OPEN ACCESS

University of BRISTOL

This electronic thesis or dissertation has been downloaded from Explore Bristol Research, http://research-information.bristol.ac.uk

Author:
Thomas Penney, Megan J

Title:
1YQ5MEG

Author:
Thomas Penney, Megan J
Title:
1YQ5MEG

Models, Interpretations, and Realism in Quantum Physics

By

MEGAN THOMAS PENNEY

Department of Philosophy
UNIVERSITY OF BRISTOL

A dissertation submitted to the University of Bristol in accordance with the requirements of the degree of DOCTOR OF PHILOSOPHY in the Faculty of Arts.

JULY 2020

Word count: 58823

ABSTRACT

From the beginnings of quantum theory in the early 20$^{\text{th}}$ Century, physicists and philosophers of physics have been struggling to interpret it. The ongoing debate about the measurement problem has lead to the creation of multiple different interpretations intended to solve it. Each interpretation presents quantum theory differently in an attempt to reconcile it with what we see in the world.

The measurement problem, and alongside it the more general problem of the quantum to classical transition, are explored in the following chapters in a journey from the first interpretation of quantum mechanics to modern Quantum Field Theory. Chapter 1 examines the basic quantum formalism and the first quantum interpretations. In Chapter 2, a chronological review of influential quantum mechanical thought experiment highlights the central issue of measurement. Chapter 3 introduces the measurement problem and divides it in two; the unique outcomes problem and the preferred basis problem, and examines how each is addressed in different interpretations. Chapter 4 investigates the theory of decoherence. Chapter 5 re-examines the interpretations from Chapter 3 in the light of decoherence as a model for the emergence of classical macrostates from quantum microstates. Chapter 6 evaluates a possible analogy between using the idealisation of an infinite limit in the quantum to classical transition and in phase transitions in statistical mechanics. It argues that the analogy fails due to strange phenomena in the quantum to classical transition which are currently explained by semiclassical mechanics. Chapter 7 considers the issue of realism in interpretations of theories beyond ordinary quantum mechanics, such as Quantum Field Theory.

This examination of quantum theory shows why interpreting it is so hard as to be ongoing after a hundred years of study, considers how we should choose between interpretations – and ultimately asks whether we should have to choose at all.

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's Regulations and Code of Practice for Research Degree Programmes and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: MEGAN THOMAS PENNEY DATE: 06/07/2020

TABLE OF CONTENTS

Page

QUANTUM MECHANICS AND THE COPENHAGEN INTERPRETATION

1.1 Introduction

This thesis explores the measurement problem in quantum mechanics from a new perspective, and shows why interpreting quantum theory is so hard as to be ongoing after a hundred years of study. It considers how we should choose between the many different interpretations that have been presented since the beginning of quantum theory, and ultimately asks whether we should have to choose between them at all. From quantum mechanical thought experiments to the theory of decoherence, and all the way to Quantum Field Theory (QFT), the issue will be approached from lesser-trod paths to gain a more nuanced understanding of the difficulties of interpreting quantum theory and the quantum to classical transition.

How do we get from the strange behaviour of quantum physics to the classical behaviour we see around us? How should we relate crucial parts of quantum theory — the wave function, superposition, entanglement — to what is actually in the world? These are the questions that physicists and philosophers of physics have been trying to answer since work on quantum theory began. Around one hundred years later, and despite huge advances in the field leading to QFT (one of our most if not the most successful empirical theories), it seems that the answer to these questions is still unclear. Though we have many interpretations to choose from about what

quantum theory says exists in the world, there is still no real consensus. We can still 'choose'.

So why is this issue ongoing after all this time? And how should we 'choose' between interpretations? Should we have to choose only one interpretation? These are the questions this thesis helps answer by providing a different perspective on the core issue of interpreting quantum theory from a variety of contexts. Reviewing important and influential thought experiments chronologically through the development of quantum theory gives a more nuanced view of the physics that is not brought out when focusing on the trilemma approach to the measurement problem. Measurement is an instance of the quantum to classical transition, and it is this transition that should be focused on when looking at the many different interpretations of quantum mechanics. This thesis explores the process of decoherence and puts decoherence theory forward as a way to pressure-test interpretations. Further examination of the transition between the quantum and classical worlds highlights the necessity of keeping classical theory in order to explain phenomena at the classical limit, and therefore the importance of a diversity of practices when interpreting quantum theory. Finally, I consider a new type of realism, Effective/Renormalisation Group (E/RG) realism, and argue that to distinguish itself from empirical structuralism it can be combined naturally with ontic structural realism. From this new kind of realism we learn that it is not necessary that we should have to choose a quantum ontology, and keeping multiple models, theories and interpretations can actually increase our explanatory power.

This thesis uses a variety of methods in order to explore the measurement problem and the more general problem of the quantum to classical transition. The history of quantum theory is important, as there is much to learn from how it was considered at its nascence and how it was first interpreted. Staying close to the actual physics is important when looking at the theory philosophically. The overlap between general philosophy of science and philosophy of physics also informs this thesis, as each has important things to teach us about the other. Chapter 1 looks at the early history of quantum theory and its interpretations to describe the issue we still have today as it was seen from the beginning. Chapter 2 examines two different conclusions from looking at quantum mechanical thought experiments: what they teach us about thought experiments in general philosophy of science, and what they teach us about quantum mechanics. Chapter 4 looks at the physics more closely when looking at the transition between quantum to

classical. Chapter 5 then applies this to the previous discussion of interpretations in chapter 3. Chapter 6 uses general philosophy of science to examine how we should think about the quantum to classical limit in the physics. Chapter 7 discusses the realism debate in advanced quantum theory such as Quantum Field Theory. This blending of different topics from philosophy of physics and more general philosophy of science reflects the blending of quantum and classical theory in the quantum to classical transition, and the blending of different inequivalent representations in Effective Field Theories. Keeping a broad perspective opens new lines of inquiry and prevents stagnation, as seen clearly in the debate in chapter 7.

The final argument of this thesis is that we should not have to limit ourselves to one interpretation, that we should not have to choose. This idea is mirrored in the many different methods used throughout the thesis, combining the history and foundations of physics, philosophy of physics, and general philosophy of science. By keeping a broad view of the different fields and looking at the intersections between them can reveal more than restricting to one or two. The same idea can be applied when looking at the quantum to classical transition.

In this thesis, I put forward several new ideas. In chapter 2, I give a new argument for the view that thought experiments must be more than arguments – they play important roles as placeholders for future experiments that are currently impossible. Following Stuart (2015, 2016), I argue that thought experiments are on a spectrum with real world experiments and allow us to perform currently or permanently impossible experiments in the imagination. I also argue that they clearly reveal the problem of measurement in quantum mechanics, and the more general problem of the quantum to classical transition. In chapters 3-5, I argue that decoherence is a good pressure-test for interpretations of quantum theory, and that this affects ψ-ontic and ψ-epistemic interpretations differently. ψ-ontic interpretations must deal with the physical process of decoherence. After a detailed exploration of the ψ-epistemic interpretation of QBism I argue that it must maintain a completely subjectivist view of quantum theory or else the incorporation of decoherence theory into the QBist framework cannot work. Chapter 6 considers the 'transition' from quantum uncertainty to classical ignorance as analogous to a phase transition in thermodynamics and statistical mechanics from an Information Theoretic perspective.[1] I argue

---

[1]This analogy was first suggested by an early copy of Bub (2016), though has since been taken out of subsequent

that this is a formal argument by analogy, but that it fails to take semiclassical mechanics into account. In Chapter 7 I argue that the new realism based on Renormalisation Group methods, Effective/RG realism, can escape empiricism by aligning with ontic structural realism.

These conclusions are reached by exploring a wide range of interpretations of quantum theory, both ψ-ontic and ψ-epistemic, and investigating the theory of decoherence, semiclassical mechanics, and the elusive quantum to classical limit, using techniques from general philosophy of science and historical accounts of the first interpretations. First it is important to lay down the foundation of basic quantum formalism that will be used throughout this thesis. This chapter goes over the formalism to define the terminology necessary for the discussion. Section 1.2 reviews the basic notation and definitions of key terms in quantum mechanics which will be used in the later chapters.

As the formalism developed, so did its interpretation. Once the basics have been covered, section 1.3 starts right at the beginning of quantum theory by discussing the first interpretation, the original Copenhagen interpretation of quantum mechanics, which appeared alongside the early work on quantum theory in the late 1920s. However, this is markedly different to what is often taken to be the orthodox Copenhagen interpretation today. Section 1.4 discusses the orthodox Copenhagen interpretation, which is agreed to have begun in the 1950s. This discussion shows that even from the start, physicists were torn over just how to interpret quantum theory, and this debate has persisted to today, only growing more complex as the years pass. Despite all these differing interpretations, there is one core issue which all of these interpretations are trying to solve; a single irritant around which the pearl of the debate formed over the years. Section 1.3 shows this to be the transition between quantum states and classical measurements.

1.2 Basic Quantum Formalism

Canonically formulated by von Neumann in 1932, the modern language of quantum mechanics takes Schrödinger's wave theory and combines it with the matrix mechanics from the work of Heisenberg and Born.[2]

---

editions.

[2] A new version of von Neumann's book, Mathematical Foundations of Quantum Mechanics, was published in 2018, and this is the edition used here (Von Neumann, 2018).

The basis of this formalism is the theory of linear algebra, which concerns linear vector spaces $\mathcal{V}$ built of vectors ($|\psi\rangle, |\phi\rangle$). A physical system in this formalism is described by a Hilbert space, $\mathcal{H}$, which is a complex vector space with an inner product and is complete and separable, and can have infinite dimension. A state of a system is represented by a ray in $\mathcal{H}$ (with norm 1 if normalised). States, as rays, have a phase, $e^{i\varphi}$, though the state $e^{i\varphi}|\psi\rangle$ is indistinguishable by any measurement from $|\psi\rangle$.

The simplest example of a system in quantum mechanics is a two-level system called a qubit. A quantum bit, or a qubit, can be described by the two states $|0\rangle$ and $|1\rangle$, which are orthonormal – they are unit vectors that have an inner product equal to zero. Because of the linearity of the Hilbert space, the qubit also has the state:

$$(1.1) \qquad\qquad a|0\rangle + b|1\rangle, \ \forall a, b \in \mathbb{C},$$

where 'a' and 'b' are such that the norm of the state is 1.[3] This is an instance of the superposition principle;

"If one possible state of an ensemble of identical systems is described by a wave function Ψ_1, and another state of this ensemble by Ψ_2, then any *linear combination*,

$$(1.2) \qquad\qquad \Psi = c_1\Psi_1 + c_2\Psi_2,$$

where c_1 and c_2 are constants, is also a wave function describing a possible state of the ensemble." (Bransden and Joachain, 1989, p. 56)

Every physical observable, e.g. O, corresponds to a Hermitian operator $\hat{O}$.[4] The order in which operators are applied is important, as for example with two operators $\hat{A}$ and $\hat{B}$, $\hat{A}\hat{B}$ is generally different to $\hat{B}\hat{A}$. The difference between these two is called the commutator:

---

[3]Therefore $|a|^2 + |b|^2 = 1$.

[4]An operator is Hermitian if and only if it is equal to its self-adjoint – equal to its complex conjugate transpose in the matrix representation.

$$(1.3) \qquad\qquad [\hat{A},\hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}.$$

When two operators commute, the commutator $[\hat{A},\hat{B}] = 0$.

Another important part of the quantum formalism, derived by Heisenberg in 1927, is the uncertainty relation.[5] Though originally the relation was about the impossibility of knowing exactly both position and momentum exactly,

$$(1.4) \qquad\qquad \Delta x \Delta p_x \geq \frac{\hbar}{2},$$

this was later generalised and can be applied to any two arbitrary non-commuting observables,

$$(1.5) \qquad\qquad (\Delta_\psi O) \odot (\Delta_\psi O') \geq \frac{1}{2}\left|\left\langle \psi \left|\left[\hat{O},\hat{O}'\right]\right|\psi\right\rangle\right|.$$

The uncertainty relation is therefore a direct result of non-commuting observables in quantum mechanics.[6]

Each operator has a set of vectors, or eigenvectors:

$$(1.6) \qquad\qquad \hat{O}\,|o_j\rangle = o_j\,|o_j\rangle,$$

and these eigenvectors form a complete basis. For any orthonormal basis, $\{|\phi_j\rangle\}_j$,

$$(1.7) \qquad\qquad \langle\phi_j|\hat{O}\,|\phi_k\rangle = \hat{O}_{jk},$$

or

$$(1.8) \qquad\qquad \hat{O} = \sum_{jk} O_{jk}\,|\phi_j\rangle\langle\phi_k|.$$

---

[5]This will be discussed in greater detail in section 2.5.1

[6]Where $\Delta A = \sqrt{\langle A^2\rangle - \langle A\rangle^2}$.

When $|\phi_j\rangle$ is equal to the eigenvector for $\hat{O}$, $|o_j\rangle$, then this becomes

$$(1.9) \qquad\qquad \hat{O} = \sum_j o_j |o_j\rangle \langle o_j|.$$

Using the example of the qubit above, the state vectors in matrix notation are

$$(1.10) \qquad\qquad |0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

The operator $\hat{A}$, where $\hat{A}|0\rangle = |1\rangle$ and $\hat{A}|1\rangle = |0\rangle$, can be written in matrix notation as

$$(1.11) \qquad\qquad \hat{A} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

and so has eigenvalues of ± 1. These are the possible measurement outcomes. The probability of finding an outcome o_j is given by the Born Rule:

$$(1.12) \qquad\qquad p(o_j) = |\langle o_j|\psi\rangle|^2,$$

where $|o_j\rangle$ is the eigenvector associated with the eigenvalue o_j, and $|\psi\rangle$ is the state of the system. The expectation value of the observable O is calculated by taking its trace:

$$(1.13) \qquad\qquad \langle\hat{O}\rangle = \langle\psi|\hat{O}|\psi\rangle = \langle\psi|\left(\sum_j o_j |o_j\rangle\langle o_j|\right)|\psi\rangle = \sum_j p(o_j)o_j.$$

For a continuous system, an integral must be used. The expectation value of $\hat{O}$ is

$$(1.14) \qquad\qquad \langle\hat{O}\rangle = \langle\psi|\hat{O}|\psi\rangle = \int do f_O(O)\langle\psi|o\rangle\langle o|\psi\rangle = \int do f_O(O)|\psi(o)|^2.$$

where the wave function is $\psi(o) = \langle o|\psi\rangle$ and $|\psi(o)|^2$ is the probability density. The probability of finding an eigenvalue of operator $\hat{O}$ in the interval o and $o + do$ given by state $|\psi\rangle$ is:[7]

---

[7] Kok (2015), p.19.

(1.15)
$$\langle \psi | (|o\rangle \langle o | do) | \psi \rangle \equiv dp(o),$$

and so

(1.16)
$$\frac{dp(o)}{do} = |\psi(o)|^2.$$

A commonly used example is that of spin, or spin angular momentum. A particle with spin-half, such as an electron, has its spin-half components described by the Pauli matrices,[8]

(1.17)
$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The spin operators representing the three spin components are $\hat{S}_x, \hat{S}_y$, and $\hat{S}_z$. These are expressed in matrix form as

(1.18)
$$\hat{S}_x = \frac{1}{2}\hbar\sigma_x,$$

and for y and z, similarly.

These operators have the following commutation relations:

(1.19)
$$[\hat{S}_i, \hat{S}_j] = i\hbar\epsilon_{ijk}\hat{S}_k,$$

and

(1.20)
$$\hat{S}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2.$$

where ϵ_{ijk} is the Levi-Civita symbol, also known as the permutation symbol, and is anti-symmetric.

---

[8]From Rae (2002).

It can be seen that, for example, the eigenvalues of $\hat{S}_z$ are $\frac{1}{2}\hbar$ and $-\frac{1}{2}\hbar$, and the corresponding eigenvectors are $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, respectively. (Rae, 2002, p. 113)

These eigenvectors and eigenvalues are for spin-half-odd integer particles called fermions. If we assume we have two fermions in a sealed box, they are indistinguishable following the Indistinguisability Postulate, as they are identical particles. We therefore give them dummy labels, $|\psi(\vec{r}_1), \phi(\vec{r}_2)\rangle$. The overall quantum state of these two particles is

$$(1.21) \qquad |\Psi(\vec{r}_1, \vec{r}_2)\rangle_{12} = \frac{|\psi(\vec{r}_1), \phi(\vec{r}_2)\rangle_{12} + e^{i\varphi}|\psi(\vec{r}_2), \phi(\vec{r}_1)\rangle_{12}}{\sqrt{2}}.$$

As these two particles are fermions with spin-$\frac{1}{2}$, they have to obey Pauli's exclusion principle and cannot be in the same state. Therefore we chose $e^{i\varphi}$ to be -1:

$$(1.22) \qquad |\Psi(\vec{r}_1, \vec{r}_2)\rangle_{12} = \frac{|\psi(\vec{r}_1), \phi(\vec{r}_2)\rangle_{12} - |\psi(\vec{r}_2), \phi(\vec{r}_1)\rangle_{12}}{\sqrt{2}}.$$

making the state anti-symmetric under swapping r_1 and r_2. These particles obey Fermi-Dirac statistics.

If the two particles were bosons, with zero or integer spin, then the state is symmetric, and the overall state is

$$(1.23) \qquad |\Psi(\vec{r}_1, \vec{r}_2)\rangle_{12} = \frac{|\psi(\vec{r}_1), \phi(\vec{r}_2)\rangle_{12} + |\psi(\vec{r}_2), \phi(\vec{r}_1)\rangle_{12}}{\sqrt{2}}.$$

These particles obey Bose-Einstein statistics (Bransden and Joachain, 1989, p. 450)

The one-dimensional time-dependent Schrödinger equation for free particles is the following:

$$(1.24) \qquad i\hbar \frac{d}{dt}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle,$$

where $H = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \hat{V}$, is a Hermitian operator with dimensions of energy, also known as the Hamiltonian. The Hamiltonian is the sum of the operators of kinetic and potential energy, $\hat{H} = \hat{T} + \hat{V}$. When the potential $\hat{V}$ is not a function of time, energy is conserved. In a closed system, the solution to the time-dependent Schrödinger equation is obtained by solving the appropriate

9

time-independent equation, and multiplying by the time-dependent phase factor (Rae, 2002, p. 19). In the case of a free particle with the origin of the potential energy chosen such that $V(x) = 0$, a solution to the time-independent Schrödinger equation is

(1.25)
$$\psi = A \exp(ikx),$$

where $k = (2mE/\hbar^2)^{1/2}$ and A is a constant. The wave function then has the form

(1.26)
$$\psi = A \exp[i(kx - \omega t)],$$

where $\omega = E/\hbar$ (Rae, 2002, p. 19). This can then be multiplied by the phase factor,

(1.27)
$$T = \exp(-iEt/\hbar).$$

The Schrödinger equation evolves via unitary time evolution, and the operator that transforms the state at t_0 to the state at time t,

(1.28)
$$|\psi(t)\rangle = U(t, t_0)|\psi(t_0)\rangle,$$

where $U(t, t_0) = exp(-\frac{i}{\hbar}At)$.

The unit operator $\hat{I}$, or the identity operator, is an operator that leaves a function unchanged (Bransden and Joachain, 1989, p. 197):

(1.29)
$$\hat{I}\Psi = \Psi.$$

A linear operator is unitary if

(1.30)
$$UU^\dagger = U^\dagger U = \mathbb{I}.$$

Unitary operators are a special kind of operator. They commute with their adjoint;

$$\hat{U}\hat{U}^\dagger = \hat{U}^\dagger\hat{U},$$

(1.31)

$$\hat{U}^{-1} = \hat{U}^\dagger.$$

These operators can be expressed in the form $U = \exp(iA)$, where $\hat{A}$ is a Hermitian operator.[9]

Each measurement outcome o_j corresponds to a projection operator, P_j, on the subspace of the eigenvectors $|o_j\rangle$, where a perfect measurement can be described as follows:

(1.32)
$$|\psi\rangle \longrightarrow \frac{P_j|\psi\rangle}{||P_j|\psi\rangle||}.$$

The expectation value of the operator $\hat{O}$ can be written as a trace,

(1.33)
$$\langle\hat{O}\rangle = Tr(|\psi\rangle\langle\psi|\hat{O}).$$

Rather than using the full operator, we can calculate the trace of

(1.34)
$$P_j = |o_j\rangle\langle o_j|,$$

which is

(1.35)
$$\langle P_j\rangle = Tr(|\psi\rangle\langle\psi|P_j) = Tr(|\psi\rangle\langle\psi||o_j\rangle\langle o_j|) = |\langle o_j|\psi\rangle|^2 = p(o_j).$$

This calculates the probability of a particular measurement outcome by taking the expectation value of the projection operator of its eigenstate. If we measure the observable O and get the eigenvalue o_j, then as soon as the measurement has taken place the system is in the corresponding eigenstate $|o_j\rangle$. This postulate, known as the 'projection postulate', due to the measurement apparently projecting the system to the eigenstate that corresponds to measured eigenvalue, can be described as a 'collapse' of the wave function. This apparent collapse is at the heart of the measurement problem in quantum mechanics.

---

[9]Bransden and Joachain (1989), p.198.

There are cases where we know the quantum state of a system exactly and must use a density operator, which can be described generally as

$$(1.36) \qquad \rho = \sum_k p_k |\psi_k\rangle \langle \psi_k|.$$

The probabilities p_k sum to one, and $|\psi_k\rangle$ are normalised states. ρ is called the density operator because it acts as a weight in the expectation value. To find the spectral decomposition of ρ, we can diagonalize it as follows;

$$(1.37) \qquad \rho = \sum_j \lambda_j |\lambda_j\rangle \langle \lambda_j|,$$

where $\{|\lambda_j\rangle\}$ forms a complete orthonormal basis. When one of the probabilities is equal to one, the others are zero and the state is no longer mixed, but is a pure state completely determined by $|\psi_j\rangle$. The density operator is used in the theory of decoherence.

If there are two systems, with two Hilbert spaces, $\mathcal{H}_1$ and $\mathcal{H}_2$, each with an orthonormal basis, $\{|\psi_j\rangle\}$ and $\{|\phi_k\rangle\}$ respectively. As $\mathcal{H}_{1+2} = \mathcal{H}_1 \otimes \mathcal{H}_2$, we get the tensor product $\{|\psi_j\rangle \otimes \{|\phi_k\rangle\}_{jk}$. So, a random pure state on $\mathcal{H}_{1+2}$ can be written as follows;

$$(1.38) \qquad |\psi\rangle = \sum_{jk} c_{jk} |\psi_j, \phi_k\rangle.$$

When the combined state is a product state, a partial trace can be taken to trace out the system we do not want to look at to leave a pure state. This is not always possible, however — if the two systems are entangled, then a partial trace will instead leave a mixed state. This means that mixtures cannot always be interpreted as ignorance of the pure state.

Entanglement implies that two or more particles have their probabilities for the outcome of measurements dependent on the state of each other, even though they are not interacting. An example of this is the singlet state:

$$(1.39) \qquad |\psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle).$$

Here we have two spin-half particles in a state where we know they have total spin 0, but we know nothing about them individually. On the measurement of one of the particles finding spin-up, any subsequent measurement of the second particle must give the measurement outcome spin-down, and vice versa (Rae, 2002, p. 256). These states are therefore correlated. A product state, however, is a separable state where there are no quantum or classical correlations between the states of the product.

This means that we do not have complete information about the system, and there is some information in the combined state that is not in the subsystems on their own. This can happen, for example, when a system interacts with the environment; taking the partial trace of the system will generally leave you with a mixed state (see chapter 4 where this is discussed in more detail).

This basic formalism does not have a full interpretation, and the difficulty of how to interpret quantum mechanics is highlighted by the projection postulate and the apparent 'collapse' of the wave function into an eigenstate of the observable after 'measurement'. Even at its nascence, quantum mechanics split the opinions of those who used it on just how to reconcile the formalism with the classical world we experience every day.

1.3 The Original Copenhagen Interpretation

The Copenhagen interpretation is generally seen as the first attempt to understand the quantum mechanical formalism. Ascribed mainly to Niels Bohr, it is often described simply in introductions to interpretations of quantum theory as taking the quantum state as a convenient fiction used to calculate measurement outcomes, and considering the system and the apparatus to be inseparable. This view contains similarities with instrumentalism, which doesn't try and understand the theory. But the Copenhagen interpretation is certainly an attempt to understand quantum mechanics, though is not very clear exactly how. There is contention on the exact definition of the Copenhagen interpretation. Asher Peres states that there "seem to be at least as many different Copenhagen interpretations as people who use that term." (Peres, 2002, p. 29).

Beginning with Bohr, Heisenberg, Born, von Neumann, and the other physicists who worked on quantum theory during its conception and development in the 1920s, the original Copenhagen

interpretation is a combination of Bohr's complementarity interpretation and correspondence principle, the completeness of quantum mechanics, and Born's statistical interpretation of the wave function. Bohr's complementarity principle states that there are certain properties of systems that are complementary – they cannot be known or measured at the same time. The general correspondence principle states that quantum mechanics gives classical mechanics in the limit as $\hbar$ goes to zero, or the number of particles goes to infinity, and when there is no longer any need of complementarity. Born's statistical interpretation of the wave function, from his work in 1926, was that the square of the absolute value of the function is a probability amplitude for a measurement outcome — now known as the Born rule.

In the 1920s the formalism of quantum mechanics was beginning to take a similar shape to the one we know today, but there was yet to be any significant work on interpreting the strange new theory. Though the Copenhagen interpretation is said to have originated from Heisenberg's work in 1927 and then discussed in talks given at the Fifth Solvay conference in 1927 (Camilleri, 2009a), much work has been done to show that the Copenhagen interpretation as the orthodox version now generally known as the 'Copenhagen interpretation' today, was a later invention of Heisenberg in the 1950s. This was done to present a united front against the new interpretations rising to oppose this 'orthodox' view, such as the hidden variables theories of de Broglie and Bohm, and also as a response to the Soviet Marxist critique of quantum mechanics' incompatibility with dialectic materialism.[10] According to Camilleri (2009a), both the defenders of the Copenhagen interpretation and its opponents are the reason for it appearing in history to be one clear and straightforward interpretation.

But the 'Bohr-Heisenberg' view did not exist at the early stage in the late 1920s – Bohr's own view was markedly different to others as he made no attempt to understand wave function collapse other than invoking the quantum/classical cut.[11] According to Gomatam (2007), Bohr's interpretation was developed to avoid the ontological wave-particle duality that the Copenhagen interpretation accepted as fundamental. Bohr had an inseparability hypothesis where the wave function was the joint state of the system and apparatus, and these two could not be separated

---

[10]A small subset of that work includes Howard (2004); Camilleri (2009a) and Camilleri and Schlosshauer (2015).

[11]For some more on Bohr's own view of the interpretation of quantum mechanics, see Gomatam (2007) and Camilleri and Schlosshauer (2015).

at the level of quantum formalism, as he did not believe the formalism to be describing the real states of the individual particles. He did not give the observer any fundamental role in measurement, and believed classical concepts to be indispensable. Gomatam describes Bohr as an anti-realist at the formalism level, while a realist at the experimentalist level (and so Bohr's original Copenhagen interpretation was not a Ψ-ontic one in the sense of chapter 3).

1.4 The Orthodox Copenhagen Interpretation

The 'orthodox' Copenhagen interpretation, a unified view presented in the 1950s to oppose the new interpretations of critics of Cophenhagen, mainly focuses on the problem of definite outcomes — the division between the quantum superposition and the classical objective measurement, the apparent 'collapse' of the wave function — and is often taken to use von Neumann's projection postulate to explain this.

In 1932, von Neumann's work on quantum measurement led to the projection postulate — that the entangled state of the quantum object and measurement apparatus collapses to a determinate state when measured (as discussed in more detail in chapter 2). This was to explain the lack of quantum superposition in macrosystems, and why when a quantum state is measured again, the same state will be measured. Von Neumann called this a type-1 process — a process that cannot be explained by quantum mechanics. Type-2 processes, however, can be explained by quantum mechanics, as they describe the unitary evolution of the quantum state in terms of the Schrödinger equation. For von Neumann, the dividing line between the quantum system and the classical measurement apparatus was a contextual one, though an observer, unlike the measurement apparatus, could never be involved in a type-2 process (Faye, 2019). Von Neumann's split therefore means that a type-1 process only takes place in the observer's consciousness. This idea of consciousness collapsing the wave function was picked up by Wigner (1961, 1967) and eventually led to the paradox of Wigner's friend (see section 2.2.5). Von Neumann, along with Heisenberg and Wheeler, accepted that there was a kind of collapse of the quantum wave function, though they all attempted to interpret this in different ways.[12] Bohr, however, did not mention

---

[12]For more details on how they did this, see Henderson (2010).

wave packet collapse in his work nor did he give the observer of a measurement any fundamental role, a significant divergence from this orthodox, 'unified' Copenhagen interpretation.

Despite all this general confusion, and the lack of a 'one true' Copenhagen interpretation, it is generally considered that it can be summarised by a few key tenets as summarised by Henderson (2010):

1. Quantum mechanics is complete.

2. Born's statistical interpretation of the wave function is correct.

3. There is wave function collapse.

4. These collapses are not able to be analysed.

Note there is disagreement over the last two between the original and orthodox, with Bohr not mentioning collapse at all, but they agree on the need not to change the quantum mechanical formalism. It depends on the notion of wave-particle duality and complementarity to interpret quantum theory, though Bohr's own version of complementarity is incompatible with this. While Bohr's own ideas can be described as "wholly epistemological" (Gomatam, 2007, p. 747), the orthodox Copenhagen interpretation is not. The orthodox view is realist at the formalism level, linking physical concepts with the quantum mechanical observables. With the view of the wave function collapsing (using the projection postulate) to an eigenstate, measurement seems to play some kind of causal role. The measurement problem is an example of the problem of quantum collapse. How do we get from quantum states to classical measurements? As an interpretation, the orthodox Copenhagen interpretation does very little to solve the problem.

1.5   Thesis Overview

As is clear even in just looking at the history of the Copenhagen interpretation, since the beginning of quantum mechanics there has been disagreement on how to interpret it – and how to solve the 'measurement problem'. The latter is presented as a trilemma by Maudlin (1995), consisting of three mutually inconsistent postulates of quantum mechanics:

1. The wave function is complete.

2. The wave function evolves unitarily according to the Schrödinger equation.

3. Measurements have unique outcomes.

Three types of interpretation stem from which of the postulates are dropped. Relinquishing postulate 1 is to propose a hidden variable theory – they state that there is more than just the wave function that governs the quantum world, though we cannot see what that is. Bohm theory is the first and most established of these theories. If instead you abandon postulate 2, you might adopt a collapse theory, where measurement occurs due to some non-unitary dynamic. The Many Worlds interpretation drops postulate 3, and instead has every possible outcome occur. This is the original splitting of interpretations of quantum physics with each choosing a different postulate to abandon in order to make sense of the theory.

More recently the measurement problem is instead divided into two distinct parts; the unique outcomes problem and the preferred basis problem. Each interpretation has to deal with these two problems, though often the problem of unique outcomes is the primary concern.[13] These are discussed in more detail in chapter 3.

One of the tools often used by these interpretations is decoherence theory, which is discussed in detail in chapter 4. Though there were key concepts of decoherence in the work of physicists from the late 1920s, the modern beginnings of decoherence theory is generally taken to be with Zeh (1970) and Zurek in the early 1980s.[14] As the Copenhagen interpretation was begun in the late 1920s and 'unified' in 1955 and was therefore formulated before decoherence theory was developed, it does not use decoherence theory as an aid to solve either measurement problem in its original form, though theoretical and experimental physicists took decoherence as a 'completion' or 'justification' of Bohr's views in the 1980s and 90s (Camilleri, 2009*b*).

Both Zeh and Zurek do not believe that decoherence by itself can solve the measurement problem; Zurek uses decoherence as the basis for his own 'existential interpretation', but believes that it only solves the preferred basis problem, and not the unique outcomes problem. Most

---

[13]There are notable exceptions to this – Bohmian mechanics, for example, is built around position as a solution to the preferred basis problem.
[14]See Zurek (1981, 1982).

other interpretations of quantum mechanics incorporate decoherence theory into their solutions of the measurement problem somehow – as a very successful research program with excellent empirical results, it is hard for an interpretation to discount decoherence and still be taken seriously. In this way, how well an interpretation incorporates decoherence is a good way of pressure-testing it. However, this is not true for all interpretations. Chapter 5 explores how the interpretations discussed in chapter 3 use decoherence theory in their frameworks, plus it explains an interpretation based entirely on decoherence theory. Some interpretations were developed parallel to decoherence theory and do not need it to solve for the measurement problem, though they are compatible with it. Collapse theories such as the GRW model have developed their own solution, but share similarities in their formalism with decoherence theory, though the theory itself is not used in their interpretation. Some interpretations, such as QBism, are affected very differently by decoherence theory due to their epistemic view of the wave function.

Decoherence is not the only process found at the transition between quantum and classical – chapter 6 explores a potential analogy for the transition, comparing it to a phase transition in statistical mechanics. I argue that this is a formal analogy, and it fails due to being unable to capture semiclassicality at the classical limit. Semiclassical theory shows the importance of keeping multiple representations of a theory, even if one, like classical theory, is considered fictional. This idea is taken further in chapter 7, where it is shown that many interpretations of quantum theory struggle to apply to more than non-relativistic quantum mechanics. Inequivalent representations of quantum phenomena in QFT and quantum statistical mechanics lead to the Coalesced Structures approach by Ruetsche (2011), where rather than searching for one true interpretation of quantum theory, we keep many inequivalent interpretations for different situations in order to increase our understanding. This leads to problems for realism; Ruetsche asks what good Structural Realists are supposed to believe in when we take up contents of theories and cast them away for different situations. Chapter 7 explores a new kind of realism put forward by Williams (2017) and Fraser (2018*b*) for this issue, and argues that to save it from empiricism it can combine with Ontic Structural Realism.

Decoherence, the measurement problem, and realism are revisited later; first, the investigation of quantum mechanics begins in an interpretation-free way to get more engaged with the

physics, using a technique that has been employed since the very start of the theory – thought experiments. Exploring quantum mechanical thought experiments through the development of quantum theory and learning interpretation-independent ways to think about the core issues in quantum mechanics in chapter 2 shows that there is more nuance to the problem of the quantum to classical transition than is found in the trilemma.

THOUGHT EXPERIMENTS

2.1 Introduction

One of the most famous thought experiments in physics is Schrödinger's cat. Its evocative and rather morbid imagery has captured the imaginations of scientists and non-scientists alike since 1935 and it is often the first thing people learn about the interpretation of quantum mechanics. Thought experiments more generally have been extremely important to the development of quantum theory, though exactly what role they play has been the subject of much discussion in the literature. It is debated whether they provide knowledge and whether they are necessary, with staunch empiricists like Norton on one side and platonists like Brown on the other. Between those two peaks lies a whole range of other views. One thing that cannot be argued with is that thought experiments are used very often in physics, and are notably prevalent in the history of quantum theory.

Section 2.2 looks at the development of quantum theory by reviewing quantum mechanical thought experiments from 19435 to 2018. The insight they provide into the foundational issues of quantum theory is as interpretation-free as possible which is why quantum mechanical thought experiments are so important in quantum physics. Section 2.3 looks at how thought experiments are discussed in general philosophy of science and the evolving ideas of what thought experiments

can give us. A new argument against Norton's view by Stuart (2015, 2016) is considered in conjunction with the idea of thought experiments being akin to analogue experiments. In section 2.4 it is asked what quantum mechanical thought experiments can teach us about thought experiments in general. Finally, section 2.6 asks what we learn about quantum mechanics from thought experiments.

2.2 The Development of Quantum Mechanics in Thought Experiments

Quantum mechanical thought experiments come in many forms. Some are exclusive to teaching a better understanding of quantum mechanical effects, such as quantum games, while others are like Schrödinger's cat and highlight problems in the theory. Thought experiments play a major role in the development of the other revolutionary part of early 20$^{\text{th}}$ century, namely in relativity theory. For Einstein, the *Chasing the Light* thought experiment provided a crucial stepping stone to Special Relativity, and the *Einstein's Elevator* thought experiment did much the same for General Relativity. These thought experiments helped lead to new theories, just as Galileo's ship did for the classical principle of relativity back in the 17$^{\text{th}}$ century. The thought experiments of Einstein and Galileo bear little resemblance to successful thought experiments in modern quantum mechanics. Newer examples of quantum mechanical thought experiments are much harder to visualise – something that no doubt restricts our ability to use them to transform our thinking, which is discussed in more detail in section 2.3. They often directly challenge our intuitions even more than older ones, but they still contain that same spark of suggestiveness.

Not all theory development involves thought experimentation; the Bohr model of the atom and Planck's model of radiation before that were not developed using revelations from thought experiments. Thought experiments are used for specific purposes; they were used by Einstein, for example, to challenge the empirical and pragmatic limitations of the uncertainty relation. This is an argument all sides agree he lost. Hence, while during the Einstein-Bohr debate there were many prominent thought experiments which did mark revolutionary theory change, that debate was conducted in terms of classical physics and uncertainty, and those thought

experiments are now only seen as being of historical interest.[1] While this chapter does not cover many of the Einstein-Bohr debate thought experiments it does cover some of the ones that were used after quantum theory was developed to better understand it and its implications. The thought experiments reviewed below do not mark a revolutionary theory change – they mark turning points *in* a theory, unlike, for example, Galileo's famous Leaning Tower of Pisa thought experiment.

What makes these quantum mechanical thought experiments special? A review of famous quantum mechanical thought experiments from the nascence of quantum theory, through to the modern day will give us an insight into foundational issues in quantum mechanics by showing us how the theory was always and still is debated and understood prior to any philosophical analysis.

2.2.1 EPR Paradox

The EPR paradox is used by Einstein, Podolsky and Rosen (1935) to show that the wave function in quantum mechanics is not a complete description of physical reality.

The paradox is set out in their paper as follows; either quantum theory is incomplete, or two non-commutating observables cannot have simultaneous reality. Using the assumption that the wave function is complete, they show that two non-commutating physical quantities can have simultaneous reality using entangled particles. Since this cannot be true (due to the uncertainty principle) then the wave function cannot give a complete description of physical reality.

Take two particles which have interacted and are in an entangled state of position and momentum, but have since moved away from each other. The measurement of one of the non-commutating quantities on particle A, such as momentum, would lead to the assumption of that particle having a definite momentum and an uncertain position. However, due to the entangled nature of the particles A and B, despite being far away, this suggests that somehow particle B 'knows' that it must have definite momentum and an uncertain position, otherwise the two quantities technically have simultaneous elements of reality. The wave function cannot then

---

[1]Peacock (2016) discusses this debate and the thought experiments involved in some detail.

be complete, as it allows for 'spooky action at a distance', or two particles which are space-like separated somehow knowing about measurements done on their entangled partners.

This famous thought experiment has remained pertinent in the wider quantum debate, and two qubits in a Bell state are called an EPR pair.[2] Real versions of the EPR set-up are now routinely performed.

Einstein's deep discomfort with the implications of quantum mechanics sparked a debate between him and Bohr, and this can be considered the ancestor of the debate now in the foundations of quantum mechanics — the debate over just how to interpret the theory. This thought experiment is not as picturesque as Schrödinger's cat, nor as striking as quantum immortality, which are both discussed later in this review, but it nevertheless is at the beginning of a debate that is still ongoing. It is an argument, but it is also, crucially, an experiment. Whether or not Einstein, Podolsky, and Rosen were right does not so much matter in the greater view of the debate. It was clearly meant to point out a contradiction in the theory that still has ramifications today. As seen with other thought experiments, it has also led to many real-world experiments, despite this not being its creators' main purpose.

2.2.2 Schrödinger's Cat

One of the most famous thought experiments, Schrödinger's cat, from Schrödinger (1935), was designed to discuss the work done in the EPR paper.[3] It has been very important for how physicists think about quantum mechanics, and interpretations of quantum mechanics all attempt to solve this paradox, as is discussed in chapter 3. Schrödinger meant as an attack on the Copenhagen interpretation by showing the ridiculousness of a superposition of macroscopic physical states.

In doing so he created the term 'verschränkung', or 'entanglement' — a term which is integral to quantum mechanics.[4] This thought experiment, despite being created to point out quantum theory's flaws, instead led to a greater understanding of the theory. It helped clarify the notion of entanglement and the measurement problem in a way that anyone can conceptualise, and it has

---

[2]Being in a Bell state is equivalent to being maximally entangled.
[3]Translated by Trimmer (1980).
[4]Entanglement is defined in chapter 1, p.7.

captured popular imagination. It also led to the thought experiment Wigner's friend, which is discussed in section 2.2.5.

Schrödinger tells us to imagine a cat in a box. Also inside this box is an atom with a fifty-fifty chance of decaying, a Geiger counter, and a vial of poison gas linked to the Geiger counter. When the atom decays, the Geiger counter will click and the gas will be released, killing the cat.

If the box is closed, and there is no way to look inside to see if the cat is alive or dead, the normalised quantum state describing this system is

$$(2.1) \qquad |\Psi_{Box}\rangle = \frac{1}{\sqrt{2}}(|\text{Undecayed}\rangle_{atom}\,|\text{Alive}\rangle_{cat} + |\text{Decayed}\rangle_{atom}\,|\text{Dead}\rangle_{cat}).$$

When the box is opened, the state collapses and the cat is seen to be either alive or dead. This, Schrödinger is arguing, is patently absurd — how can a cat be alive and dead at the same time, simply because it has not yet been measured? This is different to the discussion in EPR, as this is about entanglement. Quantum mechanics is linear because of the tensor product structure of combined systems, and so states can be in superpositions such as this one, as was stated in section 1.2.

This phenomenon of entanglement, however, is indeed found in the real world. This thought experiment has led to many real world experiments, looking into the 'cat state' in other systems, such as the singlet state from section 1.2,

$$(2.2) \qquad |\Psi\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle),$$

for example, which could signify two electrons in an entangled state of up and down spin states. In quantum computing, two qubits in a superpositions where they are either both 0 or both 1 would be described as

$$(2.3) \qquad |\Psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle).$$

Real experiments have been done by Leibfried et al. (2005) to see if "cat states" can be achieved by anything from photons to beryllium ions, and it is even posited by Li and Yin (2016)

that bacteria could be put into a cat state. As this experiment is achieved with larger and larger macrostates, it pushes what we think we know about the limit between quantum and classical theory. This thought experiment has led to real experiments which would have been impossible at the time, making a significant contribution to experimental progress.

2.2.3 Heisenberg's Microscope

This famous thought experiment uses classical optics to approximate the uncertainty principle used in quantum mechanics (Heisenberg, 1949). It is a very simple experiment using a microscope. One imagines looking at an electron by bouncing a photon off it. A microscope can only resolve an electron's position to an accuracy of

$$\Delta x = \frac{\lambda}{\sin \epsilon}, \tag{2.4}$$

where ϵ is the angle between the electron and the light ray. This inaccuracy is due to the Compton recoil.[5] This gives a momentum $\propto \frac{h}{\lambda}$,

$$\Delta p_x \approx \frac{h}{\lambda} \sin \epsilon. \tag{2.5}$$

Putting together these two equations gives the approximate uncertainty principle,

$$\Delta x \Delta p_x \approx \left(\frac{\lambda}{\sin \epsilon} \right) \left(\frac{h}{\lambda} \sin \epsilon \right) = h. \tag{2.6}$$

This thought experiment attempts to explain the uncertainty principle in quantum mechanics by using classical optics. This is not a flawless argument by any means; Heisenberg's thought experiment has been criticised by many including Bohr, his mentor. One of its main problems is assuming classical collisions in a system that will not behave classically. It also assumes a definite position for the electron, which cannot be attributed according to quantum theory — the theory Heisenberg is trying to explain using the thought experiment.

---

[5]Caused by the Compton scattering, the inelastic scattering of photons off of free charged particles. This collision causes the scattered photon's wavelength to change – the Compton shift – and the charged particle to recoil.

Heisenberg initially used this thought experiment as an argument for the uncertainty principle. It could also be used to illuminate the theory. Due to the fact that the argument stems from classical optics, however, this has since been found unsatisfactory, and so it can also be used to point out the weaknesses of the theory.

2.2.4 Renninger Negative-Result Experiment

The purpose of this thought experiment is to show just how unintuitive wave function collapse and measurement is in quantum theory. In a vein similar to Schrödinger's cat, which demonstrates the oddity of superposition and entanglement, Renninger uses the negative result experiment to show that while a measurement is usually thought of as the detection of a particle, it can also be the lack of detection of a particle. This seems counter-intuitive, but failure to observe something can indeed be a form of measurement.

Instead of using a cloud chamber to detect a particle, two hemispherical detectors are used in Renninger (1953)'s set-up. A radioactive atom about to decay is placed in the centre of the two detectors. It is assumed that the detectors are 100% efficient, and that the emitted α-particle is always detected.

The α-particle has a 50% probability of being detected by either detector — if one detector clicks, the other will not. Therefore no click from one detector is just as good as a click from the other; a non-detection is equivalent to a detection.

This set-up can be altered so that one detector is much larger than the other, so that it is further away from the radioactive atom, and so it takes the emitted particle much longer to be detected. In this case, the null result from the smaller detector is enough to show that the emitted particle will be detected by the other, larger one, despite it having not reached it yet. There is a lack of detection, but still a measurement.

Using standard quantum formalism, this is understood as a partial collapse of the wave function. The wave function of the α-particle collapses to a hemispherical shape when there is no 'click' from the other detector. It is known that the particle is not on one side, but it is still unknown where it exactly is on the other. But is this a paradox? There is wave function collapse of one side of the circular function, despite no actual measurement! How can there be wave function

collapse in the absence of a detection — how does the wave function interact with the smaller detector without being detected?

While there are a few objections to this thought experiment, most are simply due to a misunderstanding of quantum mechanics, and others have strong responses. For example, one objection is that the particle was always travelling on a classical trajectory in a superposition of plane waves, and the wave function is in fact the probability distribution – this would then solve the problem, as there would be no collapse – the particle would have always been travelling in a straight line. This view is confusing mixed states with pure states, however, and does not take into account that the α-particle is a pure state with a spherical wave function, and therefore cannot be a superposition of plane waves. It is well known that a spherically symmetric wave function collapses to a straight line trajectory, and was first observed in 1929. Mott (1929) noted that the behaviour of an $\alpha-$particle in a cloud chamber was that of a straight line trajectory despite its spherically symmetric wave function. Renninger's negative-result thought experiment is essentially a reformulation of this paradox with extended consequences.

This thought experiment aims to point out contradictions in quantum theory – to clarify just exactly what the contradictions were and describe them in a clear and understandable way.

2.2.5 Wigner's Friend

The Schrödinger's cat experiment was extended by Eugene Wigner (1961). In his thought experiment, Wigner placed the cat in a box and then placed that box within a larger box — a laboratory. He also placed an experimenter within that box, and himself outside of it. This, he believed, showed that wave function collapse was caused by consciousness. If the experimenter inside the closed system of the laboratory opened the smaller box containing the cat, this would then collapse the wave function for the experimenter, but not for Wigner outside of the laboratory. Instead, for Wigner, the experimenter friend would then be in an entangled state with the cat. This idea that a human could be in an entangled state seems impossible – the experimenter looking into the cat box must collapse the state for himself and for Wigner. This supports Wigner's argument that it is consciousness that collapses the wave function.

Though not explicitly discussed by Wigner, later versions of this thought experiment use a

basic state called the GHZ state:[6]

$$(2.7) \qquad |\Psi\rangle = \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle).$$

This is a maximally entangled quantum state, and will be discussed in more detail in section 2.2.11. It has since been discussed by many, including Penrose (2006), who argued that there is no paradox in having a superposition of mental states, and that this thought experiment does not show consciousness as having any part to play in quantum theory. Just as multiple states of a particle have no interaction with one another, simultaneous mental states have no need to interact either — there is nothing in quantum mechanics that disallows these simultaneous mental states.

In more current literature, both Wigner's friend and Schrödinger's cat are used by philosophers of physics in order to describe their own interpretations of quantum theory, and show how their view in particular renders the mysteries of quantum mechanics transparent. Richard Healey uses his pragmatist view to meet the 'challenge' of Wigner's friend;

> "The paradox of Wigner's friend challenges the objectivity of description in quantum theory. A pragmatist interpretation can meet this challenge by judicious appeal to decoherence." (Healey, 2013, p. 434)

It is also used by David Mermin in his discussion of QBism, or Quantum Bayesianism;

> "... in QBism Wigner's Friend is transformed from a paradox to a fundamental parable." (Mermin, 2017, p. 10)

The very central idea in QBism, as is discussed later in section 3.8 and 5.6, is that quantum mechanics is subjective – that there is no unique, observer-independent 'stuff' in the world. Wigner's friend, then, does nothing but support this interpretation and the lack of objectivity it seems to show in quantum theory.

---

[6]This is an entangled quantum state of more than 2 subsystems, and is named after Greenberger, Horne and Zeilinger, who first studied it in 1989 (Greenberger et al. (1989)).

This thought experiment has evolved from being a litmus test for the whole of quantum theory (as Schrödinger's Cat) and an effort to prove a fact about quantum theory (as Wigner's Friend) to now being used as an explanatory tool for new interpretations of the theory.[7] If interpretations can solve existing problems in the theory, it seems only right that they should gain more support. This thought experiment created to highlight a failing in quantum theory now shows a failing not in the theory itself but instead in how it is interpreted. It seems that Wigner then adapted Schrödinger's Cat to suit his argument that the wave function is collapsed by consciousness, that consciousness is necessary to the measurement process in quantum mechanics. Now it is often used to describe quantum theory to the uninitiated.

This thought experiment has the defining characteristic of being used as an inspiration and foundation for many real-world experiments, and lines of experimentation. It is perhaps this versatility which has given this thought experiment such lasting power in the thoughts and minds of scientists and non-scientists alike.

2.2.6 Bell's Inequalities

One of the more recent influential theorems in quantum mechanics, and once even called the "most profound discovery of science" (Stapp, 1975, p. 271), Bell's theorem is 'no-go' theorem which states that local hidden variable interpretations of quantum mechanics cannot replicate all the predictions of quantum mechanics.

John Stewart Bell (1964) wrote a paper based on the EPR paper, but following Bohm he used spin states rather than the more complicated position and momentum states. Bell writes that the problems of causality and locality could not be solved by bringing in additional local hidden variables. He shows that using these variables leads to predictions incompatible with quantum mechanics, and he does this using an inequality involving the correlations between results of measurements on the two entangled particles. Only non-local hidden variables satisfy the quantum mechanical predictions.

Bell's inequality rests on the assumption of both reality (real properties giving pre-determined measurement outcomes) and locality (no action at a distance). This inequality is violated when

---

[7]For example, it is used by Bub (2018) and Fuchs (2017) to defend their interpretations of quantum theory.

one or both of these assumptions are — as is found in quantum mechanics. This inequality was later generalised to the CHSH inequality by Clauser et al. (1969).

The set-up generally consists of the preparation of two maximally entangled particles, which are then measured separately by 'Alice' and 'Bob'. The correlations between the two measurements are shown to be stronger than classical correlations.

$$(2.8) \qquad S = |C(A,B) + C(A,B') + C(A',B) - C(A',B')|,$$

where A, A', B, B' are the detector settings for Alice and Bob respectively, C is the quantum correlation of the pair of particles, and the maximal classical value of the inequality is the following:

$$(2.9) \qquad |S| \leqslant 2,$$

As A, A', B, B' are physical variables, this follows:

$$(2.10) \qquad E(A,B) = P_{AB}(1,1) + P_{AB}(-1,-1) - P_{AB}(1,-1) - P_{AB}(-1,1).$$

Quantum mechanics violates the inequality and gives $|S| \leqslant 2\sqrt{2}$, which is known as Tsirelson's bound.[8] This is the maximum violation of the inequalities that quantum theory can give – any more, and there would have to be faster than light signalling.

These inequalities are still used in modern experimentation. Recent work has been done to explore the possibility of the Bell inequalities being violated by cosmological observables, which could then be used in the argument for cosmic scale features being "produced by quantum mechanical effects in the very early universe" (Maldacena, 2016, p. 10). Hardy (2017) has also recently written about Bell inequalities with retarded settings[9]. Bell inequalities are still present in the cutting edge of quantum theory research.

The Bell inequalities originate from Bell's work on the paradox posed by EPR, and they are akin to a litmus test. Though they started off as a thought experiment – an experiment that could

---

[8]This is named after the author of the paper where the first one was derived, Cirel'son (1980).

[9]Retarded settings are the value the setting would have taken if there had been no external interference, whether from humans or computers.

not be done in real life, and so done in the mind – they became a whole class of real experiments and are used to this day as a measure to see whether a system is quantum or classical. If a system violates the inequalities, it cannot be classical. It might be more precise to say that Bell's theorem is the thought experiment, the argument against local hidden variable theories that is described in the fashion of an experiment, and the inequalities are the product of the formation of this argument. In this case, this thought experiment has produced a better understanding of quantum and classical systems in the form of the inequalities. Bell created it with the intention of showing that local hidden variable theories violated quantum theory, though it has since grown beyond this task.

2.2.7 Wheeler's Delayed-Choice Experiment

Wheeler (1978) and Wheeler, Zurek and Ballentine (1984) used several thought experiments which can be grouped under the title of 'delayed-choice' (Marlow, 2012). Wheeler's aim was to explore the wave-particle duality of quantum phenomena, and how they 'know' when to behave as particles and when to behave as waves. In the basic two slit experimental set-up, the interference fringes vanish if there is a detector at one of the slits and it is known which slit the electron passes through. Does the electron 'know', before it is emitted, what apparatus it will be passing through? Does it 'decide', retroactively, whether to travel as a particle or as a wave?

One of the more famous thought experiments that Wheeler (1978) uses is the cosmic interferometer. While there were some arguments that the choice of detector in an interference experiment could affect how the light was propagated, Wheeler used this thought experiment to show that arguments to be wrong. Light is emitted from a quasar, far away from Earth, and due to gravitational lensing from a galaxy in-between, appears to detectors on Earth to be coming from two different directions. This light can be measured as two distinct streams of photons — as particles — by pointing two detectors in the two respective directions. It can also be measured with an interferometer — as a wave — combining the two directions. Wheeler used this thought experiment to argue that there is no retrocausality — the light cannot retroactively decide to travel as a wave or as a particle once it knows what detector awaits it on Earth. While in a small interferometer experiment, there might be a possibility that the type of detector could

be somehow transmitted to the photons emerging from the source and therefore influencing how they propagate, this cannot be the case with light from the Quasar. The light cannot have retroactively changed how it propagated through the universe light years previously from the scientists switching between detectors on Earth. There is no retrocausality – the choice of detector determines if interference is seen, and if the light is detected as a wave or a particle.

While this thought experiment cannot be easily turned into a real world experiment, the main point Wheeler wanted to prove can be shown using real-world experiments, such as in Jacques et al. (2007). These show that light does not decide, retroactively or otherwise, whether to travel as a wave or a particle, but rather travels in an indeterminate state until the moment of measurement. Wheeler puts forward these thought experiments to explore the wave-particle duality of photons, and though they can be used as arguments against any retroactive decisions on the part of the photon, they have led to the construction of many real-world experiments.[10] This thought experiment is critical of wave-particle duality and shows that there is no retrocausality, though it also illuminates the details of wave-particle duality.

2.2.8 Leggett-Garg Inequality

Similar to how Bell's inequalities are a measure of correlations between spatially separated systems, the Leggett-Garg inequalities (or LGIs) are a measure of a single system's correlations at different times (Emary et al., 2013). The inequality is, in its simplest form:

$$(2.11) \qquad\qquad K \equiv C_{21} + C_{32} - C_{31} \leqslant 1,$$

where the variables are $Q_{\alpha,\beta} = \pm 1$ at times $t = t_\alpha$ and $t = t_\beta$ and are dichotomous, with the correlation function of the variables being $C_{\alpha\beta} = \langle Q_\alpha(t_\alpha) Q_\beta(t_\beta) \rangle$ (Emary et al. (2013), p.1). The violation of this inequality implies that there is no reality to the macroscopic description of the system.

Leggett and Garg (1985) state that most physicists have a basic idea of a macroscopic reality, which requires two assumptions;

---

[10]The example of Jacques et al. (2007) has already been mentioned. Ma et al. (2016) gives a review on many realisations of Wheeler's delayed choice thought experiments.

1. If a macrosystem has two (or more) distinctly different macrostates available to it, it will be in one or the other at all times. It therefore has a unique, well-defined, pre-existing value when measured (Emary et al., 2013).

2. It is possible to perform a non-invasive measurement on a macrosystem.

These assumptions of a macroscopic reality, however, are not adhered to when quantum mechanics is "extrapolated to the macroscopic level". The idea of there being macroscopic coherence — as explored in the Schrödinger' cat thought experiment — is completely unintuitive compared to our understanding of our macroscopic world. The inequality was created to test macroscopic systems for quantum coherence — if a system violates the inequality, it cannot be measured without being disturbed, or has no realistic description (Emary et al., 2013). The Leggett-Garg inequalities, therefore, can be used as a test of a system's quantum or classical basis — if it violates the inequalities, it cannot be considered real in a macroscopic sense.

In their 1985 paper, Leggett and Garg construct the inequalities by considering an as of yet unobserved macroscopic quantum coherence in an rf SQUID.[11]

The inequalities have used to test a number of different systems since their conception in 1985, but are now more generally used as a measure of the 'quantumness' of the system. When it is not clear whether a system is behaving classically or quantum-mechanically, the LGIs can be used to explore quantitatively the transition between the quantum and classical worlds. Kofler and Brukner (2007) used the LGIs to show that the classical world and laws emerge out of coarse-grained quantum theory.

They are also used in current experimentation. Work done by Formaggio et al. (2016) on the violation of the LGI by neutrino oscillations shows that quantum coherence can 'apply broadly to microscopic systems across macroscopic distances'. As neutrinos exist in a superposition of 'flavours' — electron, mu and tau neutrinos — with this superposition remaining coherent over astrophysical distances, such as between the sun and the Earth, this kind of experiment is very different to testing on photons or similar, with short coherence lengths. This particular experiment tested a violation of 6σ over 735km — the longest distance the violation of the LGI or

---

[11]A radio frequency Superconducting Quantum Interference Device, or an rf SQUID, is a very sensitive detector used to measure small magnetic fields.

even Bell's inequalities have been tested over. The results show that alternatives to quantum mechanics are strongly constrained.

While they can be used as a real world experimental measure, the LGI were 'thought up' by Leggett and Garg. In this way, they are a thought experiment. This could not have been done in a real-world experiment, but only through exploring the concept of a macrorealistic system rather than focusing on the details of any particular example of such a system. Thought experiments, after all, are devices of the imagination (Brown and Fehige, 2017); only through mental investigation on what exactly macroscopic realism is, and what you would expect should you experiment, could lead to the LGIs. This is sufficient for the LGIs to be called a thought experiment. The fact that they can be used in real-world experiments also sets them apart from mere arguments – Leggett and Garg in essence created a tool that can be used in experimentation.

Leggett and Garg used this to show that quantum mechanics is not a macrorealistic theory. It also can be used as a measure of the "quantumness" of a system. It can be used to pick out quantum systems and therefore backs quantum theory. It has also been used in experiments, and can be used as a litmus test for 'quantumness'. It could be argued that it is more of an indicator than an illuminator of quantumness.

2.2.9 Hardy's Paradox

Unlike the previous thought experiments, which involved interaction-free measurement, Hardy (1992)'s paradox is interaction without consequences. The idea of the thought experiment is to show that it is possible for a particle-antiparticle pair to interact without annihilating. It uses two Mach-Zehnder interferometers in its set-up, and uses electron-positron for the particle-antiparticle pair. This is a different arrangement to before, as the electron-positron pair travel in curves rather than straight lines.

The paradox is that there is a non-zero (P$=\frac{1}{16}$) probability of both particles being detected at one of the detectors, which means there was an obstructing particle — they have interacted without annihilating.

Hardy's thought experiment is intended to point out a contradiction in a theory — to illustrate the paradox of an interaction without consequences in quantum mechanics. A real-world

experiment using weak measurement has since been performed to investigate this paradox by Lundeen and Steinberg (2009), where the paradox was apparently resolved. Like the Elitzur-Vaidman thought experiment, it has provided an invaluable link between theory and real-world experiments — and important feature of these quantum mechanical thought experiments.

2.2.10 Elitzur-Vaidman Bomb Tester

Like the Renninger thought experiment discussed previously, this thought experiment also involves 'interaction-free' measurement, although with marked differences. In the Renninger case, the purpose of the measurement or detection is the preparation of the state — information is known prior to the measurement, as one knows that the emitted particle will be detected at one out of two places. In the Elitzur-Vaidman (EV) thought experiment, the emphasis of the 'interaction-free' measurement is placed on being able to find out information about the state without interaction.

The thought experiment can be described both in the quantum mechanical formalism, using photons and a particle interferometer, and in a more 'dramatic' way which does not rely on any specific interpretation. Interestingly, then, this picturesque form is used by Elitzur and Vaidman to divorce the thought experiment from any potential biases that could appear in the more technical version, rather than to be more understandable or more vivid. While it may also help with those aspects, it is not the main purpose. It also does not assume any meaning of 'interaction-free' and similar concepts.

The thought experiment is as follows;

There is a stock of bombs, all with a new kind of sensor. When the sensor detects a single photon, it explodes. Some of the stock of bombs are faulty, and do not explode when a photon hits the sensor. They can, of course, be tested by attempting to activate them and seeing whether or not they explode, but this would lead to the loss of all the working bombs. Is there a way to test the bombs without wasting the ones that work? A kind of 'interaction-free' test?

It turns out that this is indeed possible. It uses an interferometer (similar to the Mach-Zehnder interferometer used in classical optics) where the photon first passes through a beam splitter. The photon continues on half the time (to the right) and the other half of the time it is

sent vertically upwards. Both of these paths reflect off a mirror, which brings the paths together again at another beam splitter. Two detectors after the second beam splitter detect the photon (See Fig. 2.3).

The apparatus can be arranged so that, due to destructive interference, only D1 detects a photon. Then, by putting a 'blocker' on one of the arms, both detectors will have equal probability of detecting the photon once more. As we know that we set up the interferometer so that only D1 should detect the photon, we then know that there is something blocking the arm, and have found out information about the existence of an object there without actually interacting with it.
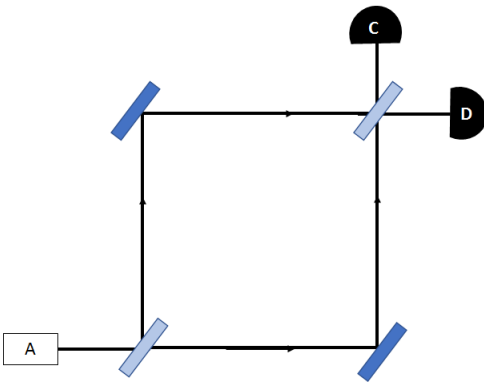
Figure 2.1: The interferometer without anything blocking the arms.
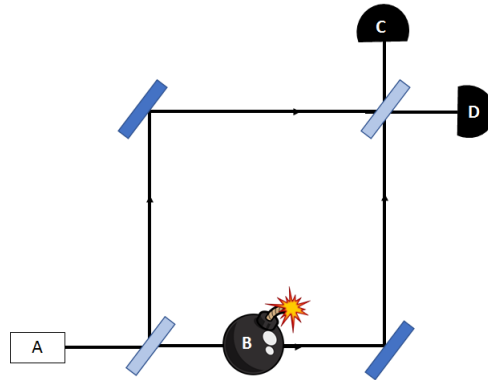
Figure 2.2: The interferometer, same as (a), but with the bomb B blocking one of the arms.

Figure 2.3: The Elitzur-Vaidman bomb testing experiment, where A is the photon emitter, B is the bomb, and C and D are the detectors. The mirrors are blue, with the paler mirrors being semi-transparent.

Putting the bomb into the interferometer as the blocker, as seen in (b), and after sending one photon through, we know that either D will click or the bomb will explode. There are three outcomes if the bomb is not broken;

1. p(bomb explodes) $= \frac{1}{2}$

2. p(C clicks) $= \frac{1}{4}$

3. p(D clicks) $= \frac{1}{4}$

If the bomb is not destroyed, we send another photon. To improve the apparatus, we make the first mirror almost transparent and the second almost not transparent. If the bomb is broken D never clicks. We can find a good bomb without destroying it approximately half the time, thus not exploding half of the working bombs — a much better result than testing all of them in the normal way and losing all the working bombs to explosions. In this way, we can measure the existence of an object of an unstable system without disturbing the internal quantum state — a measurement without an interaction.

This thought experiment was constructed by Elitzur and Vaidman (1993), and then performed as a real-world experiment by Kwiat et al. (1995). It could be argued that it undermines quantum mechanics, showing that there can be measurement without interaction, something quantum theory does not generally allow. But this thought experiment is using quantum theory to show this interaction-free measurement, and so is actually pointing out a quirk of the theory rather than undermining it. It posits a potential consequence of the theory, but does not necessarily clarify — if anything, it muddies the waters further.

2.2.11 PR Boxes

Bell's theorem, as described previously, states that any hidden variable theories of quantum mechanics must be non-local. Popescu and Rohrlich (1994) explored this further by taking nonlocality as a quantum principle. They combine this with the principle of relativistic causality. A local variable is, by definition, relativistic, as it is only influenced by events in its past light cone and can only influence events in its future light cone. Assuming that quantum mechanics preserves relativistic causality, these two principles can be taken as two possible axioms for quantum mechanics, echoing the two postulates Einstein uses for Special Relativity.

As quantum mechanics is the only theory which is consistent with both relativistic causality and nonlocality, Popescu and Rohrlich use these as their two axioms. However, when applied to the CHSH inequality, the nonlocality axiom predicts that quantum mechanics should maximally violate this inequality with a value of 4, the algebraic maximal violation — much greater than the actual upper bound (Tsirelson's bound) of $2\sqrt{2}$ quantum mechanics actually gives us. Yet this maximal violation still obeys the nonlocality "no-signalling" principle. The second axiom

of relativistic causality does not constrain the violation to $2\sqrt{2}$ either. Popescu and Rohrlich discover that neither of their axioms restrict the maximal violation of 4 to Tsirelson's bound.

It is clear that those two axioms are not enough to completely determine quantum mechanics — Popescu and Rohrlich write that the nonlocality axiom does not encompass effects like the AB effect. Rather than being simply a nonlocal theory, according to PR, quantum mechanics could instead be the "only causal theory with nonlocal equations of motion".

In an effort to derive quantum mechanics from minimal physical principles, Popescu and Rohrlich instead show that quantum mechanics is not the only nonlocal theory, or even the most nonlocal. There are also theories with 'superquantum correlations' which do not violate relativistic causality. The two principles of NL and RC are not by themselves enough to fully derive quantum mechanics.

These superquantum correlations, named PR box correlations after the two authors, are still used today in quantum theory. These theoretical stronger correlations are used to ask why quantum mechanics does not violate the inequalities more strongly, why we only see quantum correlations in the world. Work by Oppenheim and Wehner (2010) has been done on how the uncertainty principle restrains how nonlocal quantum mechanics can be. Other potential candidates for principles to constrain general no-signalling theories to quantum correlations are also explored in the literature, such as the principle of information causality (IC) by Pawłowski et al. (2009).

A strong candidate for the third axiom necessary to whittle the superquantum correlations down to normal quantum correlations is the use of the classical limit as an axiom. Rohrlich (2014) uses the classical limit — where macroscopic observables commute — as a minimal third axiom. This classical limit is a minimal axiom in the sense that if quantum mechanics is necessarily constrained by it, then so must any theories which are generalisations of quantum mechanics — even superquantum correlations. In this limit, however, superquantum correlations in fact violate the axiom of relativistic causality (RC). Rohrlich shows that with unlimited resources, there can be superluminal signalling.

The normal quantum correlations can then be found by using the three axioms to derive Tsirelson's bound, which does not assume quantum mechanics.

This statement is tested in the more recent paper by Rohrlich and Hetzroni (2016). As it was shown that the bipartite superquantum correlations violate the RC axiom in the classical limit, this paper looks instead at GHZ correlations, which are considered to be "the tripartite version of PR-box correlations". As discussed previously, a Greenberger-Horne-Zeilinger state is an entangled quantum state of more than two subsytems, the simplest of which being the three qubit state

$$|\Psi_{GHZ}\rangle = \frac{|000\rangle + |111\rangle}{\sqrt{2}},$$

(2.12)

with the more general formalism being

$$|\Psi_{GHZ}\rangle = \frac{|0\rangle^{\otimes M} + |1\rangle^{\otimes M}}{\sqrt{2}},$$

(2.13)

where the number of subsystems is > 2.

As GHZ correlations are quantum correlations, if they also violate RC in the classical limit then the argument in Rohrlich (2014) must be wrong. It is found that GHZ correlations do not signal superluminally in the classical limit, and the argument in Rohrlich (2014) passes the test. In fact Rohrlich and Hetzroni argue that the nonlocality axiom is not necessary once the classical limit is brought into play, as quantum mechanics is cannot be any more nonlocal than it is without violating the RC axiom. The NL axiom is therefore extraneous, and one only needs the RC axiom and the minimal classical limit axiom in order to derive quantum mechanics.

This thought experiment shows quite clearly how, in the words of Mach, thought experiments are techniques of discovery. There is no way to do real world experiments on superquantum correlations, as they cannot exist. They are merely an alternative theory that can be held up against quantum mechanics as a comparison, in order to tell us more about quantum mechanics. While we can show null results in real world experiments, proving that superquantum correlations are unphysical, there is little else they can show us without first moving into the realm of thought experimentation and examining this alternative theory more closely. Thought experiments such as PR boxes, looking into the realm of quantum theory and its possible consequences that are

beyond capabilities of real world experimentation, bring real epistemic value that cannot be explained away.

Popescu and Rohrlich created this thought experiment when trying to derive quantum theory from axioms, though in a way PR boxes also deny certain phenomena created from use of previous thought experiments. The superquantum correlations that resulted from their work have gone on to be used in much theoretical discussion, but as they are non-physical, they can never be tested experimentally — only their absence can be. This thought experiment helps us to learn more about quantum theory without using real-world experimentation, and in fact could not be done at all except for in thought.

2.2.12 Quantum Suicide and Immortality

The strangeness of quantum theory has led to some very striking thought experiments. Much as how quantum interpretations aim to solve the challenges posed by thought experiments such as Wigner's friend, thought experiments are also to attempt to distinguish between interpretations of quantum theory and their implications. Quantum suicide builds once more on Schrödinger's cat, but instead of being from the point of view of the experimenter friend, it is instead from that of the cat inside the box.

As explained in more detail in chapter 3, in the Many Worlds Interpretation (MWI), the universe 'branches' out, and all terms in a superposition correspond to a measurement outcome in a branch. When the box is opened, the world branches into two; one world where the cat is alive, and one where the cat is dead. Both outcomes are therefore real, but which one is measured depends on which branch you are in.

Quantum immortality, as written about by Tegmark (1998b, 2007), refers to the idea that the consciousness of the cat follows the path where it remains alive, and therefore, from the point of view of the cat, it always lives. It is this idea — that one always stays in the world where one is still alive — that leads to the idea of some sort of immortality despite the odds. The idea produced by this thought experiment is that while an observer cannot distinguish between interpretations of quantum theory, someone whose life depends on a binary quantum event can; that is, if they never die, then we live in a world where the MWI is true.

This is not, of course, a particularly watertight argument; as Max Tegmark points out, one's survival very rarely, if at all, depends on a binary quantum event. It is only within the abstract thought experiment that the person playing the role of the cat realises that they are immortal. It may not be a good test of which interpretation is correct but it adds another issue for the various quantum interpretations to address, especially Everettianism.

2.2.13 Quantum Games

2.2.13.1 Quantum Pseudo-Telepathy

This thought experiment, with another evocative title, was used to show the difference between quantum and classical theory by Brassard et al. (1999) — to show how much more communication via hidden variables is needed in classical game theory in order to correspond to the same communication in quantum game theory (QGT). It was later given the name 'pseudo-telepathy' by Brassard et al. (2003, 2005) in order to adequately describe the difference.

The basic idea of this thought experiment is as follows;

In a game of coordination, quantum players will have an anomalously high success rate — due, of course, to the quantum phenomenon of entanglement. In contrast, for two classical players to achieve the same success rate, they would have to be communicating telepathically.

Does this count as a thought experiment? It is a picturesque way of describing phenomena found in quantum theory, but it is not an argument. Nevertheless, the 'games' described, with their classical analogues, showcase the non-local characteristics of quantum theory, and these can be shown in real world experiments (see Freedman and Clauser (1972) and Weihs et al. (1998) for some examples). One can learn about quantum theory from engaging with this experiment, as it displays complex theoretical concepts in a form much more easily digested.

2.2.13.2 Quantum Tic Tac Toe

There are specific examples of games in QGT, such as the Mermin-Peres magic squares game, but one created specifically for the purpose of teaching quantum entanglement by Goff (2006) is quantum tic tac toe.

This game uses three quantum phenomena in its construction: superposition, entanglement, and collapse. By contrasting the normal 'classical' game of tic tac toe with the quantum alternative, Goff is attempting to make these three unintuitive concepts easier to grasp conceptually. The game is essentially an abstracted quantum experiment on entangled qubits in superpositions, which eventually collapse to return to the classical value. Goff calls it a 'teaching metaphor', where students can learn to understand the phenomena of quantum theory without having to deal with difficult mathematics. This, he states, gives a conceptual foundation in quantum mechanics whilst also being fun to play.

Once again, this is not an argument. It is not even particularly picturesque — it is just tic tac toe, but made much more complex, with quantum features such as decoherence and interference.[12] Though as we have seen previously, with thought experiments such as Bell's inequalities, being picturesque is not a necessary condition for something being a thought experiment. If it is anything apart from a classroom game, it is an incredibly abstracted quantum system, where none of the details remain except the behaviour. If anything, the picturesque-ness of the experiment have been taken away, to leave only the bare bones and phenomena.

It can be called an experiment, however; even if the only apparatus might be a piece of paper and some pens, or a whiteboard, it is nevertheless the manipulation of superpositions of 'qubits' to lead to the creation of entanglement and collapse. The rules of the phenomena are essentially the same as they would be in the real world, only abstracted.

These games are used often as teaching material, though quantum games are also used in theoretical physics as a way to investigate the powers of the limitations of non-locality in quantum theory; for example, Cleve et al. (2004) use nonlocal games to investigate quantum entanglement.

2.2.14 PBR Theorem

The PBR theorem (named after the three authors Pusey, Barrett, and Rudolph), while not explicitly called a thought experiment, is nevertheless a 'no-go' theorem that made waves in the foundations of quantum theory. Pusey, Barrett and Rudolph (2012) aimed to prove the reality of

---

[12]For a full description and rules see Goff (2006).

the quantum state by presenting a 'no-go' theorem to show that viewing the state as epistemic led to contradictions with quantum theory.

They do this by first making two assumptions;

- The system has a real, physical state independent of an observer,

- Independently prepared systems have independent physical states.

They take a quantum system with a physical state, λ, that determines the probabilities for each of the different measurement outcomes for the system, according to some model. The system is then prepared so that it is in the pure quantum state $|\Psi\rangle$, which does not fix λ but has a probability distribution $\mu_\Psi(\lambda)$ associated with it. This probability distribution represents an agent's state of knowledge of the system. Using the second assumption, if one independently prepares two states, $|\Psi\rangle$ and $|\Phi\rangle$, if they are ontic they cannot overlap, as they should have independent physical states. If the two probability distributions do overlap, then the two states cannot be uniquely inferred and share some physical outcomes.
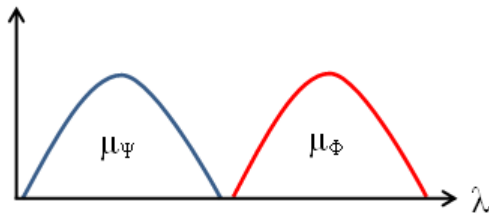
Figure 2.4: Without overlap.   Figure 2.5: With overlap.
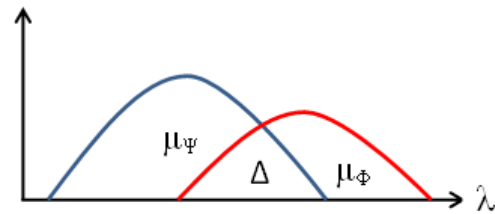
Figure 2.6: The probability distributions of two quantum states, one where there is no overlap and they can be considered real, and one with overlap Δ where they have to be considered as epistemic.

Pusey et al. (2012) make the argument that the overlap $\Delta \neq 0$ contradicts quantum theory. They do this by preparing two systems independently and bringing them together for an entangled measurement. This measurement uses the bases $|0\rangle$ and $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, giving probability $q > 0$ that the preparation of either state gives a physical state λ from the overlap in the probability distribution. This then gives a probability $q^2 > 0$ that the two physical states λ_1 and λ_2 are from the overlap area Δ.

Each of the four orthogonal states that the entangled measurement projects onto is orthogonal to a possible state of the systems, and quantum theory gives us a probability of each:

$$
\begin{aligned}
|\phi_1\rangle &= \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle) \rightarrow \quad \text{probability zero when} \quad |00\rangle, \\
|\phi_2\rangle &= \frac{1}{\sqrt{2}}(|0-\rangle + |1+\rangle) \rightarrow \quad \text{probability zero when} \quad |0+\rangle, \\
|\phi_3\rangle &= \frac{1}{\sqrt{2}}(|+-\rangle + |-0\rangle) \rightarrow \quad \text{probability zero when} \quad |+0\rangle, \\
|\phi_4\rangle &= \frac{1}{\sqrt{2}}(|+-\rangle + |-+\rangle) \rightarrow \quad \text{probability zero when} \quad |++\rangle.
\end{aligned}
$$

(2.14)

Therefore least q^2 of the time the theory gives a probability of zero, while the measuring device is uncertain and could therefore give a non-zero probability due to the overlap between the two states. The two states can therefore not overlap. If there is any overlap between them, then there is a non-zero probability of the theory conflicting with experimental results.

Accepting this result is, of course, dependent on accepting the two assumptions.

> "One [assumption] is that a system *has* a 'real physical state' — not necessarily completely described by quantum theory, but objective and independent of the observer."
> (Pusey et al., 2012, p. 1)[13]

If one does not believe that a system has a state independent from an observer – as some quantum interpretations do, such as QBism – then this no-go theorem holds no weight.[14] However, the PBR theorem was an extremely influential no-go theorem when it was published.

This is another 'no-go' theorem, but it can be called a thought experiment for the following reasons; it is the visualisation of an experiment that was done 'in thought' rather than the real world, and it also provided a 'result', new knowledge that would not have been gained had the experiment not been done.

2.2.15 Brukner's No-Go Theorem for Observer-Independent Facts

Recently the Wigner's friend thought experiment has experienced something of a renaissance in the literature, and is used by Brukner in the proof of his no-go theorem (Brukner, 2015,

---

[13]Italics for emphasis from original source.
[14]The QBist interpretation will be discussed in detail in chapters 3 and 5.

2018). Following on from his 2015 paper where he argues that "facts" can only exist relative to the observer, Brukner uses Deutch's extended version of Wigner's friend in his proof to show that, to be compatible with the assumption of observer-independent facts, quantum theory must disregard one of the following three assumptions;[15],[16]

1. Locality: The choice of measurement settings has no effect on the choice of settings of distant observers.

2. Freedom of Choice: There is freedom of choice of measurement settings that is independent of the rest of the experiment

3. Universality: Quantum theory can be applied at any scale.

He shows this by taking the famous thought experiment but with one key alteration – while Wigner, outside the isolated lab, does not know his friend's measurement outcome, he does know that they have observed a definite outcome. This means that Wigner can choose between two different ways of measuring the lab; he can take his own measurement or ask his friend directly what measurement outcome they recorded. These two statements should, in a world where facts are observer-independent, should both be able to be assigned truth values — however, in the proof, Brukner shows that this is not the case, and this choice of measurement violates the CHSH inequality. This means that Wigner's measurement, and the friend's measurement, cannot both be local objective properties of the world.

> "...there is no theoretical framework where one can assign jointly the truth values to observational propositions of different observers (they cannot build a single Boolean algebra) under these assumptions." (Brukner, 2018, p. 7)

To have them be objective properties, Brukner, argues, one must therefore drop one of the three assumptions listed previously. Either they are not *local* objective properties, Wigner has no choice of measurement basis, or quantum theory cannot be applied universally (over all scales).

---

[15]Where Brukner means "facts" here as direct experiences of observers in terms of detector clicks or pointer positions — where the Born rule gives a probability of one.
[16]Found in Deutsch (1985).

This extended Wigner's friend is a thought experiment and as a 'no-go' theorem it points out a problem with a theory. And yet it is also expanding our understanding of quantum theory. But it seems to be more than just a thought experiment, similar to PBR theorem, as it produces an apparent result – that one of those assumptions must be dropped if there are objective properties.

2.2.16 Frauchiger and Renner's Inconsistent Quantum Theory

Yet another extended Wigner's friend thought experiment has recently been published; in Frauchiger and Renner (2018), they describe their own no-go theorem which states that quantum theory cannot have all of the following characteristics;

- Universally valid.

- Consistent between agents.

- Measurements have single outcomes (from the viewpoint of an agent carrying out those measurements).

Their set-up consists of a double Wigner's friend; two friends in two different isolated labs and two Wigners. On top of that, the initial spin state measurement depends on a random value, which is known to the agents outside of the isolated lab.

The thought experiment roughly follows the following steps:[17]

- The first friend, F_1, measures the random value and sends to F_2.

- F_2 measures the spin state.

- Wigner$_1$ measures Lab$_1$ in a certain basis regarding the random generated value, and gets a result of either ok or fail.

- Wigner$_2$ measures Lab$_2$ in a certain basis regarding the spin state and gets a result of either ok or fail.

---

[17]For a full, detailed proof go to Frauchiger and Renner (2018).

In their proof, Frauchiger and Renner show that with this set-up, when the Friends and the Wigners use their measurements and quantum theory to make statements about what they believe the others will have measured, these will be inconsistent with each other. According to Frauchiger and Renner, this shows that to make sure quantum theory can consistently describe the use of itself, it cannot satisfy all the assumptions above. Which of the assumptions are invalid depends on which interpretation of quantum theory is applied. For example, the third assumption — that measurements have single outcomes — is violated by the Many-Worlds interpretation.

Much like the previous thought experiment, this is also an extended Wigner's friend that shows (in this case, literal) inconsistencies in the theory. It is also a 'no-go' theorem with results that further our understanding of interpretations of quantum theories and of quantum theory itself. And much like the previous thought experiment, it seems go beyond the normal use of a thought experiment.

2.3 Thought Experiments in the Philosophy of Science

Before we look at the lessons we can learn from the thought experiments used in the development of quantum theory, it is important to review how thought experiments are considered in general philosophy of science.

In 1897 Ernst Mach wrote an article arguing for the importance of thought experiments. Mach was a strong proponent of thought experiments, which he believed were performed on a higher intellectual level than real world experiments, but were nevertheless conducted using the same method; the method of variation. For Mach and Hiebert (1976), thought experiments were indispensable in the study of physics, as they brought with them the greatest transformation in our thinking and revealed the most significant paths of investigation. Even when they led to mere conjecture as opposed to a definite result, this was not a waste of time but rather a excellent exercise for cognitive development and perfect for teaching. Mach was incredibly influential on many physicists of that era, who went on to use thought experiments in their work around the time of the birth of quantum theory. The most notable of those influenced is Einstein, who described many thought experiments in his writing.

Karl Popper created a taxonomy for thought experiments, using mainly physics thought experiments from the Einstein-Bohr debate, in 1959. This taxonomy divided up thought experiments by what role they performed for theory choice — whether they illuminated, criticised, or supported a theory. But as the previous review of modern quantum physics thought experiments shows, they do not fit well into these three categories. They no longer resemble famous thought experiments such as Einstein's Elevator, by being a suggestive, imaginative scenario to illuminate or spark change. They are more mathematical, more technical — some, like PR boxes, even seem to provide us with knowledge we could not have ever learnt from real-world experiments. Bell's inequalities seem also to do this, though not to the empiricist. They are a tool that can be used to discover new knowledge from the world, but it is in itself not knowledge. This, however, is debatable. Without Bell's theorem, would the inequalities have been formed? Would quantum theory be the same today without the "most profound discovery of science"?

Popper's three main categories are as follows;

1. HEURISTIC — These are thought experiments which help to illuminate/clarify a theory.

2. CRITICAL — These are thought experiments which criticise or point out contradictions in a theory.

3. APOLOGETIC — These are thought experiments which support a certain theory.

A thought experiment which helps illustrate or explain a theory is useful, such as *Einstein's Elevator*, as is a thought experiment which criticises a theory, such as *Galileo's falling bodies*.[18] While the first two are "valuable" uses, Popper warns against using a thought experiment in an apologetic and defensive way. Heisenberg's microscope, used by Heisenberg to argue for the uncertainty principle using classical optics, is an example of an apologetic thought experiment.

Popper (1959) focuses mainly on thought experiments in quantum physics in his chapter on imaginary experiments, and his types are very much based around the function of theory choice. Critics of Popper's taxonomy include Krimsky (1973), who argues that Popper's 'apologetic' category does not have to be a misuse of a critical thought experiment; for example, Popper

---

[18]Popper's examples italicised.

states that any idealisations used in a thought experiment in defence of a theory should be acceptable to the opponent, while Krimsky argues that this smells of subjectivism. Popper also, in his description of apologetic thought experiments, does not allow a second theory to be used to rescue a first.[19] Krimsky disagrees, and points to the history of physics to show that theories are not self-contained or self-consistent.

In the more recent taxonomy by Brown (2011), he first divides thought experiments into two groups; constructive and destructive. Like Popper's, this taxonomy mainly covers theory choice rather than other functions of thought experiments. There are two main categories, which then split into smaller subcategories:

- **Constructive**

    - Direct: These start with unproblematic phenomena and end with a theory.[20]

    - Conjectural: These start with thought-experimental phenomena and end with a new theory to support them.[21]

    - Meditative: These start with a previously held theory and draw new conclusions from it.[22]

- **Destructive**

    - Ones that draw out contradictions in theories;

    - Ones that show conflicted beliefs;

    - Ones which undermine a central premise of an argument;

    - Counter thought experiments which deny phenomena of previous thought experiments.

This taxonomy has also faced criticism.[23] It seems that both Popper's and Brown's categories now fail to adequately capture both past and modern thought experiments in quantum mechanics.

---

[19]Popper (and Krimsky) use the example of Einstein's Photon Box, which was used to try and show that the uncertainty principle could be violated, and brought in Einstein's gravitational theory to 'rescue' quantum theory.

[20]An example Brown gives of this type is Einstein's Elevator.

[21]Brown's example is Newton's Bucket.

[22]Brown's example is Maxwell's thought experiment.

[23]Such as by Norton (1993), who argues that Brown's taxonomy is arbitrary.

There is debate on what exactly it is that thought experiments give us. Do they, like real-world experiments, bring us new knowledge? In the literature, there are two clear extreme views. On one peak, there is Norton and empiricism; on the other rests Brown and Platonism. In the trough between lie the many middle-ground views, such as the mental models view of Nersessian (1992) or experimentalism, as put forward by Sorensen (1992).

The Platonist view, championed by Brown (2014, 2011), argues that thought experiments allow us to directly access the laws of nature, which reside in a Platonic world. A good thought experiment gives us the ability to see into this world and learn from the laws of nature, from 'God's-eye-view' (Brown, 2014).

PLATONIC VIEW:

– Thought experiments allow us to access new knowledge of the laws of nature, just as real-world experiments do.

– The laws of nature reside in a Platonic world.

Main thesis: Thought experiments allow us to access the Platonic world of the laws of nature.

On the other extreme is Norton. Norton's staunch empiricist view has remained one of the main pillars of the debate for over two decades.[24] Looking for new knowledge by turning inwards to only thought is a hopeless endeavour, and thought experiments are not a special case. They are merely "ordinary argumentation disguised in vivid pictorial or narrative form" (Norton, 2004, p. 2). For Norton and empiricists, one can only gain knowledge via interaction with the natural world. As thought experiments do not interact with the natural world, they therefore cannot give us any new knowledge of it.

EMPIRICIST VIEW:

– Thought experiments do nothing more than transform pre-existing knowledge, as new knowledge can only be gained by interaction with the natural world.

---

[24]Empiricism others have understood roughly as the claims all knowledge comes from the senses of the natural world — there is no *a priori* knowledge.

– Arguments, both inductive and deductive, transform knowledge.

Main thesis: Thought experiments are arguments.

Norton concedes that they can transform what we already know — he states that the only way to do this is through both inductive and deductive argumentation. Therefore, in Norton's view, thought experiments are merely dressed-up arguments. Norton supports this by stating that all thought experiments can be reconstructed as arguments. If they are not arguments, he argues, then it is a worrying coincidence that they are so similar to them, as one would not want to state that arguments can be a channel to new knowledge (Norton, 2004). Norton also states that thought experiments can come in thought experiment and anti-thought experiment pairs (TE and $\overline{\text{TE}}$) — where one thought experiment supports one result and another thought experiment supports the opposite result. One or both must be false, or include a fallacious step. Norton asks how thought experiments can be treated as reliable when these pairs exist. Any epistemology of thought experiments must explain how to decide between these two thought experiments how to identify the false one — essentially, how to work out the reliability of thought experiments.

There are many middle-ground positions on thought experiments between these two extremes, and in this no-man's-land between Norton and Brown there lie many appealing views. The Mental Models view by Nersessian (1992), for example, uses mental modelling in cognitive science to back its claims, and thought experiments are described as mental simulations which are grounded in experience. Experimentalism argues that, as thought experiments mimic real world experiments — which are the only epistemic channel to the world for empiricists — this provides them with some kind of epistemic power (Sorensen, 1992). Norton, however, believes that middle-ground views of these kinds are unsatisfactory, and that one must either commit to either his argument or Brown's. This is due to Norton's argument that there is no mysterious 'X-factor' that thought experiments possess, and arguments do not.

In his effort to reduce thought experiments to merely 'picturesque arguments', Norton explores the various different candidates for the so-called 'X-factor' that transforms a thought experiment from an argument to something more — all of which he inevitably disregards. Evaluating other

stances on thought experiments, Norton concludes that none show the X-factor necessary for thought experiments to transcend empiricism. In all cases, he has four reasons why the view fails in the face of his empiricism:

a) Denial – no X factor.

b) Incorporation – the proposed X factor is also found in arguments.

c) Epistemic Irrelevance – the X factor has no bearing on the epistemic power of thought experiments.

d) Unreliability – the X factor cannot be used reliably by the thought experiment.

The Unreliability condition is dependent on first accepting Norton's Reliability thesis, which states that for thought experiments to be reliable epistemically, then they must be arguments that justify their conclusions, or able to be reconstructed as such arguments (Norton, 2004). This is a question which many ask of thought experiments — must they be reliable? A startling number of thought experiments in physics are simply plain wrong, or have since been discovered to be wrong, despite their leading to what is considered the correct theory. A key example of such thought experiments is Einstein's 'chasing the light' thought experiment (Norton, 2013). Thought experiments which have led to some of the most validated theories of the 20$^{\text{th}}$ century have not been reliable — but they have enabled the conceptual reshuffling necessary to form these new theories. Is it perhaps the process of the thought experiment rather than the actual outcome that has the most impact, and if so, is reliability a necessary component of a thought experiment?

The others can be seen as arguments against both Brown's platonic view (which falls foul of b) and d) in Norton's opinion) and also any middle-ground views, including Mental Models and Experimentalism. Any 'new' knowledge gained from the thought experiment is just transformed knowledge that could have been received from the reconstructed thought experiment in argument form. Norton's most common reply to middle-ground views is a) Denial — as not all thought experiments share that particular X-factor, then it cannot be a necessity and therefore cannot elevate thought experiments from mere arguments.

It is necessary to look more closely at Norton's 'denial' response, as he uses this to quickly dismiss any other properties of thought experiments, simply by pointing out that they are not universal features of thought experiments and therefore cannot be this magical X-factor. The Constructivist view by Gendler (1998) argues that thought experiments reveal problems in theories, and allow change and reform by teaching scientists about their 'mental apparatus' – something that arguments cannot do. Norton brands this with Denial — not all thought experiments perform this function. Also accused of this is the version of Experimentalism by Laymon (1991), as Norton points out that not all thought experiments idealise limiting cases of real-world experiments.

And what about the Elimination thesis from Norton (1991) — that thought experiments can in principle be replaced, throughout history, with plain arguments? Even Norton himself says that this would be very difficult, as thought experiments are so dependent on the context of the time and are simple where a purely logical argument may be unintuitive and 'cumbersome'. The "vehicle of a thought experiment might make easier the introduction of certain philosophical principles or facilitate certain inductive moves" (Norton, 1991, p .131)

It is precisely for this reason that Norton's Elimination thesis cannot be correct. Thought experiments are contextual; they are designed and shaped particularly for the problems of the time. They are narratives crafted to deliver, in a palatable form, the connections between dense and sometimes indecipherable theory. Is the act of imagining the scenario the 'X-factor'? When presented with such a vivid idea as 'a cat in a box, both dead and alive', one cannot help but try and imagine it. Thought experiments are often narratives which involve concrete, everyday objects (e.g. Galileo's ship, Einstein's elevator, Schrödinger's cat), although this can vary depending on your version of everyday (Kösem and Özdemir, 2014). Their narrative structure is a tool to enhance understanding, and connects abstract theory with what we see in the actual world. To help us with these connections, thought experiments can therefore fulfil a multitude of different purposes. They can act as theory litmus tests (Bokulich, 2001; Laymon, 1991). They can help us increase our general understanding through using our imaginations and more easily travel from the concrete to the abstract (Cartwright, 2010; Stuart, 2015). They can also be experiments in possible worlds, the chance to perform an experiment that could not be done in the actual world

— a form of analogue experiment (Hopp, 2014). All of this, of course, does not mean that they provide some sort of window into the abstract realm of the laws of nature. While they may help some knowledge become more epistemically accessible, and help us choose between theories, this does not grant them the magical epistemic power that Norton disdains.

Does their supposed 'X-factor' come from the fact that thought experiments are tuned to how our brains think — to a narrative, or a story? Reconstructing thought experiments as arguments is a worthy occupation; it can help to formalise thought experiments and strengthen their reliability. But we must not forget that thought experiments did not start as organised lines of logical argument — and that is perhaps precisely where their 'X-factor' lies. There are various arguments for why this act of imagination is so much more powerful than strings of logic. Hopp (2014) uses phenomenology to describe how one can 'behold properties' through imagination and hallucination; while they may not allow us to see actual properties in the actual world, but we are using our prior knowledge of the actual world to view properties in possible worlds.

While there may be problems with Hopp's account — not least of which that thought experiments very rarely deal with abstracts, but rather with concrete objects that we must then abstract from to find the 'moral' of the story (Cartwright, 2010) — it is nevertheless an interesting enquiry, and the literature of thought experiments is starting to become more involved with phenomenology.

More recently, Stuart (2015, 2016) has put forward his own interpretation of thought experiments in science — a kind of Structuralist viewpoint according to which thought experiments (or at least some good thought experiments) can "increase the empirical content of theoretical structures... for an agent" (Stuart, 2015, p. 1). Stuart argues for the importance of matching theoretical structures with empirical content, citing Einstein, Mach, and Heisenberg, and that thought experiments are the vehicle through which this happens. It is achieved via "imaginative use of idealised models." Stuart argues that empirical content can only enable understanding up to a threshold. Thought experiments as mental models help increase understanding beyond empirical content, marrying that content with the theoretical structure. Stuart (2015) talks of how a thought experiment is a performance, a set of cognitive actions which helps forge a connection between theoretical structure and empirical data, or experience. Whether or not

thought experiments are reliable, or successful, they do not need to be either in order to be useful in a way other than mere argumentation. Thought experiments can be used as a first step — a jumping off point in constructing a real world experiment in order to increase empirical data of a theory. With Stuart's view of thought experiments as adding to the structure of the theory, thought experiments can be described as the first step towards a real world experiment.

Thought experiments act as a launch pad, in essence — the necessary performance of imagination so that the abstract concepts are more easily grasped and understood. In a recent study, it was found that students of all levels of physics knowledge utilised thought experiments in order to solve conceptual physics problems (Kösem and Özdemir, 2014). The three groups of students — A-level students, second year undergraduate students and PhD students — were all given the same problems, and talked through how they solved them. It was found that while the PhD students often used thought experiments to explain, and used scientific concepts, the two lower level groups used simple cases in their thought experiments and everyday concepts in an effort to conceptualise the problem. It seems, from these results at least, that thought experiments can be engineered to connect our everyday lives and experiences with those that we struggle to understand — and in doing so, help us to better grasp those abstract concepts.

Kösem and Özdemir quote Rescher (2005) in their paper;

> "...all [a thought experiment] can offer us is a new way of conceptualising and reinterpreting the old..." (Rescher, 2005, p. 32)

This may not be the peek into the abstract laws of nature which Brown wants, but it is nevertheless far from being simply a picturesque argument. While no new empirical data about the actual world is found from running a thought experiment — and it can also be argued that this is true also for computational models and simulations — it nevertheless adds to the links between theory and the real world.

Stuart tries to weaken Norton's empiricist view by arguing that, by Norton's standards, no thought experiment can be justified by its logical form (Stuart, 2016). In Norton's argument, a good thought experiment is good because of certain properties in its form of a logical argument.

Stuart refutes this, claiming that a thought experiment cannot be justified "in any deep sense" due to these properties, and should be evaluated as an experiment, rather than logical schema.

Stuart argues that when thought experiments are reconstructed as logical arguments, while Norton's view allows that thought experiments can provide transformed knowledge, which can be argued to be new knowledge, this cannot be from the deductive arguments in the thought experiment. These merely arrange what we already know, and it cannot be from the inductive arguments either, as these are justified by their relation to real world properties, and not their logical structure.

It is hard to argue that thought experiments do not provide new knowledge — Norton claims that this is instead transformed knowledge, from the logical arguments hidden by the picturesque coating of thought experiment. Stuart argues that Norton cannot claim this due to his empiricist standards in regards to deductive and inductive arguments — neither of which can transform knowledge like a thought experiment can. Instead, Stuart believes that thought experiments can be evaluated like real world experiments, as while thought experiments seem to be entrenched in scientific work and progress, it is still unclear what exact role they play. Putting forward his material theory of thought experiments, Stuart claims that the true justification of a thought experiment lies in the same thing that makes real world experiments reliable. Using the five criteria for a good experiment from Franklin (1989), Stuart then applies these to thought experiments.

Real world experiment	Thought experiment
a well isolated system	isolated in the imagination, using idealisation and abstraction
no experimental bias	no cognitive and confirmation bias
sources of error known and accounted for	need to be coherent with all our current knowledge
calibrated instruments	well-calibrated imaginations, able to deal with abstract situations
a theory of instruments	a theory of inference-making

Stuart does not ignore the obvious differences between real world experiments and thought experiments. One main objection that Stuart considers is that a thought experiment does not directly intervene in extra-mental systems, something which is necessary in a real world experiment. In reply, Stuart writes that intervention is not a necessary condition of an experiment.

While Stuart is still in favour of reconstructing thought experiments in order to tell the good inferences from the bad, he argues that something extra is needed in an epistemology of thought experiments in order to know what properties justify them. Stuart's "fount" of justification lies in what makes the thought experiment a good experiment, just as one would evaluate a real world experiment. This, he states, could be used by empiricists, naturalists and rationalists, for their epistemology of thought experiments. Analogue experiments could bridge the gap between thought experiments and real-world experiments, and for Stuart they could therefore highlight the difference between thought experiments and arguments.

Some thought experiments may never become a real world experiment in their entirety, such as Einstein's elevator. No one will ever ride a Willy Wonka–esque lift into space to test the equivalence principle. Analogue experiments can be done in its place, using the equivalence principle illuminated by the elevator thought experiment. This idea can be captured with an extra category, extending Popper's original taxonomy:

1. HEURISTIC — thought experiments which help to illuminate/clarify a theory.

2. CRITICAL — thought experiments which criticise or point out contradictions in a theory.

3. APOLOGETIC — thought experiments which back a certain theory.

4. PLACEHOLDER — thought experiments which act as placeholders for as-yet impossible experiments.

This use of thought experiments, as a placeholder or a stepping stone to potential real-world experiments, is just as valuable as one used for illustrative or explanatory purposes. A thought experiment need not be restricted to just one use either, but can have multiple uses at once, dependent on perspective, or just throughout its lifetime. Placeholder thought experiments may start off life in the position of an as yet untenable real-world experiment, but as technology progresses, may then take on a heuristic role. Thought experiments used at first to criticise one theory could later be used to support and illustrate another. As long as one heeds Popper's warning — that to mould a thought experiment completely to your own purposes is to render it useless — this fluid taxonomy need not cause problems.

Mach's support for thought experiments as laying the groundwork for real world experiments supports the view that some thought experiments provide a placeholder position, and within the realm of the thought experiment leaps of intellect can be made prior to being tested in the real world. In fact, he writes of them being used both for critical evaluation of a theory and as a technique of discovery. Are they a technique of discovery of their own, or are they tied inexorably to real world experiments? Or are they two types of the same thing, under the common umbrella of experimentation?

It could argued that thought experiments leading to real-world experiments are simply heuristic thought experiments — they illuminate a theory and increase understanding, so that progress can be made. However, thought experiments such as quantum tic tac toe and other games are heuristic — they are teaching devices. This is a very different purpose to that of the Elitzur-Vaidman thought experiment, for example, which is constructed in the same way as a real-world experiment. This link between the theory and the natural world is the link that Stuart

wishes to highlight when he argues that we should evaluate thought experiments as experiments, not as arguments.

If thought experiments are like real-world experiments in that they can bring us new knowledge by having some kind of epistemic power, then Norton and his empiricism is wrong. Thought experiments and real-world experiments can be connected by the bridge of analogue experiments. Analogue experiments and Placeholder thought experiments share a similarity in being the only access we have to the actual system we wish to gain knowledge about. Recent work into black holes has been done instead in 'black hole laboratories' on Earth by Torres et al. (2016, 2017), for example, in apparatus that looks not unlike a large bath. It might seem ridiculous to empiricists like Norton that we could possibly gain knowledge about a system by merely imagining an experiment; it seems equally ridiculous that we could possibly learn about black holes by using a plughole in a bath. Nevertheless, these experiments take place using 'dumb holes' — fluid dynamical systems which are purported to be analogues to the target system of a black hole. As explored in more detail by Dardashti et al. (2015), it is argued that analogue experiments are a new type of scientific inference 'with the potential to be confirmatory.'

Systems like black holes cannot be measured directly, and it is unlikely that this will change in the next few decades. While effects of black holes, such as gravitational waves, can be observed using telescopes, it is currently impossible to examine in more detail other phenomena, like Hawking Radiation. Work has been done by Dardashti et al. (2015), however, to create analogue systems with a 'robust *syntactic isomorphism*[25] between the relevant modelling frameworks' in order to carry out experimentation on the phenomena, with real empirical results. Using this technique, phenomena such as superradiance has been recorded in water vortexes by Torres et al. (2016, 2017) — believed to be an analogue of black hole dynamics.

When real world experimentation is impossible, the only way forward is to work with what is possible. Here, it can be seen that using models to conceptualise phenomena and turn them from abstract or distant to concrete and everyday, experimentation can still be done without the original phenomena. It may not be as strong as real empirical data from the real system, but that is beyond our reach.

---

[25]Their emphasis.

The process of going from black hole to dumb hole — of going from real world phenomena of black holes to the abstract theory to concrete, everyday phenomena of a water vortex in an emptying bathtub, is incredibly similar to the process of thought experiments. The main difference with analogue experiments is that they do produce real, empirical data — but only from the analogue system. Nevertheless is it argued by Dardashti et al. (2015) that this data can go some way towards gaining knowledge about the target system. Thought experiments can be described as the starting point for these analogue experiments; the imagining of a hard-to-grasp concept by using more easily understandable phenomena.

(2.15)   TE → computer simulation → analogue experiment → real world experiment

There is, of course, one crucial difference. Analogue experiments are crafted so that they are possible to perform. One can physically fill a tub with water and observe the vortex caused by the water draining out. Einstein, with the best technology of his day (and perhaps even with the best technology of our day) could not accelerate a lift in a vacuum, away from gravitational effects, in order to demonstrate the equivalence principle. Galileo would have seen the cannonball and musket ball hit the ground at different times due to air resistance and human error. Thought experiments are unlimited, in that one can perform whatever experiment one chooses with any set-up — vacuum or no vacuum, empty space, perfect spheres, friction-less surfaces, and so on. Analogue experiments are tied down to the real world, or at least to the technology of the time in which they are performed. If thought experiments are taken to be narratives of experiments which transform abstract concepts to everyday objects and phenomena, then analogue experiments are the subset of possible-in-the-real-world thought experiments.

This satisfies Stuart's arguments against Norton's thesis — placeholder thought experiments are not merely arguments, but can be linked to real-world experiments by analogue experiments.

Though not exhaustive, this section has detailed a history of the literature of thought experiments in philosophy of science. The debate over what thought experiments give us is still ongoing. The famous attempts at taxonomies have been unsuccessful, though the addition of the placeholder category has been shown to ameliorate this, and the new category also captures an

important role played by many thought experiments. The previous section explored the development of quantum theory by reviewing chronologically the thought experiments used by physicists. Now, taking what has been learnt, the next two sections explore what these quantum mechanical thought experiments can teach us about thought experiments in general, and also what they can teach us about quantum theory itself.

2.4   What do Quantum Mechanical Thought Experiments teach us about Thought Experiments in general?

So what can we learn from our previous exploration of quantum mechanical thought experiments? It seems that they are not restricted to only one category in a given taxonomy but can move between them as science progresses, showing their evolving nature as a device of inquiry. For example, one thought experiment may be first used in a illustrative way but later used to 'apologise' for a theory. Schrödinger's cat was created to criticise quantum theory but is now mainly used heuristically. Popper's taxonomy, despite being built around thought experiments in quantum physics, is not enough to categorise and understand modern quantum mechanical thought experiments. Brown's later taxonomy also cannot achieve this goal. It is obvious that some thought experiments do belong in more than one category in their lifetimes, showing their evolving nature as a device of inquiry, and perhaps their reliability. It has also become apparent that there could be a need for a fourth category in these taxonomies, one which is important for thought experiments as being link to real-world experiments in accordance with Stuart's view, and which could be used as an argument against the empiricist view of thought experiments.

There seems to be an important role for thought experiments which is not included in Popper's taxonomy or in Brown's — thought experiments which are used in the place of impossible (even if only for the time) experiments. This includes whether they could perhaps be performed in the future or are simply beyond human capability entirely. These thought experiments act as a placeholder for potential experiments, or simply as analogues of impossible ones. They also act as a bridge between theory and the natural world, a tool that can be used in and tested by real-world experiments, but that had originated from thought rather than experimentation. This

is similar to Stuart's view of thought experiments as a jumping-off point, or first step, towards real world experiments. Having such a use does not exclude a thought experiment from belonging to any of Popper's types, either; EPR experiments were beyond the physics of the time, but are now used often in quantum experiments and research.

The majority of the thought experiments in quantum mechanics talked about here have been able to be put in this new fourth category; Wheeler's delayed choice, EPR, Schrödinger's cat (including Wigner's friend), Elitzur-Vaidman, Bell's inequalities, LGI, and PR box thought experiments can all be placed in this fourth category. In particular, the ones such as Bell's and LGI would be very difficult to reconstruct as arguments, as they are, essentially, tools, created in thought from theory, in order to apply to the natural world.

Recall Norton's empiricist view as summarised above;

EMPIRICIST VIEW:

1. Thought experiments do nothing more than transform pre-existing knowledge, as new knowledge can only be gained by interaction with the natural world.

2. Arguments, both inductive and deductive, transform knowledge.

3. Thought experiments are arguments.

In regards to the first point — it can be argued that thought experiments create a link or a bridge between theory and the natural world, which is more than transforming pre-existing knowledge. Bell's inequalities are used in real-world experiments, but did not come from experiments. Bell worked out the inequalities from a theorem. The third claim can now be called into question, as while these could perhaps be reconstructed into arguments, their main role is to provide a link between the theory and the natural world, and the argument in itself is not the important part. It may be possible still to reconstruct all these thought experiments into arguments, but this does not take away the fact that they cannot be reduced to merely arguments and describing them as such is to reduce them too far — their status as bridging the gap between theory and the empirical data is innate to their theoretical conception and experimental form.

Do we need a taxonomy of thought experiments? Brown argues that since thought experiments can evolve and change with time, making their significance depend on certain historical facts,

then it is important to know what each thought experiment was meant to accomplish. Krimsky also takes this line of argument, stating that thought experiments should be judged from their purpose. But these modern quantum mechanical thought experiments align much more closely with the new, pluralist views of thought experiments as seen in the current literature. There has been a significant shift towards looking for an identity of thought experiments as a whole, rather than a taxonomy. Sorensen (1998) argues that thought experiments are limiting cases of real world experiments, and should be studied as if they were real experiments. Buzzoni (2012) argues that empirical thought experiments are conceivable as preparing and anticipating real experiments — as can be seen from the review above, this has definitely been the case in recent quantum theory. This view has only been strengthened by the work done by Chandrasekharan, Nersessian, and Subramanian, and Lenhard, in comparing thought experiments to simulation experiments.[26] McComb (2012) uses the family resemblance concept from Wittgenstein and lists the relevant resemblances for thought experiments, including 'involving our own activity'— something in common with a real world experiment. Bokulich and Frappier (2017) discuss the evolution of thought experiments over time — something that is clearly seen in the examples above — and about whether the thought experiment evolves or becomes a different thought experiment entirely. The clear evolution of the examples above also shows their similarity with real world experiments.

From examining these thought experiments, it seems clear that a new fourth category, a Placeholder category, can be added into thought experiment taxonomies. It is also apparent that quantum mechanical thought experiments support Stuart's view that thought experiments should be evaluated as experiments, not as arguments, and that Norton's view that thought experiments are merely picturesque arguments does not give thought experiments enough credit. There is lots of scope for future work to look more closely at their use as a link between theory and concrete models, and at how this could affect the empiricist versus rationalist debate as a whole.

As our scientific knowledge has increased over the past few decades, thought experiments in

---

[26]Chandrasekharan et al. (2012) talk of thought experiments being replaced in certain cases by simulation experiments, and similarly Lenhard (2017) states that simulation experiments are the newborn sibling of thought experiments.

quantum mechanics have become more technical and less picturesque and intuitive. This shift is mirrored in the literature on the view of general thought experiments, as instead of searching for a taxonomy for them as their own category, they are now more commonly considered a sub-type of real world experiments. Though these two things do not appear to be linked specifically, quantum mechanical thought experiments followed the general trend of thought experiments becoming more technical, as seen by the rise of simulation experiments and computational modelling — it is even argued that such experiments are replacing thought experiments in modern science (Chandrasekharan et al., 2012).

This new, common view of thought experiments as being some kind of subset of, or being in continuum with, real world experiments, and with simulation experiments also belonging in this group, perfectly fits the thought experiments discussed above. As thought experiments have become more technical and less intuitive over the years, their similarity to experiments has become clearer. This view can account in some way for these more modern quantum mechanical thought experiments being about measurement. It can also support their evolution through the years. Schrödinger's cat is a well-known and well-used example, as are many of Einstein's thought experiments. Interestingly, however, the more modern examples — and arguably the more influential ones in the current theory, such as PR boxes — are rarely used, if at all, in the wider literature. In the discussion of thought experiments in quantum mechanics by Peacock (2016), he does not discuss in detail any thought experiments much more recent than Bell's theorem and Wheeler's interferometer. More recent thought experiments such as PR boxes, as ground-breaking as they have been in the quantum world, have not been touched. Despite being perfect examples of why we should not be categorising thought experiments, but instead treating them as existing on the same spectrum of real world experiments, their potential has remained untapped. Why this is, however, is not clear. Perhaps these recent thought experiments are too removed from the classic ideal of a thought experiment as an intuitive, imaginative scenario, to be easily used as an example. Regardless, they suit the new view well, and show that thought experiments are still crucial even in a technical science such as quantum theory.

Kösem and Özdemir (2014) state that one of the most cited definitions of a thought experiment comes from Sorensen (1992); "a limiting case of real experiments performed by mentally

stimulating the course of events in the mind" (Kösem and Özdemir, 2014, p. 867-868). Some thought experiments, however, are not a limiting case, but involve physics that could never exist in our world. PR box correlations, for example, are superquantum correlations which could never be possible in the real world. In that particular thought experiment, they are used as an example of a theory which can be derived from the same two axioms that quantum mechanics can be derived. This leads to the question of what third axiom is necessary to restrict theories to simply quantum mechanics — the one we know to be true. By learning what quantum mechanics is not, and what postulates it specifically needs, teaches us about quantum mechanics without interacting with the real world. In fact, what we learn, we could not have learnt from the real world; superquantum correlations do not exist in the world, and we cannot do experiments on them. All we can know from experiment is that there are no superquantum correlations. If we want to know what axioms restrict these superquantum correlations to the normal quantum correlations, we must work exclusively in the imagination.

After applying what was learnt in section 2.2's review of quantum mechanical thought experiments to the debate on thought experiments in general philosophy of science, this section tells us that while thought experiments can be thought of as arguments, that cannot be their only role; they also play very important roles as placeholders for experiments that are currently impossible (or potentially always impossible), much like analogue experiments. They cannot be easily taxonomised as they evolve over years of use in a theory, though allowing category change and using a Placeholder category could help with this. Following Stuart (2015, 2016), it is clear that thought experiments are on the same spectrum of real world experiments, as is seen clearly by the necessity of the Placeholder category, and are indispensable in areas such as quantum physics where certain experiments are impossible to perform except in the imagination. Norton's view that they are nothing more than picturesque arguments cannot be true – thought experiments are too much like real-world experiments to only be considered as arguments.

2.5 What do we learn about Quantum Mechanics from Thought Experiments?

As thought experiments are very prominent in quantum mechanics, and so much of the theory was or is still beyond the capabilities of real-world experiments, it is important to review these thought experiments and look closely at how they are used. Thought experiments in quantum mechanics remain alive and well, and seem to be an untapped resource for those discussing the identity of thought experiments in the general philosophy of science. It can be argued that thought experiments have always been close to real world experiments — in the modern quantum mechanical thought experiments, some of which are stripped of picturesque scenery, this is just clearer to see.

Previous famous thought experiments worked on intuition. It is easier to imagine standing on Galileo's ship and watching the shore move, yet watching butterflies flutter around below-decks, unaffected by the motion of the ship. Even imagining accelerating upwards in Einstein's elevator is more natural for the mind than even the most basic of quantum mechanical thought experiments such as Schrödinger's cat. Quantum mechanical thought experiments seem to be more akin to real-world experiments than others, as they revolve around the process of measurement. They are still idealisations — isolating labs as quantum systems with macrosystems such as humans inside is, as of yet, impossible — hence the extended Wigner's friend real-world experiment could not replicate the thought experiment exactly and involve the universality assumption. Many thought experiments involve situations that decoherence would make impossible in practice, making them the ideal breeding grounds for new developments and experiments in quantum theory. But these kinds of thought experiments will always be ill-fitting in any thought experiment taxonomies because they (at least currently) cannot provide an intuitive description, clarification, or criticism about measurement.[27] Without putting an interpretation onto the theory, measurement in quantum mechanics is a black box. These thought experiments give results and mainly revolve around the process of measurement, quite literally experimenting with the process of

---

[27]Or, at least, they cannot do so without using an interpretation of quantum theory, rather than quantum theory alone.

measurement in a more easily accessible environment than the real world.[28]

As we have seen, not all of the thought experiments above fit comfortably in Popper's and Brown's categories. Some belong to multiple categories — Schrödinger's cat could have been in all three categories at various points in its lifetime. Some, such as PR boxes, don't really belong in any of Popper's three categories. Some, such as the Elitzur-Vaidman Bomb tester and Hardy's paradox, have since been implemented as real-world experiments.

One reason that they do not fit comfortably into a taxonomy is because this ignores one very important aspect of their construction; they are all about measurement, which is itself an instance of the quantum to classical transition. All the thought experiments discussed above deal either directly or indirectly with the quantum wave function collapsing into a definite measurement outcome. This again highlights a point made in the previous section – thought experiments should be considered as on a spectrum with real world experiments. The new category of 'placeholder' fits these thought experiments much better, as it takes into account their focus on measurement by taking them as future experiments, as yet (or forever) impossible.

The measurement problem has been a main subject of discussion for physicists and philosophers since the theory's conception, as seen previously in sections 1.3 and 1.4, in the discussion on the Copenhagen interpretations. And while the debate has led to the creation of many different interpretations of quantum theory, there is still no consensus.

Wheeler's delayed choice thought experiment dealt with the issues around choosing a measurement basis. Schrödinger's cat and Wigner's friend inquired into the problem of entanglement and measurement collapse. The Renninger negative result experiment also looked at the unintuitive way wave function collapse works in quantum theory. The Elitzur-Vaidman Bomb tester looked at the potential of 'interaction-free' measurement — essentially looking at the possibility of 'collapse-free' measurement. Hardy's paradox also deals with interaction without consequences — without collapse. PR boxes explores the connections between the axioms of quantum theory and the strange correlations in its measurements. The last two no-go theorems, descendants of Wigner's friend, explore how quantum theory describes itself in regard to measurements of different observers. Other thought experiments, such as EPR and Bell's inequalities, bring out

---

[28]Results, in this case, not meaning data such as one would get from a real-world experiment, but for example like the PR box result — a conclusion drawn about quantum theory that had not been known previously.

conceptually interesting aspects of quantum theory other than the measurement problem such as entanglement and superposition, which exist only in the quantum microworld and not the macroworld. Quantum mechanical thought experiments are focused on the differences between the quantum realm and the classical realm, and the transition between the two. Real world experiments focusing on this area are much harder to perform, and sometimes (in the case of PR boxes) are impossible. Decoherence, which is discussed later in more detail in chapter 4, prevents phenomena like entanglement and superposition from being seen in the everyday macroworld. It is difficult today to prevent decoherence long enough to perform experiments on quantum systems, and in the early days of quantum theory it was practically impossible. Quantum mechanical thought experiments were the way in which physicists could explore quantum phenomena before it was possible to do so in the lab.

This highlights the importance of thought experiments in quantum theory. The transition between quantum theory and classical theory, measurement, and problems with measurement, both conceptual and physical, are the root of many issues in quantum theory. As Peacock states,

> "...quantum mechanics will not permit us to ignore... that it is impossible to observe and measure the properties of a particle without physically interacting with it." (Peacock, 2016, p. 16)

Measurement, in quantum theory, is a black box. There are interpretations of what happens before wave function collapse, but these are added onto the theory rather than in the theory itself. Collapse limits what can be done in the lab, as quantum systems must be isolated or they decohere and become classical.[29] In thought experiments, it is much easier to isolate a lab — one simply states it, and uses the relevant calculations. Nowhere is this divide between thought experiments and real-world experiments in quantum mechanics more apparent than in Proietti et al. (2019)'s recent experiment, designed to produce results like Brukner's no-go theorem.

Proietti et al. (2019) devised and ran the experiment in order to reject the assumption of observer-independent facts in the quantum world by using a three source, six photon set-up. Using the extended Wigner's friend thought experiment where Wigner knows that the friend has

---

[29]Decoherence itself is one key to this discussion of measurement and the quantum – classical cut, which is explored in more detail in chapters 4 and 5.

made a direct observation, but not what that observation was, they created an experiment with four entangled photons. The friend's measurement consists of a second pair of entangled photons which, through interaction with the original pair, then record the polarisation measurement non-destructively. This information is stored in the polarisation state of the second photon, which acts as the friend's record or memory. The results of this experiment were as follows; the probability distribution of all four "facts" — the two friends and the two Wigners, being the two original entangled photons and the two photons with the recorded measurement in their polarisation state — violates the CHSH inequality. This indicates that observer-independent facts cannot exist in a version of quantum theory that has the three following assumptions;

- Locality

- Freedom of choice

- Observer-independent facts of the world.

This experiment — despite being subject to a few potential loopholes — is an attempt at a real-world reproduction of Brukner's thought experiment/no-go theorem, and is similar to Frauchiger and Renner's.[30] However it is dubious whether the photons used as the 'Wigner's friend' in this experiment can take the place of a human observer, or even just a macroscopic non-human observer such as a detector or a computer. Facts about photons are not the same as human or macroscopic observer-independent facts. Losing the universality assumption — one of the key parts of Brukner's no-go theorem — means that this experiment does not accomplish the same results of the thought experiment it is attempting to recreate. Photons are very different to macrosystems — the photons in this experiment have not yet collapsed and decohered, whereas it is still unknown whether an open thermodynamic system, such as a human or a macroscopic measurement device, could be in a superposition of states. This is the very question the Schrödinger's cat thought experiment poses, and it is still unanswered. Currently, only in thought experiments can we isolate Wigner's friend in their lab.

So what do we learn about quantum mechanics from quantum mechanical thought experiments? Throughout the development of quantum mechanics, thought experiments have been

---

[30]This is explicated in their discussion section (Proietti et al., 2019).

used to probe the implications and limitations of the theory. As we have seen, the core underlying problem that runs through almost all of them, from the birth of quantum theory to the modern day, is measurement and reconciling what quantum theory tells us to what we actually see in the world – the transition between quantum and classical. In the next chapter, the measurement problem is more clearly defined, and different solutions to the problem, offered by different interpretations, are explored.

3

THE MEASUREMENT PROBLEM AND INTERPRETATIONS I

3.1   The Measurement Problem

An exploration of the development of quantum theory over the years through thought experiments makes it clear that the key foundational issue is measurement and the transition from quantum state to classical result. In quantum theory measurement is a black box.

The measurement problem is a quandary that has occupied quantum physicists since the very beginning. The standard formulation of it, as discussed briefly in chapter 1, is that the following three claims of quantum mechanics are inconsistent:[1]

1. Quantum theory is complete.

2. The wave function always evolves in accordance with unitary time evolution, as described by the Schrödinger equation.

3. Measurements have unique outcomes.

As detailed by Schlosshauer (2005), the measurement problem can actually divided into two distinct problems: the problem of definite outcomes, and the problem of the preferred basis.

---

[1]From Maudlin (1995).

Together these cover the issue of how the quantum world gives us apparently classical measurement results. The first two sections of this chapter look at what these two problems are. The following sections explore five different interpretations of quantum mechanics: Bohmian mechanics, Collapse theories, modal interpretations, Oxonian Everettiansim, and Qbism.[2] They explain how these interpretations attempt to solve the the general measurement problem, and which of the two measurement problems they are specifically addressing.

Before exploring the more detailed differences between interpretations, it is important to mention that they can be split into two main schools of thought on the ontological status of the wave function. There is a Ψ-ontic interpretation, where the quantum state has a real, objective status in the world and is an intrinsic property of the system being observed. Then there is a Ψ-epistemic interpretation, where the state represents only knowledge or information about the observed system and therefore has no real, objective status. Harrigan and Spekkens (2010) claim to show that the state can be either ontic or epistemic, but cannot be both. Since Harrigan and Spekkens (2010) set up the definitions of epistemic and ontic so that they are mutually exclusive, along with most of the literature, this is true. Most of the interpretations below are Ψ-ontic, though the last one considered, QBism, is Ψ-epistemic. The reason most of the interpretations explored below are ontic is that epistemic interpretations solve the measurement problem in a very radical way that arguably conflicts with the way physics works — this will become clear when looking at QBism in detail.

3.2 The Problem of Unique Outcomes

This part of the measurement problem arises from quantum superposition. As seen in chapter 1, the superposition principle states that the linear combination of any two or more quantum states is itself a valid quantum state. This is expressed in the linearity of the vector space. Hence we have a quantum system that can be prepared in one of two states, $|0\rangle_S$ or $|1\rangle_S$, of some operator $\hat{O}$ corresponding to observable 0, then when these states interact with the apparatus used to measure the system, the apparatus simply gives the results of 0 or 1 respectively;

---

[2]This particular flavour of the Many Worlds interpretation will be focused on in this work. This is the interpretation begun by Saunders (1993) and championed by Wallace (2010) while they were both at Oxford, and is therefore labelled Oxonian Everettianism.

$$|0\rangle_S |ready\rangle_A \longrightarrow |0\rangle_S |\text{``0''}\rangle_A \,,$$

(3.1)

$$|1\rangle_S |ready\rangle_A \longrightarrow |1\rangle_S |\text{``1''}\rangle_A \,.$$

However, if the quantum system is in a superposition of states of what is being measured,

(3.2) $$\psi(0)_S = a |0\rangle_S + b |1\rangle_S \,,$$

then if a and b are both nonzero, the unitary evolution of the Schrödinger equation implies that the product of the system superposition and the apparatus leads to the entangled system-apparatus,

(3.3) $$|\psi(0)\rangle_S |ready\rangle_A \longrightarrow a |0\rangle_S |\text{`}0'\rangle_A + b |1\rangle_S |\text{`}1'\rangle_A \,.$$

Now neither the apparatus nor the system are in a pure state no matter what bases we use for them. The system is not in an eigenstate of the observable we are measuring, and the apparatus is not in a definite state either. We obviously do not see the apparatus to be in a superposition; we see one measurement result. We do not see superpositions of macrostates and the eigenstate-eigenvalue link does not explain why we see unique outcomes rather than superpositions.

The problem of unique outcomes is generally just referred to in the literature as the measurement problem itself. However, the preferred basis problem is argued by Schlosshauer (2005) to be equally as important as the problem of unique outcomes.

3.3 The Preferred Basis Problem

The preferred basis problem, as described in Zurek (1981), also called basis degeneracy, is as follows:

When measuring observable $\hat{O}$ on a quantum system s with state vector $|\psi\rangle$ with apparatus A with state vector $|A\rangle$, unitary evolution is enough to establish a nonseparable correlation between the two state vectors.

$$|A_0\rangle \otimes |\psi\rangle = \left\{ \sum_s a_s |A_s\rangle \right\} \otimes \left\{ \sum_s c_s |s\rangle \right\}$$

(3.4)

$$\longrightarrow \sum_s c_s |A_s\rangle \otimes |s\rangle$$

where $c_s = \langle s|\psi\rangle$, and $|A_s\rangle$ is the device state associated with the eigenvalue s and the eigenstate $|s\rangle$ of $\hat{O}$.[3]

Here $|A_s\rangle$ and $|s\rangle$ are basis vectors for the apparatus and the system, with $|A_0\rangle$ the initial state of the apparatus. It seems here, from equation 3.4, that the observable $\hat{S} = \sum_s e_s |s\rangle \langle s|$ has been measured. But the state of the apparatus being used to measure the system can be written in a different orthonormal basis, $\{|A_r\rangle\}$, which is composed of a superposition of $|A_s\rangle$:

(3.5)
$$|A_r\rangle = \sum_s \langle A_s|A_r\rangle |s\rangle,$$

and then the combined system of the original quantum system and the apparatus can be expressed as

$$\sum_s c_s |A_s\rangle |s\rangle = \sum_r |A_r\rangle \sum_s c_s \langle A_r|A_s\rangle |s\rangle$$

(3.6)

$$= \sum_r d_r |A_r\rangle |r\rangle$$

So here $\{|r\rangle\}$ are normalised but generally not mutually orthogonal states of the original system, s, relative to the chosen basis set $\{|A_r\rangle\}$ of the apparatus. Does this mean that after the measurement, the quantum system will end up in one of the states $\{|r\rangle\}$ instead of $|s\rangle$? This is a particular problem when all the coefficients c_s have the same magnitude. Since the apparatus is correlated with the quantum system, it contains all of the information about the observable $\hat{S}$ and also all the other observables that are defined on the Hilbert space of the system (for example, the observable $\hat{R} = \sum_r e_r |r\rangle \langle r|$, even though generally $\hat{S}$ and $\hat{R}$ do not commute). Quantum mechanics does not allow us to simultaneously measure two non-commuting variables with arbitrary accuracy. Our experience tells us that the choice of observable that our apparatus measures cannot be made arbitrary after the system and apparatus have interacted

---

[3]As in Zurek (1981).

and correlated. Our apparatus measure the observable we make them to measure. There appears to be a unique basis set for the apparatus that records the corresponding relative states of the quantum system being measured, bases such as position or momentum, but how is this basis determined? Quantum mechanics by itself cannot answer this.

As quantum theory by itself cannot solve these two measurement problems, the problem of the preferred basis or the problem of unique outcomes, interpretations of quantum theory are employed to attempt to solve the measurement problem. This chapter looks at five different interpretations of quantum theory and which measurement problem they attempt to solve.

3.4 Bohmian Mechanics

Bohmian mechanics is the first and best known hidden-variable interpretation of quantum mechanics, started by Louis De Broglie (1928) and developed by Bohm (1952). The wave function in quantum mechanics in this interpretation is not complete; there are extra specifications not in the wave function that detail the actual positions of the particles described. The positions of the particles evolve via the 'guiding equation', and this fills in the information that is missing from the quantum wave function.

Bohmian mechanics takes quantum mechanics to be, rather than fundamentally about the behaviour of the wave function, fundamentally about the behaviour of particles. This behaviour is described by field configurations. A complete description of a quantum system, in this interpretation, is given by the configuration Q, defined by the positions of its particles, together with the wave function (Dürr et al., 1992, 1995).

In Bohmian mechanics, therefore, the complete state of the system for N particles is described by

$$(3.7) \qquad\qquad\qquad (Q, \psi),$$

where

$$(3.8) \qquad\qquad\qquad Q = (\mathbf{Q}_1, ..., \mathbf{Q}_N) \in \mathbb{R}^{3N}.$$

Here $\vec{Q}_1, ..., \vec{Q}_N$ are the positions of the particles and

$$(3.9) \qquad \psi = \psi(q) = \psi(\mathbf{q}_1, ..., \mathbf{q}_N)$$

is the wave function of the system. The initial state (Q_0, ψ_0) evolves into (Q_t, ψ_t) at time t, and this evolution is defined by first-order differential equations; the Schrödinger equation:

$$(3.10) \qquad i\hbar \frac{\partial \psi_t}{\partial t} = H\psi_t,$$

as the evolution equation for ψ, and for $\mathbf{Q}$,

$$(3.11) \qquad \frac{dQ_t}{dt} = v^{\psi_t}(Q_t).$$

Here,

$$(3.12) \qquad v^{\psi} = (\mathbf{v}_1^{\psi}, ..., \mathbf{v}_N^{\psi})$$

is a vector field on the configuration space $\mathbb{R}^{3N}$. The wave function in Bohmian mechanics has the job of generating the motion of particles. For a one-particle system,

$$(3.13) \qquad \mathbf{v}^{\psi} = \frac{\hbar}{m} \mathrm{Im} \frac{\nabla \psi}{\psi},$$

where ∇ is the configuration-space gradient and needed for rotation invariance, ψ the wave function is the denominator for homogeneity, Im from time-reversal invariance, and the constant in front is required by covariance under Galilean boosts. For a many-particle system,

$$(3.14) \qquad \mathbf{v}_k^{\psi} = \frac{\hbar}{m_k} \mathrm{Im} \frac{\nabla_k \psi}{\psi}.$$

In the limit of $\frac{\hbar}{m} \longrightarrow 0$, the motion Q_t approximates classical motion (Dürr et al., 1992).

Therefore, as Bohmian mechanics is a deterministic theory of particles in motion, there are two equations that completely describe the motion of a particle (or particles) guided by a wave. For a non-relativistic system of particles, with the state of the system being (Q, ψ), the state evolves according to

$$\frac{dQ}{dt} = \operatorname{Im} \frac{\nabla \psi}{\psi}(Q),$$

(3.15)

and

$$i\frac{\partial \psi}{\partial t} = H\psi,$$

(3.16)

where H is the Schrödinger Hamiltonian.

More specifically, for an N-particle universe of spinless particles with mass m_k, interacting via a potential energy function $V = V(q)$, the equations are

$$\frac{d\mathbf{Q_t}}{dt} = \mathbf{v}_k^{\psi}(\mathbf{Q}_1, ..., \mathbf{Q}_N) \equiv \frac{\hbar}{m_k} \operatorname{Im} \frac{\nabla_k \psi}{\psi}(\mathbf{Q}_1, ..., \mathbf{Q}_N),$$

(3.17)

and

$$i\hbar \frac{\partial \psi_t}{\partial t} = -\sum_{k=1}^{N} \frac{\hbar^2}{2m_k} \nabla_{q_k}^2 \psi_t + V\psi_t$$

(3.18)

Bohmian mechanics is manifestly nonlocal, as can be seen in these equations; the velocity of one particle typically depends on the positions of the other particles. However, a Bohmian universe with potential V is completely specified by equations 3.17 and 3.18 – there are no further postulates needed. In this way, the Bohmian interpretation solves the preferred basis problem by adding in these 'hidden variables', the actual positions of the particles, hence essentially choosing the basis to be position. The problem of unique outcomes is also no longer a problem when using the Bohmian framework because the quantum state is seen to be incomplete both before and after Schrödinger time evolution. With the Schrödinger equation and the guiding equation forming a complete description of the quantum world, there is no issue with macroscopic superpositions.

79

In order to make Bohmian mechanics empirically equivalent to ordinary quantum mechanics one needs a "quantum equilibrium" distribution, $|\psi|^2$. The quantum formalism gives a probability distribution of $\rho = |\psi(q)|^2$ for a wave function at configuration q. Dynamical systems have statistical behaviour, with the statistics given by the stationary probability distribution for the dynamics. In Bohmian mechanics, an equivariant probability distribution is needed, rather than a stationary one.[4] A probability distribution ρ^{ψ} on configuration space, depending on ψ, is equivariant if:

$$(3.19) \qquad (\rho^{\psi})_t = \rho^{\psi_t}.$$

Here, the dependence on t on the right is from Schrödinger's equation, and on the left the dependence arises from the evolution on probability densities from equation 3.15. $\rho^{\psi} = |\psi|^2$ is equivariant, as this is an immediate consequence of the fact that the probability current, $J^{\psi} = \rho^{\psi} v^{\psi}$, where v^{ψ} is the right hand side of 3.15. At t_0,

$$(3.20) \qquad \rho_{t_0}(q) = |\psi_{t_0}(q)|^2$$

goes to

$$(3.21) \qquad \rho_t(q) = |\psi_t(q)|^2$$

for all t.

This is considered in Dürr et al. (1992, 1995) to be roughly analogous to classical thermodynamic equilibrium:[5]

$$(3.22) \qquad \rho \sim e^{-\beta H_{\text{classical}}}$$

---

[4] A function is an equivariant map when computing the function and applying a transformation in either order gives the same result.

[5] For further work on this, see Valentini (1991).

There are many well-known objections with Bohmian mechanics not covered here, which are discussed in detail by Goldstein (2017). These include the argument from Leggett (2005) that Bohmian mechanics as an interpretation does not actually solve the measurement problem, but merely reformulates it.

3.5 Collapse Theories

Collapse theories of quantum mechanics focus on the problem of definite outcomes — the 'macro-objectification' problem. The Dynamical Reduction Program takes the unitary dynamics of quantum theory and adds the postulate of Wave Packet Reduction, the dynamics of which are nonlinear and stochastic (Ghirardi, 2018). There is a physical collapse where the state vector is seen as a physical state, which is reduced to one of the eigenstates in the superposition. In the original collapse theory, considered by Ghirardi (2018) to have begun in Pearle (1976, 1979), the wave-packet reduction postulate guarantees that measurements have outcomes. There is still the preferred basis problem, however — in this approach, in which basis does this physical reduction happen? It also does not answer the question of why this reduction happens more at the macroscopic level than at the microscopic level.

A model of a collapse theory with spontaneous localisations (Quantum Mechanics with Spontaneous Localisations, or QMSL) was first developed by Ghirardi, Rimini and Weber (1986) in order to solve the preferred basis problem — the most troubling superpositions at the macroscopic level are taken to be those of the positions of macroscopic objects. The model uses spontaneous processes causing the wave function to become spatially localised. These dynamical processes are instantaneous and happen at the microscopic level, leading to spatial objectivity at the macroscopic level.

The QMSL is usually called GRW after its developers. Simplified, it assumes that each particle (of a system of N distinguishable particles labelled by index i) experiences a sudden spontaneous localisation with mean frequency λ^i. This is described by

$$(3.23) \qquad\qquad |\psi\rangle \longrightarrow |\psi_x^i\rangle = L_x^i |\psi\rangle,$$

where L_x^i is a norm-reducing, positive, self-adjoint, linear operator representing the localisation of particle i around position x. This operator contains a length parameter $\frac{1}{\sqrt{\alpha}}$ that measures the localisation volume. This is chosen to be a Gaussian function in Ghirardi et al. (1986). Rewriting the description above to be norm-conserving and non-linear gives

$$|\phi\rangle \longrightarrow |\phi_x^i\rangle = \frac{|\psi_x^i\rangle}{||\psi_x^i||},$$

(3.24)

$$|\psi_x^i\rangle = L_x^i |\phi\rangle .$$

Contextually assuming that for the x occurring the probability density is

(3.25)
$$P_i(x) = ||\psi_x^i||^2,$$

this requires that

(3.26)
$$\sum \vec{d}^3 \vec{x}(L_x^i) = 1.$$

This is a Markov process in Hilbert space (Ghirardi et al., 1990). The sudden finite changes of the state vector are called a "hitting" process. In GRW, the collapses occur universally and spontaneously, happening instantaneously at the microscopic level, and they tend to suppress the linear superpositions of differently localised states. A key assumption of this account is that each constituent of a physical system could at any time be subjected to the random and spontaneous localisation processes, or 'hittings', around appropriate positions (Ghirardi, 2018).

In a system of N distinguishable particles, with the wave function $F(\vec{q}_1, \vec{q}_2, ..., \vec{q}_N)$ of the state vector, if a hitting occurs for the i-th particle at x, the wave function is multiplied by a Gaussian function

(3.27)
$$G(\vec{q}_i, \vec{x}) = K \exp\left[-\left\{\frac{1}{(2d^2)}\right\}(\vec{q}_i - \vec{x})^2 \right],$$

where d is the localisation accuracy. Immediately after the localisation, the wave function (not normalised) is

(3.28)
$$L_i(\vec{q}_1, \vec{q}_2, ..., \vec{q}_N; \vec{x}) = F(\vec{q}_1, \vec{q}_2, ..., \vec{q}_N) G(\vec{q}_i, \vec{x}).$$

In accordance with standard quantum theory, the hittings occur with higher probability at the places where the standard quantum probability of finding that particle is higher. These hittings happen at random times according to a Poisson distribution. This mechanism is enhanced when there are a number of particles in a superposition further apart, spatially, than d. When one has a superposition $|S\rangle = |H\rangle + |T\rangle$, with H and T corresponding to two macroscopic pointer positions, with N constituents, the state can be written as follows:

(3.29)
$$|S\rangle = \left[|1 \text{ near } h_1\rangle ... |N \text{ near } h_N\rangle + |1 \text{ near } t_1\rangle ... |N \text{ near } t_N\rangle \right].$$

The states in the first term on the right hand side of this equation have coordinate representations different from zero only when their arguments $(1, ..., N)$ are all near H, and vice versa with the second term and T. If any of the particles (i) undergoes hitting near t_i, the first term will be suppressed – and conversely if the ith particle underwent the process near h_i, the second term will be suppressed. Two parameters are chosen to avoid macroscopic superpositions while making sure that quantum predictions remain valid; localisation accuracy d and mean localisation frequency f. The GRW-model chooses $f = 10^{-16} s^{-1}$ and $d = 10^{-5} cm$. These allow for microsystems to undergo a localisation every hundred million years, while macrosystems undergo one every 10^{-7} seconds.

GRW was later improved upon by Pearle (1989) and Ghirardi, Pearle and Rimini (1990), since it has the major problem of not being able to account for indistinguishable particles and preserve symmetry properties of the state vector. The new version, the Continuous Spontaneous Localisation (CSL) model has continuous stochastic evolution in the Hilbert space, rather than the discontinuous jumps in GRW, where one is assuming a stochastic evolution equation

(3.30)
$$d|\psi\rangle = [-iHdt + dh - \frac{1}{2}\overline{(dh)^2}]|\psi\rangle,$$

where dh is a random, self-adjoint, linear operator which is built up with operators representing the density of particles around all points of space, containing a length parameter $\frac{1}{\sqrt{\alpha}}$ and a strength parameter γ. When conserving the norm,

$$(3.31) \qquad d\,|\phi\rangle = [-iHdt + dh_\phi - \frac{1}{2}\overline{(dh_\phi)^2}]\,|\phi\rangle\,.$$

Since dh_ϕ depends on $|\phi\rangle$, this is norm-conserving and nonlinear (Ghirardi et al. (1990)).

For the simplified version of this model, there need to be three assumptions; that you are dealing with only one kind of particle, that the standard Schrödinger term in the evolution can be disregarded, and that you can divide the whole space into cells with volume d^3. Then, denoting a Fock space with n_i in cell i as $|n_1,n_2,...\rangle$, you consider a superposition between the state $|n_1,n_2,...\rangle$ and a state $|m_1,m_2,...\rangle$. Using the previous assumptions, it can be proved that the rate of suppression between the two states is controlled by the quantity

$$(3.32) \qquad exp\{-f[(n_1-m_1)^2 + (n_2-m_2)^2 + ...]t\},$$

with all the cells of the universe appearing in the sum within the square bracket in the exponent (Ghirardi, 2018). Though there is some slight differences concerning the identity of the constituents, the physics of the CSL is similar to that of the GRW. In principle, different collapse models are empirically indistinguishable. There are other collapse theories where the collapse is not spontaneous as in GRW and CSL, but instead due to some particular kind of interaction – for example, Penrose (1989) relates gravitational effects to reduction methods.

There are many critical responses to collapse theories, and the GRW and CSL model specifically, not least there being the issue of expanding the theory to the relativistic level.[6] Nevertheless it solves both the measurement problems and gives a macrorealistic position without violating the current evidence for quantum mechanics. How satisfying the solutions are depends on how one feels about solving the preferred basis problem by fiat, since the collapse model solution of unique outcomes stems from this, and on whether a collapse model is empirically confirmed in a departure from what quantum mechanics predicts.

---

[6]There is work going on in this area. Tumulka (2006) has successfully proposed a relativistic version of the GRW-model for N non-interacting distinguishable particles.

3.6 Modal Interpretations

Modal interpretations of quantum mechanics use the standard quantum formalism with the projection postulate left out. This is done to remove the strange asymmetry in the dynamics with 'collapse', and to deny any special status to measurement. In this interpretation, the quantum wave function is the dynamical state that evolves unitarily according to the Schrödinger equation and never 'collapses' — the state determines what properties a system may have and the probabilities of those. There is also a 'value state', described by van et al. (1991) as representing the actual physical properties of the system at the time in question. The dynamical state and the value state have a probabilistic relationship, and quantum mechanics is complete. Realist modal theories employ an 'actualisation rule' which picks out the definite-valued observables from the dynamical state (Lombardi and Dieks, 2017).

The first kind of Modal interpretation was the Biorthogonal Decomposition Modal Interpretation (BDMI). When there is a state vector representing a composite system of a quantum object and the rest of the system, then there will almost always be a unique bi-orthogonal decomposition of that state vector (Dieks, 1989). Once one has this unique decomposition, the partial system of just the quantum object can be said to possess one of the values of the physical quantity corresponding with the set of basis states from the decomposition. The probabilities of these possible values being 'realised' depend on the amplitude. There are two different views on how to understand the probabilities in this interpretation; Kochen (1985) puts forward the view that all the properties were relational, and a system possess a property when it is 'witnessed' by another system. The other view, held by Dieks, is that the probabilities are simply a measure of our ignorance — though this leads to issues with the BDMI due to the seemingly arbitrary splitting of a total system up into components.

The BDMI was generalised to become the SDMI (Spectral-decomposition Modal Interpretation), and does so by fixing the definite-valued properties in terms of multi-dimensional projectors when the bi-orthogonal decomposition is degenerate (Vermaas and Dieks, 1995). The SDMI uses the spectral decomposition of the reduced density operator. The preferred context of the system is defined by the projectors of the definite-valued properties and their corresponding observables,

and the diagonalised elements of the reduced density state give the definite-valued properties and their probabilities.

There is still the problem of how to decompose the state, as in the non-relational version of the BDMI. This is the problem of the preferred basis. This can clearly be seen with the BDMI/SDMI when they are applied to non-ideal measurement. Here, the rules for selecting the definite-valued properties of the system do not always work, and sometimes disagree with experimental results (Bacciagaluppi and Hemmo, 1996). The Perspectival Modal Interpretation (PMI), to solve the issues of the previous two interpretations, has the properties take a relational character (Bene and Dieks, 2002). In this account, all the different relational descriptions are equally objective.

For those who reject the relational modal interpretation, the Modal-Hamiltonian Interpretation (MHI) is an option. The MHI takes the view that quantum states constrain possibilities, rather than actualities, and gives the Hamiltonian of a system the determining role, both for defining the systems and subsystems and in selecting the preferred context. The Hamiltonian of the system defines actualisation of properties, and any observable which doesn't have the symmetries of the Hamiltonian cannot acquire an actual definite value, as that would break the symmetry of the system. In the MHI, the basic elements of the theory are observables — states are logically posterior (Da Costa et al., 2013). The Actualisation Rule depends on the Hamiltonian of the system, and there exist universal type-properties (for example, energy), of which any instance has a possible case-property. The quantum system is a bundle of instances of these type-properties and only one will become actual. Rather than the relational view of the PMI, and the interpretation of a measure of our ignorance in the BDMI/SDMI, the MHI takes a possibilist view of probabilities. In this view, possible events are a basic ontological category, and the probability measure is seen as the representation of a certain ontological quality of a possible quantum event to become actual (Lombardi and Dieks, 2017).

There are some issues with formulating a relativist form of the modal interpretation; Lombardi and Dieks (2017) give a few of the attempts and the difficulties faced with this extension of the interpretation, and consider it an open problem.

The modal interpretation, in all its many forms, solves the unique outcomes problem in a direct way. However the non-relational modal interpretations are left with the preferred basis

problem. There are ways to deal with this in the modal interpretation framework and these include using decoherence. This is briefly discussed in chapter 5, and the theory of decoherence is explored in detail in chapter 4.

3.7 Oxonian Everettianism (Many Worlds)

Begun by Everett III (1957) when he proposed that all of the terms of a quantum state corresponded to a physical state once a measurement had been completed, the many-worlds interpretation was taken further by DeWitt (1970) and Deutsch (1985). They understood the physical state of each term in the superposition as corresponding to a universe constantly splitting into branches, each branch a simultaneous world where the measurement has yielded a different outcome. Another view — the many-minds interpretation – describes the physical states as relative to a particular mind, rather than a particular world, of a conscious observer (Lockwood, 1996).

The Many Worlds interpretation of Saunders (1993); Saunders et al. (2010) and Wallace (2010) takes quantum theory literally and therefore simply consists of the theory of time evolution of a quantum state of the single universe (the Schrödinger equation), and of a prescription which gives us the correspondence between the quantum state of the universe and our own experiences.

It is in this simple fashion that the MWI solves the unique outcomes problem — the MWI says that all the outcomes occurred, though they all occurred in different quasi-classical worlds. The world around us is one of these countless worlds. The mathematical formalism of quantum theory remains unchanged by this interpretation. The correspondence between it and our experience — why we get classical measurements at the macrolevel — is approximate, which is fine for all practical purposes (FAPP).

Though it supplies a simple answer to the definite outcomes problem, the MWI has to deal with the preferred basis problem and the probability problem. There is no probability problem with the relativistic extension of the MWI, as it is simply quantum field theory with unitary evolution. The preferred basis problem, according to Saunders (1993) and Wallace (2010), can be dissolved by using decoherence – a relatively recent theory in the timeline of quantum theory,

which is explored in chapter 4.

3.8 QBism

QBism, or Quantum Bayesianism, is a relatively recent interpretation introduced by Caves et al. (2002*a*,*b*). It is an interpretation which QBists claim is realist about the world but anti-realist about the structure of the theory; quantum theory, and science in general, is merely a user manual for an agent to apply to their external world. There is no unique, observer independent, quantum-state 'stuff' in the world.

As the name 'Quantum Bayesianism' indicates, this interpretation takes a subjective Bayesian view of the probabilities in quantum theory, and also applies that subjectivity to quantum state assignments and measurements. For a QBist, there is no such thing as 'the' quantum state, and there can be as many quantum states as there are 'agents', or 'users'.[7] The subjective Bayesian is one agent using probability to make decisions and not wanting to be incoherent.

The first tenet of QBism, as given by Fuchs (2017), is that quantum theory is normative, not descriptive — it is a user manual for agents, but does not describe the underlying structure of the world. Probability theory can be thought of as a normative theory of rational belief, and Fuchs and Stacey (2019) argue that quantum theory is an addition to probability theory. Fuchs (2016) states that determinism had vanished from the world since the advent of quantum theory, and illustrates this with a quote from J.A. Wheeler,

"...the only law of nature is that there is no law" (Fuchs, 2016, p. 3),

though of course, the disappearance of determinism allows for there being probabilistic laws. QBism uses a particular type of Bayesianism — the subjective Bayesian interpretation which has probability as degree of personal belief.[8]

QBism focuses on agents, who by measurement bring about outcomes that are not previously-existing properties. This interpretation is a mixture of subjective probability and objective

---

[7]While Fuchs and Schack prefer the word 'agent' because it better highlights that the user of quantum theory takes actions and experiences the consequences of those actions, Mermin (2017) prefers the term 'user', as it makes more of the fact that science is a user manual.

[8]This uses de Finetti's theorem, which can be seen in more detail in Caves et al. (2008).

indeterminism. The probabilities of measurement outcomes given by the quantum state are interpreted as dependent on the agent's subjective experience, following the laws of Bayesian probability. Quantum theory is only describing the world insofar as quantum theory tells us how to correctly update our beliefs, and quantum states are mental constructs.

From a subjective Bayesian's perspective, "probability is the degree of belief as measured by action" (Fuchs and Schack, 2013, p. 6). If you have assigned a probability $p(A)$ to an event A, before knowing what the value of A is, then you are therefore going to want to either buy or sell a lottery ticket (worth £1 if A) for £$p(A)$. This definition of probability has nothing to do with the event A itself, or any properties of the event, but merely the gambler's own subjective belief about that event. The only rule, Fuchs writes, is that one should aim to never gamble on an outcome of an event that one believes is a sure loss — something that seems much like common sense. This is called Dutch-book Coherence, as discussed by Caves (2000). Fuchs and Schack (2011) write that the Born rule in QBism is merely an empirical addition to Dutch-book coherence. It states that there is no such thing as the absolute, correct quantum state of the system, but rather that every measurement is subjective to an agent's experience.

Fuchs uses Wheeler (1979)'s twenty questions analogy to describe QBism's participatory realism. In a normal game of twenty questions, one person in the group leaves the room while the others decide on the answer. The person comes back in and proceeds to ask twenty questions (or fewer), hopefully reaching the correct answer. This is Wheeler's 'classical' game. His quantum game, however, has a slightly different format. This time, the group of people don't decide on a word for the answer, but simply answer yes or no to each question as they see fit — the only caveat being that they must have a word in mind that could possibly fit both their yes/no answer and all the questions that had gone before. When the questioner gets to the end of their twenty questions and gives an answer, that word is technically correct — but it did not exist as an answer until it was reached. Wheeler highlights the fact that not only was the word not 'out there' from the beginning of the questioning, but also that it came into existence *because* of the person asking the questions. In this way, it is subjective to the person asking the questions. The person is participating in the creation of this world, and QBism is participatory realism carried to its 'logical extremes' (Fuchs, 2016).

In QBism, quantum mechanics is characteristic of the world but not descriptive of it. Fuchs and Schack rewrite the Born rule in terms of SICs (symmetric informationally complete observables) and POVMs.[9] They describe an experiment where there is a fixed 'sky measurement' with outcomes $i = 1, ..., n$ and a potential 'ground measurement' in a lab with outcomes $j = 1, ..., m$, with the sky measurement being a SIC with $n = d^2$ outcomes and the ground measurement being a POVM. The following equation is fundamental to QBism,

$$(3.33) \qquad q(j) = (d+1) \sum_{i=1}^{d^2} p(i) r(j|i) - \frac{1}{d} \sum_{i=1}^{d^2} r(j|i),$$

where the probability distributions $p(i)$ and $r(j|i)$ represent how the agent in the sky would gamble if there were a conditional lottery based on the outcome of the sky measurement, and $q(j)$ is what the agent would gamble instead if there were no sky measurement and conditional lottery. This equation is equivalent to the Born Rule — one of its terms is equivalent to the classical Law of Total Probability,

$$(3.34) \qquad P(A) = \sum_n P(A|B_n) P(B_n),$$

and the other depends on the sum of the conditional probabilities.

> "Quantum theory is conditioned by the character of the world, but yet is not a theory
> directly of it." (Fuchs and Schack, 2013, p. 26)

It is instead a tool, or a user manual, with which we can organise our own personal experiences. Some might worry that, in a purely subjective interpretation, there appears to be no reason why there is such agreement on values widely regarded to be objectively real in the world. QBism has a response; while everyone's reality is subjective, there can be a common reality created by sharing experience with others through language, a 'common body of reality' (Fuchs et al., 2014). There is a world independent of that experience, but it is not described by quantum theory. Even something considered to be an objective fact from theory, such as the ground state

---

[9]A POVM is a Positive Operator Value Measure, and their values are are non-negative self-adjoint operators on a Hilbert space. Their integral is the identity operator.

of the hydrogen atom, is just a widely shared degree of belief.[10] However, sharing experience is limited to language. Personal experience is much more complex. Therefore, the shared reality is somewhat less 'real' than the personal reality created by personal experience (Fuchs et al., 2014). For QBists, quantum mechanics does not deal directly with the world, but it is an objective fact that it can be used by agents. The second tenet of QBism is that the agent's probabilities do not tell the world what to do. A calculation might tell an agent to set the probability of a particular outcome to 1, but that does not mean the world should comply with that result. It merely suggests that it would be the best gamble the agent could make. 'Nature does what it wants,' and there is no underlying mechanism, but quantum theory gives us the tools with which to gamble on what nature will do (Fuchs, 2017, p. 19). This also has the added benefit of QBism being a completely local interpretation of quantum theory.

The third tenet of QBism is that there is no agent independent fact about what particular measurement a particular device performs — the detectors are as much part of the agent as the agent's hands. The emphasis that QBism places on the *subjective* nature of measurements and personal experience is what drives this tenet; a measurement is a subjective experience, and so the detectors and devices used are also part of that experience.

The three tenets of QBism, summarised, are as follows:

1. Quantum theory is normative, not descriptive.

2. The agent's probabilities do not determine what will happen in the world.

3. There is no agent independent fact about what particular measurement a particular device performs.

Quantum systems still have real existence outside of the agent, despite the subjective nature of quantum states. But for QBism, quantum states must be completely subjective degrees of belief, not degrees of truth — for if there is some external way to judge the right or wrong-ness of the quantum state, Fuchs and Stacey (2019) argue it must have some physical properties, and quantum states with physical properties fall afoul of Wigner's friend and lead to 'spooky action at a distance'.

---

[10]For more about this argument in detail, see Fuchs (2017).

Several critics of QBism believe the subjectivist stance is too strong; Timpson (2008) argues that QBism's subjective probabilities run into trouble when encountering pure states, and Bacciagaluppi (2014) argues that subjective probabilities in QBism need to be supported by an ontic quantum state. The differences between subjective vs. objective vs. intersubjective are subtle; could it be that QBism need not be subjective?

There are other problems with the subjectivist position; it is not clear whether, in the QBism picture, it applies to only quantum mechanics or to the whole of science. There is a difference of opinion within QBism about this; while Schack argues that all theories are user theories like quantum mechanics, and therefore completely personal to each agent, Pusey argues that other theories can have an unproblematic realist reading, if that reading is the best explanation for why that theory is so useful and successful. For quantum mechanics, it is simply that for the QBist the best reading of it is as a user manual. But it seems odd that one can call quantum mechanics completely subjective but have a realist, objective view of classical physics, since quantum physics transitions to classical physics in certain limits.

To find out whether QBism really needs to be subjective, then, a good place to start looking is at the quantum–classical divide. How QBism deals with quantum decoherence and the transition between quantum and classical phenomena reveals whether subjectivity is required for their view, and whether this must apply to the whole of science or just to quantum mechanics. And not just QBism – all the interpretations are looking at the transition between the quantum world and the classical, and so how they deal with or use the process of decoherence is important to how they solve the measurement problem.

The next chapter explores the theory and history of decoherence, before coming back to each of these interpretations in turn and looking with fresh eyes at their solutions to the measurement problems in the light of what is learnt about decoherence.

4

DECOHERENCE

The study of decoherence began relatively late in the timeline of quantum physics – it started with Zeh (1970), and then was later picked up by Zurek (1981, 1982). Zeh's paper is now considered the modern beginnings of decoherence theory and proposes that the openness of quantum systems is the reason that the unitary dynamics of the Schrödinger equation vanish at the macroscopic level. However, at the time it was mostly ignored by theoretical and experimental physicists, and was repeatedly rejected by journals. This was, according to Camilleri (2009*b*), due to Zeh's starting assumption of a universal wave function and an Everettian view of quantum theory. After Zurek's own related work was published in 1981, there was renewed interest in Zeh's work from the previous decade. Zurek's work was much more easily accepted into the scientific community – his work on the preferred basis problem and superselection rules based on environment-induced decoherence was not "directly motivated by 'interpretational' issues" (Camilleri, 2009*b*, p. 291), and instead focused more on the dynamics. It did not hurt that his work on decoherence, rather than being an explicit rival to the Copenhagen interpretation like Zeh's Everettian interpretation, was not only compatible with the Copenhagen interpretation but could even be seen as positive confirmation of many of Bohr's thoughts on quantum theory. This enabled physicists to be able to use Zurek's formalism as a "completion" or "justification" of what they saw as the original Copenhagen interpretation, making Zurek's decoherence much more palatable (Camilleri, 2009*b*).

Decoherence is now formally thought of as a dynamical filter on the space of quantum states, picking out those states of the system that are stable enough to be prepared and maintained, while barring others, such as nonclassical superposition states of macroscopic systems.

Decoherence deals with the transition between the quantum world and the classical world, where only a select few of the states allowed by the quantum mechanical superposition principle actually make it through to becoming a 'classical' state. The emergence of the classical world from quantum theory therefore is at the core of both the measurement problems, and it is environment-induced decoherence theory that many interpretations employ to solve them. Decoherence theory showed the importance of system-environment correlations in emergent classicality, and so shows itself to be a useful tool in examining the quantum to classical transition by providing a "quantitative, dynamic account of the boundary between quantum and classical physics" (Schlosshauer, 2014, p. 1).

This chapter explores the theory of decoherence, and an interpretation of quantum theory produced parallel to decoherence and which does not need decoherence to solve the measurement problem due to its removal of measurement as the core issue of quantum mechanics. This emphasises the fact that decoherence is a good pressure-test for other interpretations of quantum theory that do focus on measurement. To solve the big measurement problem is to deal with the quantum to classical transition, and decoherence is a major part of that transition.

4.1   Basic Decoherence Formalism

Coherence is definedin terms of phase relations staying constant over time. As stated in section 1.2, all states have phases. This is a contrast to classical mechanics, where only waves have phases. These phases are the reason we see quantum interference, as in the famous double-slit experiment. Decoherence at its most basic destroys this coherence between phases and makes the interference pattern go away.

An example is a simplified two-slit experiment. The probability of a particle landing on the screen going through one slit or the other, from applying the Born rule to the state

(4.1)
$$|\psi\rangle = a\,|A\rangle + b\,|B\rangle,$$

is given by

(4.2)
$$|a|^2 + |b|^2 + a^* b + a b^*,$$

where the last two terms are the interference terms that cause the fringe pattern on the screen. When a second system is added it entangles with the particle state and therefore the environment states cannot be traced out (as stated in section 1.2). The interference terms will disappear, and therefore so will the interference pattern on the screen.

For another example, consider the pure state $|+\rangle$, which can be written in the basis $\{|0\rangle, |1\rangle\}$;

(4.3)
$$|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle).$$

The density operator, as mentioned in 1.1, is $\rho = |+\rangle\langle+|$. Due to time evolution, $|0\rangle$ goes to

(4.4)
$$|+\rangle \longrightarrow \frac{1}{\sqrt{2}}(|0\rangle + e^{i\omega t}\,|1\rangle).$$

The exponential term, the phase, is in the off-diagonal terms of the density matrix, and this will give an expectation value,

(4.5)
$$\langle+|\,\rho(t)\,|+\rangle = \frac{1}{2} + \frac{1}{2}\cos(\omega t),$$

where $\rho(t)$ is the density matrix at t, representing the density operator. The density matrix can also be written as

(4.6)
$$\rho(t) = \frac{1}{2}\begin{pmatrix} 1 & e^{i\omega t} \\ e^{-i\omega t} & 1 \end{pmatrix}.$$

This is called the coherence of the system. However, this coherence does not last in real physical systems, and decays exponentially at rate λ. This then gives an expectation value of

$$\langle +|\rho(t)|+\rangle = \frac{1}{2} + \frac{1}{2}e^{-\gamma t}\cos(\omega t), \tag{4.7}$$

where γ is the rate at which the coherence decays exponentially.

This de-phasing of the system is decoherence. The coherence has decayed and the rate at which it does, or the value of γ, depends on the mechanism that leads to decoherence.

Decoherence is a general process that is very fast, ubiquitous, and spontaneous — it works against experimentalists who want to retain coherence in the quantum systems they use. Decoherence can be seen clearly experimentally in the case of diffracting large molecules, using a experimental set-up similar to a two-slit experiment. The wave-particle duality of even large molecules such as C_{60} and C_{70} has been explored since the late 1990s, when it was first shown by Arndt et al. (1999) that it is not only electrons that give interference patterns when passed through two slits. At that time, these 'almost classical bodies' were the largest molecules to have been shown to give interference patterns. However, this superposition depends entirely on the molecules not passing definitely through one slit or the other. As the temperature is increased, the molecules which were previously well-isolated from the environment gain a higher internal energy. This leads to emitted and absorbed radiation, scattering, collisions and other such processes, which shows the molecules passing through one of the two slits, and leading to decoherence. This experiment was performed by Hackermüller et al. (2004). It was shown that C_{70} molecules transition between quantum and classical behaviour gradually – the interference fringes losing their visibility – as their internal temperature was increased to 3000K. There are other kinds of decoherence, such as environment-induced decoherence as talked about by Zeh and Zurek, decoherence from scattering, and relaxation to ground state by spontaneous emission.

4.2 Environment-Induced Decoherence

Environment-induced decoherence can be described as a physical process in which the environment monitors a system – the environment is constantly performing small, indirect measurements

on a quantum system. Decoherence is a very efficient process, and once a system is entangled with the environment it is usually irreversible for all practical purposes. Using the example of a two-slit set-up from Schlosshauer (2014), the particle in question passes through slit one or two, $|s_1\rangle$ and $|s_2\rangle$ respectively. When this particle interacts with another system, E[1], the quantum state of E will become either $|E_1\rangle$ or $|E_2\rangle$, depending on the state of the particle.

For the particle, in the initial superposition state $\alpha|s_1\rangle + \beta|s_2\rangle$, the final composite state will be

$$\text{(4.8)} \qquad |\Psi\rangle = \alpha|s_1\rangle|E_1\rangle + \beta|s_2\rangle|E_2\rangle.$$

We can use the reduced density matrix to find out the statistics of all possible local measurements,

$$\text{(4.9)} \qquad \begin{aligned} \rho_s &= Tr_E(\rho_{SE}) = Tr_E|\Psi\rangle\langle\Psi| \\ &= |\alpha|^2|s_1\rangle\langle s_1| + |\beta|^2|s_2\rangle\langle s_2| + \alpha\beta^*|s_1\rangle\langle s_2|\langle E_2|E_1\rangle + \alpha^*\beta|s_2\rangle\langle s_1|\langle E_1|E_2\rangle. \end{aligned}$$

If, for example, we then set up a detector screen to measure the particle's position, we get the particle probability density,

$$\text{(4.10)} \qquad \begin{aligned} p(x) &= Tr_s(\rho_{s^x}) \\ &= |\alpha|^2|\psi_1(x)|^2 + |\beta|^2|\psi_2(x)|^2 \\ &\quad + 2Re\{\alpha\beta^*\psi_1(x)\psi_2^*(x)\langle E_2|E_1\rangle\}, \end{aligned}$$

where $\psi_1(x) = \langle x|s_i\rangle$. The last term in the probability density is the contribution the environment makes to the interference pattern. How visible the interference is depends on the value of $\langle E_2|E_1\rangle$ – if it is 0, then there is no interference pattern observed and the results of the experiment are the classical predictions of a particle going through one slit or the other. If, however, $\langle E_2|E_1\rangle = 1$, the interaction between the environment and the particle has not resolved which path the particle took, and there will be an interference pattern on the detector screen. There is still full coherence.

[1]This other system could be the environment or a detector.

In between these two cases, when $0 < |\langle E_2 | E_1 \rangle| < 1$, the interference pattern on the screen is less visible.

In this example, the environment, through interaction with the particle, encodes information about which way the particle appears to go in the case of $\langle E_2 | E_1 \rangle = 0$. When this happens, we can predict with certainty that $|s_2\rangle$ if we find out that E is in $|s_2\rangle$, for example.

A more general, idealised account of this decoherence interaction has the form[2]

$$(4.11) \qquad \left(\sum_i c_i \, |s_i\rangle \right) |E_0\rangle \longrightarrow \sum_i c_i \, |s_i\rangle \, E_i(t).$$

The time parameter t is brought in where $t = 0$ at the start of the interaction with the environment, where it is generally assumed that the system and the environment are uncorrelated at $t < 0$. When there is only one particle in the environment interacting with the quantum system in question, this is generally not enough to resolve which state the system is in. Environmental states of large numbers of particles which are effectively orthogonal (and with many degrees of freedom) are sufficient, but not necessary, to resolve a system into one component of the superposition state. This build-up of interaction events decreases the overlap between their different joint states, $|E_i(t)\rangle$, and this decrease is often exponential. This gives

$$(4.12) \qquad \langle E_i(t) | E_j(t) \rangle \propto e^{-\frac{t}{\tau_d}} \qquad \text{for } i \neq j,$$

where τ_d is the characteristic decoherence timescale, which changes depending on the different parameters chosen for each decoherence model. There are several developed models of decoherence, including collisional decoherence, where one large quantum particle scatters many environmental particles.[3] Work on modelling dissipative systems was done by Caldeira and Leggett in the 1980s, at the same time as Zurek was working on decoherence.[4]

First discussed in Zurek (1981), environment-induced decoherence is based on the idea that quantum correlations are so ubiquitous as to correlate almost every physical system with its

---

[2] As seen in Schlosshauer (2014).

[3] First looked by Joos and Zeh (1985), and later refined by Hornberger and Sipe (2003).

[4] See Caldeira and Leggett (1981, 1983, 1985), and Leggett (1996) for their work.

environment (barring, of course, isolated microscopic systems where linear quantum mechanics and superpositions of states can be empirically measured).[5] Because of these correlations of the system with the environment there is no real such thing as a 'pure' quantum world in the majority of our experiences, and the interaction with the environment imposes "effective superselection rules" (Schlosshauer, 2005, p. 7) onto the space of observable states, which lead to the classical properties we experience in the world. These 'rules' prevent the selection of two coherent eigenstates, meaning that once the quantum system is decohered, there is only classical coherence between two eigenstates.

In this way, decoherence theory can be applied to the preferred basis problem. It is a technical result following from empirical evidence of the dynamics and measurement statistics of open quantum systems, and can explain when quantum probability distributions become the probability distributions that are classically expected. It shows how the 'preferred' states of the system — the ones which emerge as objective measurement outcomes — are the ones which are the most 'robust' to interaction with the environment. The idea is that the states most immune to decoherence with the environment — pointer states — satisfy the stability criterion. It does this through environment-induced selection of a preferred basis, which is where "the interaction between the apparatus and the environment singles out a set of mutually commuting observables" (Schlosshauer, 2005, p. 12).

The environment-induced framework takes the total composite system of the quantum system and the environment as the Hilbert product space, plus the interaction Hamiltonian,[6]

$$(4.13) \qquad\qquad \mathscr{H} = \mathscr{H}_S + \mathscr{H}_E + \mathscr{H}_{\text{int}}.$$

Taking a basis $\{|s_i\rangle\}$ so that the composite state of the system and environment stays in the product form $|s_i\rangle|E_i(t)\rangle$ for all times after $t > 0$ under the interaction Hamiltonian, the product starts as $|s_i\rangle|E_0\rangle$. This gives us

---

[5] See Joos et al. (2013) for a detailed look at environment-induced decoherence.

[6] In most cases, the interaction Hamiltonian dominates and $\mathscr{H} \approx \mathscr{H}_{\text{int}}$, which is the quantum-measurement limit of decoherence (Schlosshauer, 2014, p. 5).

(4.14)
$$e^{-iH_{\text{int}}t}|s_i\rangle|E_0\rangle = \lambda - i|s_i\rangle e^{-iH_{\text{int}}t}|E_0\rangle \equiv |s_i\rangle|E_i(t)\rangle,$$

where $|s_i\rangle$ are eigenstates of $\mathcal{H}_{\text{int}}$, with eigenvalue λ_i. These are the pointer states – they are stationary under $\mathcal{H}_{\text{int}}$, and the pointer observable commutes with $\mathcal{H}_{\text{int}}$,

(4.15)
$$\left[\mathcal{H}_{\text{int}}, \sum_i \pi_i |s_i\rangle\langle s_i|\right] = 0,$$

or

(4.16)
$$[\mathcal{H}_{\text{int}}, \hat{\Pi}_s] = 0.$$

This is the commutativity criterion from Zurek (1981); the measurement is performed by the environment, and so the interaction Hamiltonian selects the pointer observable as the one that will not be perturbed by the measurement. In this sense, it is 'robust'.

4.3 Consistent Histories

First proposed in 1984 by Griffiths, then later discussed by Omnés (1988), this formalism is also known under the name 'Decoherent histories' from Gell-Mann and Hartle (1990). Consistent Histories, or simply 'histories', has two main ideas at its core:

1. A new kind of logic that is compatible with the Hilbert space structure of quantum mechanics, but different to classical logic, and the quantum logic of Birkhoff and Von Neumann (1936).

2. A new dynamics that is inherently random.

Since in Consistent Histories all quantum time development is a stochastic process all of the time, not just when a measurement is taken, there is no special role for measurement – it is treated as the same as other physical processes. Griffiths (2019) also divides the measurement problem into two:

M1 : "How can the macroscopic *outcome* of the measurement, traditionally thought of as a pointer position, be described in quantum terms?"

M2 : "How is this outcome related to the earlier microscopic property the apparatus was designed to measure?"

M1 is another way to describe the definite outcomes problem – quantum theory leads us to an entangled state, which is very different to the results seen by experimental physicists. And following M1, if the apparent 'collapse' of the measurement can be explained, comes M2 – how is the macroscopic outcome of the pointer position pointing one way rather than another related to the microscopic state it is measuring? As M1 was another way to state the definite outcomes problem, M2 is another way to state the preferred basis problem.

The answer to these two measurement problems, according to the Histories approach, lies in a consistent quantum ontology which uses a new kind of logic to be compatible with Hilbert space structure, and new dynamics for the quantum world, where all quantum dynamical processes follow stochastic dynamics.[7] The price of this way of solving the two measurement problems is using this new logic, and abandoning determinism.

The new logic Griffiths uses to create this consistent quantum ontology starts with the problem of the difference between classical phase space and quantum Hilbert space. In classical phase space, a property is represented by an indicator function and is 1 at all points where the property is true, and 0 everywhere else. It is either true or false. In Hilbert space however, projectors of a property project onto the subspace corresponding to that property, with eigenvalues 0 or 1. Any ket in this subspace is an eigenvector with eigenvalue 1. The negation of the property is the orthogonal complement of the subspace. However, unlike in classical phase space, where any point in the system is either in the set defining a property or not, in a Hilbert space there can be many rays that are neither in the subspace of a property nor in the orthogonal complement of that subspace, meaning that the property is neither true nor false, but undefined. As Griffith writes, this is a central problem for any ontology using the Hilbert space. He restates it as follows; the product of two quantum projectors P and Q, PQ, is only a projector itself iff $PQ = QP$, i.e. if

---

[7]See Griffiths (2013) for more details.

they commute. When $PQ \neq QP$, then P AND Q is meaningless.[8] So, if the projectors P and Q do not commute, then their P AND Q and P OR Q must be excluded from any meaningful discussion.

An example of this divergence between classical phase space and quantum Hilbert space is the 2-d Hilbert space of a spin-half particle. The particle's spin along the z-axis and along the x-axis, as described by the Pauli spin matrices $[\sigma_z^+], [\sigma_z^-]$ and $[\sigma_x^+], [\sigma_x^-]$, do not commute as discussed in chapter 1. This means that using the conjunction "$S_x = +\frac{1}{2}$ AND $S_z = +\frac{1}{2}$" is meaningless, not false – it is not just that the opposite, or negation, is not true. The disjunction is just as meaningless; "$S_x = +\frac{1}{2}$ OR $S_z = +\frac{1}{2}$"

Continuing with the comparison between classical phase space and the quantum Hilbert space, the Histories approach describes the quantum version of a classical sample space as a 'framework'. This framework is a projective decomposition (PD) of the identity operator $\hat{I}$; a collection of mutually orthogonal projectors $\{P^j\}$ which sum to $\hat{I}$, and satisfy the same condition as a collection of classical indicators:

(4.17)
$$P^j P^k = \delta_{jk} P^j ; \sum_j P^j = \hat{I}.$$

$P^j P^k$ vanishes for $j \neq k$, as the quantum properties are mutually exclusive. Two frameworks, $\{P^j\}$ and $\{Q^k\}$ are compatible if all the projectors in one commute with all the projectors in the other. If not, they are incompatible. This leads to the single framework rule – when using probabilistic reasoning to come to conclusions about a quantum system from some data, a single framework must be used or the reasoning will be invalid. This is especially the case if two incompatible frameworks are used. As well as this rule, the Histories approach uses several principles to guide its logic:

- Principle of Liberty – a physicist can choose whichever framework they want to use.

- Principle of Equality – all frameworks are equally acceptable.

- Principle of Incompatibility – one cannot combine incompatible frameworks, as has been seen above.

---

[8]Here, it is meaningless in the sense that it can be assigned no truth value.

- Principle of Utility – not all frameworks are equally as useful in every case.

These principles, together with the single framework rule, together form the 'consistent quantum ontology' of the Histories approach. This 'ontology', according to its proponents, both provides a more realistic understanding of the quantum world and solves many of the paradoxes and problems within it. As well as this, Histories also has a specific view on quantum time development. The time dependence of a quantum system is a random or stochastic process.

To construct a quantum sample space for a Hilbert space at successive times, take the tensor product of copies of its Hilbert space. For a spin-half particle,

$$(4.18) \qquad \breve{\mathcal{H}} = \mathcal{H}_0 \odot \mathcal{H}_1 \odot \mathcal{H}_2,$$

which describes properties at three successive times, $t_0 < t_1 < t_2$.[9] For a spin-half particle that has the property F_m at t_m, it can be described as the following using a history projector,

$$(4.19) \qquad F_0 \odot F_1 \odot F_2,$$

where each F_m is a projector on the 2-d single time Hilbert space. For example, each F_m is either $[z^+]$ or $[z^-]$. Each element in the sample space is called a 'history', and together they make a 'family of histories', $\mathscr{F}$. Another family could be where each F_m is either $[x^+]$ or $[x^-]$, but these two families would be incompatible.

A sample space $\{Y^\alpha\}$ of histories of the form:

$$(4.20) \qquad Y^\alpha = F_0^\alpha \odot F_1^\alpha \odot ... F_f^\alpha,$$

where α are the different elements of the sample space, and F_m^α is a projector representing the property of the quantum system at time t_m. The histories projectors must satisfy the condition:

$$(4.21) \qquad \sum Y^\alpha = \hat{I},$$

[9]Here $\odot$ is formally the same as $\otimes$, but the different symbol is used to show that it is the product of one system at a sequence of times, not the product of several systems at the same time.

so that $Y^\alpha Y^\beta = 0$ if $\alpha \neq \beta$. This ensures that two distinct histories belonging to the same sample space are mutually exclusive. The simplest type of history family is the unitary family, where they all have the probability of 1 or 0, and for this family there is a deterministic quantum dynamics.

In Histories, the technical term for the wave function that satisfies Schrödinger's equation and therefore undergoes unitary time evolution is the *uniwave*. The uniwave is afforded special status in other interpretations of quantum theory, such as Everett's Many Worlds, but in Histories the uniwave merely describes reality for the unitary family, and following the principle of Equality, the unitary family is one of many families that could be used to describe a closed quantum system. Griffiths describes this consequence of the principle of Equality as "dethroning the uniwave" (Griffiths, 2013, p. 101).

The next most trivial situation for a closed system after the uniwave is a family at only two times, t_0 and t_1. Griffiths gives the example of the histories framework being given an orthonormal basis $|\psi_0^j\rangle$ at t_0 and another, different orthonormal basis, $|\phi_1^k\rangle$, at t_1. With $T(t_0, t_1)$ as the corresponding time development operator, the Born Rule can be used to assign a weight to the history of $[\psi_0^j] \odot [\phi_1^k]$.[10] This is interpreted by Histories as the conditional probability of $|\psi_0^j\rangle$ at t_0 given $|\phi_1^k\rangle$ at t_1, and vice versa.

Using the Born Rule,

$$(4.22) \qquad |\langle \phi_1^k | T(t_0, t_1) | \psi_0^j \rangle|^2 = |\langle \psi_0^j | T(t_0, t_1) | \phi_1^k \rangle|^2.$$

In the Histories approach, since $|\psi_0^j\rangle$ will in general be incompatible with $|\phi_1^k\rangle$, $|\psi_0^j\rangle$ cannot be spoken of as being a property of the system at t_1. Instead it is thought of as a 'pre-probability' – a mathematical tool that need not be represented in the physical reality of the system.

Since the uniwave no longer holds an important position in the Histories approach, the Born weights can be calculated without needing any reference to it. Defining

$$(4.23) \qquad |\phi_0^k\rangle = T(t_0, t_1)|\phi_1^k\rangle,$$

---

[10]In this case, with a time-independent Hamiltonian, this corresponding time development operator is given by $e^{-\frac{i(t_0 - t_1)H}{\hbar}}$.

one can integrate Schrödinger's equation backwards with $|\phi_1^k\rangle$ as the initial state. This then gives

$$(4.24) \qquad |\langle\phi_1^k|T(t_0,t_1)|\psi_0^j\rangle|^2 = |\langle\psi_0^j||\phi_0^k\rangle|^2,$$

and so the weights can be calculated by using $|\phi_0^k\rangle$ as pre-probabilities.

This is the simplest case – for more than two times, the Histories approach needs an additional necessary consistency condition.[11] Additional assumptions need to be made to go from weights to a joint probability function for three or more events, as can be seen in classical stochastic processes. According to Griffiths, families which do not satisfy this necessary condition are of the kind of 'mysterious instantaneous nonlocal influences', and families that do satisfy it are enough to cover what physicists need to know about systems. Once one has a single framework, a single consistent family of histories, then the sample space is a collection of mutually exclusive possibilities. This sample space is represented by the appropriate PD of the history identity. Only one of these possibilities actually occurs.

The Histories approach solves the first measurement problem of definite outcomes by side-stepping measurement completely and by having indeterminism rather than having unitary time evolution. With the uniwave dethroned, and physicists free to choose a framework where the projectors in the decomposition refer to different pointer positions, there is no problem with macroscopic superpositions. But what of the second measurement problem according to Griffiths (2019) – how is this measurement outcome related to the microstate of the quantum system the apparatus was designed to measure?

The model for measurement in the Histories approach takes a closed system of the measurement apparatus and the system (a particle, for example), $\mathcal{H}_m$ and $\mathcal{H}_s$, interacting in the time t_1 to t_2. Assuming

$$(4.25) \qquad |\psi_0\rangle = \sum_j c_j |s_j\rangle, |M_0\rangle,$$

the unitary time development from t_0 to t_1 to t_2 is given by

---

[11]The details on this condition, and a more detailed look at the Histories approach, can be found in Griffith's *Consistent Quantum Theory* (Griffiths, 2003).

$$(4.26) \qquad |s^j\rangle \otimes |M_0\rangle \longrightarrow |s^j\rangle \otimes |M_1\rangle \longrightarrow |s^1\rangle \otimes |M^j\rangle .$$

The initial state of the particle from equation 4.25 also goes through a unitary time evolution,

$$(4.27) \qquad |\Psi_0\rangle = |\psi_0\rangle \otimes |M_0\rangle \longrightarrow |s^1\rangle \otimes \left(\sum_j c_j |M_j\rangle \right)$$

from t_0 to t_2. Projectors $\{P^j\}$ are chosen to form a PD on $\mathscr{H}_m$ so that

$$(4.28) \qquad P^j |M^j\rangle = |M^j\rangle .$$

After applying the Born Rule,

$$(4.29) \qquad Pr(P^j \text{ at } t_2) = |c_j|^2.$$

With the selection of a framework to make sense of the pointer position, and the dissolution of the first measurement problem, the second problem can now be dealt with. Griffiths states that the experimentalist has designed the apparatus such that a particular pointer position corresponds to a prior state of the particle, P^j and $|s^j\rangle$. Compared to the quantum Histories description of what has occurred, where the description of the family

$$(4.30) \qquad [\Psi_0] \odot \{[s^j]\} \odot \{P^k\},$$

leading to the probability

$$(4.31) \qquad Pr(s^j \text{ at } t_1 \text{ AND } P^k \text{ at } t_2) = \delta_{jk} |c_j|^2,$$

and the conditional probability

$$(4.32) \qquad Pr(s^j \text{ at } t_1 | P^k \text{ at } t_2) = \delta_{jk}.$$

It is clear that if one has P^k at t_2, it follows with certainty that the particle was in the corresponding state $|s^j\rangle$ at t_1.

Once the measurement is over, the state of the particle has no interesting connection with the state of the pointer – as Griffiths writes, the particle's final state has nothing to do with the interpretation of the experiment. All that matters is that its state at t_1 corresponds with the measurement outcome P^j, or $|M^j\rangle$.

The key move made by the Consistent Histories approach is to, in Griffiths' words above, 'dethrone the monarch' that is the uniwave – it becomes one in a sea of many potential choices, as the physicist is free to choose any framework, following the principles of Liberty and Equality. This, according to proponents of the approach, then takes away the issues of definite outcomes and preferred basis.

The framework of the Histories approach can be applied directly to the classical limit. The Hilbert space can be coarse-grained with a suitable PD, where the subspaces correspond to macroscopic descriptions of the world. Coarse-grained projectors can be used to represent macroscopic properties where the subspace is of sufficiently large dimension.[12] A well-chosen quasiclassical projective decomposition (PD) could, for all practical purposes, represent the different possibilities in a macroscopic situation with adequate accuracy. Once a quasiclassical PD has been chosen, a family of quasiclassical histories, or a framework, can be chosen. The consistency condition poses less of a problem than it does with quantum systems and can be satisfied, as the size of the quasiclassical coarse-grainings leaves a lot of room for adjustments to be made where the macroscopic property remains the same, for all practical purposes. With suitable coarse-graining and time-intervals, Griffiths (2013) states that classical dynamics will emerge as an approximation of the quantum dynamics. This can explain why everyday 'classical' logic is so common – our macroscopic world only needs a single quasiclassical framework, and so within this framework ordinary logic applies.[13]

This view of the quasiclassical world doesn't need decoherence in the way that other interpretations do; it is not needed for the classical limit, nor is it needed to solve the measurement

---

[12]Griffiths suggests that the size be 10^v, where v is in itself a large number, such as 10^{10}.

[13]Or, as Griffiths states, all frameworks of this type give the same results for macroscopic properties for all practical purposes.

problems. It can be used in the Histories approach, and simple decoherence models where there is a weak interaction with a suitable environment can be used to render some histories consistent where they may have otherwise been inconsistent.

All interpretations of quantum mechanics need to take decoherence into account. Often, this is to their betterment – for example, Oxonian Everettianism uses decoherence to solve the preferred basis problem. They cannot ignore it because the decoherence program has been enormously influential since its beginning in the 70s, and should not be overlooked, as it deals with the quantum to classical transition in a general way and so opens up the black box of measurement. It is important to go back over the interpretations discussed in chapter 3 and look at their solutions to the measurement problems in the light of decoherence, for how well an interpretation assimilates decoherence provides a way of assessing it based on physics.

INTERPRETATIONS II: IN LIGHT OF DECOHERENCE THEORY

As seen in the review of quantum mechanical thought experiments in chapter 2, the measurement problem and the quantum to classical transition have been foundational problems in quantum theory since its beginning. Throughout the exploration of measurement and the different solutions to the two measurement problems from different interpretations of quantum theory, it has become clear that a big part of this transition is to do with decoherence. As study into decoherence only began in 1970, interpretations of quantum mechanics developed without it. Now, it is an important tool we can apply to the problem of measurement and the quantum to classical transition.

In his thorough examination of whether it is justified to use decoherence theory to explain the quantum to classical transition, Schlosshauer (2005) concludes that while it cannot solve the problem of unique outcomes in the standard interpretation of quantum mechanics, it can offer a very promising picture of how to solve the preferred basis problem, and that it does this via the robustness criterion. Different interpretations use decoherence theory to motivate potential answers to the problem of the preferred basis, though some need it more than others. It is a useful way of pressure-testing an interpretation. How easily they incorporate it into their formalism, as well as whether it helps or hinders their solutions to the measurement problems, are useful checks when evaluating which interpretations are worth future work.

This chapter looks at specific ways in which different interpretations of quantum theory use

decoherence to solve the measurement problems in their interpretation. For a more comprehensive review of how these interpretations of quantum mechanics use decoherence theory to solve the measurement problems by applying it to the quantum to classical transition, see Schlosshauer (2005). Schlosshauer argues that the use of decoherence theory in different quantum mechanical interpretations leaves them empirically identical, and therefore the choice between them as, quoting Tegmark, "purely a matter of taste" (Tegmark, 1998a, p. 855).

5.1 Bohmian Mechanics

Bohmian mechanics solves both the problem of unique outcomes and the problem of preferred basis, starting with making position the preferred basis. While it does not explicitly use decoherence in its formulation, it does not ignore it; according to Goldstein (2017), Bohm wrote about the effects and the importance of decoherence in 1952, though he did not call it 'decoherence' in his work. Despite this, decoherence is not required in Bohmian mechanics to solve the measurement problem. Work has been done in looking at how environmental decoherence could be the motivation for choosing position as the preferred basis, since, as seen in chapter 4, a system interacting with the environment leads to the diagonalisation of the reduced density matrix in position space. As Schlosshauer (2005) points out that there has been much done on how Bohmian mechanics can utilise the theory of decoherence, including work by Appleby (1999), Allori (2002), Bacciagaluppi (2016), Sanz and Borondo (2003), and others.[1] Bohmian trajectories are in general highly non-classical, and applying the theory of environmental decoherence to such cases is a potential way to recover quasiclassical trajectories. This does require certain additional assumptions, however, and is an ongoing project.

Recent work on this has been done by Toroš et al. (2016); they describe how the Bohmian trajectory of a particle is classicalised when the particle continuously interacts with an external agent. This happens as the conditional wave function of the particle beings to follow a dynamics that corresponds to the dynamics of the GRW model (as discussed in section 3.5). Toroš et al. (2016) show that GRW can be derived from the underlying continuous interaction of a quantum system with an external agent, and that these dynamics recover classical trajectories of particles. They

---

[1]For a more detailed list, see Schlosshauer (2005).

state that this makes the GRW collapse model "the course-grained version" of that continuous interaction. This marriage of the GRW model and Bohmian mechanics may help to explain how non-classical Bohmian trajectories can be recovered using decoherence, but it rests in the acceptance of both Bohmian mechanics and the GRW collapse model – something that may make it seem like a less satisfying choice of interpretation, since another interpretation needs to be added on.

5.2 Collapse Theories

As discussed previously, the most successful collapse theory – the CSL model from Pearle (1989) and Ghirardi et al. (1990) – allows the choice of the parameters localisation accuracy d and mean localisation frequency f with the GRW-model choosing $f = 10^{-16}s^{-1}$ and $d = 10^{-5}cm$. There has been recent work on the limits of these parameters within collapse models – some of the most recent of these being Toroš and Bassi (2018) and Vinante et al. (2016).

For collapse models the transition between quantum and classical, or between the microscopic and the macroscopic, is governed by the number of well-localised particles at positions further apart than $10^{-5}cm$ in the two states which have their coherence dynamically suppressed. This transition regime depends heavily on these parameter values d and f. The localisation accuracy parameter d in particular can have very little variation – any smaller and there would be undesirable effects on internal dynamics, any larger and macrosystems would become too inaccurately localised. There is a little more variation allowed of f, though some very large variations have been experimentally falsified (Adler, 2003). Hence collapse models such as the CSL do not need decoherence in order to solve the measurement problem or explain the quantum to classical transition.

However, there are similarities between decoherence theory and GRW; Joos commented in 1987 on the GRW equation for motion of a free particle in 1-d space from their 1986 paper, which in its simplest form is

$$(5.1) \qquad i\frac{\partial \rho(x, x', t)}{\partial t} = \frac{1}{2m}\Big(\frac{\partial^2}{\partial x'^2} - \frac{\partial^2}{\partial x^2}\Big) - i\Lambda(x - x')^2 \rho.$$

The first term on the right hand side corresponds to unitary evolution, and the second term to the destruction of coherence between two different positions. This equation is not unique to the GRW collapse model, however – it is the direct consequence of interaction of a macroscopic system with its environment being taken into account in a global Schrödinger equation, as seen in Joos and Zeh (1985). Joos (1987) comments that there is no need to postulate different fundamental dynamics in order to motivate to obtain the same results than can be gained from environmental-induced decoherence.

Of course, this similarity can also be argued to protect collapse theories from empirical disproof. Though decoherence has the advantage of being directly derived from standard quantum theory, rather than needing a postulated new fundamental law of nature, Schlosshauer believes that there is the potential for a 'fruitful union' between the two theories (Schlosshauer, 2005, p. 31), including some potential proposals for gravity being the collapse-inducing universal 'environment' (Diósi, 1989; Penrose, 1996; Breuer and Petruccione, 1999). Equation 3.32 shows how there may be an option to refer to gravitational effects as causing the suppression of coherence. The question can be posed of how many nucleons should occupy different cells in order for the superposition to be suppressed within the characteristic time of human perceptual processes (Ghirardi, 2018). With the time of $10^{-2}s$ and $f = 10^{-16}s^{-1}$, the number of displaced nucleons is of the order of 10^{18}. This corresponds closely to a Planck mass.

So while collapse theories such as the CSL model do not use decoherence to solve the measurement problem and explain the quantum to classical transition, they nevertheless share many similarities with the decoherence program, and in some places have the exact same formalism. Though they were developed in parallel to decoherence in general, and environment-induced decoherence in particular, rather than within the program, this does not rule out future unification, though currently they seem to be two distinct interpretations of similar formalism.

5.3   Modal Interpretations

Granting that modal interpretations can solve the first measurement problem, they are left with the preferred basis problem. A common answer has been to use decoherence, and it helps the

BDMI/SDMI when they are faced with non-ideal measurements. Decoherence is also used in the MHI implicitly, as the measuring system is considered a macroscopic object with a huge number of degrees of freedom. These play the role of a decohering internal environment by making the bases that generate the definite-valued observables as the ones picked out by the state. Decoherence is then necessary within this interpretation in order to define a measuring system, though it was not explicitly stated originally.

5.4 Oxonian Everettianism

In this version of the Many Worlds interpretation, according to Saunders (1993); Saunders et al. (2010), environment-induced decoherence theory is a dynamical process which leads to two components of the quantum state evolving independently, since the interference terms on the off-diagonal have been reduced to have negligible effects. This occurs due to "high level, emergent consequences of the particular dynamics and initial state of our Universe" (Wallace, 2010, p. 10).

For Wallace, the heart of the modern measurement problem lies in the issue of treating quantum states as probabilistic at the microscopic level. At the macroscopic level, there is no issue with treating states as probabilistic. At the microscopic level, however, doing so leads to issues with interference phenomena. In a double slit experiment set-up, where the particle has the state

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|s_1\rangle + |s_2\rangle), \tag{5.2}$$

there is a 50% chance of the particle going through the first slit or the second. However, rather than seeing a 50/50 split, an interference pattern is seen, since amplitudes have phases as well as magnitude. This means that, at the microlevel, depending on the relative phases, there will actually be a range of outcomes between $100\%|s_1\rangle$ and $100\%|s_2\rangle$. With interference phenomena at the microlevel, one cannot understand the state space of quantum theory as a space of probability distribution, but as a space of physical states. The disconnect between the physical state view at the microlevel and the probabilistic view at the macrolevel, and the transition between the two is, for Wallace (2012), the modern measurement problem. This differs from

Maudlin's characterisation of the measurement problem (which was discussed in section 3.1), which focusses on the three incompatible claims of quantum mechanics rather than specifically the difference between the microworld and the macroworld. This difference is instead implicit in the claim of unique outcomes, rather than a main point.

Different interpretations deal with this problem in various ways – the Many Worlds interpretation chooses to dissolve the problem rather than solve it by taking the quantum state to be always physical. Decoherence plays an important role in how it does this. It uses the environment-induced decoherence framework to measure the quantum system state, causes the system to be decohered by the environment with respect to a basis, and then the probability interpretation can then be applied to that basis.[2] The suppression of off-diagonal terms by the environment leads to the ability to interpret the state probabilistically. Since decoherence is rapid – compared to the dynamical timescales of the dynamics of the system – we can assume that the effective evolution equation of the state reduces to just its diagonal terms.

Wallace (2010) describes the issues of decoherence theory being both too strong and too weak to solve the measurement problem; too strong in it only approximately diagonalises the density operator, doesn't occur instantaneously but over short timescales, and too weak in that it picks out too many system-environment splits and not all are predicted by classical mechanics. Having a necessary criterion be only approximate seems like a lot to ask for when looking for definitions for new fundamental laws of physics. It is also too weak, as decoherence picks out too many bases, which leads to too many options of histories that all give the same classical probabilities. However, Wallace believes this problem can be avoided by treating decoherence as a criterion for quasi-classical emergence rather than as a law of dynamical collapse. The Many Worlds' 'worlds' are established — as quasi-classical, dynamically isolated structures instantiated within the quantum state — by decoherence theory.

5.5 Zurek's Existential Interpretation

The previous interpretations all incorporate decoherence in some fashion into their frameworks. But one interpretation is based almost entirely on decoherence.

---

[2]As seen previously, and in more detail in Joos et al. (2013).

As discussed in chapter 4, Zeh and Zurek were the first to explore decoherence in the 1970s and 80s. However, while Zeh started from the assumption of a universal wave function, and therefore with the foundation of an Everett-type interpretation, Zurek remained more agnostic and was much less disposed to bending decoherence towards one particular interpretation. He did put forward the idea that decoherence could be the bridge between Bohr's Copenhagen interpretation and Everett's Many Worlds in 1991 (Zurek (2003) being the updated version of this), but otherwise kept his discussion of decoherence free of any particular flavour of interpretation. Camilleri (2009*b*) gives this as a reason why Zeh's work was not taken up by the community until a decade later, when Zurek published his own work in the early 80s. But Zurek has not remained as agnostic as he began – as well as bridging the gap between Bohr and Everett, Zurek (1993) also proposed his own interpretation of quantum theory – the Existential interpretation.

Zurek's Existential interpretation takes decoherence, environment-induced superselection, and envariance (also known as environment-assisted invariance) and adds these to standard 'no-collapse' quantum theory (Zurek, 1993, 1998, 2005). These are supposed to allow the existence of unique outcomes despite the superposition principle, and Zurek's interpretation uses them to give us an origin of probability and the emergence of objective existence.

Zurek highlights 'insensitivity to measurement' as a defining feature of classicality. The objectivity of systems is completely in the eye of the observer — a 'relatively objective' existence, as we as observers are limited and can only record states that are not 'censored' by einselection – the quantum state vectors become 'real' when einselection effectively turns off the superposition principle. This happens when there is a transfer of information about selected observables into the environment. This kind of environment-induced objective existence has einselection effectively 'switching off' the superposition principle, allowing quantum state vectors to be 'real'.

Preferred pointer states from decoherence are a necessary precondition for the emergence of a classical world in Zurek's approach. The environment-induced superselection of these preferred pointer states is, according to Zurek (2005), understood through appealing to quantum correlations and envariance.

Zurek (2005) also uses envariance to show how probabilities arise in quantum mechanics. Envariance is a kind of symmetry of composite quantum states. It is when $|\Psi_{SE}\rangle$, the composite

state of quantum system and environment, can be transformed by $U_s = u_s \otimes \mathscr{I}_\epsilon$ acting solely on S,

$$(5.3) \qquad\qquad U_s |\Psi_{SE}\rangle = (u_s \otimes \mathscr{I}_\epsilon)|\Psi_{SE}\rangle = |\eta_{SE}\rangle,$$

but this can be undone by choosing an appropriate $U_\epsilon = \mathscr{I}_s \otimes u_\epsilon$,

$$(5.4) \qquad\qquad U_\epsilon |\eta_{SE}\rangle = (\mathscr{I}_s \otimes u_\epsilon)|\eta_{SE}\rangle = |\Psi_{SE}\rangle.$$

So $|\Psi_{SE}\rangle$ is envariant under U_s. This is an 'assisted' symmetry, and classical states can never be envariant.

Zurek (2005) gives an example using the Schmidt decomposition of $|\Psi_{SE}\rangle$,

$$(5.5) \qquad\qquad |\Psi_{SE}\rangle = \sum_{k=1}^{N} a_k |s_k\rangle |\epsilon_k\rangle,$$

where $\{|s_k\rangle\}$ and $\{|\epsilon_k\rangle\}$ are complex and orthonormal. Any unitary transformation with the Schmidt eigenstates $\{|s_k\rangle\}$,

$$(5.6) \qquad\qquad u_s = \sum_{k=1}^{N} \exp(i\phi_k)|s_k\rangle \langle s_k|,$$

is envariant and can be 'undone' by a countertransformation,

$$(5.7) \qquad\qquad u_\epsilon = \sum_{k=1}^{N} \exp(-i\phi_k + 2\pi l_k)|\epsilon_k\rangle \langle \epsilon_k|,$$

where l_k are arbitrary integers (Zurek, 2005, p. 3). This purely quantum symmetry of SE implies the observer's ignorance about the outcomes of their future measurements on the quantum system S. The emergence of objective properties, as we see in our everyday classical experience, arises as a consequence of the system entangling with the environment. Zurek derives the Born Rule using the envariance but not using the elements which already assume the Born Rule. This derivation is discussed by Schlosshauer and Fine (2003). Zurek (2005) writes that envariance justifies the observer's ignorance in an objective way, without having to appeal to any subjective

lack of knowledge. The observer knows perfectly the composite state of SE. Multiple records of that state are spread throughout the environment, which makes it "operationally objective" (Zurek, 2005, p. 9).

Einselection from decoherence theory is used to solve the preferred basis problem for Zurek's interpretation, through superselection of pointer states. This leaves the first measurement problem, the problem of unique outcomes. The existential interpretation attempts to solve this problem also using einselection – it follows the physical state of the observer of the system by tracking the evolution of their state. A complete description of the observer's state includes the physical state of their memory, and this updates whenever they acquire new records from observation. This is following on from Quantum Darwinism (Zurek, 2009, 2014), where not only are the preferred pointer states the 'fittest', or the most robust, but the environment records many redundant copies of information about these pointer states. The number of copies of these records is a measure of objectivity. The observer decides what observable they will measure, and so this determines the selection of possible events. Pointer states, or record states, will preserve correlation with the recorded state of the quantum system despite of the environment. The observer's memory is effectively classical as einselection makes sure that only records in the pointer basis can be kept. This focus on the relative state of the observer and on correlations has a branch-like structure, as in the various Many Worlds interpretations. An observer observing themselves in a particular pointer state then will assign a state, consistent with their records, to the rest of the Universe. The different objective states correspond to different branches, and these branches are defined by the robust states einselected by the environment. The problem of unique outcomes is solved, in Zurek's interpretation, because the observer and their memory are inseparable, and their memory is a physical state that records the state of the system

It is argued that the existential interpretation is justified by decoherence, as decoherence limits the set of states that can maintain useful correlations with the pointer states.[3]

Considering the above five ψ-ontic interpretations, it appears that all of them benefit from including decoherence in their interpretations. Zurek's Existential interpretation uses decoherence

---

[3]Because of this, as Schlosshauer (2005) notes, this interpretation intrinsically requires the notion of open systems. In an isolated system, if the observer doesn't know the observables that commute with the state of the system and performs a measurement with an observable that does not commute, then the state of the system is reprepared.

as its foundation. Some are able to include decoherence more easily than others; Oxonian Everettianism uses decoherence to solve the preferred basis problem in its interpretation, whereas the modal interpretation uses the formalism of decoherence within its structure implicitly. Bohmian mechanics uses decoherence to justify the position basis as the preferred basis, but needs to work more on the union between its formalism and decoherence, as while it could be beneficial it needs extra assumptions added on. Collapse theories have a different issue; they contain similar formalism to decoherence, but interpret it differently within their framework and might empirically diverge from decoherence theory. When choosing between interpretations, it seems that an interpretation which not only easily includes decoherence theory but can use it to its explicit benefit is a better option. Decoherence is such a critical, experimentally supported theory in physics that not only lends support to an interpretation but deals with the transition between quantum and classical physics – the very part of quantum theory that all the interpretations are attempting to understand.

But these are all ψ-ontic interpretations. Decoherence is often seen to be a physical process, and experiments such as the Hackermüller et al. (2004) seem to strongly support this view. It seems obvious that a physical process would work better with an ontic wave function. QBism, the ψ-epistemic interpretation discussed in chapter 3, does not have the advantages the ψ-ontic interpretations had when incorporating decoherence. So how does QBism deal with decoherence and the quantum to classical transition?

5.6 QBism

As we have seen, decoherence and the quantum to classical transition is a way of pressure-testing an interpretation of quantum mechanics. This section discusses how the Quantum Bayesian interpretation (QBism), a ψ-epistemic interpretation, deals with this transition. This is done first by examining a recent perspective on quantum decoherence from Fuchs and Stacey (2019), and then considering how the interpretation deals with the quantum-classical divide more generally.

As we saw in chapter 3, in the earlier discussion about QBism, it has de Finetti's subjective Bayesianism as its basis, taking to heart the statement in the preface of his book; "Probability

does not exist". Their probabilities — the quantum state — are subjective. This was chosen because, in the words of Fuchs and Stacey, without *personal* quantum state assignments, action-at-a-distance cannot be removed. With personal quantum state assignments, quantum states are not objective, agent-independent, physical properties, and the problem of action-at-a-distance disappears.

This personalist Bayesian formalism also means there are no external criteria for declaring a quantum state right or wrong. However Objective Bayesianism by Jaynes (2003) sounds like it should also work for the QBist — we are all generally in agreement about quantum state assignments, and do sometimes consider there to be a right or wrong assignment in some cases. Nonetheless, QBists argue vehemently for subjective Bayesianism because it requires the same subjectivist extremism that de Finetti employs — that the only rational constraint on prior probabilities is probability coherence. Objective Bayesianism requires rational constraints that uniquely determine prior probabilities in every situation. For example, if there were a bag of white and black dice, and no more information was known about the bag or its contents, a subjectivist of de Finetti's ilk would have no rational constraints on prior probabilities. Objective Bayesians are required by their rationality constraints to have a prior probability of $P(\text{White Die}) = \frac{1}{2}$. This is because they must follow the principle of indifference when assigning their probabilities — without any prior knowledge, they must split their credence equally over all the possible outcomes. In this case, they must assign equal probability for black and white dice.[4]

This lack of rational constraints apart from coherence for prior probabilities is needed for the QBist to fit the theory of decoherence into their interpretation. Fuchs and Schack (2012) start their description of QBism's perspective on decoherence in quantum mechanics by looking at the difference between synchronic and diachronic Dutch Book arguments. Simplistically, synchronic Dutch Books involve wagers you would take all at the same time, whereas diachronic Dutch Books concern wagers you would take at different times. Hacking argued that there is no coherence argument to force an agent to take earlier probabilities into account while setting

---

[4]There is a potential argument that the Quantum Law of Total Probability, discussed previously as being equivalent to the Born Rule, provides a priori restraints on an agent's priors that a subjective Bayesian should not have, as needing $q(i)$ to be a proper probability distribution one must place restrictions on $p(i)$ and $r(j|i)$. However, Fuchs and Schack dismiss these worries by stating that these restrictions do not limit the role prior beliefs play in the agent's assignments of the distributions $p(i)$ and $r(j|i)$.

later ones. While he was writing about synchronic Dutch Book arguments, the same was later said for diachronic Dutch Book arguments by Teller (1973). Synchronic Dutch Book arguments are not suitable for the QBism description of decoherence — for one, they do not imply Bayesian conditioning, and for another, they do not connect probability assignments between times.

Diachronic Dutch Book arguments are therefore the place to start when looking for QBism's perspective on decoherence, as without time passing there can be no decoherence. To show how they connect the probability assignments over time, the Reflection Principle from van Fraassen (1984) is used. The part QBism is concerned about is what a person believes about how his beliefs will change. A decohered quantum state is, in QBist language, an agent's belief about how the state will change over time, since all quantum states are personal probability assignments.

Van Fraassen's Reflection Principle, in its simplest form, is as follows; if an agent assigns a particular subjective probability for a proposition A in the future, then their present subjective probability for that proposition A must equal the future subjective probability.

One way to satisfy this principle to ensure an agent's beliefs are not incoherent is to use Bayesian conditioning to set your future probabilities using your current probabilities. The Reflection Principle can therefore be used to detect potentially disastrous consequences, the famous example of Ulysses and the Sirens being used. Ulysses knows his future probability of making the bad decision to steer his ship towards the sirens upon hearing their song, and therefore also into the rocks. But rather than the Reflection Principle making him endorse this catastrophic decision in Ulysses' present, it allowed him to avoid it by telling his men to bind him to the mast, making his own probability assignments incoherent but also not allowing him to act on them. His crew maintained coherence by stuffing their ears with wax, so that their probability of hearing the sirens became zero. If coherence can't be achieved, then the option for the agent is to not act on their incoherent probability assignments to avoid catastrophic consequences. The Reflection Principle marks these incoherencies so that they can be avoided, though it doesn't supply the way to remove this incoherence. For the agent, Fuchs and Schack write, avoiding incoherence is the goal (Fuchs and Schack, 2012, p. 8).

Fuchs and Schack apply the Reflection Principle and the updating of future probabilities via Bayesian conditioning to quantum decoherence. They take a standard quantum measurement

situation from a QBist perspective, where all the quantum states are the agent's degrees of belief about the future measurement outcomes of the system. Their agent plans to make two measurements; the first is a collection of trace-decreasing completely positive maps $\{\mathscr{F}_i\}$, and the second is a POVM. At time $t = 0$, they has assigned the quantum state ρ_0. The first measurement is at time $t = \tau > 0$, and the second at $t = \tau'$. When asked at $t = 0$ their probability for the state at $t = \tau'$, how should they gamble?

Using the Reflection Principle;

$$(5.8) \qquad\qquad \rho_0' = \sum_i \mathscr{F}_i(\rho_0),$$

where

$$(5.9) \qquad\qquad \mathscr{F}_i(\rho_0) = \Pi_i \rho_0 \Pi_i,$$

where Π_i are projection operators and ρ_0' has the form of a decohered state.

From this QBist perspective, there is no physical process of decoherence, but rather a coherent state assignment at $t = 0$ which imitates a belief in decoherence. The Reflection Principle means that the agent must use the state ρ_0' at $t = 0$ in order to make any decisions about the second measurement. This 'decoherence' takes place even before the first measurement and is a consequence of both the belief of that first measurement, and the application of the Reflection Principle. The physical process is the agent acting on the quantum system, and its consequence is the new state of belief of the agent about the system.

Fuchs and Schack contrast this QBism view of quantum decoherence with the view of quantum decoherence by Zeh, Zurek, Schlosshauer, and others.[5] Their view of decoherence follows the von Neumann view that during measurement, the quantum state entangles with the state of the measurement device and then one of the outcomes is selected. They then add on that the specific interaction with the environment chooses how the joint system-device quantum state decomposes. Fuchs and Schack argue that this 'selection' step is mysterious, and there is

---

[5]As described in more detail in Schlosshauer (2007).

not enough structure to describe the details of the selection. In the QBist version of quantum decoherence, the decohered state is predicated at $t = 0$ in a belief that takes into account the future probabilities for the quantum state at a later time $t = \tau$. There is no mysterious selection, but simply a recognition of future possibilities.

It is therefore key, for QBism's perspective of decoherence, that the probabilities be subjective and personal. Decoherence is the product of an agent using Bayesian conditioning to update their current probabilities in accordance with the Reflection Principle, so as to avoid incoherence. The only rational constraint on the prior probabilities is coherence.

Bacciagaluppi (2014) argues that there could be an ontic view of the quantum state in QBism with subjective probabilities as long as the state is not a probabilistic entity but instead the objective properties of a system. While this clashes with the first tenet of QBism instantly, it also would invalidate the QBism retelling of quantum decoherence. If the quantum state is in any way objective, then it can no longer just represent the agent's states of belief of the quantum system, and then it makes no sense to say that it has decohered before even the first measurement. If QBist decoherence is predicated on a $t = 0$ belief about the possibilities of the future quantum state at a later time, then describing the quantum state as ontic in this way — with objective properties — there is no way to explain why these objective properties would change before any measurement, simply because an agent is conditionalising on their belief about future possibilities in order to maintain coherent probabilities. For this view on quantum decoherence, QBism requires the quantum state to be subjective — otherwise the problem of Wigner's friend once again rears its head. How can Wigner and his friend formulate two contradictory quantum states about a system if the states are objective? While an objective Bayesian could argue that Wigner and his friend could have separate, objectively correct probability distributions, the issue arises if the state needs to be ontic. One would have to be wrong. Fuchs and Stacey argue that there is no mathematical fact in the quantum state that cannot be in a chosen set of probabilities — there is nothing extra about the state that is ontic.

This perspective on decoherence might make it tricky for QBism to interpret some experimental results which deal with the transition between quantum and classical theory. For example, the wave-particle duality of large molecules such as C_{60} and C_{70} experiment performed by Hacker-

müller et al. (2004), as mentioned in chapter 4. As a reminder, it was shown that the interference fringes of the C_{70} molecules gradually lose their visibility as their internal temperature was increased to 3000K.

This experiment seems to show gradual decoherence as a physical process — can QBism's decoherence picture explain it? If one accepts the full QBism perspective of Fuchs and Schack, this experiment can be explained by QBism's decoherence picture. The theory used by Hackermüller et al. (2004) involves changing the fullerene centre-of-mass state by taking a trace of the emitted photon state, leading to a loss of coherence. Here, in accordance with the QBism perspective, the experimentalist is updating the state of the molecule to reflect the future state after the excited molecule emits a photon. This, in accordance with the Reflection Principle, leads to a decohered state. Using decoherence theory to calculate that the fringe visibility is reduced exponentially whenever an emitted photon has a wavelength small enough to resolve path separation. This gives a decoherence curve that fits the experimental results.

QBism's first tenet — that quantum theory is normative, not descriptive — means that nothing about this experiment gets in the way of the QBist's decoherence perspective. The experimenters used quantum theory to calculate the expected decoherence curve, and their experimental results happened to match that calculation. The theory told them how to 'bet', and their bet was well-placed. Nowhere does QBism state why quantum theory is so good at telling us how to bet. But QBism never claims that it can, or at least, that it can *yet*.

The third tenet of QBism is perhaps the most revealing when it comes to where the classical-quantum divide resides in this interpretation. It is here that QBism clearly and decisively draws the distinction between its interpretation and the Copenhagen interpretation of quantum theory. In his defence of QBism, Mermin (2017) highlights this difference. The Copenhagen interpretation describes measurement in quantum mechanics as taking place between a quantum system and a classical system independently of an observer. This is impossible in QBism, as the classical system in question — the measurement device, the pointer, the detector, etc., is indivisible from the agent using it to take a measurement. The split between classical and quantum, in QBism, is specific to each agent, just as their experience and measurement are specific to them. Mermin states that the split, for a QBist, is "between that user's directly perceived internal experience,

and the external world that that user infers from her experience" (Mermin, 2017, p. 9). In fact, Mermin states that the word 'classical' has no place in the QBism interpretation, and can be replaced by 'experience'.

Fuchs and Stacey talk of the quantum–classical divide in their recent work on QBism — although they talk more of the lack of such a divide and transition in the first place.

> "QBism says 'No. Experience is neither classical nor quantum. Experience is experience with a richness that classical physics of any variety could not remotely grasp.' Quantum mechanics is something put on top of raw, unreflected experience." (Fuchs and Stacey, 2016, p. 31)

For Fuchs' and Stacey's QBism, investigating the quantum to classical transition is unnecessary, and even halts progress. Instead, looking into the transition from classical to quantum is more important, as then one can work out how we use quantum theory in our everyday lives, taking our normal everyday measurements, such as seeing and hearing. The idea of an ideal instantaneous measurement is considered on the same spectrum as listening out for a tea trolley (Fuchs and Stacey, 2016, p.31), though only further research will tell us what POVM elements should be written for a measurement like that.

Rather than the emergence of the classical world from the quantum, the QBist must instead look at the emergence of the new and better theories we use to measure the world around us — ones that allow us to see entanglement and quantum interference where previously we only saw classical results. The next step is to be able to give a consistent quantum description to an everyday measurement procedure — such as listening out for a noise.

Mermin (2017, 2018), in his work on QBism, also writes about the quantum–classical divide and transition.

> "We are able to navigate the world better because it is bettable. This is obvious for the laws of quantum mechanics, which are explicitly probabilistic. It also holds for classical physics, though bettability can be obscured in classical physics by a widespread misunderstanding of probabilities that are often very close to zero or one." (Mermin, 2018, p. 8)

Much like in Fuchs and Schack's view, classical 'certainties' are not certainties in the same way as a probability-1 in QBism is not certain in quantum mechanics — it is just an extremely strongly held belief. This, Mermin states, is where the shifty quantum – classical split occurs — there is no place where belief becomes certain for his view. The use of 'objectiveness' is just a way to escape the subjectivity of science. Probabilities are 'subjective personal judgements'. For Mermin's brand of QBism, there is no classical world.

'The roles of classical objects are played by the personal experiences on which an understanding of the world external to the QBist rests. The roles of quantum objects are played by the entities that the QBist hypothesizes to comprise that world on the basis of that experience.' (Mermin, 2018, p. 24)

Because there is no classical world, the classical objects are instead the personal experiences that the user of quantum theory uses to understand the external world. The quantum objects become the entities that the QBist, using their personal experience, hypothesises the world to be made up of. The lack of the usual quantum–classical divide — the objective/subjective distinction — then means that QBism can be applied in a classical setting. Mermin (2014) calls this 'CBism', and uses it to solve the problem of the 'Now'.

The 'Now' is my current state of affairs. Mermin argues that the existence of the 'Now' is undeniably real, as we all experience it. Yet, in physics, there is no way to describe this singular point of 'Now'-ness that seems so evident to us. It cannot be found on a world-line. This, Mermin states, is due to the removal of the physicist from the physics — the exclusion of the scientist — something that, in QBism, is rectified by the intense focus on personal experience.

In an early study of QBism by Timpson (2008), he considers a somewhat weaker view than the QBists have since developed. Using Cartwright's capacities and powers approach in order to provide the ontological structure for QBism, Timpson states that one can have unproblematically stateable truths and generalisations about macroscopic systems. The relations between the (roughly) classical macro and micro cannot be reducible, because it is impossible to have laws about the micro according to QBism. Therefore the relationship between the two can be no

125

stronger than supervenience.[6] It might be even weaker, as it could be that when microsystems come together and interact to form macro-objects, they might possess metaphysically emergent properties which may or may not be consistent with supervenience. For Timpson, Cartwright's capacities and powers view can give QBism the ontological 'furniture' it needs, and also account for there being relations and general laws at the macro (classical) level, with lawlessness beneath (as, he says, can be seen in thermodynamics and kinetic theory). As Cartwright (1999) believes, the world can be made up of a patchwork of laws, and she warns against quantum 'imperialism'. A dappled world composed of many different constrained systems need not worry about consistency.

Timpson describes how, with QBism, there is no 'shifty' classical/quantum split when dealing with quantum systems and apparatuses — there is just the user of quantum theory and the quantum world. The user applies quantum theory to learn about the quantum world, and so can decide which objects are apparatus. Items that get to be apparatuses have stateable facts about them at all times (so we don't have to use quantum theory on them) and are used to probe the objects being treated as systems. There is no issue with moving the split and treating the item as a quantum system later on, Timpson states — the use of quantum theory and the assignment of quantum states does not change any facts about the object, such as the pointer position of a detector. It is this belief in classical, stateable facts that we can ascribe to items such as apparatuses where Timpson's view of QBism seems to diverge from the more recent explanations of QBism. Fuchs and Stacey (2016) treat measurement devices as 'prosthetic hands' of the agent, an integral part of the agent that therefore has no connection to the quantum system it is measuring other than making the action on the system that the agent is using it for. The simple divide of 'agent' and 'world' is much simpler than Timpson's apparatus/system plus user, and once again places the agent in the heart of things. But there are no stateable facts about classical systems in Fuchs and Stacey's QBism; in Timpson's, there are. These stateable facts provide the ontological furniture he believes QBism needs.

Timpson's way of connecting QBism with the world — his way of circumventing the first tenet of QBism — runs into trouble when decoherence is considered. As seen previously, QBism's view of decoherence works with the gradual decoherence of large fullerene molecules, so long

---

[6]This roughly meaning that any change in the macroworld could only happen if there were change in the microworld.

as it says nothing about the workings of the world and remains only a user manual for placing bets. Once ontological furniture is added into the mix this account no longer works. QBism uses subjective Bayesiansism so that the probabilities are degrees of belief, not any kind of general knowledge or information, for that would suggest there is a right and a wrong, and the right corresponds to something physical and objective in the world. QBism claims that quantum theory is not descriptive and so does not describe physical objects and processes in the world. Once ontological furniture is added, such as Timpson's emergent properties and unproblematically stateable truths, this first tenet must be dropped. This has dire consequences for the decoherence picture in QBism. Decoherence spans both quantum and classical in a continuous way that is much more concrete than supervenience, as can be seen in the experiment by Hackermüller et al. (2004). Once the fullerene molecules have decohered and become essentially classical, they are macrosystems that, on Timpson's view, could have unproblematically stateable truths applied to them. Yet, according to the QBist decoherence perspective, there has been no known underlying change — simply, the experimenters have predicted the curve as fringe visibility is lost as the molecules are heated. Can there be a difference between macro and microsystems great enough that one can have truths about it and the other is lawless? It is clear that if one tries to add some kind of ontological furniture to QBism, then its current decoherence picture cannot explain the move between the lawless microworld and the more ordered macroworld.

This is where Timpson's version of QBism misses the mark when compared to the original QBist view of Fuchs, Schack, and Caves. If you follow the QBist doctrine of Fuchs and Schack, then the problem vanishes. There is no divide at all. The *stuff* beneath the phenomena we experience is currently unknowable. Decoherence is just us obeying the Reflection Principle and updating our future probabilities. But this in part depends on what lies beneath the phenomena we see in the world. If one believes, as Fuchs does, that there is an underlying *stuff* of the world, then it seems there will need to be an explanation of the changes in behaviour between the macro and the micro beyond the QBist perspective of decoherence, once the *stuff* of the world is explored. The problem of decoherence is pushed forward to be a future problem.

So what of the fundamental *stuff* of the world? QBism is sure that it cannot be as described by quantum theory. There are a few options for what this fundamental stuff may be. In his talk,

Pusey (2016) suggests that QBism can be divided into three potential approaches; the Minimal quantum world, the Mundane quantum world, and the Jamesian quantum world. The Minimal quantum world is the view that all we can know about the quantum world is that it gives rise to experience, and this experience is such that it is best predicted by quantum theory. This view he hesitantly ascribes to Mermin, where agents (or 'users') are primitive. In Pusey's view, this Minimal quantum world can do no further work — it has said all it can.

The Mundane view — Pusey's own — is that the quantum world is made of some fundamental stuff which behaves according to some fundamental laws, and that when this stuff is arranged to form an agent, quantum theory is the best way for that agent to know how their interactions with other stuff will go. Here, contrary to one main QBism idea, the agent is a derivable concept, not a fundamental one. Pusey (2016), while agreeing with QBism that the quantum theory is a user manual, does not agree that other theories in physics, such as Newtonian physics, also are. These theories, according to Pusey, can be read unproblematically in a realist manner, because such a reading gives the best explanation as to why those theories are so successful and useful. He believes this kind of realist reading cannot work for quantum theory because of the no-go theorems. But just how this works, given the quantum-classical limit, is unclear. Classical theory, recovered from quantum theory in a limit, or even the gradual decoherence of fullerene molecules as mentioned previously, links classical and quantum theories in a way that conflicts with having completely separate interpretations. Following Pusey's view, however, means that something during decoherence has caused the molecule to go from indescribable to describable. Extra work would be needed to explain how the reading of the molecule, as it decoheres from a quantum system to a classical one, gains a realist character. Decoherence would somehow be powerful enough to make real those systems which it affects. Can the fullerene molecules which are not hot enough to decohere be treated as mere subjective experience, when, on Pusey's view, they can become objective once they have been heated to a certain temperature? Once again, we hit the problem of including realist, objective parts in a theory which is built around the concept of being completely subjective. Pusey himself believes that it may be possible to sort out this problem with further work, and suggests one's treatment of 0 and 1 probabilities, but as of now, there seems to be no way to reconcile QBism with anything objective.

Fuchs' quantum world, Pusey states, is a Jamesian one, where there is no fundamental *stuff* as such, and the agent remains front and centre in the theory. Fuchs and Stacey (2016) suggests that though the world does not seem to consist of quantum states, it is maybe quantum measurement, or 'pure experience', that is the stuff of the world (Fuchs and Stacey, 2016, p. 35). Schack (2016) argues in his talk that the world is malleable, and changes when agents act on it. Pusey believes there is more to discover in both his own 'Mundane' quantum world reading of QBism, and in Fuchs' Jamesian quantum world. Fuchs and Stacey also believe that there is a much larger research programme that QBism is merely the beginning of — that the story of agents as gamblers using quantum theory is the start, not the end. This might answer some of the worries of Timpson (2008) after his study of QBism, one of the worries being that it is not clear what the 'ends' are if quantum theory is just a user manual and not descriptive of the world. Fuchs and Stacey state that this research programme has started by applying the lessons learnt from QBism to our other physical theories, such as special relativity and renormalisation in field theory. The next step from there has yet to be revealed.

> "Quantum states, QBism declares, are not the stuff of the world, but quantum measurement might be." (Fuchs and Stacey, 2016, p. 35)

This research programme, while still in its nascence, has many big questions ahead of it. Two of those questions are needed to justify to the realist exactly why we should follow QBism down this path in the first place: why do we keep on finding better ways to 'bet' on our measurement outcomes, and why should we try and ascribe quantum formalism to 'common day' measurement procedures if it does not give us any better results than the theories we already employ and if quantum theory itself is not tracking something in the world? And if the fundamental *stuff* of the world is pure experience, as according to Fuchs' Jamesian quantum world, then can these questions be answered?

Quantum Bayesianism, as it is presented by Fuchs, Schack, and Stacey, can account for the quantum-classical transition. In this completely subjectivist reading of quantum theory, there is no divide between quantum and classical systems, and decoherence is simply the result of an agent using Bayesian conditioning to update their future probabilities via van Fraassen's

Reflection Principle. However, for this account to work, one must also apply this subjectivity not only to quantum theory but to classical theory as well. In doing so, one sacrifices any connection with the world and therefore any explanations our theories give us. To assuage any worries about this huge explanatory loss from a realist perspective, QBism would need to provide some answer to what the fundamental stuff of the world.

QBism can accommodate quantum decoherence and the quantum–classical divide, but in doing so pays a heavy price for those wanting a realist interpretation; its complete subjectivity gives not only quantum but also classical theory no connection with the external world and therefore they cannot give explanations for anything. If we attempt to keep some objectivity and some connection with the world, as Timpson and Bacciagaluppi argue for, then the QBist perspective on quantum decoherence and on the quantum-classical divide no longer works. In fact, the QBist view of Fuchs and Schack is based entirely on the idea that quantum theory cannot explain the external world, or at least not in the way that a realist would want. QBism as an interpretation cannot have even a hint of objectivity, as the very idea is anathema to the interpretation's foundations. To account for the quantum-classical transition in QBism, or simply for the QBist interpretation more generally, it is subjectivist or bust.

In this chapter the interpretations introduced in chapter 3 were revisited in the light of decoherence theory, and a new interpretation was discussed. It is clear that every ψ-ontic interpretation mentioned here can either use decoherence theory to its benefit, or contains something similar to decoherence theory within its formalism already. In agreement with Schlosshauer (2005), decoherence theory in most cases can be employed to help solve the preferred basis problem. However, when looking at QBism – a ψ-epistemic interpretation – decoherence theory is treated very differently. Decoherence is not a physical process and decoherence theory does not help explain the quantum to classical transition on this view, so this counts against QBism to some extent. But QBism has no stock in the idea of quantum theory explaining anything anyway, and so is set apart from the other interpretations.

There is more to learn about the quantum to classical transition outside of the measurement problem. The next chapter looks more closely at the quantum to classical transition in general, and how we can understand it. This is done by assessing another ψ-epistemic interpretation and

a potential analogy for the quantum to classical transition from statistical mechanics.

AN ANALOGY FOR THE QUANTUM TO CLASSICAL TRANSITION

6.1 Introduction

Chapters 3 to 5 focused on the measurement problem in quantum theory and how it is solved by different interpretations. Decoherence provides both a possible solution to the preferred basis problem and a way of pressure-testing these interpretations. The measurement problem relates to the issue of the transition between quantum theory and classical theory. Chapter 4 discusses how decoherence can be used to explain some instances of this transition, but it is a complex process not yet fully understood. This chapter looks at a potential way to understand the quantum to classical transition – is it analogous to a phase transition in statistical mechanics? Though it is found that this is not a good analogy to use when looking at the quantum to classical transition, much can be learnt from its specific failings.

This analogy was developed by Bub (2016).[1] Bub (2011) suggests that information can be interpreted as changing its physical structure when transitioning between the quantum and classical realms, and that this change occurs in the infinite limit. This leads to an interesting analogy between phase transitions in statistical mechanics and the quantum to classical transition that

---

[1]This analogy is only found in the first edition of this book, and private correspondence confirms this has later been taken out. Despite this, the analogy is still worth exploring in the hope of learning important features of the quantum to classical transition.

he uses to explain the necessity of this infinite limit as an idealisation. This chapter examines this analogy and explores how far it can be taken by reconstructing it in detail. It is shown to be a formal analogy and that as an argument by analogy it is unsound, due to the falsity of two of its premises. Section 6.5.3 looks particularly at the analogy's failings in regards to semiclassical mechanics, as represented by Alisa Bokulich (2008*b*) and other experimental examples. Section 6.6 argues that despite its failings this analogy is useful to highlight just how rich and complex the quantum to classical transition is, and its worthiness of further investigation. Despite the use of 'fictions', Bokulich (2016, 2018) uses her Eikonic conception of explanation to show how idealisations and simplified representations of physical systems, such as semiclassical mechanics, can still increase our understanding.

6.2   The Information-Theoretic Interpretation

At the end of his book, Bub (2016) collates all the previous chapters and puts forward his Information-Theoretic (IT) interpretation as a potential solution to the conceptual problems in quantum theory. He divides the measurement problem into a 'big' and a 'small' problem, and employs two different arguments by analogy in order to describe how the IT interpretation dissolves these problems.

Bub (2016) writes that the quantum to classical transition is analogous to a phase transition as both need an infinite limit. He uses this infinite limit in order to solve what he calls the 'small' measurement problem, in the sense of the definitions from Pitowsky (2006):

- THE BIG MEASUREMENT PROBLEM: the 'ordinary' measurement problem as debated in the field of quantum theory, where unitary evolution is inconsistent with the collapse of the wave function, leading to no unique outcomes.

- THE SMALL MEASUREMENT PROBLEM: how quantum uncertainty becomes classical ignorance in the classical limit, or decoherence and the quantum-classical transition.

These two measurement problems are similar to the two discussed in chapter 3 – the first is clearly the same as the unique outcomes problem, while the second is more specifically about

the quantum to classical transition, rather than the preferred basis problem. Like QBism, the IT interpretation is a ψ-epistemic interpretation, and thus rejects the idea of the reality of the wave function. While he has previously gone into detail on how the 'big' measurement problem is avoided by the IT interpretation championed by Bub and Pitowsky (2010), it was highlighted by Timpson (2010) that there was little talk of the transition between classical ignorance and quantum uncertainty. In Bananaworld, however, Bub offers an analogy with phase transitions in classical statistical mechanics as a potential answer to the small problem. This chapter will not be focusing on the actual formalism of Bub's IT interpretation, but examining in closer detail the short argument by analogy. This analogy is no longer used in more recent publications of Bananaworld — however, by examining this analogy in greater detail there is much to be learnt about the quantum to classical transition, and so it is reconstructed in this chapter.[2]

The transition between quantum and classical measurements — or the jump between the question "what will be there when I measure it?" to "what's there?" — occurs, according to Bub's IT interpretation, when the number of microsystems making up a macrosystem is so large as to be idealised as infinite. This then leads to commuting Boolean algebras which can be associated with macro-observables with definite values, as in classical physics. In a microsystem, or in only a small number of microsystems, the non-commuting observables lead to the structure of quantum mechanics. This transition leads to "what's there" in Bub's words, and that it is "like a phase transition".[3] In the IT interpretation, macrosystems must be described as infinite quantum systems. As Bub states, such "unavoidable" idealisations are commonly used in classical statistical mechanics for phase transitions.

The analogy is not as straightforward as it first appears. The subtleties of phase transitions in classical statistical mechanics have been discussed thoroughly in the literature. Is this argument by analogy a viable one? And does it help describe the quantum to classical transition?

Section 6.3 discusses the necessary background of what constitutes a material or formal analogy, using the definitions from Hesse (1963) of a material and formal analogy and the more recent discussion by Bailer-Jones (2009). Section 6.4 first goes over, in very broad strokes, the technicalities of phase transitions in statistical mechanics, and then quantum-classical transition

---

[2] Private correspondence with Bub confirms that he no longer defends this analogy.
[3] As was found in Bub (2016), though this has now been taken out.

itself is discussed to see whether the similarities between the two allow for a material analogy.

After finding the material analogy wanting, section 6.5 reconstructs the phase transition analogy as a formal analogy, before putting it to two tests — the role of the infinite limit, and the problem of semiclassical mechanics. Section 6.6 concludes by discussing the general effectiveness of the phase transition analogy, and looks at the broader debate and the future prospectus for the quantum to classical transition.

The quantum to classical transition has a lot of interesting phenomena surrounding it which have yet to be as thoroughly philosophically examined as they deserve, as Bokulich (2008*b*) discusses. With all the debate surrounding quantum decoherence and quantum theory, it is clear that more closely examining the transition between our classical world and the quantum realm, a much more complex area than the phase transition analogy can account for, could reveal new information about quantum theory, which could go on to inform further work in the area.

6.3   The Phase Transition Analogy: Material or Formal?

Before looking closer at the phase transition analogy, it is first worth working out what kind of analogy it is and its intended purpose.

Bailer-Jones (2009) explores the use of analogies in science and physics. Analogies have also been linked with scientific modelling to the point where models have been confused with analogies and vice versa. Analogies allow us to link relatively disparate branches of science, and are extremely useful in arguments. Historic scientists have all had their own strong opinions on analogies and what they give us: analogies between unexplained phenomena and explained phenomena lead to new hypotheses and experiments (Herschel), analogies are tools to aid understanding (Maxwell), analogies are essential, *foundational*, to theory formulation (Campbell).

The main theme through all discussions of the use of analogies, as Bailer-Jones highlights, is that they generate knowledge or help us gain deeper knowledge about systems and phenomena.

Analogies are seen as being linked to the explanatory function of models, focusing on new discoveries and predictions. As Bailer-Jones writes, they can be a catalyst to aid modelling, taking relations within one domain and applying them to another. They also can be used in arguments –

arguments by analogy are, according to Bailer-Jones, a lot like thought experiments.

Hesse (1963) described the two types of analogy to be found in physical theories: formal analogies and material analogies. For Hesse, if there is a one-to-one correspondence between the "different interpretations of the same formal theory" it is a formal analogy (Hesse, 1963, p. 75). An example would be if two domains share the same mathematical form.[4]

Material analogies, on the other hand, have observable similarities "which enable predictions to be made from a model" (Hesse, 1963, p. 75). These material analogies are pre-theoretic, and are connected by similarity relations. Hesse uses the example of the wing of a bird and the fin of a fish. While these two appendages are not identical — they do not share the same purpose, structure, or composition — they nevertheless perform a similar 'flapping' motion and act to propel their respective bodies through a medium. Bailer-Jones (2009) writes that these analogies are more based on 'visual properties', and are often more distant than formal analogies. This idea of visualisation links closely to the visualisation in thought experiments, as explored in chapter 2.

Regardless of these types of analogy, one important feature of an analogy is that the model it generates or supports must live up to describing the phenomena in question. According to Bailer-Jones (2009), the analogy itself is not the model. Which of these types of analogy does the phase transition analogy belong to? There are arguments to be made for each, which the next section explores.

6.4 Material Analogy?

6.4.1 Phase Transitions

If material analogies share observable similarities, then it is important to consider both phase transitions and the quantum to classical transition separately. Phase transitions are found everywhere in physics — and despite being a common phenomenon, they are hard to define. The first attempt to classify the types of phase transitions was done by Ehrenfest (1933), after a new type of transition was discovered in liquid helium.[5] A strange 'jump' discontinuity was seen in the

---

[4]This is nomic isomorphism, which is a special case of formal analogy.
[5]This is discussed in Jaeger (1998).

temperature dependence of the specific heat of helium at a critical value.[6] Ehrenfest then began to classify phase transitions by using the jumps in the derivatives in the free energy function – whether this jump was in the first or second order derivatives. While the classifications have, since the 1930s, been adapted and changed with each new discovery, the two distinct types of first or second order phase transition are still used today as a classification.

More recent definitions from Binney et al. (1992) state that first order phase transitions involve latent heat, whereas all other phase transitions are generally called continuous phase transitions, and involve a discontinuity in the rate of change of a property or properties rather than a discontinuity in the properties themselves. First order phase transitions often reduce to a continuous phase transition at a critical point, although not all continuous phase transitions have an associated first order phase transition.[7]

Continuous phase transitions have an order parameter (usually denoted as ϕ), which has a thermal average of zero in one phase and non-zero in the other. A particular order parameter is specific to each system. This break in symmetry is a feature of phase transitions. Common features of phase transitions are therefore a discontinuity either in properties or the rate of change of properties, with an order parameter specific to the system, and some kind of symmetry breaking.

In thermodynamics, a phase transition is a discontinuity in a property, or a rate of change of a property. In statistical mechanics, the more fundamental theory, phase transitions are modelled as non-analyticities in the free energy of the system. Unfortunately, a system with a finite number of particles has no non-analyticities — hence the need of the thermodynamic limit, where the number of particles $N \to \infty$, the volume $V \to \infty$, and the particle density $N/V = $ const.

The phase transition analogy may relieve any discomfort with the necessity of infinite idealisations in the quantum to classical transition, but from the outset it does little to assuage any potential concerns. In the wider literature on reducing thermodynamics to statistical mechanics, it is not clear that the infinite limit needed for phase transitions is unavoidable or unproblematic. The thermodynamic limit is brought in to solve this problem, as discussed by Callender (2001),

---

[6]This jump was named the 'lambda' jump, on account of its graph having a shape similar to the Greek letter λ.

[7]An example of this kind of continuous phase transition, as mentioned by Binney et al. (1992), is the ferromagnetic phase transition undergone by iron in low temperatures.

and provides a non-vanishing partition function, but this needs an infinite idealisation.

6.4.2 Quantum to Classical Transitions

While classical mechanics is a limiting case of quantum mechanics, Landau and Lifshitz (2013) writes that it is nevertheless unusual in the respect that it is impossible to formulate quantum mechanics without first having classical. This special relation sets quantum mechanics apart from other physical theories, and there is no place this is shown more than in semiclassical mechanics. Bokulich (2008*b*) details the shared structure of classical and quantum theory and describes how efforts to find the classical limit have been not fully satisfactory. Reducing the classical limit to large quantum numbers oversimplifies the problem. To jump between quantum uncertainty and classical ignorance, one cannot simply reach a point where suddenly the number of microsystems is enough that it can be idealised as infinite, and allow commuting observables. There is a whole mesoscopic realm between macro and micro, in which quantum systems can be approximated to a startling precision with classical methods. Semiclassical mechanics is used often when the calculations are not computable by quantum approaches.

An important equation in regards to semiclassical mechanics is the Wentzel-Kramers-Brillouin (WKB) approximation, also known as the quasi-classical approximation. According to Karnakov and Krainov (2013), the WKB approximation is one of the most important approximation methods in quantum mechanics. This approximation expands the static solution to the one-dimension Schrödinger equation as a series in powers of $\hbar$. In the classical regime, where $\hbar \to 0$, all terms of $\hbar^2$ are neglected. $\hbar \to 0$ is not an accurate way of describing the limit, as $\hbar$ is a constant and cannot go to 0 – instead quantum mechanics goes to classical when $\hbar$ is negligible. It is more accurate to state that "one considers the limit of a dimensionless quantity formed by the ratio of Planck's constant to some other quantity of the same dimensions" (Bokulich, 2008*b*, p. 14). The term with $\hbar$ will vanish in the limit and give the classical equation. This, of course, makes it very system-specific — it is not one size fits all — and it is often a singular limit. A singular limit causes behaviour approaching the limit to be very different to behaviour at the limit, which means there cannot be a straightforward reduction of quantum to classical, or indeed of thermodynamics to statistical mechanics, as discussed by Berry et al. (1991); Berry (1995),

Batterman (1995, 2001), and Butterfield (2011*b*).

There are further problems of the $\hbar \to 0$ limit when it comes to classical chaos. The structure of classically chaotic systems only emerges in the long time limit $t \to \infty$. Unfortunately, it turns out that the two limits of $\hbar \to 0$ and $t \to \infty$ do not commute, leading to different results depending on which limit is taken first.

There are a few other contenders for the classical limit of quantum physics. A common one is the limit of large quantum numbers, or $N \to \infty$. N is usually the principle quantum number, describing the energy eigenvalues of the system. This limit also has its problems, as discussed by Liboff (1984); not all quantum systems depend on N, and thus would not go to classical in the limit, and even some systems which do depend on N do not go to classical behaviour in the limit.

There is the option of combining both the $\hbar \to 0$ and $N \to \infty$ limits. Hassoun and Kobe (1989) use this successfully for three quantum mechanical systems; the harmonic oscillator, the hydrogen atom, and the particle in a box. These are well known systems and simple to calculate, however — whether this combined limit would work with problem cases such as the chaotic helium atom has not yet been explored.

6.4.3   A Material Analogy?

The analogy between phase transitions and the quantum to classical transition can be considered a material analogy if one looks at the observable similarities between two examples of each domain. What does this mean for the physics of a macrosystem? It might suggest that when the number of microsystems is large enough as to be idealised as infinite, and their unitary time evolution at some point, at a random time, becomes non-unitary. Or it might be that there is a group decoherence event at a random point as the number of microsystems goes to infinity.[8] It would align with a critical point in a phase transition, but unlike phase transitions, for the quantum to classical transition there is no definable order parameter to mark this point. There is no derivative of a quantity which is discontinuous.

One easy example of a phase transition in statistical mechanics is the transition of water from liquid to gas. In a kettle, water boils when heat energy is added. This added heat causes the

---

[8]As discussed in chapter 4, decoherence occurs when the microsystems interact with the environment and are no longer dynamically isolated, and their states become entangled with the state of the environment.

water to transition from its liquid phase to its gaseous phase. For the phase transition analogy to be a material analogy, there should be observable similarities between the two systems. The first task is to find an experiment which models the quantum to classical transition in the same way that water in a kettle models a liquid-gas phase transition.

This can be seen in the experimental example of Hackermüller et al. (2004), which was discussed in chapter 4. As a reminder, the wave-particle duality of even large molecules such as C_{60} and C_{70} has been explored since it was first shown that it is not only electrons that give interference patterns when passed through two slits by Arndt et al. (1999). The experiment by Hackermüller et al. (2004) shows how increasing the temperature means that the molecules which were previously well-isolated from the environment have a higher internal energy. This leads to emitted and absorbed radiation, scattering, collisions and other such processes, which shows the molecules passing through one of the two slits, and leading to decoherence. It was shown that C_{70} molecules transition between quantum and classical behaviour gradually – the interference fringes losing their visibility – as their internal temperature was increased to 3000K.

A good material analogy should have some kind of explanatory power. This phase transition analogy needs to share enough similarities with the target system in order to be a suitable argument by analogy, but it must also stand alone as a well-understood system itself. Hesse (1963) took Aristotle's general criteria for evaluating analogical arguments, and these are summarised clearly by Bartha (2016). The most important ones in regards to strength of the phase transition analogy are as follows;

1. The more observable similarities between the domains, the stronger the analogy.

2. Those analogies involving structural similarities are stronger than those involving other, superficial, similarities.

3. The analogy is weaker if not much is known about either domain.

The analogy should satisfy these criteria enough that it has the predictive power required from it. As this analogy is used for explanatory purposes rather than predictive purposes, it will be called the "explanatory power" required. Using the previous example, it can be seen that the

analogy would suggest that, since an infinite limit is used in the water transitioning from one phase to another when thermodynamics is reduced to statistical mechanics, an infinite limit is used in the case of the C_{60} and C_{70} molecules transitioning from quantum to classical behaviour. This must be due to their observable similarities for it to be a material analogy. However, no infinite limit is needed in the case of the C_{60} and C_{70} molecules.

What if one takes the degrees of freedom to be the microsystems 'N' in question? As the molecule is heated the number of 'N' increases, and an infinite limit could eventually be used. As the molecule is heated and it begins to emit thermally, the number of modes increases as it radiates thermal photons. If microsystems could be defined so as to include degrees of freedom such as vibrational modes and emitted photons, then does 'N' go to infinity?

It appears that the answer is still no. Hackermüller et al. (2004) describe the three mechanisms by which a C_{70} molecule cools down: emission of a thermal photon, emission of an electron, or emission of a C_2 molecule. The second two mechanisms are discounted, as the recoil caused by these emissions is so great that the molecule would miss the detector and therefore would not contribute to the fringe. Furthermore, C_{70} molecules have 204 vibrational modes with a finite heat capacity, and need only emit around two or three photons on average to cause the interference fringe to fade. There is no need for the idealisation of the infinite limit for this decoherence and appearance of classical behaviour, therefore showing the analogy to be unhelpful in this case.

Also, apart from this example, phase transitions and the quantum to classical transition show almost no observable similarities. A phase transition in statistical mechanics is just that, a 'transition' between phases. An infinite limit (the thermodynamic limit) is necessary for a transition, such as ice melting, but once all the ice is water, it is no longer needed. For this analogy to be more useful, there would be a 'classical' phase and a 'quantum' phase, with an infinite idealisation needed to get between the two, rather than a characteristic of the classical 'phase'. To have classical and quantum theory separated in such a way would be a very contentious view, and it makes no sense to say that the number of microsystems is no longer infinite after the transition. In the IT quantum world, the infinite idealisation is necessary not only for the transition between quantum and classical probabilities, but for the entirety of the classical world.

The IT quantum to classical phase transition is not a true transition, but stuck in the idealised transition, never to reach the other phase. As a material analogy, the phase transition analogy is useless and has little to no explanatory power.

6.5 Formal Analogy?

6.5.1 A Formal Analogy?

As the phase transition analogy as a material analogy does not work, it seems better approached as a formal analogy. As a formal argument from analogy, it can be set up as follows:

P1. The quantum-classical transition in the IT interpretation can only be understood using an infinite limit.

P2. Phase transitions in statistical mechanics can be understood using an infinite limit.

P3. The infinite limit for phase transitions in statistical mechanics is both explanatorily unproblematic and unavoidable.

C. Therefore we should treat the infinite limit for the quantum-classical transition as explanatorily unproblematic and unavoidable.

There are three arguments against this analogy — one against the first premise, one against the third, and one against its validity. First, its validity: as it is written above, this argument from analogy is invalid, as it has a missing premise – that the infinite limit for phase transitions is the same as the one for the quantum to classical transition. Rather than put this missing premise in my reconstruction, I will assume that it is implicit, or supposed. The infinite limit is the shared mathematical structure the two transitions share, which is what allows for this analogy in the first place.

The problems with premise three and the wider literature on phase transitions are discussed next. Then the problems with the first premise are explored, and it is shown how it is not compatible with semiclassical mechanics.

6.5.2 The Infinite Limit

For a phase transition in thermodynamics to be correctly represented in statistical mechanics, it must be described as a non-analyticity in the free energy. Unfortunately, all the functions in statistical mechanics are analytic; no phase transition can be defined using the first definition (refer to Griffiths (1972) for a more detailed proof). It was shown by Onsager (1944) that the phase transition can be described in infinite particle systems by using an Ising model with no external magnetic field; this works in the thermodynamic limit, where both the number of particles and volume of the system go to infinity and the density ($\rho = \frac{N}{V}$) is fixed. Renormalisation group theory can also be used to describe thermodynamics, and also uses the Ising model. This theory, like statistical mechanics, requires an infinite limit. This could possibly be explained away by saying that while the definition of phase transition in thermodynamics is that it is a singularity in the partition function, it does not need to share that definition in statistical mechanics. According to Callender (2001), this problem arises from 'taking thermodynamics too seriously', and that we need not have a direct translation between thermodynamics and statistical mechanics. In this case the infinite limit is potentially avoidable.

Batterman (2001, 2005) argues that real physical systems do have "genuine physical discontinuities" and that to fully understand these we must use asymptotic reasoning and singularities. In this view, the infinite limit is unavoidable and explanatorily important. On the other side of the debate, Butterfield argues that the N $\rightarrow \infty$ is not physically real as emergent behaviours can occur before the limit, when N is large but still finite. There are also 'crossover' phenomena where these behaviours appear before the limit is reached, but the system can be altered so that this behaviour vanishes in finite N. Butterfield endorses this proposal from Mainwood (2006); we can say that phase transitions can occur in a finite system with N degrees of freedom and free energy F_N and a thermodynamic limit $F_N \rightarrow F_\infty$, if F_∞ has non-analycities. The $\hbar \rightarrow 0$ limit, analogously, does not need to be physically real. Instead Butterfield argues that the physically real behaviour is the weaker sense of emergence at a finite N.

There are, of course, conditions for this 'Butterfield Principle', as Palacios (2018) discusses. She summarises the infinite limit debate, and concludes that the use of infinite limits is pragmatically justified by the fact that, as Butterfield writes, the limit behaviour emerges before the limit.

Palacios revises Butterfield's principle by using the notion of infinite limits as controllable approximations – that is, as mathematical models that approximate the behaviour of a real finite system where the behaviour at the limit emerges approximately before the limit, and for realistic values for the system.

Recent consensus in the debate seems to put the infinite limit firmly into the camp of an approximation – as Norton (2012) writes, thermodynamics requires approximations rather than idealisations. The difference between them is subtle; Norton (2012) writes that approximations are just inexact descriptions of the target system, after all, and therefore give the necessary results to match what is observed while just making a few alterations. Idealisations, however, are more akin to thought experiments – they are the use of another system which has properties that provide an inexact description of the target system. If the infinite limit is a mathematical model, required as a sensible substitute for a real physical system, this suggests that it is not a real physical phenomenon in the world and that it is potentially, in the future, avoidable. This also suggests that it is unproblematic – as long as we remember that it is only a substitute for the real finite system by tailoring it to fit with the behaviour and values of that real system.

The phase transition analogy is based on the similarity of the necessary infinite limit. Both classical mechanics and thermodynamics reduce to quantum and statistical mechanics respectively, with the addition of an infinite limit (at least for the IT interpretation of quantum mechanics). In the literature on statistical mechanics, it is not clear that this infinite limit is either unavoidable or explanatorily unproblematic. Even for Batterman, while it is important, it is not necessarily unproblematic. Regardless, if the thermodynamic limit is an approximation, a substitute, rather than a real physical phenomenon, then for the analogy to be satisfactory then it would have to be similar in the quantum to classical case.

6.5.3 The Problem of Semiclassical Mechanics

Premise one of the reconstructed phase transition analogy rests on the assumption that, as the IT interpretation states, the quantum to classical transition needs an infinite limit in order to have commuting observables. However, this ignores the strange phenomena which take place around the quantum to classical limit — the mesoscopic regime — which can only be understood

fully by using semiclassical mechanics.[9] As quantum mechanics cannot yet explain the behaviour in the mesoscopic regime by itself, classical theory must be used to fill the gaps. At any transition between the quantum world and the classical world, semiclassical effects currently require classical theory to be explained.

The IT interpretation describes a system having its number of constituent microsystems going to infinity as becoming a macrosystem with classical behaviour. This is used to explain how non-commuting observables in quantum mechanics transition to commuting macro-observables in classical macrosystems. As we have seen, this limit is far from being the answer to the long debate over the classical limit. Semiclassical mechanics uses classical methods to explain quantum phenomena such as quantum chaos; something that should not be possible. Semiclassical mechanics is extremely useful, not only in explaining phenomena in quantum systems that has no other theoretical backing, but also, according to Heller and Tomsovic (1993), "semiclassical ideas are better able to support and suggest new experiments" (Heller and Tomsovic, 1993, p. 39), giving us more physical insight into quantum mechanics and, importantly, is much easier for us to conceptualise. There are processes and phenomena that appear in both classical and quantum domains, and — at the moment — can only be explained by combining the two.

There is one key bit of experimental evidence which cripples the phase transition analogy. It relies on the number of microsystems being large enough, relative in scale to the system or phenomenon of interest, to be comfortably idealised as infinite, in order to give commuting macro-observables — as is seen for phase transitions in statistical mechanics. The diffraction of large C_{60} and C_{70} molecules has already been discussed previously. There is another damning example of a system where this analogy cannot work — the Rydberg atom.

A Rydberg atom is one where the outer electron is excited to a very high energy level. This highly excited state gives the atom extremely large dimensions, almost to the size of a fine grain of sand. This is no longer precisely a microsystem, and yet it does not have a huge number of microsystems in order to satisfy the IT interpretation's condition for commuting observables. Not only is the Rydberg atom seemingly between the micro and the macro — in the mesoscopic realm

---

[9]Semiclassical mechanics does not only happen in this mesoscopic regime, but occur anywhere in an empirically relevant region — where part of the system is treated as quantum and the other classical. For example, one can get semiclassical effects in black holes - which are decidedly not mesoscopic.

— it also shows strange effects that cannot be explained by quantum mechanics alone.

While they behave as expected when there are no external fields present, it was shown by Garton and Tomkins (1969) that a Rydberg atom in a strong magnetic field has energy absorption peaks in its spectrum beyond the ionisation energy. The experiment was first done with highly excited barium atoms. These 'anomalous' resonances in the spectrum turned out to have a spacing independent of the type of atom, and were named 'quasi-Landau' resonances. Bokulich (2008*b*) notes that almost twenty years later there was still no theoretical explanation for why the atom in a magnetic field should show oscillations in its spectrum beyond the ionisation energy. Then it was found by Welch et al. (1989) that a higher resolution of the anomalous spectra showed irregular patterns, and that there was a link between these resonances and classical chaos.

It was discovered by Main et al. (1986) that taking a Fourier transform of these irregular spectra yielded a orderly set of peaks in the time domain. These peaks, strangely enough, corresponded exactly to the "periods, or transit times, of the classically allowed closed orbits for the electron moving in the combined Coulomb and magnetic fields" (Bokulich, 2008*b*, p. 117). In the mesoscopic realm of the Rydberg atom, quantum theory is not enough to describe it; classical theory must also be brought in to explain phenomena beyond the scope of current quantum mechanics.

This Rydberg atom, a mesoscopic system, does not have the required huge number of microsystems necessary to have classical observables in the IT interpretation — and yet to fully describe it, both classical and quantum theory must be used. It resides in that nebulous regime between an infinitely idealised macrosystem and a quantum microsystem. It is a finite system, and therefore, according to the IT interpretation it is still a quantum one with no need for classical theory. However, it in fact requires classical theory to partly explain its features and does not need an infinite limit for this. There is no sharp transition between quantum and classical mechanics in the actual world; there is no sudden step from one to the other. In truth, there is a whole realm in which they are overlapped — whether or not the classical methods are considered fictional or not, there is still apparent, unexpected classical behaviour. As it is, in semiclassical mechanics, classical theory is brought in to plug the gaps that quantum theory cannot fill. The phase transition analogy cannot account for this. The Rydberg atom is therefore direct evidence

against the phase transition analogy – it shows classical behaviour despite not being able to be idealised as having infinite microsystems.

One objection to the argument from semiclassical mechanics is that while classical theory is employed to explain the phenomena now, if classical theory is eventually eliminable, there will no longer be this problem of classical behaviour in mesoscopic systems, as it could be explained by quantum theory. If semiclassical mechanics were so important and theory-breaking in the way that it would always need to be used in mesoscopic systems, it would affect other quantum theory interpretations. In this way, one does not need to deny the semiclassical behaviour, but merely state that it will eventually be explicable without having to refer to the then potentially extinct classical theory. This of course assumes that in the future we will be able to describe semiclassical behaviour using quantum theory, which is not possible at this moment in time. If there comes a time when quantum theory can be used to explain phenomena where for now classical theory is used, then the semiclassical behaviour will no longer be a problem for this analogy. But in the current time, where semiclassical phenomena cannot be explained without recourse to classical theory, then this analogy fails. Even if this objection to semiclassical mechanics showing the analogy to be lacking the required explanatory power is taken seriously, it does not counteract the previous argument using the C_{70} decoherence experiment.

Semiclassical mechanics shows that the quantum to classical transition can not be simplified to a transition using an infinite limit. To use an argument from analogy, as Bailer-Jones (2009) writes, the model must accurately represent the phenomena. This phase transition analogy, modelling the quantum to classical transition using an infinite limit, does not completely and accurately represent all the phenomena at the transition.

6.6 The Infinite Limit Idealisation

While this chapter focuses mainly on examples which highlight the weaknesses of the phase transition analogy, as first presented by Bub (2016) and reconstructed in detail in section 6.5, all the many cases where it is true that classical macro-observables are reached by summing over an infinite number of microsystems have been ignored — not everything is a Rydberg atom. The

analogy would quite clearly work at some level as a formal analogy in these cases.

This, however, relies on an infinite idealisation (and thus the analogy with classical statistical physics) being sufficient but not necessary for the emergence of classical behaviour. This is a line one could take to defend the analogy, though it can be considered an instrumentalist use of the infinite limit and would not work for the Information-theoretic interpretation. In the Information-theoretic interpretation the structure of information is a physical thing in the world and so the change in the structure must also be, and therefore the use of the infinite idealisation should be necessary to explain the emergence of classical behaviour and observables, rather than simply sufficient. The emergence here, using the distinction from Butterfield (2011a) for types of emergence, would be ontological emergence. It seems clear, then, that when semiclassical mechanics and the decoherence of C_{70} molecules are considered and the analogy fails, the use of the infinite limit becomes an instrumentalist tool.

Quantum mechanics being an instrumentalist theory would also fly in the face of quantum mechanics as a principle theory – a 'top down' theory comprised of a set of laws, principles, or axioms, such as Special Relativity or Thermodynamics – something that those championing IT interpretations, like Clifton et al. (2003), want quantum mechanics to emulate. Others, however, such as Brown and Timpson (2006), think it is a mistake to look for a small number of postulates or axioms to cover the whole of quantum theory. The opposite of a principle theory is a constructive theory – a 'bottom up' approach that involves invisible foundations and processes. It is interesting, then, that the phase transition analogy requires the infinite limit as used in statistical mechanics – considered a constructive theory – when the IT approach aligns quantum mechanics more closely with a principle theory such as Thermodynamics. As Brown and Timpson (2006) argue, it is a mistake to try and reformulate quantum mechanics as a principle theory, as this move sacrifices much of the 'messier' insight a more constructive theory can provide. Perhaps, as seen in this discussion of the phase transition analogy, the quantum to classical transition is indeed messy, like statistical mechanics – and yet, by looking more closely at the mess rather than only viewing the cleaner, more idealistic infinite limit, there is much more one can learn about it.

Though the Information-theoretic interpretation needs the infinite limit for there to be a physical change in structure, there is a way for other interpretations without such any such prior

149

commitments to view the infinite limit as an idealisation without having to be instrumentalist. The issue with the infinite limit being an idealisation is that such idealisations do not seem to have the same level of explanatory power as explanations which are "full-bodied" things, an argument discussed briefly above. This is how Craver (2014) describes an explanation – that they, or rather the objects and phenomena they involve, should be fully objective and exist in the world.[10] This links back to the discussions of realism in regards to the various interpretations of quantum mechanics from chapters 3 and 5. The interpretations which aimed to explain the quantum-classical transition – all of them except for QBism – based their explanations on picking out real, objective properties in the world and their interactions. For example, Bohmian mechanics picks the trajectories of particles as being ontic and quantum mechanics as describing these trajectories. Bokulich (2018) calls this the 'ontic conception' of explanation, and she argues that this is "fundamentally misguided". Looking at semiclassical mechanics, it is clear that it involves phenomena which are not as full-bodied as the ontic conception of explanation would require. Even the status of the mesoscopic regime, where the semiclassical phenomena take place, is hardly ontic enough to satisfy, as it is hoped that quantum mechanics will eventually be able to explain such phenomena without the need of the 'fictional' classical theory. The mesoscopic regime is in itself a fictional tool to describe the realm between the two theories that will eventually be rendered unnecessary. Why should we even consider the mesoscopic regime when discussing the quantum-classical transition, if it is a fiction?

Bokulich (2018, 2016) puts forward her Eikonic conception of explanation in opposition of the ontic conception as one in a long tradition of epistemic accounts of explanation. She argues that explanations in science are an epistemic activity, and that scientists very rarely study phenomena in full, but instead from a certain lens or view where a more simplified representation is used. Examples of using this simplified representation of a system can be seen throughout section 2.2 in the review of quantum mechanical thought experiments – systems are considered to be isolated so as to ignore any interaction that is not directly important to the theory. Bokulich uses the example of how an electron is viewed, depending on which field or programme using its representation; an electron can be a classical particle, a quantum particle, or even an excitation

---

[10]This is also supported by Salmon et al. (1989).

of a quantum field, depending on whether the scientist looking at the electron is doing so with classical theory, quantum theory, or quantum field theory. Of course, this is not to say that an electron stops being an excitation of the quantum field, or a quantum particle, if looked at with a classical lens, but that those aspects of it are not useful in the context of classical theory. In fact, Bokulich (2018) argues that full-bodied 'ontic' objects in the world are never actually in explanations, but rather their representations are.

This Eikonic conception of explanation allows much more leeway when using idealisations such as the infinite limit. Bokulich (2016) argues that resistance to using 'fictions' such as idealisations comes from the ontic conception of explanation, and that fictional models of systems can give us genuine explanations. They look only at relevant parts of the system and remove the irrelevant aspects, rather than considering the whole system. For Bokulich, the one goal of explanations is to give us more understanding. Though a model or idealisation may be a fiction, it can nevertheless increase our understanding of a system – if it is a good representation.[11] All the semiclassical examples above in section 6.5.3 are technically fictions, as classical theory has been shown to be a approximation of quantum mechanics in certain limits. Assuming that quantum mechanics will one day be able to fully explain the transition between the microworld and the macroworld, the mesoscopic regime in this case is also a fiction. But these fictions are still necessary in certain fields to explain phenomena that cannot be explained by quantum theory alone. Removing the classical theory would lead to a loss of understanding. As seen above in section 6.5.3, without applying the classical theory of electron orbits to a Rydberg atom there would still be no explanation for the anomalous spectra. Though classical theory has already been proven to be a 'fiction', this does not remove the importance of the semiclassical explanation for Bolukich – she argues that not all explanations must be ontological, but those like semiclassical explanations can instead offer new information about patterns of dependence, highlight important features, and increase factive understanding. For Bokulich, explanations are fundamentally a sort of representation.

This does not help the Information-theoretic interpretation when it comes to the quantum to classical transition, as it needs the infinite limit to be necessary for there to be a physical change

---

[11]Bokulich (2016) writes that a fictional mechanism is a good representation or model if it captures relevant structural features of the true mechanism, or the patterns of counterfactual dependence of the true mechanism.

in structure. However, this does link to a previous argument discussed at the end of section 6.5.3 against semiclassical mechanics. Semiclassical mechanics being eventually eliminable and therefore being fictional and having little explanatory power does not imply it should therefore be abandoned given Bokulich's Eikonic conception of explanation. The thermodynamic limit from statistical mechanics, as well, can give us a better understanding without having to be ontological.

This could be argued to cause problems for those seeking a realist interpretation of quantum mechanics, or simply a realist picture in general. QBism, of course, is unaffected by this, as the QBist view does not expect quantum mechanics to be explaining the world anyway. For the other interpretations, however, there is more of an issue. Using such fictions as semiclassical mechanics and the mesoscopic regime in explanations does not fit into a realist picture if one is looking for explanations guided by objects and properties which are real and in the world, as the other interpretations are. However, when it comes to quantum mechanics, there are still so many of these practically equivalent interpretations to choose from in terms of describing the quantum to classical transition. Very little is lost by, in essence, hedging your bets, and keeping such fictions allows us to keep explanations we would otherwise have to be rid of. As Bokulich argues, these explanations which use fictions increase our knowledge overall, and in order to learn more about quantum mechanics and the quantum to classical transition it seems important that we have as much information and understanding as possible.

6.7 Conclusion

This chapter took the idea from Bub (2016) of considering the classical limit as an idealisation of infinite microsystems, and compares the 'transition' from quantum uncertainty to classical ignorance with a phase transition in thermodynamics and statistical mechanics. Reconstructing this as an formal argument by analogy shows that while the two transitions may have some similarities, the analogy does not have the explanatory power required to solve the big measurement problem. This reconstructed phase transition analogy is a formal analogy that has two false premises. The most important failing is that the analogy cannot account for the combination of

classical and quantum theory in the mesoscopic realm, where the two theories overlap. However, the analogy does work up to a point — just not universally. By making this analogy sharper, this chapter reveals that the quantum to classical transition is much more complex than just classical emergence in the infinite limit. The idea of explanatory fictions from Bokulich and her Eikonic conception of explanation show that idealisations and simplified representations such as semiclassical mechanics and infinite limits can further our understanding of science.

The quantum–classical limit is far from fully explored. As seen in this chapter, there is the potential to learn more about quantum theory, infinite limits, the importance of semiclassical mechanics as used in physics today, and about how we should formulate quantum mechanics as a theory. Following on from this, chapter 7 looks at the problems of interpretations and realism caused by phenomena in quantum theory beyond ordinary quantum mechanics. These phenomena involve the thermodynamic limit which also appears in quantum statistical mechanics, and it leads to a whole host of problems to do with inequivalent interpretations being used for different physical situations. This then leads to issues with realism — how can we be realist about our best theories if we have to use unitarily inequivalent representations for certain phenomena? Chapter 7 looks at new approaches developed specifically to deal with these issues and evaluates and amends a new kind of scientific realism.

INTERPRETATIONS, QUANTUM FIELD THEORY, AND REALISM.

Introduction

So far, this thesis has explored different solutions to the measurement problem (or problems) from different interpretations of quantum theory, considering both ψ-ontic and ψ-epistemic interpretations. In chapter 6, the Information-theoretic interpretation is discussed alongside the quantum to classical transition potentially using an infinite limit, analogous to a phase transition in statistical mechanics. However, it turns out that this analogy is not suitable when discussing the quantum to classical transition. This is not the end of the story for the combination of infinite limits and interpretations of quantum theory.

It was seen previously that not all of the interpretations considered can be readily applied beyond non-relativistic quantum mechanics. In fact, they have yet to be expanded well enough to cover all known quantum phenomena. This leads to different interpretations and representations being used for specific situations, and arguably no one interpretation or representation of quantum theory is applicable to all situations. Ruetsche (2011) defends this view using quantum field theory as an example. Ruetsche also considers how quantum statistical mechanics is taken to the thermodynamic limit, as it then falls outside of ordinary quantum statistical mechanics — this means that unitarily inequivalent representations of the microphysics involved are available

to use, as ordinary quantum statistical mechanics cannot deal with the symmetry-breaking equilibrium states. This inequivalence is where the problem lies. Two theories are physically equivalent if and only if the set of possible worlds according to the first theory is the same as the set possible according to the second. If two theories are inequivalent, then there will be some worlds possible in one but not in the other. Yet, as is clear in Ruetsche's example, we must use inequivalent interpretations of theories in order to account for different quantum phenomena. Not only must we use different interpretations for different physical situations, but the ones used cannot ever be combined due to their inequivalence. This stymies the effort to find one single interpretation of quantum theory. This is a particular issue with one view of quantum field theory – it is one of our most successful empirical theories, and yet there is no single interpretation that can account for it, and the different representations are unitarily inequivalent.

Ruetsche uses her own Coalesced Structures approach to help deal with this issue. Rather than taking one single interpretation of quantum statistical mechanics, Ruetsche's approach allows one to choose from a selection of inequivalent representations to deal with specific situations. But this casual accretion or discarding of theory content relative to the phenomena of interest leads to issues with realism. How can we be realist about our best scientific theories if we take up content with one hand and cast some away with the other? Ruetsche asks what 'good' structural realists are supposed to believe in now? This chapter assesses Ruetsche's approach to this realism issue.

Fraser (2017, 2018*a,b*, 2020) and Williams (2017) propose a kind of 'selective realism', which both uses the good qualities of quantum field theories and attempts to negate the bad. This new kind of realism (called 'Effective' realism by Williams, and Renormalisation Group (RG) realism by Fraser) gives criteria for how one can be a realist about QFT, which is one of our most successful scientific theories, if not the most successful, since its accuracy is better than anything except maybe General Relativity. There a few problems with this fledgling view, as have already been discussed by Fraser (2018*a,b*) and Ruetsche (2018, 2020). By their own admission, none of the authors above have done enough to distinguish Effective/RG realism from empiricism. This chapter argues, after exploring different kinds of realism that could be used to save Effective/RG realism from empiricism, that the way to do this is to treat it as a form of ontic structural realism.

The combination of these new realism positions together with ontic structural realism will then be shown to answer Ruestche's question of what good structural realists are supposed to believe in – the methods used by Effective/RG realism pick out the robust entities and structure from QFT that structural realists are supposed to be realist about.

7.1 Quantum Field Theories and Realism

Quantum Field theories have recently been the focus of some philosophers of physics when discussing scientific realism. Although these theories exhibit all the hallmarks of a theory rich for interpretation for a scientific realist – incredibly empirically successful and producing very accurate predictions – they come with a set of characteristics generally considered anathema for a realist view.

Ruetsche (2011) argues that no single interpretation of quantum mechanics can cover the many theories needed to properly describe quantum phenomena. She specifically uses quantum field theory (QFT) to highlight her point. In QFT, there are many occasions where inequivalent representations must be used in order to explain many phenomena, such as symmetry breaking or ergodicity. Ruetsche puts forward her 'Coalescence Approach' to theory interpretation to account for this. Explanatory success is a virtue, she argues, and the best way to be explanatorily successful is to have all available representations at one's disposal – especially for physicists who use those theories and their interpretations. This is very similar to the argument from Bokulich (2018) as discussed in chapter 6, where in her Eikonic conception of explanation representations of the systems being discussed do not have to be objective, full-bodied things in the world. This allows explanatory pluralism, and different representations of the same system in different contexts and programmes, which raises problems for realism.

French (2012) calls this Ruetsche's 'Swiss Army Knife Approach', where one can pick and choose the theory and interpretation relevant to the particular system and situation one is investigating. In this way, these interpretations remain relevant to current physics, and allow for greater explanatory success. This is perhaps impractical, but Ruetsche acknowledges this and believes that it has a healthy moral – that science is still under construction.

As a self-described exercise in liberatory particularism, Ruetsche (2011) wants to show that, in essence, what Rawls concludes about moral theory is both correct and relevant to more than just moral theory. Rawls, she writes, argued that there can be no general theory until each specific area included in moral theory, such as epistemology and philosophy of mind, have their own problems resolved.

In this way, Ruetsche believes that there can be no general, unified theory of quantum mechanics until the problems in each inequivalent interpretation of quantum mechanics are solved. Gathering quantum field theory (QFT) and quantum statistical mechanics in the thermodynamic limit under the term QM_∞, Ruetsche aims to take Rawls' argument and apply it to interpretations of these theories, with the goal of showing that no general interpretation is possible before the smaller, more specific problems are dealt with. Those questions asked in 'lofty abstraction' – such as "what are the laws of nature?" – can only be answered, Ruetsche writes, once the particulars are dealt with.

The particulars in question are the many inequivalent interpretations of those theories under the QM_∞ title. Ruetsche warns of the dangers of choosing particular theories and interpretations from lofty heights. Interpretations of theories are expected to provide predictions, explanations, and further understanding. But if one single interpretation is chosen from this lofty general abstraction, Ruetsche states that the not all of these virtues may be accomplished, and a successful theory may not have all virtues under one interpretation – such as is seen in QM_∞.

In the last chapter of her book, where Ruetsche takes the discussion of QM_∞ to bear on more general problems in philosophy of science, she argues that her QM_∞ contains theories which underdetermine their own interpretation. This, she writes, leads to 'disintegration of virtue' – the overall explanatory burden is shared by inequivalent interpretations. There is also a 'disunity of virtue' – where an individual virtue is not supported by one single pristine interpretation.[1]

The 'rampant' issues — disintegration and disunity of virtue – found in QM_∞ lead to the failure of arguments for scientific realism in the case of quantum mechanics, according to Ruetsche. Despite this stumbling block for scientific realism, Ruetsche believes that we should be pragmatic when it comes to theories in QM_∞. Following Earman's strong empiricism, Ruetsche writes that

---

[1]Where by 'pristine' Ruetsche is talking of the ideal of pristine interpretation, where the full extent of possible worlds are only those that are possible according to a theory T and its laws.

the Coalesced Structures approach 'naturalises' physical possibility and also pragmatises it. This line of thought — the naturalisation of physical possibility — aligns closely with the view of structural realism. Structural realism is a position that states our best scientific theories are approximately true and should be taken to be literally talking about and successfully referring to unobservable entities and processes in the world. Ruetsche describes structural realism, after Worrall (1989), as a position that supports both the No Miracles Argument and Pessimistic Meta-induction. The No Miracles Argument comes from Putnam et al. (1975), where he argues that scientific realism is the only way to make the success of scientific theories non-miraculous. Pessimistic meta-induction, conversely, argues that since successful historical scientific theories have since been proven false, there is nothing to suggest our current successful theories won't also be eventually found to be false.[2] Scientific realism, however, as Ruetsche states from Worrall, is 'the best of both worlds'. The structural realist believes in the realism of form or structure that continues through our best theories, therefore avoiding the pessimistic meta-induction argument, and also looks for empirical success to avoid the No Miracles argument.

But, Ruetsche argues, what are 'good Structural Realists supposed to believe in now?' (Ruetsche, 2011, p. 348), if different and rival interpretations of a theory's structure possess virtues such as explanation or prediction? This, Ruetsche argues, sets her approach apart from the structural realist position – the Coalesced Structures approach does not even need continuity of structure, as theories contained in QM_∞ cannot be attributed to a single type of structure. Unlike in the example from Worrall that Ruetsche uses to describe continuity of structure – the move between Fresnel's theory to Maxwell's – she states that in QM_∞, there exists no such comparison between C* algebras and concrete von Neumann algebras. Something similar to this is argued by French (2012). In his case study of algebraic quantum field theory (AQFT) the inequivalent representations of the net of algebras yields the superselection rules. The structural realist here cannot choose which structure is 'real' and which is 'surplus'. A proponent of Ruetsche's Coalesced Structures approach, however, does not have this problem.

These new approaches towards scientific realism are motivated by the thorny problems posed by quantum field theory and the complexity and variety of modern physics as a whole. Ruetsche

---

[2]This argument was first made by Laudan (1981).

compares her Coalesced Structures approach to two established structural realism positions, Epistemic Structural Realism (ESR) and Ontic Structural Realism (OSR). But is the coalesced structures, or the swiss army knife approach, and other such arguments, so very different from structural realism at all? The best reasons for belief in theories are reasons for belief in their structure. However, no continuity of structure with C* leads Ruetsche to put forward that not even this continuity of structure remains between separate interpretations of QM_∞. As well as this, Ruetsche's case against the Coalesced Structures approach being like structural realism includes the fact that structural realism doesn't tell us what to believe in *now* (what Ruetsche calls the Content Question). While it tells us what structure has survived through theory change, e.g. Fresnel to Maxwell, it doesn't help us choose which interpretation to use at the present time. Ruetsche states that the Coalesced Structures approach tells us to use whichever interpretation gives the most benefit for each specific phenomena, and this provides a varied breeding ground and leads to greater adaptability for the theory in the future. One of Ruetsche's examples of where phenomena diverges from standard quantum mechanical interpretations is that of when quantum statistical mechanics (QSM) is taken to the thermodynamic limit. This is done to explain phase structure, as 'ordinary' QSM does not have the discontinuities required to do this – these discontinuities only happen in an infinite limit, such as the thermodynamic limit. Going to the thermodynamic limit, however – taking the number of microsystems to infinity – is considered to be an idealisation, something that a backer of extremist interpretations would argue you don't want in your interpretation. If you take an 'extremist' interpretation then in Ruetsche's view you lose this explanation and do not have as much explanatory success as the Coalesced Structures approach would have. Ruetsche asks what the good structural realist is supposed to believe in now. How can we be realist about structures if these structures seem to change with each thorny quantum phenomena problem? This issue has led to new kinds of realism founded specifically to answer this question.

7.2 Effective Field Theories

As well as Ruetsche's Coalesced Structures Approach (or the Swiss Army Knife, in French's words), there are other kinds of realism which attempt to explain the multiplicity of theories and interpretations which run rampant in modern science, and nowhere more prominently than in quantum field theory.

A new kind of scientific realism, Effective or Renormalisation Group (RG) realism, has recently been put forward by Williams (2017) and Fraser (2017, 2018*b*). This realist position is motivated by perturbative quantum field theories and effective field theories, which are incredibly successful empirical theories that, due to their limited applicability and degradation over short length/high energy scales, do not fit so comfortably into usual foundational discussions. Rather than put these effective field theories aside as not much more than mathematical artefacts, there has instead been a few efforts to use them as the building blocks of a new kind of approach to scientific realism specific to the QFT context.

Fraser (2018*b*) states that QFT meets the empirical criteria for scientific realism. QFTs have some of the most accurate predictions in history. Yet when it comes to the actual interpretation of these QFTs, the prospectus does not look so good. Parts of QFTs lack any clear characterisation of their physical content, making it hard to justify the mathematical techniques used to give such accurate results (as shown by Fraser (2018*b*) in regard to quantisation). This makes it hard to extract an ontology from QFTs. They also lack some explanatory power; certain mathematical moves made in perturbative calculations, as well as normalisation, are completely *ad hoc*, done for no reason other than to provide good predictions.

Enter 'Effective' or 'Renormalisation Group' realism. These new context-dependent realism views use the technique of renormalisation group analysis on the effective field theories in order to find what quantities we should be realist about. By characterising these quantities, the realist can start to worry less about the lack of epistemic content in QFTs. The RG analysis selects the 'robust' quantities that represent the ontological features of the theories. This kind of selective realism uses renormalisation group analysis as the tool to pick out the parts of the QFTs we should be realist about.

This context-dependent realism – Effective or Renormalisation Group – shows that there is potential in future realist views of QFTs. But this new kind of realism comes with its own problems (as elucidated by Fraser (2018*a*,*b*) and Ruetsche (2018)), but also one major step still needs to be taken. As Fraser writes, this realism is only a small step from empiricism – and in fact it can be argued that it is currently indiscernible from empiricism. Using RG analysis is the best hope they have in fitting QFTs into the traditional realist view. More work, of course, needs to be done. The robust entities picked out by the analysis need to be precisely characterised in a realist framework. This important first step is to show exactly where this 'realism' comes from, to properly draw the line between this new realism and empiricism – otherwise, any progress will stall. Section 7.3 lays out the Effective/RG realism positions, and then section 7.4 explores this issue of where the realism for Effective/RG realism comes from and section 7.5 explores what kind of realism this is.

7.3   QFTs and Renormalisation Group Theory

QFT was presented back by Redhead (1982) as a cautionary tale for philosophers, where he warns about the dangers of dismissing or abandoning programmes too quickly when faced with inconsistencies and computational issues blocking novel predictions. QFT started off life as a technical programme for QED (Quantum Electrodynamics), with problems with divergences and infinities. These problems were patched up in the late 1940s by renormalisation theory, but even then QFT failed to work with both the strong and weak forces. Later work with gauge theories solved these problems, and QFT evolved from apparent technical artifice to one of our most successful theories. Redhead wrote that though its conceptual foundation was somewhat obscure, something that has not changed much over the years, it is nevertheless important to look at what QFT has to say about reality, and he was concerned with the additional interpretive problems QFT brought with it.

When QFT was first starting to be used, there arose difficulties from the many mathematical problems associated with using it, as Li (2015) describes, such as divergences in perturbative expansions. Perturbation theory is an approximation method, used when a problem cannot

be solved exactly. This accounts for almost all problems in quantum mechanics apart from the few solvable ones (such as a harmonic oscillator or a ground state hydrogen atom). To approximately solve these problems, the basic approach is to choose an operator for which the eigenproblem can be solved exactly, and then use the basis of eigenfunctions of this operator to expand the Hamiltonian of the system we want to solve. In its most basic form, perturbation theory involves dividing the Hamiltonian of the system into two; the solvable part, and the rest — the perturbation. In more complicated problems, described by Fraser (2017), the perturbations can lead to integrals with infinite divergences.

Renormalisation was used to solve these divergences in the late 1940s, and while these techniques worked, nobody completely understood why. To interpret this thorny new area of physics, there were those who turned to axiomatic QFT in the 1950s – a subset of which, algebraic QFT, is now the more common name for such QFTs according to Wallace (2006) – in order to clear up the interpretative difficulties. Proponents of axiomatic QFT include Fraser (2009) and Baker (2009). Axiomatic – or the now more generally used 'algebraic' – (A)QFT was not without its problems either. While the idea was to start from some necessary and mathematically rigorous axioms, and then find QFTs which satisfied these axioms, it was soon found that very few of the QFTs used by physicists actually satisfied the AQFT axioms. As Wallace (2006) writes, the only known QFTs which do satisfy the axioms in four dimensions are interaction-free, and therefore not compatible with the QFTs currently in use. AQFT is still used in foundational discussions despite not being the theory used in mainstream physics.

In opposition to the axiomatic theorists, some began to support 'constructive' (or 'conventional') QFT, a decade after axiomatic QFT began to be used (Li, 2015, p. 3). CQFT was created in an attempt to mathematically justify perturbative QFT. Wallace, however, states that constructive QFT that uses the RG (or Lagrangian QFT) is not actually in opposition to AQFT at all, but that one could have a successful AQFT program with an exact, fully defined theory at all scales, and still define effective QFTs with cut-offs. This seems correct, but as there is not yet a well-defined, successful AQFT, constructive theorists believe that there is more to be gained from looking at EFTs that have more empirical success, and that are used in mainstream physics. If one supports AQFT then one believes it to be the best guess, and see it as a prediction that an actual quantum

theory in the future will satisfy these axioms and accurately describe modern physics. This could be the case – the final form of quantum theory could very well satisfy the axioms used by AQFT theorists, but there are currently reasons to remain agnostic.

Li (2015) writes that the Renormalisation Group (RG) was first developed unrigorously within perturbative QFT as a way to provide an account of the change in dynamics of QFTs over different length and energy scales.

> "These changes are manifested as changes in the value of the coupling parameters in a theory's Lagrangian." (Li, 2015, p. 3)

The RG describes the perturbative renormalisation non-perturbatively, and so can be used to justify the use of perturbative QFT.

Renormalisation has recently been described as one of the great ideas and successes of twentieth-century physics by Butterfield and Bouatta (2015), as well as being a strong influence on how physicists think about theories. Renormalisation was so successful and influential because it didn't apply to just one specific area of physics — not only was it used in QFT, but it was also significant in advances in many other fields. In their book, Benfatto and Gallavotti (1995) give many examples of problems all over physics to which Renormalisation group theory can be applied, including; the KAM (Kolmogorov-Arnold-Moser) theory of Hamiltonian stability, universality theory of the critical point in statistical mechanics, and the Kondo problem for magnetic impurities in nonmagnetic metals. This last problem, as discussed in an article by Fisher (1998), "was widely advertised as a significant, major issue in solid state physics" (Fisher, 1998, p. 656). It was solved using the new technique of Renormalisation Group theory. So not only was RG theory a massive advancement in QFT circles, it also could be applied – with remarkable success – to many other, often completely disparate areas of physics. Though it could potentially be described as an "instrument or computational device" (Fisher, 1998, p. 653), this is nevertheless a tool which can be applied to a remarkably wide range of theories. It is a general way of solving the problem of theories which do not apply over all length scales, as the exact cut-off of length scale used in the technique is unimportant.

RG methods in perturbative quantum field theory provide a physical picture as to why the coupling parameters have to be changed in order to avoid the divergences which caused so much trouble early on in QFT's use. They also explain the empirical success of perturbative QFT – something which previously had been hard to determine. While those who support axiomatic QFT, such as D. Fraser, believe that RG methods elucidate the empirical content of quantum field theory and link the empirical and theoretical content, they do not believe that it has any foundational significance. Others, such as Li and constructive field theorists, do believe that RG methods have foundational significance, as they provide a method for proving which models of quantum field theory that satisfy the OS axioms exist, and which do not.[3]

The use of the RG for a realist interpretation has been used by Wallace (2006), Williams (2017), and Fraser (2018*b*). The two most recent works on the RG method and its potential for use in the QFT realism debate, and the wider scientific realism debate, suggest a new kind of realism based solely on QFT and RG methods.

7.3.1 Williams' Effective Realism

The highly successful perturbative QFTs have motivated a new kind of realism when interpreting physical, particularly quantum, field theories. This new movement, called 'Effective Realism' by Williams (2017), and later called 'Renormalisation Group Realism' by Fraser (2018*b*) and Ruetsche (2018), uses the specific quirks of perturbative QFTs in order to guide their search for what to be realist about.

Williams (2017) argues that standard realism fails to properly interpret modern quantum field theories, and that this "leads philosophers astray". Instead, he details a new approach to interpreting these theories, like Effective Field Theories, which is more tailored to the specific context of quantum field theory. He compares his 'Effective Realism' to the standard way of interpreting physical theories and applies them both to the effective field theories that are so successful in modern physics. These approximate, effective theories — they are very accurate at their specific length or energy scale, but break down rapidly at shorter distances or higher

---

[3]The Osterwalder-Schrader, or Wightman, axioms, "specify the properties that a theory's Wightman functions or Schwinger functions, respectively, must satisfy to define a QFT" (Li, 2015, p. 3). Any observable can be computed from these functions, which is what makes them important enough to be built into the axioms.

energies — are incredibly empirically successful, and therefore are ripe for interpretation. And yet, on the Standard Account of Theory Interpretation, as laid out by Williams, these EFTs do not tell us what to be realist about. The Standard Account requires theories to be true and complete descriptions of the world in all respects – and theories such as EFTs, which are only reliable at specific length scales and are only approximately true, do not satisfy this condition. The Standard Account of Theory Interpretation also requires a theory to be mathematically rigorous – something which 'effective' field theories are not. They break down over short distances and give only approximations. When using the Standard Interpretation, these very empirically successful theories are declared unfit for duty.

As Williams writes, we want our philosophical commitments about the world to be grounded in our most successful physical theories – the naturalistic approach to scientific realism. With EFTs, while they fulfil the criteria for being empirically successful, we cannot interpret them if we stick with the Standard Account. This requires interpreters to resort to toy quantum field theories, involving 'remote counterfactual worlds' rather than using the EFTs which are used for empirical applications. These distant theories are further divorced from the actual world, which weakens any realism commitments one might get from their interpretation.

Williams' 'Effective Realism' offers another way of interpretation. It lacks the core commandments of the Standard Account – one must give up having a true and complete description of the world in all respects, and there is no motivation to search for a deeper, fundamental ontology, as EFTs cannot offer reliable information of the world at fundamental length scales. One must also give up the wish for mathematical rigour. We also cannot interpret these EFTs in isolation – since EFTs are reliable only in their specified length scales, we need more than just one for our interpretation.

So what do we get in return for giving up these previously considered core tenets for theory interpretation? How can we know what to be realist about when the physical theories we interpret are unable to give us a true, complete description of our world?

The payoffs, according to Williams, are many. 'Effective Realism' is not without pedigree – it fits with the *divide et impera* approach to realism as described by Psillos (2005). There may not be any fundamental ontology worth finding, but there are certain 'robust' entities, properties,

and relations, which are accessible via a few different ways, the most general and common being through Renormalisation Group (RG) analysis.[4] These 'robust' entities are, using RG analysis, defined by being invariant under the RG flow. These robust features of the theories are what we should take into our ontology, as genuine representations of the world with real physical content.

Effective Realism outplays the Standard Account by allowing EFTs, our successful empirical theories, to discharge their scientific duties, as demanded for by Ruetsche (2011). Not only that, but EFTs 'announce' their foibles – they are only approximate, literally as it says on the can. They do not hide their weaknesses. They make no claim to be a complete description of the world in all respects. They come with their own warning label, which Effective Realism takes into account. Analysis of different EFTs can be used to evaluate the many roles played by different elements and therefore their physical significance – the *divide et impera* approach of selective realism.

RG theory shows that the 'fundamental' short-distance structure in an EFT is not as important as its physical content in its empirically viable domain. Rather than digging for the fundamental ontology, it is more useful to look at the robust parts of the theory in order to clarify the difference between them and the mere mathematical artefacts. These 'higher-level' entities are necessary in order for a QFT to discharge its scientific duty (as mentioned earlier), as otherwise explanations are impossible. Williams gives the example of QCD; the behaviour of hadrons in the LHC cannot be explained only by the behaviour of the fundamental level of interacting quarks. These higher-level, robust entities are also insulated against short-distance changes.

7.3.2 Fraser's Renormalisation Group Realism

Fraser (2017, 2018*b*) also discusses the details of realism when applied in the specific context of quantum field theory. For Fraser, 'Renormalisation Group Realism' is the most promising strategy for working towards a realist interpretation of quantum field theory.

He, like Williams, is focusing on realist interpretations of quantum field theories because these theories are very empirically successful. To be realist about quantum field theories using RG Realism, one uses selective realism via RG theory, and identifies the robust entities from the

---

[4]Williams uses Wimsatt's definition of robust here to include in one's ontology entities which are 'robust', they should be "accessible (detectable, measurable, derivable, definable, producible, or the like) in a variety of independent ways" (Wimsatt, 2007, p. 95).

mathematical artefacts. These quantities are preserved by RG coarse-graining transformations and encode the long-distance structure of a QFT-model. But potential issues towards being realist about these 'effective', or cutoff quantum field theories, are that they are neither mathematically rigorous enough, nor explanatory.

Ruetsche (2018), while believing that Renormalisation Group Realism is right, or at least on the right track, brings up two potential objections. These lead to RG realism not completely disarming the sceptical possibilities of the Pessimistic Meta-Induction argument, but only forcing it to retreat a level.

The first objection is that RG realism does not answer the anti-realist arguments from historical theory change, regarding empirically successful but false theories. RGR rests its hopes on future physical theories being in the same space on which the RG transformation acts – and this is not obvious *a priori*.

Fraser (2018*b*) replies by stating that RG results cannot answer this scepticism by themselves. RG methods began in the 50s in solid state physics and were carried over to be used in quantum field theory, though ideas of scale transformations and invariance have been around since Pythagoras' time. RG analysis is a 'novel epistemological source' that had not been available to previous physicists. There is no need to abandon it, but rather more carefully examine the 'epistemological mileage' of the results gained from RG methods. This reply clearly does not provide a suitable response to Ruetsche's objection, as not enough work has been done to respond satisfactorily. It is worth noting, however, that RG realism is a relatively new position and is still being developed. As RG realism is worked on further, a better reply may be forthcoming. For now this reply is not satisfactory.

The second objection regards just what is the nature of the representational success enjoyed by RG methods. These robust entities are given representational status due to their invariance over the RG, but what exactly does this mean? What features of the world are these quantum field theories connecting to? To solve this problem there would have to be a much more detailed characterisation of the non-fundamental entities and properties – the descriptive content – of the QFT-models. But this is hard to do as the non-fundamental ontology is under-theorised.

This is not a problem specific to Renormalisation Group Realism – as Fraser writes, this is a

controversial issue outside of quantum field theory realism.

> "In philosophy more generally, there is no agreed-upon framework for regimenting claims about the non-fundamental, with different approaches gaining currency in different sub-disciplines." (Fraser, 2018*b*, p. 15)

Fraser mentions that turning to an approach found in these sub-disciplines, such as the structural realist approach to this problem, the use of appealing to 'real patterns', could help a RG Realist clarify their position, but that "doing so will only be as reassuring as these approaches are well motivated" (Fraser, 2018*b*, p. 16). But will hitching the RG realism carriage to one of the already available positions in philosophy actually go any way to avoid this objection? To answer this question, there are two other questions we need to ask.

Firstly: where is the realism?

Secondly: what kind of realism is Renormalisation Group Realism?

7.4 Where is the Realism?

From Worrall's first introduction of structural realism in 1989, as well as the different types to spring up from it, it was also adopted by the empiricist van Fraassen (1997, 2006, 2010) for his own position, Empiricist Structuralism. This position is a non-realist, empiricist form of structuralism. While van Fraassen agrees with some points made by Worrall and carried on by structural realists – that there is continuity of structure through our best scientific theories at the empirical level – the big problem with structural realism as he sees it is that there is a difference between the mathematical structure of scientific theories and the physical structure it supposedly represents in the world, and this cannot be explained in solely structural terms.

Empiricist Structuralism has two key views; that our best scientific theories represent empirical phenomena using abstract structures/theoretical models, and that these abstract structures are describable only up to structural isomorphism. All scientific representation, therefore, is mathematical at some level, and all we can know is structure. But this is not a realist position – this structure we learn about from empirical phenomena isn't telling us about what is there, objectively, in the world, but simply telling us what science is. Van Fraassen denies

the underlying objective modal structure that the structural realists require for their realism, and instead remains agnostic about whether our best scientific theories are telling us about the world or not.

Williams' arguments for Effective Realism over his Standard Account of Theory interpretation, as mentioned previously, are the following;

- The theorist is allowed to make use of different EFTs rather than having to come up with a toy theory which covers all length scales;

- The theorist is allowed, through using RG analysis, to pick out the robust entities of the theory that are invariant under the RG flow – the theorist can work out the structure that accurately represents phenomena.

- These Effective theories announce their weaknesses from the outset, rather than claiming to be a complete description of the world in all respects.

It is hard to see, here, where this account and van Fraassen's own non-realist position differ. It appears that the empirical structuralist would have no issue with the problem of EFTs in the same way that the 'Standard Account' does. The empirical structuralist would also seem to have no issue with the method of picking out the robust entities, as that aligns with the abstract structure/theoretical models that make up part of their view. In fact, it seems that, apart from the word realism, there is nothing the empirical structuralist would disagree with. The only difference would be that, instead of describing these robust entities as representations of physical content, the structural empiricist would call them representations of empirical phenomena.

There seems to be nothing in the accounts of Effective/Renormalisation Group Realism that justifies the jump from empirical phenomena to physical phenomena, without linking the account to a realist position such as structural realism. The argument that the physical entities, the 'robust' entities (in that they are entities which are accessible in a variety of independent ways as mentioned before), does not cause a problem for the empirical structuralist – empirical data should also fit this definition of robust. Without this link to structural realism, there is no source for the realism to spring from – it is indistinguishable from van Fraassen's empiricist agnosticism.

Since the Effective/RG realist does in fact call themselves a 'realist', then they need to pinpoint the source of their realism, or at least distinguish the difference between their realism and empiricist structuralism. Williams may have provided desiderata for a realist position based on EFT, but he has not yet truly defended a realist position. And though Fraser goes some way to shoring up that position in his forward look to such a realist position, there is still a final leap needed from empiricist structuralism to realism.

7.5   What kind of Realism?

There are many and varied realism frameworks already in the philosophical literature, some of which could provide that extra edge over empiricism for the Effective/RG realist.

Fraser gives a option of a potential framework for the Renormalisation Group realist to turn to in order to properly state their realist position and characterise the non-fundamental quantities and structures found by RG analysis. This option consists of Structural Realism.

This Effective/RG realism falls under the umbrella of selective realism. Since the antirealists' argued against realism, citing radical discontinuity in scientific theories through history, Harker (2012) writes that realists have moved more towards narrowing their realist commitments to only certain parts of a theory rather than a theory in entirety . Our theories are merely approximately true, and successful novel predictions indicate that there are at least parts of the theory that deserve realist support. This selective and more modest realist thesis has been taken on in various ways by realists since, where the actual constituents of the theory deserving of a realist commitment varies depending on the position. In entity realism, for example, it is the entities in our physical theories that warrant our realist commitment.

Our understanding of the world, Harker writes, comes to us piecemeal, and not always in the best order. This kind of selective realism, then, holds more weight against antirealist objections than previous realist positions, and there are many selective realist programmes to choose from. This new Effective/RG realism can very easily join their ranks, but just what kind of selective realism is it?

7.5.1 Exemplar Realism

Perhaps the easiest position to work with would be exemplar realism as presented by Saatsi (2017). Rather than viewing realism as a theory that is testable, Saatsi argues for realism to be seen as more of a particular attitude towards scientific theories. By changing from theory to attitude, the 'theory' or kind of realism in question no longer needs to fulfil the same criteria as another theory – such criteria being testable, explanatory, predictive, and others. As an attitude, realism can be relaxed to the point that formal frameworks meant to be able to be applied over all science are not needed.

Saatsi does this in order to escape the many kinds of 'recipe realism', where all the many different frameworks for scientific realism use different examples of scientific theories latching onto unobservable reality – case studies specific to certain areas of science – to then be generalised to all other scientific theories. These 'exemplars', however, can point to different ways that the theory is latching on to the world – sometimes even in opposite ways. This suggests that the exemplars, while they work locally, should not be generalised to all of scientific theory (i.e. an example case from QFT should not necessarily be expected to work in the context of behavioural science).

Rather than look for a generally applicable 'recipe' for realism, Saatsi suggests instead his 'exemplar' realism, which uses the exemplars locally only. Globally, an exemplar realist takes all examples of empirically successful theories as being so successful due to latching on to some part of unobservable reality. The specific case-by-case examples describe how these theories do represent the real physical content, and this may be different for different theories.

This kind of realism does not provide the epistemic content the Effective/RG realist might look for in an attempt to escape empiricism, as Saatsi points out. It is not so much an argument for realism against anti-realism, but a 'positive attitude'. This fits perfectly with the Effective/RG realist view – Williams, Fraser, and any others who are looking for a realist interpretation of QFT are quite obviously approaching QFT, an empirically successful theory, with a realist mindset. However it is not enough to escape empiricism, as this kind of realism from Saatsi does not actually provide a source of the realism for the Effective/RG realist. Rather, it states that what we can know of unobservable reality is that our best and most successful theories are 'latching'

onto it, and we can see more clearly how our theories do this by looking locally at exemplars.[5]

7.5.2 Perspectival Realism

Another potential position to shore up the foundations of RG realism is perspectival realism. Championed by Massimi (2016, 2018), perspectival realism is described as scientific perspectivism by Giere (2010) – our scientific knowledge is perspective-dependent and compatible with realism. Perspectival realism looks at ways in which we can understand truth in perspectival terms – it focuses on getting things right by using perspectival-truth.[6] Massimi's example is of the viscosity of water – this is a fundamental property in hydrodynamics but only a derivative one in statistical mechanics — the truth-value of knowledge claim that viscosity being a fundamental property is "sensitive to the perspectival *context of use*"[7] (Massimi, 2018, p. 350). In this way properties and quantities are not true or false, but simply change truth-conditions in different perspectives.

Is this a good position for the RG realist? The use of epistemic pluralism in perspectival realism – something Massimi calls a ubiquitous phenomenon in science – is definitely something RG realism uses in its foundation. The many EFTs used in QFT, and indeed in physics generally, is one of its selling points as according to Williams; as he writes, there is no interpretation in isolation, and the many different EFTs used in conjunction with RG analysis to find the robust entities is a core part of RG realism. So in this, Effective/RG realists will agree with perspectival realists. But what of perspectival-truths? How does that fit into the Effective/RG realism ethos? It seems that the fact that the truth conditions changing between perspectives, changing with context of use, is the opposite to what Effective/RG realists would want – they are searching for the invariant entities, the robust parts of the theory that do not change between different EFTs. Being able to assess the properties from between perspectives is a part that perhaps the Effective/RG realist would agree with, but is perspectival realism strong enough in regards to the realism of certain robust entities for the Effective/RG realist to take up its position?

---

[5]Where 'latching onto reality' is cashed out by Saatsi as "a theory's veridicality accounting for its empirical accuracy" (Saatsi, 2019, p. 18).

[6]"Knowledge claims in science are perspective-dependent when their truth-conditions (understood as rules for determining truth-values based on features of the context of use) depend on the scientific perspective in which such claims are made. Yet such knowledge claims must also be assessable from the point of view of other (subsequent or rival) scientists" (Massimi, 2018, p. 354).

[7]Italics are Massimi's own.

7.5.3 Entity Realism

The talk of invariant, robust 'entities' with physical content could work well with entity realism. Begun by Cartwright and McMullin (1984) and Hacking (1983), entity realism was put forward as the middle ground between empirical antirealism and scientific realism. Entity realism advocates belief in entities (even unobservable entities) without necessary belief in the scientific theories describing those entities. This isn't to say that entity realists believe in entities blindly without needing any kind of theory to provide explanation – Clarke (2001) states that they simply need make no commitment to a global or complete theory.

There are two different axes of entity realism – Hacking (1983) takes experimentation to be the best proof of scientific realism about entities, as opposed to theorising. To believe in the reality of entities found in science, on Hacking's view, requires those entities to be manipulable.[8] Hacking employs this 'manipulability criterion' as, historically, theoretical entities which do not end up being used (or manipulated) in experimentation in order to find other entities, generally turn out to be mistakes. He uses the example of the aether – long held to be a genuine entity, the aether was debunked and could not be used in experiment to explore other theoretical entities. Despite the classic objection to this argument – that before the aether was disproved, scientists genuinely believed that they were manipulating the aether in their experiments – this is nevertheless a position on entities that effective realists may like to take. Their entities, picked out of theory by RG analysis, are invariant quantities and a criterion like the manipulability criterion could be a way to justify the behaviour of those higher-level entities that the Effective/RG realist does not currently have.

However, as well as the common objection to this view, this kind of entity realism also focuses too much on the entities in a way which Effective/RG realism does not – the entity realist does not agree with the scientific realist in the truth, or just approximate truth, of explanatory scientific theories. The RG realist may find this an uncomfortable position when their own realism is entirely motivated from EFTs which are, as Williams states, approximately true.

---

[8]After using the example of the aether, Hacking goes on to make analogy with black holes, of which he admits being sceptical about – he shares with Leibniz a certain dislike of 'occult powers', of theoretical entities that are so far from our reach as to be untouchable. Black holes certainly fulfil that level of 'occult', though whether we should be sceptical of them or not would no doubt cause contention today, especially since the recent results of gravitational waves from LIGO, see Abbott et al. (2016).

If Hacking's version of entity realism is a little too flatfooted and physical for the Effective/RG realist, then Cartwright's version may be a good option. Cartwright posits entity realism not as a separate position from scientific realism but merely as an emphasis – a kind of anti-fundamental realism. For example, Cartwright is very clear that she is only anti-realist about fundamental laws – not about phenomenological ones. Since this is exactly the view that Williams takes, that non-fundamental ontology is the kind of realism RG analysis brings to QFTs, and that the higher-level entities that are robust and insulated against any changes over short distances are the ontological objects of realism, not the fundamental theories, then this appears, at the outset, to be a potentially beneficial alliance for the Effective/RG realist.

Cartwright's entity realism seems to ring very close to Effective/RG realism, which also moves past the fundamental laws to instead focus on the non-fundamental invariant entities picked by RG analysis. As with Cartwright's definition of phenomenological laws, these are the ones closely connected with the empirical data. As with Effective realism, the non-fundamentals are important – but in Cartwright's position, she describes there being a dichotomy between the fundamental and the non-fundamental, a disunity that the effective realist may be loath to accept. Effective realism uses scale invariance and symmetry to back their realism, which would be tricky to map onto the dichotomy between fundamental and non-fundamental quantities in Cartwright's entity realism.

There are other issues with this position – it has been much maligned over the years and is often dismissed in the literature, such as being called 'untenable middle ground' by Reiner and Pierson (1995). It is moved aside in favour of the two bigger positions of empirical anti-realism and scientific realism. If Fraser is wary to ascribe to a current metaphysical position because of the extra weight of supporting the motivation behind that position as well as the one behind RG realism, then taking on entity realism and its naysayers might be too much.

7.5.4 Structural Realism

The more obvious path to take would be that of structural realism. First started by Worrall (1989), structural realism offers a way to have both the 'no miracles' argument and the Pessimistic Meta-Induction argument because structures in scientific theories can be retained through otherwise

radical ontological theory change. Since Worrall's work, there have been branches along the structural realist line; Worrall (2007)'s own position states that our successful scientific theories are structurally correct, and that this is the strongest epistemic claim we can make about them. Sometimes called Epistemic Structural realism (ESR), this simply puts forward the idea that the structure retained through theory change in our best theories is what we have the best epistemic claim for being realist about.

There is also Ontic Structural Realism (OSR), championed by Ladyman and Ross, which agrees with Worrall about the continuity of structure but takes a stronger position than ESR in that the structure in our scientific theories is what exists in the world, and that this is objective modal structure.[9] This view can be taken further — in Eliminativist OSR, from French (2014), there are no objects in the world, only relational structure. Objects are 'eliminated' to leave only a structural metaphysics.

There is a bountiful selection of structural realism flavours for the RG realist to choose from, though some fit better than others. One might want to steer clear of Eliminativist OSR, for example, originally put forward by French and Ladyman (2003), and later continued by French (2014) — if, as Fraser states, the problem is the non-fundamental nature of the entities and properties which are invariant under the RG flow, then a position which eliminates entities would not exactly provide the strongest foundations for realism. *Prima facie*, this kind of OSR is not suitable for a view advocating for robust entities. Also, French's OSR that eliminates objects is motivated by looking for the fundamental level in theories — this, of course, would not work with QFT, which is not about the fundamental level, or fundamental *stuff*. Similarly with ESR, the RG realist would want a stronger grounding to escape the problem of just where the realism comes from.

Fraser mentions the kind of structural realism that seems to be the best option for the RG realist — the 'Everything Must Go' flavour of realism from Ladyman et al. (2007), where structures are real and in the world, and is this objective modal structure. This structural realism advocates the scale relativity of the ontology as well as modal structure. Effective/RG realism works well with OSR – the objective modal structure of the world fits with the robust, invariant

---

[9]See Ladyman et al. (2007).

quantities picked out by RG methods which represent genuine physical content. The RG analysis used on EFTs becomes, then, a method through which the objective modal structure of the world can be picked out. RG techniques fit with objective modal structure, sharing qualities such as invariance and modal robustness, the characteristics Williams and Fraser state that RG analysis picks out of EFTs as the ontology of the theory, as the higher-level entities that represent genuine physical content. The RG flow is the flow through space of possible theories.

If Effective/RG realists align with this kind of structural realism, there are more benefits to be had. The OSR of Ladyman and Ross claims that the theoretical structure of our best scientific theories represents the objective modal structure of the world, and that this modal structure is independent of the way that we describe it – it commits to natural necessity.[10] Berenstain and Ladyman (2012) argue that realists are *prima facie* committed to objective modality – that there is a structure of natural necessity in the world, independent of us, and that this distinguishes structural realism from structural empiricism. An example used by Berenstain and Ladyman (2012) is that of the speed of light, c – material objects cannot go faster than c, as they would become infinitely heavy when approaching c and would therefore need infinite force to keep accelerating. Only with both these two commitments – objective modality and natural necessity – Berenstain and Ladyman (2012) argue the realist is distinct from structural empiricism. In this case, the robust quantities preserved under RG transformations would have to be regarded as modal in order to escape the structural empiricism already threatening Effective/RG realists. Following the arguments of Berenstain and Ladyman (2012), for Effective/RG realists to remain realist, they would have little choice but to make those two commitments.

In this way, Ontic Structural Realism seems to be the perfect mast for the RG realist to lash their position to, though neither Williams nor Fraser make that connection, despite Fraser (2018*b*) clearly mentioning it as an option. Fraser is concerned that any extra support that structural realism or another position could give would only be 'as reassuring as these approaches are well-motivated'. However, OSR is the best option for Effective/RG realists as an available realism position – its scale relativity and modal structure align practically seamlessly with that of Effective and RG realism, and its commitments to objective modality and natural necessity

---

[10]Natural laws are necessarily true.

distinguish it from structural empiricism. Attaching to OSR would also help Effective/RG realism in responding to Ruetsche (2020)'s recent claim that the quantities picked out by RG methods live too close to the empirical surface of the theory. Ruetsche once again highlights the fact that Effective/RG realism, as it currently is set out, is indistinguishable from empirical structuralism. If Effective/RG realism were to join with OSR, the very commitments needed to do this would distinguish it from empirical structuralism, though work would need to be done to align the two.

7.5.5 Semirealism

If neither structural realism nor entity realism satisfy the Effective/RG realist, then there is always semirealism. Proposed by Chakravartty (1998), it is put forward as a combination of structural realism and entity realism. Chakravartty (1998) describes entities and structure as "metaphysically separable but epistemically interwoven" (Chakravartty, 1998, p. 407), and that the two positions entail each other. For semirealists, the entities are causal substances with two kinds of properties: detection properties and auxillary properties. The detection properties are the causal regularities that the entities have, allowing us to infer their existence, and auxillary properties are other regularities which, after experimentation, either turn out to be detection properties or 'chimeras', therefore playing the part of methodological catalyst. That is the entity realism part of semirealism; the relations between the detection properties define the structure of the theory containing them, which is where the structural realism comes in. The structure is preserved through theory change as the mathematical structure faithfully represents the relations between real objects in the world.

Chakravartty's semirealism could be a good fit for the Effective/RG realist – the semirealist is committed to the truth of our best, most empirically successful theories, due to the preserved parts of theories as they evolve and change. Since parts of our old theories are retained in our new ones, there must be some small subset of claims the theories made that are true. Since, as Chakravartty writes, the aim of scientific enquiry is to preserve and increase truth claims in our scientific theories, then the semirealist should believe in the objects or entities and the structure they can, and should work on the rest.

Is semirealism a realistic option for Effective/RG realists? Semirealism requires the claim

that well-confirmed relations between detectable physical properties are reasonable evidence of causal activity. This therefore requires at least that notion of causality. Also, as a combination of both structural and entity realism, the Effective/RG realist will have to take on objections to both of those positions, and defend both the reality of non-fundamental entities but also the relations between those entities, which Effective/RG realism currently does not have to do. Again, as with kinds of entity realism, semirealism involves believing in entities and structure without fully believing the theories they belong to (the doubt cast on auxillary properties as an example), might cause a problem for Effective/RG realists. Semirealism also involves a specific commitment to a kind of modal structure – causal structure – which may not suit the RG realist's aims. There are parts of quantum theory, the quantum teleportation protocol, for example, which are not so easily described as causal.

7.6 Conclusion and Prospectus

Using Renormalisation Group analysis on Effective Field Theories is without doubt, as Fraser describes, a promising strategy towards refining the realist position in the context of QFT. But even disregarding the problems with Effective/RG realism that Ruetsche highlights, and without even attempting to take on the measurement problem, this new realist position still has a long way to go. There needs to be a definitive difference between this position and empiricism, or it will struggle to hold on to the title.

There are options; the Effective/RG realist could agree with Saatsi's take on realism being an 'attitude' rather than a theory, and nullify the need for a source of their realism.[11] This, however, does not necessarily solve the problem of showing a clear distinction between Effective/RG realism and empiricism – just the different state of mind with which the realists approach renormalisation group analysis compared to empiricists. If the realist wants to defend a stronger position against empiricism, they need to pinpoint the source of their realism, and characterise the non-fundamental ontology of their position.

The Renormalisation Group Realist, therefore, has the option of hitching a ride with structural realism, or at least one of the kinds of scientific realism detailed above, or even to just having

---

[11]This is also the view of Chakravartty (2017).

objective modal structure, in order to properly describe their kind of realism. Regardless of which position the RG realist chooses, without this extra step, it is not clear to see where otherwise the realism of Effective or RG realism comes from. There is a big difference between defending a realist position and providing desiderata for that theory. While Williams has been successful in setting out these desiderata, as Fraser (2018*a,b*) rightly discusses later, this position needs to detail precisely how the non-fundamental ontological content given by the Renormalisation Group analysis is characterised. If the Renormalisation Group realist was to bite the bullet and tie their realism to another position, then it is clear that the view most aligned with their aims and beliefs is that of ontic structural realism. As previously stated, the very characteristics that RG analysis picks out of theories are those that Ontic Structural Realism takes to be the objective modal structure of the world. There is no extra step needed to merge these two positions.

Further work in this area will perhaps show a more natural alignment of Effective/RG realism with one of the frameworks mentioned previously. Currently, it seems that ontic structural realism is the most successful in allowing Effective/RG realism to distinguish itself from structural empiricism.

Section 7.1 poses Ruestche's question to structural realists; what are good structural realists supposed to believe in now? Sections 7.2-5 explore the Effective/RG realist position, and it is argued that ontic structural realism is the best source of realism for the Effective/RG realist to be certain of an escape from empiricism. The Effective/RG realist, when aligning with OSR, can therefore very easily answer Ruestche's question. What, when confronted with multiple inequivalent representations of physical structure, should the structural realist believe in? According to the Effective/RG realist, robust entities which are picked out by RG methods and are invariant across the different theories. With the added realism of OSR, Effective/RG realism not only agrees with Ruetsche's Coalesced Structures approach but also gives structural realists the tools to pick out those robust entities and structures about which we should be realist.

Conclusion

It was clear from the very beginning of quantum theory that it is incredibly difficult to interpret. From the disagreements between Bohr and Einstein to the many different interpretations available to choose from today, it seems that despite the physics advancing in both scope and empirical accuracy so that it is one of the best scientific theories we have, interpreting quantum mechanics has only become more and more convoluted. With so many different and influential interpretations to choose from, it is unclear how we should choose.

This thesis, first looks at models of quantum theory in the form of thought experiments and how they not only teach us about thought experiments and mental models more generally but also help to reveal how the core issue of interpretation is the quantum to classical transition, of which measurement is a type. A discussion of different interpretations and their solutions to the measurement problems is followed by a detailed look at decoherence theory, to show how it can be used as a tool to pressure-test interpretations. The study of the quantum to classical transition and the quantum–classical limit shows the importance of semiclassical theory when considering the quantum to classical transition, and how, following Bokulich (2008*a*), even 'fictional' theories are important as they can have explanatory power. Finally, it is argued that, as Ruetsche (2011)

argues, having multiple interpretations is actually more explanatorily useful than just choosing one, especially in Quantum Field Theory and Quantum Statistical Mechanics. This poses a problem for realism, however – having multiple unitarily inequivalent representations leads Ruetsche to pose the question of what a structural realist should believe in now. However, a new kind of realism from Williams (2017) and Fraser (2017, 2018*a,b*) was shown to be compatible with ontic structural realism, which answers Ruetsche's question.

8.1 Models

Chapter 2 starts as interpretation-free as possible by chronologically exploring quantum mechanical thought experiments from 1949 to 2018, to gain insight into the development of quantum theory via these mental models and representations. These thought experiments included famous historical examples such as Schrödinger's cat, and influential modern day thought experiments such as the PBR theorem.

The debate in the wider thought experiment literature is also discussed in chapter 2. Do they, like real-world experiments, bring us new knowledge? In the recent literature, there are two clear extreme views; Norton's empiricism and Brown's Platonism. In between there are the many middle-ground views, such as the mental models view of Nersessian (1992) or the experimentalism view by Sorensen (1992). The Platonist view from Brown (2014, 2011) argues that thought experiments allow us to directly access the laws of nature from a Platonic world, while Norton's empiricist view states that thought experiments bring us no new knowledge and are simply picturesque arguments.

I argue against Norton's empiricism and state that while thought experiments can be thought of as arguments, they also play very important roles as placeholders for experiments that are currently impossible, similar to analogue experiments. Their evolution over years of theory development means they cannot be easily taxonomised, so we must allow category change and use Placeholder category. The argument also supports the view put forward by Stuart (2015, 2016) that thought experiments are on the same spectrum of real world experiments and are vital in quantum mechanics, as some experiments such as PR boxes are impossible to perform except

in the imagination. Norton's view that they are nothing more than picturesque arguments cannot be true – thought experiments are too much like real-world experiments to only be considered as arguments.

Chapter 2 also shows that thought experiments have been used to probe the implications and limitations of quantum mechanics throughout its lifetime, and that the main issue that most of the thought experiments focus on is the issue of the quantum to classical transition. Measurement, in quantum mechanics, is a black box, and is an instance of the quantum to classical transition. In highlighting the conceptually interesting aspects of the theory more than just measurement, quantum mechanical thought experiments pick out the differences between what quantum mechanics tells us the world is like and what we see in the classical. This is the core of the difficulties we face when trying to interpret the theory. While the trilemma from Maudlin (1995) identifies the measurement problem, quantum mechanical thought experiments show us that the problem is the wider issue of the quantum to classical transition.

8.2 Interpretations

While it is useful that the trilemma identifies the problem of measurement, measurement is just a special case of the larger problem of the quantum to classical transition. Chapter 3 considers a more recent discussion of the measurement problem by Schlosshauer (2005), where it is divided into two; the unique outcomes problem and the preferred basis problem. Five different interpretations of quantum mechanics are introduced and briefly discussed. In examining how each interpretation attempts to solve the measurement problem, chapter 3 analyses which of the two measurement problems the interpretations are actually solving. There are three different kinds of solutions. It is seen that in Bohmian mechanics and collapse theories, the focus is on solving the unique outcomes problem, and the preferred basis is selected *ad hoc* to suit this solution. The preferred basis as a selectable parameter with the unique outcomes problem being used interchangeably with the full measurement problem is not satisfactory.

Another tactic seen from the more modern interpretations, Modal and Oxonian Everettianism, is to solve the unique outcomes problem first, leaving the preferred basis problem still to deal

with. Rather than choosing it, these interpretations employ decoherence theory to deal with that part of the measurement problem.

A third category of solution belongs to QBism, which is the main epistemic interpretation. This interpretation solves both problems by taking a subjectivist view of quantum theory and turning objective outcomes into widely shared degrees of belief. However, there is disagreement between proponents of the QBism interpretation about just how subjectivist one needs to be about quantum theory. Chapter 3 argues that the question of 'how subjective' can be answered by looking more closely at the quantum to classical transition. This can also be applied more generally to the other interpretations, as looking at how well they incorporate decoherence theory into their frameworks could help us choose between them.

Chapter 4 looks at the history and the details of decoherence theory. It also explores the consistent histories interpretation of quantum mechanics, which unlike the other interpretations, solves the measurement problems by removing the emphasis placed on measurement. With measurement no longer the 'black box', this interpretation does not need decoherence theory in the same way that the other interpretations do, in order to solve the measurement problem. This contrast with the other interpretations shows just how important it is that they should be able to incorporate decoherence theory in their accounts.

This idea is taken further in chapter 5, where the five interpretations discussed in chapter 3 are revisited in the light of decoherence. Using decoherence theory as a pressure-test for these interpretations – interpretations that can more successfully incorporate decoherence in their frameworks gain extra credibility. Bohmian mechanics and collapse theories, both of which select a preferred basis, do not explicitly use decoherence theory in their approach, though they share similar formalism with it. They interpret the decoherence formalism differently in order to fit it within their own frameworks, and though there is the potential for future work to combine them (as Schlosshauer (2005) argues for Bohmian mechanics), they currently do not satisfactorily incorporate decoherence in their interpretations. Modal interpretations use decoherence implicitly, and it is sometimes used to solve the preferred basis problem. Oxonian Everettianism, however, explicitly uses decoherence theory as a criterion for quasi-classical emergence to establish its 'worlds' as quasi-classical, dynamically isolated structures instantiated within the quantum state.

This chapter also looked at an interpretation not discussed in chapter 3, as it is entirely based on decoherence. Zurek's Existential interpretation uses decoherence theory as its foundation. All of these interpretations seem to benefit, or could benefit, from incorporating decoherence,and the more successful they are, the more inclined we should be to choose them.

These are all ψ-ontic interpretations, however. The final interpretation from chapter 3, QBism, is a ψ-epistemic interpretation, and so has a markedly different relationship with decoherence theory. Chapter 5 discusses QBism's subjectivity and argues that for QBism to work as an interpretation, decoherence theory cannot be a physical process, as it can be in the ψ-ontic interpretations above. QBism must take decoherence as just us obeying the Reflection Principle and updating our future probabilities, and it must be entirely subjective. If any objectivity is kept, then QBism's explanation for decoherence no longer works.

Chapter 6 also looks more closely at the quantum to classical transition by examining an analogy given by Bub (2016) in the context of the Information-theoretic interpretation, another ψ-epistemic interpretation. The analogy considers the classical limit as an idealisation of infinite microsystems, and compares the transition from quantum to classical with a phase transition in statistical mechanics. It is shown that the analogy is a formal analogy, not a material analogy, and can be set up as follows;

P1. The quantum-classical transition in the IT interpretation can only be understood using an infinite limit.

P2. Phase transitions in statistical mechanics can be understood using an infinite limit.

P3. The infinite limit for phase transitions in statistical mechanics is both explanatorily unproblematic and unavoidable.

C. Therefore we should treat the infinite limit for the quantum-classical transition as explanatorily unproblematic and unavoidable.

This analogy, aside from issues with its validity, fails due to two false premises – the first and the third. The first premise states that the analogy is based on the similarity of the necessary

infinite limit. Both classical mechanics and thermodynamics reduce to quantum and statistical mechanics respectively, with the addition of an infinite limit (at least for the Information-Theoretic interpretation of quantum mechanics). However, in the literature for statistical mechanics, it is not clear that this infinite limit is either unavoidable or explanatorily unproblematic.

The analogy's biggest failing is that it does not take semiclassical theory into account. Semiclassical theory uses classical theory to explain certain strange quantum behaviour, such as in the Rydberg atom, as quantum theory cannot as of yet explain these behaviours alone. This may at first seem counter-intuitive – classical theory is known to be fictional, or at least merely an approximation of quantum theory in certain limits. Semiclassical mechanics shows that the quantum to classical transition cannot be simplified to a transition using an infinite limit and that classical theory is not yet eliminable in the context of the quantum-classical limit. In agreement with the Eikonic conception of explanation from Bokulich (2016), taking away explanatory fictions such as semiclassical theory would reduce our understanding of the quantum to classical transition.

8.3 Realism

In chapter 7 the infinite limit is explored once again, this time in the context of Quantum Statistical Mechanics and Quantum Field Theory. Many of the interpretations discussed in chapters 3, 5 and 6 are not able to discuss, in any great detail, quantum theory beyond ordinary non-relativistic quantum mechanics. Once in the realm of Quantum Statistical Mechanics and Quantum Field Theory, there lie more issues to do with interpretation. The Coalesced Structures approach from Ruetsche (2011) addresses the issue of having many unitarily inequivalent representations for phenomena. These inequivalent interpretations cause problems for the goal of finding or choosing the 'true' interpretation of quantum mechanics, as by dropping any we face losing explanations. Ruetsche uses her approach to argue that there cannot even be continuity of structure, as with all the inequivalent representations, 'what is a good structural realist supposed to believe in now?'

This issue of realism, and what we should believe to be in the world when we use these

inequivalent representations, led to a new kind of scientific realism. Williams (2017) and Fraser (2017, 2018*b*) put forward Effective and Renormalisation Group realism respectively. These new context-dependent realism theories use the technique of Renormalisation Group analysis on the effective field theories in order to find what quantities we should be realist about. These approximate, effective theories are very accurate at their specific length or energy scale but break down rapidly at shorter distances or higher energies. Effective/Renormalisation Group realism finds certain 'robust' entities, properties, and relations in a variety of ways with the most general and common being through Renormalisation Group (RG) analysis. These robust entities are, using RG analysis, defined by being invariant under the RG flow. These entities are the real physical content we should take to be in the world from these effective field theories.

It is argued by Ruetsche (2018) that this new kind of realism is indistinguishable from empiricism unless it joins with an already existing realism programme, as it seems that an empirical structuralist would have no issue with the problem of effective field theories in the same way that the Standard Account of interpretation does. Though Fraser (2018*b*) believes that by aligning with an already existing realism approach means taking on all the particular objections that come with it, I argue that combining with another realism programme saves the Effective/Renormalisation Group Realist from empiricism. After going through several different approaches, I argue that the best option for the Effective/Renormalisation Group realist is ontic structural realism from Ladyman et al. (2007), as its scale relativity and modal structure align perfectly with the idea of scale-dependent effective field theories and robust entities picked out by RG analysis. Together, these answer Ruetsche's question of what good structural realists are supposed to believe in now — the robust entities picked out by Renormalisation Group methods.

8.4 Models, Interpretations, and Realism in Quantum Physics

From the very beginning of this thesis, the emphasis has been on the difficulty of interpreting quantum physics in a way that reconciles it with what we see in the classical world. Starting with the debate over the original interpretation from the nascence of the theory, and finishing with the difficulties with interpretation in modern Quantum Field Theory, it seems as though in

the hundred years since quantum theory was first discovered we have only come up with more interpretations instead of finding the 'correct' interpretation.

Exploring the development of quantum theory through thought experiments gives a better perspective on the core issue of measurement and the quantum to classical transition and teaches us more about quantum mechanics than the trilemma approach. For example in the PR box thought experiment, Popescu and Rohrlich (1994) explore the connections between the axioms of quantum theory and the strange correlations in its measurements. Going through various interpretations' solutions to the two measurement problems showed that decoherence theory can be used as a pressure-test for interpretations, though ψ-ontic and ψ-epistemic interpretations treat it very differently. ψ-ontic interpretations all could benefit from including decoherence theory, though some such as Oxonian Everettianism do better than others at incorporating decoherence theory into their framework. This can be used as a way to choose between interpretations – decoherence as a process is empirically well-tested and is extremely important to the quantum to classical transition, which is what these interpretations are purporting to describe. This decoherence test does not go well for ψ-epistemic interpretations – a detailed look at QBism and its framework shows that it must be completely subjective and so decoherence within QBism must also be subjective and cannot be a real physical process.

A closer look at the quantum to classical transition via an analogy with phase transitions in statistical mechanics shows the importance of semiclassical theory when dealing with the quantum–classical divide. This in turn demonstrates the explanatory gains we get by keeping theories that are considered fictions. This is also shown when looking at the issue of inequivalent representations of phenomena in Quantum Field Theory and Quantum Statistical Mechanics. We need to keep these inequivalent representations despite the lack of a coherent interpretation because they provide us with explanations we would otherwise lose, just as we keep semiclassical theory for its explanatory power despite classical theory being 'fictional'.

This, of course, poses a problem for realism. Ruetsche argues for pragmatism with her Coalesced Structures Approach, as we can then keep all the virtues from each inequivalent representation. But then, as she asks, what should the good structural realist believe in when there is no continuity of structure between representations? Effective/Renormalisation Group

Realism offers a solution to this problem; Renormalisation Group analysis can pick out the robust entities and structures that persist through effective field theories, and can be used as a tool for the realist to pick out the parts of the theory which are real and in the world. Though this new kind of realism struggles to differentiate itself from empirical structuralism, I suggest that it could align itself with ontic structural realism in order to making this difference clearer.

Decoherence theory can be used as a pressure-test for interpretations and so it can be a tool we can use to go some way to deciding between them. But it seems clear that the lesson learnt from semiclassical theory, Quantum Field Theory, and Quantum Statistical Mechanics is that it is in our best interests to have all of these interpretations to hand. Choosing between them would blinker us in the same way as would ignoring semiclassical explanations of phenomena at the quantum to classical transition. While the quantum to classical transition is still the core issue in quantum physics, we need all the explanatory power we can get.

Future study of the quantum to classical transition will no doubt shed more light on decoherence theory and interpretations of quantum mechanics. As technology improves, we are able to isolate larger and larger quantum states from the environment, preventing collapse.[1] As Effective/Renormalisation Group realism is still so new, it will be interesting to see how it develops and how it replies to Ruetsche's Pessimistic Meta-Induction objection, which was not discussed in detail here. Overall, however, it is clear that the quantum to classical transition is fertile ground for learning more about quantum physics and how we should interpret it, while inequivalent representations in quantum theory show us that it is worth keeping multiple interpretations so as to foster the development of physics.

---

[1]As discussed in section 2.2.2.

Abbott, B. P., Abbott, R., Abbott, T., Abernathy, M., Acernese, F., Ackley, K., Adams, C., Adams, T., Addesso, P., Adhikari, R. et al. (2016), 'Observation of gravitational waves from a binary black hole merger', *Physical review letters* **116**(6), 061102.

Adler, S. L. (2003), 'Why decoherence has not solved the measurement problem: a response to pw anderson', *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* **34**(1), 135–142.

Allori, V. (2002), 'Decoherence and the classical limit of quantum mechanics'.

Appleby, D. (1999), 'Bohmian trajectories post-decoherence', *Foundations of Physics* **29**(12), 1885–1916.

Arndt, M., Nairz, O., Vos-Andreae, J., Keller, C., Van der Zouw, G. and Zeilinger, A. (1999), 'Wave–particle duality of c60 molecules', *nature* **401**(6754), 680–682.

Bacciagaluppi, G. (2014), A critic looks at qbism, *in* 'New Directions in the Philosophy of Science', Springer, pp. 403–416.

Bacciagaluppi, G. (2016), The role of decoherence in quantum mechanics, *in* E. N. Zalta, ed., 'The Stanford Encyclopedia of Philosophy', fall 2016 edn, Metaphysics Research Lab, Stanford University.

Bacciagaluppi, G. and Hemmo, M. (1996), 'Modal interpretations, decoherence and measurements', *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* **27**(3), 239–277.

Bailer-Jones, D. M. (2009), *Scientific models in philosophy of science*, University of Pittsburgh Pre.

Baker, D. J. (2009), 'Against field interpretations of quantum field theory', *The British Journal for the Philosophy of Science* **60**(3), 585–609.

Bartha, P. (2016), Analogy and analogical reasoning, *in* E. N. Zalta, ed., 'The Stanford Encyclopedia of Philosophy', winter 2016 edn, Metaphysics Research Lab, Stanford University.

Batterman, R. W. (1995), 'Theories between theories: Asymptotic limiting intertheoretic relations', *Synthese* **103**(2), 171–201.

Batterman, R. W. (2001), *The devil in the details: Asymptotic reasoning in explanation, reduction, and emergence*, Oxford University Press.

Batterman, R. W. (2005), 'Critical phenomena and breaking drops: Infinite idealizations in physics', *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics* **36**(2), 225–244.

Bell, J. S. (1964), 'On the einstein podolsky rosen paradox', *Physics Physique Fizika* **1**(3), 195.

Bene, G. and Dieks, D. (2002), 'A perspectival version of the modal interpretation of quantum mechanics and the origin of macroscopic behavior', *Foundations of Physics* **32**(5), 645–671.

Benfatto, G. and Gallavotti, G. (1995), *Renormalization group*, Vol. 1, Princeton University Press.

Berenstain, N. and Ladyman, J. (2012), Ontic structural realism and modality, *in* 'Structural Realism', Springer, pp. 149–168.

Berry, M. (1995), 'Asymptotics, singularities and the reduction of theories', *Studies in Logic and the Foundations of Mathematics* **134**, 597–607.

Berry, M. V. et al. (1991), 'Some quantum-to-classical asymptotics', *Chaos and Quantum Physics, M.-J. Giannoni, A. Voros, and J. Zinn-Justin (eds.), Elsevier Pub., NY* pp. 252–303.

Binney, J. J., Dowrick, N. J., Fisher, A. J. and Newman, M. (1992), *The theory of critical phenomena: an introduction to the renormalization group*, Oxford University Press, Inc.

Birkhoff, G. and Von Neumann, J. (1936), 'The logic of quantum mechanics', *Annals of mathematics* pp. 823–843.

Bohm, D. (1952), 'A suggested interpretation of the quantum theory in terms of "hidden" variables. i', *Physical review* **85**(2), 166.

Bokulich, A. (2001), 'Rethinking thought experiments', *Perspectives on Science* **9**(3), 285–307.

Bokulich, A. (2008*a*), Explanatory fictions, *in* 'Fictions in Science', Routledge, pp. 99–117.

Bokulich, A. (2008*b*), *Reexamining the quantum-classical relation*, Cambridge University Press Cambridge.

Bokulich, A. (2016), 'Fiction as a vehicle for truth: Moving beyond the ontic conception', *The Monist* **99**(3), 260–279.

Bokulich, A. (2018), 'Representing and explaining: The eikonic conception of scientific explanation', *Philosophy of Science* **85**(5), 793–805.

Bokulich, A. and Frappier, M. (2017), 'On the identity of thought experiments: Thought experiments rethought', *The Routledge Companion to Thought Experiments* .

Bransden, B. H. and Joachain, C. J. (1989), *Introduction to quantum mechanics*, Longman Group UK Limited.

Brassard, G., Broadbent, A. and Tapp, A. (2003), Multi-party pseudo-telepathy, *in* 'Workshop on Algorithms and Data Structures', Springer, pp. 1–11.

Brassard, G., Broadbent, A. and Tapp, A. (2005), 'Quantum pseudo-telepathy', *Foundations of Physics* **35**(11), 1877–1907.

Brassard, G., Cleve, R. and Tapp, A. (1999), 'Cost of exactly simulating quantum entanglement with classical communication', *Physical Review Letters* **83**(9), 1874.

Breuer, H.-P. and Petruccione, F. (1999), *Open Systems and Measurement in Relativistic Quantum Theory: Proceedings of the Workshop Held at the Istituto Italiano per gli Studi Filosofici, Naples, April 3–4, 1998*, Vol. 526, Springer Science & Business Media.

Brown, H. R. and Timpson, C. G. (2006), Why special relativity should not be a template for a fundamental reformulation of quantum mechanics, *in* 'Physical Theory and its Interpretation', Springer, pp. 29–42.

Brown, J. R. (2011), *The laboratory of the mind: Thought experiments in the natural sciences*, Routledge.

Brown, J. R. (2014), 'Explaining, seeing, and understanding in thought experiments', *Perspectives on Science* **22**(3), 357–376.

Brown, J. R. and Fehige, Y. (2017), Thought experiments, *in* E. N. Zalta, ed., 'The Stanford Encyclopedia of Philosophy', summer 2017 edn, Metaphysics Research Lab, Stanford University.

Brukner, C. (2015), 'On the quantum measurement problem', *arXiv preprint arXiv:1507.05255* .

Brukner, Č. (2018), 'A no-go theorem for observer-independent facts', *Entropy* **20**(5), 350.

Bub, J. (2011), 'Quantum probabilities: an information-theoretic interpretation', *Probabilities in physics* pp. 231–262.

Bub, J. (2016), *Bananaworld: Quantum Mechanics for Primates*, Oxford University Press.

Bub, J. (2018), 'In defense of a "single-world" interpretation of quantum mechanics', *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* .

Bub, J. and Pitowsky, I. (2010), 'Two dogmas about quantum mechanics', *Many worlds* pp. 433–459.

Butterfield, J. (2011*a*), 'Emergence, reduction and supervenience: a varied landscape', *Foundations of Physics* **41**(6), 920–959.

Butterfield, J. (2011*b*), 'Less is different: emergence and reduction reconciled', *Foundations of Physics* **41**(6), 1065–1135.

Butterfield, J. and Bouatta, N. (2015), 'Renormalization for philosophers', *Metaphysics in Contemporary Physics* **104**, 437–485.

Buzzoni, M. (2012), Thought experiments from a kantian point of view, *in* 'Thought Experiments in Science, Philosophy, and the Arts', Routledge, pp. 104–120.

Caldeira, A. and Leggett, A. J. (1985), 'Influence of damping on quantum interference: An exactly soluble model', *Physical Review A* **31**(2), 1059.

Caldeira, A. O. and Leggett, A. J. (1981), 'Influence of dissipation on quantum tunneling in macroscopic systems', *Physical Review Letters* **46**(4), 211.

Caldeira, A. O. and Leggett, A. J. (1983), 'Path integral approach to quantum brownian motion', *Physica A: Statistical mechanics and its Applications* **121**(3), 587–616.

Callender, C. (2001), 'Taking thermodynamics too seriously', *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics* **32**(4), 539–553.

Camilleri, K. (2009*a*), 'Constructing the myth of the copenhagen interpretation', *Perspectives on Science* **17**(1), 26–57.

Camilleri, K. (2009*b*), 'A history of entanglement: Decoherence and the interpretation problem', *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* **40**(4), 290–302.

Camilleri, K. and Schlosshauer, M. (2015), 'Niels bohr as philosopher of experiment: Does decoherence theory challenge bohr's doctrine of classical concepts?', *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* **49**, 73–83.

Cartwright, N. (1999), *The dappled world: A study of the boundaries of science*, Cambridge University Press.

Cartwright, N. (2010), Models: Parables v fables, *in* 'Beyond Mimesis and Convention', Springer, pp. 19–31.

Cartwright, N. and McMullin, E. (1984), 'How the laws of physics lie', *AmJPh* **52**(5), 474–476.

Caves, C. (2000), 'Probabilities as betting odds and the dutch book', *info. phys. unm. edu/~caves/reports/reports. html* **3**.

Caves, C. M., Fuchs, C. A. and Schack, R. (2002*a*), 'Quantum probabilities as bayesian probabilities', *Physical review A* **65**(2), 022305.

Caves, C. M., Fuchs, C. A. and Schack, R. (2002*b*), 'Unknown quantum states: the quantum de finetti representation', *Journal of Mathematical Physics* **43**(9), 4537–4559.

Caves, C. M., Fuchs, C. A. and Schack, R. (2008), 'Erratum: Unknown quantum states: The quantum de finetti representation [j. math. phys. 44, 4537 (2002)]', *Journal of Mathematical Physics* **49**(1), 4537.

Chakravartty, A. (1998), 'Semirealism'.

Chakravartty, A. (2017), *Scientific ontology: Integrating naturalized metaphysics and voluntarist epistemology*, Oxford University Press.

Chandrasekharan, S., Nersessian, N. and Subramanian, V. (2012), '14 computational modeling', *Thought experiments in philosophy, science, and the arts* **11**, 239.

Cirel'son, B. S. (1980), 'Quantum generalizations of bell's inequality', *Letters in Mathematical Physics* **4**(2), 93–100.

Clarke, S. (2001), 'Defensible territory for entity realism', *The British Journal for the Philosophy of Science* **52**(4), 701–722.

Clauser, J. F., Horne, M. A., Shimony, A. and Holt, R. A. (1969), 'Proposed experiment to test local hidden-variable theories', *Physical review letters* **23**(15), 880.

Cleve, R., Hoyer, P., Toner, B. and Watrous, J. (2004), Consequences and limits of nonlocal strategies, *in* 'Proceedings. 19th IEEE Annual Conference on Computational Complexity, 2004.', IEEE, pp. 236–249.

Clifton, R., Bub, J. and Halvorson, H. (2003), 'Characterizing quantum theory in terms of information-theoretic constraints', *Foundations of Physics* **33**(11), 1561–1591.

Craver, C. F. (2014), The ontic account of scientific explanation, *in* 'Explanation in the special sciences', Springer, pp. 27–52.

Da Costa, N., Lombardi, O. and Lastiri, M. (2013), 'A modal ontology of properties for quantum mechanics', *Synthese* **190**(17), 3671–3693.

Dardashti, R., Thébault, K. P. and Winsberg, E. (2015), 'Confirmation via analogue simulation: what dumb holes could tell us about gravity', *The British Journal for the Philosophy of Science* p. axv010.

De Broglie, L. (1928), 'La nouvelle dynamique des quanta', *Rapport et discussions du Vˆ e Conseil de Physique Solvay* .

Deutsch, D. (1985), 'Quantum theory as a universal physical theory', *International Journal of Theoretical Physics* **24**(1), 1–41.

DeWitt, B. S. (1970), 'Quantum mechanics and reality', *Physics today* **23**(9), 30–35.

Dieks, D. (1989), 'Quantum mechanics without the projection postulate and its realistic interpretation', *Foundations of Physics* **19**(11), 1397–1423.

Diósi, L. (1989), 'Models for universal reduction of macroscopic quantum fluctuations', *Physical Review A* **40**(3), 1165.

Dürr, D., Goldstein, S. and Zanghi, N. (1992), 'Quantum chaos, classical randomness, and bohmian mechanics', *Journal of Statistical Physics* **68**(1-2), 259–270.

Dürr, D., Goldstein, S. et al. (1995), 'Bohmian mechanics and the meaning of the wave function', *arXiv preprint quant-ph/9512031* .

Ehrenfest, P. (1933), Phase changes in the ordinary and extended sense classified according to the corresponding singularities of the thermodynamic potential, *in* 'Proc Acad Sci Amsterdam', Vol. 36, pp. 153–157.

Einstein, A., Podolsky, B. and Rosen, N. (1935), 'Can quantum-mechanical description of physical reality be considered complete?', *Physical review* **47**(10), 777.

Elitzur, A. C. and Vaidman, L. (1993), 'Quantum mechanical interaction-free measurements', *Foundations of Physics* **23**(7), 987–997.

Emary, C., Lambert, N. and Nori, F. (2013), 'Leggett–garg inequalities', *Reports on Progress in Physics* **77**(1), 016001.

Everett III, H. (1957), '"relative state" formulation of quantum mechanics', *Reviews of modern physics* **29**(3), 454.

Faye, J. (2019), Copenhagen interpretation of quantum mechanics, *in* E. N. Zalta, ed., 'The Stanford Encyclopedia of Philosophy', winter 2019 edn, Metaphysics Research Lab, Stanford University.

Fisher, M. E. (1998), 'Renormalization group theory: Its basis and formulation in statistical physics', *Reviews of Modern Physics* **70**(2), 653.

Formaggio, J., Kaiser, D., Murskyj, M. and Weiss, T. (2016), 'Violation of the leggett-garg inequality in neutrino oscillations', *Physical Review Letters* **117**(5), 050402.

Franklin, A. (1989), *The neglect of experiment*, Cambridge University Press.

Fraser, D. (2009), 'Quantum field theory: Underdetermination, inconsistency, and idealization', *Philosophy of Science* **76**(4), 536–567.

Fraser, J. D. (2017), 'The real problem with perturbative quantum field theory', *The British Journal for the Philosophy of Science* .

Fraser, J. D. (2018*a*), 'Renormalization and the formulation of scientific realism', *Philosophy of Science* **85**(5), 1164–1175.

Fraser, J. D. (2018*b*), 'Towards a realist view of quantum field theory'.

Fraser, J. D. (2020), 'of quantum field theory', *Scientific Realism and the Quantum* p. 276.

Frauchiger, D. and Renner, R. (2018), 'Quantum theory cannot consistently describe the use of itself', *Nature communications* **9**(1), 3711.

Freedman, S. J. and Clauser, J. F. (1972), 'Experimental test of local hidden-variable theories', *Physical Review Letters* **28**(14), 938.

French, S. (2012), 'Unitary inequivalence as a problem for structural realism', *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* **43**(2), 121–136.

French, S. (2014), *The structure of the world: Metaphysics and representation*, Oxford University Press.

French, S. and Ladyman, J. (2003), 'Remodelling structural realism: Quantum physics and the metaphysics of structure', *Synthese* **136**(1), 31–56.

Fuchs, C. A. (2016), 'On participatory realism', *arXiv preprint arXiv:1601.04360* .

Fuchs, C. A. (2017), 'Notwithstanding bohr, the reasons for qbism', *Mind and Matter* **15**(2), 245–300.

Fuchs, C. A., Mermin, N. D. and Schack, R. (2014), 'An introduction to qbism with an application to the locality of quantum mechanics', *American Journal of Physics* **82**(8), 749–754.

Fuchs, C. A. and Schack, R. (2011), 'A quantum-bayesian route to quantum-state space', *Foundations of Physics* **41**(3), 345–356.

Fuchs, C. A. and Schack, R. (2012), Bayesian conditioning, the reflection principle, and quantum decoherence, *in* 'Probability in physics', Springer, pp. 233–247.

Fuchs, C. A. and Schack, R. (2013), 'Quantum-bayesian coherence', *Reviews of Modern Physics* **85**(4), 1693.

Fuchs, C. A. and Stacey, B. C. (2016), 'Qbism: Quantum theory as a hero's handbook', *arXiv preprint arXiv:1612.07308* .

Fuchs, C. A. and Stacey, B. C. (2019), 'Qbism: Quantum theory as a hero's handbook', *Foundations of Quantum Theory* **197**, 133.

Garton, W. and Tomkins, F. (1969), 'Diamagnetic zeeman effect and magnetic configuration mixing in long spectral series of ba i', *The Astrophysical Journal* **158**, 839.

Gell-Mann, M. and Hartle, J. (1990), 'Complexity, entropy and the physics of information', *SFI Studies in the Sciences of Complexity* **8**, 425.

Gendler, T. S. (1998), 'Galileo and the indispensability of scientific thought experiment', *The British Journal for the Philosophy of Science* **49**(3), 397–424.

Ghirardi, G. (2018), Collapse theories, *in* E. N. Zalta, ed., 'The Stanford Encyclopedia of Philosophy', fall 2018 edn, Metaphysics Research Lab, Stanford University.

Ghirardi, G. C., Pearle, P. and Rimini, A. (1990), 'Markov processes in hilbert space and continuous spontaneous localization of systems of identical particles', *Physical Review A* **42**(1), 78.

Ghirardi, G. C., Rimini, A. and Weber, T. (1986), 'Unified dynamics for microscopic and macroscopic systems', *Physical review D* **34**(2), 470.

Giere, R. N. (2010), *Scientific perspectivism*, University of Chicago Press.

Goff, A. (2006), 'Quantum tic-tac-toe: A teaching metaphor for superposition in quantum mechanics', *American journal of physics* **74**(11), 962–973.

Goldstein, S. (2017), Bohmian mechanics, *in* E. N. Zalta, ed., 'The Stanford Encyclopedia of Philosophy', summer 2017 edn, Metaphysics Research Lab, Stanford University.

Gomatam, R. (2007), 'Niels bohr‚Äôs interpretation and the copenhagen interpretation -‚Äî are the two incompatible?', *Philosophy of Science* **74**(5), 736–748.

Greenberger, D. M., Horne, M. A. and Zeilinger, A. (1989), Going beyond bell's theorem, *in* 'Bell‚Äôs theorem, quantum theory and conceptions of the universe', Springer, pp. 69–72.

Griffiths, R. B. (1972), 'Rigorous results and theorems', *Phase transitions and critical phenomena* **1**, 7–109.

Griffiths, R. B. (2003), *Consistent quantum theory*, Cambridge University Press.

Griffiths, R. B. (2013), 'A consistent quantum ontology', *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* **44**(2), 93–114.

Griffiths, R. B. (2019), The consistent histories approach to quantum mechanics, *in* E. N. Zalta, ed., 'The Stanford Encyclopedia of Philosophy', summer 2019 edn, Metaphysics Research Lab, Stanford University.

Hackermüller, L., Hornberger, K., Brezger, B., Zeilinger, A. and Arndt, M. (2004), 'Decoherence of matter waves by thermal emission of radiation', *Nature* **427**(6976), 711–714.

Hacking, I. (1983), *Representing and intervening*, Vol. 279, Cambridge University Press Cambridge.

Hardy, L. (1992), 'Quantum mechanics, local realistic theories, and lorentz-invariant realistic theories', *Physical Review Letters* **68**(20), 2981.

Hardy, L. (2017), Bell inequalities with retarded settings, *in* 'Quantum [Un] Speakables II', Springer, pp. 261–272.

Harker, D. (2012), 'How to split a theory: Defending selective realism and convergence without proximity', *The British Journal for the Philosophy of Science* **64**(1), 79–106.

Harrigan, N. and Spekkens, R. W. (2010), 'Einstein, incompleteness, and the epistemic view of quantum states', *Foundations of Physics* **40**(2), 125–157.

Hassoun, G. Q. and Kobe, D. H. (1989), 'Synthesis of the planck and bohr formulations of the correspondence principle', *American Journal of Physics* **57**(7), 658–662.

Healey, R. (2013), 'Observation and quantum objectivity', *Philosophy of Science* **80**(3), 434–453.

Heisenberg, W. (1949), *The physical principles of the quantum theory*, Courier Corporation.

Heller, E. J. and Tomsovic, S. (1993), 'Postmodern quantum mechanics', *Physics Today* **46**(7), 38–46.

Henderson, J. R. (2010), 'Classes of copenhagen interpretations: Mechanisms of collapse as typologically determinative', *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* **41**(1), 1–8.

Hesse, M. (1963), 'Models and analogies in science', *Sheed and Ward, London* .

Hopp, W. (2014), 'Experiments in thought', *Perspectives on Science* **22**(2), 242–263.

Hornberger, K. and Sipe, J. E. (2003), 'Collisional decoherence reexamined', *Physical Review A* **68**(1), 012105.

Howard, D. (2004), 'Who invented the "copenhagen interpretation"? a study in mythology', *Philosophy of Science* **71**(5), 669–682.

Jacques, V., Wu, E., Grosshans, F., Treussart, F., Grangier, P., Aspect, A. and Roch, J.-F. (2007), 'Experimental realization of wheeler's delayed-choice gedanken experiment', *Science* **315**(5814), 966–968.

Jaeger, G. (1998), 'The ehrenfest classification of phase transitions: introduction and evolution', *Archive for history of exact sciences* **53**(1), 51–81.

Jaynes, E. T. (2003), *Probability theory: The logic of science*, Cambridge university press.

Joos, E. (1987), 'Comment on "unified dynamics for microscopic and macroscopic systems"', *Physical Review D* **36**(10), 3285.

Joos, E. and Zeh, H. D. (1985), 'The emergence of classical properties through interaction with the environment', *Zeitschrift für Physik B Condensed Matter* **59**(2), 223–243.

Joos, E., Zeh, H. D., Kiefer, C., Giulini, D. J., Kupsch, J. and Stamatescu, I.-O. (2013), *Decoherence and the appearance of a classical world in quantum theory*, Springer Science & Business Media.

Karnakov, B. M. and Krainov, V. P. (2013), Wkb approximation in quantum mechanics, *in* 'WKB Approximation in Atomic Physics', Springer, pp. 1–30.

Kochen, S. (1985), A new interpretation of quantum mechanics, *in* 'Symposium on the foundations of modern physics', Vol. 50, World Scientific, Singapore, pp. 151–169.

Kofler, J. and Brukner, Č. (2007), 'Classical world arising out of quantum physics under the restriction of coarse-grained measurements', *Physical review letters* **99**(18), 180403.

Kok, P. (2015), University Lecture notes, `http://www.pieter-kok.staff.shef.ac.uk/index.php?nav=teaching&sub=quantum`.

Kösem, Ş. D. and Özdemir, Ö. F. (2014), 'The nature and role of thought experiments in solving conceptual physics problems', *Science & Education* **23**(4), 865–895.

Krimsky, S. (1973), 'The use and misuse of critical gedankenexperimente', *Zeitschrift für allgemeine Wissenschaftstheorie* **4**(2), 323–334.

Kwiat, P., Weinfurter, H., Herzog, T., Zeilinger, A. and Kasevich, M. (1995), 'Experimental realization of interaction-free measurements', *Annals of the New York Academy of Sciences* **755**(1), 383–393.

Ladyman, J., Spurrett, D., Ross, D., Collier, J., Spurrett, D., Collier, J. G. et al. (2007), *Every thing must go: Metaphysics naturalized*, Oxford University Press on Demand.

Landau, L. D. and Lifshitz, E. M. (2013), *Quantum mechanics: non-relativistic theory*, Vol. 3, Elsevier.

Laudan, L. (1981), 'A confutation of convergent realism', *Philosophy of science* **48**(1), 19–49.

Laymon, R. (1991), 'Thought experiments by stevin, mach and gouy: thought experiments as ideal limits and as semantic domains', *Horowitz and Massey* pp. 167–91.

Leggett, A. (2005), 'The quantum measurement problem', *science* **307**(5711), 871–872.

Leggett, A. J. (1996), The superposition principle in macroscopic systems, *in* 'Foundations Of Quantum Mechanics In The Light Of New Technology: Selected Papers from the Proceedings of the First through Fourth International Symposia on Foundations of Quantum Mechanics', World Scientific, pp. 27–35.

Leggett, A. J. and Garg, A. (1985), 'Quantum mechanics versus macroscopic realism: Is the flux there when nobody looks?', *Physical Review Letters* **54**(9), 857.

Leibfried, D., Knill, E., Seidelin, S., Britton, J., Blakestad, R. B., Chiaverini, J., Hume, D. B., Itano, W. M., Jost, J. D., Langer, C. et al. (2005), 'Creation of a six-atom 'schrödinger cat' state', *Nature* **438**(7068), 639–642.

Lenhard, J. (2017), Thought experiments and simulation experiments. exploring hypothetical worlds, *in* 'The Routledge Companion to Thought Experiments'.

Li, B. (2015), 'Coarse-graining as a route to microscopic physics: the renormalization group in quantum field theory', *Philosophy of Science* **82**(5), 1211–1223.

Li, T. and Yin, Z.-Q. (2016), 'Quantum superposition, entanglement, and state teleportation of a microorganism on an electromechanical oscillator', *Science Bulletin* **61**(2), 163–171.

Liboff, R. L. (1984), 'The correspondence principle revisited', *Physics Today* **37**(2), 50–55.

Lockwood, M. (1996), "many minds'. interpretations of quantum mechanics', *The British journal for the philosophy of science* **47**(2), 159–188.

Lombardi, O. and Dieks, D. (2017), Modal interpretations of quantum mechanics, *in* E. N. Zalta, ed., 'The Stanford Encyclopedia of Philosophy', spring 2017 edn, Metaphysics Research Lab, Stanford University.

Lundeen, J. and Steinberg, A. (2009), 'Experimental joint weak measurement on a photon pair as a probe of hardy's paradox', *Physical review letters* **102**(2), 020404.

Ma, X.-s., Kofler, J. and Zeilinger, A. (2016), 'Delayed-choice gedanken experiments and their realizations', *Reviews of Modern Physics* **88**(1), 015005.

Mach, E. and Hiebert, E. N. (1976), On thought experiments, *in* 'Knowledge and error', Springer, pp. 134–147.

Main, J., Wiebusch, G., Holle, A. and Welge, K. (1986), 'New quasi-landau structure of highly excited atoms: The hydrogen atom', *Physical review letters* **57**(22), 2789.

Mainwood, P. (2006), 'Is more different? emergent properties in physics'.

Maldacena, J. (2016), 'A model with cosmological bell inequalities', *Fortschritte der Physik* **64**(1), 10–23.

Marlow, A. (2012), *Mathematical foundations of quantum theory*, Elsevier.

Massimi, M. (2016), 'Bringing real realism back home: A perspectival slant', *The Philosophy of Philip Kitcher* .

Massimi, M. (2018), 'Four kinds of perspectival truth', *Philosophy and Phenomenological Research* **96**(2), 342–359.

Maudlin, T. (1995), 'Three measurement problems', *topoi* **14**(1), 7–15.

McComb, G. (2012), Thought experiment, definition, and literary fiction, *in* 'Thought Experiments in Science, Philosophy, and the Arts', Routledge, pp. 221–236.

Mermin, N. D. (2014), 'Physics: Qbism puts the scientist back into science', *Nature News* **507**(7493), 421.

Mermin, N. D. (2017), Why qbism is not the copenhagen interpretation and what john bell might have thought of it, *in* 'Quantum [Un] Speakables II', Springer, pp. 83–93.

Mermin, N. D. (2018), 'Making better sense of quantum mechanics', *Reports on Progress in Physics* **82**(1), 012002.

Mott, N. F. (1929), The wave mechanics of α-ray tracks, *in* 'Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences', Vol. 126, The Royal Society, pp. 79–84.

Nersessian, N. J. (1992), In the theoretician's laboratory: Thought experimenting as mental modeling, *in* 'PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association', Vol. 1992, Philosophy of Science Association, pp. 291–301.

Norton, J. (1991), 'Thought experiments in einstein's work', *Horowitz and Massey* **1991**, 129–148.

Norton, J. D. (1993), 'Seeing the laws of nature', *Metascience* **3**, 33–38.

Norton, J. D. (2004), Why thought experiments do not transcend empiricism, *in* 'C. Hitchcock (ed.), Contemporary Debates in the Philosophy of Science', Oxford: Blackwell, pp. 44–66.

Norton, J. D. (2012), 'Approximation and idealization: Why the difference matters', *Philosophy of Science* **79**(2), 207–232.

Norton, J. D. (2013), Chasing the light: Einstein's most famous thought experiment., *in* 'M. Frappier, L. Meynell, & J. R. Brown (Eds.), Thought Experiments in Philosophy, Science, and the Arts', New York & London: Routledge., pp. 123–140.

Omnés, R. (1988), 'Logical reformulation of quantum mechanics. i. foundations', *Journal of Statistical Physics* **53**(3-4), 893–932.

Onsager, L. (1944), 'Crystal statistics. i. a two-dimensional model with an order-disorder transition', *Physical Review* **65**(3-4), 117.

Oppenheim, J. and Wehner, S. (2010), 'The uncertainty principle determines the nonlocality of quantum mechanics', *Science* **330**(6007), 1072–1074.

Palacios, P. (2018), 'Had we but world enough, and time... but we don't!: Justifying the thermodynamic and infinite-time limits in statistical mechanics', *Foundations of Physics* **48**(5), 526–541.

Pawłowski, M., Paterek, T., Kaszlikowski, D., Scarani, V., Winter, A. and Żukowski, M. (2009), 'Information causality as a physical principle', *Nature* **461**(7267), 1101–1104.

Peacock, K. A. (2016), 'Happiest thoughts: Great thought experiments of modern physics'.

Pearle, P. (1976), 'Reduction of the state vector by a nonlinear schrödinger equation', *Physical Review D* **13**(4), 857.

Pearle, P. (1979), 'Toward explaining why events occur', *International Journal of Theoretical Physics* **18**(7), 489–518.

Pearle, P. (1989), 'Combining stochastic dynamical state-vector reduction with spontaneous localization', *Physical Review A* **39**(5), 2277.

Penrose, R. (1989), 'The emperor's new mind (new preface (1999) ed.)'.

Penrose, R. (1996), 'On gravity's role in quantum state reduction', *General relativity and gravitation* **28**(5), 581–600.

Penrose, R. (2006), *The road to reality: A complete guide to the laws of the universe*, Random house.

Peres, A. (2002), 'Karl popper and the copenhagen interpretation', *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics* **33**(1), 23–34.

Pitowsky, I. (2006), Quantum mechanics as a theory of probability, *in* 'Physical theory and its interpretation', Springer, pp. 213–240.

Popescu, S. and Rohrlich, D. (1994), 'Quantum nonlocality as an axiom', *Foundations of Physics* **24**(3), 379–385.

Popper, K. (1959), 'On the use and misuse of imaginary experiments, especially in quantum theory', *The logic of scientific discovery* pp. 442–56.

Proietti, M., Pickston, A., Graffitti, F., Barrow, P., Kundys, D., Branciard, C., Ringbauer, M. and Fedrizzi, A. (2019), 'Experimental rejection of observer-independence in the quantum world', *arXiv preprint arXiv:1902.05080* .

Psillos, S. (2005), *Scientific realism: How science tracks truth*, Routledge.

Pusey, M. (2016), 'Is qbism 80% complete, or 20%?', `https://www.youtube.com/watch?v=_9Rs61l8MyY`.

Pusey, M. F., Barrett, J. and Rudolph, T. (2012), 'On the reality of the quantum state', *Nature Physics* **8**(6), 475–478.

Putnam, H. et al. (1975), *Mathematics, Matter and Method: Volume 1, Philosophical Papers*, CUP Archive.

Rae, A. I. M. (2002), *Quantum Mechanics*, Institute of Physics Publishing.

Redhead, M. L. (1982), Quantum field theory for philosophers, *in* 'PSA: proceedings of the biennial meeting of the Philosophy of Science Association', Vol. 1982, Philosophy of Science Association, pp. 57–99.

Reiner, R. and Pierson, R. (1995), 'Hacking's experimental realism: an untenable middle ground', *Philosophy of Science* **62**(1), 60–69.

Renninger, M. (1953), 'Zum wellen-korpuskel-dualismus', *Zeitschrift für Physik* **136**(3), 251–261.

Rescher, N. (2005), *What if?: thought experimentation in philosophy*, Transaction Publishers.

Rohrlich, D. (2014), Pr-box correlations have no classical limit, *in* 'Quantum Theory: A Two-Time Success Story', Springer, pp. 205–211.

Rohrlich, D. and Hetzroni, G. (2016), 'Ghz states and pr boxes in the classical limit', *arXiv preprint arXiv:1606.04274* .

Ruetsche, L. (2011), *Interpreting quantum theories*, Oxford University Press.

Ruetsche, L. (2018), 'Renormalization group realism: The ascent of pessimism', *Philosophy of Science* **85**(5), 1176–1189.

Ruetsche, L. (2020), 'Perturbing realism', *Scientific Realism and the Quantum* p. 293.

Saatsi, J. (2017), 'Replacing recipe realism', *Synthese* **194**(9), 3233–3244.

Saatsi, J. (2019), 'What is theoretical progress of science?', *Synthese* **196**(2), 611–631.

Salmon, W. C. et al. (1989), 'Four decades of scientific explanation', *Scientific explanation* **13**, 3–219.

Sanz, A. and Borondo, F. (2003), 'A bohmian view on quantum decoherence', *arXiv preprint quant-ph/0310096* .

Saunders, S. (1993), 'Decoherence, relative states, and evolutionary adaptation', *Foundations of physics* **23**(12), 1553–1585.

Saunders, S., Barrett, J., Kent, A. and Wallace, D. (2010), 'Chance in the everett interpretation', *Many worlds* pp. 181–205.

Schack, R. (2016), 'Participatory realism', `https://www.youtube.com/watch?v=lhbK-_mGJs8`.

Schlosshauer, M. (2005), 'Decoherence, the measurement problem, and interpretations of quantum mechanics', *Reviews of Modern physics* **76**(4), 1267.

Schlosshauer, M. (2014), 'The quantum-to-classical transition and decoherence', *arXiv preprint arXiv:1404.2635* .

Schlosshauer, M. A. (2007), *Decoherence: and the quantum-to-classical transition*, Springer Science & Business Media.

Schlosshauer, M. and Fine, A. (2003), 'On zurek's derivation of the born rule', *arXiv preprint quant-ph/0312058* .

Schrödinger, E. (1935), 'The present status of quantum mechanics', *Die Naturwissenschaften* **23**(48), 1–26.

Sorensen, R. A. (1992), *Thought experiments*, New York: Oxford University Press.

Sorensen, R. A. (1998), *Thought experiments*, Oxford University Press on Demand.

Stapp, H. P. (1975), 'Bell's theorem and world process', *Il Nuovo Cimento B (1971-1996)* **29**(2), 270–276.

Stuart, M. T. (2015), Thought experiments in science, PhD thesis, University of Toronto.

Stuart, M. T. (2016), 'Norton and the logic of thought experiments', *Axiomathes* **26**(4), 451–466.

Tegmark, M. (1998*a*), 'The interpretation of quantum mechanics: Many worlds or many words?', *Fortschritte der Physik: Progress of Physics* **46**(6-8), 855–862.

Tegmark, M. (1998*b*), 'Quantum immortality', `https://space.mit.edu/home/tegmark/quantum.html#immortality`.
Accessed: 2020-04-28.

Tegmark, M. (2007), 'Many lives in many worlds', *Nature* **448**(7149), 23–24.

Teller, P. (1973), 'Conditionalization and observation', *Synthese* **26**(2), 218–258.

Timpson, C. (2010), 'Rabid dogma? comments on bub and pitowsky', *MANY WORLDS?* p. 460.

Timpson, C. G. (2008), 'Quantum bayesianism: a study', *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* **39**(3), 579–609.

Toroš, M. and Bassi, A. (2018), 'Bounds on quantum collapse models from matter-wave interferometry: calculational details', *Journal of Physics A: Mathematical and Theoretical* **51**(11), 115302.

Toroš, M., Donadi, S. and Bassi, A. (2016), 'Bohmian mechanics, collapse models and the emergence of classicality', *Journal of Physics A: Mathematical and Theoretical* **49**(35), 355302.

Torres, T., Patrick, S., Coutant, A., Richartz, M., Tedford, E. W. and Weinfurtner, S. (2016), 'Observation of superradiance in a vortex flow', *arXiv preprint arXiv:1612.06180* .

Torres, T., Patrick, S., Coutant, A., Richartz, M., Tedford, E. W. and Weinfurtner, S. (2017), 'Rotational superradiant scattering in a vortex flow', *Nature Physics* **13**(9), 833–836.

Trimmer, J. D. (1980), 'The present situation in quantum mechanics: A translation of schrödinger's "cat paradox" paper', *Proceedings of the American Philosophical Society* pp. 323–338.

Tumulka, R. (2006), 'A relativistic version of the ghirardi–rimini–weber model', *Journal of Statistical Physics* **125**(4), 821–840.

Valentini, A. (1991), 'Signal-locality, uncertainty, and the subquantum h-theorem. ii', *Physics Letters A* **158**(1-2), 1–8.

van Fraassen, B. C. (1997), Structure and perspective: Philosophical perplexity and paradox, *in* 'Logic and scientific methods', Springer, pp. 511–530.

van Fraassen, B. C. (2006), 'Structure: Its shadow and substance', *The British Journal for the Philosophy of Science* **57**(2), 275–307.

van Fraassen, B. C. (2010), *Scientific Representation: Paradoxes of Perspective*, OUP Oxford.

van, Fraassen, B. C. et al. (1991), *Quantum mechanics: An empiricist view*, Oxford University Press.

van Fraassen, C. (1984), 'Belief and the will', *The Journal of Philosophy* **81**(5), 235–256.

Vermaas, P. E. and Dieks, D. (1995), 'The modal interpretation of quantum mechanics and its generalization to density operators', *Foundations of Physics* **25**(1), 145–158.

Vinante, A., Bahrami, M., Bassi, A., Usenko, O., Wijts, G. and Oosterkamp, T. (2016), 'Upper bounds on spontaneous wave-function collapse models using millikelvin-cooled nanocantilevers', *Physical review letters* **116**(9), 090402.

Von Neumann, J. (2018), *Mathematical foundations of quantum mechanics: New edition*, Princeton University Press.

Wallace, D. (2006), 'In defence of naiveté: The conceptual status of lagrangian quantum field theory', *Synthese* **151**(1), 33–80.

Wallace, D. (2010), 'Decoherence and ontology: Or: How i learned to stop worrying and love fapp', *Many worlds* pp. 53–72.

Wallace, D. (2012), 'Decoherence and its role in the modern measurement problem', *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **370**(1975), 4576–4593.

Weihs, G., Jennewein, T., Simon, C., Weinfurter, H. and Zeilinger, A. (1998), 'Violation of bell's inequality under strict einstein locality conditions', *Physical Review Letters* **81**(23), 5039.

Welch, G. R., Kash, M. M., Iu, C.-h., Hsu, L. and Kleppner, D. (1989), 'Positive-energy structure of the diamagnetic rydberg spectrum', *Physical review letters* **62**(17), 1975.

Wheeler, J. A. (1978), The "past" and the "delayed-choice" double-slit experiment, *in* 'Mathematical foundations of quantum theory', Elsevier, pp. 9–48.

Wheeler, J. A. (1979), Frontiers of time, *in* 'Problems in the Foundations of Physics'.

Wheeler, J., Zurek, W. and Ballentine, L. E. (1984), 'Quantum theory and measurement', *American Journal of Physics* **52**, 955–955.

Wigner, E. (1961), 'Remarks on the mind-body question. reprinted in wheeler and zurek, eds., quantum theory and measurement'.

Wigner, E. P. (1967), *Symmetries and reflections*, Indiana University Press.

Williams, P. (2017), 'Scientific realism made effective', *The British Journal for the Philosophy of Science* .

Wimsatt, W. C. (2007), *Re-engineering philosophy for limited beings: Piecewise approximations to reality*, Harvard University Press.

Worrall, J. (1989), 'Structural realism: The best of both worlds?', *Dialectica* **43**(1-2), 99–124.

Worrall, J. (2007), 'Miracles and models: Why reports of the death of structural realism may be exaggerated', *Royal Institute of Philosophy Supplements* **61**, 125–154.

Zeh, H. D. (1970), 'On the interpretation of measurement in quantum theory', *Foundations of Physics* **1**(1), 69–76.

Zurek, W. H. (1981), 'Pointer basis of quantum apparatus: Into what mixture does the wave packet collapse?', *Physical review D* **24**(6), 1516.

Zurek, W. H. (1982), 'Environment-induced superselection rules', *Physical review D* **26**(8), 1862.

Zurek, W. H. (1993), 'Preferred states, predictability, classicality and the environment-induced decoherence', *Progress of Theoretical Physics* **89**(2), 281–312.

Zurek, W. H. (1998), 'Decoherence, einselection and the existential interpretation (the rough guide)', *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* **356**(1743), 1793–1821.

Zurek, W. H. (2003), 'Decoherence and the transition from quantum to classical–revisited', *arXiv preprint quant-ph/0306072* .

Zurek, W. H. (2005), 'Probabilities from entanglement, born's rule from envariance', *Physical Review A* **71**(5), 052105.

Zurek, W. H. (2009), 'Quantum darwinism', *Nature Physics* **5**(3), 181.

Zurek, W. H. (2014), 'Quantum darwinism, classical reality, and the randomness of quantum jumps', *arXiv preprint arXiv:1412.5206* .