

The Copenhagen Interpretation Born Again

Hollowood

November 7, 2014 arXiv

Abstract

An approach to quantum mechanics is developed which makes the Heisenberg cut between the deterministic microscopic quantum world and the partly deterministic, partly stochastic macroscopic world explicit. The microscopic system evolves according to the Schrödinger equation with stochastic behavior arising when the system is probed by a set of coarse grained macroscopic observables whose resolution scale defines the Heisenberg cut. The resulting stochastic process can account for the different facets of the classical limit: Newton's laws (ergodicity broken); statistical mechanics of thermal ensembles (ergodic); and solve the measurement problem (partial ergodicity breaking). In particular, the usual rules of the Copenhagen interpretation, like the Born rule, emerge, along with completely local descriptions of EPR type experiments. The formalism also re-introduces a dynamical picture of equilibration and thermalization in quantum statistical mechanics and provides insight into how classical statistical mechanics can arise in the classical limit and in a way that alleviates various conceptual problems.

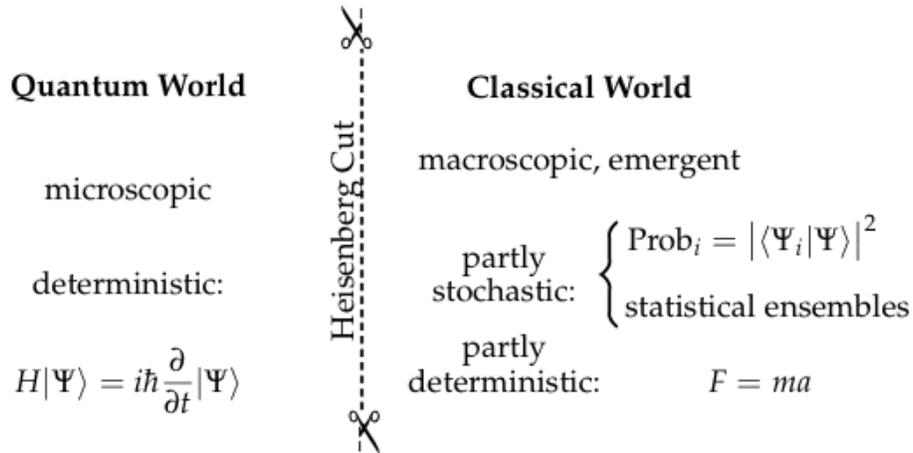
1. Introduction

Quantum mechanics is an amazingly successful and universal framework that describes physics over a huge range of scales, from the quantum theory of gravity provided by fundamental strings, through particle physics, nuclear physics, atomic physics to condensed matter physics. There can be little doubt that we know how to interpret quantum mechanics from a practical point of view by employing the rules of the Copenhagen interpretation.¹

The Copenhagen interpretation has all the tools that are needed to unlock the microscopic quantum world in the laboratory. It works by hypothesizing the existence of the Heisenberg cut, a division between the deterministic microscopic quantum world governed by the Schrödinger equation and the classical, partly deterministic, partly stochastic, macroscopic world governed by classical mechanics and the Born rule. It is worth amplifying this point: it is not the microscopic world that is probabilistic and the classical world

¹It can be argued that there is no single, historically coherent version of the Copenhagen interpretation. However, the rules are well known to, and used by, most physicists and also form the basis of most textbook treatments of quantum mechanics.

deterministic, the reality is the opposite as is apparent when one listens to a Geiger counter. Moreover, the classical world includes statistical mechanics and hence is stochastic even if the role of probability is not conceptually well understood in that context. It is important that the scale of the Heisenberg cut can be changed somewhat without altering the macroscopic predictions as long as it is kept below (as a scale in position space) macroscopic scales.



The Copenhagen interpretation posits a kind of stochastic dynamics that kicks in whenever a measurement is made on a microscopic quantum system. So during a measurement process where an initial state vector $|\Psi(0)\rangle$ evolves into the linear combination

$$|\Psi(T)\rangle = \sum_i c_i |\Psi_i\rangle \tag{1.1}$$

and where the component states $|\Psi_i\rangle$ are macroscopically distinct, the final state is then interpreted as describing an ensemble with probabilities are given by the Born rule

$$\text{prob}_i = |c_i|^2 \tag{1.2}$$

This begs some questions of which two key ones are:

1. When does the underlying pure state become interpreted as an ensemble, i.e., where is the Heisenberg cut to be placed?
2. Why are the macroscopically distinct states $|\Psi_i\rangle$ picked out as special?

The conventional answer to (1) is that it doesn't matter where we put the Heisenberg cut as long as it is somewhere between the microscopic and macroscopic while (2) is solved by fiat. The Copenhagen interpretation in its present form has shown itself sufficient to use quantum mechanics in practice. Is there, therefore, any merit in trying to go beyond the conventional rules in order to provide more convincing answers to the questions above? The danger is that this might add to the feeling that the Copenhagen interpretation is inadequate which is not our intention. On the other hand, as we will try to argue, there is the potential to unify both the deterministic and stochastic aspects of the classical world, the latter including both quantum measurement scenarios as well as statistical mechanics (SM).

The enhanced formalism we are after will determine what are the possible outcomes of a quantum measurement without having to resort to identifying them by hand. In addition, it will provide a better understanding of the Heisenberg cut and therefore how the classical world emerges from quantum mechanics. In this regard, the existence of a classical world is not guaranteed by the conventional classical limit and the correspondence principle alone. It will also yield a formalism where the reduction of the state vector is just a non-compulsory piece of book keeping.

In addition, an important element of the proposal will be to provide a new way to understand how classical SM can emerge from an underlying quantum system. It may come as a surprise to learn that classical SM still has many unresolved conceptual problems;² including:

1. The conceptual status of probability in a deterministic theory.
2. Definition of entropy and the second law.
3. Description of non-equilibrium and the approach to equilibrium.
4. The role of ergodicity and the relation between ensembles and long, or infinite, time averages.

The idea of building SM directly on quantum mechanics has its genesis in the work of Schrödinger [3] and von Neumann [4, 5] back in the 1920s. This

²Wallace [1] has a nice review of these conceptual difficulties and how they are related to those of quantum mechanics. Some of the outstanding conceptual issues of classical SM set in a historical context are discussed in the excellent review of Uffink [2].

kind of approach, taken up (or re-invented) in the modern era,³ misses out classical mechanics completely, and thereby avoids its conceptual problems. But we would surely be missing something by not understanding how classical SM can arise as a limit of quantum mechanics. In this regard, it would be wasteful not to put the probabilities of quantum mechanics to good use to provide ab initio probabilities for SM. Wallace [1] has pointed out that in this context that there is no need to introduce probabilities twice, firstly for quantum purposes and then for statistical mechanical purposes. We intend to elaborate on this point and then suggest that if we can understand the stochastic element of quantum mechanics more formally then we may at the same time provide a proper foundation for classical SM as the classical limit of quantum mechanics. There will be a considerable pay off because, as we will see, issues to do with probabilities, entropy, ensembles and ergodicity potentially become simpler in the quantum context.

With this motivation, we will construct a framework for quantum mechanics for which the following are true:

1. The underlying system is described by a state vector evolving according to the Schrödinger equation.
2. The role of observers is played by sets of observables that describe the macroscopic interactions of a macro-system.⁴
3. A classical regime and the usual Copenhagen rules of measurement emerge at macroscopic scales: definite macroscopically distinct outcomes are predicted with probabilities that are captured by the Born rule if the measuring device is accurate.
4. The classical limit for macroscopic systems displays both deterministic and stochastic features which can explain why macroscopic objects have definite positions and obey Newton's laws whilst, at the same time, their internal degrees of freedom are in thermal equilibrium. In the classical limit, a dynamical picture of classical SM emerges but with a stochastic dynamics on a discrete set of states for which issues concerning ergodicity are conceptually simpler to investigate.

³For a selection of papers see [6-18] and references therein.

⁴There is no need for Alices, Bobs, agents, brains, minds, consciousness or users — human, feline or otherwise.

The present work is a development of earlier work [19-21] which themselves are built on some aspects of the class of modal interpretations of quantum mechanics. The key idea is that Heisenberg cut is identified implicitly by the resolution scale of a set of relevant macroscopic degree of freedom of a system. In [19-21] this was done by assuming that the Hilbert space has an explicit factorization

$$\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_E \tag{1.3}$$

where \mathcal{H}_A describes the accessible, or macroscopically relevant, degrees of freedom and \mathcal{H}_E , the environment, encompasses all the remaining degrees of freedom. The tensor product factorization is therefore a manifestation of the Heisenberg cut. This kind of split is useful in toy models for describing the way the formalism works but in realistic situations we cannot expect such a clean split of the Hilbert space and, in addition, it is difficult to see how the Heisenberg cut could be varied in this context.

The fact that the Heisenberg cut can be varied as long as it is kept at sub-macroscopic scales is very reminiscent of the Wilsonian cut off in quantum field theory (QFT) [22]. In the modern era, QFTs are viewed as effective theories that have an explicit finite resolution, or coarse graining, scale known as the Wilsonian cut off.⁵ They are only valid for describing phenomena at larger distances than this scale. The important point is that phenomena on larger length scales are insensitive to changing the Wilsonian cut off as long as the latter is kept below the scale of the phenomena. The analogy with the Heisenberg cut is intriguing. One feature is worth emphasizing: there is a subjective element in choosing the Wilsonian cut off because it depends on the distance scale of the phenomena that you want to focus on. But this kind of subjectivity is completely natural: we are free to choose the length scale of the phenomena we want to investigate and each scale has its own effective theory.

The lesson for the Heisenberg cut is that it would be natural to associate it to the resolution scale of a set of observables that are needed to describe phenomena on, in this case, macroscopic scales: macroscopic phenomena re-

⁵Even our most fundamental QFT, the standard model, is only an effective theory which needs an explicit cut off to be well defined. It is much more stringent requirement to ask that a QFT has a continuum limit for which the cut off or resolution scale is taken to zero in position space. This kind of limit can only be taken if there is a UV fixed point of the renormalization group.

quire a sub-macroscopic Heisenberg cut. In the present approach, motivated for the need to describe more realistic systems, and with the notion of an effective QFT in mind, we propose to change the focus from tensor product factors and reduced density operators, to the set of relevant — or macroscopically accessible, or coarse grained — observables $\mathcal{A} = \{\mathcal{O}_n\}$. These are a set of observables that are coarse grained on sub-macroscopic scales that are needed to faithfully describe the macroscopic interactions of the system with other macroscopic systems. Implicit in the definition of the set \mathcal{A} , therefore, is the resolution or coarse graining scale which is nothing other than the Heisenberg cut. (See Figure 1) We will find the original tensor product and reduced density operator formalism will emerge as a special case of the more general formalism that we develop. In this sense, the present approach subsumes the earlier work [19-21].

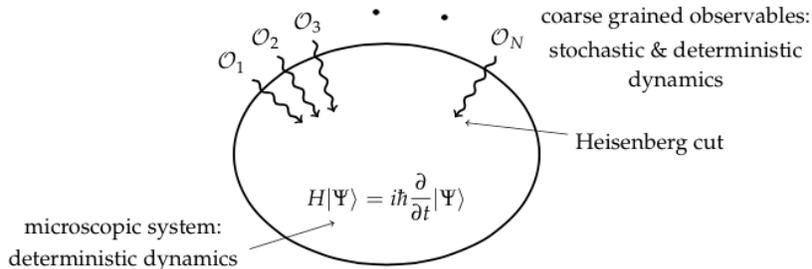


Figure 1: A quantum system evolves deterministically according to the Schrödinger equation. In our proposal, the stochastic element arises when the macroscopic interactions of the system are described by a set of probes in the form of a set of coarse grained observables $\mathcal{A} = \{\mathcal{O}_n\}$ whose resolution scale defines the Heisenberg cut.

In our approach to the quantum mechanics, the set of coarse grained observables \mathcal{A} has a perspective on the underlying quantum state with the Heisenberg cut identified with the resolution scale of the observables. Like the conventional Copenhagen interpretation, the perspective of \mathcal{A} involves a kind of stochastic dynamics, but unlike the conventional Copenhagen interpretation there is nothing ad hoc about its definition. In particular, the Born rule will emerge approximately in realistic situations with efficient measuring devices and state vector reduction will be seen to be a pragmatic — but ultimately unnecessary — piece of book keeping that is performed locally on the state vector.

It is a key idea of our approach that although the microscopic state of the system is the state vector $|\Psi\rangle$ the latter is not directly the macrostate of a system. This is very natural from the perspective of both QFT and quantum SM:

1. In QFT the macroscopic, or IR, state of a system is effectively determined by the correlation functions of the low momentum, or coarse grained, operators

$$\langle\Psi|\mathcal{O}_{n_1}\cdots\mathcal{O}_{n_p}|\Psi\rangle$$

2. In quantum statistical mechanics, the properties of equilibration and thermalization are are properties of the expectation values of coarse grained observables

$$\langle\Psi|\mathcal{O}_n|\Psi\rangle$$

In particular, they are *not* properties of the underlying state vector that does not equilibrate or thermalize itself.⁶

This paper is organized as follows. In section 2, we set out the main features of the *emergent Copenhagen interpretation* along with a limited commentary. Later sections expand upon this discussion beginning in section 3 which addresses the definition of the quantum microstates in detail. This is followed up in section 4 by a discussion of the stochastic process that controls the dynamics of the quantum microstates. It is important to show that the stochastic process is local in order to ensure that the formalism is causal. In section 5 we show how the formalism applies to a simple quantum system consisting of a particle moving in one dimension. Simple though it is, this example draws out some important details. A crude model of measurement is then described in section 6. This model is not supposed to be realistic but is simple enough to see how the proposal works in practice. In section 7, we show how the formalism produces a description of the EPR thought experiment using qubits that involves only local interactions between the qubits and the measuring devices: there is no quantum non-locality. Section 8 describes how the formalism has implications for the quantum mechanical underpinning of classical statistic mechanics. Finally, in section 9, we draw some conclusions.

⁶It is worth pointing out, and discussed at further length in section 8, that the idea of defining the effective state of a system through the expectation values of a set of coarse grained observables plays an important role in quantum approaches to SM going back to von Neumann [4].

2. The Emergent Copenhagen Interpretation

In this section, we summarize the three key elements of the emergent Copenhagen interpretation previously introduced in [19-21] but here generalized. More detail is given in the following sections, particularly 3 and 4.

1. *The underlying laws of quantum mechanics hold unchanged: in other words, the underlying state $|\Psi\rangle$ evolves according to the Schrödinger equation*

$$H|\Psi\rangle = i\hbar \frac{\partial}{\partial t} |\Psi\rangle \quad (2.1)$$

It is possible to relax the condition of the system having a pure state with the generalization to mixed states described in appendix A.

2. *The stochastic element that we observe of quantum system arises from the way the underlying quantum system interacts with other systems. This involves defining a set of probes in the form of coarse grained observables $\mathcal{A} = \{\mathcal{O}_n\}$ that describe the potential macroscopic interactions of the system with other systems. At time t , the state that is actually realized, from the point of view of the set \mathcal{A} , i.e., the macrostate, is encoded in the expectation values*

$$\langle \Psi_{i(t)} | \mathcal{O}_n | \Psi_{i(t)} \rangle \quad , \quad n = 1, 2, \dots, \dim \mathcal{A}$$

where $|\Psi_{i(t)}\rangle$ is one of the set of quantum microstates $\mathcal{M} = \{|\Psi_j\rangle\}$ that is realized at time t . These are a discrete set of orthonormal states (defined fully in section 3) that depend implicitly on both $|\Psi\rangle$ and \mathcal{A} with the property that the expectation values of operators in the set \mathcal{A} are effectively captured by an ensemble:

$$\langle \Psi | \mathcal{O}_N | \Psi \rangle = \sum_{i=1}^{\dim \mathcal{M}} |c_i|^2 \langle \Psi_i | \mathcal{O}_N | \Psi_i \rangle \quad , \quad \forall \mathcal{O}_n \in \mathcal{A} \quad (2.2)$$

that has maximal entropy and whose probabilities satisfy the “microscopic Born rule”

$$|c_i|^2 = |\langle \Psi_i | \Psi \rangle|^2 \quad (2.3)$$

Defining the set of observables \mathcal{A} and the set of quantum microstates \mathcal{M} will form the subject of section 3. In the simplest case, \mathcal{A} contains a single observable \mathcal{O} and then the set \mathcal{M} consists of the eigenstates of \mathcal{O} that appear in the expansion of $|\Psi\rangle$. However, in general realistic sets \mathcal{A} will consist of coarse grained observables in some local region of space that describe the possible macroscopic interactions of the system with other macroscopic systems.

Our approach incorporates the notion of *observer complementarity* implicit in the choice of the set of observables \mathcal{A} : different sets define different perspectives.⁷ The question that naturally arises is whether this introduces a kind of subjectivity into the approach: in other words, is the choice of the set \mathcal{A} determined or is it arbitrary? To add to the discussion in section 1, clearly there is some freedom to choose \mathcal{A} , however, this kind of non-uniqueness is familiar in physics. If we want to investigate phenomena of a system at a certain length scale the theory should be couched in terms of a set of relevant observables that are determined by the possible probes, or interactions with external systems, at that scale. The choice of scale is subjective but the details of the effective theory relevant at that scale is not. Since we want to investigate how a classical world can emerge from quantum mechanics the relevant scale is macroscopic and \mathcal{A} are a set of sub-macroscopically coarse grained observables that are relevant for describing macroscopic interactions. There will be some residual freedom in the choice of \mathcal{A} and the coarse graining scale—the Heisenberg cut—but we expect that the macroscopic predictions will be insensitive to changing this scale, in direct analogy to the insensitivity of the IR limit of effective QFT to changing the Wilsonian cut off or the way the observables are actually coarse grained.

The expression (2.2) for the expectation values of operators in the set \mathcal{A} in the state $|\Psi\rangle$ imply that they can be interpreted as a statistical ensemble of the quantum microstates with probabilities $|c_i|^2$. The fact that the probabilities equal $|\langle\Psi_i|\Psi\rangle|^2$ is a statement of what we call the *microscopic Born rule*.

⁷The fact that perspective plays an important role in quantum mechanics is implicit in the Copenhagen interpretation: local observers reduce the state vector according to the results of local measurements. It is also important for quantum approaches to statistical mechanics, where phenomena like equilibration and thermalization happen relative to a choice of macroscopic coarse grained observables.

It will be seen to be crucial to derive the Born rule itself in a measurement scenario.

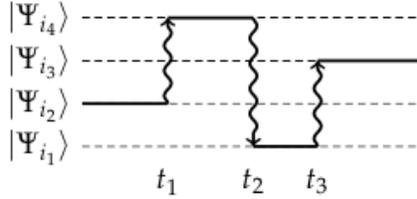
3. *The actual evolution of the system from the perspective of the set of observables A is encoded in the expectation values*

$$\langle \Psi_{i(t)} | \mathcal{O}_n | \Psi_{i(t)} \rangle, \quad \forall \mathcal{O}_n \in \mathcal{A}$$

for a sequence of quantum microstates $|\Psi_{i(t)}\rangle$ related by discontinuous transitions; e.g.,

$$i(t) = \begin{cases} i_2 & t < t_1 \\ i_4 & t_1 < t < t_2 \\ i_1 & t_2 < t < t_3 \\ \dots & \dots \end{cases}$$

as shown in the figure below:



described by a continuous time Markov chain with transition rates for $|\Psi_j\rangle \rightarrow |\Psi_i\rangle, i \neq j$

$$T_{ij} = \max\left(-\frac{2}{\hbar} \text{Im} \left[\frac{c_i}{c_j} \langle \psi_j | H | \Psi_i \rangle \right], 0\right) \quad (2.4)$$

The macrostate of the system is associated to the macroscopic time averages of $\langle \Psi_{i(t)} | \mathcal{O}_n | \Psi_{i(t)} \rangle$.

The stochastic process defined by the transition rates (2.4) allows only one way transitions between each pair of states, i.e., if $|\Psi_i\rangle \rightarrow |\Psi_j\rangle$ has a non-vanishing rate then the reverse $|\Psi_j\rangle \rightarrow |\Psi_i\rangle$ has a vanishing rate, and vice-versa.

The transition rates (2.4) ensure that the stochastic process is compatible with the Schrödinger equation and microscopic Born rule in the sense that if $p_{ij}(t, t')$ is the integrated transition probability that state $|\Psi_{j(t')}\rangle$ evolves via the process $|\Psi_{i(t)}\rangle$:

$$\begin{array}{ccc}
 |c_i(t)|^2 = \sum_j p_{ij}(t, t') |c_j(t')|^2 & & \\
 \uparrow & & \uparrow \\
 |\Psi(t)\rangle = \sum_i c_i(t) |\Psi_i\rangle & \xleftarrow{H|\Psi\rangle = i\hbar \frac{\partial}{\partial t} |\Psi\rangle} & |\Psi(t')\rangle = \sum_i c_i(t') |\Psi_i\rangle \\
 & & \downarrow
 \end{array} \tag{2.5}$$

where the time dependence of $c_i(t)$ follows implicitly from underlying Schrödinger equation (2.1). It is interesting that, in our context, it is the distribution $|c_i(t)|^2$ that is fundamental while the stochastic process is slave to it, rather than the more conventional situation where the transition rates would given and one would have to solve (2.5) for the $|c_i(t)|^2$. The result (2.5) means that, *if* the quantum microstates were distributed with probabilities $|c_i(t')|^2$ at time t' , then they will distributed with probabilities $|c_i(t)|^2$ at all later times $t > t'$. It might seem artificial to take this as an initial condition, after all the system really, from the perspective of \mathcal{A} , has expectation values

$$\langle \Psi_{i(t')} | \mathcal{O}_n | \Psi_{i(t')} \rangle$$

associated to a particular microstate at time t' . However, if the system is in equilibrium and the process is ergodic, then the $|c_i|^2$ are approximately time independent and after a suitable time, $p_{ij}(t, t')$ becomes effectively independent of j , i.e. the process forgets its starting state. In this situation, $|c_i(t)|^2$ *does* become interpreted as the probability that the quantum microstate is $|\Psi_{i(t)}\rangle$ at time t and the expectation values calculated from long time averages of the stochastic process are equal the underlying expectation value.

On the other hand, it is important also that in other situations ergodicity of the process can be broken, or partially broken, in order that Newton's laws can emerge and the measurement problem can be solved. There are three kinds of behavior that are needed to successfully account for the classical limit of macro-systems:

1. Deterministic: *this occurs when the matrix elements $\langle \Psi_i | H | \Psi_j \rangle$ are vanishingly small which can be expected when the quantum microstates are macroscopically distinct. In this case, over realistic macroscopic time scales the ergodicity of the stochastic process is broken. The possible stochastic trajectories then localize around the deterministic trajectories of classical mechanics.*

2. Stochastic: *this occurs when the matrix elements $\langle \Psi_i | H | \Psi_j \rangle$ are generic and the stochastic process is ergodic. This is exactly the situation needed to describe the thermal properties of macroscopic systems.*

3. Quantum measurement: *(a mixture of 1 and 2) during a measurement on a microscopic quantum system, the final state of the macroscopic measuring device splits up into subsets of quantum microstates $\mathcal{E}_a \subset \mathcal{M}$ for which the matrix elements $\langle \Psi_i | H | \Psi_j \rangle$ with $|\Psi_i\rangle \in \mathcal{E}_a$ and $|\Psi_j\rangle \in \mathcal{E}_b$, for $a \neq b$, are vanishingly small because the quantum microstates in different subsets are macroscopically distinct. The subsets \mathcal{E}_a become ergodically disjoint and therefore a definite measurement outcome occurs depending on which of the ergodic subsets the system lies in. In this context, the reduction of the state vector to the ergodic component corresponding to the measurement outcome that is actually realized,*

$$\sum_{|\Psi_i\rangle \in \mathcal{M}} c_i |\Psi_i\rangle \rightarrow \sum_{|\Psi_j\rangle \in \mathcal{E}_b} c_j |\Psi_j\rangle \quad (2.6)$$

is just an efficient piece of book keeping to reflect the fact that over macroscopic time scales T the integrated probabilities $p_{ij}(t+T, t)$, for $|\Psi_j\rangle \in \mathcal{E}_b$, $a \neq b$, are vanishingly small.

3. Quantum Microstates

In this section, we define the set of quantum microstates $\mathcal{M} = \{|\Psi_i\rangle\}$ that are key for our approach. It is important to recognize that these states depend implicitly on both the underlying quantum state $|\Psi\rangle$ and the set of observables \mathcal{A} .

Following the classic work of Jaynes [23], if we have incomplete knowledge of an underlying quantum state $|\Psi\rangle$ in the form of the expectation values of a set of observables

$$\langle \Psi | \mathcal{O}_n | \Psi \rangle \quad , \quad \mathcal{O}_n \in \mathcal{A}$$

then the most unbiased description of the state is the density operator ρ with

$$\langle \Psi | \mathcal{O}_n = \text{Tr}(\rho \mathcal{O}_n) \quad , \quad \forall \mathcal{O}_n \in \mathcal{A} \quad (3.1)$$

which maximizes the von Neumann entropy

$$S = -\text{Tr}(\rho \log \rho) \quad (3.2)$$

subject to the constraints (3.1). In a sense, we can view ρ as describing an ensemble of its eigenstates

$$\rho = \sum_i |c_i|^2 |\Psi_i\rangle \langle \Psi_i| \quad (3.3)$$

with the eigenvalues $|c_i|^2$ interpreted as probabilities. In the following, we will make this ensemble interpretation precise and will find that it emerges as the macrostate of a system in thermal equilibrium.

There are, of course, many density operators that can reproduce the expectation values of a set of observables \mathcal{A} , but the maximal entropy requirement picks out a unique one. We will make a refinement of this notion, by adding an additional constraint that we call the *microscopic Born rule*. This leads us to propose the definition of the set of quantum microstates $\mathcal{M} = \{|\Psi_i\rangle\}$:

1. Microscopic Born rule: *for each observable $\mathcal{O}_n \in \mathcal{A}$*

$$\langle \Psi | \mathcal{O}_n | \Psi \rangle = \sum_i |c_i|^2 \langle \Psi_i | \mathcal{O}_n | \Psi_i \rangle \quad ' \quad c_i = \langle \Psi_i | \Psi \rangle \quad (3.4)$$

2. Non-degeneracy: *for every pair of states $|\Psi_k\rangle$ and $|\Psi_l\rangle$, there exists at least one $\mathcal{O}_n \in \mathcal{A}$ for which*

$$\langle \Psi_i | \mathcal{O}_n | \Psi_j \rangle \neq \mu_n \delta_{ij} \quad , \quad i, j \in \{k, l\} \quad (3.5)$$

If the contrary were true then we simply reduce the dimension of \mathcal{M} by replacing the pair $|\Psi_k\rangle$ and $|\Psi_l\rangle$ with the, suitably normalized, linear combination

$$c_k |\Psi_k\rangle + c_l |\Psi_l\rangle \quad (3.6)$$

This condition, ensures that the quantum microstates are distinguishable with respect to the set of observables \mathcal{A} and it acts to limit the size of \mathcal{M} .

3. Maximal entropy: *subject to the conditions above, the coarse grained density operator has maximal entropy*

$$S = -\text{Tr}(*\rho \log \rho) = -\sum_i |c_i|^2 \log |c_i|^2 \quad (3.7)$$

This condition, roughly speaking, acts to make the set \mathcal{M} as big as possible subject to the other constraints.

As an example, suppose that \mathcal{A} consists of a complete set of observables acting on a tensor product factor \mathcal{H}_A of a Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_E$. In this case, discussed fully as Example 2 in appendix B, the quantum microstates $|\Psi_i\rangle$ are the states appearing in the Schmidt decomposition

$$|\Psi\rangle = \sum_i c_i |\Psi_i\rangle \quad , \quad |\Psi_i\rangle = |\psi_i\rangle \otimes |\tilde{\psi}_i\rangle \quad (3.8)$$

These states form the basis for the reduced density operator formalism developed in [19-21].⁸

3.1. Realistic sets of observables

To start with, let us suppose that \mathcal{A} consists of a set of commuting operators. In this case, the quantum microstates are simply the mutual eigenstates that appear in the decomposition of $|\Psi\rangle$. It is clear that the microscopic Born rule (3.4) is satisfied. More precisely, the non-degeneracy condition (3.5), ensures that there is only one quantum microstate for each degenerate eigenspace $\mathcal{H}_i \subset \mathcal{H}$. The maximal entropy condition requires that each degenerate eigenspace contributes separately.

In order to describe a macro-system, one expects that the set of observables \mathcal{A} contains coarse grained versions of the microscopic observables. Each such

⁸In those works, the states $|\psi_i\rangle$ of \mathcal{H}_A were called ontic states and $|\tilde{\psi}_i\rangle$ their mirrors. Here, given the more general context and the relation to statistical mechanics, we prefer the more general term *quantum microstate* for $|\Psi_i\rangle$ the states of $\mathcal{H}_A \otimes \mathcal{H}_E$.

observable \mathcal{O} has a finite working range $\Delta(\mathcal{O})$ and a finite resolution $\delta_{\mathcal{O}}$. We can define a set projection operators associated to the intervals $[\lambda_{i-1}, \lambda_i]$ which cover the working range of \mathcal{O} . with $\lambda_i = \lambda_0 + i\delta_{\mathcal{O}}$, $i = 1, 2, \dots, N$, where $N\delta_{\mathcal{O}} = \Delta(\mathcal{O})$, defined as

$$\Pi_i^{(\mathcal{O})} = \sum_{\lambda_a \in [\lambda_{i-1}, \lambda_i]} |\psi_a\rangle \langle \psi_a| \quad , \quad \Pi_i^{(\mathcal{O})} \Pi_j^{(\mathcal{O})} = \delta_{ij} \Pi_i^{(\mathcal{O})} \quad (3.9)$$

where $|\psi_a\rangle$ and λ_a are the eigenstates and eigenvalues of \mathcal{O} . For operators with continuous spectra, the sum is replaced by an integral. One can then define a class of coarse grained observables by taking linear combinations of these projection operators $\sum_i \alpha_i \Pi_i^{(\mathcal{O})}$, for example a coarse grained version of \mathcal{O} can be written

$$\mathcal{O}_{c.g} = \sum_{i=1}^N \left(\lambda_0 + \left(i - \frac{1}{2}\right) \delta_{\mathcal{O}} \right) \Pi_i^{(\mathcal{O})} \quad (3.10)$$

These operators have a discrete spectrum and are generally highly degenerate. A quantum microstate is associated to each degenerate subspace \mathcal{H}_i of the form

$$|\Psi_i\rangle = \sum_{\lambda_a \in [\lambda_{i-1}, \lambda_i]} f_{ia} |\psi_a\rangle \quad (3.11)$$

for coefficients f_{ia} determined by the state $|\Psi_i\rangle$.

Of course, we expect that the set \mathcal{A} will contain many observables \mathcal{O}_n and then the issue becomes how to treat the situation where the microscopic observables do not commute. If one builds the coarse grained observables out of the projection operators defined above then these projection operators will not commute. As an example, consider the position and momentum of a particle with associated projection operators

$$\Pi_i^{(x)} = \int_{x_{i-1}}^{x_i} dx |x\rangle \langle x| \quad , \quad \Pi_i^{(x)} \Pi_j^{(x)} = \delta_{ij} \Pi_i^{(x)} \quad (3.12)$$

for $x_i = x_0 + i\delta_x$, and

$$\Pi_i^{(p)} = \int_{p_{i-1}}^{p_i} dp |p\rangle \langle p| \quad , \quad \Pi_i^{(p)} \Pi_j^{(p)} = \delta_{ij} \Pi_i^{(p)} \quad (3.13)$$

with $p_i = p_0 + i\delta_p$. These projection operators do not commute even though in the realistic limit $\delta_x \delta_p \gg \hbar$, the lack of commutation is small.

In our approach there is nothing to prevent us building coarse grained observables on these projection operators leading to a non-abelian set \mathcal{A} , as we describe in appendix B. Although this creates no conceptual difficulties, it does create technical difficulties in finding explicit expression for the quantum microstates. However, we can use the freedom that we have in defining the observables to slightly deform the operators $\Pi_i^{(x)}$ and $\Pi_j^{(p)}$ to get new coarse grained projection operators which *do* commute. Intuitively one expects that this should be possible when the volume of the cell in phase space $\delta_x \delta_p \gg \hbar$. The idea of deforming the position and momentum operators so that they commute, but which are suitably close to the original operators, goes back to von Neumann [4, 5]. For example, Halliwell [26] has made an explicit construction of operators which are effectively deformations of our phase space projectors which do commute and are still projectors and so would be suitable for our purposes.⁹ If we now take \mathcal{A} to consist of the deformed operators the quantum microstates are just the simultaneous eigenstates. These eigenstates are very accurately—but not exactly—localized in the cell

$$[x_{i-1}, x_i] \cap [p_{j-1}, p_j]$$

in phase space when $\delta_x \delta_p \gg \hbar$. It is reasonable to expect that working with a slightly different set of observables \mathcal{A} will not lead to any significant differences in the resulting analysis.¹⁰

The same kind of deformation can be generalized to any non-abelian set \mathcal{A} where the non-commutativity is small and so it is sufficient to assume that the set of observables \mathcal{A} needed to describe the classical limit of a quantum system are abelian.

⁹It is important that the projectors of Halliwell are not exhaustive on phase space to avoid the consequences of the Balian-Low theorem. However, this is not a restriction for us because there is no requirement that \mathcal{A} is generated by an exhaustive set of projection operators. It is more natural for us that \mathcal{A} consists of operators that only cover a local region of phase space.

¹⁰This kind of freedom is familiar in effective QFT. For example, one can define coarse grained observables by defining the theory on a lattice or one can work in the continuum and cut off Fourier modes at high frequency. The IR behavior of the QFT is insensitive to these differences in regularization.

3.2. Local observables

In general, we expect that \mathcal{A} contains coarse-grained observables that are localized in some compact region of space A . To reflect this, the Hilbert space has a tensor product decomposition $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_E$. The set of accessible operators \mathcal{A} act locally on the \mathcal{H}_A factor, i.e., are of the form $\mathcal{O}_n \otimes 1_E$. If there were no coarse graining, so that \mathcal{A} were a complete set of observables on \mathcal{H}_A , then this would correspond to Example 2 in appendix B. More realistically we expect some amount of coarse graining so that the set \mathcal{A} is not complete. In this more realistic situation, therefore, the set of observables \mathcal{A} has both an ultra-violet cut off, the scale of the coarse graining, and an infra red cut off, the scale of the compact region in space.

Taking \mathcal{A} to consist of a set of commuting observables, the quantum microstates will generally be entangled with states of the environment, i.e., of the form

$$|\Psi\rangle = \sum_{i=1}^N c_i |\Psi_i\rangle \quad , \quad |\Psi_i\rangle = \sum_a |\psi_{ia}\rangle \otimes |\tilde{\psi}_{ia}\rangle \quad (3.14)$$

where $|\psi_{ia}\rangle$ are the orthonormal eigenstates of the abelian set \mathcal{A} with $a = 1, 2, \dots, \dim \mathcal{H}_i$.¹¹ These states look superficially like the Schmidt states (3.8), however, there is no requirement here that the states $|\tilde{\psi}_{ia}\rangle$ are orthogonal. The only constraint on them is the normalization condition

$$\sum_a \langle \tilde{\psi}_{ia} | \tilde{\psi}_{ia} \rangle = 1 \quad (3.15)$$

4. The Stochastic Process and Ergodicity

In this section we discuss in more detail the stochastic process defined by the transition rates (2.4). The first point, amplifying what we said earlier in section 2, the process is consistent with the Schrödinger equation in the sense that if we temporarily were to interpret $|c_i|^2$ as the probability that the system from the point of view of \mathcal{A} is in the quantum microstate $|\Psi_i\rangle$ at time t , then the stochastic process preserves this interpretation at later time, as in (2.5). So viewing the quantum microstates as an ensemble with probabilities $|c_i|^2$ then the ensemble average over the quantum microstates evolving according to the stochastic process equals the ordinary expectation

¹¹For the continuous spectrum case the sum is replaced by an integral.

value for any observable in \mathcal{A} .

From the Schrödinger equation, we find

$$\frac{d|c_i|^2}{dt} = -\frac{2}{\hbar} \sum_{j \neq i} \text{Im} \left[c_i c_j^* \langle \Psi_j | \left(H - i\hbar \frac{\partial}{\partial t} \right) | \Psi_i \rangle \right] \quad (4.1)$$

In realistic cases, the observables \mathcal{A} are time independent and it follows that

$$\langle \Psi_j | \frac{\partial}{\partial t} | \Psi_i \rangle = 0 \quad , \quad i \neq j \quad (4.2)$$

The generalization to the time dependent case is described in appendix C. Equation (4.1) can be written in the form of a master equation of a stochastic process

$$\frac{d|c_i|^2}{dt} = \sum_{j \neq i} [T_{ij}|c_j|^2 - T_{ji}|c_i|^2] \quad (4.3)$$

where T_{ij} is the transition rate for $|\Psi_i\rangle \rightarrow |\Psi_j\rangle$.

Our task is to solve (4.3) given the known time dependence of $c_i(t)$ that follow from the Schrödinger equation in (4.1). Note that in the theory of stochastic processes, one would usually be presented with the problem the other way around, i.e. given the transition rates T_{ij} one would have to solve for the instantaneous probabilities $|c_i|^2$. There is a very natural solution for the transition rates of the form¹²

$$T_{ij} = \max \left(-\frac{2}{\hbar} \text{Im} \left[\frac{c_i}{c_j} \langle \Psi_j | H | \Psi_i \rangle \right], 0 \right) \quad (4.4)$$

which is the expression quoted in (2.4). Note that if T_{ij} is non-vanishing, then $T_{ji} = 0$ and vice-versa. The form makes clear that it is the mis-alignment between the energy eigenstate basis and the quantum microstate basis that

¹²It might be thought unnecessary to define the stochastic process as a continuous process in time and indeed in earlier work [19, 20] we defined the process with some temporal cut off chosen to be much smaller than any characteristic time scales in the problem. This was done to avoid a problem with ‘‘crossovers’’ discussed in detail in [19, 20]. But the problem with crossovers only arises from taking an unrealistic choice for the observables as in Example 2 in appendix B. Realistic choices as described in section 3.1 do not suffer from this problem. So in the end whether to take a continuous or discrete version of the stochastic process is simply a matter of computational expedience.

drives the transitions.

The transition probabilities (4.4) are actually not the most general set of probabilities that are consistent with (4.1). The freedom involves

$$T_{ij} \rightarrow T_{ij} + \frac{\Omega_{ij}}{|c_j|^2}$$

for any symmetric $\Omega_{ij} > 0$. However, it is our hypothesis that $\Omega_{ij} = 0$ because T_{ij} is then not only the simplest and most natural choice depending as it does on just the matrix elements between the initial and final quantum microstates $\langle \Psi_j | H | \Psi_i \rangle$ but also the choice has some important implications for the ergodicity of the stochastic process that will be important for obtaining a sensible classical limit.¹³

4.1. Ergodicity and the classical limit

The key feature of the stochastic process is that it preserves the equality of the ensemble average with respect to the coarse grained density operator ρ with the expectation value (2.2). The question is how is the quantum expectation value $\langle \Psi | \mathcal{O} | \Psi \rangle$, related to the time sequence of expectation values $\langle \Psi_{i(t)} | \mathcal{O}_n | \Psi_{i(t)} \rangle$ of the stochastic process? The situation here is very similar to the relation between ensemble averages and long time averages in classical SM. The issue rests on the ergodicity of the stochastic process. In the ergodic situation ensemble averages capture long time averages. In classical SM, it is generally very difficult to establish ergodicity because classical mechanics is deterministic and classical motion corresponds to a trajectory in phase space. It is then a delicate matter to argue that the trajectory samples phase space in an ergodic way. The present quantum situation is much simpler, partly because the quantum microstates are discrete, and, partly because they follow stochastic dynamics. Moreover, the dynamics is a conceptually simple continuous time Markov chain for which ergodicity is—conceptually at least—straightforward to establish.

¹³The transition rates (4.4) are identical to written down by Bell in his theory of beables [24] while the generalization with time dependent $|\Psi_i\rangle$ described in appendix C was considered by Bacciagaluppi and Dickson [25]. In Bell's case, the $|\Psi_i\rangle$ were taken to be eigenstates of some set of preferred observables, the *beables*. In our approach, the set of observables is not fixed once and for all in some global way, rather it is dictated locally by the way the quantum system interacts macroscopically with its surroundings.

The issue of ergodicity, when it holds and when it is broken, is the key to understanding how a classical world can emerge from quantum mechanics in our proposal. When the system is in thermal equilibrium, the expectation values $\langle \Psi | \mathcal{O}_n | \Psi \rangle$ are approximately constant. We then argue in section 8 that the probabilities $|c_i(t)|^2$ are also approximately constant, a fact that actually follows from von Neumann's quantum ergodic theorem [4]. The stochastic process will effectively be ergodic over a long time T if the integrated transition probabilities $p_{ij}(t+T, t)$ become independent of j so that memory of the initial state is lost. It follows from (2.5) that $|c_i|^2$ is precisely the probability that the system has expectation values $\langle \Psi_i | \mathcal{O}_n | \Psi_i \rangle$. In particular, from a macroscopic point of view which involves a long time average compared with microscopic time scales, the long time average is captured by the ensemble average over the quantum microstates \mathcal{M} with probabilities $|c_i|^2$:

$$\textbf{Equilibrium: } \quad \frac{1}{T} \int_0^T dt \langle \Psi_{i(t)} | \mathcal{O}_n | \Psi_{i(t)} \rangle \approx \sum_i |c_i|^2 \langle \Psi_i | \mathcal{O}_n | \Psi_i \rangle = \langle \Psi | \mathcal{O}_n | \Psi \rangle \quad (4.5)$$

for large enough T . In the above, the last equality is the microscopic Born rule (3.4). We will investigate ergodicity in the context of SM more fully in section 8.

However, we do not expect that the stochastic process of a system will always be ergodic; indeed if it were so then the classical limit would not involve any deterministic dynamics. So the issue of ergodicity breaking is key to understanding how macroscopic systems can obey Newton's laws. Ergodicity breaking occurs when pairs of quantum microstates $|\Psi_i\rangle$ and $|\Psi_j\rangle$ are macroscopically distinct because it would then follow that the matrix elements $\langle \Psi_j | H | \Psi_i \rangle$ are vanishingly small and so the transition rate (2.4) between the two states and the integrated probability $p_{ij}(t, t')$ would be hugely suppressed and ergodicity over macroscopic time scales would inevitably be broken. For any given initial quantum microstate $|\Psi_i\rangle$, the stochastic process would then effectively only explore a more limited ergodic subspace of \mathcal{M} . This subspace consists of quantum microstates that are tightly clustered around a classical trajectory, as we illustrate in section 5.

A macroscopic object will typically have elements of ergodic and non-ergodic behavior: the position of its moving parts in space will involve a breaking of ergodicity whereas, at the same time, it will typically be in thermal equilibrium, behavior that involves an ergodic component corresponding to the

appropriate thermal ensemble.

4.2. Locality

The issue of locality in quantum mechanics is full of misunderstandings.¹⁴ It cannot be over emphasized that quantum mechanics is a perfectly local theory. This is particular clear in QFT where locality and causality are built into the fabric of the formalism: interaction terms in the Hamiltonian are spatially local; and causality is manifested in the fact that space-like separated operators commute.

The worry might be that with the stochastic process we have introduced some non-locality. We need to show that the stochastic dynamics associated to a set of observables $\mathcal{A} = \{\mathcal{O}_n\}$ that act on spatially localized degrees of freedom is only driven by interactions that are local to these degrees of freedom. In the case of a spatially localized subsystem \mathcal{A} , we can expect the Hilbert space to factorize $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_E$ with the observables acting as $\mathcal{O}_n \otimes 1_E$ and a Hamiltonian with the form

$$H = H_A \otimes 1_E + 1_A \otimes H_E + H_{int} \quad (4.6)$$

We will consider the physically realistic situation where \mathcal{A} consists of a set of commuting operators acting on \mathcal{H}_A factor as described in section 3.2. In this case, the quantum microstates on the total system take the form (3.14), where the states $|\psi_{ia}\rangle$ are the simultaneous eigenstates of the commuting set \mathcal{A} . The key point is that the set of states $\{|\psi_{ia}\rangle\}$ are orthogonal while the set $\{|\tilde{\psi}_{ia}\rangle\}$ are not. Therefore,

$$\langle \Psi_i | 1_A \otimes H_E | \Psi_j \rangle = 0 \quad , \quad i \neq j \quad (4.7)$$

¹⁴We refer to [27] for an insightful discussion of these kinds of misunderstandings. In particular, quantum mechanics famously violates Bell’s theorem but is completely consistent with experiments. So the theorem holds for theories that are not realized in nature. It is hard to see that it has any implications for quantum mechanics itself and, in particular, it most certainly does not imply that quantum mechanics is non-local. All “delayed choice” type experiments designed, apparently, to highlight the spooky, non-local nature of quantum mechanics simply reveal the correlations between entangled states: there is no non-locality.

on account of the fact that $\langle \psi_{ia} | \psi_{jb} \rangle = \delta_{ij} \delta_{ab}$; hence, the transition rates take the form

$$T_{ij} = \max \left(-\frac{2}{\hbar} \text{Im} \left[\frac{c_i}{c_j} \langle \Psi_j | \{ H_A \otimes 1_E + H_{int} \} | \Psi_i \rangle \right], 0 \right) \quad (4.8)$$

It follows that the stochastic process depends on the internal dynamics of the subspace \mathcal{A} , via H_A , as well as the local interactions of A with the environment E , via H_{int} . Importantly, though, the dynamics of the environment involving H_E are irrelevant. This proves that the stochastic dynamics is local to the degrees of freedom on which the observables \mathcal{A} act.

4.3. Patching together a classical world

The discussion above has implications for the case when the set of operators $\mathcal{A} \cup \mathcal{B}$ where the subsets consists of local observables acting on spatially separated degrees of freedom. The question is how the perspectives of the system provided individually by \mathcal{A} and \mathcal{B} are related to the global perspective provided by $\mathcal{A} \cup \mathcal{B}$. If the macroscopic world is to be classical then it must be the different views can be integrated together, at least to high precision: a classical world is, after all, an approximation.

The first point is that the quantum microstates of $\mathcal{A} \cup \mathcal{B}$ can naturally be given a pair of labels $|\Psi_{i_1 i_2}\rangle$ to reflect the decomposition of $\mathcal{A} \cup \mathcal{B}$. So the label i is split into the pair i_1, i_2 which labels the eigenstates of the sets \mathcal{A} and \mathcal{B} , respectively. The quantum microstates of the individual localized sets \mathcal{A} or \mathcal{B} are given by the partial sums

$$c_{i_1} |\Psi_{i_1}\rangle = \sum_{i_2} c_{i_1 i_2} |\Psi_{i_1 i_2}\rangle \quad , \quad c_{i_2} |\Psi_{i_2}\rangle = \sum_{i_1} c_{i_1 i_2} |\Psi_{i_1 i_2}\rangle \quad (4.9)$$

and

$$|c_{i_1}|^2 = \sum_{i_2} |c_{i_1 i_2}|^2 \quad , \quad |c_{i_2}|^2 = \sum_{i_1} |c_{i_1 i_2}|^2 \quad (4.10)$$

Notice that the partial sums in (4.9) are forced by the non-degeneracy condition (3.5) because

$$\langle \Psi_{i_1 i_2} | \mathcal{O}_n \otimes 1_B | \Psi_{i_1 j_2} \rangle = \mu_n \delta_{i_1 j_2} \quad , \quad \forall \mathcal{O}_n \in \mathcal{A} \quad (4.11)$$

When $c_{i_1 i_2} \neq c_{i_1} c_{i_2}$ the state has entanglement between the two spatially separated regions. Entanglement could occur if the two subsystems were the

two measuring devices in an EPR type experiment as we describe in section 7.

The stochastic process of the total set $\mathcal{A} \cup \mathcal{B}$ is just a refinement of the individual stochastic processes of the two subsets with

$$T_{i_1 j_1}^{(\mathcal{A})} = \sum_{i_2, j_2} T_{i_1 i_2, j_1 j_2}^{(\mathcal{A} \cup \mathcal{B})} \left| \frac{c_{j_1 j_2}}{c_{j_1}} \right|^2, \quad T_{i_2 j_2}^{(\mathcal{B})} = \sum_{i_1, j_1} T_{i_1 i_2, j_1 j_2}^{(\mathcal{A} \cup \mathcal{B})} \left| \frac{c_{j_1 j_2}}{c_{j_2}} \right|^2 \quad (4.12)$$

In order that the two stochastic processes associated to \mathcal{A} and \mathcal{B} are genuinely independent requires that:

1. The only non-vanishing transitions of the global system are $|\Psi_{j_1 i_2}\rangle \rightarrow |\Psi_{i_1 i_2}\rangle$ and $|\Psi_{i_1 j_2}\rangle \rightarrow |\Psi_{i_1 i_2}\rangle$.
2. The rate $|\Psi_{j_1 i_2}\rangle \rightarrow |\Psi_{i_1 i_2}\rangle$ is independent of i_2 and that for $|\Psi_{i_1 j_2}\rangle \rightarrow |\Psi_{i_1 i_2}\rangle$ is independent of i_1 .

These conditions together imply that the only non-vanishing transition rates of the total system are

$$T_{i_1 i_2, j_1 i_2}^{(\mathcal{A} \cup \mathcal{B})} = T_{i_1 j_1}^{(\mathcal{A})}, \quad T_{i_1 i_2, i_1 j_2}^{(\mathcal{A} \cup \mathcal{B})} = T_{i_2 j_2}^{(\mathcal{B})} \quad (4.13)$$

Recalling the formula for the transition rates (4.4), the first condition above follows from the fact that realistic Hamiltonians only couple spatially local degrees of freedom and so cannot drive transitions at two spatially separated regions simultaneously; hence

$$\langle \Psi_{i_1 i_2} | H | \Psi_{j_1 j_2} \rangle = 0 \quad \text{or} \quad i_i \neq j_1 \quad \text{and} \quad i_2 \neq j_2 \quad (4.14)$$

The second condition above is more exacting. To investigate the issues, we need to consider the more detailed form of the quantum microstates. Let us take the Hilbert space to have the factorized form $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_E \otimes \mathcal{H}_2$ to reflect the spatial locality of the observables with elements of \mathcal{A} acting as $\mathcal{O} \otimes 1_E \otimes 1_2$ and elements of \mathcal{B} as $1_1 \otimes 1_E \otimes \mathcal{O}$. The subsystem \mathcal{H}_E is the environment. We can write the quantum microstates of $\mathcal{A} \cup \mathcal{B}$, generalizing (3.14), more concretely as

$$|\Psi_{i_1 i_2}\rangle = \sum_{a_1 a_2} |\psi_{i_1 a_1}\rangle \otimes |\tilde{\psi}_{i_1 a_1, i_2 a_2}\rangle \otimes |\phi_{i_2 a_2}\rangle \quad (4.15)$$

which shows that the subsystems are generally entangled with the environment. Note that the sets $\{|\psi_{i_1 a_1}\rangle\}$ and $\{|\phi_{i_2 a_2}\rangle\}$ are orthonormal but the set $\{|\tilde{\psi}_{i_1 a_1, i_2 a_2}\rangle\}$ is not. So even though H cannot couple degrees of freedom in \mathcal{H}_1 and \mathcal{H}_2 directly, the question is whether it can couple degrees of freedom in \mathcal{H}_1 and those in the environment that are entangled with those in \mathcal{H}_2 , or vice versa?

Any such coupling must be very small because, by locality, degrees of freedom in \mathcal{H}_1 can only couple locally with the environment and the latter is a much bigger system than the subsystems \mathcal{H}_1 and \mathcal{H}_2 and it is unlikely the degrees of freedom of the environment that are entangled with \mathcal{H}_2 are also spatially local to \mathcal{H}_1 . So, for realistic systems, to a high degree of accuracy, the second condition above is likely fulfilled and the stochastic process associated to $\mathcal{A} \cup \mathcal{B}$ consists of two independent processes \mathcal{A} and \mathcal{B} and the perspectives of \mathcal{A} and \mathcal{B} can be integrated into the global view $\mathcal{A} \cup \mathcal{B}$.

We will discuss more aspects of situation with two sets of spatially separated observables in section 7.

5. A Particle

In this section, we develop our formalism for the case of a particle moving in one dimension, although the generalization to more dimensions is then obvious. In this crude model, we will simply ignore the internal structure of the particle and take it to be a quantum particle with a wave function $\Psi(x)$. We will also simplify the situation even further by taking \mathcal{A} to consist of a suitably coarse grained definition of the position operator only. The momentum will be added in later.

The observables consist of the set $\mathcal{A} = \{\Pi_i^{(x)}\}$ defined in (3.12). In this case, the observables are mutually commuting and so the quantum microstates are simultaneous eigenstates. More precisely, the eigenstates have the degenerate eigenspaces that are spanned by the eigenstates $|x\rangle$ in each interval $x \in [x_{i-1}, x_i]$, with

$$\Pi_i^{(x)} = \begin{cases} |x\rangle & x \in [x_{i-1}, x_i] \\ 0 & \text{otherwise} \end{cases} \quad (5.1)$$

The quantum microstates follow from the decomposition of $|\Psi\rangle$:

$$|\Psi\rangle = \sum_i c_i |\Psi_i\rangle \quad , \quad |\Psi_i\rangle = \frac{1}{c_i} \int_{x_{i-1}}^{x_i} dx \Psi(x) |x\rangle \quad (5.2)$$

with

$$|c_i|^2 = \int_{x_{i-1}}^{x_i} dx |\Psi(x)|^2 \quad (5.3)$$

The quantum micro states in this case, are spatially localized in the interval $[x_{i-1}, x_i]$ and represent states that have a good classical interpretation on scales $\gg \delta_x$.

5.1. The stochastic process

The transition rates are given in (4.4). In order to calculate them, we note that

$$c_j \Psi_j(x) = (\theta(x - x_{j-1}) - \theta(x - x_j)) \Psi(x) \quad (5.4)$$

and therefore the matrix elements $\langle \Psi_j | H | \Psi_i \rangle$ are only non-vanishing if $i = j \pm 1$. Taking $i = j + 1$, we have

$$\text{Im} \left[\frac{c_{j+1}}{c_j} \langle \Psi_j | H | \Psi_{j+1} \rangle \right] = -\frac{\hbar^2}{M|c_j|^2} \int dx \theta(x - x_j) \delta(x - x_j) \text{Im} \left[\Psi^* \frac{\partial \Psi}{\partial x} \right] \quad (5.5)$$

with a similar expression for $i = j - 1$. Evaluating the integral¹⁵ gives

$$\text{Im} \left[\frac{c_i}{c_j} \langle \Psi_j | H | \Psi_i \rangle \right] = \frac{\hbar^2}{2M|c_j|^2} \text{Im} \left(\Psi^* \frac{\partial \Psi}{\partial x} \Big|_{x=x_{j-1}} \delta_{i,j-1} - \Psi^* \frac{\partial \Psi}{\partial x} \Big|_{x=x_j} \delta_{i,j+1} \right) \quad (5.6)$$

So the non-vanishing transitions rates involve $|\Psi_j\rangle \rightarrow |\Psi_{j\pm 1}\rangle$ with

$$T_{j\pm 1,j} = \max \left(\pm \frac{\hbar^2}{M|c_j|^2} \text{Im} \left[\Psi^* \frac{\partial \Psi}{\partial x} \Big|_{x=x_j, x_{j-1}} \right], 0 \right) \quad (5.7)$$

The expression in the brackets here is, of course, just the Schrödinger probability current evaluated at the interface between the j^{th} cell and one of adjacent cells divided by the probability to be in the j^{th} cell. The transitions only go in the direction of the current.

¹⁵Using $\int dx \theta(x - x_j) \delta(x - x_j) f(x) = \frac{1}{2} f(x_j)$.

We can now investigate this process in certain interesting limits. An important issue to consider is the relative magnitude of the width of the wave function of the particle Δ and the resolution scale δ_x . For a macroscopic object, we expect that the wave function will be very narrow compared with the coarse graining scale $\delta_x \gg \Delta$. We will consider this case first. However, it is also interesting to consider situations where the wave function is smooth on the resolution scale $\delta_x \ll \Delta$.

5.2. Narrow wave function $\delta_x \gg \Delta$

This limit, which is relevant to a macroscopic particle, is very simple because if at some time t_1 the center of the wave packet lies within a region $[x_{i-1}, x_i]$, the quantum microstate $|\Psi_i\rangle$ has a probability that is $|c_i(t_1)|^2 \approx 1$. Essentially in this case, the underlying state $|\Psi\rangle \approx |\Psi_i\rangle$ is the only quantum microstate that has any appreciable probability. The center of wave packet will, on account of Ehrenfest's theorem, follow a classical trajectory $x(t)$ and so as the wave packet moves across into a neighboring region $[x_{i+1}, x_{i+2}]$, at, say time t_2 , the quantum microstate changes to $|\Psi_{i+1}\rangle$ with a probability close to unity. It is clear, therefore, that with a probability close to unity, the resulting sequence of quantum microstates just describes a coarse graining of the classical trajectory; that is, at time t the micro state is $|\Psi_{i(t)}\rangle$ where $i(t)$ is determined by the condition $x(t) \in [x_{i(t)-1}, x_{i(t)}]$. This is illustrated in figure 2.

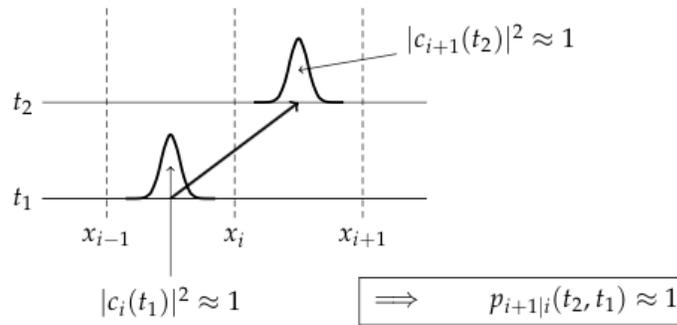


Figure 2: When the wave function is narrow compared with the resolution scale, the quantum microstate just follows the peak of $|\Psi(x)|^2$ and the stochastic process effectively leads to deterministic dynamics.

5.3. Smooth wave function $\delta_x \ll \Delta$

The other limit, occurs when the width of the wave function is much greater than the resolution scale $\delta_x \ll \Delta$. In this case, the resulting dynamics is fundamentally stochastic. However, even in this case, if the wave function has the typical WKB form

$$\Psi(x) = R(x)e^{iS(x)/\hbar} \quad (5.8)$$

then a classical trajectory emerges. In the semi-classical limit $S(x)$, which is a smooth function on macroscopic scales, is identified with the classical action of the particle and the velocity of a classical trajectory is simply

$$v = \frac{1}{M} \frac{\partial S}{\partial x} \quad (5.9)$$

We are assuming that the amplitude $R(x)$ is a smooth function on the scale δ_x . In this case, the probabilities associated to the quantum microstates $|\Psi_i\rangle$ are $|c_j|^2 \approx R(x_j)^2 \delta_x$.

The transition rates (5.7) are exact but now we can use the form (5.8) for the wave function and the approximation $|c_j|^2 \approx R(x_j)^2 \delta_x$. This yields the transition probabilities

$$\begin{aligned} T_{j-1,j} &\approx \frac{1}{M\delta_x} |S'(x_j)|, \quad T_{j+1,j} = 0 \quad \text{if } S'(x_j) < 0 \\ T_{j-1,j} &= 0, \quad T_{j+1,j} \approx \frac{1}{M\delta_x} S'(x_j) \quad \text{if } S'(x_j) > 0 \end{aligned} \quad (5.10)$$

Given these transitions rates, it is clear that the stochastic process is quite simple: the particle can hop from the interval $[x_{j-1}, x_j]$ to a neighboring one, depending on the sign of $S'(x_j)$. Notice that \hbar has completely dropped out of the expression for T_{ij} . In fact this stochastic process of the type we have derived is identical to the one considered by Vink [28] whose starting point and general approach is rather different,¹⁶ but whose analysis we can still borrow.

The question is whether the resulting stochastic process gives a behavior for

¹⁶In Vink's approach, the set up is rather different with the wave function being coarse grained on a spatial lattice whose spacing is taken to zero at the end. In our approach, our coarse graining scale is the spatial resolution scale which is kept finite.

the position that is recognizably classical. This does not seem at all obvious because classically there should just be a single trajectory. What we will argue is that the stochastic trajectories are clustered around the classical trajectory.

In order to investigate the process we introduce a small temporal cut off δt and then investigate the process starting at, say $x = x_i$ over a time interval T which involves a large number $N = T/\delta t$ of time steps, but is not too large, so that $S'(x_j) = vM$, $|i - j| < N$, is approximately constant and let us suppose $v > 0$. For each time step, we have must have

$$p \equiv p_{i+1,i} = \frac{v}{\delta_x} \delta t \ll 1 \quad (5.11)$$

which can be achieved by having a small enough δt . After the N steps, the position x_j with $j = k + i$ has a binomial probability distribution

$$\binom{N}{k} p^k (1-p)^{N-k} \quad (5.12)$$

whose average position is

$$\langle x \rangle = x_i + vT \quad (5.13)$$

which shows that v is, indeed, interpreted as the velocity of the average of the distribution. The variance of the distribution is

$$\langle x^2 \rangle - \langle x \rangle^2 = vT \delta_x \left(1 - \frac{v}{\delta_x} \delta t \right) \approx X \delta_x \quad (5.14)$$

where $X = vT$ is the distance moved in our small interval of time and we have used (5.11). So for sub-macroscopic resolution δ_x the variance is small.

What we have shown here is that even when the wave function is wide compared with the resolution a classical trajectory can emerge. This suggests that there should be a way to relate the situation with a fixed wave function with different resolutions scales by changing the Heisenberg cut.

5.4. Changing the Heisenberg cut

One could view the results of sections 5.2 and 5.3 as pertaining to a given wave packet with fixed width but with a coarse and finer resolution scale

δ_x , respectively. A coarser view of the dynamics appears more deterministic than the finer view. The question then is, more generally, how does the stochastic dynamics behave as one varies the Heisenberg cut. In the present example, there is a simple way to do this which involves a kind of blocking transformation that doubles the Heisenberg cut $\delta'_x = 2\delta_x$. The new coarser description involves projection operators

$$\Pi_i^{(x)'} = \Pi_{2i-1}^{(x)} + \Pi_{2i}^{(x)} \quad (5.15)$$

It follows that the new quantum microstates are simply

$$c'_i |\Psi'_i\rangle = c_{2i-1} |\Psi_{2i-1}\rangle + c_{2i} |\Psi_{2i}\rangle \quad (5.16)$$

where $|c'_i|^2 = |c_{2i-1}|^2 + |c_{2i}|^2$. The blocking transformation has a simple effect on the stochastic dynamics; the transition rates of the new process are related to the old one via

$$T'_{ij} = \frac{1}{|c'_j|^2} (T_{2i-1,2j-1} |c_{2j-1}|^2 + T_{2i-1,2j} |c_{2j}|^2 + T_{2i,2j-1} |c_{2j-1}|^2 + T_{2i,2j} |c_{2j}|^2) \quad (5.17)$$

Performing this kind of blocking transformation allows us to relate the two distinct regimes in sections 5.2 and 5.3.

5.5. Including the momentum

The simple model that we have described can only be partly correct because we have avoided any mention of the particle's momentum. Classically the particles state is specified by both a position and momentum and so we need to consider enlarging \mathcal{A} to include coarse grained momentum operators.¹⁷

The wave function of a macroscopic object is expected to be narrow in both position and momentum space compared with the resolution scales δ_x and δ_p and so there is only one quantum microstate, the original state itself $|\Psi\rangle$, with a probability close to one. In this classical limit, the stochastic process becomes essentially deterministic and the stochastic trajectory localizes around a classical trajectory in phase space as illustrated in figure 3.

¹⁷In fact, without the momentum it would be possible for two apparently decoherent wave packets that cross in space to interfere. This is because classical trajectories in configuration space can intersect. But once momentum is included the trajectories in phase space cannot intersect and these kind of interferences are excluded.

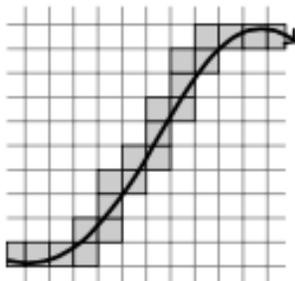


Figure 3: The classical limit where the wave packet is much narrower than the size of the cells the sequence of quantum microstates associated to cells in phase space just follows the expectation values of x and p which Ehrenfest's theorem implies is the classical trajectory.

6. Quantum Measurement

In this section, we consider a very simple measuring scenario to see how the formalism can lead to definite macroscopically distinct outcomes with probabilities that are given by Born's rule (at least in the case of an efficient measuring device).

In the model, the total system includes a microscopic system that has an observable A that we want to measure. Let $|\psi_a\rangle$ be the eigenstates of A and the initial state of the microscopic system is $\sum_a \lambda_a |\tilde{\zeta}_a\rangle$. The system is coupled to a measuring device with a Hamiltonian that ensures an evolution of the combined system of the form¹⁸

$$\left[\sum_a \lambda_a |\tilde{\zeta}_a\rangle \right] \otimes |\psi_0\rangle \rightarrow \sum_a \lambda_a |\tilde{\zeta}_a(t)\rangle \otimes |\psi_a(t)\rangle \quad (6.1)$$

Note that in general the states of the microscopic system will be changed as a result of the interaction with the measuring device. The states $|\tilde{\zeta}_a(t)\rangle$ remain normalized but are not necessarily orthogonal.

¹⁸In the following, we have assumed that the measuring device is 100% efficient. Making the generalization to a more realistic measuring device does not change the central conclusion that we come to below that definite outcomes are obtained. However, the resulting probabilities will no longer be equal to $|\lambda_a|^2$. So the derivation of the usual Born rule relies on the fact that the measuring device is 100% efficient.

The idea is that the states of the measuring device $|\psi_a(t)\rangle$ become macroscopically distinct for $t \geq T$. After this time a measurement could be said to have occurred. In our simple model, we will suppose that the measuring device consists of a pointer with a position. So the states of the measuring device that are entangled with microscopic system are wave functions $\psi_a(x, t)$. The idea is that the measuring device is designed in such a way that for $t > T$, each component $\psi_a(x, t)$ is a narrow wave packet peaked around a distinct position $x^{(a)}$. Note that each component $\psi_a(x, t)$ is separately normalized.

If we now apply our interpretation by using the coarse-grained position operators $\mathcal{A} = \{\Pi_i^{(x)}\}$ as in section 5. At some time $t > T$, the quantum microstates are

$$|\Psi_i\rangle = \frac{1}{c_i} \int_{x_{i-1}}^{x_i} dx \sum_a \lambda_a |\tilde{\zeta}_a(t)\rangle \otimes \psi_a(x, t) |x\rangle \quad (6.2)$$

with

$$|c_i|^2 = \int_{x_{i-1}}^{x_i} dx \sum_{ab} \lambda_a^* \lambda_b \langle \tilde{\zeta}_a(t) | \tilde{\zeta}_b(t) \rangle \psi_a(x, t)^* \psi_b(x, t) \quad (6.3)$$

For the measuring device to be accurate, we need its spatial resolution scale δ_x to be much smaller than the separation of $x^{(a)}$. It is also realistic to assume that the spatial resolution scale δ_x is much larger than the width of the individual wave packets $\psi_a(x, t)$. Given these assumptions, it follows that $|c_i|^2$ is only appreciable if some $x^{(a)} \in [x_{i-1}, x_i]$, for some particular i which we can call i_a . So only the diagonal term $a = b$ contributes in (6.3). Note that only one $x^{(a)}$ can lie in any interval in which case, it follows for this interval that

$$|c_{i_a}|^2 \approx |\lambda_a|^2 \quad (6.4)$$

Since the initial quantum microstate is unique, $|\lambda_a|^2$ represents the probability that the quantum microstate is $|\Psi_{i_a}\rangle$ at time $t > T$. This amounts to a derivation of the usual Born rule. The matching of the components of the underlying state to the quantum microstates is illustrated in figure 4.

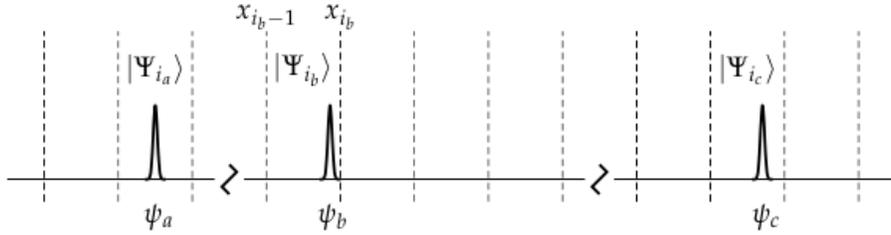


Figure 4: The measurement scenario. In the final state, each measurement outcome corresponds to a distinct quantum microstate which are ergodically disjoint.

The last important point is that once the system is in the quantum microstate $|\Psi_i\rangle$ there is only a minute probability for making a transition to a macroscopically distinct state $|\Psi_j\rangle$. So it is the breaking of ergodicity that leads to distinct classical outcomes and hence solves the measurement problem. The situation is illustrated in figure 5.

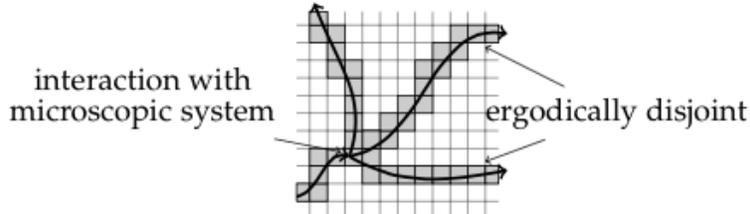


Figure 5: A picture of a quantum measurement in the phase space of a very simple measuring device with one degree of freedom. The system follows a coarse graining of the classical trajectory. At the moment of interaction with the microscopic system, the trajectory splits into several distinct trajectories that soon become ergodically disjoint. The measurement outcome is determined by which of the disjoint trajectories the system ends up in.

At the end of the measurement, from the point of view of \mathcal{A} it is prudent to reduce the underlying state vector to the component $|\tilde{\zeta}_a(t)\rangle \otimes \psi_a(x, t)$ that is actually realized as a quantum microstate.

7. EPR

In this section, we give an account of Bohm's version [29, 30] of the classic thought experiment of Einstein, Podolsky, and Rosen (EPR) [31]. Anticipating the result, we find that is no non-locality or breakdown of causality, the experiment simply reveals the correlation of an entangled state of two qubits.

In order to describe spin measurements on the qubits, we introduce a very simple measuring device with three quantum microstates $|A_0\rangle$ and $|A_{\pm}\rangle$. In this simple model the quantum microstates are actually the possible macrostates of the measuring device; in other words, in order to focus on the essential details, we avoid the fact that, in reality, the measuring device is a complicated thermodynamic object with many degrees of freedom. The measuring device is designed so that the solution of the Schrödinger equation for its interaction with a qubit from an initial time $t = 0$ to a final time T is of the form

$$|A_0\rangle \otimes (c_+ |z^+\rangle + c_- |z^-\rangle) \rightarrow c_+ |A_+\rangle \otimes |z^+\rangle + c_- |A_-\rangle \otimes |z^-\rangle \quad (7.1)$$

The stochastic process associated to the observables of the measuring device yield integrated probabilities

$$p_{+|0}(T, 0) = |c_+|^2 \quad , \quad p_{-|0}(T, 0) = |c_-|^2 \quad (7.2)$$

In this simple model, we can identify the quantum microstates in the final state with the reduced state vectors of the Copenhagen interpretation. In a realistic measuring device, the reduced state vector only emerges when the ergodic components in (2.6) become identifiable. This happens in a macroscopically short time but is ultimately a dynamical question.

The experiment is described in figure 6

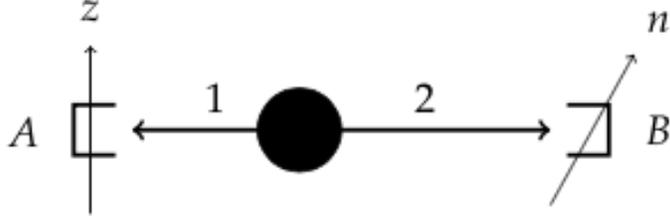


Figure 6: The EPR-Bohm thought experiment. Two qubits in the entangled state $|\Phi\rangle$ are produced at the source and then recoil back-to-back towards 2 qubit detectors A and B designed to measure the component of the spin along directions z and n , respectively. In our set up, we choose an inertial frame for which the interaction between A and 1 happens before B and 2.

with the qubits being produced in the singlet state

$$|\Phi\rangle = \frac{1}{\sqrt{2}}(|z^+z^-\rangle - |z^-z^+\rangle) \quad (7.3)$$

and so the initial state of the overall system is

$$|\Psi(t_1)\rangle = |A_0B_0\rangle \otimes |\Phi\rangle \quad (7.4)$$

In a certain inertial frame, at time $t_A > t_1$, the measuring device A , set to measure the z component of the spin, interacts with qubit 1. After a short time, the state of qubit 1 becomes entangled with A and the state of the total system becomes, for $t_2 > t_A$,

$$|\Psi(t_2)\rangle = \frac{1}{\sqrt{2}}(|A_+B_0z^+z^-\rangle - |A_-B_0z^-z^+\rangle) \quad (7.5)$$

Then at a time t_B the measuring device B , set to measure the component of the spin along a vector n at an angle 2θ to the z axis, interacts with qubit 2

producing, for $t_3 > t_B$, a state¹⁹

$$|\Psi(t_3)\rangle = \frac{1}{\sqrt{2}} (\sin \theta |A_+ B_+ z^+ n^+\rangle + \cos \theta |A_- B_- z^+ n^-\rangle - \cos \theta |A_- B_+ z^- n^+\rangle + \sin \theta |A_+ B_- z^- n^-\rangle) \quad (7.6)$$

Just to emphasize, the states $|\Psi(t_i)\rangle$, $i = 1, 2, 3$ follow from solving the Schrödinger equation of the system with a Hamiltonian that only has local interaction terms between the measuring devices and their associated qubit, A with 1 and B with 2.

In this system, there are three different perspectives associated to the observables of the two measuring devices \mathcal{A} and \mathcal{B} as well as the global view associated to $\mathcal{A} \cup \mathcal{B}$. It is important to appreciate that the descriptions \mathcal{A} and \mathcal{B} cannot answer questions about correlations between the measuring devices, it is only the more refined global view $\mathcal{A} \cup \mathcal{B}$ that can do that. Figure 7 shows the three Markov processes.

The quantum microstates of \mathcal{A} and \mathcal{B} are related to those of $\mathcal{A} \cup \mathcal{B}$ precisely as in (4.9). It is important that the changes in the quantum microstates are driven by local interactions: A with qubit 1 at t_A and B with qubit 2 at t_B . It is worth remarking on the fact that the quantum microstates associated to \mathcal{A} and \mathcal{B} by focussing on, say, \mathcal{A} , we forego any knowledge of \mathcal{B} , and vice-versa. The only way to ask about joint properties of the two measuring devices is via the quantum microstates associated to $\mathcal{A} \cup \mathcal{B}$. The quantum microstates of the latter are always refinements of those of \mathcal{A} and \mathcal{B} as is clear from the last column in figure 7.

¹⁹The relation between the basis of qubit states along the z and n axes is

$$|z^+\rangle = \cos \theta |n^+\rangle - \sin \theta |n^-\rangle$$

and

$$|z^-\rangle = \sin \theta |n^+\rangle + \cos \theta |n^-\rangle$$

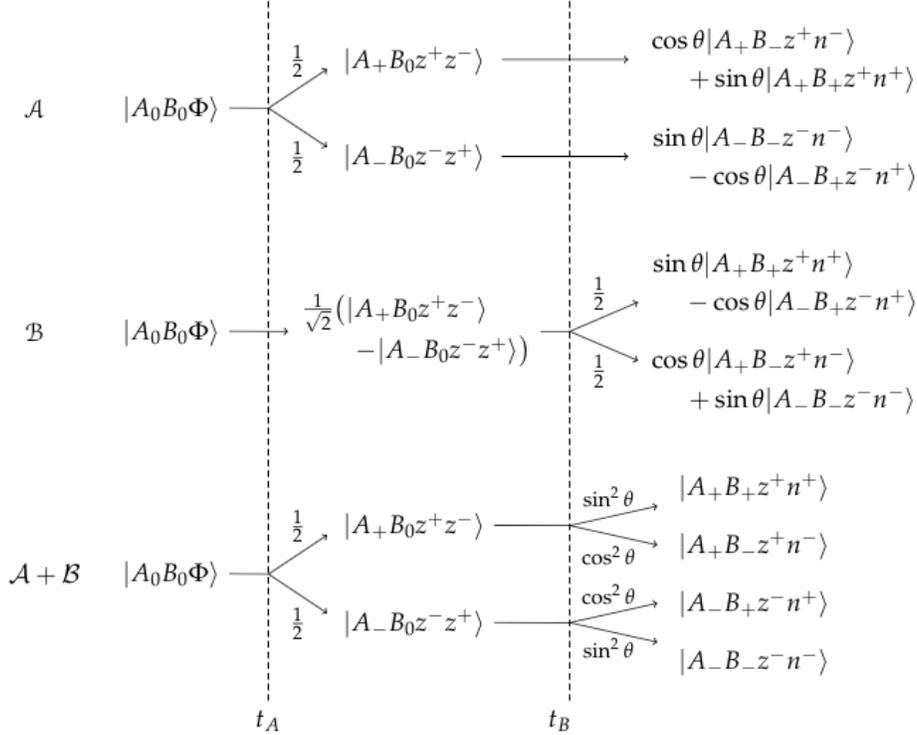


Figure 7: The Markov chains of the quantum microstates (the macrostates or reduced state vectors in this simple model) of the sets of observables \mathcal{A} and \mathcal{B} associated to the two measuring devices and the global view $\mathcal{A} + \mathcal{B}$ at times t_1 , t_2 and t_3 between the measurements at t_A and t_B . Note that the quantum microstates of \mathcal{A} and \mathcal{B} always have the factored form $|A_i\rangle \otimes |\psi\rangle$ and $|B_i\rangle \otimes |\psi\rangle$ respectively, for $i = 0, \pm$ and for some $|\psi\rangle$ in the appropriate complementary subsystem. Also shown are the probabilities of the Markov chains.

,

With our simple description of the measuring devices, we can identify the quantum microstates of \mathcal{A} and \mathcal{B} with the reduced state vectors of the Copenhagen interpretation. It then becomes apparent that different local observers, in this case \mathcal{A} and \mathcal{B} , perform their own reduction of the state vector to take account the results of their local measurements.²⁰

²⁰The fact that each local observer has their own reduced state vector after making local measurements has been emphasized in [27].

One way to think of the relation between the refined view $\mathcal{A} \cup \mathcal{B}$ and, say \mathcal{A} is the following. When the measurement is made at B , the quantum microstates of $\mathcal{A} \cup \mathcal{B}$, for example $|A_+B_0z^+z^- \rangle$ splits into the pair $|A_+B_\pm z^+n^\pm \rangle$. But the set of observables \mathcal{A} cannot distinguish this pair of states:

$$\langle A_+B_i z^+ n^i | \mathcal{O}_n \otimes 1_B \otimes 1_1 \otimes 1_2 | A_+B_j z^+ n^j \rangle = \mu_n \delta_{ij} \quad , \quad \forall | \mathcal{O}_n \in \mathcal{A} \quad (7.7)$$

it is only observables in \mathcal{B} that can. Hence, the pair of states do not satisfy the non-degeneracy condition (3.5) and that is why the quantum microstate with respect to \mathcal{A} is the linear combination

$$\cos \theta |A_+B_- z^+ n^- \rangle + \sin \theta |A_+B_+ z^+ n^+ \rangle \quad (7.8)$$

The stochastic process associated to the set of observables \mathcal{A} for measuring device A gives the non-vanishing integrated probabilities

$$p_{+|0}^A(t_2, t_1) = p_{-|0}^A(t_2, t_1) = \frac{1}{2} \quad , \quad p_{+|+}^A(t_3, t_2) = p_{-|-}^A(t_3, t_2) = 1 \quad (7.9)$$

In particular nothing happens when the measurement is made at B . While that associated to \mathcal{B} gives

$$p_{0|0}^B(t_2, t_1) = 1 \quad , \quad p_{+|0}^B(t_3, t_2) = p_{-|0}^B(t_3, t_2) = \frac{1}{2} \quad (7.10)$$

In this case, nothing happens when the measurement is made at A . Just to be completely clear, in figure 7, the change in the quantum microstate of \mathcal{B} from t_1 to t_2 , whilst the measurement is being made at A , is just evolution by the Schrödinger equation. Likewise is the change in either of the quantum microstates of \mathcal{A} from t_2 to t_3 .

The global view $\mathcal{A} \cup \mathcal{B}$ provides a more refined description of the system that yields the correlations between measurements at A and B :

$$\begin{aligned} p_{+0|00}^{\mathcal{A} \cup \mathcal{B}}(t_2, t_1) &= p_{-0|00}^{\mathcal{A} \cup \mathcal{B}}(t_2, t_1) = \frac{1}{2} \\ p_{++|+0}^{\mathcal{A} \cup \mathcal{B}}(t_3, t_2) &= p_{--|-0}^{\mathcal{A} \cup \mathcal{B}}(t_3, t_2) = \sin^2 \theta \\ p_{+-|+0}^{\mathcal{A} \cup \mathcal{B}}(t_3, t_2) &= p_{-+|-0}^{\mathcal{A} \cup \mathcal{B}}(t_3, t_2) = \cos^2 \theta \end{aligned} \quad (7.11)$$

giving the overall probabilities

$$\begin{aligned} p_{++|00}^{\mathcal{A} \cup \mathcal{B}}(t_3, t_1) &= p_{--|00}^{\mathcal{A} \cup \mathcal{B}}(t_3, t_1) = \frac{1}{2} \sin^2 \theta \\ p_{-+|00}^{\mathcal{A} \cup \mathcal{B}}(t_3, t_1) &= p_{+-|00}^{\mathcal{A} \cup \mathcal{B}}(t_3, t_1) = \frac{1}{2} \cos^2 \theta \end{aligned} \quad (7.12)$$

which agree with the probabilities calculated in the Copenhagen interpretation using the Born rule.

So in this account of the EPR-Bohm thought experiment, the stochastic process yields the conventional quantum mechanical predictions in a way that involves only local interactions between measuring devices and qubits. There is no mystery or spooky action-at-a-distance, on the contrary there is simply the unveiling of an underlying correlation carried by the qubits—albeit of a kind that cannot be mimicked by a classical system.

8. The Emergence of Classical Statistical Mechanics

It should be apparent that our interpretation has some important implications for the way that classical SM arises from quantum mechanics. The first point, as we discussed in the introduction, is that there are many issues in the theory of classical SM that are not completely settled. Take the issue of probability. Essentially, as reviewed, for example, by Wallace [1], there are two distinct points of view called the *inferential* and *dynamical*, sometimes portrayed as the Gibbs view versus the Boltzmann. The former views probabilities as arising from our incomplete knowledge of the microstate given the macroscopic information that we have about a system. On the other hand, in the dynamical point of view, there is no fundamental need for probabilities because classical mechanics is deterministic. However, one can talk about the relative number of particles moving with such and such velocity and the dynamics is complicated and it is not possible to follow individual trajectories of microstates and so one has to talk about averages over suitably long times. In ergodic systems it is possible to link these points of view directly, but ergodicity is problematic to formulate and difficult to prove in general.

Probabilities, of course, naturally arise in quantum mechanics in the conventional Copenhagen interpretation and Wallace [1] makes the point that there is then no need to introduce probabilities in quantum SM twice. It would be much more natural if the probabilities in SM and the Copenhagen interpretation were essentially one and the same. This is what our formalism achieves with the potential to illuminate and solve the conceptual problems of classical SM.

In fact, a completely quantum approach to SM has its roots in the work of Schrödinger [3] and von Neumann [4] in the 1920s.²¹ We will be particularly interested in the approach of von Neumann leading to his *quantum ergodic theorem*. The idea is to focus on expectation values of suitable coarse grained observables with the aim of showing that the resulting values behave in the way expected in SM. One can either consider a system in isolation or one can focus on a small subsystem of a much larger system, corresponding to the microcanonical and canonical situations, respectively.

There are two important issues: does a system *equilibrate* and then if so does it *thermalize*? The issue of equilibration, concerns whether the expectation values of various coarse grained observables become constant, or at least constant for the vast majority of time after some initial transient regime. Note that this is a non-trivial question because the underlying quantum state of the system certainly does not approach a constant. The issue of thermalization is whether the equilibrium state gives expectation values that one would expect of the microcanonical or canonical ensembles. So the properties of equilibration and thermalization are properties of the state of the system with respect to a set of coarse grained observables $\mathcal{A} = \{\mathcal{O}_n\}$ rather than the underlying quantum state itself.

Let us consider an isolated system and suppose that it is described by a density operator $\rho(t)$, which includes the possibility of a pure state, and consider an observable $\mathcal{O} \in \mathcal{A}$. The idea is that it is not the state $\rho(t)$ but the expectation value $\text{Tr}[\mathcal{O}\rho(t)]$ that equilibrates in the sense that the long time average of it is very close to $\text{Tr}[\mathcal{O}\rho_0]$ where ρ_0 is the time independent component of $\rho(t)$. The latter is just the diagonal component of $\rho(t)$ in the energy eigenstate basis:²²

$$\rho_0 = \sum_n p_n |n\rangle \langle n| \quad , \quad p_n = \langle n | \rho(t) | n \rangle \quad (8.1)$$

where $|n\rangle$ are the energy eigenstates and the p_n are time independent probabilities $\sum_n p_n = 1$. Since many energy eigenstates are expected to be occupied each with a small probability, roughly $p_n \lesssim 10^{-N}$, where N is the number of degrees of freedom of the system, e.g., $N = O(10^{23})$.

²¹There is now a large literature but a limited set of references are [6-18].

²²For simplicity, we assume that there are no degeneracies in the E_n and also in the gaps $E_n - E_m$ which is reasonable in a realistic interacting non-integrable quantum system.

There are various bounds that can be established. For example, one can prove [10, 12] that the long time average of the variance of the expectation value $\text{Tr}[\mathcal{O}\rho(t)]$ about the equilibrium $\text{Tr}[\mathcal{O}\rho_0]$ is bounded above by a quantity of order²³

$$\Delta(\mathcal{O})^2 \sum_n p_n^2 \quad (8.2)$$

where $\Delta(\mathcal{O})$ is the range of eigenvalues of \mathcal{O} . Since the probabilities p_n are very small, (8.2) is certainly much smaller than the square of the resolution scale $\delta_{\mathcal{O}}^2$. Hence the time average of the expectation value of \mathcal{O} is captured to high degree of the accuracy by the equilibrium state ρ_0 .

Another interesting quantity is the fraction of time that the system spends away from the equilibrium state as measured by the inequality [13]

$$|\text{Tr}[\mathcal{O}\rho(t)] - \text{Tr}[\mathcal{O}\rho_0]| > \delta_{\mathcal{O}} \quad (8.3)$$

where $\delta_{\mathcal{O}}$ is the resolution scale defined in section 3.1. It can be shown that this fraction of time is bounded above by a quantity of order

$$\left(\frac{\Delta(\mathcal{O})}{\delta_{\mathcal{O}}}\right)^2 \max_n p_n \quad (8.4)$$

Although the ratio $\Delta(\mathcal{O})/\delta_{\mathcal{O}}$ can be large, it is expected to be dominated by the smallness of the eigenvalues p_n and so the implication is that the system spends the vast majority of time close to the equilibrium state with only occasional excursions away. It is important to emphasize, though, that the equilibrium state is not strictly time independent due to these fluctuations.

The issue of thermalization rests on the extent to what expectation values calculated from the equilibrium state ρ_0 are close to the microcanonical ensemble

$$\rho_{mc} = \frac{1}{N} \sum_{E_n \in [E, E + \Delta E]} |n\rangle \langle n| \quad (8.5)$$

This is a more refined question that requires additional hypotheses. For instance, it follows if the initial state is pure and has a sharply defined energy in the band $[E, E + \Delta E]$ from von Neumann's quantum ergodic theorem [4]. More generally, it follows from the *eigenstate thermalization hypothesis*

²³We refer to the original work for the precise statements.

discussed, for example, in [12-16]. There are similar statements that can be made for a sub-system A interacting with a larger system E . In this case, it is the expectation values of operators on the subsystem that equilibrate and can give values that are close to the canonical ensemble [6-11].

What is striking about this quantum approach to SM is that it completely cuts classical SM and any mention of microstates or probability out of the picture. So whilst the approach is very successful, there does not seem to be a way to see how classical SM can arise in the classical limit.

However, there are some parallels between our approach to the Copenhagen interpretation and the quantum approach to SM. Both place the expectation values of a set of realistic coarse grained observables \mathcal{A} center stage in the sense that the macroscopic state of the system is determined by both the underlying quantum state *and* the observables. In particular, in the SM context the underlying quantum state by itself is not the relevant object to measure equilibration and thermalization—it certainly does not have these properties—rather it is the expectation values of the observables $\text{Tr}[\mathcal{O}_n \rho(t)]$, $\mathcal{O}_n \in \mathcal{A}$.

In our proposal, the expectation values of the observables $\text{Tr}[\mathcal{O}_n \rho(t)]$ as a function of time should be replaced by the expectation values in the sequence of quantum microstates determined by the stochastic process $\langle \Psi_{i(t)} | \mathcal{O}_n | \Psi_{i(t)} \rangle$. The key question for us is:

To what extent is the long time average of $\langle \Psi_{i(t)} | \mathcal{O}_n | \Psi_{i(t)} \rangle$ captured by the microcanonical equilibrium expectation value $\text{Tr}[\mathcal{O}_n \rho_{mc}]$?

In order to answer this, let us make us some assumptions:

1. We will take the overall state to be the pure state $\rho = |\Psi\rangle\langle\Psi|$ with sharply defined energy in a microcanonical shell $[E, E + dE]$ defined by a subspace of the Hilbert space $\mathcal{H}_{mc} \subset \mathcal{H}$ with a large dimension N with ΔE macroscopically small but microscopically large so that N is large.
2. We will assume that the observables \mathcal{O}_n are explicitly time independent and commuting, so that the quantum microstates are the simultaneous eigenstates. In addition, none of the \mathcal{O}_n commutes with the Hamiltonian (except the coarse grained energy: see below) . If the contrary

were true, there would be super-selection sectors which would break ergodicity and in that case one would have to restrict to a particular eigenspace.

3. Finally we assume that the microcanonical energy shell defining a subspace of the Hilbert space \mathcal{H}_{mc} has a decomposition in terms of the eigenspaces of the observables

$$\mathcal{H}_{mc} = \bigoplus_i \mathcal{H}_i \ , \ d_i = \dim \mathcal{H}_i \ , \ \sum_i d_i = N \quad (8.6)$$

This seems reasonable because it is natural that \mathcal{A} contains a coarse grained version of the energy consisting of the projector onto \mathcal{H}_{mc} .²⁴

When the system, with respect to the set \mathcal{A} , equilibrates in the sense described above, the occupation probabilities of the quantum microstates $|c_i(t)|^2$ also equilibrate. In fact, von Neumann’s quantum ergodic theorem [4] (see also [17]) states that, for any $|\Psi(t)\rangle \in \mathcal{H}_{mc}$ and most times in the long run,

$$|c_i(t)|^2 = |\langle \Psi_i | \Psi(t) \rangle|^2 = \langle \Psi(t) | \Pi_i | \Psi(t) \rangle \approx \frac{d_i}{N} \quad (8.7)$$

where Π_i is the projector onto \mathcal{H}_i . The exact conditions for the equality are to be found in [4, 17] but they include the requirement that the dimensions d_i are suitably large. Note that (8.7) implies that the expectation values take their microcanonical form:

$$\langle \Psi | \mathcal{O}_n | \Psi \rangle = \sum_i |c_i|^2 \langle \Psi_i | \mathcal{O}_n | \Psi_i \rangle \approx \text{Tr}[\mathcal{O}_n \rho_{mc}] \quad (8.8)$$

where the first equality is the microscopic Born rule (3.4). The result also follows from the eigenstate thermalization hypothesis [16]. To emphasize, being in equilibrium does not mean that the $|c_i(t)|^2$ are strictly constant: the equilibrium state allows for fluctuations and very occasional larger excursions away from the average.

So in order to answer the question posed above, we need to establish that the stochastic process in equilibrium is also ergodic so that the long time

²⁴Von Neumann [4] calls the family of spaces $\{\mathcal{H}_i\}$ the “macro-observer”. In our approach we associate the “macro-observer” with the macroscopic time average of the expectation values $\langle \Psi_{i(t)} | \mathcal{O}_n | \Psi_{i(t)} \rangle$

average of $\langle \Psi_{i(t)} | \mathcal{O}_n | \Psi_{i(t)} \rangle$, as in (4.5), is equal to the equilibrium values $\text{Tr}[\mathcal{O}_n \rho_{mc}]$. We discussed the issue of ergodicity in section 4.1. In equilibrium, the stochastic process is effectively a homogeneous Markov chain for which ergodicity requires that over a long enough time the integrated transition probabilities $p_{ij}(t+T, t)$ becomes independent of j .

The transition rates of the stochastic system are given in (4.4). In order to address the question of ergodicity, we need to investigate the matrix elements $\langle \Psi_j | H | \Psi_i \rangle$, with $i \neq j$. In an interacting and non-integrable, i.e. chaotic, quantum system, it is reasonable to expect that the energy basis and the basis of quantum microstates are related by a random unitary matrix. With this hypothesis, one finds approximately that for each pair i, j , the non-vanishing transition rate T_{ij} , say, is randomly distributed over a range $[0, d_i \Delta E / (\hbar \sqrt{2N})]$ while $T_{ji} = 0$.

A homogeneous Markov chain with these random transition rates is demonstrably ergodic in numerical simulations. A suitable measure for the time scale for the system to be ergodic is how quickly the system becomes independent its initial state; one finds for large T

$$|p_{ij}(t+T, t) - p_{ik}(t+T, t)| \sim \exp[-\mu T] \quad (8.9)$$

where μ is of order $\sqrt{N} \Delta E / \hbar$. This shows that the process is ergodic over very short time scales of order $\hbar / (\sqrt{N} \Delta E)$. The implication is that the long time average as in (4.5) over macroscopic time scales T , of the expectation values that follow from the stochastic process reproduce those of the microcanonical ensemble:

$$\frac{1}{T} \int_0^T dt \langle \Psi_{i(t)} | \mathcal{O}_n | \Psi_{i(t)} \rangle \underset{\substack{\text{ergodicity (4.5)} \\ \text{microscopic Born rule (3.4)}}}{\approx} \sum_i |c_i|^2 \langle \Psi_i | \mathcal{O}_n | \Psi_i \rangle \underset{\substack{\text{quantum ergodic theorem (8.8)}}}{=} \langle \Psi | \mathcal{O}_n | \Psi \rangle \underset{\text{quantum ergodic theorem (8.8)}}{\approx} \text{Tr}[\mathcal{O}_n \rho_{mc}]. \quad (8.10)$$

The stochastic dynamics of the expectation values $\langle \Psi_{i(t)} | \mathcal{O}_n | \Psi_{i(t)} \rangle$ re-instates a dynamical, i.e., Boltzmann, picture of a thermal ensemble but now in the context of quantum SM. Now we can interpret the density operator as being an ensemble of quantum microstates as in (3.4) with a clear understanding of

the role of probability as a fundamental property of the Markov process. So the conceptual picture is clear and the quantum microstates provide a quantum precursor of the microstates of classical SM. There remains the program of investigating the stochastic dynamics in detail for particular examples, like a gas of microscopic particles, to find to what extent the stochastic dynamical picture is related to the classical mechanical picture. We have already shown how the latter can arise from the former for a macroscopic particle in section 5 but one would need to apply the same ideas to the whole gas.

9. Discussion

We have described an approach to quantum mechanics which offers the possibility to understand the different deterministic and stochastic elements of the classical limit of quantum mechanics in a more unified way. The key idea is that the underlying quantum system evolves according to the standard rules of quantum mechanics, i.e. the Schrödinger equation, while the stochastic behavior enters at the level of the set of coarse grained observables that are needed to construct an effective theory that describes the interactions of the system with other macroscopic systems. The central element of this effective theory is a stochastic process that preserves the microscopic Born rule. All the different facets of the classical limit follow from properties of the stochastic process and in particular to what extent the process is ergodic. When ergodicity is broken, the behavior can be effectively deterministic and the classical limit gives rise to classical mechanics. When the process is ergodic, the behavior is characteristic of statistical mechanics and we have shown that the quantum system gives a dynamical picture of a thermal ensemble. Finally, when ergodicity is partly broken with large ergodic subsets of states, the formalism provides a natural solution to the measurement problem.

To finish, we can give an account of the fate of Schrödinger's famous cat. The evolution of the quantum state leads to a final state vector that involves a linear combination of an alive and a dead cat:

$$|\Psi(0)\rangle = |cat\ alive\rangle \rightarrow |\Psi(T)\rangle = c_+ |cat\ alive\rangle + c_- |cat\ dead\rangle \quad (9.1)$$

In this simple analysis, we need not explicitly describe the other parts of the system, the measuring device, microscopic quantum system and environment: they are assumed to be included in the states $|cat\ alive\rangle$ and $|cat\ dead\rangle$.

The effective macroscopic description of the cat involves a set of sub-macroscopically coarse grained observables \mathcal{A} . The set of quantum microstates $\{|\Psi_i\rangle\}$ associated with \mathcal{A} splits into two subsets $\{|\Psi_{i_+}\rangle\}$ and $\{|\Psi_{i_-}\rangle\}$,

$$|cat\ alive\rangle = \sum_{i_+} c_{i_+} |\Psi_{i_+}\rangle \quad , \quad |cat\ dead\rangle = \sum_{i_-} c_{i_-} |\Psi_{i_-}\rangle \quad (9.2)$$

corresponding to an alive and a dead cat, respectively. The macroscopic state of the system is then described as the macroscopic time average of the expectation values $\langle \Psi_{i(t)} | \mathcal{O}_n | \Psi_{i(t)} \rangle$. The important point is that for realistic Hamiltonians, the integrated probability over a macroscopic time scale for making a transition for a dead to alive, or vice versa, over a macroscopic time scale $t_2 - t_1$

$$p_{i_{\pm}|j_{\mp}}(t_2, t_1) \ll 1 \quad (9.3)$$

So there is no practical prospect of an alive state $|\Psi_{j_+}\rangle$ making a transition to dead state $|\Psi_{i_-}\rangle$, and vice versa. So the states of an alive cat and dead cat are ergodically disjoint. Once the quantum microstate of the cat lies in a given ergodic component, which it does with probability $|c_+|^2$ or $|c_-|^2$, it remains there, either dead or alive, with definite positions for its moving part but also with its more microscopic degrees of freedom in approximate thermal equilibrium. It makes sense to then remove the ergodically disjoint parts of the state vector that can never be realized, a harmless procedure that is nothing other than the reduction of state vector:

$$|\Psi(T)\rangle \rightarrow |cat\ alive\rangle \quad \text{or} \quad |cat\ dead\rangle \quad (9.4)$$

Appendices

A. Mixed State Generalization

In this appendix, we show how to generalize the formalism to the case when the state is mixed described by a density matrix ρ . We will also focus on the realistic case with a commuting set of observables \mathcal{A} .

As in the case of a pure underlying state, the quantum microstates are just the simultaneous eigenstates of the set \mathcal{A} . The underlying mixed quantum state can always be expressed as

$$\rho = \sum_{ij} A_{ij} \rho_{ij} \quad , \quad \sum_i A_{ii} = 1 \quad (A.1)$$

where ρ_{ij} is an operator which maps $\mathcal{H}_j \rightarrow \mathcal{H}_i$ with $\text{Tr}\rho_{ij} = \delta_{ij}$. The microscopic Born rule in this case is expressed as

$$\text{Tr}(\mathcal{O}_n \rho) = \sum_i A_{ii} \text{Tr}(\rho_{ii} \mathcal{O}_n) \quad \forall \mathcal{O}_n \in \mathcal{A} \quad (\text{A.2})$$

The role of the quantum microstate is now played by the diagonal elements ρ_{ii} and A_{ii} is the generalization of the probability $|c_i|^2$.

We now define the stochastic process by the same logic as in (2.5). Using von Neumann's equation of motion for the density operator, one finds

$$\frac{dA_{ii}}{dt} = \frac{2}{\hbar} \sum_{j \neq i} \text{Im} \left[A_{ij} \text{Tr}_i \left([H, \rho_{ij}] - i\hbar \frac{\partial \rho_{ij}}{\partial t} \right) \right] \quad (\text{A.3})$$

where the trace is over \mathcal{H}_i . This yields the transition rates

$$T_{ij} = \max \left(\frac{2}{\hbar} \text{Im} \left[\frac{A_{ij}}{A_{jj}} \text{Tr}_i \left\{ [H, \rho_{ij}] - i\hbar \frac{\partial \rho_{ij}}{\partial t} \right\} \right], 0 \right) \quad (\text{A.4})$$

In the case of a pure state $A_{ij} = c_i c_j^*$ and $\rho_{ij} = |\Psi_i\rangle \langle \Psi_j|$ we recover the expression (4.1) using, for $i \neq j$,

$$\text{Tr}_i \left\{ [H, \rho_{ij}] - i\hbar \frac{\partial \rho_{ij}}{\partial t} \right\} = - \langle \Psi_j | \left(H - i\hbar \frac{\partial}{\partial t} \right) | \Psi_i \rangle \quad (\text{A.5})$$

If our quantum system is in a mixed state *rho* and is interacting with a larger quantum system then the time evolution of ρ will be non-trivial. It would be interesting to generalize our formalism to this case, but we will leave such an endeavor to future work.

B. Non-Abelian Sets of Observables

In this appendix, we consider several examples where the set \mathcal{A} contains operators that do not commute.

Example 1: Suppose that \mathcal{A} consists of the complete set of observables on the finite dimensional Hilbert space \mathcal{H} . If the dimension of \mathcal{H} is n then the complete set of observables has dimension n^2 . It quickly becomes apparent in this case that the set of quantum microstates states \mathcal{M} consists of just $|\Psi\rangle$ itself. In an intuitive sense, if we know the expectation values of all the

observables on \mathcal{H} then we have complete knowledge of the quantum state and so the coarse grained density operator must be equal to $\rho = |\Psi\rangle\langle\Psi|$. The stochastic process will be trivial and the evolution of the system with respect to \mathcal{A} is just the same as the deterministic microscopic dynamics encoded in the Schrödinger equation.

So an omnipotent observer \mathcal{A} sees the system evolve according to the Schrödinger equation.

Example 2: Suppose that \mathcal{A} consists of the complete set of observables which as operators act on a factor \mathcal{H}_A of a tensor product of finite dimensional Hilbert spaces $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_E$. These operators act as $\mathcal{O}_n \otimes 1_E$. The dimension of \mathcal{A} is equal to $(\dim \mathcal{H}_A)^2$.

In this example, the quantum microstates are just the conventional Schmidt states

$$|\Psi\rangle = \sum_{i=1}^N c_i |\Psi_i\rangle \quad , \quad |\Psi_i\rangle = |\psi_i\rangle \otimes |\tilde{\psi}_i\rangle \quad (\text{B.1})$$

where $\langle\psi_i|\psi_j\rangle = \langle\tilde{\psi}_i|\tilde{\psi}_j\rangle = \delta_{ij}$. The states $|\Psi_i\rangle$ and $|\tilde{\Psi}_i\rangle$ are eigenstates of the reduced density matrices $\rho^{(A)} = \text{Tr}_E |\Psi\rangle\langle\Psi|$ and $\rho^{(E)} = \text{Tr}_A |\Psi\rangle\langle\Psi|$, respectively. Their number equals $N = \min(\dim \mathcal{H}_A, \dim \mathcal{H}_E)$.

The microscopic Born rule (3.4) follows from the orthogonality of $|\tilde{\Psi}_i\rangle$:

$$\langle\psi_i| \otimes \langle\tilde{\psi}_i| (\mathcal{O}_n \otimes 1_E) |\psi_j\rangle \otimes |\tilde{\psi}_j\rangle = \langle\psi_i| \mathcal{O}_n |\psi_j\rangle \langle\tilde{\psi}_i|\tilde{\psi}_j\rangle = \delta_{ij} \langle\psi_i| \mathcal{O}_n |\psi_i\rangle \quad (\text{B.2})$$

which implies (3.4). The non-degeneracy condition rules out having more than one quantum microstate associated to each $|\tilde{\psi}_i\rangle$ since having two potential microstates $|\tilde{\psi}_i\rangle \otimes |\tilde{\psi}'_i\rangle$ and $|\tilde{\psi}_i\rangle \otimes |\tilde{\psi}''_i\rangle$ violates the non-degeneracy condition (3.5).

This example forms the basis of the reduced density operator formalism developed in [19-21].

Example 3: Consider the case of a small finite dimensional set of non-commuting observables \mathcal{A} in a large Hilbert space in a generic state $|\Psi\rangle$. We will assume that \mathcal{A} forms a Lie algebra, or can be completed into such an algebra. To be more precise, we will consider the case that the observables generate a $U(N)$ algebra. This example is not supposed to be realistic but

it just illustrates how the formalism can cope with a non-abelian set of operators.

The complexified generators can be labelled by a pair $i, j \in \{1, 2, \dots, N\}$ with²⁵

$$[\mathcal{O}_{ij}, \mathcal{O}_{kl}] = \delta_{jk}\mathcal{O}_{il} - \delta_{il}\mathcal{O}_{jk} \quad (\text{B.3})$$

We are assuming that the dimension of the algebra N^2 is much smaller than the dimension of the Hilbert space \mathcal{H} . This means that a typical state $|\Psi\rangle$ is not annihilated by any element of \mathcal{A} :

$$\mathcal{O}|\Psi\rangle \neq 0, \quad \forall \mathcal{O} \in \mathcal{A} \quad (\text{B.4})$$

In such a situation, the N^2 dimensional space $\mathcal{H}_{\mathcal{A},\Psi} \subset \mathcal{H}$ spanned by the independent states

$$\mathcal{O}_{ij}|\Psi\rangle, \quad \forall i, j \quad (\text{B.5})$$

carry an adjoint representation of the Lie algebra. A basis for this space consists of the orthonormal states $|ij\rangle$ labelled by a pair i, j with

$$\mathcal{O}_{ij}|kl\rangle = \delta_{jk}|il\rangle \quad (\text{B.6})$$

The generators and states can always be redefined by a unitary $U(N)$ transformation:

$$\mathcal{O}_{ij} \rightarrow U\mathcal{O}_{ij}U^{-1}, \quad |ij\rangle \rightarrow U|ij\rangle \quad (\text{B.7})$$

Given the typical state $|\Psi\rangle$, we can always use this freedom to perform a $U(N)$ transformation to align the algebra with the state $|\Psi\rangle$ in the sense that

$$|\Psi\rangle = \sum_{i=1}^N c_i |ii\rangle + c_{N+1} |\Psi^\perp\rangle \quad (\text{B.8})$$

In this expression, $|\Psi^\perp\rangle$ is the component of $|\Psi\rangle$ lying orthogonal to $\mathcal{H}_{\mathcal{A},\Psi}$.

The expression above identifies the $N + 1$ quantum microstates as

$$|\Psi_i\rangle = |ii\rangle, \quad i = 1, 2, \dots, N, \quad |\Psi_{N+1}\rangle = |\Psi^\perp\rangle \quad (\text{B.9})$$

²⁵The Hermitian combinations are $\mathcal{O}_{ii}, \mathcal{O}_{ij} + \mathcal{O}_{ji}$ and $i(\mathcal{O}_{ij} - \mathcal{O}_{ji})$. The complexified operators satisfy an algebra under multiplication as well as commutation.

The quantum microstates are precisely the eigenstates of the Cartan subalgebra generated by the mutually-commuting operators \mathcal{O}_{ii} , $i = 1, 2, \dots, N$ and, in particular, $|\Psi_{N+1}\rangle$ has zero eigenvalues for all the Cartan generators:

$$\mathcal{O}_{ii}|\Psi_j\rangle = \delta_{ij}|\Psi_j\rangle \quad , \quad j = 1, 2, \dots, N \quad , \quad \mathcal{O}_{ii}|\Psi_{N+1}\rangle = 0 \quad (\text{B.10})$$

Note that the Cartan generators are the only operators with non-vanishing expectation values

$$\langle\Psi|\mathcal{O}_{ij}|\Psi\rangle = |c_i|^2\delta_{ij} \quad (\text{B.11})$$

C. Time Dependent Observables

In this appendix, we discuss how to general the formalism to the case of explicitly time dependent observables. The transition rates in (4.4) as generalized to

$$T_{ij} = \max\left(-\frac{2}{\hbar}\text{Im}\left[\frac{c_i}{c_j}\langle\Psi_j|\left(H - i\hbar\frac{\partial}{\partial t}\right)|\Psi_i\rangle\right], 0\right) \quad (\text{C.1})$$

The discussion of locality in section 4.2 is modified in the following way, without any of the conclusions being changed. First of all, (4.7) generalizes to

$$\langle\Psi_i|1_A \otimes \left(H_E - i\hbar\frac{\partial}{\partial t}\right)|\Psi_j\rangle = 0 \quad , \quad i \neq j \quad (\text{C.2})$$

where the derivative term here acts only on $|\psi_{ja}\rangle$. The final local expression for the transition rates is

$$T_{ij} = \max\left(-\frac{2}{\hbar}\text{Im}\left[\frac{c_i}{c_j}\langle\Psi_j|\left\{H_A - i\hbar\frac{\partial}{\partial t}\right\}|\Psi_i\rangle \otimes 1_E + H_{int}\right], 0\right) \quad (\text{C.3})$$

which generalizes (4.8).

References

- [1] D. Wallace, Inferential vs. Dynamical Conceptions of Physics, arXiv:1306.4907 [quant-ph].
- [2] J. Uffink, Compendium of the foundations of classical statistical physics, <http://philsci-archive.pitt.edu/2691/>. 2006. Chapter for Handbook for Philosophy of Physics, J. Butterfield and J. Earman (eds).

- [3] E. Schrödinger, Energieaustausch nach der Wellenmechanik, *Annalen der Physik* 388(15) (1927) 956. English translation by J. F. Shearer, *The Exchange of Energy according to Wave Mechanics*, pp. 137146 in E. Schrödinger: *Collected Papers on Wave Mechanics*. Providence, R.I.: AMS Chelsea (1928 and 1982)
- [4] J.von Neumann Beweis des Ergodensatzes und des H-Theorems in der neuen Mechanik, *Zeitschrift für Physik* 57 (1929) 30, English translation by R. Tumulka in J. von Neumann: *Proof of the ergodic theorem and the H-theorem in quantum mechanics*, *European Phys. J. H* 35 (2010) 201 [arXiv:1003.2133].
- [5] J.Von Neumann, *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, 1955.
- [6] S. Popescu, A. J. Short, and A. Winter, The foundations of statistical mechanics from entanglement: Individual states vs. averages, arXiv quant-ph/0511225.
- [7] S.Popescu,A.J.Short,andA.Winter, Entanglement and the foundations of statistical mechanics, *Nature Physics* 2 (2006) 754.
- [8] S.Goldstein, J.L. Lebowitz, R. Tumulka, and N. Zanghì, Canonical Typicality, *Phys. Rev. Lett.* 96 (2006) 050403.
- [9] N. Linden, S. Popescu, A. J. Short, and A. Winter, Quantum mechanical evolution towards thermal equilibrium, *Phys. Rev. E* 79 (2009) 061103 [arXiv:0812.2385 [quant-ph]].
- [10] A.J.Short, Equilibration of quantum systems and subsystems, *New Journal of Physics*13 (2011) 053009.
- [11] J. Gemmer, M. Michel, and G. Mahler, *Quantum Thermodynamics: Emergence of Thermodynamic Behavior within Composite Quantum Systems*, Berlin: Springer, 2004.
- [12] P. Reimann, Foundation of Statistical Mechanics under Experimentally Realistic Conditions, *Phys. Rev. Lett.* 101 (2008) 190403 [arXiv:0810.3092 [cond-mat.stat-mech]].

- [13] P. Reimann and M. Evstigneev, On the Foundation of Statistical Mechanics under Experimentally Realistic Conditions: A Comparison between the Quantum and the Classical Case, *Phys. Rev. E* 88 (2013) 052114 [arXiv:1311.2732 [cond-mat.stat-mech]].
- [14] M. Srednicki, Thermal fluctuations in quantized chaotic systems, *J. Phys. A* 29 (1996) L75 [arXiv:chao-dyn/9511001].
- [15] M. Srednicki, The approach to thermal equilibrium in quantized chaotic systems, *J. Phys. A* 32 (1999) 1163 [arXiv:cond-mat/9511001 [cond-mat.stat-mech]].
- [16] M. Rigol and M. Srednicki, Alternatives to Eigenstate Thermalization, *Phys. Rev. Lett.* 108 (2012) 110601 [arXiv:1108.0928 [cond-mat.stat-mech]]
- [17] S. Goldstein, J.L. Lebowitz, C. Mastrodonato, R. Tumulka and N. Zanghi, NormalTypicality and von Neumanns Quantum Ergodic Theorem, *Proceedings of the Royal Society A* 466(2123) (2010) 3203 [arXiv:0907.0108 [quant-ph]].
- [18] C. Gogolin, Pure State Quantum Statistical Mechanics, [arXiv:1003.5058 [quant-ph]].
- [19] T.J. Hollowood, The Copenhagen Interpretation as an Emergent Phenomenon, *J. Phys. A* 46 (2013) 325302 [arXiv:1302.4228 [quant-ph]].
- [20] T.J. Hollowood, The Emergent Copenhagen Interpretation of Quantum Mechanics, *J. Phys. A* 47 (2014) 185301 [arXiv:1312.3427 [quant-ph]].
- [21] T. J. Hollowood, New Modal Quantum Mechanics, arXiv:1312.4751 [quant-ph].
- [22] K. G. Wilson, The renormalization group and critical phenomena, *Rev. Mod. Phys.* 55 (1983) 583.
- [23] E. .T. Jaynes, Information Theory and Statistical Mechanics II, *Phys. Rev.* 108 (1957) 171.

- [24] J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics: Collected papers on quantum philosophy*, Cambridge University Press, 2004.
- [25] G. Bacciagaluppi and W.M. Dickson, Dynamics for Modal Interpretations, *Found. Phys.* 29 (1999) 1165.
- [26] J.J. Halliwell, Commuting position and momentum operators, exact decoherence, and emergent classicality, *Phys. Rev. A* 72 (2005) 042109.
- [27] B-G. Englert, On Quantum Theory, *Eur. Phys. J. D* 67 (2013) 238 [arXiv:1308.5290 [quant-ph]].
- [28] J.C. Vink, Quantum potential interpretation of the wave function of the universe, *Nucl. Phys. B* 369 (1992) 707.
- [29] D. Bohm, *Quantum Theory*, Courier Dover Publications, 1951.
- [30] D. Bohm and Y. Aharonov, Discussion of Experimental Proof for the Paradox of Einstein, Rosen, and Podolsky, *Phys. Rev.* 108, 1070 (1957).
- [31] A. Einstein, B. Podolsky, and N. Rosen, Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Phys. Rev.* 47, 777 (1935).