1 Introduction

One of the most important goals of physics is to come up with theories which successfully predict observable physical phenomena. But even when a theory is successful in the empirical sense, there remain many questions to be answered. What, if anything, does a successful predictive algorithm tell us about the way the world is? What ontology does the theory suggest? How does the theory relate to other successful physical theories? How does it relate to our own experiences? What does the theory tell us about familiar concepts like space, time, and matter?

In attempting to answer these questions, we are engaging in what is known as foundations of physics. Quantum foundations is the branch of foundations of physics that relates to quantum mechanics, the physical theory developed in the early twentieth century which applies to phenomena at very small scales. While there are certainly interesting foundational problems to be addressed for all of our major physical theories,[1] quantum foundations has attracted an unusual amount of attention because there seems to be something uniquely puzzling about quantum mechanics. Most famously, there exists no consensus about the physical interpretation of the predictive algorithm that quantum mechanics gives us, and many of the alternatives under consideration paint a picture of a physical world which is radically different from the classical world of our pre-quantum imaginings.

Although there is a significant intersection between foundational questions and the traditional domain of philosophy, quantum foundations is distinct from the philosophy of quantum mechanics in that the two fields typically involve different methodological approaches. One of the important innovations made by Bell in proving his famous theorem, which we will discuss in Section 3, was that it is possible to address foundational questions in a *quantitative* way by means of mathematical proof, and subsequently the field of quantum foundations has applied similar quantitative approaches to many other conceptual questions. It is this mathematical approach which makes quantum foundations a branch of physics rather than philosophy – although of course researchers in quantum foundations have many interests in common with their peers in the philosophy of quantum mechanics, and the field can be understood as sitting at the intersection of physics, maths, and philosophy.

Before we go further, we should address head-on the not uncommon view that making empirical predictions is the *only* goal of physics, and that it is

[1] For an overview of some interesting problems in the foundations of special and general relativity, see Maudlin (2012); for statistical mechanics, see Sklar (1993); and for classical mechanics, see Sklar (2012).

therefore pointless to worry about interpretations or to ask any of the other conceptual questions with which quantum foundations is concerned. As a first response, we observe that many people with an interest in science are driven not only by the desire to make predictions but also by the desire to understand how the world works, and quantum foundations exists in part to satisfy this intellectual thirst. Of course it is true that we can never be completely certain that the conclusions we reach in this endeavour are correct, but then one can never be completely certain that one's predictive theories are universally correct either.

Furthermore, quantum foundations would be important even if the only purpose of science *were* to make predictions, because thinking deeply about the nature of quantum mechanics and coming to a better understanding of the physical reality from which it arises is likely to lead to progress on outstanding problems in physics. This is particularly important right now because many physicists have come to believe that fundamental physics is in a state of stagnation, with little meaningful progress having been made in the last few decades (Hossenfelder 2018). It seems entirely possible that this has arisen because physicists never properly got to grips with what quantum mechanics tells us about the world, and therefore all subsequent physics has been based on an improper understanding of the earlier theory, leading us into a dead end. Thus quantum foundations is not merely an exercise in intellectual curiosity – it may be the best hope we have of breaking out of the impasse that physics seems to have found itself in. After all, many historic advances in physics have resulted from thinking deeply about conceptual questions; for example, Einstein's theory of special relativity is the result of asking foundational questions about the nature of time and simultaneity (Einstein 1905).

This Element offers a short tour through some important topics in quantum foundations. The next section introduces the basics of quantum mechanics together with some ideas and notation that we will use throughout this Element. Since the motivating principle of quantum foundations is to address conceptual questions with mathematical methods, in Sections 3, 4, and 5 we introduce an important mathematical result from the field and then discuss the conceptual issues linked to it. In Section 6 we give a brief summary of some other areas of quantum foundations that due to considerations of space cannot be covered in detail here. Finally, in the concluding section, we give an assessment of the current state of quantum foundations and make some suggestions about what its future might look like.

2 Preliminaries

2.1 What Is Quantum Mechanics?

Quantum mechanics has its origins in a number of apparently minor problems that puzzled physicists in the late nineteenth and early twentieth centuries. One of these was the 'black-body problem', which was solved by Max Planck in 1900 using the hypothesis that energy is radiated and absorbed in discrete packets known as quanta. Another was the 'photoelectric effect', for which Einstein proposed a similarly quantum solution in 1905. Subsequently, in 1913, Niels Bohr came up with a quantised theory of atomic structure to explain Ernest Rutherford's experimental observations. These early quantum ideas were developed over the first half of the twentieth century by physicists such as Schrödinger, Heisenberg, Born, von Neumann, Dirac, Pauli, Hilbert, and many others, and the theory that emerged has become known as quantum mechanics.[2]

Quantum mechanics is in many ways more a methodological prescription than a concrete scientific theory (Gell-Mann 1980, Nielsen and Chuang 2011), since it sets out a mathematical framework for the construction of physical theories that must be supplemented with detailed experimental work to determine which specific mathematical objects represent the actual physical systems whose behaviour we would like to predict. However, the field of quantum foundations is largely concerned with this abstract structure rather than with any specific realisation of it, and hence for us it is sufficient to regard quantum mechanics as being characterised by the following four postulates (Nielsen and Chuang 2011) (see Strang (2016) for an introduction to the linear algebra terminology used in these postulates):

1. To every physical system we ascribe a Hilbert space, $\mathscr{H}$, known as the state space of the system.[3] At any given time, the system is completely described by its state vector, which is a unit vector $|\psi\rangle$ in the state space.
2. Closed quantum systems evolve by unitary transformations.[4] In particular, a closed quantum system can be associated with a fixed Hermitian operator[5] H, and the time evolution of the state of the system is then given by $H|\psi(t)\rangle = i\hbar \frac{d|\psi(0)\rangle}{dt}$ (i.e., the Schrödinger equation).

[2] See Lindley (2008) for an engaging account of the early history of quantum mechanics.
[3] A Hilbert space is a complex vector space equipped with an inner product.
[4] A unitary transformation is a transformation that preserves the value of the inner product; unitary operators U satisfy $U^\dagger U = \mathbb{I}$, where $U^\dagger$ denotes the conjugate transpose of U and $\mathbb{I}$ denotes the identity operator.
[5] A Hermitian operator is an operator that is equal to its own conjugate transpose.

3. A measurement is described by a projective measurement, which is a Hermitian operator on the state space of the system. The projective measurement can be decomposed into a set of operators $\{P_m\}$ sum to the identity operator, each associated with a measurement outcome. When a system is prepared in the state $|\psi\rangle$ and the measurement $\{P_m\}$ is performed, the probability of obtaining the outcome associated with P_m is equal to $Tr(P_m|\psi\rangle\langle\psi|)$, where $Tr(..)$ denotes the trace and $|\psi\rangle\langle\psi|$ denotes the outer product of the state vector $|\psi\rangle$ with itself; after this result has been obtained, the state of the system is $\frac{P_m|\psi\rangle}{\sqrt{Tr(P_m|\psi\rangle\langle\psi|)}}$.
4. When we combine two physical systems, the state space for the resulting composite system is the tensor product of the individual state spaces; if we combine n systems prepared in states $|\psi_1\rangle, |\psi_2\rangle, \ldots, |\psi_n\rangle$, the resulting joint state is $|\psi_1\rangle \otimes |\psi_2\rangle \otimes \ldots \otimes |\psi_n\rangle$.

2.1.1 Quantum States

Let's consider an example. Take a quantum system with a Hilbert space of dimension two, such as a particle that can have spin pointing up or down. We will write the state 'spin up' as the vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ or equivalently the ket $|0\rangle$ and the state 'spin down' as the vector $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ or equivalently $|1\rangle$. The inner product of the vectors $|\psi\rangle$ and $|\phi\rangle$, written as $\langle\phi|\psi\rangle$, is given by multiplying each entry in $|\psi\rangle$ by the conjugate transpose of the corresponding entry in $|\phi\rangle$, so for example we have $\langle 0|1\rangle = 1\times 0 + 0\times 1 = 0$. By convention, quantum states are normalised so that for any state $|\psi\rangle$, $\langle\psi|\psi\rangle = 1$.

Suppose we apply some measurement $M = \{P_0, P_1\}$ to the system when it is in state $|0\rangle$. The measurement operators P_0 and P_1 are both represented by 2×2 matrices. One possible measurement we can perform on this system is a spin measurement, where the operator $P_0 = |0\rangle\langle 0|$ is associated with the property 'spin up' and the operator $P_1 = |1\rangle\langle 1|$ is associated with the property 'spin down'. If we perform this measurement when the system is in state $|0\rangle$, the probability of obtaining the outcome 'spin up' is $Tr(P_0|0\rangle\langle 0|)$, which can be rearranged to $\langle 0|P_0|0\rangle = \langle 0|0\rangle\langle 0|0\rangle$ which equals 1 due to the normalisation of quantum states. So the probability of obtaining the outcome 'spin up' when the system is in the state 'spin up' is equal to 1, and a similar calculation shows that the probability of obtaining the outcome 'spin down' when the system is in the state 'spin down' is equal to 1, exactly as we would expect.

Now if this were a classical system, then spin up and spin down would be the only possibilities for the state. But because this is a quantum system, the state can also be a vector like $\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$ or, in quantum-mechanical notation, $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. In this state, the particle's spin is neither up nor down; it is an equal superposition of both, and when we measure its spin the probability of getting a 'spin up' result is $Tr(P_0 \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\frac{1}{\sqrt{2}}(\langle 0| + \langle 1|)$, which comes out to 0.5. That is, we have a 50 per cent chance of getting the result up and a 50 per cent chance of getting the result down. Superpositions like this allow a continuous range of states between 'spin up' and 'spin down'. Note that by applying the formula in postulate three, we find that after the result 'spin up' is obtained, the system will always be in the state $|0\rangle$, which corresponds to 'spin up', so even if the particle did not have a definite state of spin before the measurement, after the measurement it is in the 'spin up' state corresponding to the observed measurement outcome.

If this were a classical system, then we would always be able to find out its state with certainty, provided we were able to perform sufficiently sensitive measurements. But in quantum mechanics, this is no longer true. That is, given two arbitrary quantum states ψ_1, ψ_2, it will not always be possible to find a measurement $\{M_1, M_2\}$ such that if we perform this measurement when the system is in state ψ_1, we are guaranteed to get outcome M_1, and if we perform the measurement when the system is in state ψ_2, we are guaranteed to get outcome M_2. Usually, the best we will be able to do is to come up with a measurement $\{M_1, M_2\}$ such that if we perform this measurement when the system is in state ψ_1, we get outcome M_1 with some probability p_1, and if we perform this measurement when the system is in state ψ_2, we get outcome M_1 with probability $p_2 < p_1$. Thus, when we do perform the measurement and get outcome M_1, we will be able to conclude that it is more likely the state was ψ_1 rather than ψ_2.

However, there are certain sets of states for which it is possible to find distinguishing measurements. In particular, this is the case for any set of states that are all orthogonal to each other, where we say two states $|a\rangle$ and $|b\rangle$ are orthogonal if their inner product is zero, i.e. $\langle a|b\rangle = 0$. In real vector spaces with two or three dimensions, 'orthogonal' means the same as 'perpendicular', and in other vector spaces, it can be thought of as a generalisation of that notion. The states $|0\rangle$ and $|1\rangle$ in our example are orthogonal, since $\langle 0|1\rangle = 0$, and indeed they can be perfectly distinguished by the spin measurement $\{P_0, P_1\}$, since we will always get the outcome 'spin up' when the system is in state $|0\rangle$ and we will always get the outcome 'spin down' when the system is in state $|1\rangle$.

The fact that quantum states in general cannot be perfectly distinguished poses obvious practical challenges when it comes to performing and analysing quantum experiments, since it means we must draw our conclusions from the statistical ensemble of the results over a large number of experiments, rather than the result of any individual experiment. This is also a first indication that quantum states may have properties quite unlike classical states, and in forthcoming sections we will see that this is borne out in a variety of interesting ways.

2.1.2 Mixed States and Density Operators

On the face of it, the postulates we have set out suggest that the state of a quantum system must always be describable by a state vector in some Hilbert space. But, in fact, there are two ways in which we can obtain different types of states within this formalism. First, we can select a state from among a set of states $\{|\psi_i\rangle\}$ with some set of probabilities $\{p(\psi_i)\}$, giving rise to a probabilistic mixture $\rho = \sum_i p(\psi_i)|\psi_i\rangle\langle\psi_i|$; this is known as a proper mixture (d' Espagnat 1971, Busch et al. 1996). Second, we can take two systems A, B in an entangled state ψ (see Section 3) and then throw away the information about the state of A, leaving B in a reduced state $\rho = Tr_A(|\psi\rangle\langle\psi|)$; this is known as an improper mixture (d' Espagnat 1971, Busch et al. 1996). Happily, it turns out that these two methods of preparation give rise to exactly the same type of mathematical object – to wit, a density matrix, a positive Hermitian matrix of trace one. States that can be represented as state vectors are known as 'pure states', and these states can also be represented as density matrices, since the state vector $|\psi\rangle$ corresponds to the density matrix $|\psi\rangle\langle\psi|$, whereas states that can only be written as density matrices and have no state vector representation are known as 'mixed states'.

From the original four postulates set out in Section 2.1, we can derive statements about the behaviour of density matrices. The set of possible evolutions is expanded to include all evolutions that can be obtained by appending some other system to the original system and then applying a unitary evolution as described in postulate two to the whole, which leads to the set of all completely positive trace-preserving (CPTP) maps[6]; the set of possible measurements is expanded to include all measurements that can be implemented by appending an extra system to the original system and applying a projective measurement as defined in postulate three to the whole, which leads to the set of all positive operator valued measures (POVMs). A POVM is a set of positive semi-definite

[6] A completely positive trace-preserving map is an operator which does not change the trace of the matrices that it acts on (this is the 'trace-preserving' property), such that if we take the tensor product of this operator with the identity matrix and apply the result to a positive semi-definite matrix, the resulting matrix will still be a positive semi-definite (this is the 'completely positive' property).

operators[7] $\{K_n\}$ that sum to the identity operator; the probability of obtaining the result associated with the operator K_n when we the measured system has density matrix ρ is given by $Tr(K_n\rho)$, which is similar to the corresponding formula for a projective measurement. In postulate three, we were also able to give a formula for the state that a system will be in after a projective measurement, but unfortunately, this cannot be done for POVMs, because any given POVM can be implemented in a number of different ways and the post-measurement state depends on the particular implementation (Nielsen and Chuang 2011, Paris 2012, Landau and Lifshitz 2013). In fact, this is true of projective measurements as well (for example, sometimes in a projective measurement the state being measured is destroyed!), but postulate three provides a simple, natural update rule that works in most situations and is widely accepted as canonical. By contrast, for POVMs there is no such simple, canonical approach.

It should be noted that since the ideal of pure states, unitary operators, and projective measurements can seldom be perfectly realised in the laboratory, in real applications we are mostly dealing with mixed states, CPTP maps, and POVMs, rather than pure states, unitary transformations, and projective measurements (de Muynck 2007).

We pause at this point to reinforce that the pedagogical approach of deriving the existence of density matrices, CPTP maps, and POVMs from the four postulates set out in Section 2.1 – an approach known as the 'Church of the Larger Hilbert Space', (Timpson 2008) – is not entirely uncontroversial. There are also advocates of the 'Church of the Smaller Hilbert Space' who contend that the density matrix should be thought of as the fundamental object of quantum mechanics and that we have no reason to suppose that quantum systems cannot be in mixed states without being derived from either a larger pure state or a probabilistic mixture of pure states (Dürr et al. 2005, Allori et al. 2013, Weinberg 2014). We will not comment further on this question here, but we observe that the choice between the two Churches is closely related to the interpretational questions discussed in Section 5.

2.2 Ontological Models

A large part of the field of quantum foundations is concerned with understanding the nature of the reality from which quantum mechanics arises. Since we are interested in finding mathematical ways of addressing conceptual questions, it is necessary to have an appropriate mathematical framework in which to pose

[7] A matrix M is said to be positive semi-definite if for any vector v with n entries, $v^* \times M \times v$ is positive or zero. (v^* here denotes the conjugate transpose of v.)

our questions, and the framework which is most commonly used for this purpose is the *ontological models approach*. In its modern form, this approach was first put forward and developed by Rob Spekkens, but the motivating ideas for the formalism had been floating around in the field for some time previously, and thus for the sake of continuity we will prove Bell's theorem (see Section 3) using the ontological models framework, although Bell's result predates Spekkens' work and was originally proved using slightly different language. The Spekkens contextuality theorem in Section 4 and the PBR theorem in Section 5 will also be expressed using ontological models, since they were originally parsed in that framework.

The core of the ontological models approach is very simple: we suppose that any quantum system has a real underlying state, known as its 'ontic state'. The ontic state is determined by how the system was prepared and by any subsequent transformations which have been applied to it, and when measurements are made on the system, the results of those measurements can depend only on its ontic state. It is important to reinforce that the ontic state is not assumed to have anything to do with the quantum state – it might contain all the same information as the quantum state, or less information, or more information. And of course the ontological models approach is sufficiently general that we may attempt to make ontological models of theories which are not at all like quantum mechanics and which may not even have a concept of 'state'.

Mathematically, these ideas are represented as follows. For a given system, we suppose that there exists a space Λ of possible ontic states λ. To each preparation procedure P which can be performed on the system, we assign a probability distribution p_P over ontic states, such that $p_P(\lambda)$ gives the probability that the system will end up in the ontic state λ when we perform procedure P; to each transformation T which can be applied to the system, we assign a column-stochastic matrix T[8] describing how the distribution over ontic states is affected by this transformation; and to each measurement M which can be performed on the system and every possible outcome O of that measurement, we assign a response function $\xi_{M,O}$ such that $\xi_{M,O}(\lambda)$ gives the probability that outcome O will occur when we perform measurement M on a system which is in ontic state λ.

Suppose now that we aim for our ontological model to be capable of reproducing the empirical results of quantum theory. This imposes a number of constraints on the space of ontic states Λ, distributions p_P, transformations T and response functions $\xi_{M,O}$. For example, suppose we perform a preparation

[8] A column-stochastic matrix is a matrix containing only non-negative real values such that the columns of the matrix all sum to 1.

procedure P which, according to quantum mechanics, prepares the quantum state $|\psi\rangle$, and then we perform a measurement M with outcomes M_1, M_2 which, according to quantum mechanics, is represented by the POVM $\{O_1, O_2\}$: quantum mechanics tells us that the probability of obtaining outcome M_1 is equal to $Tr(O_1|\psi\rangle\langle\psi|)$. In order for our ontological model to reproduce this result, it must be the case that the probability of obtaining a given ontic state λ times the probability of obtaining outcome M_1 to measurement M when the system is in state λ, summed over all λ, gives $Tr(O_1|\psi\rangle\langle\psi|)$. That is:

$$\sum_{\lambda} p_P(\lambda)\xi_{M,M_1}(\lambda) = Tr(O_1|\psi\rangle\langle\psi|)$$

And if the set of ontic states is infinite, the sum becomes an integral:

$$\int p_P(\lambda)\xi_{M,M_1}(\lambda) = Tr(O_1|\psi\rangle\langle\psi|)$$

It is clear that imposing this requirement for all possible quantum mechanical preparations and measurements places very strong limitations on an ontological model, and much of the progress in quantum foundations over the last 50 years has essentially been concerned with following up the consequences of these limitations and attempting to understand what properties an ontological model must have if it is to faithfully reproduce quantum mechanics. The idea is that this will give us leading a better understanding of what the reality underlying quantum mechanics must look like in order to produce the empirical results that we have observed.

An important question about the ontological models approach concerns what exactly an ontic state should be ascribed to. We have said that every system is to be assigned an ontic state, but what is a 'system' in this context? For example, if we perform a joint preparation on two quantum systems and then separate them, should we assign separate ontic states to each of the particles, or a single joint state to both? We will see in Section 3 that there are good reasons to use a single joint state in at least some such cases, but this is not obvious before we start on the work of analysing quantum mechanics. Clearly, then, the ontological models approach has a certain vagueness, but this is actually a feature, not a bug: the whole idea is to have a framework general enough to encompass many different hypotheses about what the reality underlying quantum mechanics might look like, and in order to achieve this it is important not to be too prescriptive.

It should also be reinforced that one can make use of the formalism of ontological models without necessarily interpreting it as an attempt at a faithful representation of reality – indeed, Spekkens himself prefers to regard it as a classification schema which enables us to give precise mathematical definitions for concepts like contextuality (Spekkens, n.d.) (see Section 4). Nonetheless, it seems to be the case that within quantum foundations this formalism or something close to it is often regarded as a description of reality and perhaps as the only possible way of describing reality – for example, in Leifer and Pusey (2017), it is claimed that any model in which correlations are not explained by appeal to ontic states should not really be regarded as a realist model at all. Indeed we shall see that the view of reality enshrined in the ontological models formulation is so ubiquitous in quantum foundations that most of the important results of the field make sense only in that context.

2.3 Quantum Field Theory

Because particles in scattering experiments are frequently accelerated almost to the speed of light, it is not possible to neglect relativistic effects in the description of scattering experiments, and therefore in order to do particle physics accurately, it is necessary to make some adjustments to standard quantum mechanics. There exists a relativistic formulation of quantum mechanics where the Schrödinger equation is replaced by the Klein-Gordon and Dirac equations, but it turns out that this is not sufficient to allow us to study particle physics, because both non-relativistic and relativistic quantum mechanics are defined only for scenarios that can be described by a finite, constant number of degrees of freedom, whereas in particle physics it is necessary to describe fields with an infinite number of degrees of freedom and scattering processes in which particles may be created or destroyed. Thus, in order to apply quantum mechanics to this realm, it has been necessary to create an extension of the theory accommodating an infinite number of degrees of freedom. This extension, which has become known as quantum field theory, allows us to model successfully almost all features of elementary particle physics – gravity alone, among the fundamental forces, still resists being cast in this form (Peskin and Schroeder 1995, Lancaster and Blundell 2014).

One might ask at this juncture why anyone should persist in studying the foundations of quantum mechanics when quantum mechanics has already been supplanted by a more advanced theory. A first answer is that quantum field theory is essentially the result of applying quantum mechanics to a new domain and following up the varied and sometimes unexpected consequences, so the underlying principles of quantum field theory are very close to the underlying principles of quantum mechanics and one might expect that most of the interesting conceptual issues that arise in quantum field theory will already appear in quantum mechanics. Moreover, quantum field theory has many mathematical