Environmental decoherence versus intrinsic decoherence

By P. C. E. Stamp*

Pacific Institute of Theoretical Physics, Physics and Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, British Columbia, Canada V6T 1Z1

We review the difference between standard environmental decoherence and ‘intrinsic decoherence’, which is taken to be an ineluctable process of Nature. Environmental decoherence is typically modelled by spin bath or oscillator modes—we review some of the unanswered questions not captured by these models, and also the application of them to experiments. Finally, a sketch is given of a new theoretical approach to intrinsic decoherence, and this scheme is applied to the discussion of gravitational decoherence.

Keywords: decoherence; spin bath; oscillator bath; gravitation

1. Introduction

In this short paper, I was given the job of briefly reviewing current progress on the theory of environmental decoherence as applied to condensed matter systems. In order to make the topic a little more interesting, I have added some material on ‘intrinsic decoherence’. This latter term denotes a process that is considered to be generic to Nature, and that causes phase decoherence even without any ‘averaging over an environment’. As such, it is equivalent to a breakdown of quantum mechanics. Many arguments have been given to suggest that such a breakdown is something to be looked for—perhaps the most compelling are those beginning either from the clash between quantum mechanics and general relativity, or from the difficulties we have in thinking about macroscopic superpositions.

We will see that the topic of intrinsic decoherence is in its infancy, and that it is then interesting to compare it with environmental decoherence where things are a lot clearer. However, in what follows, I will also try to emphasize some of the remaining problems (as well as some of the successes) that we have with environmental decoherence.

2. Environmental decoherence: existing models

In order to say something useful about environmental decoherence, let us begin by reviewing the current understanding of how this works. This has come a very
long way since the schematic models that were used before the 1980s. One begins by writing the universe as a Hamiltonian of form

\[
H_{\text{osc}}^{\text{eff}} = H_0(P, Q) + H_{\text{int}}(P, Q; p_v, q_v) + H_{\text{env}}(p_v, q_v),
\]

(2.1)

where the \( Q, P \) are generalized coordinates and momenta for the central system of interest, and the \( \{ q_v, p_v \} \) are the same for the environmental degrees of freedom. There is then a large variety of models, for both the environment and the central system.

\( (a) \) Models of the environment

At first glance, it seems as though we ought to require a large variety of different model Hamiltonians for different environments. However, with few exceptions, it has been possible to reduce all of these to one of two models, as follows.

(i) Oscillator baths

Here, we model the environment by a bath of independent oscillators \([1,2]\), so that

\[
H_{\text{osc}}^{\text{env}} = \sum_{q=1}^{N_o} \left( \frac{p_q^2}{m_q} + m_q \omega_q^2 x_q^2 \right),
\]

(2.2)

where the \( \{ \omega_q \} \) are the frequencies of the \( N_o \) oscillators. Opinion on oscillator bath models was polarized in the early days—sometimes they were applied willy-nilly to almost any environment. And yet in their original discussion of such models, Feynman & Vernon [1] made it quite clear that the oscillator bath representation was generally valid only if the coupling between each environmental mode and the central system was very weak. The most obvious case is when the environmental modes are delocalized, extending over some large volume (so that \( N_o \) is an extensive quantity, depending in general on the UV cutoff \( \Omega_o \) we employ). Examples are phonons, photons, magnons, electron–hole pairs and other collective excitations in a fermionic system. The coupling is then taken to have the general form \([1,3]\)

\[
H_{\text{int}}^{\text{osc}} = \sum_{q=1}^{N_o} \left[ u_q(Q) x_q + v_q(P) p_q \right],
\]

(2.3)

which is restricted to being linear in the oscillator coordinates \( x_q \) and their conjugate momenta \( p_q \). The weakness of the interactions \( u_q(Q), v_q(P) \) is assured by the delocalization of the modes (so that each mode has a normalization factor \( \sim N_o^{-1/2} \), and thus \( |x_q|, |p_q| \sim N_o^{-1/2} \)); in the ‘thermodynamic limit’ (i.e. for large systems), when \( N_o \gg 1 \), this works.

Actually, this model is far more generally applicable than it seems. Often couplings to oscillator modes that are higher-order in the oscillator coordinates are important or even dominant in the real world—but these can be lumped into linear couplings to effective oscillators via temperature-dependent couplings. And while anharmonic couplings between oscillators will obviously affect the very long-time dynamics of the bath, they typically have little effect on the system dynamics over experimentally relevant time scales—which is what we are interested in.
The basic idea behind the oscillator bath model is one of linear response by the bath modes to the perturbation by the system—as such the idea goes back at least to Rayleigh’s model of friction.\(^1\)

(ii) Spin baths

Where oscillator bath models typically fail is when the environmental modes are not delocalized. This is quite common—examples include lattice defects, dislocation modes, spin impurities, dangling bonds, nuclear spins and localized phonons/vibrons; there are many others. Because these modes are localized in space, their Hilbert space is very small—most commonly they can be modelled as a ‘spin bath’ of two-level systems [6], described by a set of \(N_s\) ‘pseudospins’ \(\{\sigma_k\}\), with a Hamiltonian

\[
H_{\text{sp env}}^{\text{sp}} = \sum_k^{N_s} \mathbf{h}_k \cdot \mathbf{\sigma}_k + \sum_{k,k'} V_{kk'}^{\sigma \sigma} \sigma_k^\alpha \sigma_{k'}^\beta, \tag{2.4}
\]

where each pseudospin feels a set of external fields \(\{\mathbf{h}_k\}\), and inter-spin interactions \(V_{kk'}\). These are then coupled to the central system by an interaction:

\[
H_{\text{sp int}}^{\text{sp}} = \sum_k^{N_s} \mathbf{F}_k(P, Q) \cdot \mathbf{\sigma}_k. \tag{2.5}
\]

Unlike the case of the oscillator bath, the interactions \(\mathbf{F}_k(P, Q)\) are not necessarily weak; there is no reason to assume that \(|\mathbf{F}_k(P, Q)| \sim N_s^{-1/2}\) (indeed, in many cases, \(\mathbf{F}_k(P, Q)\) is independent of \(N_s\)!), and moreover, the localized modes may also be limited to some region in space, so that their number \(N_s\) is limited. We note also that one often must include the interactions \(V_{kk'}^{\sigma \sigma}\) between the bath spins; even though they are weak, they can have significant effects if \(N_s\) is not large.

A key dimensionless parameter for a spin bath is \(\lambda_k = |\mathbf{F}_k(P, Q)|/|\mathbf{h}_k|\), which tells us the relative strengths of the two fields acting on the bath spin \(\mathbf{\sigma}_k\). If \(\lambda_k \gg 1\), the bath spin dynamics is slaved to the central system; in many experimental systems, this is the case.

We see that the spin bath is inherently quantum-mechanical: discrete environmental states in a finite Hilbert space have no classical analogue. At low temperatures, such modes are everywhere in a solid—even very pure and well-prepared metallic or semiconducting crystal have a non-negligible concentration of defects and dislocations, as well as paramagnetic impurities, and typically their surfaces are very hard to control and will have surface impurities, dangling bonds and various kinds of lattice defect and ‘charge fluctuator’ modes, all of which can be described as spin bath modes. And this of course is quite apart from nuclear spins, which exist in profusion in almost all solids, and local vibrational modes, which in many non-crystalline and organic systems play a key role, coupling strongly to, for example, electronic modes, and behaving at low \(T\) very much like two-level systems.

Reviews of both oscillator bath and spin bath models can be found in the literature [6,7].

\(^1\)The use of a bath of oscillators to describe dissipation goes back to the nineteenth century. Rayleigh’s use of them (see Rayleigh [4]) is described concisely in Landau & Lifshitz [5].
(b) Models for central systems

These of course abound, and theory has examined a large number of them, both for their intrinsic interest, and in the context of experiments. The best-known class of models is one in which the central system is reduced to a ‘qubit’, i.e. a two-level system with Hamiltonian

$$H_o = B_o(t) \cdot \mathbf{r},$$

where $B_o(t)$ is some ‘external’ time-dependent field acting on the qubit. Most theoretical analyses of decoherence for this system have been performed assuming a static $B_o$. When coupled to an oscillator bath, via a coupling $H_{\text{int}}^z = \sum_q u_q x_q \hat{r}_z$, we have the diagonally coupled ‘spin-boson’ model [7,8]; it has been studied in the weak-coupling limit ever since the beginning of quantum mechanics (notably in atomic physics and quantum optics), and for more general coupling strengths starting in the late 1960s, in the context of RG theory and the Kondo problem. One can also add a non-diagonal coupling $H_{\text{int}}^\perp = \sum_q [u_q^\perp x_q + H.c.]$, which can be important.

When coupled to a spin bath, via a coupling $H_{\text{int}}^z = \sum_k F_k \cdot \mathbf{\sigma}_k \hat{r}_z$, the coupled system is known as the ‘central spin’ model [6]. Such models were first studied in nuclear magnetic resonance (NMR) and electron spin resonance, where again work was limited to weak coupling; more general studies began with work on tunnelling magnetic systems [9–11], and have accelerated with work on decoherence in solid-state qubit systems [6,12]. Again, one can generalize this to include non-diagonal terms, leading to a general tensor interaction.

The spin-boson and central spin models have very different dynamics [6,8]; this is inevitable, given the entirely different structure of the Hilbert spaces for the two baths. The differences are seen most easily in the lineshape (e.g. in $\chi''_{zz}(\omega)$, the Fourier transform to frequency space of the time correlation $\langle \hat{r}_z(t) \hat{r}_z(0) \rangle$). Suppose we hold $B_o(t)$ constant in time. Without the baths, $\chi''_{zz}(\omega) \propto \delta(\omega - \omega_o)$, where $\omega_o = 2B_o$. As one increases the coupling to an oscillator bath, the $\delta$-function for the spin-boson model first spreads into a narrow Lorentzian, and then with increased coupling, the peak broadens and moves to lower frequency; if the coupling becomes strong enough (as in an ohmic bath at strong coupling) the peak moves to zero energy (critical damping), and eventually collapses to a $\delta$-function at zero frequency (i.e. localization in time). On the other hand, for a spin bath, the central spin lineshape typically becomes very unconventional as one increases the coupling, developing multiple peaks and looking anything but Lorentzian. The corresponding behaviour in time also looks very different. Figure 1 illustrates this for the central spin model in a case where the dimensionless coupling to the spin bath is intermediate in strength—the contributions are divided between processes in which the net bath polarization does not change, and those for which it does. For a more detailed discussion of such results, see Prokof’ev & Stamp [6].

The reason for this difference is that (i) spins have a dynamics very different from that of oscillators (the spin dynamics is highly constrained compared with that of the oscillators) and (ii) in the spin-boson model, the bath is always in the linear response regime—each oscillator is only weakly affected by the qubit, whereas in the central spin model, the bath spin dynamics is strongly affected by the qubit, and we are typically far from a linear response regime. If $\lambda_k \gg 1$, the bath spins simply follow the central qubit (with deviations $\sim O(\lambda_k^{-1})$).

Phil. Trans. R. Soc. A (2012)
Decoherence: environmental and intrinsic

Another key difference between the spin-boson and the central spin models is related to this latter feature. In the spin-boson model (and indeed in all oscillator bath models), dissipation and decoherence are closely linked [2]; one cannot have one without the other, and both are connected to bath fluctuations via the fluctuation–dissipation theorem. However, in the central spin model, it is not only possible, but also quite typical for strong decoherence to occur with little or no dissipation. Indeed, the main decoherence mechanism in the central spin model is that of ‘precessional decoherence’, in which the bath spins precess in the time-varying field of the qubit, without any energy exchange between them [6,13]. Again, this is a fairly generic feature of spin bath decoherence.

For a more extended discussion of the differences between spin and oscillator baths, see Stamp [13].

Many other models for central systems have been discussed in the literature, including tunnelling systems, multiple qubit systems, quantum critical systems and quantum-information-processing systems. We will only look at one, i.e. a set of coupled qubits, with Hamiltonian

$$H_{0}^{\text{QIP}} = \sum_{j} B_{j}(t) \cdot \tau_{j} + \sum_{i<j} J_{ij}^{\alpha \beta}(t) \tau_{i}^{\alpha} \tau_{j}^{\beta},$$  \hspace{1cm} (2.7)$$

in which each local field $B_{j}(t)$ on the jth qubit is subject to independent control, as are the couplings $J_{ij}^{\alpha \beta}(t)$ between them. Then, to model decoherence, we couple
Figure 2. The dipolar interactions between qubits can cause resonant flip-flop transitions over long ranges—the figure depicts the qubits for which such a resonant interaction might exist in some qubit network. (Online version in colour.)

each individual qubit to the oscillator bath and the spin bath, using the same local couplings as given earlier for a single qubit (but with, in general, different couplings on each qubit). There are many possible variants here—different qubits may couple to different spin baths (being localized, these baths may be quasi-independent), and by changing the form of the coupling to oscillator baths (ohmic, super-ohmic, etc.) we can radically alter the dynamics.

However, a new kind of ‘internal decoherence’ enters with the introduction of the inter-qubit interactions. Much depends on the form of the $J_{ij}$ (in particular, whether they are long-range or short-range in space); we will be interested particularly in long-range dipolar interactions, which we can write in the form $J_{ij}^{a\beta} = J_{ij}^0 D_{ij}^{a\beta}$, where the interaction strength is $J_{ij}^0 = (\mu_0/\pi)\mu_B^2/r_{ij}^3$ for spins separated by a distance $r_{ij}$, and the spin tensor $D_{ij}^{a\beta}$ is

$$D_{ij}^{a\beta} = \frac{g_i^{\gamma\alpha} g_j^{\delta\beta}}{4} \left[ \delta_{\gamma\delta} - \frac{3}{r_{ij}^2} r_{ij}^{\gamma} r_{ij}^{\delta} \right],$$

where the $g_i^{\gamma\alpha}$ are the effective ‘g-factor’ tensors for the qubits [14]. Now these interactions are ‘frustrating’ [15], i.e. they make it hard for the system to order, and hard for simple pairwise interactions to operate locally; and they are long-range. The radical effect that these interactions can have is illustrated in figure 2, which shows a set of typical ‘resonating pairs’ of qubits, able to perform ‘flip-flop’ transitions via the interaction $J_{ij}$. The key point here is that in a three-dimensional system, such pairs are just as likely to be found far apart from each other as nearby (the decrease in interaction strength being compensated by
the larger number of available pairs at longer range). Thus, the quantum dynamics of the system is not driven by short-range ‘quantum diffusion’ of information—the information diffusion is spread all over the system.

Why do we care about this? One reason is that it is very hard to get rid of dipolar interactions such as this—they are not easily screened. So quantum computation is faced here with the problem that unless one can actually use these interactions as part of the computation protocol itself, then they cause ‘correlated errors’ over long ranges. This is, of course, just another form of decoherence. Thus, we see that Hamiltonians such as (2.7) containing dipolar interactions describe a key aspect of all qubit-based quantum-information-processing systems.

Studies of problems such as this on real systems have until recently been limited to looking at the effect of both the dipolar interactions and the coupling to the spin bath on, for example, the quantum phase transition in these systems [16–19] (note that related problems have been discussed in the NMR literature [20,21]). However, in §4, we will discuss how it has now become possible to quantify the various decoherence processes in such a system.

3. Environmental decoherence: some questions

A number of questions about the preceding framework jump out at us. Here are some.

(1) How good are these models? Thus, under what circumstances are they valid, and how does one derive them for given systems? And when can one use an oscillator bath representation even for localized modes? What about models like ‘noise’ models, where we model the environment by a noise source? And so on.

(2) Are there other kinds of decoherence (not captured by these models)? What about decoherence caused by time variation of parameters, or decoherence existing in the interaction channels—neither of these is captured in the above. What about ‘intrinsic decoherence’, i.e. decoherence quite independent of any environment, intrinsic to Nature?

(3) What about experiment? Can we have large-scale quantum computing, or large-scale quantum devices? What are the real limits on quantum behaviour imposed by decoherence? And—is decoherence inevitable at the macroscale?

(4) What about the ‘quantum measurement problem’? If there is such a problem, does decoherence ‘solve’ it, or any of the other foundational problems in quantum mechanics?

In the rest of this section, we take a whirlwind ride through questions (1) and (2). The next section will look at (3); and the rest of the paper will take a closer look at intrinsic decoherence. As for question (4), my general feeling is that until questions (1)–(3) have been resolved, there is really nothing new that can be said about (4).

(1) How good are these models? The derivation of the oscillator bath models was discussed by Leggett et al. [2,3]. The key here is the assumption of adiabaticity—if we define instantaneous eigenstates \( \tilde{\phi}_n(x, Q) \) and energies \( \tilde{\varepsilon}_n(Q) \) for the
environment when the central system has coordinate $Q$, then the oscillator energies $\omega_q \equiv \omega_{nm} = \tilde{\epsilon}_n(Q) - \tilde{\epsilon}_m(Q)$, and we can define the couplings $u_q, v_q$ in terms of the gauge potential

$$i A_{nm} = \prod_p \int dx_n \phi^*_n(x_p) \frac{\partial}{\partial Q} \phi_m(x_p)$$

(3.1)

provided that $Q(t)$ varies slowly enough in time that $|A_{nm}| \ll |(\tilde{\epsilon}_n - \tilde{\epsilon}_m)|$. If this condition is obeyed, then we can use the oscillator model even if the original couplings to the bath are strong—the $\{u_q, v_q\}$ will still be small.

However, in many cases, there will be low-energy environmental modes for which this condition is not obeyed (indeed, some of them will ‘resonate’ with the central system, i.e. their frequencies will match the characteristic frequencies of dynamics of the central system). In this case, one must fall back on the perturbative assumption discussed earlier, i.e. that the actual couplings to the bath modes are weak. It is precisely this assumption that fails in the case of many localized modes.

How good then are the spin bath models? In some cases, the couplings $F_k(P, Q)$ and fields $h_k$ are known accurately from microscopic arguments and/or experiment (e.g. with hyperfine couplings to nuclear spins, in spin qubit systems). In other cases, we can sometimes guess only at these quantities (as with the coupling to defects in magnets or superconductors). The key assumption, however, is that the couplings $F_k(P, Q)$, or their rates of change, are small enough so that the bath modes are not excited outside their restricted Hilbert spaces. Note, incidentally, that we also assume that $V_{ab}^{kk'}$ is small compared with either $|h_k|$ or $|F_k(P, Q)|$; otherwise, the bath spins couple together into delocalized modes, which behave more like oscillators.

It is quite common in some fields to approximate the effect of an environment by coupling the system to an external ‘noise’ source, i.e. to approximate the full Hamiltonian in (2.1) by a form $H(t) = H_0(P, Q) + V_N(Q; t)$, where the time-dependent potential $V_N$ contains a random force acting on $Q$, and the effects of this force are averaged over, using some technique (Langevin, Lindblad, Fokker–Planck, etc.). But it is well known that such models can never fully capture the effect of integrating out a dynamical environment [22]. They describe only the real part of the influence functional, which itself completely characterizes the effect of the dynamic environment on the system [1,22]; moreover, when approximated by, for example, Markovian processes, they also lose the long-time correlations that are characteristic of the influence functional of many environments. Physically, it is clear that a key part of the environmental effect on the central system dynamics is via a time-retarded ‘back-reaction’ on this system. We note that the results may sometimes be quite counterintuitive. The example of a particle moving on a lattice while coupled to a spin bath [23] provides an example: the particle density matrix shows simultaneous ballistic motion and anomalous diffusion, reminiscent of weak localization. Further discussion of all this is quite technical.

A related question is that of the crossover between spin bath and oscillator bath models, as one switches on the coupling between the bath spins—when the interspin interaction dominates, we expect the spin bath to behave like an
Decoherence: environmental and intrinsic

oscillator bath. This crossover (which for a macroscopic bath should, under some circumstances, become a phase transition) has been discussed in some detail for the central spin model [6], but there has been no general analysis as yet.

(2) What other kinds of decoherence exist? A more obvious question about failures of the oscillator and spin bath models is—what do they leave out? This turns out to be an interesting line of investigation. Here are a few suggestions (the list is not meant to be exhaustive).

— Other environmental modes: it is certainly not clear that the set of all environmental modes that might cause decoherence is exhausted by oscillators and localized ‘spin bath’ modes. One possible example, found in many systems at low energies, is that of delocalized quantum soliton modes—it is by no means clear (and has certainly never been shown) that these can be described in all cases by an oscillator bath model. There may also be other possibilities

— Third-party decoherence: because environmental decoherence arises from the entanglement of a system with its environment, we can certainly envisage situations where this happens without the direct interaction of the two. For example, the state of each may be ‘prepared’ or otherwise conditioned by a third party. A simple example is that of the coupling between the vibrational and rotational modes of a buckyball, on the one hand, and its centre of mass motion, on the other, via the slits it passes through in a two-slit experiment [13]. But one can easily imagine more subtle examples, and the real question of interest here seems to be—how can we know that a system has not previously entangled with some environment, perhaps long in the past, perhaps via a third party, in such a way as to cause decoherence now? This seems to be to be an open question, relevant in principle to many experiments.

— Switching decoherence: it is quite common, particularly in quantum information theory, to write time-dependent Hamiltonians—we have done so in (2.6) and (2.7). But this raises two points. First, where is the time-dependent field coming from that is causing this time variation, and should not this source term also be included in the Hamiltonian (to take account of possible back-reaction in the source)? And second, how do we know that the time variation cannot project the system dynamics out of the Hilbert space of the effective Hamiltonian we are using? This is of course the whole problem in ‘adiabatic quantum computation’, and in the similar problem of the dynamics of quantum phase transitions, where a large number of energy gaps can decrease to very small values during the time evolution. In this case, one can also include all these extra levels in the ‘environment’ (now with time-dependent energies). This topic is ripe for investigation, and certainly there are many papers on the dynamics of driven systems coupled to a bath, including some discussions of decoherence for simple models. However, to my knowledge, no systematic study is yet available.

Some approaches to the dynamics of driven quantum systems coupled to an oscillator bath are reviewed in Grifoni & Hanggi [24]. Note that analyses of such problems using a master equation approach (e.g. Na & Carmichael [25]) can miss important environmental feedback effects on the decoherence.
— Channel decoherence: in quantum information theory, it has been common to discuss the decoherence arising during the transmission of information down a channel. But such decoherence is ignored in most of the effective Hamiltonians we write down, where coupling constants are just treated as constants—this is equivalent to assuming instantaneous interactions. The value of these interaction constants depends on the UV cutoff $\hbar \Omega_o$ we assume in the theory—but this takes no account, for a finite quantum computer, of the time $T_o$ taken for signals to cross the computer. In reality, there is a set of time scales—the time $t_o$ taken for qubits to switch, the time $\tau_o = 1/\Omega_o$ (which is a UV cutoff time) and the time scales associated with the qubit operations themselves. Unless $T_o$ is very short compared with all of these, we should not be assuming an instantaneous interaction, but rather a retarded propagator. A proper treatment of this will be equivalent to a microscopic derivation of channel decoherence.

— Intrinsic decoherence: by intrinsic decoherence, I will refer to a decoherence in Nature that is ineluctable; it does not result from averaging over any kind of external environment, and thus can even exist for an isolated system. From this, we see that it must amount to a breakdown of quantum mechanics, because the linear equation of motion on quantum mechanics preserves superpositions for all time. The possible breakdown of quantum mechanics has been investigated in many different ways in the last 40 years; some of the best-known efforts include the postulation of additional nonlinear terms \[26–28\], or else stochastic terms \[29–31\], to the Schrödinger equation, and a breakdown of this equation caused by gravitational effects \[32,33\]; this list is by no means exhaustive. In most cases, the effect of these deviations from standard quantum mechanics is to mimic decoherence—in other words, in an experiment, they would be expected to add a further contribution to the existing environmental decoherence, which may be very hard to distinguish from it.

It is fair to say that our understanding of all five of the above-mentioned decoherence mechanisms is more or less open right now—and all of them are interesting and potentially important for experiments. In the second half of this paper, I will be looking at just one of them, i.e. intrinsic decoherence. However, a key challenge for any experiment looking for intrinsic decoherence will be to distinguish all sources of environmental decoherence from any intrinsic decoherence (and if possible eliminate these sources). Therefore, we first take a look at how it is possible to nail down environmental decoherence sources in experiments (cf. question (3) given at the beginning of this section).

4. Environmental decoherence: theory versus experiment

In the past decade, many experiments have been performed looking at decoherence in both quantum optics and condensed matter systems—in the last couple of years, these have been extended to hybrid systems such as optomechanical, optomagnetic or opto-superconducting systems, in which some combination of optical cavities, superconducting or magnetic qubits and mechanical or electromechanical oscillators are coupled and put collectively into
Decoherence: environmental and intrinsic

4439

a quantum state. In some of these experiments, the decoherence mechanisms were fairly simple, involving coupling of an oscillator degree of freedom to photons or phonons, and in this case, the experimental decoherence rates did not wildly disagree with simple theory. In more complicated cases, usually involving larger systems where the environment involved spin bath degrees of freedom, the experimental decoherence rates have often been much larger than naive theory would predict.

The temptation is to simply write these discrepancies off as ‘junk’ effects, coming from some combination of defects, impurities, etc. However, the problem here is that if the experiments are supposed to be testing, for example, the applicability of quantum mechanics to large and/or complex systems, such a temptation is no longer open to us—the theory should then also be predicting the environmental decoherence rates, and a genuine experimental test of quantum mechanics will then be testing such predictions.

It is of course always possible to argue that the theoretical models are perfectly alright, but that we are analysing their dynamics incorrectly. All theoretical predictions assume that we can average over all environmental modes to get a reduced density matrix for the central system alone; the dynamics of this reduced density matrix is then described by a propagator

\[ K(2,1) = K(Q_2, Q'_2; Q_1, Q'_1; t_2, t_1), \]

so that the reduced density matrix \( \tilde{\rho}(2) = \tilde{\rho}(Q_2, Q'_2; t_2) \) evolves according to

\[ \tilde{\rho}(2) = \int d1K(2,1)\tilde{\rho}(1), \quad (4.1) \]

where \( d1 \) means \( \int dQ_1 dQ'_1 \). The propagator can be written in the general path integral form [1]

\[ K(2,1) = \int_{Q_1}^{Q_2} DQ \int_{Q'_1}^{Q'_2} DQ' e^{-i/\hbar}(S_o(Q) - S_o(Q')) \mathcal{F}[Q, Q'], \quad (4.2) \]

where \( S_o[Q] \) is the free central system action, and \( \mathcal{F}[Q, Q'] \) is the ‘influence functional’, defined in general by

\[ \mathcal{F}[Q, Q'] = \prod_v \langle \hat{U}_v(Q, t) \hat{U}_v^+(Q', t) \rangle. \quad (4.3) \]

Here, the unitary operator \( \hat{U}_v(Q, t) \) describes the evolution of the \( v \)th environmental mode, given that the central system follows the path \( Q(t) \) on its ‘outward’ voyage, and \( Q'(t) \) on its ‘return’ voyage; and \( \mathcal{F}[Q, Q'] \) acts as a weighting function, over different possible paths \( (Q(t), Q'(t')) \).

The calculation of the influence functional is not necessarily easy. It is given in general form by a functional integral over the environmental modes as

\[ \mathcal{F}[Q, Q'] = \prod_v \int Dq_v \int Dq'_v \exp \frac{i}{\hbar} \Psi(q_v, q'_v; Q, Q'), \quad (4.4) \]

where the complex phase \( \Psi(q_v, q'_v; Q, Q') \) is given by the difference in actions over the two paths in the reduced density matrix, i.e.

\[ \Psi(q_v, q'_v; Q, Q') = S_B(q_v, Q) - S_B(q'_v, Q'). \quad (4.5) \]

Phil. Trans. R. Soc. A (2012)
The ‘bath action’ $S_B(q_r, Q)$ just sums over the environmental and interaction terms, i.e. we have $S_B = \int d\tau L_B$, where

$$
L_B(q_r, Q) = L_{\text{int}}(q_r, Q) + L_{\text{env}}(q_r).
$$

(4.6)

Everything then depends on the form taken by the influence functional. The oscillator bath influence functional was derived by Feynman and co-workers [1,7]; one lets $q_r \rightarrow x_q$, and the Lagrangian $L_B(x_q, Q)$ is then

$$
L_q(x_q, \dot{x}_q; t) = \frac{m_q}{2} [\dot{x}_q^2 - \omega_q^2 x_q^2] - \gamma_q(t)x_q,
$$

(4.7)

where $\gamma_q(t)$ is the driving force acting on the $q$th bath oscillator, coming from the central system. Thus, for example, if the coupling in (2.3) is $L_{\text{int}} = -\sum_q F_q(Q)x_q$, we have $\gamma_q(t) = F_q(Q(t))$. The bath action $S_B$ is then a quadratic form in the oscillator coordinates and one can find an exact expression for $F[Q, Q']$. However, depending on the form of $F_q(q)$ as a function of $Q$, the final integrals in (4.2) may still be highly non-trivial. This topic has been studied at great length—a good review is by Weiss [7].

The spin bath influence functional is very different, and has been studied much less. Now one lets $q_r \rightarrow \sigma_k$, and the Lagrangian $L_B(\sigma_k, Q)$ takes the form

$$
L(\sigma_k, \dot{\sigma}_k; t) = A_k \cdot d\sigma_k - Y_k(t) \cdot \sigma_k,
$$

(4.8)

where $A_k$ is formally the static gauge field that is produced by a monopole of unit strength at the centre of the Bloch sphere for $\sigma_k$ (and which leads to the Berry phase for $\sigma_k$ [34]). The time-dependent field $Y_k(t)$ is given by

$$
Y_k(t) = h_k + F_k(t) + \xi_k(t),
$$

(4.9)

in which $h_k$ is the static external field acting on the $k$th bath spin, and $F_k(t)$ is the dynamic field acting on this spin, coming from the central system (cf. equations (2.4) and (2.5)). The extra $t$-dependent fluctuating field $\xi_k(t)$ is defined by its components as

$$
\xi_k^\alpha(t) = \sum_{k'} V_{kk'}^{\alpha\beta} \sigma_k^\beta(t)
$$

(4.10)

and to first approximation this can be dropped, because $V_{kk'}^{\alpha\beta}$ is a small perturbation on the field $h_k + F_k(t)$.

This Lagrangian is much harder to deal with than the oscillator Lagrangian—indeed the solution to the problem of a spin-1/2 in almost any but the simplest time-dependent fields is practically impossible to determine (something that always surprises students). In spite of this, detailed expressions have been produced for the dynamics of systems coupled to spin baths, which are believed to be accurate in a fairly wide parameter range, at least up to a certain time [6,12]. Nevertheless, given the uncertainties here, one obvious way that one might try
to explain the discrepancies between theoretical and experimental environmental decoherence rates in so many experiments is to simply argue that the theoretical calculations of spin bath decoherence rates are just way off target.

What is obviously required here is a confrontation between theory and experiment where one has accurate knowledge of the physical sources of the decoherence (notably the spin bath modes). In what follows, I would like to briefly describe one specific experiment, which did explicitly attempt such a test. The experiment looked at a macroscopic crystal of tunnelling magnetic molecules, and specifically was looking for the macroscopic spin polarization wave that should propagate in a region of the sample if the decoherence rate is low. The experiment is interesting, because it checks out not only decoherence from local coupling to both oscillator and spin baths, but also a more insidious decoherence that comes from the long-range dipolar coupling between all of the tunnelling molecular spins.

The experiment was performed on a crystal of ‘Fe-8’ molecules; each molecule contains eight Fe ions in a small core, which are surrounded by a complicated sheath of ligand molecules, containing H, C, N, O and Br ions (figure 3). The spins in the central magnetic core are locked together by superexchange interactions into a spin-10 system, whose dynamics is governed by a biaxial crystal field. Depending on the different isotopic concentrations in the molecules, each one contains up to 216 nuclear spins. For the discussion here, we can just imagine that at sufficiently low temperature (below about 2.5 K), we have a set of spin qubits, each of which couples locally to both a phonon oscillator bath and a nuclear spin bath. These couplings, including all of the hyperfine couplings to the different nuclear spins, are very well understood quantitatively. The same is true...
of the dipolar coupling between the qubits, and so we can model the system very accurately by the effective Hamiltonian

$$H = H_o + H_{SB} + H_{osc},$$

where

$$H_o = \sum_j \Delta_o (\mathbf{H}_\perp) \hat{\tau}_x + \sum_{i<j} J_{ij}^{\alpha\beta} (\mathbf{H}_\perp) \hat{\tau}_i^\alpha \hat{\tau}_j^\beta$$

(4.11)

sums over all crystal lattice sites, and where $H_{SB}$ and $H_{osc}$ describe the nuclear spin and phonon terms, respectively. The temperature $T$ of the system is defined by the phonons; there is also a magnetic field $\mathbf{H}_\perp$ applied in such a direction that we can control the ‘tunnelling splitting’ $\Delta_o$ over a very wide range (some seven orders of magnitude), without adding any terms $\propto \hat{\tau}_j^z$. The effective interaction energy $J_{ij}^{\alpha\beta} (\mathbf{H}_\perp) \hat{\tau}_i^\alpha \hat{\tau}_j^\beta$ is actually dominated by longitudinal terms $J_{ij}^{zz}$; so $H_o$ then becomes the famous dipolar quantum Ising model. If the interaction were short-ranged, with strength $J_o$, we would get a standard zero-$T$ quantum phase transition when $|\Delta_o/J_o| \sim O(1)$. However, the interaction here is dipolar, and moreover we also have the couplings to the oscillator and spin baths—this changes things fundamentally.

This system is a key test case for quantum computing, for the following reason. We expect almost any solid-state gate-based quantum computer to possess local couplings between the qubits and spin and oscillator baths; and moreover, we expect long-range dipolar interactions to exist between the qubits, either electric or magnetic (or even strain-mediated). Such long-range interactions are very dangerous for quantum computation because conventional error correction routines do not work for them—indeed, it is not yet clear how one would correct errors arising from them, although there are arguments that concatenated error correction could do the job [35]. In any case, we see that this model contains all of the agents of decoherence that we may expect to arise in a multi-qubit quantum computer or quantum register (apart perhaps from the switching decoherence and channel decoherence noted earlier). In this sense, the experiment to be described is the first one to test all of these decoherence mechanisms together. A key feature of this experiment was that all the hyperfine interactions to the nuclear spin bath were known accurately, and so one can in principle directly test the theory of spin bath decoherence, and the calculational methods used to derive decoherence rates.

The theory for the decoherence rate as a function of temperature, applied field and for different isotopic concentrations of different nuclear spins was actually framed in 2006, and gave detailed predictions for the Fe-8 system [14], in a spin echo configuration suggested in that paper. The basic idea was that one could measure the ‘1-spin’ decoherence rate by doing a global spin echo experiment, in which the entire spin system was set into a spatially uniform precessional motion. One would, of course, like to measure the dynamics of multi-spin density matrices, and this will be a key goal for the future. In any case, one finds theoretically that the dimensionless decoherence rate (defined as $\gamma_{\phi} = 1/T_2 \Delta_o$, where $T_2$ is the 1-spin decoherence time extracted from spin echo measurements) should vary enormously depending on the field (which controls $\Delta_o$) and on the temperature (figure 4). The phonon decoherence rates increase quickly as the qubit tunnelling frequency $\Delta_o$ increases (because the phase space for phonon emission is increasing). On the other hand, the nuclear spin bath decoherence rates were predicted to fall rapidly at high $\Delta_o$ because the nuclear spins can no longer follow the qubit dynamics. Thus, for a single qubit, one arrives at an
optimal $\Delta_0$ (the ‘coherence window’ [36]), where the decoherence will be at a minimum. However, the dipolar decoherence, mediated by magnons, messes this up completely—and the only way to suppress it is by going to low temperatures (the qubit flip-flop processes that excite the magnons, and that are caused by dipolar interactions, are gapped in energy; they are then exponentially suppressed at low $T$).

A number of attempts were made by different experimental groups to test this theory, but they were rendered very difficult by the strong dipolar decoherence, which forced a very low-$T$ experiment when conventional fields were used. This problem was circumvented by Takahashi et al. [37], who went to fields in excess of 11 T, and also changed the field configuration. A very satisfactory agreement with theory was then found, with no adjustable parameters in either the theory or the experiment. This experiment actually accomplished some other notable goals—it was the first experiment to see macroscopic spin precession in an ordered array of qubits, and also achieved a lower decoherence rate than any previous molecular spin qubit system (the decoherence ‘Q-factor’ or inverse decoherence rate went to $1.5 \times 10^6$, and if these experiments had gone to lower $T$, it should have gone to $6 \times 10^7$). However, what interests us here is that it tested all three decoherence mechanisms simultaneously (including the first-ever measurement of dipolar decoherence). This is because by varying $T$ and $H_\perp$, one can independently vary the three mechanisms. So the obvious question is: what was being tested in this experiment?

*Phil. Trans. R. Soc. A* (2012)
The two most obvious conclusions from the results are (i) that because apparently all three decoherence mechanisms checked out, there could have been only negligible ‘extrinsic’ effects in the experiment—no hidden decoherence from dirt effects such as hidden defects, dangling bonds, etc., so that the only spin bath modes were the nuclear spins and (ii) that the calculations of both spin bath decoherence and dipolar decoherence are actually reliable. To this, we can add that the agreement with experiment for phonon decoherence was to be expected—this is a well-understood process. The success of the predictions for spin bath and dipolar decoherence is welcome—for a long time, it was not obvious if disagreements between theory and experiment came from experimental dirt or inadequate theory.

But let us now return to the question (3) posed in §3. What can we now say about decoherence for large-scale quantum processes, including the kind that would be involved in quantum computing with many qubits? Indeed, to what extent can one say that, in doing experiments such as this, we are testing fundamental theories of quantum mechanics on large systems, in the same way that experiments on, say, coherence in superconducting quantum interference devices (SQUIDs) [38] are supposed to be doing?

It seems to me that there are two basic approaches that one can take to this question. One is to probe the predictions of quantum mechanics in situations where its results are very counterintuitive, and apparently violate deep prejudices we have about Nature. Examples of this approach are the inequalities derived by Bell [39,40], which probe non-local entanglement, and have now been tested in many different ways on microscopic systems (and indeed constitute a ‘resource’ for quantum computation); and the generalization of these to time inequalities, to discuss possible experiments testing ‘macroscopic realism’ on SQUIDs [41]. It is clear that the experiments described earlier do not yet test quantum mechanics at the macroscopic level in this sort of way, but they certainly could be adapted to do so, and probably will be in the near future.

However, a second way of testing quantum mechanics, rather than just probing its consequences, is to develop alternative theories, and then do experimental tests that compare these directly with quantum mechanics, and indeed distinguish explicitly between them. This is precisely what theories of intrinsic decoherence aim to do; and so we now turn to these theories.

5. Intrinsic decoherence: theoretical framework

As noted earlier, by ‘intrinsic decoherence’ we mean a process that looks like decoherence (i.e. one in which phase coherence or phase interference is destroyed with time) but where this process is intrinsic to Nature, i.e. not obtained by simply averaging over ‘external’ degrees of freedom that happen to be entangled with the system of interest. As such, an investigation into intrinsic decoherence amounts to a search for a breakdown of quantum mechanics, of a kind that mimics (at least up to a point) conventional environmental decoherence. Thus, we are looking for a breakdown of quantum mechanics, of a specific kind.

For many physicists, the first reaction to this idea is probably to ask why anyone would bother looking at all! The two main reasons are (i) the well-known difficulties attending macroscopic superpositions of states in quantum mechanics
Decoherence: environmental and intrinsic

and (ii) the clash between quantum mechanics and general relativity, which has yet to be resolved—the success and generality of each of these theories makes a resolution of this clash perhaps the most important single problem in physics today.

As noted earlier, most previous discussions of the possible breakdown of quantum mechanics have started from the Schrödinger equation, and added extra nonlinear terms. In what follows, I will very briefly sketch a different framework. The purpose of this sketch is not to give anything like a full exposition of this theory, for which there is no space here, but simply to indicate how this theoretical framework can be set up. Full details will appear elsewhere.

Our choice of theoretical framework is motivated by (i) the observation that the quantum phase, via its connection to the action, plays a fundamental role in the time evolution of all known physical systems and (ii) that the transition amplitude (or just the ‘quantum amplitude’) plays a similarly fundamental role. It is hard to see how these would no longer be incorporated in any new theory intended to go beyond quantum mechanics. The emphasis on the quantum phase also focuses attention on phase coherence.

Our starting point will be the standard Feynman path integral formulation of quantum mechanics, and in this short paper, we will discuss only the non-relativistic version of this, for ordinary particles. The great advantage of the path integral formulation is that it makes explicit the idea that at any particular point on its worldline, a particle can choose between all possible paths. We will now generalize this idea as follows. We define a propagator $G(R, R')$ for a quantum system between two space–time points $R' = (r', t')$ and $R = (r, t)$, which we will assume to have the form

$$G(R, R') = G_o (R, R') + \Delta G (R, R'),$$

where $G_o (r, r'; t, t')$ is given by the usual path integral expression in non-relativistic quantum theory, i.e.

$$G_o (r, r'; t, t') = \int r' d x(\tau) \exp \frac{i}{\hbar} \int_t^{t'} d \tau L(x, \dot{x}; \tau),$$

whereas the new term $\Delta G (r, r'; t, t')$ will be taken, in the first instance, to have the form

$$\Delta G (r, r'; t, t') = \int r' d x_1 (\tau) \int_{r'} r'' d x_2 (\tau) \kappa [x_1, x_2] \exp \frac{i}{2\hbar}$$

$$\times \int\int_{t'} d \tau [L(x_1, \dot{x}_1; \tau) + L(x_2, \dot{x}_2; \tau)]$$

in which the interpath correlation functional $\kappa [x_1, x_2]$ (a functional of the two paths $x_1(t), x_2(t)$) remains to be determined. This definition is heuristic in the same way that the usual path integral is, in that the normalization has yet to be determined (i.e. the measure for the paths has yet to be defined): this point is

Discussion of some simple inconsistencies resulting from quantization of gravity is given in Unruh [43]; see also Penrose [33]. These inconsistencies are quite separate from the technical problems of non-renormalizability, etc., that arise when trying to quantize gravity. String theory and similar enterprises such as loop gravity attempt to find generalizations of quantum field theory which incorporate gravity, but are faced with many currently unresolved problems.

Phil. Trans. R. Soc. A (2012)
discussed later. The system is clearly non-relativistic—we have separated space and time, and the correlator $\kappa[1,2]$ depends only on integrals over the spatial variables $x_1, x_2$, so that the correlations it mediates are instantaneous.

The correction term $\Delta G(R, R')$ describes pairs of paths, and it allows communication or correlations between different paths. The physical interpretation of (5.3) is not that of a self-interaction—in conventional quantum mechanics, self-interactions for a particle are described by single particle paths, in which an interaction leaves and then rejoins the particle path, and then one sums over the different possible paths. Here, we are summing over all possible pairs of paths between the initial and final state, with the correlator $\kappa[1,2]$ weighting pairs of paths (figure 5); the total transition amplitude is now produced by summing not only over paths but over correlations between them. We can, if we wish, think of this as some kind of communication between different branches of the wave function. Clearly, there is no particular reason for stopping at pairs of paths—indeed, it would be more elegant to assume a sum over all possible $n$-tuples of paths, with $0 < n < N$ and $N \to \infty$. However, we will not discuss this more general formulation any further here, because we are simply going to assume for the moment that $K[2,1]$ is very small, so that we can, to first approximation, ignore all higher $n$-tuples. We then write

$$\kappa[1,2] = \gamma J[1,2] \quad (\gamma \ll 1),$$

(5.4)

where $\gamma$ is dimensionless and we note that both $J[1,2]$ and $\kappa[1,2]$ have dimensions of inverse length (or inverse interval, in a relativistic version). We will also typically assume that $J[1,2]$ takes the form

$$J[x_1, x_2] = \exp i\chi[x_1, x_2],$$

(5.5)

where the ‘phase’ $\chi[1,2]$ will in general be complex.

Figure 5. Contributions to $G(R, R')$ from different paths. In (a), we show the contributions in conventional quantum mechanics—paths sum independently. In (b), we show the intrinsic decoherence contribution, in which the correlator $\kappa[1,2]$ connects different branches of the propagator. (Online version in colour.)
In this theoretical framework, we will assume that both the amplitude \( \psi(r, t) \) and the density matrix still appear; however, the normalization of the path integral is not so obvious, and in fact has to be handled rather carefully. The amplitude obeys the equation of motion

\[
\psi(r, t) = \int \! dr' G(r, r'; t, t') \psi(r', t'). \tag{5.6}
\]

However, the density matrix is no longer given by \( \rho(r_1, r_2, t) = \psi(r_1, t) \psi^*(r_2, t) \). This is because we must allow correlations between different paths of the density matrix, in line with the assumptions made earlier for the one-particle propagator. This means that we can still write the equation of motion for the density matrix \( \rho(X, Y) = \rho(x, y, t) \) in the form \( \rho(2) = \int \! dK(2,1) \rho(1) \) as before, but now the propagator \( K(X, Y; X'Y') \) contains correlations between all possible paths connecting the end points \( X, Y, X', Y' \). If we then make the same expansion in powers of \( \lambda \) that is implicit in (5.3), we get an expression for the propagator of the form

\[
K(X, Y; X'Y') = \bar{K}(X, Y; X'Y') + \Delta K(X, Y; X'Y'), \tag{5.7}
\]

where \( \bar{K}(X, Y; X'Y') = G(X, X')G^*(Y, Y') \) is the product of the two independent propagators (each containing internal corrections from the correlator \( \kappa[1,2] \), as already mentioned), and the correction \( \Delta K(X, Y; X'Y') \) contains cross-correlations between the forward and reverse paths in the density matrix propagator, with the lowest-order term having the expression

\[
\Delta K(X, Y; X'Y') = \int_x^X \! D\bar{x}(\tau) \int_y^Y \! D\bar{y}(\tau) \kappa[x, y] \exp \left( \frac{i}{\hbar} \int_{\tau_1}^{\tau_2} \! d\tau' [L(x, \dot{x}; \tau) - L(y, \dot{y}; \tau)] \right) \tag{5.8}
\]

and so on. This last term, because it correlates the two parts of the density matrix propagator, can then have the effect of generating what looks like decoherence.

One can also define a reduced density matrix in the case where one deals with a system for which, for one reason or another, it is necessary to average over some of the degrees of freedom. Suppose the total density matrix (for the ‘universe’) is \( \rho(Q, Q', t) \), where the coordinates \( Q = (R, \{x_k\}) \), with \( k = 1, 2, \ldots, N \), and we are only interested in the ‘central coordinate’ \( R \). Then we simply write that the reduced density matrix is given as

\[
\bar{\rho}(R_1, R_2, t) = \text{Tr}_{\{\text{Env}\}} \rho(Q, Q', t) = \prod_k \int \! dx_k \rho(Q, Q', t) \tag{5.9}
\]

in the usual way.

We may then, by analysing the dynamics of the reduced density matrix, examine the corrections to its dynamics coming from the extra term \( \Delta K(2,1) \) in the propagator of the full density matrix. Suppose, for example, that we have a Lagrangian of form \( \mathcal{L}(R, \{x_k\}) = L_o(R) + L_f(R, \{x_k\}) \), where the ‘fast’ Lagrangian is written as

\[
L_f(R, \{x_k\}) = L_f(\{x_k\}) + L_{\text{int}}(R, \{x_k\}) \tag{5.10}
\]
with the fast variables now playing the role of the environment. Then, we can also make such a separation in the propagator $G(2, 1)$. The ‘bare’ term has the standard form

$$G_0^{f}(2, 1) = \int_{1}^{2} \mathcal{D}R \, e^{i\frac{\hbar}{\hbar} \int dt L_{0}(R, t)} \, G_{0}^{f}(2, 1), \quad (5.11)$$

where $G_{0}^{f}(2, 1) = G_{0}^{f}(\{x^{(2)}_{k}, x^{(1)}_{k}\}; t_{2}, t_{1}[R(t)])$ is the propagator for the fast variables, conditional on some path $R(t)$ for the slow ‘central’ variable (i.e. the propagator for the $\{x_{k}(t)\}$ when the slow variable path is ‘frozen’ to be some specific $R(t)$). Typically, one sandwiches the fast propagator between the instantaneous eigenstates $|n(R(t))\rangle$ of the entire collection of fast variables, and then we can write that

$$G_{0}^{f}(2, 1) = \langle 2|n(R)\rangle \, G_{0}^{f}(m, n; R(t)) \langle m(R)|1\rangle, \quad (5.12)$$

where $G_{0}^{f}(n, m; R(t)) = e^{(i/h) \int dt L_{nm}^{f}(R(t))}$, with an effective Lagrangian

$$L_{nm}^{f}(R(t')) = L_{0}(R) - \epsilon_{n}(R) - i\hbar \cdot \langle n|\nabla R m\rangle$$

containing the usual ‘Berry phase’ term derived by Born and Oppenheimer.

Now suppose we add the extra term $\Delta G(2, 1)$ to the propagator. If we again make the separation between slow and fast variables, we get an extra phase term in the dynamics of the Green function, and similar extra terms in the propagator for the density matrix, whose effect is to correlate the variables between two separate paths, for both the fast and slow variables. Substituting (5.10) into (5.3), we find that these phase factors have the form

$$\Phi_{nm}[R, R'] = \int dt \mathcal{R} \cdot \langle n(R)|\nabla R |\alpha(R)\rangle \chi_{\alpha\beta}[R, R'] \langle \alpha(R')|\nabla R'|m(R')\rangle \cdot \mathcal{R}', \quad (5.14)$$

where $\chi_{\alpha\beta}[R, R']$ is given by

$$\chi_{\alpha\beta}[R, R'] = |\alpha(R))\kappa[R, R'](\beta(R'))|. \quad (5.15)$$

In analysing a macroscopic system with many degrees of freedom, this kind of technique is essential.

One may also look at various limiting cases in this formalism. The semiclassical limit $\hbar \to 0$ can be studied directly in (5.3), and we see that because $\hbar$ appears in the denominator of phase factors, the same subtleties arise as for ordinary quantum mechanics. In the same way, one may study the adiabatic limit (both for a single particle and for an $N$-particle system), and the ‘macroscopic’ limit, where the number of particles in some object is taken to be very large.

Up until now, I have said nothing about how we are supposed to interpret the density matrix, and expressions such as $\langle \hat{A} \rangle = \text{Tr}\{\rho A\}$ (in standard quantum mechanics, of course, these refer to expectation values and lead, via the Born rule, to probabilities for these). Quite apart from these interpretational questions, we see that there must be a normalization requirement on the functional integrals—this is of course a problem even in the path integral formulation of ordinary quantum mechanics. In the present case, the normalization (and the
interpretation of the interpath correlation) depends on what specific form we employ for the kernel $k(2,1)$; this has then to be determined for each case. This should not be too surprising—indeed, the measure of paths in the path integral in ordinary quantum mechanics also depends on which specific Lagrangian we choose. To give a full discussion of this would take us into too long a technical diversion [44].

At first, the modifications of ordinary quantum mechanics seem rather small, if the kernel $k(1,2)$ is assumed to be small. However, the effect on the dynamics of any quantum system actually depends critically on the form of the kernel $k(1,2)$. It is then interesting to explore one example, which is of considerable interest in its own right. Thus, we now turn to the specific mechanism of gravitational decoherence, in which a correlation between paths mediated by gravitational interactions is assumed to arise.

6. An example: gravitational decoherence

Our example of an intrinsic decoherence mechanism is taken from studies in quantum gravity. Models such as the one to be discussed have to be taken seriously, given the apparent incompatibility between quantum mechanics and general relativity. However, I immediately emphasize that what follows is not intended to do more than indicate, by way of a toy example, the sort of thing that might be carried out—it is very far from being any kind of polished theory.

The basic physical idea to be discussed is that gravitational effects must lead inevitably to some sort of intrinsic decoherence in the time evolution of isolated systems. We see that any such arguments lead us inevitably to a non-unitary time evolution for state vectors—no averaging over an environment is involved, and yet phase correlations are lost. Such proposals have been quite common in the literature, ever since Hawking proposed that such non-unitary time evolution ought to be a part of any quantum theory of gravity [46]. Some of these proposals have incorporated quantum space–time fluctuations at the Planck energy as the source of this intrinsic decoherence [47,48], even to the extent of estimating how these might cause apparent decoherence in the dynamics of a SQUID. These theoretical proposals are controversial [49,50], and there has certainly been no clear experimental test of them at all (note that the estimates given in the theory for intrinsic decoherence rates contain a number of factors $\sim O(1)$, which then end up being exponentiated).

Somewhat distinct from these ‘Planck scale’ proposals are the ideas discussed by Penrose [33], following on from Diòsi [32]. These authors argue that in analysing quantum superpositions of different space–times, coming from situations where a mass is in two different positions, a kind of ‘gravitational time–energy uncertainty principle’ must operate. In these analyses, the time scale $\Delta t_G = \hbar/\Delta E_G$ plays the role of a decoherence time; here $\Delta E_G$ is the energy scale associated with the gravitational self-interaction of an object between its different quantum states. However, this decoherence is not environmental—it can be viewed as either coming from an uncertainty in the background time, or as arising from a mismatch of the space–times generated by the gravitational field of the system in question, in its different quantum states.
How can we incorporate these ideas into the earlier-mentioned formalism? Let us, for the sake of having a concrete example, consider a correlator $\kappa(1, 2)$ of form

$$
\kappa[r, r'] = \exp \int_0^t d\tau \frac{4\pi i Gm^2}{|r(\tau) - r'(\tau)|} - 1
$$

(6.1)

so that the strength of the phase communication between different paths is now proportional to the strength of the ‘gravitational self-interaction’ between them. I make no attempt here to justify this choice for $\kappa[r, r']$ except to note that it is consistent at least with the spirit of the Diòsi–Penrose uncertainty arguments mentioned earlier. We have assumed that the gravitational self-interaction can be handled in a Newtonian approximation.

Note, however, one key difference with the simple form introduced earlier—the correlator $\kappa[r, r']$ is no longer multiplied by a small parameter $\lambda$. Indeed, the phase factor actually blows up at short distances, even when the mass $m$ is small; and no power series expansion of the exponential is valid in the short-distance limit.

Consider now a really simple problem, in which we have a free particle of mass $m$ propagating from $X' = (R', t')$ to $X = (R, t)$. In the absence of the correction $\Delta G$, we just have the usual free propagator, i.e. $G(X, X') \to G_o(X, X')$. Adding the gravitational term then gives the correction

$$
\Delta G(X, X') \propto \int Dx_1(\tau) \int Dx_2(\tau) \kappa[x_1, x_2] \exp \frac{i}{2\hbar} \int d\tau \frac{m}{2} (\dot{x}_1^2 + \dot{x}_2^2).
$$

(6.2)

This correction affects only the relative coordinate $r = x_1 - x_2$, and thus we see that

$$
\Delta G(X, X') \propto \mathcal{A}(0, 0; t, t') G_o(X, X'),
$$

(6.3)

where $\mathcal{A}(r, t; r', t')$ is just the return amplitude (starting and finishing at the origin) for a particle of inertial mass $m$ in an attractive ‘Coulomb’ field, with a ‘charge’ of strength $8\pi Gm^2$. Thus, the main effect on the one-particle propagator is just a finite renormalization, whose effects are not obviously important, because we expect them to be absorbed in the normalization of the path integral. However, the effect on the density matrix propagator is more dramatic—we have a correction term that we write out in full, i.e.

$$
\Delta K(X, Y; X' Y') \sim \int_X Dx(\tau) \int_{X'} Dy(\tau) \left[ \exp \int d\tau \frac{8i \pi Gm^2}{|x(\tau) - y(\tau)|} - 1 \right] \times \exp \frac{i}{2\hbar} \int d\tau \frac{m}{2} (\dot{x}^2 - \dot{y}^2),
$$

(6.4)

in which decoherence is now caused by what looks like a gravitational interaction between forward and backward paths of the density matrix. Indeed, the kernel $\kappa[x, y]$ is now acting as an influence functional. Decoherence in this calculation will appear in the reduced density matrix in both the momentum and position representations.

An obvious question one can ask about this calculation is that, as noted already, the intrinsic decoherence correction to the dynamics here is not small in the limit $|x(\tau) - y(\tau)| \to 0$. Thus, we have no reason to believe that these calculations, limited as they are to the lowest order in the expansion in $\kappa[1, 2]$, give a proper picture of what is happening. This might be alright if the correlator as given...
earlier is assumed to be exact, but this is hardly likely. Another question that has to be dealt with if one wishes to discuss possible experiments is to generalize this calculation to deal with a system of many particles. Thus, one can imagine doing an experiment with a large mass made from particles distributed according to some density distribution. The way to deal with this is precisely using the separation of slow and fast variables discussed earlier—the calculations are quite lengthy [44].

This, of course, leads us to the key question: how easy will it be to see departures from quantum mechanics caused by intrinsic decoherence? A key problem with many of the theoretical attempts to predict a breakdown of quantum mechanics is that, because they are attempting to describe some kind of wave function collapse, they give predictions for dynamics that look very much like standard decoherence. Thus, one can ask: how can we distinguish between intrinsic and environmental decoherence in experiments on, for example, a macroscopic system such as a SQUID, or an array of magnetic molecules?

This is clearly a research problem of great interest. On the experimental side, some very ambitious and interesting efforts exist to look for gravitational decoherence [51–53], where the aim is to set up quantum states in which reasonably massive objects are in superpositions of spatially separated states. One very nice feature of these experiments is that the gravitational decoherence rate should indeed vary with external parameters in a way rather different from what one would reasonably expect from environmental decoherence rates (at least those coming from currently understood environmental decoherence mechanisms). However, the problem, at least for the tests of gravitational decoherence, is that there seems to be no unambiguous way of calculating the gravitational decoherence rate using the simple time–energy uncertainty arguments that the Diósi–Penrose arguments use. The answers depend in a quite arbitrary and very sensitive way on how one ‘coarse-grains’ the density distribution of the objects involved—this is simply because of the divergence of the gravitational ‘self-energy’ term as two mass distributions approach each other, so that the behaviour of the density distribution at very small distances determines the intrinsic decoherence rate. One interest of the earlier-mentioned approach is that it may help us to clarify this problem.

It is clearly going to be a very important task for theory to come up with a clear quantitative prediction for gravitational decoherence rates—and in fact, because experiments of this kind are perhaps not too far from completion, there may not be too much time left to do this!

This work was supported by NSERC, CIFAR and PITP. The author thanks Bill Unruh, Roger Penrose, Dirk Bouwmeester and Igor Tupitsyn for discussions on the material.

References


Decoherence: environmental and intrinsic


44 Stamp, P. C. E. In preparation.


