *Phys. Rev.* **D45**, 2823–2842.

Penrose, R. (1981): "Time asymmetry and quantum gravity." In: *Quantum Gravity Vol. 2*, ed. by Isham, C.J., Penrose, R. and Sciama, D.W. (Clarendon, Oxford).

Penrose, R. (1986): "Gravity and state vector reduction." In: *Quantum Concepts in Space and Time"*, ed. by C.J. Isham and R. Penrose (Clarendon Press), p. 129-146.

Pfeifer, P. (1980): *Chiral Molecules - a Superselection Rule Induced by the Radiation Field*. Dissertation #6551, ETH Zürich.

Polarski, D. (1999): "Primordial fluctuations from inflation: A consistent histories approach." *Phys. Lett.* **B446**, 53–57.

Polarski, D. and Starobinsky, A.A. (1996): "Semiclassicality and decoherence of cosmological perturbations." *Class. Quantum Grav.* **13**, 377–92.

Prokopec, T. (1993): "Entropy of the squeezed vacuum." *Class. Quantum Grav.* **10**, 2295–2306.

Sakagami, M. (1988): "Evolution from Pure States into Mixed States in de Sitter Space." *Progr. Theor. Phys.* **79**, 442–453.

Santos, E. (1994): "Objectification of classical properties induced by quantum vacuum fluctuations." *Phys. Lett.* **A188**, 198–204.

Santos, E. (1995): "Reply to comments on 'Objectification of classical properties induced by quantum vacuum fluctuations'." *Phys. Lett.* **A197**, 185–186.

Schumaker, B. (1986): "Quantum mechanical pure states with Gaussian wave functionals." *Phys. Rep.* **135**, 317–408.

Seiberg, N. and Witten, E. (1999): "String theory and noncommutative geometry". *JHEP* **9909**, 032.

Shaisultanov, R.Zh. (1995): "Backreaction in scalar QED, Langevin equation and decoherence functional." Report hep-th/9509154 (unpublished).

Starobinsky, A.A. (1979): "Spectrum of relict gravitational radiation and the early state of the universe." *JETP Lett.* **30**, 682–85.

Vilkovisky, G.A. (1984): "The unique effective action in quantum field theory." *Nucl. Phys.* **B234**, 125–37.

Wald, R.M. (1984): *General relativity* (University of Chicago Press).

Walls, D.F. and Milburn, G.J. (1994): *Quantum Optics* (Springer, Berlin).

Wheeler, J.A. (1968): "Superspace and the nature of quantum geometrodynamics." In: *Battelle rencontres*, ed. by C.M. De Witt and J.A. Wheeler (Benjamin, New York), p. 242–307.

Wightman, A.S. (1995): "Superselection rules: old and new." *Nuovo Cim.* **110B**, 751–769.

Zeh, H.D. (1986): "Emergence of classical time from a universal wave function." *Phys. Lett.* **A116**, 9–12.

Zeh, H.D. (2001): *The Physical Basis of the Direction of Time* (Springer, Heidelberg).

Why do cosmological perturbations look classical to us?

Claus Kiefer*
Institut für Theoretische Physik, Universität zu Köln, Zülpicher Strasse 77, 50937 Köln, Germany

David Polarski†
Laboratoire de Physique Théorique et Astroparticules, CNRS, Université de Montpellier II, 34095 Montpellier, France

According to the inflationary scenario of cosmology, all structure in the Universe can be traced back to primordial fluctuations during an accelerated (inflationary) phase of the very early Universe. A conceptual problem arises due to the fact that the primordial fluctuations are quantum, while the standard scenario of structure formation deals with classical fluctuations. In this essay we present a concise summary of the physics describing the quantum-to-classical transition. We first discuss the observational indistinguishability between classical and quantum correlation functions in the closed system approach (pragmatic view). We then present the open system approach with environment-induced decoherence. We finally discuss the question of the fluctuations' entropy for which, in principle, the concrete mechanism leading to decoherence possesses observational relevance.

Keywords: primordial fluctuations; inflation; decoherence; entropy

I. INTRODUCTION

It is often emphasized these days that the field of cosmology has entered a golden age. There is no doubt that the main reason for this statement is the accumulation of observations of ever increasing accuracy. In this way cosmological models aiming to describe the evolution of the Universe from the Big Bang until today are no longer purely speculative: their predictions can be tested and some models can indeed be ruled out.

With the advent of inflationary models, according to which the Universe underwent a phase of accelerated expansion at a very early stage, we now have at our disposal theoretical tools to apprehend such fundamental problems as the origin of cosmological perturbations and the eventual formation of large-scale structures like galaxies. There are many ways in which inflationary models address fundamental physical theories. As inflation is supposed to take place at very high energies in the early Universe, these models offer a unique window on energy scales of the order of 10^{15} GeV. Another intriguing aspect of these models is that inflationary perturbations originate from quantum fluctuations though we do not see this quantum nature in the Universe nowadays. It is this aspect of inflationary perturbations that we want to describe in our essay.

We could, of course, as well consider non-inflationary cosmological models in which perturbations are assumed to be classical from the beginning on. However, such models are plagued with problems of causality as distant points on the last-scattering surface, about 350.000 years after the Big Bang, were never in contact before. Hence the impressive homogeneity of the Cosmic Microwave Background (CMB) would have to be put in by hand in the absence of an inflationary stage. Inflationary models are thus much more natural – and they can be observationally tested.

The main part of our essay consists of four parts. We shall first give in Sec. II a brief review of inflationary cosmology and its mechanism for the generation of perturbations. We then discuss in Sec. III the quantum-to-classical transition in the closed system approach (we call it also the pragmatic view) which focusses on the indistinguishability of quantum expectation values and classical stochastic averages. Sec. IV presents the successful observational predictions which emerge from this scenario. Sec. V, then, is devoted to environmental decoherence. We discuss the problem of the classical variables (the pointer basis) as well as the entropy of the fluctuations and its observational significance. We end with a brief conclusion.

II. INFLATION

We give here a brief review of the way in which inflationary models give an elegant solution to many fundamental problems occuring in non-inflationary Big-Bang cosmology, see, for example,[1]. As we shall see, these models do also make characteristic predictions, by which we mean that in the absence of certain observable signatures most if not all inflationary models would be ruled out. We shall first describe the evolution of the homogeneous background for inflation and then turn to the generation of perturbations.

*Electronic address: kiefer@thp.uni-koeln.de
†Electronic address: polarski@lpta.univ-montp2.fr

A. Background expansion

The crucial point here is that inflation is a stage of accelerated expansion. In this stage, proper (physical) scales are stretched by a huge factor so that scales inside the Hubble radius during inflation will eventually end up at the end of inflation far outside the Hubble radius. Today these scales can correspond to cosmological scales, and typically scales corresponding to the Hubble radius today have exited the Hubble radius during inflation about 65 e-folds before the end of inflation. Typically, inflationary stages are quasi-de Sitter stages during which the Hubble parameter is nearly constant. As we shall see below, inflation provides a mechanism for the causal generation of perturbations.

It is a basic assumption that our Universe is on large scales homogeneous and isotropic. The metric is of the form

$$ds^2 = dt^2 - a^2(t)\ \left[\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta\ d\phi^2)\right]\ ; \tag{1}$$

a spatially flat universe corresponds to $k = 0$, a closed universe to $k = 1$, and an open universe to $k = -1$. (We set the speed of light $c = 1$ throughout.) In an expanding universe, the scale factor $a(t)$ is a growing function of time, which starts close to zero at the Big Bang about 13.7 billions years ago. The dynamics of the scale factor is given by the Friedmann equations,

$$\left(\frac{\dot{a}}{a}\right)^2 = \sum_i \frac{8\pi G}{3}\ \rho_i - \frac{k}{a^2}\ , \tag{2}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\sum_i(\rho_i + 3p_i)\ , \tag{3}$$

where the index i stands for any isotropic (comoving) perfect fluid. For radiation we have $p_{\rm r} = \rho_{\rm r}/3$, for dust $p_{\rm m} = 0$. For the recent accelerated expansion caused by some smooth dark energy component we would have $p_{\rm DE} = w_{\rm DE}\ \rho_{\rm DE}$, where $w_{\rm DE} < -1/3$ is still unknown and in many models time-dependent. ¿From (3) the expansion is typically decelerated, $\ddot{a} < 0$, unless at least one of the components satifies $\rho_i + 3p_i < 0$.

A space-independent scalar field $\phi(t)$ can be viewed as a comoving perfect fluid with

$$\rho_\phi = \frac{1}{2}\dot{\phi}^2 + V(\phi)\ , \tag{4}$$

$$p_\phi = \frac{1}{2}\dot{\phi}^2 - V(\phi)\ . \tag{5}$$

Hence, a scalar field $\phi(t)$ can induce an accelerated expansion provided

$$\dot{\phi}^2 < V(\phi)\ . \tag{6}$$

The field $\phi(t)$ driving the inflationary stage is called the inflaton and evolves according to the Klein-Gordon equation

$$\ddot{\phi} + 3H\dot{\phi} + \frac{dV}{d\phi} = 0\ , \tag{7}$$

which is the form taken by the conservation of energy for a perfect fluid defined by (4) and (5), and we have introduced the Hubble parameter $H \equiv \dot{a}/a$. In most inflationary models, the inflaton field $\phi(t)$ satisfies the slow-roll conditions $\ddot{\phi} \ll 3H\dot{\phi}$, and hence

$$3H\dot{\phi} \approx -\frac{dV}{d\phi}\ . \tag{8}$$

It is easy to show that the conditions for slow-roll to hold are

$$\dot{H} \ll 3H^2\ , \quad \frac{d^2V}{d\phi^2} \ll 9H^2\ , \tag{9}$$

in which case the condition (6) is amply satisfied so that accelerated expansion – inflation – takes place.

We conclude this brief summary on the background evolution during inflation by discussing the relative evolution of physical scales. The Hubble radius $R_{\rm H} \equiv H^{-1}$ defines an important scale in cosmology. If $a \propto t^p$, we have $R_{\rm H} \propto t$, and it is clear that $R_{\rm H}$ grows faster than a physical scale $\lambda \propto a$ during a decelerated expansion, which has $p < 1$. Hence physical scales greater than the Hubble radius, which we shall call "superhorizon" or "super-Hubble" scales, will eventually enter the Hubble radius, by which we mean that they will become smaller than $R_{\rm H}$: this is the situation in standard cosmology. This picture changes dramatically during inflation; to illustrate this we take a purely de Sitter stage, which is characterized by $H =$ constant and $a(t) \propto \exp(Ht)$. Now it is clear that physical scales inside the Hubble radius, which we shall call "subhorizon" or "sub-Hubble" scales will eventually become larger than the Hubble radius.

If a scale is said to cross the "horizon" 65 e-folds before the end of inflation, this means that at the end of inflation (where $t = t_{\rm e}$) one has $a = a_{\rm e} = {\rm e}^{65}a_k$ or $N_k = 65$ with

$$N_k = \frac{a_{\rm e}}{a_k}\ ; \tag{10}$$

here, $a_k \equiv a(t_k)$ if t_k is the "horizon-crossing" time of that particular scale with physical wavelength $(2\pi/k)a$. (Sometimes the factor 2π is omitted.) In a pure de Sitter stage this would mean that $H(t_{\rm e} - t_k) = 65$. If we can compute the present physical scale evolving from the Hubble radius during inflation, we know to which physical scale today a scale with given N_k corresponds. Depending on the details of the model, the Hubble radius *today* corresponds typically to a scale with $N_k \approx 65$. It can be shown that in slow-roll models N_k can be computed from the value $\phi(t_k)$ and that it depends on the potential $V(\phi)$.

In consistent inflationary scenarios, inflation is followed by a standard cosmic expansion during which scales that went outside $R_{\rm H}$ become again smaller than $R_{\rm H}$; they "re-enter the horizon". For a given scale, the number of e-folds between the first horizon crossing time t_k during inflation and the second horizon crossing time

during the radiation or matter stage at $t = t_{k,\rm f}$ is given by the parameter r_k,

$$r_k \equiv \ln \frac{a(t_{k,\rm f})}{a_k} \equiv \ln \frac{a_{k,\rm f}}{a_k} \ . \tag{11}$$

We shall see in Sec. III that r_k coincides with the squeezing parameter for a quantum state[2]. For typical cosmological scales today, $r_k \sim 100$ and even larger. Physically this corresponds to an enormous expansion of the universe, while a given scale k was outside the Hubble radius. As we shall see below, the ensuing huge amount of squeezing for the quantum state plays a crucial role in the quantum-to-classical transition of inflationary quantum fluctuations. It also means that the quantum state originating from inflation is a very peculiar one.

B. Generation of perturbations

During an inflationary stage, quantum field fluctuations evolve according to the general principles of quantum field theory. Inflation is supposed to take place at an energy scale where space-time can be described as a classical curved space-time on which the quantum field fluctuations are defined. The inflaton fluctuations $\delta\phi(\mathbf{x}, t)$ can be treated as a massless scalar field. This is an excellent approximation when the inflaton field satisfies the slow-roll conditions (9) and it is even exact when we consider primordial gravitational waves.

It is convenient to consider the rescaled quantity $a\delta\phi \equiv y(\mathbf{x}, t)$ and to work with conformal time $\eta = \int dt/a(t)$; a prime will be used to denote a derivative with respect to η. The formalism presented here is exact for gravitational waves, but can be extended in a straightforward way to the primordial density perturbations.

The quantization of the real perturbation $y(\mathbf{x}, \eta)$ proceeds with the usual canonical quantization scheme. We start from the classical Hamiltonian describing the perturbations,

$$\begin{aligned} H &\equiv \int d^3\mathbf{x}\ \mathcal{H}(y, p, \partial_i y, \eta) \\ &= \frac{1}{2}\int d^3\mathbf{k}[p(\mathbf{k})p^*(\mathbf{k}) + k^2 y(\mathbf{k})y^*(\mathbf{k}) \qquad (12) \\ &+ \frac{a'}{a}\left(y(\mathbf{k})p^*(\mathbf{k}) + p(\mathbf{k})y^*(\mathbf{k})\right)] \ , \qquad (13) \end{aligned}$$

where p is the momentum conjugate to y,

$$p \equiv \frac{\partial \mathcal{L}(y, y')}{\partial y'} = y' - \frac{a'}{a}y \ . \tag{14}$$

In (13) we have introduced the (time-dependent) Fourier transform $y(\mathbf{k}, \eta)$ of the rescaled fluctuation $y(\mathbf{x}, \eta)$. (We sometimes keep the dependence on η.) In the Lagrangian formulation, it obeys the following classical equation of motion:

$$y''(\mathbf{k}, \eta) + \left(k^2 - \frac{a''}{a}\right) y(\mathbf{k}, \eta) = 0 \ . \tag{15}$$

Upon quantization, the Fourier transforms are promoted to operators on which we impose the canonical commutation relations,

$$[y(\mathbf{k}, \eta), p^\dagger(\mathbf{k}', \eta)] = \mathrm{i}\delta^{(3)}(\mathbf{k} - \mathbf{k}') \ . \tag{16}$$

(We set $\hbar = 1$.) We can write the Hamiltonian operator in the following way:

$$\begin{aligned} H \ = \ & \int \frac{d^3\mathbf{k}}{2}\big[k\big(a(\mathbf{k})a^\dagger(\mathbf{k}) + a^\dagger(-\mathbf{k})a(-\mathbf{k})\big) + \\ & \mathrm{i}\frac{a'}{a}\big(a^\dagger(\mathbf{k})a^\dagger(-\mathbf{k}) - a(\mathbf{k})a(-\mathbf{k})\big)\big] \ . \end{aligned} \tag{17}$$

The time-dependent annihilation operators $a(\mathbf{k})$ (we often skip the argument η for conciseness) appearing in (17) are defined as usual,

$$a(\mathbf{k}) = \frac{1}{\sqrt{2}}\left(\sqrt{k}\ y(\mathbf{k}) + \frac{\mathrm{i}}{\sqrt{k}}p(\mathbf{k})\right) \ , \tag{18}$$

so that

$$y(\mathbf{k}) \ = \ \frac{a(\mathbf{k}) + a^\dagger(-\mathbf{k})}{\sqrt{2k}} \ , \tag{19}$$

$$p(\mathbf{k}) \ = \ -\mathrm{i}\sqrt{\frac{k}{2}}\left(a(\mathbf{k}) - a^\dagger(-\mathbf{k})\right) \ . \tag{20}$$

It is easily seen from (16) that a and $a^\dagger$ satisfy the commutation relations

$$[a(\mathbf{k}, \eta), a^\dagger(\mathbf{k}', \eta)] = \delta^{(3)}(\mathbf{k} - \mathbf{k}') \ . \tag{21}$$

Let us consider the time evolution of these operators. ¿From the Hamiltonian (17) we get

$$\begin{pmatrix} a'(\mathbf{k}) \\ (a^\dagger(-\mathbf{k}))' \end{pmatrix} = k \begin{pmatrix} -\mathrm{i} & \frac{aH}{k} \\ \frac{aH}{k} & \mathrm{i} \end{pmatrix} \begin{pmatrix} a(\mathbf{k}) \\ a^\dagger(-\mathbf{k}) \end{pmatrix} . \tag{22}$$

The second piece of the Hamiltonian (17), which is proportional to a'/a, is responsible for a mixing between creation and annihilation operators. In the Heisenberg representation it corresponds to a Bogolubov transformation; physically it means that particles are produced in pairs with opposite momenta. For reasons that will become clear later, this phenomenon is called squeezing in the Schrödinger picture; the corresponding squeezing parameter r_k turns out to be given by the expression (11) above. ¿From (22) one can see that mixing of creation and annihilation operators is efficient when the off-diagonal terms dominate, in other words, on super-Hubble scales when $aH/k \gg 1$.

Using (20) and (22), one obtains after a little algebra,

$$y(\mathbf{k}, \eta) \equiv f_k(\eta)\ a_{\mathbf{k}} + f_k^*(\eta)\ a_{-\mathbf{k}}^\dagger \ , \tag{23}$$

where $a_{\mathbf{k}} \equiv a(\mathbf{k}, \eta_0)$, and the field modes f_k obey Equation (15) and satisfy $f_k(\eta_0) = 1/\sqrt{2k}$. At the initial time

η_0, the field modes are deep inside the Hubble radius. Equation (23) can be written in the suggestive way

$$y(\mathbf{k},\eta) = \sqrt{2k}\ f_{k1}(\eta)\ y_{\mathbf{k}} - \sqrt{\frac{2}{k}}\ f_{k2}(\eta)\ p_{\mathbf{k}}\ , \qquad (24)$$

where $y_{\mathbf{k}} \equiv y(\mathbf{k},\eta_0)$ and $p_{\mathbf{k}} \equiv p(\mathbf{k},\eta_0)$, $f_{k1} = \Re f_k$, $f_{k2} = \Im f_k$. We have in an analogous way momentum modes $g_k(\eta)$, with $g_k(\eta_0) = \sqrt{k/2}$,

$$p(\mathbf{k}) = \sqrt{\frac{2}{k}}\ g_{k1}(\eta)\ p_{\mathbf{k}} + \sqrt{2k}\ g_{k2}(\eta)\ y_{\mathbf{k}}\ . \qquad (25)$$

We shall now address the first step in understanding why and to which extent these quantum field modes appear classically.

III. QUANTUM-TO-CLASSICAL TRANSITION: THE PRAGMATIC VIEW

In the last section we have described the evolution of the quantum modes in the Heisenberg representation, in which operators evolve in time and quantum states do not. While the quantum-to-classical transition is in general formulated in the Schrödinger picture, for the inflationary perturbations the Heisenberg picture provides deep insight, too.

To see this, let us assume that there is a limit in which f_{k2} and g_{k1} (or f_{k1} and g_{k2}) vanish. Then it is clear from (24) that the non-commutativity of the operators $y_{\mathbf{k}}$ and $p_{\mathbf{k}}$ is no longer relevant. What is the physical meaning of such a limit? Let us consider a classical stochastic system where the dynamics is still described by equations of the form (24), but with now $y(\mathbf{k},\eta_0)$ and $p(\mathbf{k},\eta_0)$ representing random initial values (c-numbers). If f_{k2} and g_{k1} vanish, we get

$$p(\mathbf{k},\eta) \equiv p_{\rm cl}(y(\mathbf{k},\eta)) = \frac{g_{k2}}{f_{k1}}\ y(\mathbf{k},\eta)\ . \qquad (26)$$

This is true for the quantum system (in the operator sense) and for the classical stochastic system (in the c-number sense). Therefore, for a given realization of the perturbation $y(\mathbf{k},\eta)$, the corresponding momentum $p_{\rm cl}(\mathbf{k},\eta)$ is fixed and equal to the classical momentum corresponding to this value $y(\mathbf{k},\eta)$. Then the quantum system is effectively equivalent to the classical random system, which is an ensemble of classical trajectories with a certain probability associated to each of them[3].

This is, in fact, what happens for the primordial fluctuations. The field modes obey (15), and this equation has, on super-Hubble scales, solutions that become dominant and solutions that become negligible (so-called "growing" and "decaying" modes). Eventually the decaying mode can be neglected and one in left with the growing mode. It turns out that f_{k2} and g_{k1} are decaying modes, and one is left with (26).

From the Heisenberg representation it follows that the operational equivalence with the classical stochastic system does not depend on the initial state; this was indeed shown explicitly for a wide class of initial states (and extended to some gauge-invariant quantities)[4].

We now look at the problem in the Schrödinger representation where the state evolves in time, while the operators are fixed. The initial quantum state of the perturbations is the vacuum state $|0,\eta_0\rangle$ satisfying

$$a_{\mathbf{k}}|0,\eta_0\rangle = 0 \qquad\qquad \forall \mathbf{k}\ . \qquad (27)$$

At later times, due to the creation of particles, the time-evolved state is annihilated by a more complicated operator,

$$\left\{y_{\mathbf{k}} + {\rm i}\gamma_k^{-1}(\eta) p_{\mathbf{k}}\right\}|0,\eta\rangle = 0\ . \qquad (28)$$

The corresponding (Gaussian) wave function reads

$$\Psi[y_{\mathbf{k}}, y^*_{\mathbf{k}}, \eta] = \frac{1}{\sqrt{\pi|f_k|^2}} \exp\left(-\frac{|y_{\mathbf{k}}|^2}{2|f_k|^2}\{1 - {\rm i}2F(k)\}\right)$$
$$\equiv \left(\frac{2\Omega_{\rm R}(\eta)}{\pi}\right)^{1/4} \exp\left(-[\Omega_{\rm R}(\eta) + {\rm i}\Omega_{\rm I}(\eta)]|y_{\mathbf{k}}|^2\right)\ . \qquad (29)$$

In (28,29), we have

$$\gamma_k \ = \ \frac{1}{2|f_k|^2}[1 - 2{\rm i}F(k)]\ ,$$
$$F(k) \ = \ \Im f_k^* g_k \ = f_{k1}g_{k2} - f_{k2}g_{k1}\ . \qquad (30)$$

At the initial time $\eta = \eta_0$, $\gamma_k(\eta_0) = k$, and hence $F(k) = 0$; in other words, we have a minimum uncertainty wave function. This is no longer so later, as $|F(k)|$ becomes very large; the probabilities, however, remain Gaussian. Another way to exhibit the physical meaning of our state is to consider the Wigner function, W, which can be considered as a kind of quasi-probability density in phase space. For Gaussian wave functions, W has the property to be positive definite. For the wave function (29) one obtains

$$W \stackrel{|r_k|\to\infty}{\longrightarrow} |\Psi|^2\ \delta^{(2)}\ (p_{\mathbf{k}} - p_{\rm cl}(y_{\mathbf{k}}))\ . \qquad (31)$$

The dynamics of the fluctuations leads to the large-squeezing limit $|r_k| \to \infty$. One gets a highly elongated ellipse whose large axis is oriented along the line $p_{\mathbf{k}} = p_{\rm cl}(y_{\mathbf{k}})$ and whose width becomes negligible. This is a direct vizualisation of the classical stochastic behaviour of our system: the variable $y_{\mathbf{k}}$ can take any value with corresponding probability $|\Psi|^2$, while $p_{\mathbf{k}}$ takes the corresponding value $p_{\mathbf{k}} = p_{\rm cl}(y_{\mathbf{k}})$. Instead of being essentially located in phase space around one physical trajectory, as for coherent states, the system behaves as if it followed an infinite number of classical trajectories with a definite probability to be on each of them. Interestingly, an analogous situation happens for a free non-relativistic particle[5] possessing an initial Gaussian minimal uncertainty wavefunction. As is well known, $F \propto t$ and becomes very large. At very late times, the position does no longer depend on the initial position,

$$x(t) \simeq \frac{p_0}{m} t\ . \qquad (32)$$

We get an equivalence with an ensemble of classical particles obeying (32), where p_0 is a random variable with probability $P(p_0) = |\Psi|^2(p_0)$. This illustrates the kind of classicality we are dealing with. Moreover, when (32) holds, position operators at different times approximately commute (which, in quantum-optical language, corresponds to a quantum-nondemolition situation).

Using the canonical commutation relations, the quantum coherence between the growing and decaying mode can be expressed as

$$f_{k1}g_{k1} + f_{k2}g_{k2} = \frac{1}{2} \ . \tag{33}$$

Clearly, when f_{k2}, g_{k1} are unobservable, this coherence becomes unobservable as well. This is the case when the decaying mode is so small that we have no access to it in observations. For the ratio of the growing to the decaying mode one has

$$\frac{f_{k2}}{f_{k1}} \propto e^{-2|r_k|} \ , \tag{34}$$

which is why a large squeezing parameter r_k in the Schrödinger picture implies a vanishing decaying mode in the Heisenberg representation. The width of the Wigner function is given by

$$\langle (p_{\mathbf{k}} - p_{\rm cl}(y_{\mathbf{k}}))^2 \rangle = g_{k1}^2 \ , \tag{35}$$

which becomes unobservable like the decaying mode. A further consequence is that the typical phase-space volume occupied by the system becomes negligible, too.

Let us take the concrete and important example of a perturbation on de Sitter space $a \propto {\rm e}^{Ht}$, with H being constant. The exact solution of (15) with the correct initial condition (ground state for initial sub-Hubble modes) then reads up to an unimportant *constant* phase factor

$$f_k \ = \ \frac{-{\rm i}}{\sqrt{2k}} \ {\rm e}^{-{\rm i}k\eta}\big(1 - \frac{\rm i}{k\eta}\big) \ , \tag{36}$$

$$g_k \ = \ -{\rm i}\sqrt{\frac{k}{2}}{\rm e}^{-{\rm i}k\eta} \ , \quad \eta \equiv -\frac{1}{aH} < 0 \ . \tag{37}$$

Modes initially inside the Hubble radius become much larger than the Hubble radius during inflation solely as a result of their dynamics to satisfy $k\eta \ll 1$: here we have the limit mentioned above! This can be shown also to correspond to the large-squeezing limit. Actually, this is a particular case of the general situation when an equation like (15) has a growing-mode solution and a decaying-mode solution. Here the decaying mode becomes vanishingly small; when it is neglected we are in the limit of a random stochastic process. Perturbations are then given by

$$\delta\phi(\mathbf{k}, \eta) = \frac{H}{\sqrt{2k^3}} \ e_{\mathbf{k}} \ . \tag{38}$$

We have set here $\sqrt{2k} \ y_{\mathbf{k}} = e_{\mathbf{k}}$, which assumes the role of a classical Gaussian random variable with unit variance. From (38) we see that the perturbations tend to a constant value (they become "frozen"). One should realize that the true reason for the quantum-to-classical transition in the sense discussed here is that the decaying mode becomes vanishingly small. Primordial gravitational waves follow exactly the behaviour (38) (up to some factor)[6], but after re-entering the Hubble radius they will start oscillating. They retain their classical appearance because the decaying mode (which oscillates as well by then!) is negligible[3].

IV. OBSERVATIONAL PREDICTIONS

The perturbations produced during inflation have remarkable properties which can be confronted with observations. This confrontation makes essential use of the effective classical behaviour discussed in the last section.

Primordial inflaton fluctuations generate a primordial Newtonian potential and the corresponding energy-density fluctuations $\delta\rho$. A central quantity is the power spectrum, $P(k)$, of the quantity $\delta \equiv \delta\rho/\rho$,

$$\langle \delta(\mathbf{k}) \ \delta^*(\mathbf{k}') \rangle = P(k) \ \delta^{(3)}(\mathbf{k} - \mathbf{k}') \ . \tag{39}$$

When the statistical properties are isotropic, the power spectrum depends only on $k \equiv |\mathbf{k}|$. It can be shown that the power spectrum is the Fourier transform of the correlation function (in space), and it can be defined for any quantity. Deep in the matter-dominated stage, $P(k)$ has the following expression on "super-horizon" scales in slow-roll single-field inflation,

$$P(k) = \frac{1024}{75}\pi^3 G^3 \left(\frac{V^3}{V'^2}\right)_{t_k} (aH)^{-4} \ k, \tag{40}$$

where V' is the derivative of the inflaton potential with respect to the inflaton ϕ, and the fraction has to be evaluated at the Hubble-radius crossing time $k = a(t_k)H(t_k)$ *during* inflation. Because of the quasi-exponential inflationary expansion, it depends very weakly on k. Neglecting this dependence, we get

$$P(k) \propto k \ , \tag{41}$$

which is the scale-invariant "Harrison–Zeldovich" spectrum that plays a crucial role in these investigations. This spectrum is called scale-invariant for the following reason: if we compute the r.m.s. relative mass fluctuations $\langle (\delta M/M)^2 \rangle$ at the time t_k when a scale eventually re-enters the Hubble radius, the same value is obtained for all scales.

Using the expansion (23) and the commutation relations (21), it is straightforward to show that

$$\langle \delta\phi^2 \rangle = \frac{1}{2\pi^2} \int_0^\infty dk \ k^2 \ |\delta\phi_k(\eta)|^2 \ , \tag{42}$$

with $f_k(\eta) = a \ \delta\phi_k(\eta)$. This means that the power spectrum of $\delta\phi$ is just given by $|\delta\phi_k(\eta)|^2$. However, the average on the left is a quantum average; it is only by virtue

of the quantum-to-classical transition mentioned above that we can consider $|\delta\phi_k(\eta)|^2$ as the power spectrum of a classical random variable, whose time evolution is consistent with probabilities conserved along classical trajectories. In the opposite case this would be impossible due to quantum interferences. We note also the result in the limit (38), which gives

$$\frac{d\langle\delta\phi^2\rangle}{d\ln k} = \left(\frac{H}{2\pi}\right)^2 , \qquad (43)$$

where the derivative is with respect to some cut-off value.

Primordial fluctuations leave their imprint on the CMB and this provides the best constraint on their properties and on the inflationary models in which they were presumably produced. While the CMB is remarkably homogeneous with a black body spectrum, perturbations induce very tiny inhomegeneities of the order 10^{-5}. In this regime, linear perturbation theory is very accurate so that precise predictions can be made. The measurement of the temperature anisotropies angular power spectrum, the C_ℓ's,

$$C_\ell = \langle|a_{lm}|^2\rangle , \quad \frac{\Delta T}{T}(\vartheta,\varphi) = \sum_{l,m} a_{lm}\, Y_{lm} , \qquad (44)$$

(which are in the isotropic case independent of m) will culminate with the Planck satellite (ESA). The exquisite data we have thus far, in particular those collected by the WMAP collaboration (NASA), show excellent agreement with a flat universe and adiabatic perturbations[7,8]. Such perturbations respect the equation of state of the background; for the baryon–photon plasma this is when $\frac{\delta T}{T} = \frac{1}{3}\frac{\delta n_B}{n_B}$, where n is the baryon number density. This is a natural outcome of single-field inflation.

Before decoupling, the baryon–photon plasma is tightly coupled and its density oscillates on scales inside the Hubble radius, yielding oscillations similar to pressure waves. These are often called acoustic oscillations. The location of the first (Doppler) peak gives roughly the angular scale of the Hubble radius at decoupling and is consistent with a flat universe. The pattern of the angular power spectrum is in agreement with primordial adiabatic fluctuations. After decoupling, the baryons retain the primordially induced acoustic "Sakharov" oscillations, the baryonic acoustic oscillations (BAO); these were detected in the galaxy power spectrum and are presently used in order to constrain dark energy models.

To parametrize the departure from scale invariance, one introduces the spectral index n with $P(k) \propto k^n$. Latest CMB data constrain n to be very close, but slightly lower than one[7]. Finally we see no clear evidence for non-Gaussianity in the statistics of the perturbations. All these data are in surprisingly good agreement with the simplest single-field slow-roll inflationary models (see e.g.[9]).

Let us return in more detail to the acoustic oscillations. They arise because of the standing-wave behaviour of the perturbations inside the Hubble radius. There are always two modes that are solutions to the equations and they will both oscillate. One of the modes matches the growing (dominant) mode, and the other the decaying (subdominant) mode. For modes sufficiently long outside the Hubble radius, the decaying mode disappears and the growing mode will match the corresponding oscillating mode inside the Hubble radius. At decoupling, each mode has a given oscillation phase, and this gives rise to the acoustic oscillations seen in the C_ℓ's. If we had a way to generate classical perturbations that would evolve outside the Hubble radius for very long, just the same would be true. If these perturbations had random initial conditions, obeying the same statistics as our initially quantum fluctuations, both systems would be indistinguishable. Hence the presence of acoustic oscillations is in no way connected to the quantum nature of the perturbations but rather to their primordial origin. But the quantum-to-classical transition can only take place in a system where the decaying mode is negligible enough so that acoustic oscillations *do* arise. It is interesting that a similar standing-wave behaviour is present in the primordial stochastic gravitational waves background produced during inflation. Unfortunately, to detect it in a direct detection experiment today would require a resolution in frequency of about 10^{-18} Hz,[3] clearly beyond present or foreseeable capabilities. The same property yields also small superimposed oscillations in the power spectra of the CMB temperature anisotropy and polarization. This is similar to the acoustic oscillations but with a period approximately twice as small (solely due to the difference between the light velocity and the sound velocity in the baryon–photon plasma at the recombination time)[3]. Their observation is very difficult but not hopeless if the parameter characterizing the tensor-to-scalar ratio in the CMB temperature anisotropy is not too small, see[10] for detailed estimates of the CMB polarization B-mode produced by primordial gravitational waves only.

We finally mention that calculations done for the creation of matter by parametric resonance after inflation use the description of perturbations in terms of classical stochastic fields. All the predictions mentioned above and which were confirmed by observations are done in the closed-system approach, that is, by taking the perturbations as an isolated system. Similar results were obtained in various disguise by several authors[11,12,13] and even extended beyond the linear regime[14]. In this approach the system becomes indistinguishable, in an operational sense, from a classical stochastic system solely by virtue of its peculiar inflationary dynamics.

From a purely pragmatic point of view, the closed-system approach is sufficient. In astrophysical observations one measures certain classical correlation functions for which the above line of thought shows that they are indistinguishable from the fundamental quantum expectation values. Still, in the next section we shall go beyond the closed-system approach by taking into account the interaction of the modes with other, "environmental",

degrees of freedom. This has several reasons. First, the environment-induced decoherence process is generally invoked in order to explain the appearance of classical behaviour in quantum theory[15]. Second, since an environment is expected to be present anyway, it is important to consider whether it does not spoil the successful predictions from the closed-system approach. It should, in particular, not erase the acoustic oscillations. Moreover, invoking large non-linear effects might irremediably modify the CMB angular power spectrum and induce large non-Gaussianity. Finally, there is the question about the entropy of the perturbations which by definition cannot be addressed inside the closed-system approach.

We shall see that these questions and problems can be successfully dealt with without spoiling the successful predictions of the closed-system approach including the quantum-to-classical transition in the pragmatic approach adopted in this section.

V. QUANTUM-TO-CLASSICAL TRANSITION: DECOHERENCE

A. Decoherence and pointer basis

In the last section we have described the primordial fluctuations in cosmology by a collection of independent quantum states labelled by the wave number k. Since no interaction between different k or between the fluctuations and other fields have been considered, we deal with a pure quantum state for each k. The initial condition for each quantum state is the harmonic-oscillator ground state with respect to k. During inflation, modes with wavelengths larger than the Hubble scale H^{-1} assume a squeezed Gaussian state. We focussed attention on the modes far outside the Hubble scale, which experience an enormous squeezing. For these highly-squeezed modes, which are the ones relevant for cosmological observations, all expectation values containing the field-amplitudes or their momenta are indistinguishable from classical stochastic averages[3]. It is this approximate coincidence between quantum and classical expectation which is the basis of the pragmatic approach to the quantum-to-classical transition discussed above for the primordial fluctuations.

One can, however, adopt a more fundamental point of view. It is far from realistic to assume that a primordial fluctuation with wave number k is exactly isolated. We must take into account its interaction with other degrees of freedom (called the 'environment' for simplicity). The main reason is the following. As one knows from standard quantum theory, even a tiny interaction with other degrees of freedom can become important, in the sense that an entanglement of a system with its environment can form even without direct disturbance of the system. If the environmental degrees of freedom are inaccessible to observations (as they usually are), the ensuing entanglement with the system leads to *decoherence* – interference terms can no longer be observed at the system itself and the system *appears* classical[15]. This is the fundamental origin of the quantum-to-classical transition. The phenomenon of decoherence is by now theoretically well understood and has been experimentally tested with high precision[15,16,17]. Decoherence leads to an *apparent ensemble* of wave packets for the observable with respect to which the interferences vanish. A paradigmatic example is the localization of a quantum particle due to scattering with photons, air molecules, or other particles[15,17,22]. There the position basis of the particle is the approximate basis distinguished by the scattering process. The basis distinguished by the environment is generally called the *pointer basis*; the corresponding observable is called pointer observable. Interferences between different members of the pointer basis are suppressed by the decohering influence of the environment.

One would expect, therefore, that decoherence is of crucial importance for the primordial fluctuations, too. This expectation is, moreover, supported by the fact that the system by itself evolves into a highly squeezed state in which squeezing is in the field momentum and broadening is in the field amplitude (corresponding to the position variable in quantum mechanics): one knows from quantum theory that highly squeezed states are extremely sensitive to any environment[15]. This is the reason why they are so difficult to generate in the laboratory – it is very hard to isolate them from any environment. In view of their huge squeezing, this argument should apply to the cosmological fluctuations *a fortiori*.

But could it be imaginable that the cosmological fluctuations, in contrast to a typical quantum-mechanical situation, are indeed strictly isolated? The answer is definitely no.

Firstly, in any fundamental theory (such as string theory) there is an abundance of different fields with different interactions. Among them it will not be difficult to find appropriate candidates for environmental fields generating decoherence for the primordial fluctuations.

Secondly, even if one assumes to have no such fields, there are two processes which cannot be neglected. The first one is the interaction between modes with different k; recall that the full theory is non-linear and that, therefore, the various modes cannot be treated independently of each other. Such non-linear interactions concern both the interaction with the modes of the inflaton and the perturbations of the metric (containing, in particular, gravitational waves).

The second process is the entanglement of the modes' quantum state between different *spatial regions*: even if the modes are independent in k-space, the Gaussian wave functions for the amplitudes in real space are highly correlated over spacelike regions (as in the Einstein–Podolsky–Rosen situation). This leads, in particular, to an entanglement between the regions inside and outside the Hubble radius. Famous non-cosmological examples are the Hawking and the Unruh effects, where the thermal appearance of the corresponding radiation

can be understood from the entanglement between inside and outside the event horizons and the tracing out of the correlations into the horizon[18]. Even for spacelike surfaces which stay outside the horizon, the thermal nature of Hawking and Unruh radiation can be understood from the entanglement with other fields, leading to decoherence[19].

The process of decoherence is, moreover, needed to justify the results from the isolated (closed) system in the first place. Even if the classical and quantum expectation values are indistinguishable, the presence of a pure state means that one has a quantum superposition of all possible field amplitudes, *not* an ensemble of stochastically distributed classical values. This situation is similar to Schrödinger's cat. In the pragmatic point of view of Sec. III, the approximate coincidence of the expectation values suffices. Such a coincidence is, however, not sufficient for a realistic interpretation. Only decoherence can eventually justify the pragmatic point of view in that it leads to an apparent ensemble of wave packets for the system variables itself (which, in our case, are the field amplitudes). The insufficience of approximately equal classical and quantum expectation values for a fundamental interpretation has recently been clearly emphasized in a different context (the quantum mechanics of classically chaotic systems) by Schlosshauer[20]. In the presence of a pure state one can always find an observable for which no classical counterpart exists, that is, for which the comparison of quantum and classical expectation values is meaningless.

The quantum-to-classical transition happens for the highly-squeezed modes whose wavelengths exceed the Hubble scale. It is for these modes where environmental decoherence is most efficient[21]. How can this happen? Would one not expect that no causal interaction can occur on scales larger than the Hubble scale? This is true only for a direct disturbance of the system. But the crucial point is that quantum entanglement can form without direct disturbance. And this is all one needs for decoherence! In the context of the quantum measurement process, the sole formation of entanglement is referred to as an 'ideal measurement' or a 'quantum non-demolition measurement': the system remains undisturbed, but the environment is affected through the formation of entanglement. The general mechanism is as follows[15].

Consider a quantum system which is initially in the state $|n\rangle$ and a 'measurement device' (here: the environment) which is in some initial state $|\Phi_0\rangle$. (We assume that $|n\rangle$ belongs to a set of eigenstates of a system observable.) The evolution according to the Schrödinger equation is in the special case of an 'ideal measurement' given by

$$|n\rangle|\Phi_0\rangle \stackrel{t}{\longrightarrow} \exp\left(-\mathrm{i}H_{\mathrm{int}}t\right)|n\rangle|\Phi_0\rangle = |n\rangle|\Phi_n(t)\rangle \ , \quad (45)$$

where H_{int} denotes the interaction Hamiltonian (assumed here to dominate over the free Hamiltonians) which correlates the system state with its environment without changing the system state.

In the general case, the quantum system can be in a superposition of different eigenstates of the system observable. Then, due to the linearity of the time evolution, an initial product state with $|\Phi_0\rangle$ develops into an entangled state of system plus apparatus,

$$\left(\sum_n c_n|n\rangle\right)|\Phi_0\rangle \stackrel{t}{\longrightarrow} \sum_n c_n|n\rangle|\Phi_n(t)\rangle \ . \quad (46)$$

But this is a highly non-classical state! Since the environmental states $\{|\Phi_n\rangle\}$ are not accessible, they have to be traced out from the full quantum state. One thereby arrives at the reduced density matrix ρ_{S} which contains all the information that is available at the system itself. Since the environmental states $\{|\Phi_n\rangle\}$ can be assumed as being approximately orthogonal (otherwise they would not be able to serve as a 'measurement device'), the reduced density matrix is of the form

$$\rho_{\mathrm{S}} \approx \sum_n |c_n|^2|n\rangle\langle n|, \quad (47)$$

that is, it assumes the form of an *approximate ensemble* for the various system states $|n\rangle$, each of which occurs with probability $|c_n|^2$.

In our case, the cosmological fluctuations represent the system to be decohered. The environmental states $\{|\Phi_n\rangle\}$ can be other fields or inaccessible parts of the fluctuations themselves (see below). The system states $|n\rangle$ are given by the field-amplitude states $|y_{\mathbf{k}}\rangle$. The interaction with the environment can, in the ideal-measurement case, be described by the multiplication of an initial density matrix $\rho_0(y,y')$ with a Gaussian factor in $y-y'$ (omitting here and in the following the index $\mathbf{k}$ in $y_{\mathbf{k}}$),

$$\rho_0(y,y') \longrightarrow \rho_\xi(y,y') = \rho_0(y,y')\exp\left(-\frac{\xi}{2}(y-y')^2\right). \quad (48)$$

Here, the parameter ξ encodes the details of the interaction between the modes and their environment. Given a specific model with a specific interaction, ξ can be calculated. The special decoherence process (47) is typical for the description of localization in quantum mechanics[15,17,22].

One recognizes from (48) that interferences between different values of the field amplitude y have been suppressed by interaction with the environment. This is decoherence. So far we have just assumed without derivation that $|y\rangle$ is the pointer basis, that is, the relevant robust system basis which is distinguished by the environment. This must, of course, be justified. A detailed derivation for the field-amplitude basis to be the pointer basis has been presented in[21] and[25]. We review here the main arguments and refer the reader to these references for more details.

According to the classical equations, for modes with very large wavelength one has $y \propto a$, that is, the physical fluctuations $\delta\phi$ are approximately constant ('frozen'). In the Heisenberg picture of the quantum theory, this

means that the operator $\widehat{\delta\phi}$ approximately commutes with the Hamiltonian. Now comes the crucial point. Additional (environmental) fields coupling with the cosmological fluctuations are expected to couple field amplitudes, not canonical momenta of field amplitudes; that is, the coupling is expected to involve $\widehat{\delta\phi}$, *not* its momentum. Consequently, the fluctuations $\widehat{\delta\phi}$ commute with the *whole* Hamiltonian of system plus environment. Such a variable is a pointer observable par excellence[15,16,17]. It is stable (robust) in time because of this commutativity which holds for the wavelengths much bigger than the Hubble scale. The phenomenological expectation (48) is thus fully justified. One must keep in mind, though, that $\widehat{\delta\phi}$ is only an approximate pointer observable: although the non-diagonal terms in (48) become exponentially suppressed, they never vanish exactly, as would be the case if the $\widehat{\delta\phi}$ were the exact pointer observable. In fact, the reduced density matrix can be decomposed into narrow Gaussians in $\delta\phi$-space. The whole situation is in strong analogy to the localization of a massive particle by scattering with the environment[15,17,22].

The approximate commutativity of $\widehat{\delta\phi}$ with the full Hamiltonian means in particular that the kinetic term, that is, the p^2-term, of the system becomes irrelevant in the large-squeezing limit. If this term were relevant (as it is for modes with smaller wavelength), the pointer basis would not be the field-amplitude basis, but the coherent-state basis[25]. But this is not the case here. The coherent-state basis is, in particular, unstable under the time evolution.

So far we have restricted our attention to a special initial state: the vacuum state. This is, however, not necessary. In[25] we have presented a formalism that is general enough to encompass a wide range of initial states and interactions. A central role in this formalism is played by a master equation for the reduced density matrix, which is of the Lindblad form. More concretely, the density matrix is assumed to satisfy[15]

$$\frac{\mathrm{d}\hat{\rho}}{\mathrm{d}t} = -\mathrm{i}[\hat{H},\hat{\rho}] + \hat{L}\hat{\rho}\hat{L}^{\dagger} - \frac{1}{2}\hat{L}^{\dagger}\hat{L}\hat{\rho} - \frac{1}{2}\hat{\rho}\hat{L}^{\dagger}\hat{L}\ , \qquad (49)$$

where $\hat{L}$ is the Lindblad operator. Most of the particular models discussed in the literature lead to a master equation of this form. It is thus of interest to study this equation as general as possible. We have assumed that the Lindblad operator is linear in our variables p and y, but kept it general otherwise. The Hamiltonian $\hat{H}$ is given by the expression (17).

The results of our discussion in[25] can be summarized as follows. It turns out that the behaviour of the master equation is qualitatively different for modes outside the Hubble radius (as is the case here) and the modes inside. The decoherence time t_{d} for the modes with wavelengths much bigger than the Hubble radius is during inflation of the order

$$t_{\mathrm{d}} \sim H_{\mathrm{I}}^{-1} \ln \frac{H_{\mathrm{I}}^{-1}}{t_0}, \qquad (50)$$

where H_{I} is the (approximately constant) Hubble parameter during inflation, and t_0 is a typical time characteristic for the details of the interaction. We emphasize that (50) is approximately independent of these details. It is basically given by the Hubble time, with the details only entering logarithmically. The time t_{d} also gives the timescale for the Wigner function to become positive. The reduced density matrix can then be decomposed into an apparent ensemble of narrow Gaussians for the values of the field amplitude, cf.[26] for a general discussion. For the large-wavelength modes in the radiation-dominated phase one obtains instead

$$t_{\mathrm{d}} \sim \frac{H_{\mathrm{I}} t_{\mathrm{L}}^2}{2}, \qquad (51)$$

where t_{L} depends again on the details of the interaction. One has now a more sensitive dependence on the interaction. Moreover, for $H_{\mathrm{I}} t_{\mathrm{L}} \gg 1$ one has a much longer decoherence time than during inflation. This means that, depending on the interaction, decoherence can be much less efficient than during inflation.

For modes smaller than the Hubble scale, the situation is very different[25]. Taking as a representative example a photon bath as the environment (realized e.g. by the CMB), the decoherence time is independent of the Hubble parameter and strongly dependent on the coupling to the bath. Dissipation now becomes the dominant source of influence, in contrast to the case of the super-Hubble modes for which only entanglement occurs.

Decoherence is often connected with symmetry breaking[15], see also[27], section 6.1. This is also the case here. The initial de Sitter-invariant vacuum state for the fluctuations is highly symmetric. But the observed classical fluctuations are certainly non-symmetric. This can easily be understood and does not require new physics (as e.g. demanded in[28]). The initial vacuum state develops into a squeezed vacuum, which can be understood as a superposition of different field-amplitude eigenstates. Decoherence then makes this indistinguishable from an ensemble of (approximate) field-amplitude eigenstates, each of which is highly inhomogeneous. The situation resembles the case of spontaneous symmetry breaking in field theory, where the symmetric initial state evolves into a superposition of 'false vacua'. After decoherence one is left with an apparent ensemble of different false vacua, one of which corresponds to our observed world.

B. Entropy

In Sec. III the primordial fluctuations were treated as isolated and thus described by a pure (squeezed) state. Consequently, they possess zero entropy: all information is contained in the system itself. But as we have seen, the primordial fluctuations are an *open* quantum system; they are entangled with their environment. Because of this entanglement, the fluctuations are described by the reduced density matrix (48). They thus possess positive

entropy because the information about the correlations with the environment are unavailable in the system itself. The local entropy is calculated from the standard von Neumann formula,

$$S = -\mathrm{tr}(\rho_\xi \ln \rho_\xi) \ , \tag{52}$$

where ρ_ξ is given in (48), and where we have set $k_\mathrm{B} = 1$. Considering one (real) mode with wave number k, the maximal entropy, S_max, would be $2r_k$, where r_k is again the squeezing parameter[29] (we skip again the index k in the following). We have calculated and discussed the entropy for the fluctuations in[23,25]. To display the result, it is convenient to introduce the dimensionless parameter $\chi = \xi/\Omega_\mathrm{R}$, where Ω_R is the width of the Gaussian (29); it controls the strength of decoherence. (In the case of pure exponential inflation one has $\chi = \xi(1+4\sinh^2 r)/k$.) Inserting (48) into (52), one gets the explicit expression[25]

$$\begin{aligned} S &= -\ln\frac{2}{\sqrt{1+\chi}+1} - \frac{1}{2}\left(\sqrt{1+\chi}-1\right)\ln\frac{\sqrt{1+\chi}-1}{\sqrt{1+\chi}+1} \\ &= \ln\frac{1}{2}\sqrt{\chi} - \sqrt{1+\chi}\ln\frac{\sqrt{1+\chi}-1}{\sqrt{\chi}} \ . \end{aligned} \tag{53}$$

One recognizes that the entropy vanishes for $\xi \to 0$, as it must for a pure state. In the limit $\chi \gg 1$ (large decoherence) one gets

$$S = 1 - \ln 2 + \frac{\ln\chi}{2} + \mathcal{O}(\chi^{-1/2}) \ . \tag{54}$$

This asymptotic value is readily attained.

As we have emphasized above, modes with wavelength bigger than the Hubble scale can only experience pure entanglement, not direct disturbance. In such a case the entropy obeys the bound

$$S < \frac{S_\mathrm{max}}{2} = r \ . \tag{55}$$

The same bound follows from the general discussion of the Lindblad equation[25]. It can also be interpreted in the following way[25]: in spite of decoherence, some squeezing compared to the vacuum state (which has $\Omega_\mathrm{R} = k$) remains. In the language of the Wigner function it means that the Wigner ellipse is not smeared out to become a circle, but still exhibits an elongated and a squeezed part. And this has important consequences for observation! If the bound (55) were violated, there would no longer be any coherences between the field amplitude and the momentum and, consequently, no coherences in the coupled baryon–photon plasma (Sec. IV). There would then not be any acoustic peaks in the anisotropy spectrum of the CMB – in contrast to observation! The fundamental questions of the quantum-to-classical transition have thus observational relevance.

The upper bound $S_\mathrm{max}/2$ corresponds to the case when the pointer basis is the exact field-amplitude basis. (For $S = S_\mathrm{max}$, the pointer basis would be the particle-number basis.) As our pointer basis consists of narrow packets in field amplitudes, the entropy of the fluctuations approaches the upper bound asymptotically.

The existence of the bound (55) shows, again, how peculiar the case of fluctuations in an inflationary universe is. According to a theorem by Page[30] (see also[31]), if a total quantum system with dimension mn is in a random pure state, the average entropy of a subsystem of dimension $m \leq n$ is almost maximal. But this is not the case for our system: the situation for the fluctuations during inflation *is* very special, and their entropy cannot exceed half of the maximal entropy, which leaves enough information for the formation of the acoustic peaks.

Our results for the entropy in[23] and[25] also yield the following simple formula for the entropy production during inflation:

$$\dot{S} \approx \dot{r} \approx H_\mathrm{I} \ . \tag{56}$$

For chaotic systems, the entropy production rate is proportional to the Lyapunov parameter. This would correspond in our case to the Hubble parameter H_I. However, our system is not chaotic, but only classically unstable, so the analogy is not complete.

Using (50), one can find the amount of entropy produced after the decoherence time t_d,

$$S \sim H_\mathrm{I} t_\mathrm{d} \sim \ln\frac{H_\mathrm{I}^{-1}}{t_0} \ . \tag{57}$$

In the radiation-dominated phase following inflation, a relation similar to (56) holds, with H_I replaced by the Hubble parameter $H \propto t^{-1}$. The entropy thus only increases logarithmically in time, not linearly as in inflation.

C. Specific models

So far, we have kept the discussion as general as possible. We have reviewed the arguments which lead to the result that cosmological fluctuations appear like a classical ensemble of field amplitudes. Necessary requirements are the inflationary expansion of the universe and the focus on modes that are highly squeezed. An interaction with some environment is needed, but the details of it are unimportant. Still, it is of interest to discuss specific examples for such interactions. Our paper[25] gives an extended list of references; here we shall restrict ourselves to some recent examples.

The purely spatial entanglement between the modes inside the Hubble scale and outside the Hubble scale was discussed in[32], see also[33]. It was shown there that this entanglement is, by itself, sufficient to produce the desired decoherence. This is analogous to the black-hole case where the decoherence from the tracing out of the modes behind the horizon leads to the thermal radiation of the Hawking effect[18,19]. The authors of[32] also showed that the entropy scales with the volume inside

the Hubble scale and satisfies an upper bound of $S \approx r$ per mode, which coincides with the upper bound (55) discussed above. It is thus not in conflict with the observed acoustic peaks in the cosmic microwave background.

Instead of pure spatial entanglement one can consider the entanglement of our strongly squeezed super-Hubble modes with sub-Hubble modes (which then play the role of the environment). This was discussed, for example, in[34]. The authors take the short-wavelength modes to be in their ground states and find that decoherence is not sufficient during inflation. This happens because vacuum states are usually ineffective to lead to decoherence[15]. Our arguments above and in[25] can thus only be applied to this model if at least some modes are not in their ground states. But such modes can be found: one can interpret the fluctuations with wavelengths $\lambda \gtrsim H_{\rm I}^{-1}$ as an appropriate environment; they assume a role intermediate between ground state and state with large squeezing. Ideas similar to the ones in[34] have been pursued in[35,36], and elsewhere, with results that are consistent with our general discussion above. A variant of this system-environment split is presented in[37] using a two-field model of inflation. There, the system consists of curvature perturbations, and the environment consists of isocurvature modes. Finally, another possible source of sub-Hubble modes being in non-vacuum states is the *secondary* gravitational wave background ("foreground" in astronomical terminology) emitted by matter after the end of inflation[23].

VI. CONCLUSION

Inflation is a robust scenario which gives an elegant solution to some oustanding problems of Big-Bang cosmology, and its predictions are in agreement with present observations, in particular the accurate CMB anisotropy data. It is gratifying that this scenario offers also the possibility to deal with such fundamental and subtle questions as to why quantum perturbations produced in the early Universe give rise to classical inhomogeneities today. We believe that this aspect is no less fascinating than its other successful predictions.

We expect that models of the quantum-to-classical transition for the primordial fluctuations will continue to appear in the literature. But we are convinced that the general mechanism of this transition presented in this essay will hold true for all scenarios based on inflation.

Acknowledgements

We are happy to thank our collaborator Alexei Starobinsky for his crucial input in obtaining the results presented here and for his comments on our manuscript.

We kindly acknowledge financial support from The Foundational Questions Institute (http://fqxi.org) for the visit of D. P. to the University of Cologne.

[1] A. R. Liddle and D. H. Lyth, Cosmological inflation and large-scale structure, Cambridge University Press, Cambridge **(2000)**.
[2] L. P. Grishchuk and Y. V. Sidorov, Class. Quantum Grav. 6, L161 **(1989)**.
[3] D. Polarski and A. A. Starobinsky, Class. Quantum Grav. 13, 377 **(1996)**.
[4] J. Lesgourgues, D. Polarski, and A. A. Starobinsky, Nucl. Physics B 497, 479 **(1997)**.
[5] C. Kiefer and D. Polarski, Ann. Phys. (Leipzig) 7, 137 **(1998)**.
[6] A. A. Starobinsky, JETP Lett. 30, 682 **(1979)**.
[7] E. Komatsu, J. Dunkley, M. R. Nolta, C. L. Bennett, B. Gold, G. Hinshaw, N. Jarosik, D. Larson, M. Limon, L. Page, D. N. Spergel, M. Halpern, R. S. Hill, A. Kogut, S. S. Meyer, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright, arXiv:0803.0547v1 **(2008)**.
[8] D. Baumann and H. Peiris, Cosmological Inflation: Theory and Observations, contribution to this issue.
[9] W. H. Kinney, E. W. Kolb, A. Melchiorri, and A. Riotto, arXiv:0805.2966v2 **(2008)**; J. Lesgourgues, A. A. Starobinsky, and W. Walkenburg, JCAP 0801:010 **(2008)**.
[10] J. Lesgourgues, S. Prunet, and A. A. Starobinsky, Astron. Astrophys. 359, 414 **(2000)**.
[11] A. H. Guth and S.-Y. Pi, Phys. Rev. D 32, 1899 **(1985)**.
[12] D. H. Lyth, Phys. Rev. D 31, 1792 **(1985)**.
[13] A. Albrecht, P. Ferreira, M. Joyce, and T. Prokopec, Phys. Rev. D 50, 4807 **(1994)**.
[14] D. H. Lyth and D. Seery, Phys. Lett. B 662, 309 **(2008)**.
[15] E. Joos, H. D. Zeh, C. Kiefer, D. Giulini, J. Kupsch, and I.-O. Stamatescu, Decoherence and the appearance of a classical world in quantum theory, second edition, Springer, Berlin **(2003)**. See also www.decoherence.de.
[16] W. H. Zurek, Rev. Mod. Phys. 75, 715 **(2003)**.
[17] M. Schlosshauer, Decoherence and the Quantum-to-Classical Transition, Springer, Berlin **(2008)**.
[18] W. Israel, Phys. Lett. A 57, 107 **(1976)**.
[19] C. Kiefer, Class. Quantum Grav. 18, L151 **(2001)**.
[20] M. Schlosshauer, arXiv:quant-ph/0605249v3 **(2008)**.
[21] C. Kiefer, D. Polarski, and A. A. Starobinsky, Int. J. Mod. Phys. D 7, 455 **(1998)**.
[22] E. Joos and H. D. Zeh, Z. Phys. B 59, 223 **(1985)**.
[23] C. Kiefer, D. Polarski, and A. A. Starobinsky, Phys. Rev. D 62, 043518 **(2000)**.
[24] C. Kiefer, J. Lesgourgues, D. Polarski, and A. A. Starobinsky, Class. Quantum Grav. 15, L67 **(1998)**.
[25] C. Kiefer, I. Lohmar, D. Polarski, and A. A. Starobinsky, Class. Quantum Grav. 24, 1699 **(2007)**.
[26] L. Diósi and C. Kiefer, J. Phys. A 35, 2675 **(2002)**.
[27] H. D. Zeh, The Physical Basis of the Direction of Time, fifth edition, Springer, Berlin **(2007)**.
[28] A. Perez, H. Sahlmann, and D. Sudarsky, Class. Quantum Grav. 23, 2317 **(2006)**.
[29] T. Prokopec, Class. Quantum Grav. 10, 2295 **(1993)**.
[30] D. N. Page, Phys. Rev. Lett. 71, 1291 **(1993)**.

[31] P. Hayden, D. W. Leung, and A. Winter, arXiv:quant-ph/0407049v2 **(2005)**.
[32] J. W. Sharman and G. D. Moore, JCAP 0711:020 **(2007)**.
[33] Y. Nambu, Phys. Rev. D 78, 044023 **(2008)**.
[34] C. P. Burgess, R. Holman, and D. Hoover, Phys. Rev. D 77, 063534 **(2008)**.
[35] P. Martineau, Class. Quantum Grav. 24, 5817 **(2007)**.
[36] F. C. Lombardo and D. López Nacir, Phys. Rev. D 72, 063506 **(2005)**.
[37] T. Prokopec and G. I. Rigopoulos, JCAP 0711:029 **(2007)**.

Thomas Banks

Statement

and

Readings

Quantum mechanics is widely regarded as mysterious, even among physicists. Yet no one has ever proposed a "more sensible" theory, which explains quantum mechanics in terms that its detractors find more easy to accept. Like special relativity it is a theory whose conclusions must be accepted because they conform to experimental reality, even though they seem to violate our intuition. With relativity we've come to accept that our intuitions about how velocities add have no more validity than Aristotle's "obvious" conclusion that matter's natural state of motion is to be at rest. Quantum mechanics is harder to accept because it violates the rules of classical logic that we long believed to be at the basis of all rational thought.

In my experience there are several simple ideas, all long extant in the literature, which help one to come to terms with the realities of quantum physics. The first is called *quantum logic*, a frightening term that obscures a rather simple idea: the mathematical formulation of the laws of logic, due to George Boole, contains within itself an obvious generalization. This mathematical formulation is simply the theory of a maximal set of commuting operators in Hilbert space, and once one has constructed it one cannot avoid the fact that there are Hermitian operators which do not commute with the chosen set, and that linear algebra provides us with a formula for a probability distribution for the eigenvalues of those operators, for each choice of sharp values of the maximal commuting set.

What is peculiar here is the appearance of probability in a situation which seems to have no uncertainty. We have measured, with the theorist's absolute precision, the maximal amount of information that we can extract from the system. Probability's original appearance in physics and gaming, was as a way of estimating the chances of something happening in a situation where we were ignorant of some of the details of the initial conditions. It plays a rather different, and much more intrinsic role here.

My main point though is that this mathematics is simply unavoidable. The non-commuting operators are there, even if one decides not to discuss them. Indeed, Koopman long ago showed how to reformulate classical mechanics as a special case of quantum mechanics. One simply introduces position and momentum as commuting quantum variables, each with its own canonical conjugate, $\pi_{p,q} = -i\partial_{p,q}$. If we take the quantum Hamiltonian to be the Liouville operator on phase space

$$H = i(\partial_p E \partial_q - \partial_q E \partial_p),$$

where E is the classical Hamiltonian function of the system, then, *as long as we agree not to talk about measurements of* $\pi_{p,q}$ the predictions of this quantum system agree with the predictions of classical mechanics, with initial phase space probability density

$$\rho = \psi^*(p,q)\psi(p,q).$$

The probabilistic aspect of quantum mechanics here reduces to the conventional uncertainty in initial conditions. The italicized covenant prevents us from finding out about the behavior of the phase of the wave function. Nonetheless, the theory makes predictions about hypothetical measurements of operators that do not commute with the phase space coordinates.

So, the precise, and therefore mathematical, formulation of what we mean by *logic*, simply forces quantum mechanics, with its intrinsic probabilities, not tied to the uncertainties of initial conditions, upon our attention. The world has used it, even if we would rather not. Once one has accepted that *quantum mechanics is an intrinsically probabilistic theory* the "mystery of wave function collapse" becomes familiar, rather than mysterious. Our probabilistic theory of the weather is often unable to tell whether a hurricane will hit Galveston or New Orleans. Meteorologists routinely solve the equations, *and throw out that part of the solution that predicts a disaster in Galveston, when they see people suffering in a sports stadium in New Orleans.* This is called the method of conditional probabilities. Physical theories are a means for predicting the future. When we are forced to use probabilities in our equations, some of these predictions are inevitably wrong, because we consider probabilities for two mutually exclusive macroscopic events. For convenience, we then incorporate the observations of misery in New Orleans into our equations, by defining a new probability for future events, conditional upon the fact that the hurricane hit New Orleans. This is "an artificial rule imposed on the equations", which "violates conservation of probability" (because the new conditional probability distribution predicts that New Orleans was hit with certainty). The "distribution collapses", and if we regarded the distribution as a physical field, because it satisfies *e.g.* the Fokker-Planck equation, we could expend reams of paper and giga-bytes of tortured prose exploring the philosophical implications of the conditional probability rule.

Much of the confusion about QM in the literature comes from trying to attribute the same sort of reality to the quantum wave function that we are used to attributing to Maxwell's electromagnetic fields. Real classical fields are coherent states of a huge number

of bosons, all in the same quantum state. They attain their reality (which means simply that we can make predictions about them, which have no uncertainties) because they are collective variables of macroscopic systems, whose quantum fluctuations are negligible for all practical purposes. Attempts to attribute the same sort of reality to wave functions of individual particles, as exemplified by sentences like, "The electron *is* its wave function, and satisfies a deterministic equation of motion", are conceptually wrong, because they try to hide the intrinsically probabilistic nature of QM.

Apart from the fact that probabilities are not only a consequence of uncertainty in initial measurements, QM differs from the equations of the weather in an important respect. The rule for computing probabilities involves superposition of amplitudes rather than of probabilities themselves. Some of the unease with wave function collapse has to do with the fact that this aspect of QM seems to disappear after a measurement is performed. Although this has long been understood by experts, the essence of it deserves more popular recognition. A measurement in QM consists, as von Neumann proposed long ago, of an interaction that puts a micro-variable like the state of an electron spin, into QM entanglement with *a collective coordinate of a macroscopic system called the measuring apparatus*. The phrase collective coordinate refers to some sort of an average value over a huge number of micro-states. Roughly speaking, if the system contains N atoms, there are e^{cN} microstates with the same value of the collective coordinate. In normal circumstances the microscopic state of the system wanders over this huge space of states, constantly changing its phase, without in any way altering the value of the collective coordinate. The wave function overlaps, which lead to the deviation of the quantum laws of probability from their classical counterparts, are doubly exponentially small (I use the phrase *doubly exponentially small* to denote numbers whose natural log is negative and of order 10^n with $n > 3$.). For even modest values of N 10^4, the time that it would take to see the quantum interference effects is essentially the same number when expressed in Planck times, as in ages of the universe (the two units differ by a mere factor of 10^{61}). We call this effective erasure of quantum phase interference, *decoherence*.

In making these statements, we have completely neglected the coupling of the micro-system plus measuring apparatus to their vast external environment. In practice the decoherence due to coupling with the environment is an even larger effect, and makes it even harder to see interference between states of macroscopic collective variables. Environmental decoherence is one of the practical barriers to building a quantum computer. Only when the macrosystem is kept in its ground state throughout the course of the experiment, can

we hope to see quantum interference between states of collective coordinates. This has been achieved for currents in SQUIDs, for systems consisting of $\sim 10^8$ atoms. Note that the quantum fluctuations of collective variables in the ground state are still small, but they vanish like powers, rather than exponentials of the number of atoms. The textbook explanation of why spreading of the wave function is negligible for basketballs or the moon, emphasizes these powers, while the exponential suppression due to decoherence is never mentioned.

As an example consider Schrödinger's infamous exercise in cruelty to animals. Once the intrinsically probabilistic interpretation of the wave function, and the simple facts of decoherence are taken into account, statements like "quantum mechanics predicts that the cat is alive and dead at the same time", are revealed to be silly misconceptions. Quantum mechanics simply predicts a probability that the cat will be found alive or dead in any given experiment, a prediction that can only be tested by subjecting many poor beasts to savage and useless experiments. Furthermore, *as a consequence of decoherence*, quantum mechanics predicts that once a particular cat is found dead, there is no probability that future experiments will find it alive. More precisely, that probability is ridiculously small, of order the probability for the atoms in the dead cat to spontaneously reassemble themselves to make the live cat.

These simple arguments show that our unease with QM is not so far removed from the unease Aristotle would have shown when confronted with the principle of inertia invented by Galileo and Newton. If we accept, following Darwin and Wallace, that the brain evolved, then the forces that shaped its evolution had to do with the behavior of collective coordinates of macroscopic systems. Watch that rock! Beware the jaws of that tiger! Jump for that low hanging fruit, but don't break your leg! These are the evolutionary pressures that led to the differential survival probability of those proto-humans who had the intelligence to formulate intuitions about some of the laws of physics. There was never a selection pressure for understanding that the distribution of electrons passing through two microscopically spaced slits, did not follow the same laws of probability that worked perfectly well for rocks and tigers. Indeed, our considerations above show that the likelihood that special relativistic corrections to Newtonian physics would come to our "evolutionary attention", is doubly exponentially larger than that for quantum mechanics. Perhaps this accounts for the remarkable longevity of even Nobel prize winning physicists' unease with the principles of quantum mechanics. About the only way I can think of of producing selection pressures that could lead to creatures with an intuitive understanding

of quantum mechanics, is to make the right to reproduce contingent on getting tenure in a physics department. The obvious political difficulties of such a proposal, are far from its worst aspect.

In summary, quantum mechanics is a radical departure from all previous physical theories because it abandons the hitherto sacred principles of classical logic for a more general, but less intuitive, set of rules. Yet the mathematical formulation of the rules of logic shows us that quantum mechanics is an unavoidable part of its structure. While we can artificially construct special systems, which allow one to avoid discussing non-commutative observables, one should not be surprised if nature has taken another path. The only puzzle is why our brains are not built to recognize the real quantum rules, and that is simply explained by the quantum theory of macroscopic objects. Quantum interference between states of the collective coordinates of such objects is, under almost all circumstances, an effect which is of order $\exp(-10^{20})$ or smaller. Such doubly exponentially small numbers defeat even the the most precise experiments we can conceive, and were certainly irrelevant to the macroscopic events that shaped the evolution of our brains. This understanding of the mysteries of quantum mechanics does not require us to make any kind of mystical connection between consciousness and the fundamental laws of physics. It only requires us to understand that our consciousness is an evolved collective variable of a macroscopic system, which obeys the rules of local quantum theory.

hep-th/yymmnnn
SCIPP-2006/13

Locality and the classical limit of quantum systems

Tom Banks

Department of Physics and SCIPP
University of California, Santa Cruz, CA 95064
E-mail: banks@scipp.ucsc.edu
and
Department of Physics and NHETC, Rutgers University
Piscataway, NJ 08540

Abstract

I argue that conventional estimates of the criterion for classical behavior of a macroscopic body are incorrect in most circumstances,because they do not take into account the locality of interactions, which characterizes the behavior of all systems described approximately by local quantum field theory. The deviations from classical behavior of a macroscopic body, except for those that can be described as classical uncertainties in the initial values of macroscopic variables,are *exponentially* small as a function of the volume of the macrosystem in microscopic units. Conventional estimates are correct only when the internal degrees of freedom of the macrosystem are in their ground state, and the classical motion of collective coordinates is adiabatic. Otherwise, the system acts as its own environment and washes out quantum phase correlations between different classical states of its collective coordinates. I suggest that it is likely that we can only achieve meso-scopic superpositions, for systems which have topological variables, and for which we can couple to those variables without exciting phonons.

1 Classical behavior in the non-relativistic quantum mechanics of particles

In standard texts on non-relativistic quantum mechanics the classical limit is described via examples and via the WKB approximation. In particular, one often describes the spreading of the wave packet of a free particle, and estimates it as a function of time and the particle mass M. There is nothing wrong with the mathematics done in these texts, but the implication that these estimates provide the basis for an understanding of why classical mechanics is such a good approximation for macroscopic objects is not correct and therefore misleading. In particular it leads one to conclude that the corrections to decoherence for a wave function describing a superposition of two different macroscopic states is power law in the mass. I would aver that this mistake forms part of the psychological unease that many physicists feel about the resolution of Schrödinger's cat paradox in terms of the concept of decoherence.

These estimates have also led to recent experimental proposals to demonstrate quantum superposition of states of variables which are "almost macroscopic". I will argue that no such demonstration is possible, without extreme care taken to keep the constitutents of the macrosystem in their microscopic ground state. The essence of my argument is that the essential variable that controls the approach to the classical limit, is the number of localizable constituents of large quantum system. In a macroscopic material this would be something like the number, N, of correlation volumes contained in the sample. Away from critical points, the correlation volume is microscopically small, and we are roughly counting the number of atoms.

Indeed, all previous discussions also identify this number as the crucial parameter. These discussions identify a variety of collective classical variables, like the center of mass of the system, and note that the effective Lagrangian for these variables has a factor of N in it. For the center of mass, this is simply the statement that the mass is large. The traditional argument simply studies the quantum mechanics of these collective variables and estimates the corrections to classical predictions, which are typically power law in the large, extensive parameters. Estimates based on these ideas have led to the suggestion that plausible extensions of current experiments can reach the limit of quantum coherence for collective coordinates of systems with dimensions of millimeters. The failure to observe such correlations might be taken to mean that there is some fundamental error in applying quantum mechanics to macroscopic systems, as has been proposed by Penrose, Leggett and others.

The essential point of this paper is that "small" corrections to this collective coordinate

approximation completely invalidate this argument. It is not that the classical dynamics is not a good approximation to the quantum mechanics of the collective variables. What is not a good approximation is to neglect the back reaction of the collective variables on the huge set of other degrees of freedom in the macroscopic object. Locality ensures that external forces acting on the macroscopic body affect the collective coordinate through collective interactions with individual constituents, which then give rise to terms in the Hamiltonian coupling the collective coordinate to the constituents. In effect, different classical motions of the collective coordinate give rise to different, time dependent, Hamiltonians for the constituents. These extra terms are small, inversely proportional to powers of extensive parameters.

However, typical macro-systems have a finite microscopic correlation length. The wave function of the system is a sum of terms which are products of individual cluster wave functions for a localized microscopic subset of the constituents. This idea is the basis for approximate variational calculations like the Hartree-Fock or Jastrow approximations. As a consequence of the small corrections described in the previous paragraph, the individual cluster wave functions will be modified by a small amount and the overlap between wave functions for two different classical trajectories of the collective coordinate will be proportional to $1 - \epsilon$ where ϵ is a measure of the strength of the perturbation that leads to non-uniform motion of the center of mass. However, because the full many body wave function is a product of $o(N)$ cluster wave functions, the overlap is of order $(1 - \epsilon)^N$, *which is exponentially small in the volume of the system measured in microscopic units.* In other words, *for a macroscopic body, different classical trajectories of a collective coordinate divide the system into different approximate super-selection sectors.*

One can argue, using the methods of quantum field theory and statistical mechanics, that the time that it takes to observe phase correlations between different approximate superselection sectors is of order 10^{cN} where c is a constant of order one. This is true as long as one is in a regime where the density of states of the microscopic degrees of freedom is large, *i.e.* that the state of the system is a superposition of a densely spaced set of eigenstates, which behaves in a manner describable by statistical mechanics. Note that the ratio between the current age of the universe and the Planck time is a mere 10^{61}, so that even for a moderately large system containing $N \sim 10^3$ correlation volumes, this time is *so long that it is essentially the same number of Planck times as it is ages of the universe. No imaginable experiment can ever distinguish the quantum correlations between different states of the collective coordinates of a macro-system.* The extraordinary smallness of such double exponentials defeats all of our ordinary intuitions about ordinary physics. Over such long time scales, many counter-intuitive

things could happen. For example, in a hypothetical classical model of a living organism made of this many constituents, or in a correct quantum model, the phenomenon of Poincare recurrences assures that given (exponentially roughly) this much time,the organism could spontaneously self assemble,out of a generic initial state of its constituents. So much for Schrödinger's cat.

Another way of phrasing the same arguments comes from the vast literature on *decoherence*, which also introduces an important concept I have not yet emphasized. This is the fact that an approximate superselection sector is not a single state, but actually a vast ensemble of order 10^{cN} states, which share the same value of the collective coordinate. In the decoherence literature, it is argued that rapid changes in the micro-state of a macroscopic environment wipe out the quantum phase correlations between *e.g.* states with two different positions of a macroscopic pointer, which have been put into a *Schrödinger's cat* superposition via interaction with some micro-system. Another way to state the conclusions of the previous paragraph is simply to say that the constituents of a macroscopic body serve as an environment, which serves to decohere the quantum correlations between the macro-states of collective coordinates. Unlike typical environments, which one might hope to eliminate by enclosing the system in a sufficiently good vacuum, the inherent environment of a macro-system cannot be escaped. The collective variables exist and behave as they do, because of the properties of the environment in which they are embedded. It is only when the macroscopic system is held in its ground state, during experiments in which the dynamics of the collective variables is probed, that conventional estimates of quantum coherence for the collective coordinate wave function are valid.

In this introductory section, I will fill in the argument that conventional estimates of quantum corrections to classical behavior are wrong, using standard ideas of non-relativistic quantum mechanics. In the remainder of the paper I will discuss the basis for these calculations in quantum field theory. This will also remove the necessity to resort to Hartree-Fock like approximations to prove the point directly in the non-relativistic formalism. As noted, the essential point of the argument is that we must take into account the fact that a macroscopic object is made out of a huge number, which generally I will take to be $> 10^{20}$, of microscopic constituents, in order to truly understand its classical behavior. I will argue that, as a consequence, the overlaps between states where the object follows two macroscopically different trajectories, as well as the matrix elements of all local operators[1] between such states, are of

[1]In this context local means an operator which is a sum of terms, each of which operates only on a few of the constituent particles. A more precise, field theoretic, description will be given in the next section.

order

$$e^{-10^{20}}.$$

Consider then, the wave function of such a composite of $N \gg 1$ particles, assuming a Hamiltonian of the form

$$H = \sum \frac{\vec{p}_i^{\,2}}{2m_i} + \sum V_{ij}(x_i - x_j).$$

Apart from electromagnetic and gravitational forces, the two body potentials are assumed to be short ranged. We could also add multi-body potentials, as long as the number of particles that interact is $\ll N^2$.

The Hamiltonian is Galilean invariant and we can separate it into the kinetic energy of the center of mass, and the Hamiltonian for the body at rest. The wave function is of the form

$$\psi(X_{cm})\Psi(x_i - x_j).$$

Ψ is a general function of coordinate differences. In writing the Schrodinger equation we must choose $N-1$ of the coordinate differences as independent variables. If the particles are identical, this choice obscures the S_N permutation symmetry of the Hamiltonian. One must still impose Bose or Fermi statistics on the wave functions. This is a practical difficulty, but not one of principal. We now want to compare this wave function with the internal wave function of the system when the particle is not following a straight, constant velocity trajectory. In order to do this, we introduce an external potential $U(x_i)$. It is *extremely* important that U is not simply a function of the center of mass coordinate but a sum of terms denoting the interaction of the potential with each of the constituents. This very natural assumption is derivable from local field theory: the external potential must interact locally with "the field that creates a particle at a point". So we assume

$$U = \sum u_i(x_i),$$

where we have allowed for the possibility, *e.g.* that the external field is electrical and different constituents have different charge.

To solve the external potential problem, we write $x_i = X_{cm} + \Delta_i$ and expand the individual potentials around the center of mass, treating the remaining terms as a small perturbation. We then obtain a Hamiltonian for the center of mass, which has a mass of order N, as well as a potential of order N. The large N limit is then the WKB limit for the center of mass motion.

[2]Or that the strength of k body interactions fall off sufficiently rapidly with k for $k > N_0 \ll N$.

The residual Hamiltonian for the internal wave function has small external potential terms, whose coefficients depend on the center of mass coordinate.

The Schrodinger equation for the center of mass motion thus has solutions which are wave functions concentrated around a classical trajectory $X_{cm}(t)$ of the center of mass, moving in the potential $\sum u_i(X_{cm})$[3]. These wave functions will spread with time in a way that depends on this potential. For example, initial Gaussian wave packets for a free particle will have a width, which behaves like $\sqrt{t/Nm}$ for large t, where m is a microscopic mass scale. The fact that this is only significant when $t \sim N$ is the conventional explanation for the classical behavior of the center of mass variable.

In fact, this argument misses the crucial point, namely that the small perturbation, which gives the Hamiltonian of the internal structure a time dependence, through the appearance of $X_{cm}(t)$, is not at all negligible. To illustrate this let us imagine that the wave function at rest has the Hartree-Fock form, an anti-symmetrized product of one body wave functions $\psi_i(x_i)$, and let us characterize the external potential by a strength ϵ. In the presence of the perturbation, each one body wave function will be perturbed, and its overlap with the original one body wave function will be less than one. *As a consequence, the overlap between the perturbed and unperturbed multi-body wave functions will be of order* $(1-\epsilon)^N$. This has the exponential suppression we claimed, as long as $\epsilon \gg \frac{1}{N}$. It is easy to see that a similar suppression obtains for matrix elements of few body operators. One can argue that a similar suppression is obtained for generalized Jastrow wave functions, with only few body correlations, but a more general and convincing argument based on quantum field theory will be given in the next section. Here we will follow through the consequences of this exponential suppression.

The effect is to break up the full Hilbert space of the composite object in the external potential, into *approximate super-selection sectors* labeled by macroscopically different classical trajectories $X_{cm}(t)$ (microscopically different trajectories correspond to $\epsilon \sim \frac{1}{N}$). We will argue that local measurements cannot detect interference effects between states in different super-selection sectors on times scales shorter than $e^{10^{20}}$ (we leave off the obviously irrelevant unit of time). That is to say, for all *in principle purposes*, a superposition of states corresponding to different classical trajectories behaves like a classical probability distribution for classical trajectories. The difference of course is that in classical statistical physics one avers that *in principle* one could measure the initial conditions precisely, whereas in quantum mechanics the uncertainty is intrinsic to the formalism.

[3] See [2] for a construction of such wave functions for the Coulomb/Newton potential.

The argument for the exponentially large time scale has two parts, each of which will be given in more detail below. First we argue that it takes a time of order N, for a local Hamiltonian to generate an overlap of order 1 between two different superselection sectors. Then we argue that most macroscopic objects are not in their quantum ground state. The typical number of eigen-states present in the initial state of the object, or that can be excited by the coupling to the time dependent motion of the collective coordinate is of order e^{cN}. These states are highly degenerate. The time dependent Hamiltonian induced by the coupling to the collective coordinate will induce a time dependent unitary evolution on this large space of states, with a time scale of order 1 (in powers of N). Thus, there is a rapid loss of phase coherence between the two super-selection sectors, while the Hamiltonian is generating a non-trivial overlap between them. We would have to wait for the motion on the Hilbert space to have a recurrence before we could hope to see coherent quantum interference between two states with different macroscopic motions of the collective coordinate. The shortest recurrence time is of order e^{cN}.

Two paragraphs ago, I used the phrase *in principle* in two different ways. The first use was ironic; the natural phrase that comes to mind is *for all practical purposes*. I replace *in practice* by *in principle* in order to emphasize that any conceivable experiment that could distinguish between the classical probability distribution and the quantum predictions would have to keep the system isolated over times inconceivably longer than the age of the universe. In other words, it is meaningless for a *physicist* to consider the two calculations different from each other. In yet another set of words; the phrase "With enough effort, one can in principle measure the quantum correlations in a superposition of macroscopically different states", has the same status as the phrase "If wishes were horses then beggars would ride".

The second use of *in principle* was the conventional philosophical one: the mathematical formalism of classical statistical mechanics contemplates arbitrarily precise measurements, on which we superimpose a probability distribution which we interpret to be a measure of our ignorance. In fact, even in classical mechanics for a system whose entropy is order 10^{20}, this is arrant nonsense. The measurement of the precise state of such a system would again take inconceivably longer than the age of the universe.

This comparison is useful because it emphasizes the fact that the tiny matrix elements between super-selection sectors are due to an entropic effect. They are small because a change in the trajectory of the center of mass changes the state of a huge number of degrees of freedom. Indeed, in a very rough manner, one can say that the time necessary to see quantum interference effects between two macroscopically different states is of order the Heisenberg recurrence time

of the system. This is very rough, because there is no argument that the order 1 factors in the exponent are the same, so the actual numbers could be vastly different. The important point is that for truly macroscopic systems both times are super-exponentially longer than the age of the universe.

The center of mass is one of a large number of *collective* or *thermodynamic* observables of a typical macroscopic system found in the laboratory. The number of such variables is a measure of the number of macroscopic moving parts of the system. As we will see, a system with a goodly supply of such moving parts is a good measuring device. Indeed, the application of the foregoing remarks to the quantum measurement problem is immediate. As von Neumann first remarked, there is absolutely no problem in arranging a unitary transformation which maps the state

$$\alpha|\uparrow\rangle > +\beta|\downarrow\rangle > \otimes|Ready >,$$

of a microsystem uncorrelated with the $|Ready\rangle$ state of a measuring apparatus, into the correlated state

$$\alpha|\uparrow\rangle > \otimes|+\rangle + \beta|\downarrow\rangle > \otimes|- >,$$

where $|+/- >$ are *pointer states* of the measuring apparatus. If we simply assume, in accordance with experience, that the labels $+/-$ characterize the value of a macroscopic observable in the sense described above, then we can immediately come to the following conclusions

- 1.The quantum interference between the two pieces of the wave function cannot be measured on time scales shorter than the super-exponential times described above. The predictions of quantum mechanics for this state are identical *in principle* (first usage) to the predictions of a classical theory that tells us only the probabilities of the machine reading $+$ or $-$. *Like any such probabilistic theory* the algorithm for interpreting its predictions is to condition the future predictions on any actual measurements made at intermediate times. This is the famous "collapse of the wave function", on which so much fatuous prose has been expended. It no more violates conservation of probability than does throwing out those weather simulations, which predicted that Hurricane Katrina would hit Galveston.

- 2.One may worry that there is a violation of unitarity in this description, because if I apply the *same* unitary transformation to the states $|\uparrow\rangle \otimes |Ready\rangle$ and $|\downarrow\rangle \otimes |Ready\rangle$, individually, then I get a pair of states whose overlap is not small. This seems like a violation of the superposition principle, but this mathematical exercise has nothing to do with physics, for at least two reasons. First the macro-states labeled by $+/-$ are not

single states, but huge ensembles, with e^N members. The typical member of any of these ensembles is a time dependent state with the property that time averages of all reasonable observables over a short relaxation time are identical to those in another member of the ensemble. The chances of starting with the identical $|Ready\rangle$ state or ending with the same $|+/-\rangle$ states in two experiments with different initial micro-states, is e^{-N}. Furthermore, and perhaps more importantly, the experimenter who designs equipment to amplify microscopic signals into macroscopic pointer readings, *does not* control the microscopic interaction between the atoms in the measuring device and *e.g.* the electron whose spin is being measured. Thus, in effect, every time we do a new measurement, whether with the same input micro-state or a different one, it is virtually certain that the unitary transformation that is actually performed on the system is a different one.

For me, these considerations resolve all the *angst* associated with the Schrödinger's cat paradox. Figurative superpositions of live and dead cats occur every day, whenever a macroscopic event is triggered by a micro-event. We see nothing remarkable about them because quantum mechanics makes no remarkable predictions about them. It never says "the cat is both alive and dead", but rather, "I can't predict whether the cat is alive or dead, only the probability that you will find different cats alive or dead if you do the same experiment over and over". Wave function collapse and the associated claims of instantaneous action at a distance are really nothing but the the familiar classical procedure of discarding those parts of a probabilistic prediction, which are disproved by actual experiments. This is usually called the use of conditional probabilities, and no intellectual discomfort is attached to it.

It is important to point out here that I am not claiming that any classical probability theory could reproduce the results predicted by quantum mechanics. John Bell showed us long ago that this is impossible, as long as we insist that our classical theory obey the usual rules of locality. My claim instead is that the correct philosophical attitude toward collapse of the wave function is identical to that which we invoke for any theory of probability. In either case we have a theory that only predicts the chances for different events to happen, and we must continuously discard those parts of the probability distribution, which predicted things that did not occur. The considerations of this paper show that when we discard the dead cat part of the wave function after seeing that the cat is alive, we are making mistakes about future predictions of the theory that are in principle unmeasurable.

We are left with the discomfort Einstein expressed in his famous aphorism about mythical beings rolling dice. Those of us who routinely think about the application of quantum mechanics

to the entire universe, as in the apparently successful inflationary prediction of the nature of Cosmic Microwave Background temperature fluctuations, cannot even find comfort in the frequentist's fairy tale about defining probability "objectively" by doing an infinite number of experiments. Probability is a guess, a bet about the future. What is it doing in the most precisely defined of sciences? I will leave this question for each of my readers to ponder in solitude. I certainly don't know the answer.

Finally, I want to return to the spread of the wave packet for the center of mass, and what it means from the point of view presented here. It is clear that the uncertainties described by this wave function can all be attributed to the inevitable quantum uncertainties in the initial conditions for the position and velocity of this variable. Quantum mechanics prevents us from isolating the initial phase space point with absolute precision. These can simply be viewed as microscopic initial uncertainties in the classical trajectory $X_{cm}(t)$. In the WKB approximation, the marginal probability distributions for position and momentum are Gaussian, and there is a unique Gaussian phase space distribution that has the same marginal probabilities.

If we wait long enough these uncertainties would, from a purely classical point of view, lead to macroscopic deviations of the position from that predicted by the classical trajectory we have expanded around. The correct interpretation of this is that our approximation breaks down over such long time scales. A better approximation would be to decide that after a time long enough for an initial microscopic deviation to evolve into a macroscopic one, we must redefine our super-selection sectors. After this time, matrix elements between classical trajectories that were originally part of the same super-selection sector, become so small that we much declare that they are different sectors.

Thus instead of, in another famous Einsteinian phrase, complaining that the moon is predicted to disappear when we don't look at it (over a time scale power law in its mass), we say that quantum mechanics predicts that our best measurement of the initial position and velocity of the moon is imprecise. The initial uncertainties are small, but grow with time, to the extent that we cannot predict exactly where the moon is. Quantum mechanics *does* predict, that the moon has (to an exponentially good approximation) followed some classical trajectory, but does not allow us to say which one, a long time after an initial measurement of the position and velocity.

Of course, if the constituents of the macroscopic body are kept in their ground state during the motion, then we must treat the wave function of the center of mass with proper quantum mechanical respect, and the predictions of quantum interference between different classical

trajectories should be verifiable by experiment. This is clearly impossible for the moon. In a later section, I will discuss whether it is likely to be true for mesoscopic systems realizable in the laboratory.

1.1 Bullets over Broad-slit-way

To make these general arguments more concrete, let's consider Feynman's famous discussion of shooting bullets randomly through a pair of slits broad enough to let the bullets pass through. The bullet moves in the x direction, and we assume the initial wave function of the center of mass of the bullet is spread uniformly over the y coordinate distance between the slits. Then subsequent to the passage through the slits, the wave function of the center of mass is, to a good approximation, a superposition of two Gaussian wave functions, centered around the two slit positions. A conventional discussion of this situation would solve the free particle Schrodinger equation for this initial wave function and compare the quantum mechanical probability distribution a later times, with a classical distribution obtained by solving the Liouville equation for a free particle, with initial position and momentum uncertainties given by some positive phase space probability distribution whose marginal position and momentum distributions coincide with the squares of the position and momentum space wave functions.

In the latter calculation, the term in the initial probability distribution coming from the overlap of the Gaussians centered at the two different slits is of order $e^{-(\frac{L}{w})^2}$, where L is the distance between the slits and w their width. Liouville evolution can lead to uncertainty about which slit the particle went through in a time of order $\frac{2ML}{\hbar}$, just as in the quantum calculation. However, it gives rise to a different spatial distribution of probability density, with no interference peaks. Thus, for such times, the interference terms and exact Schrodinger evolution give a different result from classical expectations with uncertain initial conditions.

Now, let us take into account the fact that the two branches of the center of mass wave function must be multiplied by wave functions of the internal coordinates, which are in different super-selection sectors. As long as the micro-state is a superposition of internal eigenstates coming from a band with exponentially large density of states, it would be highly unnatural to assume that the micro-state in the top slit is simply the space translation of that in the bottom slit. The probability for this coincidence is e^{-cN}. It then follows from our previous discussion that we will have to wait of order a recurrence time in order to have a hope that the interference term in the square of the Schrodinger wave function is not exponentially small. The difference between quantum evolution of the center of mass wave function, and classical evolution with

uncertain initial conditions is completely unobservable, except perhaps at selected instants over super-exponentially long time scales.

2 Quantum field theory

I will describe the considerations of this section in the language of relativistic quantum field theory. *A fortiori* they apply to the non-relativistic limit, which we discussed in first quantization in the previous section. They also apply to cutoff field theories, with some kind of spatial cutoff, like a space lattice. The key property of all these systems is that the degrees of freedom are labeled by points in a fixed spatial geometry, with a finite number of canonical bosonic or fermionic variables per point. The Hamiltonian of these degrees of freedom is a sum of terms, each of which only couples together the points within a finite radius[4] In the relativistic case of course the Hamiltonian is an integral of a strictly local Hamiltonian density.

Let us first discuss the ground state of such a system. If the theory has a mass gap, then the ground state expectation values of products of local operators fall off exponentially beyond some correlation length L_c. If d is the spatial dimension of the system,and V is a volume $\gg L_c^d$, define the state

$$|\phi_c, V\rangle,$$

as the normalized state with minimum expectation value of the Hamiltonian, subject to the constraint that

$$\langle \phi_c, V | \int_V d^d x \ \phi(x)/V | \phi_c, V \rangle = \Phi_c.$$

Let $N = V/L_c^d$. One can show, using the assumption of a finite correlation length, that these states have the following properties

1.The quantum dynamics of the variable Φ_c is amenable to the semi-classical approximation, with expansion parameter $\propto 1/N$.

2.The matrix elements of local operators between states with different values of Φ_c satisfy

$$\langle \Phi_c, V | \phi_1(x_1) \dots \phi_n(x_n) | \Phi_c', V \rangle \sim e^{-cN},$$

[4]Various kinds of exponentially rapid falloff are allowed, and would not effect the qualitative nature of our results.

where n is kept finite as $N \to \infty$.

3.The interference terms in superpositions between states with different values of Φ_c remain small for times of order N. This follows from the previous remark and the fact that the Hamiltonian is an integral of local operators. This remark is proved by thinking about which term in the t-expansion of e^{-iHt} first links together the different superposition sectors with an amplitude of order 1. One needs terms of order N in order to flip a macroscopic number of local clusters from one macro-state to another. This term is negligible until $t \sim N$ in units of the correlation length. For many systems, there is a technical problem in this argument, because the Hamiltonian is unbounded, but it is intuitively clear that a cutoff at high energy should not affect the infrared considerations here.

4. However, there is another important phenomenon occurring, on a much shorter time scale. The microscopic degrees of freedom are evolving according to the microscopic Hamiltonian, perturbed by the time dependent term due to the motion of the collective coordinate. In a typical situation, the macroscopic object is not in its quantum ground state, but rather in some micro-state that is a superposition of many eigenstates from an energy band where the density of states is, according to quantum field theory, of order e^N[5]. The evolution in this subspace of states is qualitatively like that of a random Hamiltonian in a Hilbert space of this dimension. It leads to thermalization and loss of quantum coherence, through rapid changes of relative phase[3]. Initial quantum correlations will reassert themselves only once a recurrence time, and the shortest recurrence time is $o(e^N)$.

In the language of the previous section, *averages of local fields over distances large compared to the correlation length are good pointer observables, whenever the system is in a typical state chosen from an ensemble where the density of states is* $o(e^N)$. It is only when a macro-system is in its ground state, and the motion of the collective coordinates is adiabatic, relative to the gap between the ground state and the region of the spectrum with exponential density of states, that conventional estimates of power law (in N) time scales for seeing quantum coherence are valid.

Typical field theories describing systems in the real world contain hydrodynamic modes like phonons, with very low energies, and one would have to consider frequencies of collective coordinate motion lower than these hydrodynamic energy scales in order to observe quantum

[5]In reality, the system is unavoidably coupled to an environment, and is not in a pure state. If nothing else, soft photon emission will create such an environment. However, since our point is that the macroscopic system decoheres itself, we can neglect the (perhaps numerically more important) environmental decoherence.

coherence for the macroscopic observables over power law time scales. Certainly the motion of the moon is not in such an adiabatic regime.

To define an actual apparatus, we have to assume that the quantum field theory admits bound states of arbitrarily large size. Typically this might require us to add chemical potential terms to the Hamiltonian and insist on macroscopically large expectation values for some conserved charge. The canonical example would be a large, but finite, volume drop of nuclear matter in QCD. We can repeat the discussion above for averages over sub-volumes of the droplet.

Of course, in the real world, the assumption of a microscopically small correlation length is not valid, because of electromagnetic and gravitational forces. Indeed, most real measuring devices use these long range forces, both to stabilize the bound state and for the operation of the machine itself. I do not know how to provide a mathematical proof, but I am confident that the properties described above survive without qualitative modification[6]. This is probably because all the long range quantum correlations are summarized by the classical electromagnetic and gravitational interactions between parts of the system[7] . It would be desirable to have a better understanding of the modification of the arguments given here, that is necessary to incorporate the effects of electromagnetism and (perturbative) gravitation. One may also conclude from this discussion that a system at a quantum critical point, which has long range correlations not attributable to electromagnetism or gravitation, would make a poor measuring device, and might be the best candidate for seeing quantum interference between "macroscopic objects". Of course, such conformally invariant systems do not have large bound states which could serve as candidate "macroscopic objects".

Despite the mention of gravitation in the previous paragraph, the above remarks do not apply to regimes in which the correct theory of quantum gravity is necessary for a correct description of nature. We are far from a complete understanding of a quantum theory of gravity, but this author believes that it is definitely not a quantum field theory. In a previous version of this article [1] I gave a brief description of my ideas about the quantum theory of gravitation. I believe that it gets in the way of the rest of the discussion, and I will omit all but the conclusions.

In my opinion, the correct quantum theory of gravity has two sorts of excitations, something

[6]In the intuitive physics sense, not that of mathematical rigor.

[7]Recall that the Coulomb and Newtonian forces between localized sources are described in quantum field theory as quantum phase correlations in the wave function for the multi-source system.

resembling conventional particles, and black holes. A given region of space-time supports only a finite amount of information, and a generic state of that region is a black hole. Black holes have very few macroscopic moving parts, and do not make good measuring devices. Low entropy states in the region can be described in terms of particles with local interactions, and are approximable for many purposes by local field theory. I have explained how the general principles of field theory lead to an understanding of approximately classical measuring devices.

The exponential approach to the classical limit allows us to understand why these conclusions will not be changed in the quantum theory of gravity. Systems describable by local field theory over a mere $10^3 - 10^4$ correlation volumes already have collective variables so classical that their quantum correlations are unmeasurable. The fact that there exist energy scales orders of magnitude below the Planck scale, when combined with these observations, show us that practically classical systems can be constructed without the danger of forming black holes.

On the other hand, these same considerations show us that *exactly classical* observables in a quantum theory of gravity must be associated with infinite boundaries of space-time. This observation is confirmed by existing string theory models, and has profound implications for the construction of a quantum theory of gravity compatible with the world we find ourselves in.

3 Proposed experiments

3.1 SchrÖdinger's drum

I was motivated to rewrite this article for publication, by a number of papers, which propose experiments to observe quantum correlations for the observables of a mesoscopic system[4]. In its simplest form the system consists of two dielectric membranes, suspended in a laser cavity. By tuning the laser frequencies, it is claimed that one can cool the motion of the translational collective coordinates of the two membranes, which are coupled through the laser modes, down to their "steady state ground state". The ground state can be engineered to be a superposition of two different relative positions for the membranes. The system has been dubbed *Schrödinger's Drum*[4]. Although state of the art experiments can not yet reach the ground state splitting, it is plausible that it can be reached in the near future.

The membranes are about one millimeter square and 50 nano-meters thick. Typical phonon

energies are thus of order 10^{-4} eV[8]. In the analysis of the proposed experiments [5] it is argued that the collective coordinates of the two membranes in a laser cavity has a pair of classically degenerate ground states, which are split by 10^{-10} eV. It is then argued that by tuning the laser frequencies, one can cool the collective coordinate system down to temperatures below this splitting. The true ground state is a superposition of two classical values for the collective coordinates, and it is claimed that one can observe the entanglement of these two states.

At first glance, one might assume that the extremely low temperature of the collective coordinates means that the considerations of our analysis are irrelevant. However, on closer scrutiny it becomes apparent that the whole process of laser cooling, depends crucially on the coupling between the collective coordinates and a source of "mechanical noise". The latter is treated as a large thermal system with a temperature (in the theoretical analysis $10^{-7} - 10^{-6}$ eV)[9]. The analysis of [5] uses a quantum Langevin equation to describe the way in which energy is drained from the collective coordinates into the reservoir of mechanical noise. This might be[10] perfectly adequate for showing that the temperature of the collective coordinates can indeed be lowered, but it does not give an adequate account of quantum phase coherence.

The very fact that the source of mechanical noise can be modeled as a system obeying the laws of statistical mechanics, implies that the collective coordinate is coupled to a large number of other degrees of freedom, in a regime where the density of states of these degrees of freedom is exponentially large. Our analysis applies, and there should be no phase coherence between superselection sectors of the noise bath. If we take the correlation length in the membranes to be 100 nano-meters, then $e^N \sim e^{10^{12}}$, and there is no hope of seeing quantum coherence. *The failure to see quantum correlations in these experiments is not an indication that quantum mechanics breaks down for macro-systems, but simply a failure to understand that the approximate two state system of the membrane collective coordinates, suffers decoherence due to its coupling to the system which cools it down to the quantum energy regime.*

Indeed, the collective coordinates are coupled to the the system that provides the mechanical noise. The very fact that it is permissible to describe this system by statistical mechanics shows that it has a huge reservoir of states through which it is cycling on a microscopic time

[8]Here I refer to the phonons of internal sound waves on each membrane, rather than the phonons associated with the relative motion of the two membranes.

[9]I suspect that when applied to actual experiments with mm. size membranes, the source of this noise is excitation of sound waves on the membrane, and the temperature is even higher.

[10]The book [6] referred to in [5] suggests that the quantum Langevin treatment is adequate only when the noise reservoirs are collections of oscillators with linear couplings to the collective coordinates. It is not clear to me that this is the case. The collective coordinates are really zero wave number phonons of the individual membranes, and I would have guessed that they are coupled non-linearly to the shorter wavelength modes.

scale. The coupling of the collective coordinates to these states, which is necessary for the cooling process, also washes out phase correlations between different classical states of the collective modes.

3.2 Josephson's flux

By contrast, an older experiment[7] seems to illustrate the fact that when the microscopic degrees of a macro-system can be kept in their ground state, the standard analysis of the quantum mechanics of collective coordinates is correct. This experiment consists of a Josephson junction, with a flux condensate composed of $o(10^9)$ Cooper pairs. By appropriate tuning, one can bring the system to a state where there is resonant tunneling between a degenerate pair of quantum levels of the Landau-Ginzburg order parameter. The author's argue, correctly I believe, that because of the superconducting gap, and because their external magnetic probes couple directly to the order parameter, they can keep the system in its quantum ground state. They verify the level repulsion of a two state quantum system, when two classically degenerate states are connected by a tunneling transition. This experiment truly achieves a quantum superposition of macro-states. The recurrence time scale e^{10^9} of a typical state of this many microscopic constituents is irrelevant to the analysis of this experiment.

There is a hint here of what is necessary in order to approach macroscopic superpositions, and it echoes an insight that has already appeared in the quantum computing literature. Kitaev[8] has emphasized that *topological order parameters* may be essential to the construction of a practical quantum computer. Quantum Hall systems and superconductors have such order parameters. In the rather abstract language of quantum field theory, what would appear to be necessary is a system whose low energy dynamics is described by a *topologicalfieldtheory*. In plain terms, this is a system whose localized excitations are separated from excitations of a selected set of topological variables by an energy gap. From a practical point of view, what we need is a system in which this gap is large enough so that one can carry out experiments, which do not excite states above the gap.

Macroscopic systems will generically have phonon excitations, with energies that scale like the inverse of the largest length scale in the macroscopic body. However, as the case of the Josephson junction shows, it may be possible to devise probes of the system, which couple directly to the topological order parameters, without exciting mechanical oscillations. If that is the case, we are in a regime which is properly analyzed by the conventional collective coordinate quantum mechanics. For appropriately mesoscopic systems, and with sufficient

protection against decoherence by coupling to a larger environment, we can achieve quantum coherence for appropriate macroscopic order parameters.

This does not change the main burden of this article, which is that for typical macroscopic objects, quantum coherence is *superexponentially* unlikely, and cannot be observed over any experimentally realizable time scale. It does however, confirm the insight of Kitaev, that there may be a topological route to practical quantum computation.

4 Conclusions

I suspect the material in this paper is well understood by many other physicists, including most of those who have worked on the environmental decoherence approach to quantum measurement. If there is anything at all new in what I have written here about quantum measurement, it lies in the statement that a macroscopic apparatus of modest size serves as its own "environment" for the purpose of environmental decoherence. In normal laboratory circumstances, the apparatus interacts with a much larger environment and the huge recurrence and coherence times become even larger. Nonetheless, there is no reason to suppose that a modestly macroscopic apparatus, surrounded by a huge region of vacuum, with the latter protected from external penetrating radiation by thousands of meters of lead, would behave differently over actual experimental time scales, than an identical piece of machinery in the laboratory.

The exception to this kind of self-decoherence that we have identified, seems to involve topological variables of systems like superconductors and quantum Hall materials. These are systems with an interesting finite dimensional Hilbert space of quasi-degenerate ground states, separated from the rest of the spectrum by a substantial gap. In addition, one must have probes which can couple directly to the topological variables, without exciting low energy phonon degrees of freedom (which are present in any macroscopic object). For such systems, one might expect to be able to create robust superpositions of states of collective variables of macroscopic systems. Kitaev has argued that these may be the key to quantum computing.

The essential point in this paper is that the corrections to the classical behavior of macroscopic systems are exponential in the size of the system in microscopic units. This puts observable quantum behavior of these systems in the realm of recurrence phenomenon, essentially a realm of science fiction rather than of real experimental science. When a prediction of a scientific theory can only be verified by experiments done over times super-exponentially longer than the measured age of the universe, one should not be surprised if that prediction is counter-intuitive

or "defies ordinary logic".

Quantum mechanics does make predictions for macro-systems which are different than those of deterministic classical physics. Any time a macro-system is put into correlation with a microscopic variable - and this is the essence of the measurement process - its behavior becomes unpredictable. However, these predictions are indistinguishable from those of classical statistical mechanics, with a probability distribution for initial conditions derived from the quantum mechanics of the micro-system. It is only if we try to interpret this in terms of a classical model of the micro-system that we realize something truly strange is going on. The predictions of quantum mechanics *for micro-systems are* strange, and defy the ordinary rules of logic. But they do obey a perfectly consistent set of axioms of their own, and we have no real right to expect the world beyond the direct ken of our senses, which had no direct effect on the evolution of our brains, to "make sense" in terms of the rules which were evolved to help us survive in a world of macroscopic objects.

Many physicists, with full understanding of all these issues, will still share Einstein's unease with an intrinsically probabilistic theory of nature. Probability is, especially when applied to non-reproducible phenomena like the universe as a whole, a theory of guessing, and implicitly posits a mind, which is doing the guessing. Yet all of modern science seems to point in the direction of mind and consciousness being an emergent phenomenon; a property of large complex systems rather than of the fundamental microscopic laws. The frequentist approach to probability does not really solve this problem. Its precise predictions are only for fictional infinite ensembles of experiments. If, after the millionth toss of a supposedly fair coin has shown us a million heads, and we ask the frequentist if we're being cheated, all he can answer is "probably". Neither can he give us any better than even odds that the next coin will come up tails if the coin toss is truly unbiased.

I have no real answer to this unease, other than "That's life. Get over it." For me the beautiful way in which linear algebra generates a new kind of probability theory, even if we choose to ignore it and declare it illogical[11], is some solace for being faced with a question to which, perhaps, my intrinsic makeup prevents me from getting an intuitively satisfying answer. On the other hand, I believe that discomfort with an intrinsically probabilistic formulation of fundamental laws is the only "mystery" of quantum mechanics. If someone told me that the fundamental theory of the world was classical mechanics, with a fixed initial probability

[11] One can easily imagine an alternate universe, in which a gifted mathematician discovered the non-commutative probability theory of quantum mechanics, and speculated that it might have some application to real measurements, long before experimental science discovered quantum mechanics.

distribution, I would feel equally uncomfortable. The fact that the laws of probability for micro-systems don't obey our macroscopic "logic" points only to facts about the forces driving the evolution of our brains. If we had needed an intuitive understanding of quantum mechanics to obtain an adaptive advantage over frogs, we, or some other organism, would have developed it. Perhaps we can breed humans who have such an intuitive understanding by making the right to reproduce contingent upon obtaining tenure at a physics department. Verifying the truth of this conjecture would take a long time, but much less than time than it would take to observe quantum correlations in a superposition of macro-states.

5 Acknowledgments

I would like to thank Michael Nauenberg, Anthony Aguirre, Michael Dine, Jim Hartle, W. Zurek, Bruce Rosenblum, and especially Lenny Susskind, for important discussions about this thorny topic. This research was supported in part by DOE grant number DE-FG03-92ER40689.

References

[1] T. Banks, "Locality and the classical limit of quantum mechanics," arXiv:0809.3764 [quant-ph].

[2] M. Nauenberg, *Quantum wavepackets in Kepler elliptic orbits* Phys. Rev. A 140, 1133 (1989); *Autocorrelation function and quantum recurrences of wave packets*, J. Phys. B. At. Mol. Opt. Physics 23 1990; *Coherent wavepackets and semiclassical approximation in the Coulomb Potential*, in "Coherent States" edited by Feng,Klauder and Stayer, (World Scientific,1994): *The Classical Atom*,(with C. Stroud and J. Yeazell) Scientific American (June 1994); *Wave Packets: Past and Present*in " The Physics and Chemistry of Wave Packets" edited by J. Yeazell and T. Uzer (Wiley, 2000); *Negative Probability and the Correspondence between Quantum and Classical Physics*, M. Nauenberg and A. Keith, in Quantum Chaos-Quantum Measurement, ed. P. Cvitanovic et al, 265 (Kluwer 1992).

[3] M.V. Berry, *Regular and irregular semiclassical wavefunctions* , J. Phys A 10, 2083 (1977); J.M. Deutsch, *Quantum statistical mechanics in a closed system* Phys. Rev. A 43, 2046 (1991); M. Srednicki, *Chaos and quantum thermalization* Phys. Rev. E 50, 888 (1994) arXiv:cond-mat/9403051; M. Srednicki, *Thermal fluctuations in quantized chaotic systems* J. Phys. A 29, L75 (1996) arXiv:chao-dyn/9511001; M. Srednicki, *The approach to thermal equilibrium in quantized chaotic systems* J. Phys. A 32, 1163 (1999) arXiv:cond-mat/9809360. See the last of these for a more extensive list of references.

[4] J. Sankey, *Schrodinger's Drum*, Physical Review Focus, 11 Nov. 2008.

[5] M.J. Hartmann, M.B. Plenio, *Steady State Entanglement in the Mechanical Vibrations of Two Dielectric Membranes*, Phys. Rev. Lett. 101, 200503

[6] C.W. Gardiner and P. Zoller, *Quantum Noise: A Handbook of Markovian and Non-Markovian Quantum Stochastic Methods with Applications to Quantum Optics*, Springer Series in Synergetics, Oct 15, 2004.

[7] J.R. Friedman, V. Patel, W. Chen, S.K. Tolpygo, J.E. Lukens, *Quantum Superpositions of Macroscopic States*, Nature, 406, July 2000.

[8] A.Yu. Kitaev, *Fault tolerant quantum computation with anyons*, quant-ph/9707021.