PROGRESS IN QUANTUM FOUNDATIONS

Andre O. Ranchin

DEPARTMENT OF PHYSICS

IMPERIAL COLLEGE LONDON

Report submitted in partial fulfillment of the requirements of the degree of MRes of Imperial College

SEPTEMBER 2012

Abstract

We present a review of the current state of affairs in the foundations of quantum meachanics. After introducing the axioms of quantum theory, the first part of the report examines in detail the trials of interpreting quantum mechanics in a realist way. The main focus of the second half of the review is the attempt to describe quantum mechanics within a larger space of hypothetical theories. A particular emphasis is given to the study of Categorical Quantum Mechanics. This leads to the derivation of a new result: the complete set of quantum circuit equations for stabilizer quantum mechanics. We also suggest a number of future directions for research.

Acknowledgements

I am grateful to my supervisors Terry Rudolph and Bob Coecke for their helpful suggestions and also wish to thank Miriam Backens, Aleks Kissinger and Mihai Vidrighin for useful discussions. I acknowlwedge financial support from the EPSRC.

Contents

Chapter 1

Introduction

Our scientific theories aim to accurately describe every phenomenon that can possibly occur in the world we live in. However, one can hope that a theory will not only explain all observable occurrences and predict new results, but will also convey an understanding of the inner workings of nature; an insight into why things are the way they are.

Of course, any theory or model put forward to explain natural phenomena will have a limited domain of validity. Even within this restricted domain, the understanding provided by our imperfect human constructions is flawed; we only obtain an approximate truth, based upon defective axioms and unreliable reasoning. It is unclear whether there even is such a thing as objective truth. Indeed, a timeless, unquestionable reality independent of perspective may be a fallacy. Even if such a thing is possible in principle, then could a theory exhibiting the objective truth of some subclass of all objects and ideas ever be fathomed by our human minds?

Our current progress in seeking out a relatively consistent theory which allows us to approximately describe most observed physical phenomena has led us to the study of two complementary theories of nature: General Relativity and Quantum Mechanics. One would expect that if we manage to produce a set of consistent physical axioms from which we can derive both General Relativity and Quantum Mechanics as approximate emerging descriptions of the world, then we would be a step closer towards obtaining a theory exhibiting objective truth. Even an anti-realist or a subjectivist might praise this achievement for elegance and consistency alone, independently of the belief in an underlying reality, or ontological realm of objects and facts, that exists independently of the mind.

In order to attain this ambitious goal, a crucial first step is to work towards a more thorough understanding of the foundations of Quantum Mechanics. In addition to the revolution in physics this theory has already caused, an improved grasp of Quantum theory has the potential to lead to another considerable shift in our perspective on how the world works. Of course, developments of quantum mechanics go hand in hand with numerous technological advances, including lasers, transistors, quantum information and computation. In the following thesis, however, we will only describe technological developments insofar as they help us understand fundamental features of nature.

Quantum foundations has a number of distinct goals, aiming to further our understanding of quantum theory or quantum-like theories of nature. One of these is the search for and analysis of non-classical or quantum effects. These may reveal important quantum-like or classical-like features which the world may or may not exhibit. Another important aim of quantum foundations is to provide an adequate interpretation and a thorough analysis of the axioms of quantum theory. Combining these, we can then determine fundamental principles from which the quantum formalism can be derived. This should eventually lead to a thorough conceptual and mathematical underpinning of quantum theory without any inconsistencies.

An alternative approach to foundations is to examine a larger space of hypothetical theories containing quantum theory. This allows us to formalize and clarify the different ways in which the world could possibly work and where quantum mechanics fits with respect to these possibilities. Such an analysis of mathematical theories of nature requires a high level of abstraction guided by deep and novel physical intuition.

Analysis of the foundations of quantum mechanics may then bring out new ways to conceptualize the theory, suggesting new possible experiments and theory that would have been difficult to imagine without foundational insight. Such an inquiry should then pave the way towards alternative approaches in cases where we are presently uncertain how to go about applying the current quantum formalism. This methodology could even provide alternative theories which would supplant quantum theory. Such theories would have to either predict correctly the result of an experiment for which quantum mechanics is wrong or provide a modification of the standard theory which can lead to accurate novel predictions not provided previously.

Let us now sketch the state of affairs in quantum foundations.

Chapter 2

Quantum mechanics as it is

2.1 Orthodox postulates

A natural starting point for an analysis of the foundations of quantum theory would be the postulates of quantum mechanics:

2.1.1 Axiom 1

The physical state $|\psi\rangle$ of the system corresponds to a normalized ray of a Hilbert space H, known as the state space of the system.

2.1.2 Axiom 2

The evolution of a closed system is a unitary transformation: $|\psi(t)\rangle = U(t, t_0)|\psi(t_0)\rangle$ (such that $U^{-1} = U^{\dagger}$) depending only on the initial time t_0 and the final time t.

2.1.3 Axiom 3

Associated with each observable property of a system is a Hermitian operator M (a Hermitian operator satisfies $M = M^{\dagger}$ and has real eigenvalues and orthogonal eigenvectors.). $M = \sum_m mP_m$, where P_m is the projector onto the eigenspace of M with eigenvalue m. The possible results of a measurement of M on the state $|\psi\rangle$ are the eigenvalues m of M. The probability of getting outcome m is: $p(m) = \langle\psi|P_m|\psi\rangle$.

2.1.4 Axiom 4

Given that outcome m occurred, the state of the system changes discontinuously as: $|\psi\rangle \to \frac{P_m|\psi\rangle}{p(m)}$.

2.1.5 Axiom 5

If two systems $|\psi_1\rangle$ and $|\psi_2\rangle$ have state spaces H_1 and H_2 respectively then if we treat these two systems as one single compound system $|\psi_1\rangle \otimes |\psi_2\rangle$, the state space of the compound system is the tensor product $H_1 \otimes H_2$.

We can immediately notice several odd features of this set of postulates. This definition of physical states as elements of an abstract Hilbert space and the use of the tensor product to form composite systems seem arbitrary. There is an immediate clash between the deterministic and continuous evolution of closed systems and the indeterministic discontinuous evolution due to measurement. One might wonder how to interpret the quantum state, where the division lies between observer and observed or where the measurer is to be found if the system in question is the whole universe.

For now, we will delay these questions and take a minimalist, operational approach to quantum theory. Using this methodology, we find more general axioms for quantum theory.

2.2 Operational axioms

A useful way of interpreting a physical theory is to forget about all the inner workings specific to the given theory. One can argue that all empirical evidence perceptible by human beings is restricted to macroscopically distinguishable initializations and outcomes expressed in classical terms.

In this operational interpretation, the only role of a physical theory is to provide a minimal explanation of experimental phenomena. This can generally be done by providing a description of physical preparation (P), transformation (T) and measurement (M) procedures which yields correct statistics for experiments that can be done. In such a setting, the axioms of quantum theory can be reformulated as:

2.2.1 Axiom 1: Preparation

A preparation P is associated to a trace one positive operator ρ, known as the density operator, acting on the Hilbert space H.

Note that:

(i) If a system preparation is associated with $|\psi_i\rangle$ with probability p_i then the density operator corresponding to the overall preparation is $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$.

(ii) A preparation ρ is called a 'pure state' if $Tr(\rho^2)$=1. Otherwise $Tr(\rho^2) < 1$ and ρ is called a 'mixed state'.

(iii) Two preparations ρ_1 and ρ_2 can be combined as before into one single compound preparation using the tensor product: $\rho_{12} = \rho_1 \otimes \rho_2$.

(iv) Conversely, we can get one of the subspreparations by tracing out the other subpreparation: $\rho_1 = Tr_2(\rho_{12})$.

2.2.2 Axiom 2: Transformation

A transformation T is associated to a completely positive trace non-decreasing map $\mathcal{E}$: $\rho \to \mathcal{E}(\rho)$,

Such that:

(i) $0 \leq Tr(\mathcal{E}(\rho)) \leq 1$ for any preparation ρ

(ii) For probabilities $\{p_i\}$: $\mathcal{E}(\sum_i p_i\rho_i) = \sum_i p_i \mathcal{E}(\rho_i)$

(iii) $\mathcal{E}(A)$ and $(I \otimes \mathcal{E})(A)$ are positive for any positive operator A (I is the identity operator).

Note that (i), (ii) and (iii) are formally equivalent to either of the following:

(KRAUS) $\mathcal{E}(\rho) = \sum_i (E_i \rho E_i^\dagger)$ where $\sum_i (E_i^\dagger E_i) \leq 1$ and E_i are the Kraus operators.

(ANCILLA) $\mathcal{E}(\rho) = Tr_E(PU(\rho \otimes \rho_0)U^\dagger P)$, where we couple the prepared system to the environment E (ancillary system ρ_0), perform a general unitary evolution U then a projective measurement P (that has some chance of failure) then trace out the environment.

2.2.3 Axiom 3: Measurement

Measurements are now a special case of axiom 2 where every measurement M is associated with a positive operator valued measure (POVM) $\{M_k\}$ such that $\sum_k M_k = I$. This is a CP map where the Kraus operators are the $\{M_k\}$.

The probability of a measurement M yielding outcome k, given a preparation P (corresponding to ρ) and transformation T (corresponding to $\mathcal{E}$), is: $p(k|P,T,M) = Tr(M_k\mathcal{E}(\rho))$.

This set of axioms aims to get rid of any mention of underlying physical states or their evolution and aspires to be as minimal as possible. The axioms of quantum theory formulated in this way are very general and have numerous applications. They provide a clear target which alternative interpretations of quantum theory, notably realist approaches, must reproduce.

A realist approach to quantum theory must aim to go further than just give an account of all the results of experiments performed. Such an interpretation must also provide an accurate, verifiable description of the underlying physical mechanisms leading to the results. We will describe how such an attempt at a realist approach reveals several unexpected features of the world.

Chapter 3

Non locality

3.1 EPR

In their 1935 paper [1], Einstein, Podolsky and Rosen raise a fundamental issue regarding quantum theory. The authors define elements of physical reality in the following way: "If, without in any way disturbing a system we can predict with certainty the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity". They also make the point that a physical theory should not just be correct but should also be complete, in the sense that: "every element in the physical reality must have a counterpart in the physical theory". EPR then make use of a quantum state $|\psi\rangle$ of two particles which have been prepared such that their relative distance $x_1 - x_2$ is arbitrarily close to L and their total momentum $p_1 + p_2$ is arbitrarily close to zero.

A measurement of x_1 then allows one to predict with certainty the value of x_2 without disturbing particle 2. Indeed, the authors assume a notion of locality along the following lines: "since at the time of measurement the two systems no longer interact, no real change can take place in the second system in consequence of anything that may be done to the first system". This means that x_2 corresponds to an element of physical reality as EPR defined.

In the same way, one can perform a measurement of p_1 instead of x_1 and determine p_2 with certainty without disturbing particle 2 in any way. This means that x_2 and p_2, which don't commute and therefore cannot be simultaneously assigned precise values by quantum mechanics, both correspond to elements of physical reality. This leads EPR to conclude that quantum mechanics, which cannot describe every element of physical reality, is not a complete theory (based on local causality). The question of whether there exists such a complete theory is left open.

3.2 Bohr response

Not long after the publication of the EPR paper, Bohr published a response [2] explaining his point of view regarding the EPR result. Bohr analyses the actual approach one takes when performing a quantum experiment. He describes the way in which an observer can use his free will to arbitrarily choose his experiments. He explains that "we are not dealing with an incomplete description characterized by the arbitrary picking out of different elements of physical reality at the cost of sacrificing other such elements, but with a rational discrimination between essentially different experimental arrangements and procedures".

In this way, Bohr safeguards quantum theory by resorting to an operational description of an experiment in which the entire phenomenon is regarded as a single and unanalyzable whole. The impossibility of controlling the reaction of the object due to the measuring device and the indivisibility of the quantum of action leads Bohr to question the classical idea of causality and criticize the EPR criterion of reality as ambiguous.

According to Bohr, the non-local nature of quantum theory means that the requirement of not disturbing the system in any way in order to define an element of physical reality is flawed. Indeed, he tells us that: "Of course there is [...] no question of a mechanical disturbance of the system under investigation during the last critical stage of the measuring procedure. But even at this stage there is essentially the question of an influence on the very conditions which define the possible types of predictions regarding the future behavior of the system".

Schrodinger [3] coined the term 'entanglement' to describe this peculiar connection between quantum systems. Indeed, the parts of a quantum system such as the EPR state cannot be separated into valid quantum states for localized subsystems ($|\psi\rangle \neq |\alpha\rangle \otimes |\beta\rangle$ for any states $|\alpha\rangle$ and $|\beta\rangle$). This leads Schrodinger to study quantum steering, this influence of the measuring procedure of one subsystem on the other subsystem, as described by Bohr.

In this way, Bohr introduced the principle of complementarity, namely that: "evidence obtained under different experimental conditions cannot be comprehended within a single picture, but must be regarded as complementary in the sense that only the totality of the phenomena exhaust the possible information about the objects". One could then interpret that all physical concepts correspond to phenomena and reality is described by the whole set of phenomena.

3.3 Hidden variables and Von Neumann's no go theorem

Bohr did not aim to construct an ontological interpretation of quantum theory nor did he decisively question Einstein's assertion [4] that: "On one supposition we should, in my opinion, absolutely hold fast: the real factual situation of the system S2 is independent of what is done with the system S1 which is spatially separated from the former". The question of whether the statistical, non deterministic element of quantum mechanics arises because quantum states are averages over better defined 'dispersion free' states, specified by 'hidden variables' as well as the quantum state, was left open.

Von Neumann gave an early analysis [5] of whether hidden variable theories can reproduce the statistics of quantum mechanics. He proves that, under certain assumptions, quantum mechanics cannot be reproduced by averaging over dispersion free states. One of Von Neumann's assumptions is that the linear combination of two (Hermitian operator) observables is an observable and that the linear combination of expectation values is the expectation value of the combination, for both the quantum mechanical states and dispersion-free states. He then shows that there must be an observable such that $< A >^2 \neq < A^2 >$ so that the dispersion for the measurement of at least one observable (for any state) must be greater than zero.

Bell showed that Von Neumann's assumption, that the linear combination of expectation values is the expectation value of the combination, is not valid for dispersion free states. This assumption breaks down since for two non commuting operators A and B, distinct experimental setups are required to measure A, B and A+B. Bell falsified this conjecture by explicitly constructing a deterministic model [6], generating results identical on average to those of quantum theory, which does not obey this assumption.

The model concerns a spin half particle and measurement of two operators $A = m \cdot \sigma$ and $B = n \cdot \sigma$, where m and n are arbitrary real three-vectors and σ has matrix components which are the Pauli matrices:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.1}$$

Quantum mechanical measurements of A and B always yield $\pm|m|$ and $\pm|n|$ respectively. The hidden variable model consists of the quantum state $|\psi\rangle$ and also a hidden variable λ which takes values between -1 and 1. For a given λ, the result of a measurement of A is deterministically:

$-|m|$ if $-1 < \lambda < -\langle\psi|A|\psi\rangle/m$, which occurs with probability $\frac{(1-\langle\psi|A|\psi\rangle/m)}{2}$

$+|m|$ if $-\langle\psi|A|\psi\rangle/m < \lambda < 1$, which occurs with probability $\frac{(1+\langle\psi|A|\psi\rangle/m)}{2}$.

The average result is then:

$< A >= \langle\psi|A|\psi\rangle = \frac{(1+\langle\psi|A|\psi\rangle/m)}{2} - \frac{(1-\langle\psi|A|\psi\rangle/m)}{2}$,

which perfectly agrees with the quantum mechanical prediction (experiments yield a uniform distribution of λ between -1 and 1). Measurement of B yields values $\pm|n|$ in the same way as measurements of A and also reproduce quantum predictions. Measurements of $A+B = (m+n)\cdot\sigma$, always gives results $\pm|m+n|$, therefore, for this hidden variable model, $< A+B >=< A > + < B >$ does not hold.

Bell's model does not in general have additive expectation values for operators and gives precise predictions for the results of all measurements whilst exactly reproducing quantum mechanical predictions if we average over the hidden variable λ. This deterministic hidden variable model exhibits a non-local character is the sense that: "an explicit causal mechanism exists whereby the disposition of one piece of apparatus affects the results obtained with a distant piece". This led Bell to explicitly ask the question of whether it is possible to construct a local hidden variable model which reproduces the predictions of quantum theory.

3.4 Bell's theorem and the CHSH inequality

Bell derived a quantitative criterion for the existence of a realistic interpretation of any local theory [7]. Consider as an example a system of two spin half particles (note that we could reformulate this in terms of boxes with switches and lights flashing such that the inequality obtained is purely about operational correlations). Suppose that both particles (if there is such a thing as particles) go towards two measuring devices which measure spin along directions a and b. The results $A(a,\lambda)$ and $B(b,\lambda)$ of the two measurements are always ± 1 and can depend on the hidden variable λ along with the setting of the corresponding measuring device a or b. Einstein locality, as we saw before, requires that A is completely independent of the measurement setting b and B of a.

The question is then whether the mean value of the product AB averaged over the hidden variable λ:

$P(a,b) = \int d\lambda\rho(\lambda)\bar{A}(a,\lambda)\bar{B}(b,\lambda)$

can reproduce the quantum statistics if we average also over instrument variables. We then have: $|\bar{A}| \leq 1$ and $|\bar{B}| \leq 1$ and count A and B as zero whenever detectors fail. If c and d are alternative instrument settings for measuring the first and second particle respectively then:

$P(a,b) - P(a,d) = \int d\lambda\rho(\lambda)[\bar{A}(a,\lambda)\bar{B}(b,\lambda) - \bar{A}(a,\lambda)\bar{B}(d,\lambda)]$

$= \int d\lambda\rho(\lambda)\bar{A}(a,\lambda)\bar{B}(b,\lambda)[1 \pm \bar{A}(c,\lambda)\bar{B}(d,\lambda)] - \int d\lambda\rho(\lambda)\bar{A}(a,\lambda)\bar{B}(c,\lambda)[1 \pm \bar{A}(c,\lambda)\bar{B}(b,\lambda)]$.

Therefore, we get:

$|P(a,b) - P(a,c)| \leq \int d\lambda\rho(\lambda)[1 \pm \bar{A}(c,\lambda)\bar{B}(d,\lambda)] + \int d\lambda\rho(\lambda)[1 \pm \bar{A}(c,\lambda)\bar{B}(b,\lambda)]$.

This then yields an inequality that cannot be violated by a local hidden variable theory first derived by Clauser, Holt, Shimony and Horne [8] (CHSH inequality):

$|C| = |P(a,b) - P(a,d)| + |P(c,d) + P(c,b)| \leq 2$.

The original form of the result, given in Bell's original paper [7] can be derived using c=d and $P(d,d) = -1$ such that the CHSH inequality becomes:

$|P(a,b) - P(a,d)| \leq 1 + P(d,b)$.

This inequality can be violated using quantum mechanics. Let the joint state of the system be the singlet state for spin half: $|\psi\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$, where $|0\rangle = (1,0)^\dagger$ and $|1\rangle = (0,1)^\dagger$ might, for example, correspond to the vertical and horizontal polarization of a photon. Let the apparatus for the first particle measure either $A = \sigma_z$ or $C = \sigma_x$, corresponding to settings a and c respectively. Similarly, let the apparatus for the second particle measure either $B = \frac{-\sigma_z - \sigma_x}{\sqrt{2}}$ or $C = \frac{\sigma_z - \sigma_x}{\sqrt{2}}$, corresponding to settings b and d respectively. In this way, we get that the averages are: $P(a,b) = P(c,b) = P(c,d) = \frac{1}{\sqrt{2}}$ and $P(a,d) = -\frac{1}{\sqrt{2}}$. This means that quantum mechanics allows us to attain $C = 2\sqrt{2}$.

Aspect performed an elaborate experiment [9] verifying this violation of the CHSH inequality using pairs of photons. Several loopholes [10] also have to be verified (in a single experiment) to make sure that the CHSH inequality is indeed violated in nature. The two measurement apparatus must be spacelike separated so that there cannot be any communication of results and update. If the detection efficiency is low [11], we must also assume that the data collected is a fair sample. Another loophole which could allow for local hidden variables is free will. If hidden variables guide which settings the measurement apparatus will use and when measurements will be performed, then the CHSH inequality may be violated. If one believes in superdeterminism then the CHSH inequality does not say much, since there can then be local hidden variables which dictate everything that will ever happen (at least if you believe everything was once in the same light cone).

3.5 Cirelson bound

Cirelson asked whether quantum theory enforces an upper limit on non-local correlations [12], corresponding to a maximal violation of the CHSH inequality. Consider four operators A, B, C and D satisfying $A^2 = B^2 = C^2 = D^2 = I$ and: $[A,B] = [B,C] = [C,D] = [D,A] = 0$.

Consider the CHSH correlation operator: C=AB+BC+CD-DA such that: $C^2 = 4 + [A,C][B,D]$. We know that for any two bounded operators S and T, we have:

$||S+T|| \leq ||S|| + ||T||$ and $||ST|| \leq ||S||||T||$

and so: $||[A,C]|| \leq 2||A||||C|| \leq 2$ and $||[B,D]|| \leq 2||B||||D|| \leq 2$.

Therefore, $||C^2|| \leq 8$ and $||C|| \leq 2\sqrt{2}$.

This is the Cirelson inequality. This shows that quantum theory cannot violate the CHSH inequality any more than the violation already achieved in the Aspect experiment. A natural question to ask next is whether it is physically possible to achieve the maximal violation of the CHSH inequality.

3.6 Popescu Rohrlich boxes

In a 1994 paper, Popescu and Rohrlich asked the question of whether non-locality can be used as an axiom for quantum theory [13]. They then proceed to note that relativistic causality, or the principle of non-signaling between space-like separated observers, does not restrict the violation of the CHSH inequality to $|C| \leq 2\sqrt{2}$ but allows for maximal violations of $|C| = 4$.

The non-local device which allows for such a maximal violation, which was previously introduced by the same authors [14], is called a PR box.

This is an operational device which has input settings $x = \{0,1\}$ and $y = \{0,1\}$ and outputs $X = \{0,1\}$ and $Y = \{0,1\}$. The PR box can be defined as satisfying:

(i)$\sum_{y=0,1} P(X,Y|x,y) = p(X|x)$ and $\sum_{x=0,1} P(X,Y|x,y) = p(Y|y)$, which correspond to the no-signaling condition.

(ii)$p(X|x) = p(Y|y) = \frac{1}{2}$ so that the marginals are completely random distributions.

(iii) The PR box acts on both inputs as: $X + Y = xy$ to give the outputs.

If we have access to a PR box, then we can get averages $P(a,b) = P(c,b) = P(c,d) = 1$ and $P(a,d) = -1$, where we take inputs $x = \{0,1\}$ to correspond to a or c, and inputs $y = \{0,1\}$ to correspond to b or d. Therefore, the PR box allows us to reach maximum violations of the CHSH inequality: $|C| = |P(a,b) - P(a,d)| + |P(c,d) + P(c,b)| = 4$.

Aharanov had conjectured (in his unpublished lecture notes) that relativistic causality together with non-locality could be used to derive quantum theory. The authors showed that this is not enough to define quantum mechanics

It then makes sense to ask why this violation is not attained by quantum theory and whether we expect nature to satisfy Cirelson's bound. It has been shown that the correlations of the singlet can be simulated by supplementing hidden variables with a single use of the PR-box [15].

Simulation of entangled states would be a bit too easy and communication complexity would become trivial if PR boxes existed in nature [16]. Indeed, maximally strong no-signaling correlations would allow one observer to have access to any m bit subset of the whole data set by just accessing one bit of that data set. If nature behaved in this way, it would violate the principle of information causality (see [17] for more details). Such extra features we expect the world to satisfy are valuable potential physical axioms for quantum theory, or even theories going beyond quantum mechanics.

3.7 Generalized CHSH inequality

We will not prove it here but for any bipartite entangled state, it is possible to find pairs of observables whose correlations violate the CHSH inequality [18].

The CHSH inequality can also be easily generalized [19] by allowing more measurement settings for each of the two observers to whom we send half of a spin half singlet state. Let the first and second observers measure the spin component along one of: $a_1, a_3, ..., a_{2n-1}$ and $b_2, b_4, ..., b_{2n}$ respectively. The results of the measurements are A_r and B_s and have values ± 1.

Averaging over many particle pairs gives a generalized CHSH inequality:

$| < A_1 B_2 > + < B_2 A_3 > + ... + < A_{2n-1} B_{2n} > - < B_{2n} A_1 > | \leq 2n - 2$

In quantum theory, letting the 2n observation directions $a_1, b_2, a_3, ..., a_{2n-1}, b_{2n}$ be chosen such that there is an angle $\frac{\pi}{2n}$ between them, then the left hand side of the inequality can be made arbitrarily close to 2n as $n \to \infty$.

It is possible to generalize the CHSH inequality in a number of ways [20] [21], with more observers, more measuring settings, more measurement results, etc. Some of these generalized Bell-type inequalities may be undiscovered and have novel features and applications (although unnecessary abstraction will most likely lead to complication for no benefit). In the next section, we will describe a generalization to three observers which is particularly elegant and interesting.

3.8 Mermin non-locality

Based on an argument of Greenberger, Horne and Zeilinger, Mermin described a new test of non-locality [22] which doesn't depend on an inequality based upon the statistics of the data accumulated in many runs but depends on the outcome of a single run.

Let a source emit a trio of particles which goes to three far-away detectors. These detectors have two switch settings 1 and 2 and emit either a red or green light (like in the operational description of the CHSH experiment). We observe that if one detector is set to 1 and the others to 2 then an odd number of red lights always flash, and if all three detectors are set to 1 then an odd number of red lights never flash.

Einstein locality would then lead us to conclude that all the information on which colour the detector will flash given settings 1 or 2 must be carried by the particle (this information may be encoded in hidden variables). The colour flashing cannot depend on the setting of the other two switches. We denote the information carried by all three particles, which determines the sets of colours flashing at each detector depending on the setting, as:

(detector1 setting1,detector2 setting1,detector3 setting1;detector1 setting2,detector2 setting2,detector3 setting2).

We can then enumerate all the allowed sets of flashing colours which correspond to an odd number of red lights flashing if one detector is set to 1 and the others to 2:

(R,R,R;R,R,R), (R,G,G;R,G,G), (G,R,G;G,R,G), (G,G,R;G,G,R),
(R,G,G;G,R,R), (R,R,R;G,G,G), (G,G,R;R,R,G) and (G,R,G;R,G,R).

However, every one of these sets of instructions results in an odd number of red flashes if all three switches are set to 1. In this way, a single run of 111, where an even number of red lights flash, is enough to show that local realism does not hold here.

This can be achieved using quantum mechanics. Indeed, let one prepare a three particles GHZ state: $|GHZ\rangle = \frac{1}{\sqrt{2}}(|000\rangle - |111\rangle)$, where $|0\rangle$ and $|1\rangle$ are spin up and spin down states along the z axis. Let us then measure σ_x or σ_y on each particle depending on whether the switch is respectively on setting 1 or 2. But we know that $\sigma_x \otimes \sigma_y \otimes \sigma_y$, $\sigma_y \otimes \sigma_x \otimes \sigma_y$ and $\sigma_y \otimes \sigma_y \otimes \sigma_x$ all commute and have eigenstate $|GHZ\rangle$ with eigenvalue one. Therefore, if we set outcomes $+1$ and -1 of the measurements as Red and Green flashes then there is always an odd number of red flashes if one detector is set to 1 and the others to 2.

What about the case when all three detectors are set to one?

In that case, we measure:

$\sigma_x \otimes \sigma_x \otimes \sigma_x = -(\sigma_x \otimes \sigma_y \otimes \sigma_y)(\sigma_y \otimes \sigma_x \otimes \sigma_y)(\sigma_y \otimes \sigma_y \otimes \sigma_x)$,

which has eigenstate $|GHZ\rangle$ with eigenvalue -1. This means that there must always be an even number of red flashes when all three detectors are set to 1. Therefore, quantum theory can be shown to violate local causality in a single run.

There is an implicit assumption we made at first, linked to Einstein locality, which is that one can associate values for the outcomes of measurements regardless of what occurs in spacelike separated regions. The measurement of σ_x for the first observer and the assignment of a value to its result requires mutually exclusive experiments if the other observers both measure σ_x or both measure σ_y. One must be careful with counterfactual assumptions concerning independence of the context in which a measurement is performed. We will now proceed to study this new notion of contextuality.

Chapter 4

Contextuality

4.1 The over-protective seer

In order to illustrate his early thoughts on the limitations of non-contextuality, Specker introduced a mathematical parable [23]. The story is that of an overprotective seer who does not wish for his daughter to marry any of her suitors. As a challenge they must overcome to earn his daughter's hand, he faced each suitor with the following task. They were given three boxes, which each may or may not contain a gem, and told to pick out either two as empty or two as full. After each suitor had made his prediction, he was ordered by the father to open any two boxes which he had predicted to be both empty or any two boxes which he had predicted to be both full. It always turned out, however, that one of these boxes was empty and the other was full. Eventually, the daughter cheated and married the suitor she fancied most (they divorced three years later, but that is another parable).

It is impossible to come up with a configuration of empty and full 'properties' to boxes such that opening any two of them reveals one full box and one empty one. The correlations described in the parable are a simple example of contextuality. Indeed, if one wishes to explain the measurements (opening a box) as revealing a pre-existing property, then one must imagine that the outcome of a measurement depends on the context of the measurement. Whether a gem is seen or not in the first box depends on whether that box was opened together with the second or together with the third. In this way, the suitors can never achieve their goal since they are asked to assign the outcomes of measurements in a non-contextual way for a system whose statistics are contextual. In fact, such a correlation is also impossible using quantum theory since in quantum theory one can implement a set of Hermitian measurement operators jointly if and only if one can implement every pair of this set jointly (when they commute).

4.2 Gleason's theorem

Gleason [24] was interested in reformulating quantum theory using a weaker set of axioms than Von Neumann's [5]. In doing so, he decided to tackle Mackay's problem of determining all measures on the closed subspaces of a Hilbert space. A measure μ on the closed subspaces is function which associates to each closed subspace a non-negative real number such that for any countable collection of mutually orthogonal subspaces A_i having closed linear span B, we get: $\mu(B) = \sum_i \mu(A_i)$.

His main result, known as Gleason's theorem, is that for a Hilbert space of dimension 3 or greater, the only possible measure of the probability of the state associated with a particular linear subspace 'a' of the Hilbert space will have the form $Tr(P(a)\rho)$, the trace of the operator product of the projection operator P(a) and the density matrix ρ for the system. This shows that if one uses Hilbert space then it is very hard to get rid of the Born rule for measurement.

In his attempt at axiomatization, Gleason treats quantum events, notably measurement outcomes, as logical propositions (yes-no questions called elementary tests), and studies the relationships and structures formed by these events. His fundamental axioms are then:

(i) Elementary tests are represented by projectors P(u) on Hilbert space vectors u.

(ii) Compatible elementary tests, which can be answered together, correspond to commuting projectors.

(iii) If P(u) and P(v) are orthogonal projector, then their sum P(uv)=P(u)+P(v), which is also a projection operator, has expectation value: $< P(uv) >=< P(u) > + < P(v) >$.

The proof of Gleason's theorem is not directly relevant to contextuality so we will only briefly mention some details. Gleason defines a frame function of weight W as a real valued function f defined on the surface of a Hilbert space H such that if $\{e_i\}$ is an orthonormal basis of H then: $\sum_i f(e_i) = W$. A frame function f is regular iff there exists a Hermitian operator T on H such that: f(x)=(Tx,x) for all unit vectors x. By finding these frame functions (using properties of spherical harmonics), Gleason shows that every non-negative frame function in three or more dimensions is regular. Gleason's theorem then follows (relatively) easily.

Although it is not directly addressed to hidden variables, Gleason's work was an important source of inspiration for the no-go theorems of Bell and Kochen-Specker.

4.3 Bell corollary of Gleason's theorem

In a paper written before his famous non-locality article, Bell derived an important corollary [6] of Gleason's work in the form of a no-go theorem against non-contextual hidden variable theories.

To do this, Bell reformulates (relevant consequences of) the Gleason axioms (i), (ii) and (iii) as:

(A) If with some vector u, $< P(u) >= 1$ for a given state, then for that state $< P(v) >= 0$ for any vector v orthogonal to u.

(B) If for a given state $< P(u) >=< P(v) >= 0$ for some pair of orthogonal vectors, then $< P(\alpha u + \beta v) >= 0$ for all α and β.

Now, let u be a normalized vector such that, for a given state, $< P(u) >= 1$ and let v be a vector such that $< P(v) >= 0$. We can write $v = u + \epsilon u'$, where u' is normalized and orthogonal to u, and $\epsilon \in \mathbb{R}$.

Let the vector space be at least three dimensional and let u" be a normalized vector orthogonal to both u and u' so that (A) gives: $< P(u') >=< P(u'') >= 0$.

Therefore (B) gives: $< P(v + \frac{\epsilon u''}{\gamma}) >=< P(-\epsilon u' + \gamma\epsilon u'') >= 0$,where $\gamma \in \mathbb{R}$.

So (B) gives: $< P(u + u''\epsilon(\gamma + \frac{1}{\gamma})) >= 0$.

But if $\epsilon \leq \frac{1}{2}$ then there exist real γ st: $\epsilon(\gamma + \frac{1}{\gamma}) = \pm 1$. This then implies, using (B) again, that:

$< P(u) >=< P(u + u'') > + < P(u - u'') >= 0$, which is a contradiction. Therefore, we have $\epsilon > \frac{1}{2}$.

This implies that $|v - u| > \frac{1}{2}|u|$ and so u and v cannot be arbitrarily close if $< P(u) > \neq < P(v) >$. But, if we consider dispersion free states (which can include hidden variables), then for each one of these states each projector must have a value 0 or 1 associated with it. But both values must occur (for at least one projector) and there must at times be arbitrarily close pairs of projection directions u and v which give different expectation values. Therefore, if we accept assumptions (A) and (B) then there cannot be dispersion free states.

If we wish to construct a realist interpretation of quantum theory using hidden variables, we can reject assumption (B). Indeed, operator $P(\alpha u+\beta v)$ commutes with P(u) and P(v) only if either $\alpha = 0$ or $\beta = 0$. This means that a measurement of $P(\alpha u + \beta v)$ generally requires a distinct experimental arrangement, meaning that (B) relates results of incompatible experiments which cannot be performed simultaneously. This criticism is similar to that Bohr made of Einstein's criterion of reality when he introduced the notion of complementarity [2].

Bell explains nicely that the danger lies in the implicit assumption that hidden variable models must be non-contextual: "It was tacitly assumed that measurement of an observable must yield the same value independently of what other measurements may be made simultaneously".

Kochen and Specker devised an algebraic proof (not involving a continuum) that any ontological description of quantum theory must not just account for non-locality but must be contextual. We will look at this next.

4.4 Kochen Specker theorem

The Kochen Specker theorem [25] asserts that any ontological deterministic theory that would attribute definite results to each quantum measurement and still reproduce the statistical properties of quantum theory must be contextual. This means that for three operators A,B and C such that [A,B]=[A,C]=0 and $[B, C] \neq 0$, the result of measuring A depends on whether A is measured alone, together with B or together with C (it depends on the context of the measurement).

A more precise statement of the Kochen-Specker theorem is that in a Hilbert space of dimension superior or equal to 3, it is impossible to associate definite numerical values $v(P_m)$ (equal to 0 or 1), with every projection operator P_m, such that if a commuting set $\{P_m\}$ satisfies $\sum_m P_m = I$, then $\sum_m v(P_m) = 1$.

The theorem can be proven by taking a (well chosen) complete set of orthonormal vectors $v_1, ..., v_N$ such that the N matrices $P_m = v_m v_m^\dagger$ are projectors in directions v_m. These projectors commute and satisfy $\sum_m P_m = I$. In order to satisfy $\sum_m v(P_m) = 1$, one must associate 1 with one of the u_m and zero with all the N-1 others (there are N ways to do this). Considering several distinct orthogonal bases which share some vectors leads us to conclude that it is not always possible to associate the value 1 or 0 to a vector which is part of more than one basis, irrespective of the choice of other basis vectors.

Kochen and Specker's original proof [25] used a set of 117 vectors in real three dimensional space but a number of proofs involve many less vectors (Conway and Kochen wound one using 31 vectors [19]). Peres came up with two particularly elegant proofs [26] using 33 rays in $\mathbb{R}^3$ and 20 rays in $\mathbb{R}^4$. In higher dimensions, the theorem can usually be proven using less vectors [27] (particularly if we restrict the analysis to a known state [28]). We will not prove the Kochen Specker theorem in detail here.

Similarly to the Bell theorem, the Kochen Specker theorem does not just apply to quantum theory. It is a geometrical statement which affects the interpretation of quantum measurements. This result has the advantage that, unlike the non locality no-go theorem, it does not involve statistical correlation over large ensembles but compares results that can be found on a single system.

In fact, the Kochen Specker result can even be recast in logical terms as a result about partial Boolean algebras within a category-theoretic framework [29]. A recent analysis which also seems worthy of brief mention here is due to Abramsky and Hardy [30]. They also use a logic approach to consider contextuality and non-locality and introduce logical Bell inequalities.

4.5 Mermin magic square

We will now conclude this chapter on contextuality by presenting an elegant result by Mermin [31].

The following square of 9 observables has the property that each row and column is a set of commuting observables that multiply to give I, except the last row which gives -I:

$$\begin{array}{ccc} I \otimes \sigma_z & \sigma_z \otimes I & \sigma_z \otimes \sigma_z \\ \sigma_x \otimes I & I \otimes \sigma_x & \sigma_x \otimes \sigma_x \\ \sigma_x \otimes \sigma_z & \sigma_z \otimes \sigma_x & \sigma_y \otimes \sigma_y \end{array} \tag{4.1}$$

An attempt to associate predetermined values ± 1, independently of the context in which the observable may be measured, leads to a contradiction. We expect the product of all the values corresponding to the 9 operators taken twice to be $+1$, since each value is ± 1. To agree with quantum predictions, however, the product of all the operators taken twice should be -1 (each row and column of the square must multiply to one except the last row, which is -1). This contradiction leads us to conclude that observables do not have pre-determined noncontextual values in quantum mechanics.

Note that we could use a similar proof to reveal the contextuality exhibited in the Mermin non-locality argument we saw above [22] (using a five-pointed star instead of a square).

Contextuality is an important topic in the foundations of quantum theory and will be a recurring theme in the following.

Chapter 5

Ontological models of quantum mechanics

Thus far, we have seen how a naive attempt at interpreting quantum theory as a realist theory of the world runs into trouble. Indeed, if one believes that quantum theory can be interpreted as a statistical theory arising as an average over an underlying ontological theory, then we have seen that such a theory must satisfy certain constraints. Indeed, such a realist attempt reveals that the world exhibits surprising features: non locality and contextuality.

5.1 Defining ontological models

If one wishes to make this quest for a realist interpretation of quantum theory more formal then one can introduce ontological models [32] [33]. These are realist models which reproduce the predictions of quantum mechanics and have the following features:

(i) All the physical properties of a system are determined by the ontic state λ, which is an element of the ontic space Λ.

(ii) The quantum state (preparation P) is an incomplete description of the underlying reality, which corresponds to some (probability-like) distribution over Λ: $|\psi\rangle \in H^{(d)} \leftrightarrow (\mu_{P,|\psi\rangle}(\lambda))$. This explains the probabilistic nature of quantum mechanics (and allows some people to sleep at night).

(iii) Measurements (M) correspond to splittings of the ontic state into distributions $\{\xi_{M,k}(\lambda)\}$ over Λ such that:

$0 \leq \xi_{M,k}(\lambda) \leq 1$ and $\sum_k \xi_{M,k}(\lambda) = 1$, for all λ.

For deterministic ontological models, these are characteristic functions which are just equal to 1 (or 0) for values of λ which do (or don't) give the corresponding outcome.

(iv) The probability of getting outcome k for a measurement M given preparation P is then given by 'averaging ' over the whole ontic space:

$p(k|P, M) =< \xi_{M,k}(\lambda)\mu_{P,|\psi\rangle}(\lambda) >_\Lambda := \int d\lambda \xi_{M,k}(\lambda)\mu_{P,|\psi\rangle}(\lambda)$

This allows us to compare the predictions of the ontological model with the operational framework we wish to consider. We can, for example, compare the results in the model with the quantum prediction: $p(k|P, M) = Tr(M_k\rho)$, where M_k is a POVM element for measurement M and ρ is the density matrix corresponding to the preparation P.

(v) We also need to account for a transformation of Λ over 'time', which can even potentially be stochastic (although a unitary-like deterministic evolution would be nicer). Also, measurements can potentially disturb the space Λ and the model must account for this.

A realist would expect it to be possible to reproduce the predictions of any accurate operational theory using such an ontological model (or perhaps a less naive ontological 'super-model' based on an improvement of the definition above).

If we perform the preparation P with setting S_P then the system will be prepared in a particular ontic state $\lambda \in \Lambda$. If one believes that the quantum states are a complete description of reality then they correspond directly to the ontic states themselves and the ontic space is just the projective Hilbert space of the system $\Lambda = H$. We call this a ψ-ontic interpretation of quantum theory.

Alternatively, the quantum state can correspond to a state of knowledge about reality. In such a ψ-epistemic interpretation of quantum theory, the preparation procedure corresponding to the quantum state corresponds to a probability distribution: $\mu(\lambda|S_P)$, satisfying $\int d\lambda \mu(\lambda|S_P) = 1$, which encodes the epistemological uncertainty about the ontic state we prepared. This situation is compatible with the case where the quantum state is an incomplete description of reality which must be supplemented by hidden variables such that: $H \subset \Lambda$.

Another option would be that the quantum state does not play a realistic role at all such that: $H \not\subset \Lambda$. We can call this a ψ-calculational interpretation of quantum theory.

Note that the ontic space Λ need not be restricted to a set and can a priori be any mathematical object. One must be careful not to discard potential realist interpretations of physics because of mathematically naive restrictions. It may be useful to illustrate the ontic space Λ as a simple generalization of the Bloch sphere, or as a real line, where we integrate over λ to reproduce statistical predictions. If are seeking out a mathematical object underlying all physical states of reality, however, we have to be careful not to restrict too stringently our analysis of potential ontic spaces.

We will now describe some of the work done on ontological models.

5.2 Examples of ontological models

As an illustration, we will now look at several examples of simple ontological models [34] [32].

(A) The first of these is the Beltrametti-Bugajski model [35]. This is an ontological model corresponding to the orthodox interpretation of quantum mechanics, with a ψ-ontic interpretation of the quantum state. The ontic space is the projective Hilbert space $\Lambda = H$ so a system prepared in a quantum state $|\psi\rangle$ is associated with a sharp probability distribution: $\mu(\lambda|\psi) = \delta(\lambda - \lambda_\psi)$ over Λ, where λ_ψ is the unique ontic state associated with $|\psi\rangle$.

Measurements correspond to the distributions:

$\xi(k|\lambda, M) = Tr(|\lambda\rangle\langle\lambda|M_k)$,

where $|\lambda\rangle$ is the unique quantum state associated with $\lambda \in \Lambda$ and M_k is the POVM quantum theory associates with measurement M.

This model trivially reproduces the quantum mechanical operational predictions since:

$$pr(k|M, \psi) = \int d\lambda \xi(k|\lambda, M)\mu(\lambda|\psi) = Tr(|\psi\rangle\langle\psi|M_k) \tag{5.1}$$

(B) The next model, which is for two dimensional Hilbert spaces, is due to Kochen and Specker [25]. The ontic states are vectors λ on the unit sphere Λ and the quantum state ψ is associated with the probability distribution:

$\mu(\lambda|\psi) = \frac{1}{\pi}\Theta(\psi \cdot \lambda)\psi \cdot \lambda$,

where Θ is the Heaviside function ($\Theta(x) = 1$ or 0, for $x \geq 0$ or $x < 0$ respectively) and ψ is the vector corresponding to the quantum state. This assigns the value $cos(\theta)$ to all the points in the hemisphere centered on ψ and zero to the points in the other hemisphere.

A measurement associated with a projector onto vector ϕ is associated with the distribution: $\xi(\phi|\lambda) = \Theta(\phi \cdot \lambda)$, such that a positive outcome occurs if the ontic state λ is in the hemisphere centered on ϕ.

This model is deterministic and reproduces two-dimensional pure state quantum theory since:

$p(\phi|\psi) = \int d\lambda \xi(\phi|\lambda)\mu(\lambda|\psi) = \frac{1}{2}(1 + \psi \cdot \phi) = | < \psi|\phi > |^2$.

Note that Bell's hidden variable model [6], which we previously described as a counter-example for Von Neumann's no go theorem, can also be expressed as an ontological model for two dimensional Hilbert space.

(C) A third example of an ontological model is that of a qutrit, or three dimensional quantum system [34]. The ontic state in this case consists of all the rank one projectors in $GL(3, \mathbb{C})$, which the general linear group of degree 3, i.e. the set of all 3x3 invertible complex matrices.

A quantum state $|\psi\rangle$ is then represented by the probability distribution:

$\mu(\lambda|\psi) = N(Tr(\lambda\lambda_\psi) - \Delta)$ if $Tr(\lambda\lambda_\psi) - \Delta \geq 0$

or $\mu(\lambda|\psi) = 0$ otherwise.

Δ is a parameter that can be played with to vary the support of $\mu(\lambda|\psi)$.

Measurements are deterministic and can be described by the characteristic functions:

$\xi_0(\lambda) = \Theta(Tr(\lambda\lambda_0) - Tr(\lambda\lambda_1))\Theta(Tr(\lambda\lambda_0) - Tr(\lambda\lambda_2))$

$\xi_1(\lambda) = \Theta(Tr(\lambda\lambda_1) - Tr(\lambda\lambda_0)\Theta(Tr(\lambda\lambda_1) - Tr(\lambda\lambda_2))$

$\xi_2(\lambda) = \Theta(Tr(\lambda\lambda_2) - Tr(\lambda\lambda_0))\Theta(Tr(\lambda\lambda_2) - Tr(\lambda\lambda_1))$

so that a state λ gives the outcome corresponding to which central element λ_0, λ_1 or λ_2 it is closest to.

Sadly this model does not reproduce the predictions of quantum mechanics but it comes really really close.

We can see that this last model, as expected if we want it to reproduce quantum theory, exhibits a form of contextuality. Indeed, there may exist some ontic states λ (called unfaithful points) which are closer to central element λ_0 then λ_1 or λ_2 but which are closer to other central elements λ'_1 or λ'_2 than to λ_0. This is a form of measurement contextuality for the ontological model, where the outcome of a measurement depends on knowledge of all three measurements which are simultaneously performed.

In model (B), we can see that the Born rule is artificially built into the model. If we wish to gain real insight into how the statistical character of quantum mechanics arises from an underlying deterministic realist theory, however, we would like to come up with a principle which accounts for this. In the next section we will see how many of the interesting features of quantum theory can be derived from a simple ontological model together with an epistemic restriction.

5.3 Spekkens toy theory

In defense of ψ-epistemic interpretations of quantum theory, Spekkens introduced a toy theory [36] which reproduces many features of quantum mechanics. The theory is based on the following knowledge balance principle: "If one has maximal knowledge, then for every system, at every time, the amount of knowledge one possesses about the ontic state of the system at that time must equal the amount that one lacks".

The ontic space in this theory is simply the set $IV := \{1, 2, 3, 4\}$ (ontic states are 1, 2, 3 and 4) for each elementary consistuent and IV^n for a compound system with n elementary consistuents. We define a canonical question set as: "a set of yes-no questions about the ontic state of a system, which has the minimum number of elements such that the answers uniquely identify the ontic state". The measure of knowledge for which the knowledge balance principle can be applied. is then the number of questions in a canonical question set to which we know the answer.

The analogue of the quantum state in our system is then the state of our knowledge about the system, or the epistemic state. For a single system (with ontic space IV), the epistemic states are: $1 \vee 2$, $1 \vee 3$, $1 \vee 4$, $2 \vee 3$, $2 \vee 4$ and $3 \vee 4$. The canonical set being unanswered corresponds to the state of maximum uncertainty: $1 \vee 2 \vee 3 \vee 4$.

Any two states whose ontic bases have an empty intersection are called disjoint (for example:$1 \vee 2$ and $3 \vee 4$). This is the analogue of orthogonal quantum states. We can also easily define formal analogues of quantum fidelity and superpositions if we make the associations:

$1 \vee 2 \leftrightarrow |0\rangle$, $1 \vee 3 \leftrightarrow |+\rangle$, $1 \vee 4 \leftrightarrow |-i\rangle$, $2 \vee 3 \leftrightarrow |+i\rangle$, $2 \vee 4 \leftrightarrow |-\rangle$, $3 \vee 4 \leftrightarrow |1\rangle$ and $1 \vee 2 \vee 3 \vee 4 \leftrightarrow \frac{I}{2}$, where $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle)$.

Transformations on the ontic states $IV \rightarrow IV$ are defined as transformations on the epistemic states which are allowed by the knowledge balance principle. Therefore the allowed transformations are the permutations of the four ontic states ,which correspond to elements of the symmetric group S_4 of 24 such permutations under composition.

Measurement in the toy theory corresponds to asking as many questions from a canonical set as the knowledge balance principle will allow you to answer. For a single system, the allowed measurement questions are:

$1 \vee 2$ or $3 \vee 4$?, $1 \vee 3$ or $2 \vee 4$? and $1 \vee 4$ or $2 \vee 3$?

Note that measurement would be deterministic if the ontic state was known, but the restriction on our knowledge of the state of the system leads to an apparent indeterminism. Also, the knowledge balance principle implies that measurement inevitably induces a disturbance on the ontic state such that the epistemic state of the system corresponds exactly to the answers of the measurement questions asked. So, if we performed the measurement $a \vee b$ or $c \vee d$? (a,b,c,d $\in$ IV) and obtained the outcome corresponding to $a \vee b$, then the epistemic state of the system must be $a \vee b$ after the measurement. This means that, in order to satisfy the knowledge balance principle, the ontic system undergoes one of the following disturbances: either nothing happens or the ontic states a and b swap, but we don't know which one of these occurs. The toy theory also reproduces analogues of non-commutative measurements and quantum interference.

The ontic state of a pair of systems is IVxIV, which corresponds to sixteen ontic states: 11, 12, 13, ..., 44. In this case a canonical question set contains four questions, which means that epistemic states of maximal knowledge correspond to one of four possibilities. Two extra constraints must be added: epistemic states must be defined such that the knowledge balance principle should apply to each constituent subsystem as well as to the overall composite system and applying any allowed operation to a state must yield an epistemic state which satisfied the knowledge balance principle.

This means that there are essentially two basic types of states which are allowed.

The first of these is of the form: $(a \vee b)(c \vee d)$, with $a \neq b$ and $c \neq d$ (for example: $13 \vee 14 \vee 23 \vee 24$), where we have maximal knowledge about the individual systems, but we know nothing about the relationship between them. These are analogous to separable quantum states. The second of these is of the form: $ae \vee bf \vee cg \vee dh$, with $a \neq b \neq c \neq d$ and $e \neq f \neq g \neq h$ (for example: $11 \vee 22 \vee 33 \vee 44$). These are analogous to maximally entangled quantum states. Further states can be introduced which are the analogues to mixed states in quantum theory, like the completely mixed state $\frac{I}{2} := (1 \vee 2 \vee 3 \vee 4)(1 \vee 2 \vee 3 \vee 4)$.

Measurements and transformations can be defined in an analogous way as before (with several complications) and the toy theory can similarly be generalized to more elementary systems. The toy theory also allows for the description of a number of features which seemed specific to quantum mechanics. These include entanglement, remote steering, no cloning, no broadcasting, superdense coding, teleportation and the monogamy of entanglement.

There are a number of quantum phenomena that are not reproduced by Spekken's toy theory. The main quantum properties that are absent from the theory are: the continuum of quantum states, the possible exponential speed up relative to classical computation and most notably non-locality and contextuality. Indeed, the toy theory is by construction a local, noncontextual hidden variable theory. This demonstrates the importance of these concepts as key ingredients of the quantum formalism.

Before moving on to describe contextuality for ontological models, let us briefly mention two generalizations of Spekkens toy theory. Larsson has introduced a contextual extension of Spekkens toy theory with a memory requirement [37]. Recently, Spekkens and Schreiber have been working on generalizing the theory to higher dimensional systems [38]. They have shown that an epistemic restriction based on a discrete version of the uncertainty principle (instead of the knowledge balance principle) allows us to extend the toy model to three dimensions, with a 9-state discrete phase space. It is possible to use this statistical theory to reconstruct a subset of three dimensional quantum mechanics: the stabilizer formalism for trits.

5.4 Contextuality for ontological models

Spekkens has introduced an operational definition of contextuality which applies to ontological models and to arbitrary operational theories [39]. This generalized notion of non-contextuality is defined by Spekkens as: "A non-contextual ontological model of an operational theory is one wherein if two experimental procedures are operationally equivalent, then they have equivalent representations in the ontological model".

This means that we can define three types of noncontextuality, corresponding to the three types of experimental procedures: preparations, transformations and measurements.

Preparation noncontextuality is the feature that the probability distribution $\mu_P(\lambda)$ over ontic states is the same for all preparation procedures in an operational equivalence class. This means that, for any pair of preparation procedures P and P' such that the probability of outcome k (given that measurement procedure M is performed) is the same for all outcomes k (and for all measurement procedures M) that are allowed in the operational model, the distribution associated with the preparations P and P' in the ontological model are the same.

Therefore: $p(k|P,M) = p(k|P',M)$ (for all M and k) $\implies \mu_P(\lambda) = \mu_{P'}(\lambda)$

An example from quantum theory of two preparation procedures in the same equivalence class would be the preparation of a maximally mixed state of a spin half system using two different bases, for example:

$\frac{I}{2} = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) = \frac{1}{2}(|+\rangle\langle +| + |-\rangle\langle -|)$, where $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$.

Similarly,transformation (or measurement) noncontextuality is the features that transformations (or measurements) are represented in exactly the same way in the ontological model, for all transformation (or measurement) procedures in an operational equivalence class.

Measurement noncontextuality can then be defined as the assumption that:

$p(k|P,M) = p(k|P,M')$ (for all P and k) $\implies \xi_{M,k}(\lambda) = \xi_{M',k}(\lambda)$.

An interesting feature of these generalized notions of contextuality is that, unlike the traditional notion of contextuality, it has been shown that for both preparation contextuality and unsharp measurement contextuality (using POVMs), proofs of contextuality can be found for two dimensions (instead of three). It is also possible to retrieve the traditional notion of contextuality along with the corresponding no-go theorems that we studied above from this generalized notion of contextuality [39]. This requires us to assume the perfect discrimination of orthogonal states, which we know is a feature of quantum theory. Therefore, ontological models must be preparation contextual in addition to measurement contextual.

An operational notion of noncontextuality is a very desirable result since it can lead us to methods of experimentally differentiating noncontextual and contextual theories. We can look for noncontextual inequalities [40] [41] [42], similar to Bell inequalities, which give an observable bound on experimental achievements of noncontextual theories. Such inequalities could be the first step towards clarifying potential applications of contextuality (for quantum computing for example) or understanding exactly what role contextuality might play in an axiomatization of quantum theory.

5.5 PBR theorem

So far, we have seen that ontological models must satisfy a number of properties. Indeed, they must exhibit both non-locality and contextuality. In addition, Lucien Hardy has presented an ontological excess baggage theorem [43], showing that the ontic space, even for a qubit, must have infinite cardinality. Montina has also proven that the manifold dimension of the ontic state space is necessarily exponential [44], assuming that the dynamics of the ontic states is Markovian.

Pusey, Barrett and Rudolph, in an attempt to clarify what a quantum state represents, introduced another no-go theorem for ontological models [45]. This theorem has a slightly different flavor to those of Bell and Kochen-Specker. It states that:

"Any model in which a quantum state represents mere information about an underlying physical state of the system must make predictions which contradict those of quantum theory".

This theorem attempts to rule out ψ-epistemic ontological models, where quantum states are epistemic and there is some underlying ontic state so that quantum mechanics is the statistical theory of these ontic states.

The PBR argument rests on the following assumptions: the physical system has a real physical state (independent of the observer) and systems that are prepared independently have independent physical states. Also, the ontic space has to be a measure space and both states and measurements need to be mathematically nice (i.e. probability distributions). Coarse graining over ontic states λ is performed by averaging, using an integration over ontic states. The proof is then the following:

Let the ontic space Λ be a measure space and preparation of each of the quantum states $|\psi_i\rangle$ give an ontic state λ from a probability distribution $\mu_i(\lambda)$ over Λ.

Assume that n systems can be prepared independently in quantum states: $|\psi_{x_1}\rangle, ..., |\psi_{x_n}\rangle$ corresponding to ontic states $\lambda_1, ..., \lambda_n$ sampled from the product distribution: $\mu_{x_1}(\lambda_1)...\mu_{x_n}(\lambda_n)$.

Assume also that the probability $p(k|\lambda_1, ..., \lambda_n)$ for outcome k of a measurement is fixed by the ontic states $\lambda_1, ..., \lambda_n$. Then the operational probabilities are:

$\int ... \int p(k|\lambda_1, ..., \lambda_n)\mu_{x_1}(\lambda_1)...\mu_{x_n}(\lambda_n)d\lambda_1...d\lambda_n$.

To reproduce quantum mechanics, the probability for each measurement outcome should be within some small $\epsilon > 0$ of the predicted quantum probability (using the Born rule). PBR have shown that (even in the presence of noise) if this is the case for a model, then for distinct quantum states $|\psi_0\rangle$ and $|\psi_1\rangle$ corresponding to distributions: $\mu_0(\lambda)$ and $\mu_1(\lambda)$ respectively, we have (see the paper [45] for details): $D(\mu_0(\lambda), \mu_1(\lambda)) = \frac{1}{2}\int |\mu_0(\lambda) - \mu_1(\lambda)|d\lambda \geq 1 - 2\epsilon^{\frac{1}{n}}$ (for some n).

This means that for small ϵ, $D(\mu_0(\lambda), \mu_1(\lambda))$ (which is a measure of distance of two probability distributions) is close to 1 so that an ontic state λ is closely associated with only one of the two quantum states. This shows that for distinct quantum states $|\psi_0\rangle$ and $|\psi_1\rangle$, if the corresponding two distributions: $\mu_0(\lambda)$ and $\mu_1(\lambda)$ overlap then there is a contradiction with the predictions of quantum theory (modulo the assumptions we stated before).

Note that Lewis, Jennings, Barrett and Rudolph recently constructed ψ-epistemic models [46], such that the probability distributions corresponding to distinct quantum states overlap, that recover the Born rule. Their paper does not contradict the PBR result since the models violate one of its assumptions: they do not have the property that product quantum states are associated with independent underlying physical states.

We could alternatively take the approach of quantum Bayesianism [47] and interpret the quantum state as representing information about possible measurement outcomes and not about the objective physical state of the system, which would violate another assumption of the PBR theorem.

We will end here with the description of ontological models and shall now proceed with a description of other explicit attempts to construct an ontological interpretation of quantum theory.

Chapter 6

Explicit attempts to construct an ontological interpretation of quantum theory

Several attempts have been made to actually construct realist theories which account for all the phenomena described by quantum mechanics. A number of these aim to go beyond quantum theory and several attempt a consistent description of quantum gravity. Any such approach should try to get rid of the arbitrary division of the world into observing objects and observed objects which arises in orthodox quantum mechanics. The fundamental role of measurement and necessity of always referring to an outside observer means that the universe as a whole is, as Bell puts it, an embarrassing concept (does the universe require the presence of a universal God-like observer which can observe itself to even exist?).

Here, we will briefly describe two of the most prominent ontological interpretations of quantum mechanics: de-Broglie Bohm theory and the many-worlds interpretation.

6.1 Bohmian mechanics

The ontic state in Bohmian mechanics [48] [49] is the quantum mechanical wavefunction $\psi(r,t)$ together with particle position ξ. This means that de-Broglie Bohm theory for a single particle is a hidden variable model with an ontic space Λ=H x $\mathbb{R}^3$.

The evolution equations for the ontic state are the Schrodinger equation:
$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi + V(r)\psi$,
where $\psi(r,t) = R(r,t)exp(\frac{iS(r,t)}{\hbar})$ (and we choose a spacetime frame [x,t], note that this is not a fundamentally Lorentz invariant theory),

along with the guidance equation:
$\frac{d\xi(t)}{dt} = \frac{1}{m}[\nabla S(r,t)]_{r=\xi(t)}$ (note that this is a first order equation).

The Hamilton-Jacobi equation (real part of the Schrodinger equation):
$\frac{\partial S}{\partial t} + \frac{(\nabla S)^2}{2m} + V + Q = 0$

now has an extra term: $Q = -\frac{\hbar^2}{2m}\frac{\nabla^2 R}{R}$, which we call the quantum potential. An ensemble of particles satisfying this quantum Hamilton-Jacobi equation has the following equation for the conservation of probability (corresponding to the imaginary part of the Schrodinger equation):
$\frac{\partial R^2}{\partial t} + \nabla\cdot(R^2\frac{\nabla S}{m}) = 0.$

The particle therefore has a well defined position $\xi(t)$ which is causally determined and varies continuously in time. The field ψ is a pilot wave which guides the particle position independently of its amplitude and there is no backlash on this wave(it is not affected by ξ). This field provides active information to the particle: very little energy directs a much greater energy. The particle satisfies Newton's equation:

$m\frac{d^2\xi(t)}{dt^2} = -\nabla\cdot[V+Q]$

with the potential V+Q instead of just V (although the work done by Q is not mechanical and is also independent on magnitude of ψ; it is perhaps linked to some form of informational energy).

We can recover the quantum mechanical probability that the particle is within $d\xi$ of ξ at time t, namely: $\rho(\xi,t) = |\psi(\xi,t)|^2$ (taking as an assumption that this relation holds at some point in time). This is true since the probability current is:

$j = \rho v = \frac{1}{m}(\rho\nabla S)$

but by the continuity equation and the imaginary part of the Schrodinger equation, if $\rho = R^2$ at the start, then:

$\frac{\partial\rho}{\partial t} = -\nabla\cdot j = -\nabla\cdot(\frac{1}{m}(\rho\nabla S)) = -\nabla\cdot(\frac{1}{m}(R^2\nabla S)) = \frac{\partial R^2}{\partial t}$.

This explains how interference arises even though particle positions are well defined.

If we split a wavefunction $\psi = \sum_j c_j\psi_j$ into orthogonal (usually spatially disjoint) component waves ψ_j (this is the j^{th} wave), then the j^{th} wave is called occupied if: $\xi \in$ spatial support of ψ_j (and unoccupied otherwise).

If only one of these component waves is occupied, then the guidance equation depends only on that wave. This means that the dynamics depends only on occupied waves, where particles have well defined positions guided by the occupied waves. This allows us to fully calculate the dynamics for any of the usual physical situations studied in quantum theory.

Bohmian mechanics can easily be extended for many particles, where the ontic space is the wavefunction on configuration space $\psi(r_1,...,r_n)$ together with all the particle positions $\xi_1,...,\xi_n$. The evolution equations are then:

$$i\hbar\frac{\partial\psi(r_1,...,r_n,t)}{\partial t} = \sum_i(-\frac{\hbar^2}{2m_i}\nabla_i^2\psi(r_1,...,r_n,t)) + V(r_1,...,r_n)\psi(r_1,...,r_n,t)$$

where: $\psi(r_1,...,r_n,t) = R(r_1,...,r_n,t)exp(\frac{iS(r_1,...,r_n,t)}{\hbar})$.

and: $\frac{d\xi_j(t)}{dt} = \frac{1}{m_j}[\nabla_j S(r_1,...,r_n,t)]_{r_i=\xi_i(t)}$, for j=1, 2, ..., n.

We can then define separable states:

$\psi(r_1,r_2,t) = R_1(r_1,t)exp(\frac{iS_1(r_1,t)}{\hbar})R_2(r_2,t)exp(\frac{iS_2(r_2,t)}{\hbar})$,

which trivially obey:

$S(r_1,r_2,t) = S_1(r_1,t) + S_2(r_2,t)$.

This means that the two separable particles evolve completely independently.

We can also define an entangled state as:

$\psi(r_1,r_2,t) = \sum_j c_j R_j^1(r_1,t)exp(\frac{iS_j^1(r_1,t)}{\hbar})R_j^2(r_2,t)exp(\frac{iS_j^2(r_2,t)}{\hbar})$.

The j^th wave is occupied if:

$(\xi_1,\xi_2) \in supp(R_j^1(r_1,t)exp(\frac{iS_j^1(r_1,t)}{\hbar})R_j^2(r_2,t)exp(\frac{iS_j^2(r_2,t)}{\hbar}))$

and is empty otherwise. If only one wave is occupied then, as before, the two particles evolve completely independently. Otherwise, the particles do not evolve independently, even if they are spacelike separated. In this way, we can see that Bohmian mechanics is a non-local hidden variable theory. Indeed the active information of the pilot wave propagates in a nonlocal way but, as in quantum theory, it is not a mechanical disturbance of the system.

Note also that in Bohmian mechanics, the results of quantum mechanical observations is determined by hidden variables of the combined apparatus and system. As Kochen and Specker noted [25], this means that this is also a contextual hidden model variable, which embodies Bohr's notion of indivisibility of the combined system of observing apparatus and observed object.

Importantly, this theory reproduces the operational predictions of quantum mechanics. Consider a measurement of a Hermitian operator A, with eigenvectors $\phi_k(r)$. This system couples to a measuring apparatus $\chi(r')$ and the environment $\mu(r_1, ..., r_n)$. After a measurement, we have:

$\sum_k c_k \phi_k(r) \chi(r') \mu(r_1, ..., r_n) \to \sum_k c_k \phi_k(r) \chi_k(r') \mu_k(r_1, ..., r_n)$,

where the apparatus and environment come into states depending on the state of the system.

The probability of the j^{th} wave being occupied is then $|c_j|^2 = |c_j \phi_j(\xi) \chi_j(\xi')|^2$. We expect distinct states of the environment to correspond to disjoint regions in configuration space ($\mu_k \mu_j = 0$ for $j \neq k$). So, if the j^{th} wave becomes occupied, we postulate an effective collapse of the guiding wave (like an update of information): $\sum_k c_k \phi_k(r) \to \psi_j(r)$.

In this way, we recover the quantum collapse, since only the j^{th} term is relevant for dynamics after the measurement (decoherence is important here).

Note that Bohmian mechanics can also be extended to relativistic quantum theory and can provide an ontological interpretation for bosonic and fermionic fields. We shall not delve further into the details of this theory but note that they are well described in Bohm and Hiley's book [50]. We will not go through objections of de-Broglie Bohm theory here, but will instead move on to a description of many-worlds theory.

6.2 Many-worlds theory

The many-worlds interpretation is an attempt to maintain the representational completeness of the quantum wavefunction, whilst getting rid of measurements completely so that the only possible evolution is the deterministic unitary one. There are a number of different versions of this theory, but we will mostly focus on the accounts given by Everett [51] and DeWitt [52].

Everett allows the universe as a whole to exist objectively and correspond to a vector in Hilbert space. He attempts to attribute subjective states to observers within the universe, which are in direct correspondence with aspects of the physical universe. These observers posses physical memories in direct correspondence with their past experience, from which deductions can be made about the subjective experience of the observer.

In this relative state formulation, the observer is considered as an automatic machine, whose future actions are determined by the memory together with its present sensory data. Let us illustrate Everett's approach by examining the measurement of spin for a particle in the state: $|\psi\rangle = a|0\rangle + b|1\rangle$. We can see that the measurement acts on the joint state of the system, the measurement apparatus M and the observer O itself as:

$(a|0\rangle + b|1\rangle)|Mready\rangle|Oready\rangle \to a|0\rangle|get0\rangle|observe0\rangle + b|1\rangle|get1\rangle|observe1\rangle$.

In this way, the memory of the observer has been entangled with the system such that the observer does not have a definite memory of the outcome in quantum theory. Therefore, in order to avoid collapse of this wavefunction, Everett assumes that each part of the observer wavefunction corresponds to a definite state of awareness of the content of the observer's memory. In this way, there is a single total awareness where each of the two partial awarenesses are unaware of the other or of the whole. This causes many possible branches to arise along with a sequence of possible partial awarenesses (unaware of each other), where the experience of a particular person is restricted to one branch.

The theory therefore relates the universe as a whole to all the various points of view of the observers contained within it, which each establish a relation between a state of awareness and some part of the universe containing the observed object. This sort of relationship is defined by Everett as the relative state of the system corresponding to a particular state of the awareness of the observer. This means that there are 'reference frames' corresponding to the memories of the various observers and that any part of the total state only makes sense relative to these frames of reference.

One of the problems we are faced with in the relative states approach, is to understand why we interpret the subjective experiences in any given basis rather than any other [53]. This could lead to subjective experiences of the form $\frac{1}{\sqrt{2}}(|observe0\rangle + |observe1\rangle)$ or $\frac{1}{\sqrt{2}}(|observe0\rangle - |observe1\rangle)$, which are not obvious to interpret. This led Kent [54] to make the following criticism: "no preferred basis can arise, from the dynamics or from anything else, unless some basis selection rule is given".

Let us now move to DeWitt's version of the theory, which is closer to the usual account of the many-worlds interpretation. One of his main goals is to introduce a minimal number of concepts into the theory. DeWitt assumes that the whole conceptual basis for quantum theory is provided by Hilbert space and the fact that "the world must be sufficiently complicated that it can be decomposed into systems and apparatuses". He then asserts that the universe is a vector in Hilbert space which is split into an astronomical number of branches, not only due to measurement but also due to many other natural processes. Unlike Everett's relative state (many minds) formulation, this interpretation doesn't just aim to explain our perceptions of the universe, since the universe is itself split into many parts (many worlds). It is not clear when the split is meant to occur and how this precisely depends on complexity.

The key issue for many-worlds theory is then to account for how probability can arise in a deterministic theory, where all possible outcomes occur and the universe is a vector in Hilbert space. The resolution of this issue is not obvious but one option is to use a modified version of many-worlds, described by Deutsch [55], which can deal with probabilities. He assumes that there is a random distribution of an infinite and constant number of universes, with probabilities corresponding to the quantum probabilities. This construction allows us to recover the quantum mechanical probabilities for events (with some caveats [56]).

Let us conclude this section with a quick comparison between many-worlds theory and the de-Broglie Bohm interpretation. First of all, the Bohmian pilot wave also has a multiplicity of realities, and therefore many-worlds is preferred by Occam's razor. In fact the additional structure of particle positions means that unlike Everett's formulation, de-Broglie Bohm's theory does not obey Lorentz covariance. It does not, however, have any issues with probabilities and we can easily interpret macroscopic phenomena in Bohmian mechanics as depending on the configuration of Bohmian particles.

Let us now proceed to an analysis of collapse models.

6.3 Collapse models

Several theories have attempted to resolve the clash between discontinuous statistical behavior of measurement and the linear unitary evolution of closed systems by including the measurement jump as part of dynamics. This has lead to an attempt at forming non-linear extensions of Schrodinger's equation. These would be expected to have a high degree of non linearity when observers are concerned, whilst still being linear in known instances and giving rise to (relativistic) classical dynamics for macroscopic objects.

It would be very interesting to either find such an equation for generalized quantum dynamics or some reason (for example, a no-go theorem for certain types of partial differential equation evolutions based on physical assumptions) why such an equation could not be constructed. In this way, one could hope to have a relatively simple ontology which possibly goes beyond quantum mechanics, where all the complications arise in the evolution equations.

Let us now briefly look at an example of a dynamic collapse model due to Ghirardi, Rimini and Weber [57]. The wave function for N particles is assumed to evolve according to the Schrodinger equation: $i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = H|\psi(t)\rangle$ at most times, but at every time interval $\frac{\tau}{N}$ on average there is a reduction in the spread of the wavefunction (spontaneous collapse):

$|\psi(t+dt)\rangle = \frac{1}{\sqrt{p(q_k)}}\sqrt{E^{(k)}(q_k)}|\psi(t)\rangle$,

where $E^{(k)}(q_k) = \int dr_k K exp(\frac{-(r_k-q_k)^2)}{\sigma^2})|r_k\rangle\langle r_k|$ is a positive operator which has expectation values: $p_k = \langle\psi(t)|E^{(k)}(q_k)|\psi(t)\rangle$ and K is a normalization constant. Also, k is chosen at random and q_k is chosen by sampling from $p(q_k)$. This introduces two new universal constants, which are the mean time between collapses for one particle $\tau \simeq 10^{16}s$, and the localization width of each particle $\sigma \simeq 10^{-7}m$. This process is like a POVM with a continuous outcome space occurring on average every $\frac{\tau}{N}$, which is like a noisy position measurement (this type of evolution is a CP map). This model exhibits non-locality and we can define entangled states of several particles similarly to quantum theory.

The GRW model also reproduces the operational quantum results for measurement without the need for any observer. Indeed, the overall wavefunction, after interaction between the observed system and the apparatus is in the superposition:

$\psi = \sum_n C_n\psi_n(x)\phi_n(y_1,...,y_R,Y)$ where x is the coordinate of the observed system, $y_1,...,y_R$ are the internal coordinates of the apparatus and Y is the macroscopic pointer setting of the apparatus. The spontaneous collapse process of a single particle will affect directly the spread of the pointer coordinate Y and will leave the single result $\phi_m(y_1,...,y_R,Y)$ with a well defined pointer reading (collapses occur very rapidly).

A consideration of an ensemble of such experiments will leave a randomly distributed selection of results where the probability of the m^{th} result is $|C_m|^2$, in agreement with quantum mechanics. With the choice of τ and σ given, this theory is experimentally plausible to date.

Pearle has devised similar models with continuous spontaneous localization instead of the discrete collapses in the GRW model [58].

It would be interesting to present a number of other interpretations of quantum theory at this point, including perhaps the consistent histories approach [59], the transactional interpretation [60] and work on quantum measure theory [61]/ causal sets [62]. One could then proceed to introduce the current formalism of quantum gravity and the various problems that ensue with combining relativity with the interpretations we presented. Maybe next time...

Chapter 7

Quantum logic and quantum Bayesianism

Another way in which one can interpret quantum theory is as a fundamental modification of classical logic or classical probability theory.

7.1 Quantum logic

Three years after the publication of his book [5], where he introduced the current mathematical formalism of quantum mechanics using Hilbert spaces, Von Neumann wrote to Birkhoff [63]:

"I would like to make a confession which may seem immoral: I do not believe absolutely in Hilbert space no more".

The natural framework to apply, in order to devise axioms for a theory, is that of mathematical logic. Indeed, a year later, Birkhoff and Von Neumann published a paper which launched the study of quantum logic.

7.1.1 Birkhoff-Von Neumann quantum logic

In that paper [64], BvN investigate the notion of a physical property and the structure imposed on these properties by the nature of quantum observations. The authors suggest that: "in any physical theory involving a phase space, the experimental propositions concerning a system S correspond to a family of subsets of its phase-space P, in such a way that x implies y (written $x \subset y$) means that the subset of P corresponding to [experimental proposition] x is contained set-theoretically in the subset corresponding to [experimental proposition] y ". Therefore, the first postulate for the propositional calculi of physical systems is that: "the physical qualities attributable to any physical system form a partially ordered system".

The next postulate states that calculi of propositions are a certain type of partially ordered system, called a lattice, which allows one to define the logical properties 'and' ($\cap$) and 'or' ($\cup$) from the relation '$\subset$'. Several further postulates (see the paper [64] for the details) show that the calculus of propositions for quantum theory is just like the usual (classical logic) calculus of propositions with respect to adequately defined 'and', 'or' and 'not' operations, except with the usual distributive law:

$a \cup (b \cap c) = (a \cup b) \cap (a \cup c), a \cap (b \cup c) = (a \cap b) \cup (a \cap c)$

replaced by a weaker orthomodular law:

If $a \subset c$ then $a \cup (b \cap c) = (a \cup b) \cap c$.

This representation of states using a lattice of properties means that quantum superposition can be defined as the strongest property which is true for two distinct states being also true for other states than the two given ones.

One of the main criticisms of this approach is that it fails to elegantly capture the composition of quantum systems. Ideally, it would be more desirable to explain quantum theory in terms of the manner in which quantum systems compose, including superpositions. One might hope that this would allow for notions of entanglement, non-locality and complementarity to arise in quantum logic.

7.1.2 Linear logic

Since the publication of the Birkhoff-Von Neumann paper, there have been numerous important advances in logic which may be very relevant to a successful interpretation of quantum theory as a fundamental modification of classical logic. In this section, we will briefly introduce one of these advances: Girard's linear logic extension of usual logic [65].

Classical logic deals with stable truths since:

if A and $A \Rightarrow B$ then B, but A still holds.

In physics, however, a causal implication leads to a modification of the premises (conditions), meaning that real implications cannot be repeated. We write such a causal implication (called a linear implication) $\multimap$ and define the exponential ! to express iterability of an action so that:

$A \Rightarrow B = (!A) \multimap B$ (A is implied by B exactly when B is caused by some iteration of A). In linear logic, there are two connectives $\otimes$ (times) and & (with) for 'and'. These correspond to the availability of two actions, where either both are done (for $\otimes$) or one action is chosen (for &) such that $!(A\&B)$ is equivalent to: (!A $\otimes$!B).

Linear logic also has a linear negation $(\cdot)^{\perp}$, which is analogous to the transpose in linear algebra (involutive duality) such that:

action of type A=reaction of type $A^{\perp}$

In addition, there are two disjunctions $\oplus$ (plus) and $\bar{\&}$ (par) for 'or' which are the duals of & and $\otimes$ respectively.

Girard gives the example of an application to chemistry, in order to illustrate the power of linear logic in describing the physical state of a system and physical processes (with no reference to time). Classical logic fails, for example, to describe physical state updating in a process like: $2H_2 + O_2 \rightarrow 2H_2O$. In linear logic, however, this can easily be written as:

$H_2 \otimes H_2 \otimes O_2 \multimap H_2O \otimes H_2O$.

Note that the classical logic principles of weakening ($A \wedge B \Rightarrow A$) and contraction ($A \Rightarrow A \wedge A$) no longer hold. Also, classical logic cannot make the distinction between stable facts (like axioms) and the description of the current state (we do not wish to update the axioms when we update the current state).

This means that we should no longer write physical theories of states and processes using only classical logic but should have:

Theory=linear logic+ axioms (written using exponentials)+ current state

Writing down all the syntax and semantics rules of linear logic here would take up far too much space (see Girard's article [65] for a description of what linear logic actually is).

The main point I wish to make is that linear logic has been applied very successfully to computation, proof theory and in other aspects of mathematics but it is, almost by construction, an appropriate logical foundation for theories of physics. Such a foundational logical approach based on processes is at the heart of the categorical approach of the next section.

7.1.3 Categorical quantum logic

We will only very quickly sketch the categorical approach to quantum logic proposed by Coecke [66]. We will return to a more detailed analysis of categorical quantum mechanics towards the end of the report.

The basic idea is that, instead of immediately worrying about physical properties and the structure imposed on these properties as was done by Birkhoff and Von Neumann, we should take as a primitive the composition of quantum systems (and processes) in order to retrieve the quantum structure. This leads us to ask what remains of the quantum formalism if we only look at how systems compose and don't worry about underlying features. The answer is that if we focus only on the structure imposed by the composition of systems, physical processes form mathematical objects called symmetric monoidal categories.

Next, one may ask what kind of structure is needed in addition to composition, in order to deduce experimentally observed phenomena, and notably operational quantum mechanics. Answering this question leads to the development of categorical quantum mechanics, where we can give additional categorical structure to these symmetric monoidal categories by defining abstract equivalents to scalars, superposition, adjoints, etc. This can then be described by a rich diagrammatic calculus with defined by a constructed syntax and semantics (similarly to linear logic). This process allows us to identify the minimal process logic of symmetric monoidal categories as a structure which we hope will arise elsewhere in our 'classical reality'.

We will conclude here this brief glimpse into quantum logic and move on to quantum Bayesianism.

7.2 Quantum Bayesianism

An important avenue of research in quantum foundations has been the attempt to formally describe quantum theory as a generalization of classical probability theory. We will now describe some of the progress made in this direction.

7.2.1 Quantum Bayesian probabilities

In the subjective Bayesian approach to probability theory [67], probability quantifies a degree of belief for a single trial. In this interpretation of probability, different individuals having the same evidence can still have different degrees of belief in a hypothesis. The probability of an event E is then a measure of the rate at which an individual is prepared to bet on E. By the Ramsey-De Finetti theorem, the standard Kolmogorov axioms of probability [68] then follow from certain coherence conditions which make sure that the various degrees of belief fit together.

The subjective Bayesian approach is usually formalized in the following way. In order to measure the degree of belief of Bob in some event E, Alice makes Bob chose a betting quotient q (which is a measure of his degree of belief in E) and choses a stake S (which can be positive or negative but is small compared to Bob's wealth). Bob then pays Alice qS in exchange for S.

If Bob has to bet on many events $E_1, ..., E_n$, then his betting quotients are said to be coherent if and only if Alice cannot chose stakes $S_1, ..., S_n$ such that she wins whatever happens (this is called making a Dutch book against Bob). The Ramsey-De Finetti theorem [67] then states that a set of betting quotients is coherent if and only if they satisfy the Kolmogorov axioms of probability.

Caves, Fuchs and Schack [69] have argued that a ψ-epistemic interpretation of quantum theory (whether or not there is an underlying ontology), where quantum states are states of knowledge, would require that all probabilities derived from a quantum state are subjective probabilities. They point out that the distinction between classical and quantum probabilities lies in the nature of the information they encode, since the maximal information in the quantum world is not complete and cannot be completed.

The authors derive a similar result to Gleason's theorem, showing that any subjective probability assignment must fundamentally obey the Born rule for quantum mechanics if one requires Dutch-book consistency for probability assignments which are faithful to the Hilbert-space structure of elementary tests. This is an elegant result, which indicates that it may be possible to formulate quantum theory as a theory of Bayesian inference [47]. In the next section, we will describe how one might attempt this.

7.2.2 Quantum theory as a causally neutral theory of Bayesian inference

We shall now introduce the work of Leifer and Spekkens [70] on the formalism of quantum conditional states. They aim to write quantum theory as a causally neutral theory of Bayesian inference, which unifies the description of experiments involving two systems at a single time and of a single system at two times. We will first describe the case of acausally related regions (A and B), where neither one has a causal influence on the other.

The quantum state ρ_A, acting on a Hilbert space H_A, is the analogue of a probability distribution P(R) assigned to a random variable R. The density operator ρ_{AB} of the composite region AB, acting on the Hilbert space $H_{AB} = H_A \otimes H_B$, is analogous to the joint distribution P(R,S). The quantum analogue of marginalization $P(R) = \sum_S P(R,S)$ is the partial trace: $\rho_B = Tr_A(\rho_{AB})$. An acausal conditional state for B given A is then a positive operator on H_{AB} satisfying:

$Tr_B(\rho_{B|A}) = I_A$.

This is the quantum analogue of the conditional probability distribution $P(S|R)$, which satisfies: $\sum_S P(S|R) = 1$.

The analogue of the multiplication law in classical probability: $P(R,S) = P(S|R)P(R)$ has to provide a method of constructing a joint state on H_{AB} from a conditional state on H_{AB} and a reduced state on H_A.

This is provided by the following rule:

$\rho_{AB} = \rho_{B|A} \star \rho_A := (\rho_A^{\frac{1}{2}} \otimes I_B)\rho_{B|A}(\rho_A^{\frac{1}{2}} \otimes I_B)$.

Note that classical probability is a trivial special case with:

$\rho_{RS} = \sum_{r,s} P(R=r, S=s)|r\rangle\langle r| \otimes |s\rangle\langle s|$

$\rho_R = \sum_r P(R=r)|r\rangle\langle r| = Tr_S(\rho_{RS})$

$\rho_{S|R} = \sum_{r,s} P(S=s|R=r)|r\rangle\langle r| \otimes |s\rangle\langle s|$.

The next step is to find an analogue of the classical law of total probability $P(S) = \sum_R P(S|R)P(R)$ for quantum belief propagation from region A to region B. This acausal quantum belief propagation is just:

$\rho_B = \mathcal{E}_{B|A}(\rho_A) := Tr_A(\rho_{B|A}\rho_A)$,

since: $\rho_B = Tr_A(\rho_{AB})$, $\rho_{AB} = \rho_{B|A} \star \rho_A$ and the trace is cyclic. The CP map $\mathcal{E}_{B|A}$ corresponds exactly to the map associated to $\rho_{B|A} : L(H_A) \to L(H_B)$ via the Jamiolkowski Isomorphism.

Let us now turn our attention to causally connected regions, where one region (say A) can have a causal influence on the other (say B). The aim is for causally related regions to be described in the same framework as the acausally related regions we saw before, since probability and correlation should be independent of causation.

We define the causal conditional state of B given A as the following operator on H_{AB}:

$\sigma_{B|A} = \rho_{B|A}^{T_A}$,
where $(\cdot)^{T_A}$ is the partial transpose in some basis on H_A.

The causal quantum belief propagation rule can be written using the CP map
$\rho_{B|A} : L(H_A) \to L(H_B)$, as:
$\rho_B = \mathcal{E}_{B|A}(\rho_A) := Tr_A(\sigma_{B|A}\rho_A)$.

We can also define the causal joint state of causally related regions A and B as the operator:
$\sigma_{AB} = \sigma_{B|A} \star \rho_A$.

Measurement using the Born rule (in region A) can be derived as a special case of quantum belief propagation, since measurement of a POVM $\{E_Y^A\}$: $P(Y = y) = Tr(E_Y^A \rho_A)$ can just be expressed as: $\rho_Y = Tr_A(\sigma_{Y|A}\rho_A)$, where $\rho_Y = \sum_Y P(Y = y)|y\rangle\langle y|$.

In this way, the authors showed that dynamics, preparation of states, measurements, state update, the Heisenberg picture and even classical-quantum hybrid models can all be described in a unified manner using the quantum belief propagation rule (which is essentially the same in the causal and acausal cases).

Leifer and Spekkens then proceed to introduce a quantum version of Bayes theorem:
$P(R|S) = \frac{P(S|R)P(R)}{P(S)}$, which relates the conditional states of B given A and of A given B.

The quantum Bayes theorem for acausal and causal conditional states are then:
$\rho_{A|B} = \rho_{B|A} \star (\rho_A \rho_B^{-1}) = \rho_{B|A} \star (\rho_A (Tr_A(\rho_{B|A}\rho_A)^{-1}))$
and: $\sigma_{A|B} = \sigma_{B|A} \star (\rho_A \rho_B^{-1}) = \sigma_{B|A} \star (\rho_A (Tr_A(\sigma_{B|A}\rho_A)^{-1}))$
respectively.

These can both be generalized as a quantum channel of the form:
$\mathcal{F}_{A|B}(\cdot) = \rho_A^{\frac{1}{2}} (\mathcal{E}_{A|B}^{\dagger}(\rho_B^{-\frac{1}{2}})(\cdot)\rho_B^{-\frac{1}{2}})\rho_A^{\frac{1}{2}}$.

One can then find a very similar Bayes theorem for classical-quantum hybrid models. This new acausal quantum Bayes rule allows the authors to study retrodiction, time symmetry and remote steering in an elegant manner.

It is then interesting to generalize Bayesian conditioning to the quantum case and see how quantum behavior may arise from simply updating information. There are still many issues to be resolved since this work is in progress (see the paper [70] for more details).

A related article by Leifer and Poulin [71] describes in more detail the mathematical structures that may underly quantum belief propagation. They study in detail belief propagation algorithms acting on graphical models.

Another interesting advance is an article by Coecke and Spekkens [72] introducing a graphical framework for Bayesian inference, based on ideas from categorical quantum mechanics. Developing such a formalism further could yield new insights and help the progress of a formulation of quantum mechanics as a theory of Bayesian inference.

We will now turn our attention to another operational interpretation of quantum mechanics in which probability theory plays a major role, namely generalized probabilistic theories.

Chapter 8

Generalized probabilistic theories

For any theory, whether it applies to nature or not, we can consider a number of important features of the theory. This allows us to understand traits of nature in a more general context than just quantum mechanics. Indeed, the study of a broad range of theories within an operational framework can yield considerable insight. This can, for example, help differentiate between different theories within the framework, simplify calculations within any of these theories or reveal novel fundamental features of the world. In the rest of this report, we will present two distinct analyses of larger spaces of hypothetical theories containing quantum theory, namely generalized probabilistic theories and generalized process theories.

8.1 Hardy's operational framework

We will start by describing a framework for convex operational theories, introduced by Hardy [73], in which quantum theory is derived from a set of five axioms. Like in most operational approaches, he considers preparation devices which prepare a system in a given state, transformation devices, and measurement devices whose distinct outcomes correspond to macroscopic events. Central to his axioms are the two integers:

K, which is the number of degrees of freedom, defined as the minimum number of probability measurements needed to determine the state.

N, which is the dimension, defined as the maximum number of states that can be reliably distinguished from one another in a single shot measurement.

Quantum theory can then be derived from the following axioms:

Axiom 1: Relative frequencies tend to the same value (called probability) for any case when a given measurement is performed on an ensemble of n systems, given some preparation, as n goes to infinity.

Axiom 2: K is a function of N which takes the minimal value allowed by the axioms, for each N.

Axiom 3: A subsystem with has support on only M states of a set of N distinguishable states, behaves like a system of dimension M.

Axiom 4: A composite system containing subsystems A and B has: $N = N_A N_B$ and $K = K_A K_B$.

Axiom 5: There exists a continuous reversible transformation on a system between any two pure states of that system.

Note that if we do not include the word 'continuous' in axiom 5 then we obtain classical probability theory (with K=N) instead of quantum theory (with $K = N^2$).

Let us now sketch how quantum theory can be derived from these axioms and introduce Hardy's operational framework for convex operational theories in the process.

The first axiom simply defines probability. This uses a frequentist approach but the framework is compatible with any of the standard interpretations of probability. It is then possible to define the state of a system as: "any mathematical object which can be used to determine the probability for any measurement that could possibly be performed on the system". In order not to over-specify the state, it is also useful to introduce a set of fiducial measurements as: "a certain minimum number K of appropriately chosen measurements which are both necessary and sufficient to determine the state". This means that the (operational) state is fully specified by a vector of probabilities $p = (p_1, ..., p_K)^T$ of getting a given outcome in each of the fiducial measurements.

Any probability p_m that can be measured, is assumed to be determined by a function f of the state p: $p_m = f(p)$. The first postulate, together with the possibility of probabilistically preparing states, means that f is linear and therefore: $p_m = r \cdot p$, where r is a vector associated with measurement. Note that the fiducial measurement vectors are the Cartesian basis vectors $r^i = e_i = (0, ..., 1, ..., 0)^T$.

Transformations of the system correspond to real KxK matrices Z such that: $p \to Zp$. The set of allowed states, measurements and transformations are all convex sets.

One can then define pure states as non zero extremal states of the convex state space S, i.e. non-zero vectors in S which cannot be written as a convex sum of other vectors in S. The identity measurements and normalization of states can be similarly defined. We also expect there to be sets of states p_n (at most N of them), called basis states, which are distinguishable from one another in a single-shot measurement, by measurement vectors r_m (which cover all outcomes) such that:

$r_m \cdot p_n = \delta_{mn}$.

In this way, we can see that physical systems are characterized by their dimension N (number of basis states) and the number of degrees of freedom K (number of fiducial measurements).

Although we will not go through the proof here, Hardy showed that, in general, the axioms imply: $K = N^r$, where r=1,2,... . The second axiom then tells us that we must take the smallest value of r which is consistent with the other axioms.

The third and fourth axioms dictate how subsystems combine to form larger systems, but we will not insist on how these work since we shall return later to the description of how separate systems combine in generalized probabilistic theories.

As we mentioned before, the fifth axiom provides the distinction between quantum theory and classical probability theory. It implies that there exists an allowable reverse transformation Z^{-1} for any input state and that the set of reversible transformations forms a compact Lie group. This means, for example, that a pure states can always be transformed to any other pure state along a continuous trajectory through pure states.

Such a thing is not possible for classical states, since the space of classical pure states corresponds to vertices of a simplex. One can then show that, in accordance with axiom 2, quantum theory and classical probability theory are both special cases of these convex operational theories satisfying: $K = N^2$ and K=N respectively.

8.2 Information theoretic constraints for quantum theory

Two years after Hardy's paper, Clifton, Bub and Halvorson [74] attempted to derive quantum theory from information theoretic axioms only. They adopt the following axioms:

Axiom 1: It is impossible to transfer superluminal information between two physical systems by performing measurements on one of them.

Axiom 2:It is impossible to perfectly broadcast the information contained in an unknown physical state.

Axiom 3: It is impossible to unconditionally perform secure bit commitment.

The authors work in a $C\star$-algebraic framework which encompasses both classical and quantum statistical theories. They argue that quantum theory can be picked out from this general $C\star$ framework by the satisfaction of certain physical constraints: kinematic independence, non commutativity, and nonlocality. They then formulate their three axioms (which are known to hold in quantum theory) in $C\star$ algebraic terms and show that they imply kinematic independence, non commutativity, and nonlocality.

This is an interesting result but it can be criticized since it assumes a framework which is not much more general than quantum theory to begin with and since it fails to establish the full structure of quantum theory. This work, however, showed that progress can be made in understanding the connection between information processing and physical principles in general by studying information processing in a wide range of theories. Such an insight was an important motivation for the study of information processing in generalized probabilistic theories, which we shall describe in the following section.

8.3 Information processing in generalized probabilistic theories

Barrett introduced a framework [75] for generalized probabilistic theories which is based on Hardy's formalism. The five Hardy axioms are now replaced by the following assumptions:

Assumption 1: The state of a single system is completely specified by the vector of probabilities for the outcomes of all fiducial measurements:

$\vec{P} = (P(a=1|X=1), P(a=2|X=1), ...; P(a=1|X=2), P(a=2|X=2), ...; ...)^T$

where $P(a=i|X=j)$ is the probability of getting outcome i when fiducial measurement j is performed on the system.

Assumption 2: The set of allowed normalized states (satisfying $|\vec{P}| = \sum_i P(a=i|X=j) = 1, \forall j$) is closed and convex. The complete set of states S is is the convex hull of allowed normalized states and $\vec{0}$.

Assumption 3: An element of the set of allowed operations $\{M_i\} \in O$ must satisfy:

$0 \leq \frac{|M_i \cdot \vec{P}|}{|\vec{P}|} \leq 1, \forall i, \vec{P} \in S$

$\sum_i \frac{|M_i \cdot \vec{P}|}{|\vec{P}|} = 1, \forall \vec{P} \in S$

$M_i \cdot \vec{P} \in S, \forall i, \vec{P} \in S$

A set of transformations $\{M_i\}$ is an element of O if and only if M_i is a element of the set of allowed transformations T (for all i) and $\sum_i \frac{|M_i \cdot \vec{P}|}{|\vec{P}|} = 1, \forall \vec{P} \in S$. We assume that such a set T exists and by definition it is convex.

Assumption 4: The final state of a joint system does not depend on the order in which operations are independently performed on on each of its subsystems.

Assumption 5: The global state of a system can be completely determined by specifying joint probabilities of outcomes for fiducial measurements performed on each subsystem. Also, if the joint state $\vec{P}^{AB}$ is in the set of allowed states S^{AB} for the joint system AB, then the reduced state $\vec{P}^A$ for system A (with outcome probabilities $P(a=i|X=j) = \sum_{i'} P(a=i, b=i'|X=j, Y=j')$) is in the set of allowed states S^A for system A.

Assumption 6: If $\vec{P}^A \in S^A$ and $\vec{P}^B \in S^B$ then $\vec{P}^A \otimes \vec{P}^B \in S^{AB}$.

Assumption 7: A theory first specifies a set of allowed states, then all transformations M_i^A that are well defined, in the sense that $(M_i^A \otimes I)\vec{P}^{AB} \in S^{AB}$ whenever $\vec{P}^{AB} \in S^{AB}$, are allowed transformations.

The first three assumptions lead to convex operational theories very similar to those derived by Hardy's axioms [73]. Although, the Barrett assumptions takes the degrees of freedom of the state as internal degrees of freedom (requires a closer analysis of the role of spacetime) and treats transformations and measurements in a unified way. The other assumptions deal with how systems combine to make other systems. These allow us to derive (as a theorem) that systems combine according to a tensor product rule, which leads to a natural definition of entanglement in these theories. Note that the no-signaling principle is a corollary of assumption 4.

Several features, which at first seem specifically quantum, arise in all these generalized probabilistic theories, except the classical one. These include the disturbance of a system on measurement [76], the multiple decompositions of a mixed state into pure states and the no-cloning theorem [77] (no-deleting theorem[78]?). The paper also describes in detail classical and quantum theory in the framework, along with GNS (general non-signaling theory), containing states giving rise to PR-box correlations [14], and GLT (generally local theory), where all states are local. Barrett asks what further assumptions would uniquely identify quantum theory, and proposes that quantum theory might be optimal for computation. A number of open questions are being addressed with regard to entropy [79], time and causal structure [80] in these generalized probabilistic theories.

For the remainder of this review, we will analyze another set of hypothetical physical theories containing quantum theory: generalized process theories.

Chapter 9

Categorical quantum mechanics

9.1 Categories and quantum mechanics

A category C consists of a class of objects $|C|$ and for each pair of objects $R, S \in |C|$, a collection (hom-set) C(R,S) of arrows. It also has a composition map $- \circ -$: C(R,S) x C(S,T)$\rightarrow C(R,T)$ for any triple of objects $R, S, T \in |C|$, which is associative, i.e: h $\circ (g \circ f) = (h \circ g) \circ f$, for all f, g, h in C(R,S), C(S,T), C(R,T) respectively, and for each object R there exists an identity arrow $id_R : R \rightarrow R$ such that: $f \circ id_R = id_S \circ f$.

Category theory was first introduced by Eilenberg and Mac Lane [81] and increasingly thorough introductions to the theory can be found in [82, 83, 84]. Due to the lack of space, we will skip to the most relevant concepts and refer the reader to the references above in order to fill in the gaps.

A symmetric monoidal category (SMC) consists of a category C, a bifunctor $-\otimes$-:CxC $\rightarrow C$, a unit object I and natural isomorphisms:
$\lambda_A : A \cong I \otimes A$, $\rho_A : A \cong A \otimes I$, $\alpha_{A,B,C} : A \otimes (B \otimes C) \cong (A \otimes B) \otimes C$ and $\sigma_{A,B} : A \otimes B \cong B \otimes A$ along with coherence conditions (see [84] for details).

Monoidal categories are ideal for describing very general compositional theories of systems and processes, since they contain two interacting modes $\otimes$ and $\circ$ of composing systems and processes. These lead to a very simple diagrammatic calculus [85] where arrows are represented by boxes and the objects are vertical inputs/outputs. The $\otimes$ and $\circ$ operations are respectively represented as boxes juxtaposed next to each other and attached in vertical sequence.

There is a theorem by Joyal and Street [86] which states that an equation in the symbolic language of SMCs holds if and only if it holds up to an isomorphism of diagrams in the graphical language. This allows us to use diagrammatic reasoning by isomorphism of diagram and substitution to undergo complex calculations. In the hope of recovering the structure of quantum theory, we must add extra algebraic structure to our SMCs.

A dagger compact symmetric monoidal category (†-CSMC) C is a SMC with an identity-on-objects involutive contravariant endofunctor $\dagger : C \rightarrow C$ such that:
$(f \circ g)^\dagger = g^\dagger \circ f^\dagger$, $(f \otimes g)^\dagger = f^\dagger \otimes g^\dagger$, $id_A^\dagger = id_A$ and $(f^\dagger)^\dagger = f$ (†-SMC),
which is also compact, meaning that each object $A \in |C|$ has a dual object $A^* \in |C|$ and arrows: $\eta_A : I \rightarrow A^* \otimes A$ and $\epsilon_A : A \otimes A^* \rightarrow I$ such that:

$(\epsilon_A \otimes id_A) \circ (id_A \otimes \eta_A) = id_A$ and $(id_{A^*} \otimes \epsilon_A) \circ (\eta_A \otimes id_{A^*}) = id_{A^*}$

(Note that we will usually use $A = A^*$).

We define a state of a system A as an arrow: $\psi : I \rightarrow A$, an effect as: $\pi : A \rightarrow I$ and scalars as $s : I \rightarrow I$. The inner product between states is then the scalar: $\psi^\dagger \circ \phi : I \rightarrow I$.

If we add to the previous graphical calculus an involutive asymmetry in the boxes representing arrows and the rule that taking the adjoint reflects the boxes vertically and the fact that $f \circ f^\dagger = id_A = f^\dagger \circ f$ then we get the following key theorem which allows us to use graphical reasoning as with SMCs:

Theorem [87, 88]: An equational statement between formal expressions in the language of †-CSMC holds if and only if it holds up to isotopy in the graphical calculus.

Important examples of †-CSMCs are the category of finite dimensional Hilbert spaces and linear maps/tensor products FHilb, the category of finite sets with relations/Cartesian products FRel and the category of open quantum systems and CP maps/tensor products CP(FHilb). Therefore, we can use the graphical calculus to derive results for any of these categories.

9.2 Important Categorical Background

In order to fully describe quantum theory using the graphical calculus, it is necessary to add a few more concepts, including measurement, observables and classical channels. Such a graphical description requires a novel approach to defining bases and observables which holds in the general context of †-CSMC and reduces naturally to the familiar concepts in the standard Hilbert space formalism. The key insight in doing so is that the contrapositive of the no cloning and no deleting theorems [77, 78, 89] states that orthonormal basis states are the only ones which can be copied and erased.

A mathematical formalization of this idea [90, 91] leads to the following algebraic definition of observable structures:

An observable structure is a †-special commutative Frobenius algebra on a †-CSMC C. This is a triple
{ A $\in |C|$, $\delta : A \to A \otimes A$ (copying map), $\epsilon : A \to I$ (erasing map) } satisfying the following:
(i) $\{A, \delta, \epsilon\}$ is a cocommutative comonoid, i.e: $(\delta \otimes id_A) \circ \delta = (id_A \otimes \delta) \circ \delta$; $\lambda_A^{-1} \circ (\epsilon \otimes id_A) \circ \delta = \rho_A^{-1} \circ (id_A \otimes \epsilon) \circ \delta = id_A$ and $\sigma_{A,A} \circ \delta = \delta$.
(ii)It satisfies the Frobenius law: $(\delta^\dagger \otimes id_A) \circ (id_A \otimes \delta) = \delta \circ \delta^\dagger$
(iii) It is special: $\delta^\dagger \circ \delta = id_A$

We know that [91] in FHilb, orthonormal bases are in a one to one correspondence with †-special commutative Frobenius algebras. This structure captures the essential information encoded by non-degenerate observables and allows us to depict classical states and measurement in a pictorial manner. This definition for observable structures has been shown [92] to be equivalent to the spider laws described in the next section.

Let $\{A, \delta, \epsilon\}$ be an observable structure. If for each state $\psi_\alpha : I \to A$ we define: $\psi_{-\alpha} := (\psi_\alpha^*)^\dagger$ (* is complex conjugation in the basis of the observable) and we restrict the states to those which respect: $\delta^\dagger \circ (\psi_\alpha \otimes \psi_{-\alpha}) = \epsilon^\dagger$, then we get an abelian group with the multiplication $\delta^\dagger$ which we call the phase group. This allows us to define spiders diagrams with phases which are a generalization of phase gates in quantum computing (S rules below).

Two observable structures $\{A, \delta_Z, \epsilon_Z\}$ and $\{A, \delta_X, \epsilon_X\}$ are complementary if the so-called eigenstates of one are unbiased for the other. In our language these observables are complementary if and only if: $\delta_Z^\dagger \circ \delta_X = \epsilon_Z \circ \epsilon_X^\dagger$.

In the rules of ZX calculus, we require strong complementarity between pairs of observables. Strong complementarity implies complementarity but the converse is untrue [93]. Strongly complementary

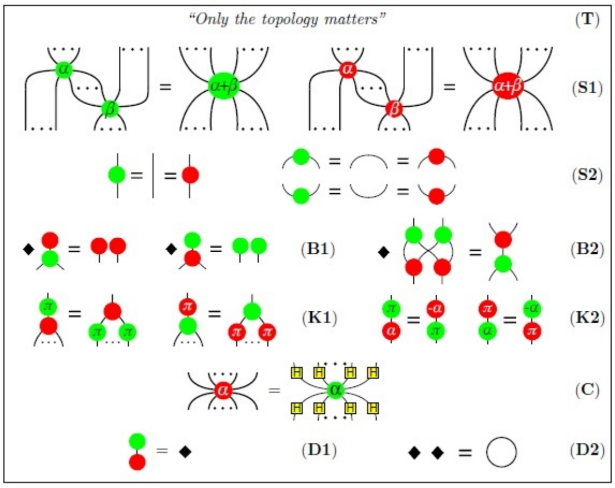

Figure 9.1: Diagrammatic rules for ZX calculus

observables form what is called a scaled Hopf algebra [94] (B rules below).

In the next section, we will describe the ZX calculus, which utilizes the notions introduced in the previous sections. The use of two complementary observables results in a two-colored pictorial calculus. In the present review, we have only had the space to present some of the most important concepts which lead up to the ZX calculus but for a more thorough derivation of the calculus itself, the reader is invited to read [93].

9.3 ZX Calculus

The rules of the ZX calculus, which uses as complementary observables the Pauli Z and Pauli X operators, are summarized in Figure 9.1. The (T) rule means that after identifying the inputs and outputs, any topological deformation of the internal structure does not matter.

This allows diagrammatic reasoning by isomorphism of diagram and substitution to be applied as before: diagrams are built out of tensor products and compositions of generators.

Quantum circuits can be constructed from the diagrams of ZX calculus as shown in Figure 9.2.

We know that if two diagrams are equal according to the rules of ZX calculus then their corresponding quantum circuits are equivalent. Note that the converse is not true. This provides a convenient way of doing calculations for qubits using the graphical reasoning. Since single qubit unitaries and C-Not gates are universal for quantum computing [95], the ZX calculus is also universal for quantum computing if we allow arbitrary phases. It is interesting to study the property of the calculus for different phase groups.

Also, measurement based quantum computing [96] can easily be described using the ZX calculus. The diagrammatic reasoning could even provide an efficient way of translating between the cluster state and circuit models of quantum computation.

The ZX calculus simplifies numerous quantum calculations. It allows us to study a number of fundamental aspects of quantum theory from a high-level mathematical point of view.

9.4 Recent work

We will now briefly mention some recent work which is relevant. First of all, a useful practical advance is the development of an automated reasoning software tool called "quantomatic" which is based on

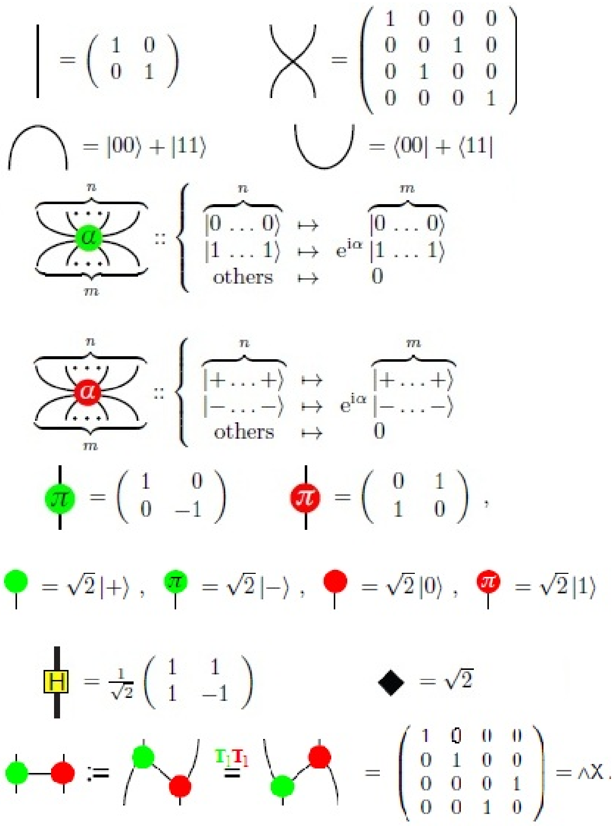

Figure 9.2: Quantum circuit interpretation of the ZX Calculus

the work done in [97]. When finished, this tool will allow us to move very efficiently between diagrams in the ZX calculus.

Another important development which deserves more mention is the diagrammatic representation of completely positive maps, described in [88]. This allows for a more general pictorial description of quantum phenomena using a generalized version of Kraus operators. To see how this construction fits in with the ZX calculus, see section 12 of [93].

Much of the recent work done in categorical quantum mechanics is dedicated to the investigation of the structure of multipartite entanglement [98]. The pictorial calculus has even been used to analyze topological quantum computing [99].

Another fruitful application of the categorical formalism has been in the study of non-locality. A novel paper [100] shows that a key difference between qubit stabilizer theory [101] (which cannot be modeled by a local hidden variable theory) and Spekkens toy theory [36] (which can be modeled by a local hidden variable theory) is their phase group. For stabilizer theory the phase group is $\mathbb{Z}_4$, whereas for Spekkens toy theory it is $\mathbb{Z}_2$x$\mathbb{Z}_2$.

Building further on [100], it was shown [102] that there is a close connection between GHZ-Mermin type non-locality [22] and strong complementarity.

9.5 Stabilizer quantum theory

A very useful subclass of quantum mechanical operations is stabilizer quantum mechanics. Stabilizer states are eigenstates with eigenvalue 1 of each operator in a subgroup of the Pauli group:

$P_n := \{\alpha g_1 \otimes ... \otimes g_n : \alpha \in \{\pm 1, \pm i\} \wedge g_k \in \{I, \sigma_x, \sigma_y, \sigma_z\}, \forall k\}$.

The Clifford group is the group of unitary operations:

$C_n := \{U : UgU^\dagger \in P_n, \forall g \in P_n\}$.

It is generated by the phase, Hadamard and C-NOT gates. The local Clifford group lC_n consists of the n-fold tensor products of single qubit Clifford operators.

Stabilizer quantum mechanics [101] includes preparations of qubits in the $|0\rangle$ state, Clifford unitaries and measurements in the computational basis. This subclass of quantum mechanics is particularly useful for quantum error correction and can be simulated using a classical computer.

An interesting piece of work, which is closely related to Edward's result on the phase group of the ZX Calculus [100], is Pusey's stabilizer notation for Spekkens toy theory [103].

Note that the stabilizer formalism can be generalized beyond two dimensional qubit systems [104]. We can generalize the Pauli group $\mathcal{P}_n$ to higher dimensions by generating it from tensor products of (ω^a times products of) d^{th} order X_d and Z_d (instead of X and Z), where:

$X_d|j\rangle = |j+1\rangle$ and $Z_d|j\rangle = \omega^j|j\rangle$, where ω is a primitive d^{th} root of unity

and $X_d Z_d = \omega^{-1} Z_d X_d$.

Stabilizer states are then eigenstates with eigenvalue 1 of each operator in a subgroup of $\mathcal{P}_n$.

The Clifford group $\mathcal{C}_n$ (at least for prime dimensions d) is then the set of operators that leave $\mathcal{P}_n$ invariant under conjugation. This is generated by the following d-dimensional gates:

(1) Discrete Fourier transform: $|j\rangle \to \sum_{s=0}^{d} \omega^{js}|s\rangle$

(2) SUM gate (generalized CNOT): $|i\rangle|j\rangle \to |i\rangle|i \oplus j[d]\rangle$

(3) Phase gate: $|j\rangle \to \omega^{j(j-1)/2}|j\rangle$

(4) Exponent gate: $|j\rangle \to |aj\rangle$.

Stabilizer quantum mechanics in d dimensions then corresponds to preparations of stabilizer states, operators in $\mathcal{C}_n$ and measurements of operators in $\mathcal{P}_n$.

9.6 Completeness of ZX Calculus for stabilizer quantum mechanics

Recently, Backens [105] has shown that the ZX calculus is complete for stabilizer quantum mechanics. This means that any equation between two ZX calculus diagrams (put into matrix mechanics) which can be shown to be true using stabilizer quantum mechanics is derivable using the rules of the ZX calculus. In this section, we shall outline her proof of this result.

A graph is an object $G = \{V, E\}$, where V is a set of vertices and E is a collection of (ordered or unordered) pairs of vertices called edges.

The adjacency matrix θ of a graph has θ_{ij}=1 if there is an edge connecting vertices i and j, and θ_{ij}=0 otherwise. Given a graph G (with n vertices), we can define a graph state $|G\rangle$ as the state formed by preparing a $|+\rangle$ state at each vertex and the applying a controlled-Z gate at each edge.

$|G\rangle$ is the unique qubit state whose stabilizer subgroup is generated by:

$X_a \otimes \bigotimes_{b \in V} Z_b^{\theta_{ab}}, \forall a \in V$.

It is a known result [106] that two graph states are equivalent under local Clifford operations iff they can be transformed into each other by local complementations. Note that, in order for it to satisfy this 'Van den Nest theorem', an extra axiom corresponding to the Euler decomposition of the Hadamard gate has to be added to the ZX calculus [107] (this paper introduces local complementation in the ZX calculus).

It was also shown in [106] that any stabilizer state is equivalent to some graph state, under local Clifford operations.

Backens proved that a direct analogue of these theorems holds within the ZX calculus. She demonstrated (using a similar approach as [108]) that the stabilizer states represented by two ZX calculus diagrams are equal, then these diagrams can always be shown to be equal within the calculus by converting them into a certain 'GS-LC' form (corresponding to graph states in the calculus with a single Clifford operator applied to each output). In this form, they can always be converted to one another by local complementations in the calculus. The proof is generalized for operators by using the Choi-Jamiolkowski isomorphism. In this way, Backens showed that the ZX Calculus is complete for stabilizer quantum mechanics.

A modified version of this proof could potentially be applied to show that other restricted subclasses of quantum theory (maybe more physical subclasses; restrictions of quantum optics, for example) which can express graph states, are complete for stabilizer quantum mechanics.

9.7 A complete set of circuit equations for stabilizer quantum mechanics

We shall now describe a novel result, presented here for the first time. In fact, thanks to Backen's proof [105] that the ZX Calculus is complete for stabilizer quantum mechanics, this result yields two new insights.

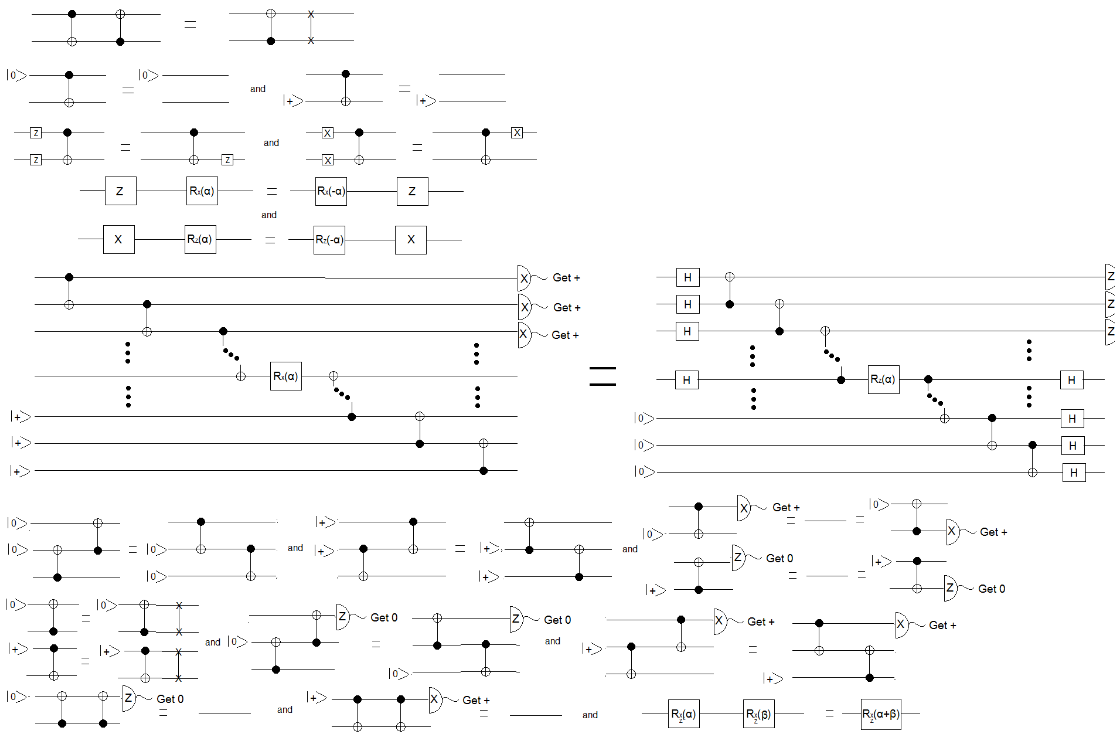

Figure 9.3: Complete set of circuit equations for stabilizer quantum mechanics.

First of all, we have found a complete set of circuit equations for stabilizer quantum mechanics, presented in 9.3.

These are a set of circuit equations which have the property that any circuit equation which can be shown to be true using stabilizer theory (which means that both quantum circuits in the equation correspond to equivalent processes in stabilizer quantum mechanics) can be derived from this set by diagrammatic reasoning, using isomorphisms of diagrams and substitution. This teaches us something about the logical aspect of the stabilizer formalism.

Note that the circuits in 9.3 have some degree of degeneracy. Many of these equations are just special instances of C-NOT gates with various input states and post-selected measurements.

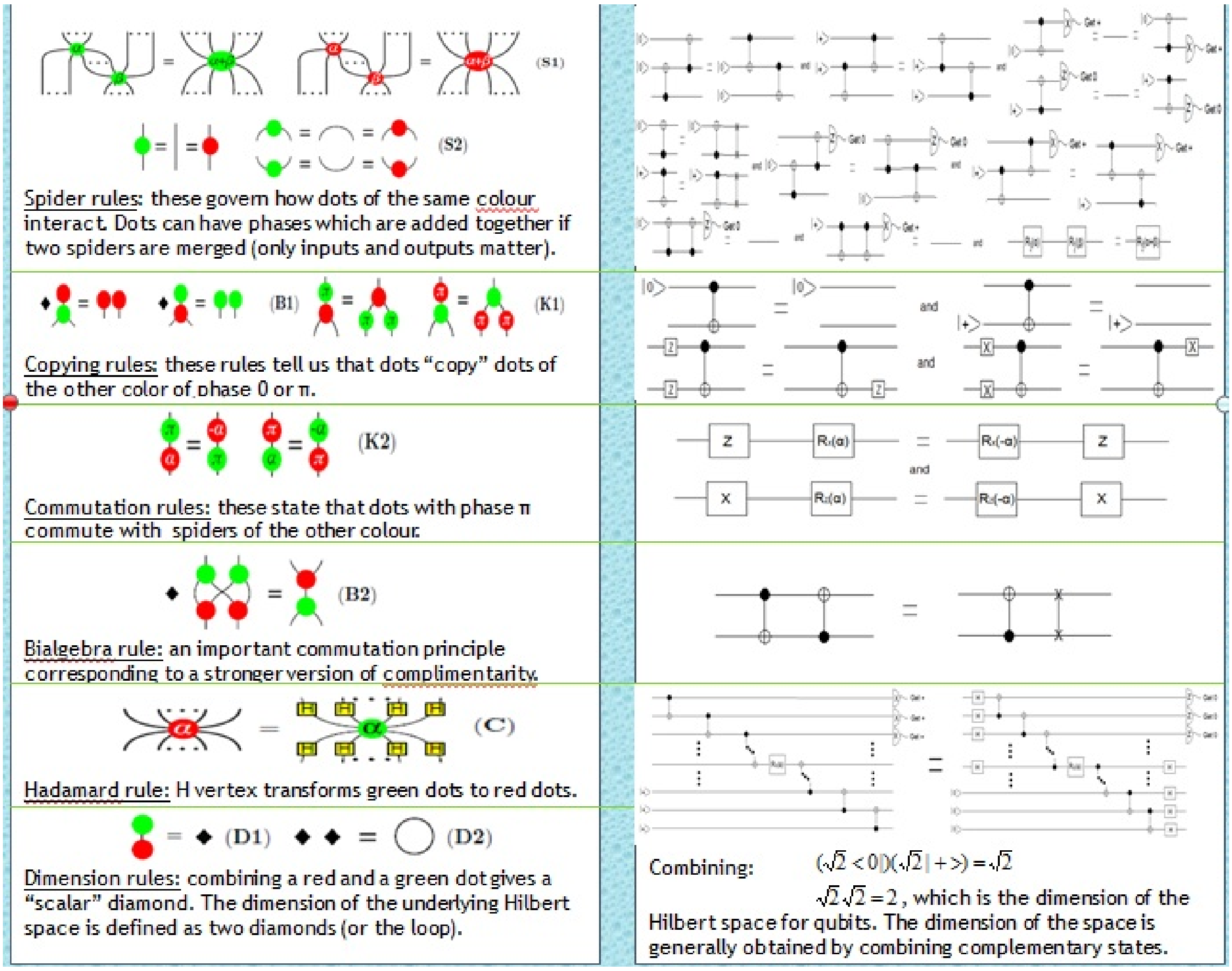

Figure 9.4: Circuit equations for each ZX Calculus axiom.

The second insight that can be obtained from this result is the following theorem:

Theorem:: There is an equivalence of categories between the symmetric monoidal categories of quantum circuits $\mathcal{F}_{SMC}(Circ)$ and of the ZX calculus $\mathcal{F}_{SMC}(ZX)$ (quotient to their axioms):

$\mathcal{F}_{SMC}(Circ)/\equiv_{Circ}\leftrightarrow \mathcal{F}_{SMC}(ZX)/\equiv_{ZX}$.

An equivalence of categories means that there exists a full, faithful, essentially surjective functor. The constructive proof of the existence of this functor required us to find an equivalent set of ZX circuits to the axioms of the ZX Calculus, which are in a form that can be directly related to quantum circuits using 9.2.

$\mathcal{F}_{SMC}(Circ)$ is a symmetric monoidal category over the monoidal signature [109] (these are the consistent 'gates' of the symmetric monoidal category):

$\mathcal{S} := \{C-NOT; SWAP; prepare|0\rangle; prepare|+\rangle; postselect|0\rangle, postselect|+\rangle, R_x(\alpha); R_z(\beta)\}$.

The axioms of the ZX Calculus ($\mathcal{F}_{SMC}(ZX)$) and the corresponding axioms for the category $\mathcal{F}_{SMC}(Circ)$ (which are like the circuit version of the ZX-Calculus axioms) are given in 9.4. This gives us a new insight into the structure of the ZX Calculus, namely an understanding of what the axioms of the calculus signify, in terms of familiar quantum circuits.

The fact that the ZX calculus is complete for stabilizer quantum mechanics means that any equation between two ZX calculus diagrams (put into matrix mechanics) which can be shown to be true using stabilizer quantum mechanics is derivable using the rules of the ZX calculus. Therefore, the theorem above proves that any circuit equation which can be shown to be true using stabilizer theory can be derived using the set of quantum circuit equations 9.3 presented earlier.

9.8 Future directions

We will now succinctly describe some of the ideas for further research which follow from the previous work we described. A natural question which ensues from [102] is whether there may be a family of different notions of complementarity which can be connected to particular computational tasks. This might allow to generalize other schemes in an analogous manner to the way in which Mermin's non-locality is generalized in [102] (quantum secret sharing [110] could be an example).

Another interesting avenue of research could be to study general process theories, which could give insights about key concepts of quantum theory (maybe in a more general setting) and could even lead to some new ideas for alternative axioms for quantum mechanics. Combining the general process theories and the complementary approach of generalized probabilistic theories [111] could also potentially lead to some interesting results.

Also, categorical quantum mechanics could be utilized to study further the structure of multipartite entanglement and could allow a fruitful analysis of the problem of MREGS [112] and LOCC reversibility from a new perspective. Maybe we could find some generalized measures of entanglement, a definition of entropy in †-CSMCs, new multipartite entanglement algebraic structures providing clues on MREGS or other hints at a general theory of multipartite entanglement.

In addition, since it has been shown to be complete for stabilizer quantum mechanics, we could use the pictorial calculus to describe error-correction protocols and fault tolerance in a general way (this is in progress...). Perhaps this could lead to an improved understanding of quantum error correction and even useful new codes. Variations of this idea could help develop error correction for particular quantum computing implementations.

An important step in determining how useful the ZX calculus would be for calculation could be the determination of the computational resources needed to compute results using the calculus. For example, what is the complexity class associated with converting from a circuit computation to a cluster state computation using the ZX calculus?

A great deal of work has gone into generalizing the uncertainty relations to a more general entropic setting [113]. These entropic uncertainty relations, which can perhaps be seen as an alternative (better?) way of looking at complementarity, have recently been shown to be closely related to non-locality [114]. Perhaps this important work could be understood in a categorical setting and extended.

The ZX Calculus may also be an ideal arena to study contextuality. This could perhaps yield insight into how contextuality arises, the role it plays in quantum-like theories and how it could be used as a resource for computation.

Next, it would be good to generalize the ZX calculus to some elegant ZXY calculus with more than just two complementary observables.

It should be valuable to extend the formalism beyond qubits. Indeed, a higher dimensional version of the calculus could exhibit numerous quantum features and hopefully incorporate both higher dimensional Spekkens toy model [38] and stabilizer theory [104] (maybe we could find a generalization in higher dimensions of the relation described in [100] [103]).

Finally, the generality of monoidal categories (both FHilb and nCob are dagger compact categories, for example) and some previous work [115, 116] shows that this framework could be very useful in the quest for a theory of quantum gravity.

Chapter 10

Conclusion

Although considerable progress has been made regarding our understanding of the foundations of quantum theory, the sheer number of conflicting interpretations demonstrates the extent of the task left ahead. Can we find an interpretation, perhaps transcending quantum theory, which may lead us to a scientific consensus? Or do our philosophical preconceptions always make such concordance unattainable?

We often seem hindered by our naive intuitions and by the ineptitude to detach ourselves from any object under analysis. Will such limitations forever cause the percipience of an underlying reality to slip through our fingers?

Take this kiss upon the brow!
And, in parting from you now,
Thus much let me avow–
You are not wrong, who deem
That my days have been a dream;
Yet if hope has flown away
In a night, or in a day,
In a vision, or in none,
Is it therefore the less gone?
All that we see or seem
Is but a dream within a dream.

I stand amid the roar
Of a surf-tormented shore,
And I hold within my hand
Grains of the golden sand–
How few! yet how they creep
Through my fingers to the deep,
While I weep–while I weep!
O God! can I not grasp
Them with a tighter clasp?
O God! can I not save
One from the pitiless wave?
Is all that we see or seem
But a dream within a dream?

Bibliography

[1] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935), URL `http://link.aps.org/doi/10.1103/PhysRev.47.777`.

[2] N. Bohr, Phys. Rev. **48**, 696 (1935), URL `http://link.aps.org/doi/10.1103/PhysRev.48.696`.

[3] E. Schrdinger, Mathematical Proceedings of the Cambridge Philosophical Society **31**, 555 (1935), URL `http://www.journals.cambridge.org/abstract_S0305004100013554`.

[4] A. Einstein, Journal of the Franklin Institute **221**, 349 (1936), ISSN 0016-0032, URL `http://www.sciencedirect.com/science/article/pii/S0016003236910475`.

[5] J. von Neumann, *The Mathematical Foundations of Quantum Mechanics* (Princeton Univ. Press, Princeton NJ, 1932), 1st ed.

[6] J. S. Bell, Rev. Mod. Phys. **38**, 447 (1966), URL `http://link.aps.org/doi/10.1103/RevModPhys.38.447`.

[7] J. S. Bell, Physics **1**, 195 (1964), URL `http://www.physics.princeton.edu/~mcdonald/examples/QM/bell_physics_1_195_64.pdf`.

[8] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969), URL `http://link.aps.org/doi/10.1103/PhysRevLett.23.880`.

[9] A. Aspect, P. Grangier, and G. Roger, Phys. Rev. Lett. **47**, 460 (1981), URL `http://link.aps.org/doi/10.1103/PhysRevLett.47.460`.

[10] J. S. Bell, *Speakable and unspeakable in quantum mechanics; Chapter 12*, Collected papers on quantum philosophy (Cambridge Univ. Press, Cambridge, 2004).

[11] A. Garg and N. D. Mermin, Phys. Rev. D **35**, 3831 (1987), URL `http://link.aps.org/doi/10.1103/PhysRevD.35.3831`.

[12] B. S. Cirel'son, Letters in Mathematical Physics **4**, 93 (1980), ISSN 0377-9017, 10.1007/BF00417500, URL `http://dx.doi.org/10.1007/BF00417500`.

[13] S. Popescu and D. Rohrlich, Foundations of Physics **24** (1994).

[14] S. Popescu and D. Rohrlich, Physics Letters A **169**, 411 (1992), ISSN 0375-9601, URL `http://www.sciencedirect.com/science/article/pii/0375960192908198`.

[15] B. F. Toner and D. Bacon, Phys. Rev. Lett. **91**, 187904 (2003), URL `http://link.aps.org/doi/10.1103/PhysRevLett.91.187904`.

[16] W. van Dam, eprint arXiv:quant-ph/0501159 (2005), `arXiv:quant-ph/0501159`.

[17] V. Scarani, AIP Conference Proceedings (Melville, New York) **844**, 309 (2006).

[18] N. Gisin and A. Peres, Physics Letters A **162**, 15 (1992), ISSN 0375-9601, URL `http://www.sciencedirect.com/science/article/pii/037596019290949M`.

[19] A. Peres, *Quantum Theory: Concepts and Methods*, Fundamental Theories of Physics (Springer, 1995), ISBN 9780792336327, URL `http://books.google.co.uk/books?id=rMGqMyFBcL8C`.

[20] A. Peres, Found.Phys. **29**, 589 (1999), `quant-ph/9807017`.

[21] L. Hardy, Phys. Rev. Lett. **71**, 1665 (1993), URL `http://link.aps.org/doi/10.1103/PhysRevLett.71.1665`.

[22] N. D. Mermin, American Journal of Physics **58**, 731 (1990), URL `http://link.aip.org/link/?AJP/58/731/1`.

[23] E. Specker, Dialectica **1**, 239 (1960).

[24] A. Gleason, Indiana Univ. Math. J. **6**, 885 (1957), ISSN 0022-2518.

[25] S. Kochen and E. Specker, Journal of Mathematics and Mechanics **17**, 59 (1967).

[26] A. Peres, Journal of Physics A: Mathematical and General **24**, L175 (1991), URL `http://stacks.iop.org/0305-4470/24/i=4/a=003`.

[27] M. Planat, European Physical Journal Plus **127**, 86 (2012), `1204.4275`.

[28] M. Kernaghan and A. Peres, Physics Letters A **198**, 1 (1995), ISSN 0375-9601, URL `http://www.sciencedirect.com/science/article/pii/037596019500012R`.

[29] J. L. Bell, philpapers (1996).

[30] S. Abramsky and L. Hardy, Phys. Rev. A **85**, 062114 (2012), URL `http://link.aps.org/doi/10.1103/PhysRevA.85.062114`.

[31] N. D. Mermin, Phys. Rev. Lett. **65**, 3373 (1990), URL `http://link.aps.org/doi/10.1103/PhysRevLett.65.3373`.

[32] N. Harrigan and T. Rudolph, ArXiv e-prints (2007), `0709.4266`.

[33] N. Harrigan, T. Rudolph, and S. Aaronson, ArXiv e-prints (2007), `0709.1149`.

[34] T. Rudolph, eprint arXiv:quant-ph/0608120 (2006), `arXiv:quant-ph/0608120`.

[35] E. G. Beltrametti and S. Bugajski, J. Phys. A: Math. Gen. **28**, 3329 (1995).

[36] R. W. Spekkens, Physical Review A **75**, 1 (2007), URL `http://link.aps.org/doi/10.1103/PhysRevA.75.032110`.

[37] J. A. Larsson, in *American Institute of Physics Conference Series* (2012), vol. 1424 of *American Institute of Physics Conference Series*, `1111.3561`.

[38] O. Schreiber and R. W. Spekkens, to be published (2012-3).

[39] R. W. Spekkens, Phys. Rev. A **71**, 052108 (2005), URL `http://link.aps.org/doi/10.1103/PhysRevA.71.052108`.

[40] R. W. Spekkens, D. H. Buzacott, A. J. Keehn, B. Toner, and G. J. Pryde, Phys. Rev. Lett. **102**, 010401 (2009), URL `http://link.aps.org/doi/10.1103/PhysRevLett.102.010401`.

[41] A. Cabello, S. Severini, and A. Winter, ArXiv e-prints (2010), `1010.2163`.

[42] Y.-C. Liang, R. W. Spekkens, and H. M. Wiseman, Physics Reports **506**, 1 (2011), ISSN 0370-1573, URL `http://www.sciencedirect.com/science/article/pii/S0370157311001517`.

[43] L. Hardy, Studies in History and Philosophy of Science Part B **35**, 267 (2004).

[44] A. Montina, Phys. Rev. A **77**, 022104 (2008), URL `http://link.aps.org/doi/10.1103/PhysRevA.77.022104`.

[45] M. F. Pusey, J. Barrett, and T. Rudolph, Nature Physics **8**, 476 (2012), `1111.3328`.

[46] P. G. Lewis, D. Jennings, J. Barrett, and T. Rudolph, ArXiv e-prints (2012), `1201.6554`.

[47] C. M. Caves, C. A. Fuchs, and R. Schack, Studies in History and Philosophy of Science Part B **38**, 255 (2007).

[48] D. Bohm, Phys. Rev. **85**, 166 (1952), URL `http://link.aps.org/doi/10.1103/PhysRev.85.166`.

[49] D. Bohm, Phys. Rev. **85**, 180 (1952), URL `http://link.aps.org/doi/10.1103/PhysRev.85.180`.

[50] D. Bohm and B. Hiley, *The undivided universe: an ontological interpretation of quantum theory*, Physics, philosophy (Taylor & Francis, 1995), ISBN 9780415121859, URL `http://books.google.co.uk/books?id=vt9XKjc4WAQC`.

[51] H. Everett, Rev. Mod. Phys. **29**, 454 (1957), URL `http://link.aps.org/doi/10.1103/RevModPhys.29.454`.

[52] B. DeWitt and N. Graham, *The Many-Worlds Interpretation of Quantum Mechanics*, Princeton Series in Physics (Books on Demand, 1973), ISBN 9780783719429, URL `http://books.google.co.uk/books?id=3fvBAAAACAAJ`.

[53] H. P. Stapp, *The basis problem in many-worlds theories* (2002).

[54] A. Kent, International Journal of Modern Physics A **5**, 1745 (1990), `arXiv:gr-qc/9703089`.

[55] D. Deutsch, Royal Society of London Proceedings Series A **455**, 3129 (1999), `arXiv:quant-ph/9906015`.

[56] D. Wallace, eprint arXiv:quant-ph/0211104 (2002), `arXiv:quant-ph/0211104`.

[57] G. C. Ghirardi, A. Rimini, and T. Weber, Phys. Rev. D **34**, 470 (1986), URL `http://link.aps.org/doi/10.1103/PhysRevD.34.470`.

[58] P. Pearle, in *Lecture Notes in Physics, Berlin Springer Verlag*, edited by H.-P. Breuer and F. Petruccione (1999), vol. 526 of *Lecture Notes in Physics, Berlin Springer Verlag*, p. 195, `arXiv:quant-ph/9901077`.

[59] F. Dowker and A. Kent, Phys. Rev. Lett. **75**, 3038 (1995), URL `http://link.aps.org/doi/10.1103/PhysRevLett.75.3038`.

[60] J. G. Cramer, Rev. Mod. Phys. **58**, 647 (1986), URL `http://link.aps.org/doi/10.1103/RevModPhys.58.647`.

[61] R. D. Sorkin, ArXiv General Relativity and Quantum Cosmology e-prints (1995), `arXiv:gr-qc/9507057`.

[62] F. Dowker, J. Henson, and R. D. Sorkin, Modern Physics Letters A **19**, 1829 (2004), `arXiv:gr-qc/0311055`.

[63] M. Redei, **27**, 493 (1997).

[64] G. Birkhoff and J. von Neumann, Annals of Mathematics **37**, 823843 (1936).

[65] J.-Y. Girard, Theoretical Computer Science **50**, 1 (1987), ISSN 0304-3975, URL `http://www.sciencedirect.com/science/article/pii/0304397587900454`.

[66] B. Coecke, ArXiv e-prints (2012), `1204.3458`.

[67] B. de Finetti, *Theory of Probability* (New York: Wiley, translation by A Machi and AFM Smith of 1970 book, 1970).

[68] A. N. Kolmogorov, *Landmark Writings in Western Mathematics 1640-1940: Chapter 75* (Elsevier Science, 1933), ISBN 9780444508713, URL http://books.google.co.uk/books?id=UdGBy8iLpocC.

[69] C. M. Caves, C. A. Fuchs, and R. Schack, Phys. Rev. A **65**, 022305 (2002), URL http://link.aps.org/doi/10.1103/PhysRevA.65.022305.

[70] M. S. Leifer and R. W. Spekkens, ArXiv e-prints (2011), 1107.5849.

[71] M. Leifer and D. Poulin, Annals of Physics **323**, 1899 (2008), ISSN 0003-4916, URL http://www.sciencedirect.com/science/article/pii/S0003491607001509.

[72] B. Coecke and R. Spekkens, Synthese **186**, 651 (2012), ISSN 0039-7857, 10.1007/s11229-011-9917-5, URL http://dx.doi.org/10.1007/s11229-011-9917-5.

[73] L. Hardy, eprint arXiv:quant-ph/0101012 (2001), arXiv:quant-ph/0101012.

[74] R. Clifton, J. Bub, and H. Halvorson, Foundations of Physics **33**, 1561 (2003), ISSN 0015-9018, 10.1023/A:1026056716397, URL http://dx.doi.org/10.1023/A:1026056716397.

[75] J. Barrett, eprint arXiv:quant-ph/0508211 (2005), arXiv:quant-ph/0508211.

[76] C. A. Fuchs and A. Peres, Phys. Rev. A **53**, 2038 (1996), URL http://link.aps.org/doi/10.1103/PhysRevA.53.2038.

[77] W. K. Wootters and W. H. Zurek, Nature **299**, 802 (1982), URL http://dx.doi.org/10.1038/299802a0.

[78] A. K. Pati and S. L. Braunstein, Nature **404**, 164 (2000), URL http://arxiv.org/abs/quant-ph/9911090.

[79] H. Barnum, J. Barrett, L. O. Clark, M. Leifer, R. Spekkens, N. Stepanik, A. Wilce, and R. Wilke, New Journal of Physics **12**, 033024 (2010), URL http://stacks.iop.org/1367-2630/12/i=3/a=033024.

[80] L. Hardy, Journal of Physics A: Mathematical and Theoretical **40**, 3081 (2007), URL http://stacks.iop.org/1751-8121/40/i=12/a=S12.

[81] S. Eilenberg and S. MacLane, Transactions of the American Mathematical Society **58**, 231 (1945), URL http://www.jstor.org/stable/1990284?origin=crossref.

[82] B. Coecke, What is category theory **45-74arXiv**, 4574 (2006), URL http://arxiv.org/abs/0808.1032.

[83] S. Abramsky and N. Tzevelekos, University Computing **813**, 96 (2011), URL http://arxiv.org/abs/1102.1313.

[84] S. Mac Lane, *Categories for the working mathematician*, vol. 5 (Springer, 1998), URL http://www.amazon.com/dp/0387984038.

[85] P. Selinger, New Structures for Physics pp. 1–63 (2009), URL http://arxiv.org/abs/0908.3347.

[86] A. Joyal and R. Street, Advances in Mathematics **88**, 55 (1991), URL http://linkinghub.elsevier.com/retrieve/pii/000187089190003P.

[87] G. M. Kelly and M. L. Laplaza, Journal of Pure and Applied Algebra **19**, 193 (1980), URL http://dx.doi.org/10.1016/0022-4049(80)90101-2.

[88] P. Selinger, Electronic Notes in Theoretical Computer Science **170**, 139163 (2007), URL http://linkinghub.elsevier.com/retrieve/pii/S1571066107000606.

[89] D. Dieks, Physics Letters A **92**, 271 (1982), URL `http://linkinghub.elsevier.com/retrieve/pii/0375960182900846`.

[90] B. Coecke, D. Pavlovic, and J. Vicary, ENTCS (2008), URL `http://arxiv.org/abs/0810.0812`.

[91] B. Coecke and D. Pavlovic, in *Mathematics of Quantum Computing and Technology*, edited by G. Chen, L. Kauffman, and S. Lamonaco (Taylor and Francis, 2007), pp. 567–604, arXiv:quant-ph/0608035.

[92] B. Coecke and E. O. Paquette, Electronic Notes in Theoretical Computer Science **210**, 15 (2006), URL `http://arxiv.org/abs/quant-ph/0608072`.

[93] B. Coecke and R. Duncan, in *Proceedings of the 37th International Colloquium on Automata, Languages and Programming (ICALP)* (2008), Lecture Notes in Computer Science, extended version: arXiv:quant-ph/09064725.

[94] P. Cartier and H. Etudes, Frontiers in Number Theory Physics and Geometry II p. 537615 (2006), URL `http://inc.web.ihes.fr/prepub/PREPRINTS/2006/M/M-06-40.pdf`.

[95] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, vol. 70 (Cambridge University Press, 2000), URL `http://link.aip.org/link/AJPIAS/v70/i5/p558/s2&Agg=doi`.

[96] R. Raussendorf, D. E. Browne, and H. J. Briegel, Physical Review A **68**, 1 (2003), URL `http://arxiv.org/abs/quant-ph/0301052`.

[97] A. Kissinger (2012), `1203.0202`, URL `http://arxiv.org/abs/1203.0202`.

[98] B. Coecke and A. Kissinger, in *Proceedings of the 37th International Colloquium on Automata, Languages and Programming (ICALP)* (2010), Lecture Notes in Computer Science, pp. 297–308, extended version: arXiv:1002.2540.

[99] C. Horsman, New Journal of Physics **13**, 18 (2011), URL `http://arxiv.org/abs/1101.4722`.

[100] B. Coecke, B. Edwards, and R. W. Spekkens, Electronic Notes in Theoretical Computer Science **270**, 15 (2011), arXiv:1003.5005.

[101] D. Gottesman, Energy **2008**, 114 (1997), URL `http://arxiv.org/abs/quant-ph/9705052`.

[102] B. Coecke, R. Duncan, A. Kissinger, and Q. Wang (2012), `arXiv/1203.4988`.

[103] M. Pusey, Foundations of Physics **42**, 688 (2012), ISSN 0015-9018, 10.1007/s10701-012-9639-7, URL `http://dx.doi.org/10.1007/s10701-012-9639-7`.

[104] D. Gottesman, Chaos, Solitons and Fractals **10**, 1749 (1999), ISSN 0960-0779, URL `http://www.sciencedirect.com/science/article/pii/S0960077998002185`.

[105] M. Backens, to be published (2012).

[106] M. Van den Nest, J. Dehaene, and B. De Moor, Phys. Rev. A **69**, 022316 (2004), URL `http://link.aps.org/doi/10.1103/PhysRevA.69.022316`.

[107] R. Duncan and S. Perdrix, ArXiv e-prints (2009), `0902.0500`.

[108] S. Anders and H. J. Briegel, Phys. Rev. A **73**, 022334 (2006), URL `http://link.aps.org/doi/10.1103/PhysRevA.73.022334`.

[109] A. Kissinger, ArXiv e-prints (2012), `1203.0202`.

[110] M. Hillery, V. Buzek, and A. Berthiaume, Physical Review A **59**, 6 (1998), URL `http://arxiv.org/abs/quant-ph/9806063`.

[111] J. Barrett, Phys. Rev. A **75**, 032304 (2007), URL `http://link.aps.org/doi/10.1103/PhysRevA.75.032304`.

[112] C. H. Bennett, S. Popescu, D. Rohrlich, J. A. Smolin, and A. V. Thapliyal, Physical Review A **63**, 13 (1999), URL `http://arxiv.org/abs/quant-ph/9908073`.

[113] S. Wehner and A. Winter (2009), `0907.3704`, URL `http://arxiv.org/abs/0907.3704`.

[114] J. Oppenheim and S. Wehner, Science **330**, 1072 (2010), URL `http://www.sciencemag.org/cgi/doi/10.1126/science.1192065`.

[115] J. C. Baez, The structural foundations of quantum gravity pp. 1–21 (2004), URL `http://arxiv.org/abs/quant-ph/0404040`.

[116] M. Atiyah, La Rivista Del Nuovo Cimento Series 3 **21**, 1 (1993), URL `http://www.springerlink.com/index/10.1007/BF02725740`.