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:

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

and

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

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

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

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 $\sim$ 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, without 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, pp$\dot{9}$3–95) 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

$$|\alpha|^2 \left| q_1, p_1 \right\rangle \left\langle q_1, p_1 \right| + |\beta|^2 \left| q_2, p_2 \right\rangle \left\langle q_2, p_2 \right| \tag{3.42}$$

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 $|\theta\rangle$ 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

$$\begin{aligned}\mathcal{T}(j,t;j',t') &= \frac{\| \widehat{P_{j'}}\widehat{U}(t',t)\widehat{P_j}\widehat{U}(t,t_0)\, |\psi\rangle\, \|^2}{\| \widehat{P_j}\widehat{U}(t,t_0)\, |\psi\rangle\, \|^2} \\ &= \frac{\langle\psi|\, \widehat{U}^\dagger(t,t_0)\widehat{P_j}\widehat{U}^\dagger(t',t)\widehat{P_{j'}}\widehat{U}(t',t)\widehat{P_j}\widehat{U}(t,t_0)\, |\psi\rangle}{\langle\psi|\, \widehat{U}^\dagger(t,t_0)\widehat{P_j}\widehat{U}(t,t_0)\, |\psi\rangle}.\end{aligned} \tag{3.46}$$

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

$$|\Psi(\Delta t)\rangle \equiv \widehat{U}(\Delta t)\,|\Psi\rangle$$

$$= \int\int \mathrm{d}p_0\,\mathrm{d}q_0\,\mathrm{d}p_1\,\mathrm{d}q_1\,\beta_1(p_1,q_1;p_0,q_0)\alpha(p_0,q_0)\,|p_0,q_0\rangle\otimes|\phi(p_1,q_1,p_0,q_0)\rangle \tag{3.51}$$

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

$$\mathcal{D}(\alpha, \beta) = \langle\psi| \, \widehat{C}^\dagger_\alpha \widehat{C}_\beta \, |\psi\rangle \,. \tag{3.63}$$

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

$$\Pr(\alpha) = \langle\psi| \widehat{C}^\dagger_\alpha \widehat{C}_\alpha |\psi\rangle = \mathcal{D}(\alpha, \alpha). \tag{3.66}$$

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 $\widehat{P}^k_i \in \mathcal{S}^k$ 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

$$\langle\psi|\, \widehat{C}^\dagger_\alpha \widehat{C}_\alpha \,|\psi\rangle = \sum_{\alpha_i,\alpha_j \in \mathrm{Dec}(\alpha)} \langle\psi|\, \widehat{C}^\dagger_{\alpha_j} \widehat{C}_{\alpha_i} \,|\psi\rangle = \sum_{\alpha_i,\alpha_j \in \mathrm{Dec}(\alpha)} \mathcal{D}(\alpha_i, \alpha_j) \tag{3.71}$$

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?