Alternative theories in Quantum Foundations

André O. Ranchin

Submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Physics of Imperial College London

July 2016

Supervised by
Bob Coecke and Terry Rudolph

Department of Physics

Imperial College London

Alternative theories in Quantum Foundations

André O. Ranchin

Submitted for the degree of Doctor of Philosophy
July 2016

Abstract

Abstraction is an important driving force in theoretical physics. New insights often accompany the creation of physical frameworks which are both comprehensive and parsimonious. In particular, the analysis of alternative sets of theories which exhibit similar structural features as quantum theory has yielded important new results and physical understanding. An important task is to undertake a thorough analysis and classification of quantum-like theories. In this thesis, we take a step in this direction, moving towards a synthetic description of alternative theories in quantum foundations.

After a brief philosophical introduction, we give a presentation of the mathematical concepts underpinning the foundations of physics, followed by an introduction to the foundations of quantum mechanics. The core of the thesis consists of three results chapters based on the articles in the author's publications page. Chapter 4 analyses the logic of stabilizer quantum mechanics and provides a complete set of circuit equations for this sub-theory of quantum mechanics. Chapter 5 describes how quantum-like theories can be classified in a periodic table of theories. A pictorial calculus for alternative physical theories, called the ZX calculus for qudits, is then introduced and used as a tool to depict particular examples of quantum-like theories, including qudit stabilizer quantum mechanics and the Spekkens-Schreiber toy theory. Chapter 6 presents an alternative set of quantum-like theories, called quantum collapse models. A novel quantum collapse model, where the rate of collapse depends on the Quantum Integrated Information of a physical system, is introduced and discussed in some detail. We then conclude with a brief summary of the main results.

Acknowledgments

Scientific research resembles the activity of an underground explorer who undertakes the arduous task of digging tunnels and constructing elaborate subterranean passages in search of elusive precious minerals. Naturally, this process – which will more often lead to a frustrating conclusion than to the launch of a fruitful enterprise – cannot be undertaken alone. Friends and family provide a pillar of strength which buffers the impact of the inevitable collapse of theoretical caverns. Fortunately, there is only a minute risk of suffocation and it is only in a metaphorical sense that one may end up covered in dirt and trapped in a confined space. Moreover, failure is a far better teacher than success. I have certainly learnt many things in the last few years.

It is my pleasure to mention some of the people who have played an essential role throughout my PhD. First of all, I would like to thank my supervisors Bob Coecke and Terry Rudolph for their insightful help and helpful insights, kindly given whenever necessary. Their patience and understanding provided an invaluable catalyst for intellectual development throughout my years of academic study in London and Oxford. I am also much obliged to Sandu Popescu for introducing me to the fascinating world of quantum physics research and encouraging me to pursue further study.

I have had the privilege of working with not just one but two exceptional circles of colleagues: the Oxford Quantum group and the Imperial Controlled Quantum Dynamics group. A particular mention should go to Miriam Backens, Raymond Lal, William Zeng, Ross Duncan, Aleks Kissinger, John Selby, Sean Mansfield, Rui Soares Barbosa, Nadish De Silva, Hugo Nava Kopp, Dan Marsden, Jonathan Barrett, Destiny Chen and to the eleven members of the CQD DTC Cohort 3. Mihai-Dorian Vidrighin and Mark Mitchison especially deserve my appreciation for their friendship and for their irreplaceable assistance and advice.

A particular word of thanks should go to the directors of the CQD DTC, particularly Danny Segal, whose compassionate words of support provided solace at a most difficult time. I must also mention the staff of the BHOC, without whose effort and care, the present work would not have seen the light of day.

Although it goes without saying, I would like to sincerely thank my parents for their constant encouragement as well as for the key role they have played in my academic and personal development.

Finally, my deepest gratitude goes to Sylvia, whose thoughtful and unfailing support has provided an invaluable bedrock for any achievement of mine.

I acknowledge financial support from the EPSRC.

Declaration of Originality:

I declare that all the work presented in this thesis is my own or is properly referenced such that the original source is clearly stated.

Author's publications

1. A. O. Ranchin, B. Coecke, "**Complete set of circuit equations for Stabilizer Quantum Mechanics**". *Physical Review A*, **90**, 012109 (2014).

2. A. O. Ranchin "**Depicting qudit quantum mechanics and mutually unbiased qudit theories**". *EPTCS Quantum Physics and Logic 2014*, **172** (2014).

3. K. Kremnizer, A. O. Ranchin "**Integrated Information-induced quantum collapse**". *Foundations of Physics*, **45** (2015).

Contents

Acknowledgments ii

Author's publications iv

List of Figures viii

1 Introduction 1

2 Background I: Mathematical tools 8

2.1 Set theory . . . 9

2.1.1 Axiomatic set theory . . . 9

2.1.2 Relations and functions . . . 11

2.1.3 Numbers . . . 12

2.2 Group Theory . . . 14

2.3 Algebraic structures . . . 19

2.3.1 Rings, Fields and Galois theory . . . 19

2.3.2 Linear Algebra and Graph theory . . . 24

2.4 Topology and Hilbert spaces . . . 28

2.4.1 Topology . . . 28

2.4.2 Topological vector spaces . . . 32

2.5 Category theory . . . 36

2.5.1 Categories and functors . . . 36

2.5.2 Limits . . . 37

2.5.3 Examples of categories . . . 38

2.5.4 Natural Transformations and adjoints . . . 39

2.5.5 Categorical quantum mechanics . . . 41

3 Background II: Quantum theory 47

3.1 Operational theories . . . 47

3.2 Quantum mechanics introduced . . . 49

3.2.1 Orthodox postulates . . . 49

3.2.2 Operational axioms . . . 51

3.3 Quantum computation . . . 52

3.3.1 Quantum circuits . . . 52
3.3.2 Other quantum computation models . . . 55
3.4 Non locality and Contextuality . . . 57
3.4.1 Realism and quantum theory . . . 57
3.4.2 EPR . . . 57
3.4.3 Bohr response . . . 58
3.4.4 Hidden variables and Von Neumann's no go theorem . . . 59
3.4.5 Bell's theorem and the CHSH inequality . . . 61
3.4.6 Cirelson bound . . . 63
3.4.7 Popescu Rohrlich boxes . . . 63
3.4.8 Generalized CHSH inequality . . . 64
3.4.9 Mermin non-locality . . . 65
3.4.10 The over-protective seer . . . 67
3.4.11 Gleason's theorem . . . 68
3.4.12 Bell corollary of Gleason's theorem . . . 69
3.4.13 Kochen Specker theorem . . . 70
3.4.14 Mermin magic square . . . 71
3.4.15 Leggett-Garg inequality . . . 72
3.5 Ontological models for quantum mechanics . . . 73
3.5.1 Examples of ontological models . . . 74
3.5.2 Spekkens toy theory . . . 77
3.5.3 Contextuality for ontological models . . . 79
3.5.4 PBR theorem . . . 81
3.6 Ontological interpretations of quantum theory . . . 83
3.6.1 Bohmian mechanics . . . 83
3.6.2 Many-worlds theory . . . 84
3.6.3 Collapse models . . . 86
3.7 Generalized probabilistic theories . . . 88
3.7.1 Hardy's operational framework . . . 89
3.7.2 Information theoretic constraints for quantum theory . . . 91
3.7.3 Information processing in generalized probabilistic theories . . . 92

4 The logic of Stabilizer Quantum Mechanics 94
4.1 Introduction . . . 96
4.2 Stabilizer quantum theory . . . 98
4.3 ZX network . . . 99
4.4 Completeness of the ZX calculus . . . 102
4.5 Quantum circuits for the ZX network axioms . . . 104
4.6 Proof of the Equivalence Lemma . . . 106
4.7 A complete set of circuit equations for stabilizer quantum mechanics . . . 116
4.8 Derivation of an equation between stabilizer quantum circuits from the complete set . . . 120
4.9 Reasoning with the ZX network is easier than using the quantum circuit calculus 122
4.10 Conclusion . . . 123

5 A periodic table of quantum-like theories **125**
5.1 Introduction 128
5.2 Explicit models of theories 130
5.2.1 Qudit stabilizer quantum mechanics 131
5.2.2 Spekkens toy theory in higher dimensions 132
5.3 Depicting qudit quantum mechanics and toy models 135
5.3.1 Derivation of the qudit ZX calculus 135
5.3.2 The ZX calculus for qudit quantum mechanics: 144
5.3.3 Mutually unbiased qudit theories 153
5.3.4 Picturing stabilizer quantum mechanics 153
5.3.5 Depicting Spekkens-Schreiber toy theory for dits 159
5.4 A periodic table of quantum-like theories 166
5.5 Topological Ontological models 170
5.6 Further work 174

6 Quantum collapse theories and Quantum Integrated Information **176**
6.1 The philosophy of consciousness 178
6.1.1 History 178
6.1.2 Philosophical positions 180
6.1.3 Problems 181
6.2 Consciousness and Integrated Information 183
6.3 Calculating the Quantum Integrated Information 185
6.4 A review of existing quantum collapse theories 187
6.4.1 Pearle's collapse equation 187
6.4.2 GRW Model 188
6.4.3 QMULP Model 190
6.4.4 Continuous Spontaneous Localization Model 191
6.4.5 Gravity-induced collapse models 191
6.4.6 Adler trace dynamics 193
6.5 Integrated Information and state-vector reduction 193
6.6 Experimental tests of Integrated Information-induced collapse 196
6.7 Conclusion 197

7 Conclusion **199**

Bibliography **201**

List of Figures

3.1 A preparation process. . . . 48
3.2 A transformation process. . . . 48
3.3 A measurement process. . . . 49
3.4 The Bloch sphere representation of a qubit. . . . 53
3.5 Examples of basic quantum gates. . . . 54
3.6 Plaquette and vertex operators on a section of the toric code. . . . 56

4.1 Generating diagrams for the ZX network. . . . 99
4.2 Diagrammatic rules for the ZX network. . . . 100
4.3 Quantum circuit interpretation of the ZX network elements. . . . 101
4.4 Alternative ZX axioms in a form resembling quantum circuit equations. . . . 107
4.5 Sound and complete set of circuit equations for stabilizer quantum mechanics. 119

5.1 Examples of symmetry in nature: the snowflake, honeycomb lattice and aloe polyphylla. . . . 125
5.2 The five Platonic Solids. . . . 126
5.3 Hall of the Abencerrajes in the Alhambra palace. . . . 127
5.4 Polya's representation of the 17 plane symmetry groups. . . . 127
5.5 Generating diagrams for the qudit ZX calculus. . . . 145
5.6 Hilbert space interpretation of the qudit ZX calculus elements. . . . 147
5.7 Depicting qutrit stabilizer theory on two tori. . . . 155
5.8 Group table for the four-dimensional stabilizer phase group. . . . 158
5.9 Periodic table of quantum-like theories. . . . 167
5.10 Plot of the fundamental polynomial for Spekkens toy theory. . . . 168
5.11 Plot of the fundamental polynomial for stabilizer quantum mechanics. . . . 169

Chapter 1

Introduction

"To those who look at the rich material provided by history, and who are not intent on impoverishing it in order to please their lower instincts, their craving for intellectual security in the form of clarity, precision, 'objectivity', 'truth', it will become clear that there is only one principle that can be defended under all circumstances and in all stages of human development. It is the principle: anything goes."

Paul Feyerabend

WHAT is the purpose of scientific research – does the knowledge obtained through scientific investigation play a singular role in human understanding? Is it possible to draw a line of demarcation, thereby allowing us to distinguish between scientific inquiry and other, non-scientific activities?

Since antiquity, it has been recognized[240,17] that the acquisition of knowledge can be achieved through three types of critical reasoning: deduction, induction and abduction. *Deductive reasoning* is the process of using unambiguous rules to reach a logically certain conclusion from one or more initial premises. Given a theory of evidence[270] and some initial premises based upon empirical observation, *inductive reasoning* indicates some degree of support for a new claim. *Abductive reasoning* is the process of using empirical evidence to construct a theory – which strives to be minimal in terms of theoretical simplicity – accounting for the observed evidence by providing a sufficient (but not necessary) explanation of the observations.

Despite the countless potential interpretations of any observed physical phenomenon, scientific progress relies upon the use of abduction to obtain a single economical explanation, a tangible theoretical narrative for the physical process under examination. If one is skeptical about dubious claims concerning demonstrable 'truths' about reality, then what distinguishes a scientific account of our experience from other alternative descriptions?

(A) For Karl Popper, the key feature of scientific theories is the possibility of empirical *falsification*. Scientific theories are abstract constructions which can never be proven and that can only be tested indirectly, by reference to their implications [244]:
"[T]here can be no statements in science which can not be tested, and therefore none which cannot in principle be refuted, by falsifying some of the conclusions which can be deduced from them."

(B) Thomas Kuhn built on Popper's claim that scientific observations and evaluations are theory-laden, in the sense that they cannot be separated from their interpretation within a particular logically consistent theoretical paradigm [194]. Moreover, he argued that it is not possible to evaluate two competing paradigms independently from each other, as neither one provides a standard by which the other can be judged. Hence, Kuhn introduced a division between normal science, which takes place within a paradigm, and extraordinary science which leads to a *paradigm shift*. The key differentiating feature of science is then due to social and subjective factors, related to the structure of the scientific community [194]:
"A paradigm is what the members of a community of scientists share, and, conversely, a scientific community consists of men who share a paradigm."

(C) Paul Thagard provided another clear demarcation criterion, which describes a theory as non-scientific if and only if [283]:
"(i) it has been less progressive than alternative theories over a long period of time, and faces many unsolved problems; but
(ii) the community of practitioners makes little attempt to develop the theory towards solutions of the problems, shows no concern for attempts to evaluate the theory in relation to others, and is selective in considering confirmations and disconfirmations."

(D) More recently, David Deutsch [109] presented another test to distinguish science and non-science, namely whether the explanatory narrative provided by a theoretical framework can be easily varied. He argues that: "easy variability is the sign of a bad explanation, because, without a functional reason to prefer one of countless variants, advocating one of them, in preference to the others, is irrational."

(E) Building upon Pierre Duhem's earlier work [116], Willard Quine introduced *epistemological holism*, which is the view that individual statements cannot be confirmed or disconfirmed by empirical tests, but only coherent sets of statements can be verified together [249]: "our statements about the external world face the tribunal of sense experience not individually but only as a corporate body."

Scientific knowledge then corresponds to a bundle of hypotheses – including all background assumptions – which can be tested against the empirical world and undergoes falsification if it fails the observational test. The Duhem-Quine thesis states that it is impossible to isolate any single hypothesis in the bundle.

(F) Inspired by Popper, Kuhn and Quine, Imre Lakatos introduced the concept of *research programs*. Satisfactorily developed methods and theories form the 'hard core' of the research program, and scientists can add auxiliary hypothesis to a 'protective belt' which defends the core of the program from falsification [196]. Arbitrary theoretical amendments in the protective belt can cause a research program to be progressive – if they enhance the program's explanatory or predictive power – or degenerative – if they have been made out of necessity, in the face of new and troublesome evidence. In this respect:

"The positive heuristic of the programme saves the scientist from becoming confused by the ocean of anomalies."

Whether or not research programs are scientific depends on their success at predicting novel facts.

(G) Larry Laudan's *pessimistic induction* [199] strongly questions the role of induction and the possibility of convergent realism in scientific theories:

"the history of science furnishes vast evidence of empirically successful theories that were later rejected; from subsequent perspectives, their unobservable terms were judged not to refer and thus, they cannot be regarded as true or even approximately true."

This leads to the idea that if science distinguishes itself from other activities at all, it is only due to pragmatic efficiency[198]: "the aim of science is to secure theories with a high problem-solving effectiveness". In this respect, scientific progress will often proceed counter-inductively:

"Indeed, on this model, it is possible that a change from an empirically well-supported theory to a less well-supported one could be progressive, provided that the latter resolved significant conceptual difficulties confronting the former."

(H) Another possible interpretation of science is as a specific language or collection of languages, either mathematical or discursive. We could then apply Ferdinand de Saussure's semiotic methodology[259] and interpret science as a collection of *signs*, comprising of:

(i) signifiers (*signifiant*), which are symbols or sounds allowing for the identification of a sign

(ii) signified (*signifié*), corresponding to the meaning of the sign acquired through the differences between signifiers.

Scientific signs can be particularly opaque since each word is at the summit of a pyramid of concepts and ideas. In this interpretation we should abandon any attempt to relate the scientific language to 'truth' or 'real objects' but think of scientific theories as self-enclosed systems in which any semantic content consists of internal interrelations between signifiers.

In this light, it seems difficult to circumvent Jacques Derrida's objection that language leads to an endless process where meaning is sought but never found[105]. We can introduce the concept of *différance scientifique*. Scientific concepts and constructions consisting of abstract signs and words only have meaning because of the contrast (différence) between these signs and possible alternatives. Moreover, meaning is never present but rather is acquired at a later stage, deferred (différé) to other signs. For Derrida: 'Il n'y a pas de hors-texte'; indeed language leads to a perpetual movement of differences in which there is no stable equilibrium and one can no longer appeal to reality as a refuge independent of language. Scientists should therefore be wary of our desire for immediate access to meaning; we must ensure that there is a process of deconstructing the 'metaphysics of presence'[105], a constant effort to avoid privileging presence over absence.

(I) A strong critique of scientific progress is laid out in Paul Feyerabend's 'Against Method'[136]. Through the use of many examples from the history of science, he argues:
(i) for the inevitable use of counter-inductive reasoning in science:
"Hypotheses contradicting well-confirmed theories give us evidence that cannot be obtained in any other way."
(ii) against the requirement that scientific theories must always be consistent:
"[T]here is not a single interesting theory that agrees with all the known facts in its domain"
(iii) for the inevitability of dissociating scientific theories and facts from their process of genesis:
"[T]he material which a scientist actually has at his disposal, his laws, his experimental results, his mathematical techniques, his epistemological prejudices, his attitude towards the absurd consequences of the theories which he accepts, is indeterminate in many ways, ambiguous, and never fully separated from the historical background."

This analysis leads Feyerabend to *epistemological anarchism* or the conclusion that there are no useful, fixed methodological rules governing the growth of knowledge or the progress of science. This means that:
"Knowledge so conceived is not a series of self-consistent theories that converges towards an ideal view; it is not a gradual approach to the truth. It is rather an ever increasing ocean of mutually incompatible alternatives, each single theory, each fairy-tale, each myth that is part of the collection forcing the others into greater articulation and all of them contributing, via this process of competition, to the development of our consciousness."

The twentieth century has initiated the process of exorcising the illusion that scientific research is revealing a profound, objective and undeniable truth. It seems that Hélène Cixous was right in saying that: "We are living in an age where the conceptual foundation of an ancient culture is in the process of being undermined by millions of moles of a species which has yet to be identified." Scientists have a perpetual duty to question the conceited and complex story they create, to live up to the motto: *'nullius in verba'*. Despite our technological advances and apparent progress in understanding, one must be reluctant to assume that the elaborate myth created by interpreting scientific knowledge is superior to other human legends.

How should our understanding of the nature of science affect and shape the work of the practicing scientist? It is essential to take on board the lessons from the philosophy of science when working in the foundations of physics. First of all, we must always ensure that there is the utmost clarity in any use of scientific language – mathematical or otherwise – and its relation to theory-laden observations. Any theoretical lack of consistency and possible gap in scientific reasoning should never be swept under the carpet but must be dealt with directly and without hypocrisy.

Moreover, the theory-laden nature of empirical observation, the importance of paradigm shifts, différance scientifique and epistemological anarchism all point to the crucial importance of **studying alternative scientific theories**. In this respect, we both have the freedom and the duty to analyze the "increasing ocean of mutually incompatible alternatives" and to show tolerance towards the diverse range of possible methodologies and theoretical constructs.

As Freud reminds us :

"Only in the study of the abnormal can we learn the true nature of the normal."

The first two chapters of this thesis will introduce the necessary background material for a thorough understanding of the foundations of quantum mechanics.

The initial chapter presents a brief exploration of the concepts and language of mathematics. The focus is to provide a succinct construction of the tower of theoretical concepts underlying a rigorous analysis of the foundations of physics. In this respect, there are many definitions and numerous omissions (notably proofs).

The second background chapter proceeds by discussing essential physics background, with a special emphasis on quantum theory. This chapter gives a concise summary of important work in the foundations of quantum mechanics.

The following three chapters present the main research results of the thesis and loosely follow the three articles in the *author's publications list*.

The first results chapter analyses the logic of Stabilizer quantum theory and provides a complete set of circuit equations for this sub-theory of quantum mechanics.

The second results chapter describes how quantum-like theories can be classified in a periodic table of theories. A pictorial calculus for alternative physical theories, called the ZX calculus for qudits, is then introduced and used as a tool to depict particular examples of quantum-like theories, including qudit stabilizer quantum mechanics and the Spekkens-Schreiber toy theory.

The final results chapter presents an alternative set of quantum-like theories, called quantum collapse models. A novel quantum collapse model, where the rate of collapse depends on the Quantum Integrated Information of a physical system, is introduced and discussed in some detail.

We then conclude the thesis with a short summary.

Chapter 2

Background I: Mathematical tools

Mathematical abstraction is at the heart of the progress in describing physical processes. New developments in Physics often go hand in hand with novel insights about the mathematical language used to narrate our evolving story about 'physical reality'.

We will now aim to introduce some mathematical tools, a language which will be the foundation of our description of physical theories. Given the elusive nature of an independent physical reality existing independently from observation, we shall stress the importance of presenting physical theories from an operational point of view. In this light, it is desirable to have a general mathematical formalism providing an abstract and broad representation of physical processes corresponding to physical preperations, transformations and measurements. We will start by introducing standard mathematical objects used as theoretical tools in Physics and shall then present **Category Theory**, which is a powerful device for creating a general formalism of physical theories.

Given our task of analyzing foundational physical theories, it is an essential responsibility for us to spend some time exploring the landscape of concepts and objects which form the basis of our theoretical analysis of physical phenomena. Aside from the occasional contact with experimental physics and empirical verification, theoretical physicists are restricted to the use of this mathematical language. In the foundations of physics, it is not sufficient to contain our analysis within a single mathematical theory of physics. Therefore, a broad knowledge and precise understanding of the mathematical objects used to define physical concepts and processes is essential.

For example, when we discuss physical ideas we give names to mathematical objects such as: elements of Hilbert spaces, completely positive trace preserving maps and Lorentzian Manifolds, calling them *quantum states*, *quantum processes*, *space-time*. What exactly makes us choose these abstract mathematical objects rather than others? When we develop physical theories and make discoveries about these theories (and their components), are we discovering something beyond the features of the mathematical language we are using as an arena for progress? Could our story be told without these complex and abstract protagonists or is physics restricted to defining arbitrary objects and proving mathematical theorems about these? We are reminded of Shakespeare's crocodile in Antony and Cleopatra [271]:

LEPIDUS: What manner o' thing is your crocodile?

ANTONY: It is shaped, sir, like itself, and it is as broad as it hath breadth. It is just so high as it is, and moves with it own organs. It lives by that which nourisheth it, and the elements once out of it, it transmigrates.

LEPIDUS: What colour is it of?

ANTONY: Of its own colour, too.

In any case, we will spend some time introducing mathematics, aiming to ensure that there is sufficient clarity in our discourse of Physics.

2.1 Set theory

2.1.1 Axiomatic set theory

We will present the Zermelo-Fraenkel axioms [302,293] which define a collection of objects called a **set**. We introduce a membership property ($\in$) such that $X \in Y$ means that X is an element of Y.

Axiom 1: (Existence) There exists a set which has no elements.

Axiom 2: (Extensionality) If every element of X is an element of Y and every element of Y is an element of X, then X=Y.

Lemma 1.1: There is a unique set $\emptyset$ with no elements, called the **empty set**.

Proof: Suppose that there exist two sets X and Y which both have no elements. If a $\in$ X then a $\in$ Y and if b $\in$ Y then b $\in$ X therefore, by Axiom 2, X=Y.

<u>Axiom 3:</u> (Comprehension) Let P(x) be a property of set x. For any A, there exists B such that x $\in$ B if and only if x $\in$ A and P(x) holds.

<u>Axiom 4:</u> (Pairing) For any sets A and B, there exists a set C such that $x \in C$ if and only x=A and x=B.

<u>Definition 1.1:</u> The set having exactly A and B as its elements is written $\{A, B\}$ and is called the unordered pair of A and B.

<u>Axiom 5:</u> (Union) For any set A, there exists a set U such that $x \in U$ if and only if $x \in X$ for some $X \in A$.

<u>Definition 1.2:</u> The **union** of A and B, written $A \bigcup B$ is the set of all elements in either A, B or both. The existence of $A \bigcup B$ follows from applying Axiom 5 to the pairing $\{A, B\}$ obtained from Axiom 4.

<u>Definition 1.3:</u> The **intersection** of A and B, written $A \bigcap B$ is the set of all elements in both A and B. The existence of $A \bigcap B = \{x \in A | x \in B\}$ follows from applying Axiom 3 to the set A and the property $\{P(x) : x \in B\}$.

<u>Definition 1.4:</u> A is called a subset of B, written as $A \subseteq B$, if every element of A is an element of B.

<u>Axiom 6:</u> (Power set) For any set A, there exists a set P such that $X \in P$ if and only if $X \subseteq A$. P is called the **power set** of A.

<u>Axiom 7:</u> (Schema of Replacement) Let P(x,y) be a property such that for every x there is a unique y for which P(x,y) holds. For every set A there exists a set B such that for every x $\in$ A there is y $\in$ B for which P(x,y) holds.

<u>Axiom 8:</u> (Infinity) There is an inductive set I, defined such that the empty set is in I and if $x \in I$, then the set formed by taking the union $x \bigcup \{x\}$ is in I, where $\{x\}$ is the singleton set, with exactly one element (x).

These eight Zermelo-Fraenkel axioms can be strengthened by adding the **Axiom of Choice**, which we state below.

<u>Axiom 9:</u> (Choice) Let A be a set whose members are all non-empty. Then there exists a function f from A to the union of the members of A (see below for the definition of a function), called a **choice function**, such that for all B $\in$ A, one has f(B) $\in$ B.

2.1.2 Relations and functions

We can then introduce maps between sets.

Definition 1.5: An ordered pair (x,y) is defined to be $\{\{x\},\{x,y\}\}$.

Definition 1.6: The Cartesian product of sets X and Y is defined as:

$$X \times Y = \{(x,y)|x \in X; y \in Y\}.$$

Definition 1.7: A set R is called a **relation** if all elements of R are ordered pairs. We denote (x,y) $\in$ R as xRy.

Definition 1.8: Let R be a relation on a set X.

(i) R is reflexive if xRx, for all $x \in X$.

(ii) R is symmetric if xRy implies yRx, for all $x, y \in X$.

(iii) R is antisymmetric if xRy and yRx imply x=y, for all $x, y \in X$.

(iv)R is transitive if xRy and yRz imply xRz for all $x, y, z \in X$.

(v) R is an **equivalence relation** (on X) if it is reflexive, symmetric and transitive.

(vi) R is an ordering of X if it is reflexive, antisymmetric and transitive.

Definition 1.9: Given an equivalence relation R on a set X we can define the equivalence class [a] of an element $a \in X$ as: $[a] = \{x \in X|aRx\}$. The set of all equivalence classes in X with respect to an equivalence relation R is denoted as X/R and is called the **quotient set** of X by R.

Definition 1.10: A relation F is called a **function** if $xFy_1 = xFy_2$ implies that $y_1 = y_2$, for all x, y_1, y_2. We write this unique $y_1 = y_2$ as F(x).

Definition 1.11: Let f: $X \to Y$ be a function.

(i) f is **injective** if for all $x_1, x_2 \in X$, $f(x_1) = f(x_2)$ if and only if $x_1 = x_2$.

(ii) f is **surjective** if for every $y \in Y$, there is an $x \in X$ such that f(x)=y.

(iii) f is **bijective** if it is both injective and surjective.

Definition 1.12: A binary operation on a non-empty set X is a mapping b:X $\times$ X $\to$ X which is defined for every pair of elements in X and which uniquely assigns an element of X to each pair of elements in X.

2.1.3 Numbers

We can construct the set of **natural numbers** $\mathbb{N}$ by using Peano's axioms[293]:

(i) There exists a distinguished element $0 \in \mathbb{N}$.

(ii) There exists an equivalence relation = on $\mathbb{N}$ (called equality) such that $\mathbb{N}$ is closed under equality (if $a \in \mathbb{N}$ and a=b then $b \in \mathbb{N}$).

(iii) There exists an injective successor function S: $\mathbb{N} \rightarrow \mathbb{N}$.

(iv) There does not exist $n \in \mathbb{N}$ such that S(n)=0 (so S is not surjective).

(v) Let K be a subset of $\mathbb{N}$ such that: $0 \in K$ and if $k \in K$ then $S(k) \in K$, $\forall k \in \mathbb{N}$. Then $K = \mathbb{N}$ (principle of induction).

We can then recursively define addition + and multiplication $\cdot$ operations:

$$\mathrm{a} + 0 = \mathrm{a} \text{ ; } \mathrm{a} + \mathrm{S(b)} = \mathrm{S(a+b)} \text{ (add)}$$
$$\mathrm{a} \cdot 0 = 0 \text{ ; } \mathrm{a} \cdot \mathrm{S(b)} = \mathrm{a} + (\mathrm{a} \cdot \mathrm{b}) \text{ (mult)}$$

$\mathbb{N}$ admits an ordering relation: $\leq$ defined as $\forall x, y \in \mathbb{N}$, $x \leq y$ if and only if there $\exists c \in \mathbb{N}$ such that x + c = y.

We can construct **integers** as equivalence classes of ordered pairs of natural numbers.

Definition 1.13: For ordered pairs of natural numbers (a, b), (c,d) $\in \mathbb{N} \times \mathbb{N}$, we define a relation $\equiv_Z$:

$$(a, b) \equiv_Z (c, d) \text{ iff } a + d = b + c \tag{2.1}$$

Note that $\equiv_Z$ is an equivalence relation.

Definition 1.14: We define the set of integers as the quotient set: $\mathbb{Z} := \mathbb{N} \times \mathbb{N} / \equiv_Z$.

Proposition 1.1: Every equivalence class [(x,y)] (with $x, y \in \mathbb{N}$) can be written as either [(0,n)] or [(n,0)] for some $n \in \mathbb{N}$.

Proof: Let $x, y \in \mathbb{N}$. Recall that if $x \leq y$ then $\exists n \in \mathbb{N}$ such that x+n = y+0 and [(x,y)]=[(0,n)]. If $y < x$ then $\exists n \in \mathbb{N}$ such that y+n = x+0 and [(x,y)]=[(n,0)]. Therefore, every equivalence class [(x,y)] contains an ordered pair with at least one zero coordinate.

For each natural number $n \in \mathbb{N}$, we write the integer [(0,n)] as -n and the integer [(n,0)] as n.

We can define the addition + and multiplication $\cdot$ operations in $\mathbb{Z}$ as:

$$[(a, b)] + [(c, d)] := [(a + c, b + d)]$$

$$[(a,b)] \cdot [(c,d)] := [(a \cdot c + b \cdot d, b \cdot c + a \cdot d)]$$

where the operations within brackets are natural number addition and multiplication as defined earlier. We can also define subtraction by adding the additive inverse, meaning that the difference between integers $a, b \in \mathbb{Z}$ is defined as: $a - b := a + (-b)$.

Note that addition and multiplication are commutative both in $\mathbb{N}$ and in $\mathbb{Z}$.

We can similarly construct **rational numbers** as equivalence classes of ordered pairs of integers.

Definition 1.15: For ordered pairs of integers (a, b), (c,d) $\in \mathbb{Z} \times \mathbb{Z}$, we define a relation $\equiv_Q$:

$$(a,b) \equiv_Q (c,d) \text{ iff } a \cdot d - c \cdot b = 0 \tag{2.2}$$

Note that $\equiv_Q$ is an equivalence relation.

Definition 1.16:: We define the set of rational numbers as the quotient set:

$$\mathbb{Q} := (\mathbb{Z} \times (\mathbb{Z} - \{0\}))/ \equiv_Q \tag{2.3}$$

Once again, we can define addition + an multiplication $\cdot$ operations:

$$[(a,b)] + [(c,d)] := [(a \cdot d + b \cdot c, b \cdot d)]$$
$$[(a,b)] \cdot [(c,d)] := [(a \cdot c, b \cdot d)]$$

We can interpret rational numbers as quotients of two integers and [(x,y)] can be chosen so that y is positive and gcd(x,y) = 1, meaning that x and y share no common factors.

The construction of **real numbers** is more delicate. We will present the Dedekind cut method, where real numbers are subsets of $\mathbb{Q}$ called cuts.

Definition 1.17: A non empty subset $C \subseteq \mathbb{Q}$ is called a cut if:

(i) $C \neq \mathbb{Q}$

(ii) If $x \in C$, $y \in \mathbb{Q}$ and y < x then $y \in C$.

(iii) If $x \in C$, then x < r for some $r \in C$.

Definition 1.18: The set of real numbers $\mathbb{R}$ is the collection of all cuts of $\mathbb{Q}$.

If $a, b \in \mathbb{R}$ then we can define addition and multiplication as the cuts:

$$a + b := \{x + y | x \in a, y \in b\}$$

$$ab := [(q \in \mathbb{Q}|q < xy \text{ for some } x \in a, y \in b, x > 0, y > 0)]$$

We can also define the cuts:

$$0 := \{p \in \mathbb{Q}|p < 0\}$$
$$1 := \{q \in \mathbb{Q}|q < 1\}$$

such that they play the role of the additive and multiplicative identity respectively. There is an ordering relation on the real numbers such that $a \leq b$ iff $a \subseteq b$. Note that the rational numbers can be embedded in the real numbers.

The set of **complex numbers** $\mathbb{C}$ can then be defined as the set of ordered pairs of real numbers $a + bi := (a, b)$. Addition and multiplication are defined as the operations:

$$(a, b) + (c, d) := (a + c, b + d)$$
$$(a, b) \cdot (c, d) := (ac - bd, bc + ad)$$

Note that all the sets of numbers we have constructed are fields (in the algebraic sense defined below) with respect to the addition and multiplication operations we introduced. One can also define hypercomplex numbers[231], which have interesting algebraic properties and no longer admit a field structure.

2.2 Group Theory

Definition 2.1: A **group** is a set G together with a binary operation $\circ$ on G. The group operation is associative, meaning that we have:

$$\forall x, y, z \in G,\ (x \circ y) \circ z = x \circ (y \circ z) \tag{2.4}$$

There exists an identity element $e \in G$ such that:

$$\forall x \in G,\ x \circ e = e = e \circ x \tag{2.5}$$

$$\forall x \in G, \exists x^{-1} \in G,\ x \circ x^{-1} = e = x^{-1} \circ x \tag{2.6}$$

A simple proof by contradiction shows that the identity element e is unique and that for each $x \in G$ there is a unique inverse $x^{-1} \in G$ satisfying $x \circ x^{-1} = e = x^{-1} \circ x$. A group is called **Abelian** if the operation is commutative, meaning that:

$$\forall x, y \in G,\ x \circ y = y \circ x \tag{2.7}$$

Definition 2.2: A **subgroup** H of a group G is a subset $H \subseteq G$ which is closed under the group operation $\circ$ in G and which forms a group with respect to $\circ$. We write $H \leq G$.

Definition 2.3: let H be a subgroup of a group G then the sets:

$$Hg := \{h \circ g : h \in H\} \tag{2.8}$$

$$gH := \{g \circ h : h \in H\} \tag{2.9}$$

where $g \in G$ are respectively the right and left **coset** of H (determined by g).

Lagrange's Theorem: Let $H \leq G$. Then $|G| = |H||G : H|$, where $|G|$ is the order of the group G (number of elements in its set) and $|G : H|$ is the index of H in G (number of distinct left cosets of H in G).

Definition 2.4: A normal subgroup $N \leq G$ is a subgroup which is invariant under conjugation by members of G. This means that N is a normal subgroup of G if and only if gN = Ng $\forall g \in G$. We write $N \lhd G$.

Definition 2.5: Let $H \lhd G$ then we can define the **quotient group** G/H by taking the set of all left cosets of H in G. The associative group operation consists of taking the product of $aH \in G/H$ and $bH \in G/H$ to be $(aH)(bH) = (a \circ b)H \in G/H$ (since H is a normal subgroup). The group identity element is eH=H and the inverse element of aH is $a^{-1}H$ $\forall aH \in G/H$.

Definition 2.6: Given a subset $X \subseteq G$, we define the subgroup $\langle X \rangle$ **generated** by X as the smallest subgroup of G containing X. This means that: $X \subseteq \langle X \rangle$ and if $X \subseteq H$, where $H \leq G$, then $\langle X \rangle \subseteq H$.

A group is finitely generated if it is generated by a finite number of its elements.

Definition 2.7: A group is **cyclic** if it is generated by one of its elements, meaning that

$\exists g \in G$ such that $\langle g \rangle = G$.

Example 2.1: Group $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$ of integers with addition modulo n. This is an Abelian cyclic group of order n with identity element 0.

Taking $\mathbb{Z}_4 = \{0, 1, 2, 3\}$ and the (normal) subgroup $H = \{0, 2\}$ then the cosets of H are $\{0, 2\}$ and $\{1, 3\}$ and we can form the quotient group $\mathbb{Z}_4/H$ which is isomorphic to $\mathbb{Z}_2$.

Note that $\mathbb{Z}_p \times \mathbb{Z}_q$ is isomorphic to $\mathbb{Z}_{pq}$ iff gcd(p,q)=1.

Definition 2.8: A group **homomorphism** is a map $\theta : G \to H$ from group G to group H such that: $\theta(g_1 \circ g_2) = \theta(g_1) \star \theta(g_2)$, $\forall g_1, g_2 \in G$ where $\circ$ and $\star$ are the group operations of G and H respectively. If θ is also a bijective map, then we say that it is an **isomorphism** and that G and H are isomorphic.

Example 2.2: General linear group $GL_n(K)$ of non-singular $n \times n$ matrices over a field F, with matrix multiplication. The determinant, which satisfies $det(M_1) \in K \setminus \{0\}$ (due to non-singularity) and $det(M_1 M_2) = det(M_1)det(M_2)$, $\forall M_1, M_2 \in GL_n(K)$, is a homomorphism from $GL_n(K)$ to the group corresponding to $K \setminus \{0\}$ with respect to multiplication.

The isomorphism theorems make explicit the relationship between quotients, homomorphisms, and sub-objects in a general algebraic sense[173,122]. We will only state the first Isomorphism theorem for groups without proof.

First Isomorphism Theorem: Let $\theta : G \to H$ be a group homomorphism then:

(i) The **kernel** of θ: $ker(\theta) := \{g \in G : \theta(g) = e\}$ is a normal subgroup of G.

(ii) The image of θ: $im(\theta) := \{h \in H : h = \theta(g),\ g \in G\ \}$ is a subgroup of H.

(iii) $im(\theta)$ is isomorphic to G/ $ker(\theta)$.

Lemma 2.1: Every subgroup of a cyclic group is cyclic.

Proof: Let $H \leq G$ and G be a cyclic group. We assume that $H \neq \{e\}$ (otherwise the result is trivial) so that $\exists k \in \mathbb{Z}$ such that $a^k \in H$. By the Euclid division algorithm[73] we can find $q, r \in \mathbb{Z}$ with $0 \leq r < m$ such that $k = mq + r$, where m is the smallest positive integer with $a^m \in H$. This gives us: $a^r = (a^m)^{-q} a^k$ which must be in H by group closure. But m is the smallest positive integer with $a^m \in H$ and $0 \leq r < m$ so we must have r=0 and k=qm. Therefore H is generated by a^m.

Corollary 2.1: The only cyclic groups (up to isomorphism) are the groups $\mathbb{Z}$ and $\mathbb{Z}_n := \mathbb{Z}/n\mathbb{Z}$ with respect to (modular) addition.

Proof: Let G $= \langle g \rangle$ be cyclic and consider the map $\theta : \mathbb{Z} \to G, \theta(n) = g^n$. θ is a homomorphism since $\theta(n+m) = g^n g^m$. Therefore, by the First Isomorphism theorem, G is isomorphic to $\mathbb{Z}/ker(\theta)$ where $ker(\theta) \triangleleft \mathbb{Z}$. The result then follows from the previous lemma showing that the subgroup of a cyclic group is cyclic.

Definition 2.9: A group G is called the direct sum of a finite set of subgroups $\{H_i\}$, where i= 1, ..., n, if:

(i) Each H_i is a normal subgroup of G.

(ii) Each H_i has a trivial intersection with $\langle H_j | j \neq i \rangle$, $\forall$ i= 1, ..., n.

(iii) G is generated by the subgroups $\{H_i\}$.

We write: $G = H_1 \oplus ... \oplus H_n$.

Classification theorem of finitely generated abelian groups: Every finitely generated Abelian group G is isomorphic to a direct sum of primary cyclic groups (whose order is a power of a prime) and infinite cyclic groups, such that:

$$G = \mathbb{Z}^n \oplus \mathbb{Z}_{p_1} \oplus ... \oplus \mathbb{Z}_{p_n} \tag{2.10}$$

where $\mathbb{Z}$ is the group of integers and $p_1, ..., p_n$ are powers of prime numbers.

We will now discuss mechanisms for analyzing groups by mapping them back to other mathematical objects, studying the action on a set or the representation in terms of matrices and linear maps.

Definition 2.10: Let G be a group and Ω a non-empty set. Let there be a unique element $\omega \cdot g \in \Omega$ for each $\omega \in \Omega$ and $g \in G$. We say that G acts on Ω if:

(i) $\omega \cdot e = \omega, \ \forall \omega \in \Omega$, where e is the identity of G

(ii) $(\omega \cdot g) \cdot h = \omega \cdot (g \circ h), \ \forall \omega \in \Omega, \forall g, h \in G$

Definition 2.11: The subgroup $G_\omega := \{g \in G : \omega \cdot g = \omega\}$ of G is called the stabilizer of ω in G.

Definition 2.12: A **representation** of group G on a vector space V over a field F (see the next section) is a group homomorphism $\rho : G \to GL(V)$, from G to the general linear group GL(V) (of automorphisms of V). V is called the representation space.

This means that we can reduce group theoretic problems to linear algebra by representing group elements as matrices (choosing a basis) and the group operation by matrix multiplication. The essential information about a group representation can be expressed in a more condensed form by studying its character.

Definition 2.13: Let $\rho : G \to GL(V)$ be group representation of a group G on a vector space V (over a field F). The **character** of ρ is a map $\chi_\rho : G \to F$, such that $\chi_\rho(g) = Tr(\rho(G))$, where $Tr(\rho(G))$ is the trace of the linear transformation representing G.

Definition 2.14: A permutation on a set Ω is a bijective function $f_\Omega : \Omega \to \Omega$. The set of permutations of a set can be shown to form a group under function composition, which we call the symmetric group on Ω, written Sym(Ω). A **permutation group** is a subgroup of Sym(Ω) for some set Ω. The set of even permutations in Sym(Ω) forms a normal subgroup Alt(Ω) which we call the alternating group on Ω.

Cayley's theorem: Every group is isomorphic to a permutation group.

Proof: Take a group G and a homomorphism θ from G to the group of permutations of the underlying set G (a permutation representation). Since G= G/$\{e\}$, $g \cdot e$ is a group action on G. Now let $k \in ker(\theta)$, then $k = k \cdot e = \theta(k) \cdot e = e$ so $ker(\theta)$ is trivial and θ is injective. The result then follows from the First Isomorphism Theorem.

Definition 2.15: A group G is **solvable** if there is a finite collection of normal subgroups $G_1, ..., G_n$ such that: $1 = G_1 \subseteq G_2 \subseteq ... \subseteq G_n = G$ and G_{j+1}/G_j is abelian for $1 \leq j < n$.

Feit-Thompson Theorem: Every finite group of odd order is solvable.

Definition 2.16: A group is **simple** if its only normal subgroups are the trivial group and itself.

Definition 2.17: A **composition series** of a group G is a finite sequence of normal subgroups:

$$1 = H_0 \triangleleft H_1 \triangleleft ... \triangleleft H_n = G \tag{2.11}$$

where each H_j is a maximal strict normal subgroup of H_{j+1}. The quotient groups H_{j+1}/H_j are called the composition factors and are simple groups. The length n of the series is called the composition length.

Jordan Hölder Theorem: Up to permutation and isomorphism, any two composition series of a given group are equivalent, meaning that they have the same composition length and the same composition factors.

This theorem shows that finite simple groups are the basic building blocks of all finite groups. We will conclude this section by briefly stating an impressive result: the classification theorem for finite simple groups[98,153].

Finite simple group theorem: Every finite simple group is isomorphic to one of the following groups:

(i) Cyclic groups of finite order.

(ii) Alternating groups of order more than 5.

(iii) A simple group of Lie type[70].

(iv) One of 26 sporadic simple groups[153].

2.3 Algebraic structures

Having introduced group theory, we will now proceed by studying other algebraic structures.

2.3.1 Rings, Fields and Galois theory

Adding another binary operation to a group leads to the definition of a ring.

Definition 3.1: Let R be a set with two operations denoted by + and juxtaposition, then R is a **ring** (with unit) if:

(i) R is an abelian group with respect to +.

(ii) Juxtaposition is associative: (xy)z = x(yz), $\forall x, y, z \in R$.

(iii) Juxtaposition is distributive over +: $x(y + z) = xy + xz$ and $(y + x)z = yz + xz$, $\forall x, y, z \in R$.

(iv) $\exists 1 \in R$ such that $1r = r = r1$, $\forall r \in R$.

Definition 3.2: A map $\theta : R \to S$ is a ring homomorphism if: $\theta(r_1 + r_2) = \theta(r_1) + \theta(r_2)$ and $\theta(r_1 r_2) = \theta(r_1)\theta(r_2)$, $\forall r_1, r_2 \in R$.

Definition 3.5: Let R be a ring (with unit) as defined above. R is called a **field** if it also satisfies:

(i) $1 \neq 0$.

(ii) $\forall a \in R, a \neq 0, \exists b \in R$ such that: ab=ba=1.

(iii) Juxtaposition is commutative.

A ring R which satisfies (i) and (ii) is called a division ring.

Example 3.1: Examples of fields include the natural, rational, real and complex numbers that we introduced previously (with respect to addition and multiplication).

We have that every field is a division ring, but there are division rings that are not fields (e.g. the quaternions); every division ring is a ring with unity, but there are rings with unity that are not division rings (e.g. the integers if you want commutativity, or the n×n matrices with coefficients in R and n>1 if you want non-commutativity); every ring with unity is a ring, but there are rings that are not rings with unity (e.g. strictly upper triangular 3×3 matrices with coefficients in R).

Note also that a **domain**, meaning a ring satisfying ab = 0 implies a = 0 or b = 0, is another object between a ring and a field.

Definition 3.6: Given a field F, the smallest integer k for which adding the multiplicative identity k times gives the additive identity, meaning that 1 + 1 + ... + 1 = 0, is called the **characteristic** of F.

Definition 3.7: A field F is algebraically closed if every non-constant polynomial ring F[X]:= $\{f \in F | f = \sum_{j=0}^{m} f_j X^j\}$ (where $f_j \in F$, $\forall j = 1, ..., m$) has a root in F, meaning that $\exists a \in F$, such that: F[a]=0.

There is an important theorem stating that the field of complex numbers $\mathbb{C}$ is algebraically closed. Note that no finite field is algebraically closed (consider $F[X] = (X - f_1)(X - f_2)(X - f_n) + 1$, where $F := \{f_1, ..., f_n\}$) and that the field of real numbers is not algebraically closed as the polynomial $x^2 + 1 = 0$ has no solution in $\mathbb{R}$.

Contrary to group theory, where the main objects of study are subgroups of a given group, field theory is mainly concerned with the analysis of field extensions containing a given field.

Definition 3.8: A field K is a **field extension** of a field F (written K/F) if F is a subfield of K, meaning that F is a subset of K which satisfies the field axioms with the same operations as in K. The dimension of K as a vector space over F is the field extension degree of K/F and is written $|K : F|$.

Definition 3.9: Let F be a field and $f(X) \in F[X]$ be a polynomial of degree $n > 0$. A field extension K of F is called a **splitting field** for f(X) over F if the polynomial f(X) decomposes into linear factors: $f(X) = a\Pi_{i=1}^{n}(X - \alpha_i)$, where $a, \alpha_i \in K$. A field extension K/F is said to be **normal** if K is the splitting field of a family of polynomials in F[X].

Definition 3.10: Let F be a field and $f(X) \in F[X]$ be a polynomial of degree $n > 0$. A field extension K of F is called **separable** over F when it is algebraic over F – in the sense that every element of K is a root of some non-zero polynomial with coefficients in F – and its minimal polynomial in F[X] has distinct roots in a splitting field over F, such that each root has multiplicity one.

Definition 3.11: Let K/F be a field extension of finite degree, then it is called a **Galois extension** iff it is both a normal field extension and a separable field extension.

Galois theory allows us to study field extensions by associating a group to each finite field extension. One can then study study polynomials over fields and important problems in field theory by using tools from Group theory.

Definition 3.12: An automorphism of the finite degree field extension K/F is an isomorphism θ from K to K such that $\theta(f) = f, \forall f \in F$. The set of all automorphisms of K/F forms a group with the operation of function composition which is called the **Galois group** (written Gal(K/F)) of K/F.

Consider a polynomial $f(X) \in F[X]$, with coefficients chosen from the field F, and the field K obtained by adjoining the roots of the polynomial $f(X)$ to the field F. Any permutation of the roots – such that any algebraic equation satisfied by the roots is still satisfied after the roots have been permuted – gives rise to an automorphism of K/F, and vice versa. Therefore, we can use the Galois group as a tool to analyze the solutions of polynomials. We can make this precise[208] by noting that subgroups of the Galois group Gal(K/F) exactly correspond with the subfields of K containing F.

Fundamental theorem of Galois theory: Let E be a Galois extension over a field F and G=Gal(E/F) be the Galois group of E/F. Let us denote the collection of intermediate fields K and subgroups of G as: $\mathcal{F} := \{K | F \subseteq K \subseteq E\}$ and $\mathcal{G} := \{H | H \leq G\}$ respectively.
Consider the map ϕ: $\mathcal{G} \to \mathcal{F}$ such that $\phi(\cdot) = Gal(E/\cdot)$. The map ϕ is a bijection which reverses containments. If ϕ(K)=H then: $|E : K| = |H|$ and $|K : F| = |G : H|$ and H is a normal subgroup of G iff K is a Galois extension of F, in which case Gal(K/F) is isomorphic to G/H.

Example 3.2: We will briefly illustrate[97] how Galois theory works through a simple example. Consider the field extension $\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q}$ of degree 4. We then have F=$\mathbb{Q}$ and K=$\mathbb{Q}(\sqrt{2}, \sqrt{3})$, whose elements can be written as:

$$(q_1 + q_2\sqrt{2}) + (q_3 + q_4\sqrt{2})\sqrt{3} \tag{2.12}$$

where $q_i \in \mathbb{Q}$, for i=1,2,3,4.
The Galois group G = Gal($\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q}$) can be determined by examining the automorphisms of K which must send $\sqrt{2}$ to either $\sqrt{2}$ or -$\sqrt{2}$, and must send $\sqrt{3}$ to either $\sqrt{3}$ or -$\sqrt{3}$, since the permutations in a Galois group can only permute the roots of an irreducible polynomial. Therefore, we can determine that the Galois group G is isomorphic to the Klein-four group $G \cong \langle \alpha, \beta | \alpha^2 = \beta^2 = (\alpha\beta)^2 = 1 \rangle$, where α exchanges $\sqrt{2}$ and $-\sqrt{2}$, β exchanges $\sqrt{3}$ and $-\sqrt{3}$, and these are automorphisms of K .

The mapping in the fundamental theorem of Galois theory gives rise to the following correspondences:
The trivial subgroup $1 \in G$ maps to K=$\mathbb{Q}(\sqrt{2}, \sqrt{3})$.
The subgroup $\{1, \alpha\}$ maps to the subfield $\mathbb{Q}(\sqrt{3})$.
The subgroup $\{1, \beta\}$ maps to the subfield $\mathbb{Q}(\sqrt{2})$.
The subgroup $\{1, \alpha\beta\}$ maps to the subfield $\mathbb{Q}(\sqrt{6})$.
The entire group G maps to F= $\mathbb{Q}$.

An important application of the fundamental theorem is to show that the general quintic equation is not solvable in $\mathbb{Q}$ (the Abel-Ruffini theorem[1]). This can be done by finding the Galois groups of radical extensions, using the fundamental theorem to show that solvable extensions correspond to solvable groups and then using the result that the symmetric group S_5 is not solvable[208]. We will conclude our discussion of Galois theory with the following theorem.

Kronecker-Weber theorem: Every finite Abelian extension of the rational numbers $\mathbb{Q}$ (which has an Abelian Galois group) is a subfield of a cyclotomic field, meaning a number field obtained by adjoining a complex primitive root of unity to $\mathbb{Q}$.

<u>Definition 3.13:</u> Let F be a field and let V be a set with two operations + and $\cdot$. Then V is a vector space over the field F if:

(i) $v_1 + v_2 = v_2 + v_1$ (commutativity)

(ii) $v_1 + (v_2 + v_3) = (v_1 + v_2) + v_3$ (associativity)

(iii) $\exists 0 \in V$ such that: $0 + v_1 = v_1 = v_1 + 0$ (identity element)

(iv) $\exists(-v_1) \in V$ such that: $v_1 + (-v_1) = (-v_1) + v_1 = 0$ (inverse element)

(v) $f_1 \cdot (v_1 + v_2) = f_1 \cdot v_1 + f_1 \cdot v_2$

(vi) $(f_1 + f_2) \cdot v_1 = f_1 \cdot v_1 + f_2 \cdot v_1$

(vii) $(f_1 \cdot (f_2 \cdot v_1)) = (f_1 f_2) \cdot v_1$

(viii) $v_1 + v_2 \in V$ and $f_1 \cdot v_1 \in V$

$\forall v_1, v_2, v_3 \in V, \forall f_1, f_2 \in F$.

Note that we will usually write the scalar multiplication as juxtaposition, omitting '$\cdot$'.

<u>Definition 3.14:</u> An **algebra** over a field F is a ring A which is a vector space over F.

<u>Definition 3.15:</u> Let A and B be algebras over F then $\theta : A \to B$ is an algebra homomorphism if:

(i) $\theta(1_A) = 1_B$

(ii) $\theta(a_1 a_2) = \theta(a_1)\theta(a_2)$

(iii) $\theta(f a_1 + a_2) = f\theta(a_1) + \theta(a_2)$

$\forall f \in F$, $\forall a_1, a_2 \in A$.

We will now define two important examples of algebras.

<u>Definition 3.16:</u> A **Lie Algebra** is a vector space $\mathfrak{g}$ over a field F together with a map $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ such that:

(i) $[a_1, a_2] = -[a_2, a_1]$

(ii) $[a_1, [a_2, a_3]] + [a_2, [a_3, a_1]] + [a_3, [a_1, a_2]] = 0$

$\forall a_1, a_2, a_3 \in \mathfrak{g}$.

<u>Definition 3.17:</u> A **Boolean Algebra** consists of a set B together with two binary operations $\wedge$ (AND) and $\vee$ (OR) and a complementation operation $(\cdot)^c : B \to B$ such that:

(i) $b_1 \wedge b_1 = b_1 \vee b_1 = b_1$ and $b_1 \vee (b_1 \wedge b_2) = b_1 \wedge (b_1 \vee b_2) = b_1$

(ii) $b_1 \wedge b_2 = b_2 \wedge b_1$ and $b_1 \vee b_2 = b_2 \vee b_1$

(iii) $b_1 \vee (b_2 \vee b_3) = (b_1 \vee b_2) \vee b_3$ and $b_1 \wedge (b_2 \wedge b_3) = (b_1 \wedge b_2) \wedge b_3$

(iv) $b_1 \vee (b_2 \wedge b_3) = (b_1 \vee b_2) \wedge (b_1 \vee b_3)$ and $b_1 \wedge (b_2 \vee b_3) = (b_1 \wedge b_2) \vee (b_1 \wedge b_3)$

(v) $b_1 \wedge (b_1)^c = \emptyset$ and $b_1 \vee (b_1)^c = (\emptyset)^c$

(vi) $\emptyset \vee b_1 = b_1$, $\emptyset \wedge b_1 = \emptyset$, $(\emptyset)^c \vee b_1 = (\emptyset)^c$ and $(\emptyset)^c \wedge b_1 = b_1$

for all $b_1, b_2, b_3 \in B$ and where $\emptyset$ is the empty set.

We conclude this section by noting that the Stone representation theorem[280] shows that every Boolean algebra is isomorphic to a pair { X, F }, where X is a set and F is a non-empty subset of the power set of X, closed under the intersection and union of pairs of sets and under complements of individual sets.

2.3.2 Linear Algebra and Graph theory

Given the direct relevance for quantum theory, we will introduce Linear Algebra in some detail.

Definition 3.22: Let V be a vector space over a field F. A subset L $\subset$ V is called **linearly independent** if, whenever $f_1, ..., f_n \in F$ and $l_1, ..., l_n \in L$, we have:
$f_1 l_1 + ... + f_n l_n = 0$ implies $f_1 = f_2 = ... = f_n = 0$.

Definition 3.23: A subset S $\subset$ V is said to **span** V if $\forall v \in V$, $\exists f_1, ..., f_n \in F$ and $\exists s_1, ..., s_n \in S$ such that: $v = f_1 s_1 + ... + f_n s_n$.

Definition 3.24: A subset B $\subset$ V is called a basis of V if it spans V and is linearly independent. The size of B is the dimension of V, written as dim(V).

Definition 3.25: Let V and W be vector spaces over a field F. A map T: V $\rightarrow$ W is a linear transformation if: $T(f_1 v_1 + f_2 v_2) = f_1 T(v_1) + f_2 T(v_2)$, $\forall f_1, f_2 \in F$, $\forall v_1, v_2 \in V$.

Theorem 3.1: Let V and W be vector spaces over a field F such that dim(V)=m and dim(W)=n. The set Hom(V,W) of linear transformations from V to W is isomorphic to the space of n $\times m$ matrices over F.

Proof: Let $B_V = \{e_1, ..., e_m\}$ and $B_W = \{e'_1, ..., e'_n\}$ be bases for V and W respectively. Let $[T]_{B_V,B_W}$ be an $n \times m$ matrix with i, j entry t_{ij} such that: $T(e_j) = \sum_{k=1}^{n} t_{kj} e'_k$. It is easy to check that the assignment $T \mapsto [T]_{B_V,B_W}$, which maps composition of linear transformations to multiplication of matrices, is an isomorphism of vector spaces from Hom(V,W) to the space of n $\times m$ matrices over F.

Note that there are isomorphism theorems for vector spaces, which are directly analogous to those for groups and rings[122]. The following is a corollary of the first isomorphism theorem for vector spaces:

Rank-Nullity theorem: Let T: V $\rightarrow$ W be a linear transformation and dim(V) be finite then:

$$dim(V) = dim(ker(T)) + dim(im(T)) \tag{2.13}$$

Note that dim(ker(T)) and dim(im(T)) are respectively called the nullity and the rank of the matrix corresponding to T.

Definition 3.26: The **determinant** of a square matrix M with entries m_{kl} is defined as:

$$det(M) := \sum_{i_1,...,i_n=1}^{n} \epsilon_{i_1,...,i_n} m_{1i_1}...m_{ni_n} \tag{2.14}$$

$\epsilon_{i_1,...,i_n}$ is the Levi-Civita symbol, which is equal to 1 if $(i_1, ..., i_n)$ is an even permutation of (1, 2, ..., n), equal to -1 if $(i_1, ..., i_n)$ is an odd permutation of (1, 2, ..., n) and equal to 0 otherwise.

Note that the determinant maps matrices to scalars and that:

$$det(A^{-1}) = det(A)^{-1} \text{ and } det(AB) = det(A)det(B)$$

Definition 3.27: The characteristic polynomial of a square matrix M is:

$$\chi_M(\lambda) := det(M - \lambda I) \tag{2.15}$$

where I is the identity matrix. **Eigenvalues** of a matrix M are defined as the roots λ of the characteristic polynomial. Therefore, λ is an eigenvalue of M iff $\exists$ $\mathbf{v} \in F^n$ such that:

$$M\mathbf{v} = \lambda\mathbf{v} \tag{2.16}$$

Vectors $\mathbf{v} \in F^n$ satisfying equation (2.16) are called **eigenvectors** of V (corresponding to eigenvalue λ).

Definition 3.28: Given a vector space V over a field F, its **dual space** $V^\star$ is a vector space of linear transformations from V to F. The elements of $V^\star$ are called linear functionals. Note that if $B = \{e_1, ..., e_n\}$ is a basis for V then $B^\star = \{e_1^\star, ..., e_n^\star\}$, where $e_i^\star(e_j) = \delta_{ij}$ (where δ_{ij} is the Kronecker delta).

Definition 3.29: Given a vector space V over a field F, a **bilinear form** on V is a function $b : V \times V \to F$ satisfying:

(i) $b(fv_1, v_2) = b(v_1, fv_2) = fb(v_1, v_2)$

(ii) $b(v_1 + v_2, v_3) = b(v_1, v_3) + b(v_2, v_3)$

(iii) $b(v_1, v_2 + v_3) = b(v_1, v_2) + b(v_1, v_3)$

$\forall v_1, v_2, v_3 \in V$, $\forall f \in F$. A map $b : V \times V \to F$, where $F = \mathbb{C}$ is sesquilinear form on V if (i) above is replaced by:

(i') $b(\bar{f}v_1, v_2) = b(v_1, fv_2) = fb(v_1, v_2)$

where $\bar{f}$ is the complex conjugate of f.

Definition 3.30: An **inner product space** is a vector space V over the field $\mathbb{C}$, equipped with an **inner product** $\langle ., . \rangle : V \times V \to F$, which is a sesquilinear form that satisfies:

(i) $\langle v_1, v_2 \rangle = \overline{\langle v_2, v_1 \rangle}$, $\forall v_1, v_2 \in V$ (conjugate symmetric).

(ii) $\langle v, v \rangle$ is positive $\forall v \neq 0$ in V (positive definite).

When we have an inner product space V over $\mathbb{C}$, we can define a dual space consisting of linear functionals:

$$\langle v, . \rangle : V \to \mathbb{C}, \text{ such that } w \mapsto \langle v, w \rangle \tag{2.17}$$

Definition 3.31: Let $B = \{e_1, ..., e_n\}$ be the basis for an inner product space V. B is called orthonormal if $\langle e_i, e_j \rangle = \delta_{ij}, \forall i, j = 1, ..., n$. Note that one can always find an orthonormal basis for an inner product space by using the Gram-Schmidt process[16].

Definition 3.32: The general linear group GL(V) of a vector space V over a field F is defined as the group of all automorphisms of V, meaning the set of all bijective linear transformations V $\to$ V, together with composition as the group operation.

Definition 3.33: Given an inner product space V and a linear transformation $T : V \to V$, we can define the following linear transformations:

(i) The inverse transformation $T^{-1} : V \to V$ satisfying $T^{-1}(T(v)) = T(T^{-1}(v)) = v, \forall v \in V$.

(ii) The adjoint transformation $T^\star : V \to V$ satisfying: $\langle v_1, T(v_2) \rangle = \langle T^\star(v_1), v_2 \rangle$.

(iii) The transpose transformation $T^t : V \to V$ defined as the complex conjugate of the adjoint: $T^t := \bar{T}^\star$.

Given a choice of basis for V, the matrix equivalents of these concepts yield the familiar notions of matrix inverse, adjoint and transpose.

Definition 3.34: A linear transformation $T : V \to V$ is called:

(i) Unitary if $T^\star = T^{-1}$

(ii) Orthogonal if $T^t = T^{-1}$

(iii) Normal if $TT^\star = T^\star T$.

The sets of all unitary/ orthogonal square n $\times$ n matrices over $\mathbb{R}/\mathbb{C}$ form a group called the orthogonal/unitary group O_n/U_n. The restriction of these groups to matrices with determinant 1 gives the special orthogonal group SO_n and the special unitary group SU_n.

The normal matrices correspond exactly to unitarily diagonalizable matrices, in the sense that N is normal iff there exists a unitary matrix U such that $D := UNU^{-1}$ is diagonal[225].

We will state without proof the following theorem which plays an important role in quantum theory.

Simultaneous diagonalization theorem: Let S, T: V $\to$ V be normal linear transformations over a finite dimensional inner product space which are commuting, in the sense that: [S, T]:= ST - TS = 0. Then, there exists a basis B whose elements are simultaneously the eigenvectors of S and of T.

We will now introduce a few concepts from elementary Graph Theory, which are closely related to Linear Algebra and will be useful later in the thesis.

Definition 3.35: A **graph** is an ordered pair G=(V,E) consisting of a set V of vertices and a set E of edges, which are two element subsets of V. A graph is *simple* if it has no self-loops (edges connecting a vertex to itself) and one edge at most connecting any two vertices and it is *undirected* if its edges are unordered pair of vertices. A graph is *finite* if V and E are finite and the number of vertices and edges are then respectively called the *order* and the *size* of the graph.

Definition 3.36: A simple undirected graph of order n can be described by a symmetric n $\times$ n matrix A with $A_{ij} = 1$ if there is an edge connecting vertices i and j and $A_{ij} = 0$ otherwise.

Definition 3.37: A **subgraph** of a graph G=(V,E) is a graph whose vertices and edges are subsets of V and E.

Definition 3.38: The **local complementation** of a graph G=(V,E) about the vertex

$v \in V$ sends G to:

$$G \star v := (V, E\Delta\{x, y\} : \{x, v\}, \{y, v\} \in E \wedge x \neq y) \tag{2.18}$$

where $X\Delta Y := (X - Y) \cup (Y - X)$ is the symmetric set difference of sets X and Y.

Definition 3.39: An isomorphism of graphs G and H is a bijective map f between the set of vertices of G and H such that any two vertices v_1 and v_2 of G are adjacent in G iff f(v_1) and f(v_2) are adjacent in H.

We will conclude our brief presentation of graph theory by mentioning that determining whether two finite graphs are isomorphic is an interesting problem in computational complexity theory[261].

2.4 Topology and Hilbert spaces

Following our introduction of algebraic concepts, we will now proceed by introducing mathematical ideas from a more analytic and geometric perspective. We shall first present some fundamental ideas from topology.

2.4.1 Topology

Definition 4.1: Let X be a set and τ be a family of subsets of X. τ is called a **topology** on X if:
(i) The empty set and X are both elements of τ.
(ii) Any union of elements of τ is an element of τ.
(iii) Any finite intersection of elements of τ is an element of τ.

$\{X, \tau\}$ is called a **topological space**. The members of τ are called (τ) *open sets* in X and subsets of X whose set complement is in τ are called *closed sets* in X (relative to τ).

Two simple examples of topologies are the trivial topology, which only includes the empty set and the entire space X, and the discrete topology, which includes all the subsets of X. Every topology is contained in the discrete topology and contains the trivial topology. If there are two topologies τ_1 and τ_2 on X such that $\tau_1 \subset \tau_2$, then each τ_1 open set is a τ_2

open set and we say that τ_1 is coarser than τ_2 (and τ_2 is finer than τ_1).

Definition 4.2: Let (X, τ) be a topological space and p be a point in X. A **neighborhood** of p is a subset V of X that includes an open set U containing p.

Definition 4.3: A point p of a subset A of a topological space (X, τ) is called an interior point iff A is a neighborhood of p. The set of all interior points of A is the interior of A, written A^0. The boundary of a subset A is the set of all point which are interior to neither A nor the complement of A in X.

Theorem 4.1: A set is open iff it contains a neighborhood of all its points.

Proof: If A is open then it trivially contains a neighborhood (A itself) of all its points. Let the set A contain a neighborhood of each of its points. The union U of all open subsets of A is an subset of A. Each member p of A belongs to an open subset of A so each p is in U and therefore A=U and A is open.

An alternative way of defining a topological structure on a set is to use the the **Kuratowski closure axioms**. Let X be a set, $\mathcal{P}(X)$ be its power set (the set of all subsets) and define the map **clo**: $\mathcal{P}(X) \to \mathcal{P}(X)$ by the following axioms:

(i) **clo**($\emptyset$)=$\emptyset$

(ii) For each A, A $\subset$**clo**(A)

(iii) For each A, **clo**(**clo**(A)) = **clo**(A)

(iv) For each A and B, **clo**(A $\cup$ B)= **clo**(A) $\cup$ **clo**(B)

We can then say that a set A is **closed** iff **clo**(A)=A.

One can show that[180] the closure operation we introduced corresponds to the closure of a subset A of a topological space (X, τ), defined as the intersection of all the members of the family of closed sets containing A (i.e. the smallest closed set containing A).

Definition 4.4: A **basis** $\mathcal{B}$ for a topology (X, τ) is a subfamily $\mathcal{B}$ of τ so that for each point p of the space X and each neighborhood U of p, there is a member V of $\mathcal{B}$ such that $p \in V \subset U$. A basis is a collection of open sets such that every open set can be written as a union of its elements.

Example 4.1: The set of real numbers can be given a **standard topology** such that the open sets are the subsets A of the real numbers which are open intervals, meaning that

$a \in A$ iff $\exists x, y \in \mathbb{R}$ such that $x < a < y$. The collection of all open intervals in the real line forms a base for the standard topology on the real line because the intersection of any two open intervals is itself an open interval or empty. The closed sets in the standard topology are the closed intervals B, such that $b \in B$ iff $\exists x, y \in \mathbb{R}$ such that $x \leq b \leq y$. In the standard topology, a set V $\subset \mathbb{R}$ is a neighborhood of a point $p \in \mathbb{R}$ if, for some $\delta > 0$, the open interval from $x - \delta$ to $x + \delta$ is contained in V. The boundary of an interval is the set whose only members are the endpoints of the interval.

Definition 4.5: A **connected** space is a topological space that cannot be represented as the union of two or more disjoint non-empty open subsets.

Definition 4.6: A **compact** space is a topological space where each open cover – defined as an arbitrary collection of open subsets of X: $\{U_j\}_{j \in J}$ satisfying $X = \bigcup_{j \in J} U_j$ – has a finite subcover, meaning that there is a finite subset $K \subseteq J$ such that $X = \bigcup_{k \in K} U_k$. Similarly, a space is Lindelöf if every open cover has a countable subcover (where K is a countable subset of J).

The topological property of **separation** provides a hierarchy of topological spaces, classified in accordance with the ability to distinguish disjoint sets and distinct points through topological methods.

Definition 4.7: **The Trennungsaxiom Hierarchy**

(0) A space is T_0 (Kolmogorov) if for every pair of distinct points p_1 and p_2 in the space, there is at least either an open set containing p_1 but not p_2, or an open set containing p_2 but not p_1.

(1) A space is T_1 (Fréchet) if for every pair of distinct points p_1 and p_2 in the space, there is an open set containing p_1 but not p_2.

(2) A space is T_2 (Hausdorff) if every two distinct points have disjoint neighborhoods.

(3) A space is $T_{2\frac{1}{2}}$ (Urysohn) if every two distinct points have disjoint closed neighborhoods.

(4) A space is T_3 (Regular Hausdorff) if it is T_0 and regular, in the sense that if V is a closed set and p is a point not in V, then V and p have disjoint neighborhoods.

(5) A space is $T_{3\frac{1}{2}}$ (Tychonoff) if it is T_0 and completely regular, in the sense that given any closed set V and any point p that not in V, then there is a continuous map f (as defined below) from X to $\mathbb{R}$ such that f(p) is 0 and f(v) is 1, $\forall v \in V$.

(6) A space is T_4 (Normal Hausdorff) if it is T_1 and normal, in the sense that any two

disjoint closed sets are separated by neighborhoods.
(7) A space is T_5 (Completely Normal Hausdorff) if it is T_1 and completely normal, in the sense that any two separated sets (disjoint from each other's closure) have disjoint neighborhoods.

Each T_j space is also a T_i space for $i \leq j$.

Definition 4.8: A map f: X $\rightarrow$ Y between two topological spaces (X, τ_X) and (Y, τ_Y) is called a **homeomorphism** if it obeys the following:
(i) It is a bijection.
(ii) It is **continuous** (with respect to τ_X and τ_Y), meaning that for each open set $U \in \tau_Y$, the inverse $f^{-1}(U) \in \tau_X$ is an open set.
(iii) The inverse function f^{-1} is continuous.

We then say that X and Y are **homeomorphic**. Two homeomorphic spaces share the same topological properties: if one of them is connected, compact or a T_i space (for some i) then the other is as well.

Definition 4.9: Let $(X_i, \tau_i)_{i \in I}$ be topological spaces, $X := \prod_{i \in I} X_i$ be a Cartesian product and π_i : X $\rightarrow X_i$ be projection maps. The **product topology** on X is the coarsest topology for which all the projections π_i are continuous maps.

Definition 4.10: Let f and g be continuous functions from a topological space (X, τ_X) to a topological space (Y, τ_Y). A **homotopy** between f and g is a continuous function H : X $\times$ [0,1] $\rightarrow$ Y from the product of the space X with the unit interval [0,1] to Y such that, if x $\in$ X then H(x,0) = f(x) and H(x,1) = g(x). We then say that f and g are homotopic. Moreover two topological spaces (X, τ_X) and (Y, τ_Y) are of the same homotopy type if there exist continuous maps f : X $\rightarrow$ Y and g : Y $\rightarrow$ X such that g $\circ$ f and f $\circ$ g are homotopic to the identity maps id_X and id_Y respectively.

An interesting recent project is the development of Homotopy Type Theory as an alternative foundation for Mathematics[246].

Definition 4.11: Let (X, τ) be a topological space and $\equiv$ be an equivalence relation on

X. The **quotient space** Q:= X/≡ is the set of equivalence classes of points in X:

$$Q = \{[x]|x \in X\} \tag{2.19}$$

together with the topology:

$$\tau_Q = \{U \subseteq Q| \bigcup_{[a]\in U} [a] \in \tau\} \tag{2.20}$$

given to subsets of X/≡.

Definition 4.12: A **metric space** is an ordered pair (M,d), where M is a set and d is a function from the Cartesian product $M \times M$ to the non-negative reals which satisfies:
(i) $d(m_1, m_2) = d(m_2, m_1)$
(i) $d(m_1, m_2) = 0$ iff $m_1 = m_2$
(i) $d(m_1, m_2) + d(m_2, m_3) \geq d(m_1, m_3)$
$\forall m_1, m_2, m_3 \in M$.

There are many interesting examples of topological spaces where the topology is derived from a notion of distance[191].

Definition 4.13: A metrizable space is a topological space that is homeomorphic to a metric space.

Urysohn metrization theorem: Every regular T_1 space which has a countable basis is metrizable[180].
The Nagata-Smirnov-Bing metrization theorem[53,297] characterizes exactly when a topological space is metrizable, namely when it is T_3 and has a countably locally finite basis.

2.4.2 Topological vector spaces

We will now present how Hilbert spaces, and therefore standard quantum theory, arise from a fusion of algebraic concepts and topological structure.

Definition 4.14: A **topological ring** is a ring R that is also a topological space (R, τ_R), such that the addition and multiplication maps (x,y) $\mapsto$ x+y and (x,y) $\mapsto$ xy from R $\times$ R $\to$ R are continuous functions (where R $\times$ R has the product topology). A topological field is a field that is also a topological ring where the inversion map is a continuous function.

<u>Definition 4.15:</u> A **topological vector space** (X, τ) is a vector space over a topological field K, where vector addition X $\times$ X $\rightarrow$ X and scalar multiplication K $\times$ X $\rightarrow$ X are continuous functions whose domains are endowed with product topologies[63].

Two important examples of topological vector spaces are Banach and Hilbert spaces.

<u>Definition 4.16:</u> Let (X, d) be a metric space. A Cauchy sequence in X is a sequence $(x_n)_{n\in\mathbb{N}}$ of elements of X such that:

$$\forall\epsilon > 0, \exists N \in \mathbb{N} \text{ such that } d(x_n, x_m) < \epsilon, \forall n, m > N \tag{2.21}$$

Note that every convergent sequence $(x_n)_{n\in\mathbb{N}}$ – which has a limit $x \in X$ such that: $\forall\epsilon > 0, \exists N \in \mathbb{N}$ such that $d(x_n, x) < \epsilon$ for $n > N$ – is a Cauchy sequence, due to the triangle inequality $d(x_1, x_2)+d(x_2, x_3) \geq d(x_1, x_3), \forall x_1, x_2, x_3 \in X$. Conversely, a metric space (X,d) where all the Cauchy sequences converge (have a limit in X) is called **complete**.

<u>Definition 4.16:</u> Let V be a vector space over a field F. A map $N : V \rightarrow \mathbb{R}^+$ is called a **norm** on V if:

(i) $N(v_1 + v_2) \leq N(v_1) + N(v_2)$

(ii) $N(f \cdot v_1) = |f|N(v_1)$

(iii) $N(v_1) = 0$ iff $v_1 = 0$

$\forall v_1, v_2 \in V, \forall f \in F$ and $|f| \in \mathbb{R}^+$.

A vector space endowed with a norm N (often denoted $||\cdot||$) is called a normed vector space.

The notion of a Cauchy sequence makes sense in the context of a topological vector space with a norm N, if we consider that for any open subset U there exists $N(U)$ such that $x_n - x_m \in U, \forall n, m > N(U)$.

<u>Definition 4.17:</u> A **Banach space** is vector space B with a norm $||\cdot||$ such that the metric space (B, d) – where the metric d is defined by taking: $d(b_1, b_2) = ||b_1 - b_2||$ for $b_1, b_2 \in B$ – is a complete metric space.

<u>Definition 4.18:</u> A **Hilbert space** is vector space H with an inner product $\langle\cdot,\cdot\rangle$ such that the norm $||\cdot|| := \sqrt{\langle h_1, h_2\rangle}$ makes H into a Banach (complete metric) space.

<u>Example 4.2:</u> A Hilbert space is always a Banach space but the converse is not true, as the following counter-example demonstrates.

Consider the space $\mathcal{C}[0,1]$ of continuous functions f: $[0,1] \to \mathbb{R}$ together with the supremum norm $||\cdot|| := \sup_{x\in[0,1]} |f(x)|$, where the supremum is the smallest positive real which is never exceeded by $|f(x)| \in \mathbb{R}^+$. Note that a Banach space (X,$||\cdot||$) which is also a Hilbert space satisfies the parallelogram law:

$$||x_1 + x_2||^2 + ||x_1 - x_2||^2 = 2(||x_1||^2 + ||x_2||^2) \tag{2.22}$$

$\forall x_1, x_2 \in X$. But consider $f_1, f_2 \in \mathcal{C}[0,1]$ such that $f_1(x) = 1$ and $f_2(x) = x$, then we get:

$$5 = ||f_1 + f_2||^2 + ||f_1 - f_2||^2 \neq 2(||f_1||^2 + ||f_2||^2) = 4 \tag{2.23}$$

Therefore $(\mathcal{C}[0,1], ||\cdot||)$ is a Banach space which is not a Hilbert space.

In order to illustrate Banach and Hilbert spaces, we will briefly introduce some basic concepts from Analysis [209].

<u>Definition 4.19:</u> A **measure space** (X, Σ, μ) consists of a set X together with a sigma-algebra Σ over X, which is a collection of subsets of X which satisfy:
(i) If $A \in \Sigma$ then the set complement $X - A \in \Sigma$
(ii) Let $\{A_1, A_2 ...\}$ be a countable family of sets in Σ then: $\bigcup_{j=1}^{\infty} A_j \in \Sigma$
(iii) $X \in \Sigma$
and a measure $\mu : \Sigma \to \mathbb{R}^+$ which satisfies:
(a) $\mu(\emptyset) = 0$
(b) $\mu(\bigcup_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} \mu(A_j)$ for a countably infinite sequence of disjoint sets in Σ.

A measure on a set provides a general method for associating a number to subsets of that set and defining integration from an abstract perspective. Measure spaces play an underlying role in the mathematical theory of probability [20].

<u>Definition 4.20:</u> Given two measure spaces (X, Σ_1, μ_1) and (Y, Σ_2, μ_2), a function $f : X \to Y$ is called measurable if: $\{x \in X | f(x) \in t\} \in \Sigma_1, \forall t \in \Sigma_2$.

<u>Definition 4.21:</u> Let (X, Σ, μ) be a measure space. An L^P space $L^P(X, \mu)$ is a set of functions $f : X \to \mathbb{C}$, together with the norm $|f|_P := (\int_X (|f|)^P)^{\frac{1}{P}}$, where all functions

$f \in L^P(X,\mu)$ are measurable and satisfy $|f|_P < \infty$. The fact that $L^P(X,\mu)$ is a vector space then follows from the inequality:

$$(|\alpha+\beta|_P)^P \leq 2^{P-1}(|\alpha|_P)^P + (|\beta|_P)^P \tag{2.24}$$

Definition 4.22: An l_p space consists of the set of sequences $x := (x_n)_{n\in\mathbb{N}}$ (with $x_n \in \mathbb{C}$) such that: $\sum_n |x_n|^P < \infty$, together with the norm $||x||_P := (\sum_n |x_n|^P)^{\frac{1}{P}}$.

Note that L^P spaces and l_p spaces are Banach spaces for all $p > 0$ and are Hilbert spaces iff p=2. An interesting result is that every Hilbert space is isomorphic to a set of the form $l_2(E)$ for some set E[195].

Definition 4.23: A linear functional ϕ on a complex Hilbert space H is a map from H to $\mathbb{C}$. A linear functional ϕ is said to be bounded if $\exists M \in \mathbb{C}$ such that: $|\phi(h)| \leq |M|||x||$, $\forall x \in H$.

The following theorem establishes the important connection between a Hilbert space and its dual space, which justifies the correspondence between bras and kets in quantum mechanics[160].

Riesz representation theorem: If ϕ is a bounded linear functional on a Hilbert space H, then there is a unique $y \in H$ such that:

$$\phi(x) = \langle y, x\rangle, \forall x \in H \tag{2.25}$$

A corollary of this theorem is the existence of a unique adjoint of a bounded operator on a Hilbert space. The adjoint $A^\star$ of a bounded operator A is defined by:

$$\langle x, Ay\rangle = \langle A^\star x, y\rangle, \forall x, y \in H \tag{2.26}$$

We can construct a larger Hilbert space by taking the tensor product of two Hilbert spaces.

Definition 4.24: Let H_1 and H_2 be Hilbert spaces with inner products $\langle\cdot,\cdot\rangle_1$ and $\langle\cdot,\cdot\rangle_2$ respectively. The tensor product $H_1 \otimes H_2$ of H_1 and H_2 is a Hilbert space with a bilinear (linear in both arguments) map $\otimes : H_1 \times H_2 \to H_1 \otimes H_2$ such that:

(i) The closed linear span of all vectors $v \otimes w$, where $v \in H_1$ and $w \in H_2$, is equal to

$H_1 \otimes H_2$.
(ii) $H_1 \otimes H_2$ has the inner product $\langle v_1 \otimes w_1, v_2 \otimes w_2 \rangle = \langle v_1, v_2 \rangle_1 \langle w_1, w_2 \rangle_2, \forall v_1, v_2 \in H_1, \forall w_1, w_2 \in H_2$.

One can show that the tensor product construction is unique up to unique isomorphism[177].

Many notions from linear algebra naturally generalize to the theory of Hilbert spaces and form the basic building blocs of quantum theory.

2.5 Category theory

2.5.1 Categories and functors

Category theory was first introduced by Eilenberg and Mac Lane[125] and increasingly thorough introductions to the theory can be found in the literature[81,5,22,213].

Definition 5.1: A **category** $\mathcal{C}$ consists of a class OBJ($\mathcal{C}$) of objects, and a class HOM($\mathcal{C}$) of arrows such that each arrow $f \in HOM(\mathcal{C})$ is associated to two objects dom(f) and cod(f), called the domain and codomain of f. This is written: $f : dom(f) \to cod(f)$.

Given arrows $f : A \to B$ and $g : B \to C$, there is an arrow $g \circ f : A \to C$ called the composite of f and g.

For each object A, there is an arrow $1_A : A \to A$ called the identity arrow of A. For every arrow $f : A \to B$, we have:

$$f \circ 1_A = f = 1_B \circ f \tag{2.27}$$

For all arrows $f : A \to B$, $g : B \to C$, $h : C \to D$, we have:

$$h \circ (g \circ f) = (h \circ g) \circ f \tag{2.28}$$

Definition 5.2: A (covariant) **functor** $F : \mathcal{C} \to \mathcal{D}$ between categories $\mathcal{C}$ and $\mathcal{D}$ maps OBJ($\mathcal{C}$) to OBJ($\mathcal{D}$) and HOM($\mathcal{C}$) to HOM($\mathcal{D}$), such that:

$$F(f : A \to B) = F(f) : F(A) \to F(B) \tag{2.29}$$

$$F(1_A) = 1_{F(A)} \tag{2.30}$$

$$F(g \circ f) = F(g) \circ F(f) \tag{2.31}$$

A contravariant functor F is defined as above except replacing equation (2.31) by:

$$F(g \circ f) = F(f) \circ F(g) \tag{2.32}$$

Definition 5.3: The **dual** category $\mathcal{C}^{op}$ of a category $\mathcal{C}$ has the same objects as $\mathcal{C}$ but each arrow $f : C \to D$ in $\mathcal{C}^{op}$ is an arrow $f : D \to C$ in $\mathcal{C}$.

Definition 5.4: In a category $\mathcal{C}$, the object:

(i) 0 is **initial** if for every object C $\in$ OBJ($\mathcal{C}$) there is a unique arrow 0 $\to$ C.

(ii) 1 is **final** if for every object C $\in$ OBJ($\mathcal{C}$) there is a unique arrow C $\to$ 1.

Definition 5.5: In a category $\mathcal{C}$, an arrow $f : A \to B$ is called:

(i) A **monomorphism**, if given any arrows g,h: C $\to$ A, $f \circ g = f \circ h$ implies g=h.

(ii) An **epimorphism** if given any arrows i,j: B $\to$ D, $i \circ f = j \circ f$ implies i=j.

These are the generalizations of the notions of injective and surjective functions (beyond the category of sets and functions).

Definition 5.6: In a category $\mathcal{C}$, an arrow $f : A \to B$ is called an **isomorphism** if it admits a two-sided inverse, meaning that there is another arrow $g : Y \to X$ in that category such that $g \circ f = 1_A$ and $f \circ g = 1_Y$.

2.5.2 Limits

Definition 5.7: Let G: $\mathcal{C} \to \mathcal{D}$ be a functor and D $\in$ OBJ($\mathcal{D}$). A **universal problem** requires one to find the 'best approximation' of D in $\mathcal{C}$. To be precise, one needs to find a **universal solution**, which is a pair {C, v} consisting of an object $C \in \mathcal{C}$ and an arrow v: $D \to G(C)$ such that, for every object $C' \in \mathcal{C}$ and every morphism f: D $\to$ G(C'), there is a unique arrow u: C $\to$ C' such that: G(u) $\circ$ v = f.

Definition 5.8: A **limit** is a universal (left) solution. Limits are unique up to isomorphism. Note that one can also define a **colimit** which is the dual notion of a limit [213].

We will now present equalizers, products and pullbacks, which are examples of limits.

Definition 5.9: Let $\mathcal{C}$ be a category containing a pair of arrows f,g : $A \to B$.

An **equalizer** of f and g is a pair $\{E, e\}$, where E $\in$ OBJ($\mathcal{C}$) and an e $\in$ HOM($\mathcal{C}$), with e: $E \to A$ such that $f \circ e = g \circ e$ and e is *universal*, in the sense that given any $z : Z \to A$ with $f \circ z = g \circ z$, there is a unique u: Z $\to$ E with $e \circ u = z$.

Definition 5.10: The **product** of two categories $\mathcal{C}$ and $\mathcal{D}$ is a new category $\mathcal{C} \times \mathcal{D}$ with objects of the form (C,D), where $C \in OBJ(\mathcal{C})$ and $D \in OBJ(\mathcal{D})$, and arrows of the form $(f, g) : (C, D) \to (C', D')$, where $f : C \to C' \in \mathcal{C}$ and $f : D \to D' \in \mathcal{D}$. Composition and units are defined component-wise.

We can define two projection functors π_i with $i = 1, 2$ such that:

$$\pi_1(C, D) = C;\ \pi_1(f, g) = f;\ \pi_2(C, D) = D;\ \pi_2(f, g) = g \tag{2.33}$$

Definition 5.11: Let $\mathcal{C}$ be a category containing a pair of arrows f: $A \to C$ and g: $B \to C$. The **pullback** of f and g consists of a pair of arrows $p_1 : P \to A$ and $p_2 : P \to B$ such that $f \circ p_1 = g \circ p_2$ and which are *universal* in the sense that: given any $z_1 : Z \to A$ and $z_2 : Z \to B$ with $f \circ z_1 = g \circ z_2$, there exists a unique arrow u: $Z \to P$ with $z_1 = p_1 \circ u$ and $z_2 = p_2 \circ u$.

Theorem 5.1: A category $\mathcal{C}$ has limits iff it has products and equalizers[22].

2.5.3 Examples of categories

We will now illustrate the definitions we have introduced by presenting examples of categories.

(A) A category with a single object is a monoid.

(B) A category with a single object in which all the arrows (group elements) are isomorphisms is a group.

(C) A category in which all the arrows are isomorphisms is a groupoid.

(D) **Set** is the category with sets as objects and functions as arrows.

In **Set**, monomorphisms are the injective functions, epimorphisms are the surjective functions and isomorphisms are the bijective functions. The empty set serves as the initial object and every singleton set is a terminal object. The product in **Set** is given by the

Cartesian product of sets and the coproduct is given by the disjoint union. An equalizer of two functions is the set of elements of the common domain where the functions are equal. The pullback of two functions f: A $\to$ C and g: B $\to$ C consists of subsets (a,b) $\in$ A $\times$ B of the Cartesian product such that the equation f(a)=g(b) holds.

(E) **Rel** is the category with sets as objects and relations as arrows.

(F) **Grp** is the category with groups as objects and group homomorphisms as arrows.

(G) **Ring** is the category with rings as objects and ring homomorphisms as arrows.

(H) $\boldsymbol{Mod_R}$ is the category with modules over a ring R as objects, and module homomorphisms as arrows. Lawvere theory[200] allows a synthetic study of the categories **Grp**, **Ring** and $\boldsymbol{Mod_R}$.

(I) $\boldsymbol{Vect_k}$ is the category with vector spaces over the field k as objects and linear maps as arrows. This is a special case of $\boldsymbol{Mod_R}$ when R is a field.

(J) **Hilb** is the category with Hilbert spaces as objects and linear maps (of norm at most 1) as arrows. **Hilb** and **FHilb**, the category with finite-dimensional Hilbert spaces as objects and linear maps as arrows, play an important role in Categorical Quantum Mechanics.

(K) **Top** is the category with topological spaces as objects and continuous functions as arrows.

Isomorphisms in **Top** are the homeomorphisms. The empty set considered as a topological space is the initial object and any singleton topological space is a terminal object. The product is given by the product topology on the Cartesian product and the coproduct is given by the disjoint union of topological spaces. Equalizers and pullbacks also resemble the equivalent notions in **Set**.

(L) **Diff** is the category with smooth manifolds as objects and smooth maps as arrows.

(M) **Cat** is the category with (small) categories as objects and functors as arrows.

In **Cat** the initial object and final object are the empty category 0 (with no objects and arrows) and the trivial category 1 (with a single object and arrow) respectively.

2.5.4 Natural Transformations and adjoints

Natural transformations provide a method of transforming one functor into another.

Definition 5.12: Let F and G be functors between categories $\mathcal{C}$ and $\mathcal{D}$. A **natural**

transformation $\eta : F \to G$ is a family of arrows $\eta_C : FC \to GC$ (where $C \in \mathcal{C}$) in $\mathcal{D}$ such that for every arrow $f : C \to C'$ in $\mathcal{C}$, we have:

$$\eta_{C'} \circ F(f) = G(f) \circ \eta_C \tag{2.34}$$

The arrow $\eta_C : FC \to GC$ (in HOM($\mathcal{D}$)) is called the **component** of η at C.

A *natural isomorphism* is a natural transformation which has a two-sided inverse, meaning that each of its components $\eta_C : FC \to GC$ ($\forall C \in OBJ(\mathcal{C})$ is an isomorphism in $\mathcal{D}$.

Definition 5.13: An **equivalence of categories** between two categories $\mathcal{C}$ and $\mathcal{D}$ consists of a pair of functors: $F : \mathcal{C} \to \mathcal{D}$ and $G : \mathcal{D} \to \mathcal{C}$ together with a pair of natural isomorphisms:

$$\epsilon_1 : (F \circ G) \to id_{\mathcal{D}} \text{ and } \epsilon_2 : (G \circ F) \to id_{\mathcal{C}} \tag{2.35}$$

The categories $\mathcal{C}$ and $\mathcal{D}$ are then said to be **equivalent**.

One can show[213] that two categories are equivalent iff there is a functor $F : \mathcal{C} \to \mathcal{D}$ which is:
(i) Full, meaning that $\forall x, y \in OBJ(\mathcal{C})$ the map $HOM(\mathcal{C})(x, y) \to Hom(\mathcal{D})(Fx, Fy)$ which is induced by F (between arrows from x to y and arrows from Fx to Fy) is surjective.
(ii) Faithful, meaning that $\forall x, y \in OBJ(\mathcal{C})$ the map $HOM(\mathcal{C})(x, y) \to Hom(\mathcal{D})(Fx, Fy)$ which is induced by F (between arrows from x to y and arrows from Fx to Fy) is injective.
(iii) Essentially surjective, meaning that $\forall y \in OBJ(\mathcal{D})$, $\exists x \in OBJ(\mathcal{C})$ such that y is isomorphic to F(x)in $\mathcal{D}$.

Definition 5.14: A pair of functors $F : \mathcal{C} \to \mathcal{D}$ and $G : \mathcal{D} \to \mathcal{C}$ are said to be **adjoint** (or form an adjunction) if there exist a pair of natural transformations:

$$\epsilon : (F \circ G) \to id_{\mathcal{D}} \text{ (counit) and } \eta : id_{\mathcal{C}} \to (G \circ F) \text{ (unit)} \tag{2.36}$$

such that:

$$(\epsilon \circ id_F) \circ (id_F \circ \eta) : F \to F \text{ and } (id_G \circ \epsilon) \circ (\eta \circ id_G) : G \to G \tag{2.37}$$

are both the identity natural transformation. We then write F $\dashv$ G and say that F is the left adjoint of G and that G is the right adjoint of F. The left or right adjoint of any functor, if it exists, is unique up to unique isomorphism.

Definition 5.15: Given a category $\mathcal{C}$ and an object $c \in OBJ(\mathcal{C})$, there is a functor Hom($\cdot$, c): $\mathcal{C}^{op} \to$ Set, called a hom-functor. Therefore, we can define the Yoneda functor Y: $\mathcal{C} \to$ Fun($\mathcal{C}^{op}$, Set), where the category of functors Fun($\mathcal{C}^{op}$, Set) is called the category of presheaves of $\mathcal{C}$.

The following result is an important representation theorem, similar in spirit to the Cayley, Stone and Riesz representation theorems which we have introduced previously.

Yoneda lemma: Let x be an object in a category $\mathcal{C}$ and F be a presheaf in Fun($\mathcal{C}^{op}$, Set). The canonical restriction map:

$$Hom_{Fun(\mathcal{C}^{op},Set)}(Y(x), F) \to F(x) \tag{2.38}$$

is an isomorphism.

Proof: We construct the inverse map $F(x) \to Hom_{Fun(\mathcal{C}^{op},Set)}(Y(x), F)$. Given f $\in$ F(x), construct a natural transformation $\eta : Y(x) \to F$ with components $\eta_y : Hom(y, x) \to F(y)$, which map an arrow h $\in$ Hom(y, x) to F(h)(f). Since F preserves composition of arrows, we can see that η is indeed a natural transformation.

One can check that F(x) $\to$ Hom(Y(x), F) $\to$ F(x) is the identity. Moreover, the naturality condition on the natural transformation η ensures that η is completely determined by the value $\eta_x(id_x) \in F(x)$ of its component on the identity morphism.

Definition 5.16: Given a functor F: $\mathcal{C}^{op} \to$ Set (a presheaf on $\mathcal{C}$), a representation of F is a natural isomorphism $\theta : Hom_{\mathcal{C}}($ $\cdot$, c) $\to$ F. By the Yoneda lemma, a representation is uniquely determined by an element of F(c), called the universal element for F.

We will conclude this section by mentioning that Category theory can be generalized to the higher order study of **n-categories**[28,76].

2.5.5 Categorical quantum mechanics

Definition 5.17: A **symmetric monoidal category** (SMC) consists of:

(i) a category $\mathcal{C}$

(ii) a functor $-\otimes$-: $\mathcal{C} \times \mathcal{C} \to \mathcal{C}$

(iii) a unit object I

(iv) natural isomorphisms (with coherence conditions[213]):

$$\lambda_A : A \cong I \otimes A,\ \rho_A : A \cong A \otimes I,\ \alpha_{A,B,C} : A \otimes (B \otimes C) \cong (A \otimes B) \otimes C,\ \sigma_{A,B} : A \otimes B \cong B \otimes A$$

Monoidal categories are ideal for describing very general compositional theories of systems and processes[3], since they contain two interacting modes $\otimes$ and $\circ$ of composition. These lead to a very simple diagrammatic calculus[267] 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.

Definition 5.18: A dagger compact symmetric monoidal category (†-CSMC) $\mathcal{C}$ is a SMC with an identity-on-objects contravariant dagger functor $\dagger : \mathcal{C} \to \mathcal{C}$ such that:

$$(f \circ g)^\dagger = g^\dagger \circ f^\dagger,\ (f \otimes g)^\dagger = f^\dagger \otimes g^\dagger,\ id_A^\dagger = id_A,\ (f^\dagger)^\dagger = f$$

which is also compact, meaning that each object $A \in OBJ(\mathcal{C})$ has a dual object $\bar{A} \in OBJ(\mathcal{C})$ (usually $A = \bar{A}$) and arrows: $\eta_A : I \to \bar{A} \otimes A$ and $\epsilon_A : A \otimes \bar{A} \to I$ such that:

$$(\epsilon_A \otimes id_A) \circ (id_A \otimes \eta_A) = id_A \text{ and } (id_{\bar{A}} \otimes \epsilon_A) \circ (\eta_A \otimes \bar{A}) = id_{\bar{A}}$$

We define a **state** of a system A as an arrow: $\psi : I \to A$, an **effect** as: $\pi : A \to I$ and scalars as $s : I \to I$. The inner product between states is then the scalar: $\psi^\dagger \circ \phi : I \to I$. The dimension of an object A is defined as: $dim(A) := \eta_A^\dagger \circ \eta_A$.

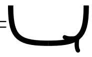
If we add to the previous graphical calculus a vertical involutive asymmetry in the boxes representing arrows and the rule that taking the adjoint reflects the boxes vertically, then we get the following key theorem which allows us to use graphical reasoning:

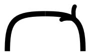
Theorem 5.2: An equational statement between formal expressions in the language of †-CSMC holds if and only if it holds up to isotopy in the graphical calculus.[181,266]

Example 5.1: Important examples of †-CSMCs are:

(i) **FHilb**, the category of finite dimensional Hilbert spaces and bounded linear maps with the usual tensor product.
(ii) **FRel**, the category of finite sets and relations with the Cartesian product of sets as the tensor product.

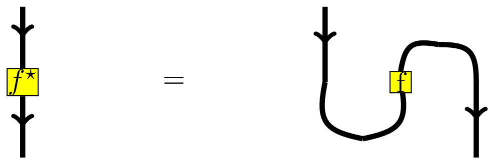
Definition 5.19: In a †-CSMC, an object A has a dual system $A^{\star}$ if there exist arrows:

$d_A : A \otimes A^{\star} \to I=$  (called cups) and

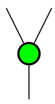
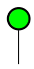
$e_A : I \to A^{\star} \otimes A=$  (called caps), such that:

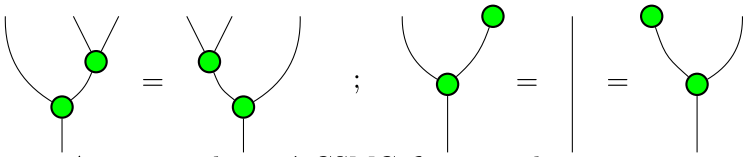
$$(d_A \otimes id_A) \circ (id_A \otimes e_A) = id_A; (id_{A^{\star}} \otimes d_A) \circ (e_A \otimes id_{A^{\star}}) = id_{A^{\star}} \tag{2.39}$$

We can then compose cups and caps to define the dual $f^{\star} : B^{\star} \to A^{\star}$ as:

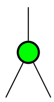
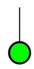


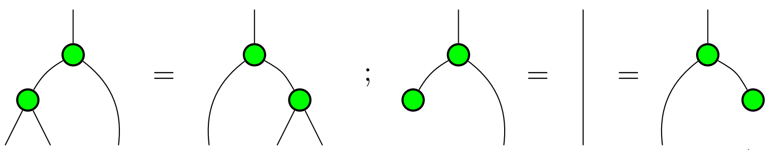
Definition 5.20: A monoid in a †-CSMC $\mathcal{C}$ is a triple:

{ A $\in OBJ(\mathcal{C})$, $\delta^{\dagger} : A \otimes A \to A$  (multipication), $\epsilon^{\dagger} : I \to A$  } (unit)

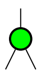
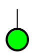
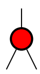
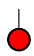
which satisfy: 

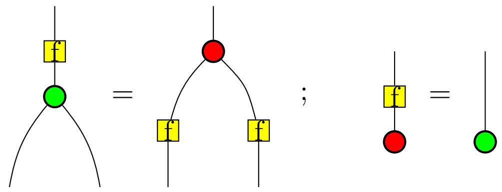
Definition 5.21: A comonoid in a †-CSMC $\mathcal{C}$ is a triple:

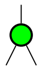
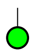
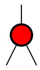
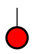
{ A $\in OBJ(\mathcal{C})$, $\delta : A \to A \otimes A$  (copying map), $\epsilon : A \to I$  (erasing map) }

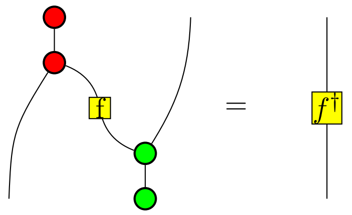
which satisfy: 

We can use differently coloured dots to represent different monoids (or comonoids) on the same object.

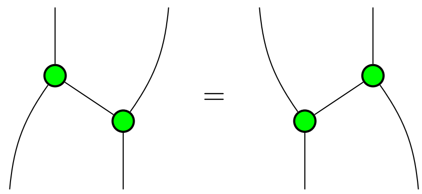
Definition 5.22: A comonoid homomorphism is a map f: (A, , ) $\to$ (A, , ) such that:



Definition 5.23: A comonoid homomorphism f: (A, , ) → (B, , ) is self conjugate if:



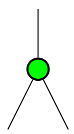
Deinition 5.24: A dagger-Frobenius algebra in a †-CSMC is a pair of a monoid and a comonoid which satisfy the following equation:



If the multipication map of the monoid is commutative then the dagger-Frobenius algebra is commutative.

We can describe bases and observables in the general context of †-CSMC by noting that the contrapositive of the no cloning [301] and no deleting theorems [229] states that orthonormal basis states are the only ones which can be copied and erased.

Definition 5.25: An observable structure is a †-special commutative Frobenius algebra on a †-CSMC $\mathcal{C}$ [88]. This is a triple:

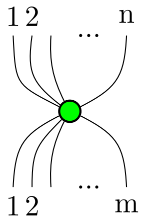
{ A $\in OBJ(\mathcal{C})$, $\delta : A \to A \otimes A$  (copying map), $\epsilon : A \to I$  (erasing map) }

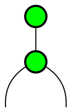
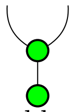
satisfying:

$$(\delta \otimes id_A) \circ \delta = (id_A \otimes \delta) \circ \delta;\ \lambda_A^{-1} \circ (\epsilon \otimes id_A) \circ\ \delta = \rho_A^{-1} \circ (id_A \otimes \epsilon) \circ \delta = id_A;\ \sigma_{A,A} \circ \delta = \delta;$$

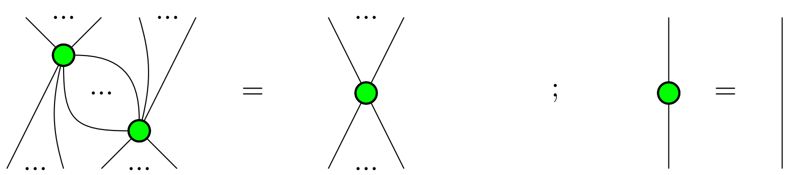
$$(\delta^\dagger \otimes id_A) \circ (id_A \otimes \delta) = \delta \circ \delta^\dagger;\ \delta^\dagger \circ \delta = id_A$$

Theorem 5.3 (Spider Theorem): Given a classical structure on A, then any process $A^{\otimes n} \to A^{\otimes m}$ built from the maps $\{\delta, \delta^\dagger, \epsilon, \epsilon^\dagger\}$ which has a connected graph is equal to the spider with n inputs and m outputs[185,83]:

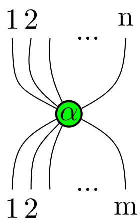


Note that each classical structure on A can be used to make A dual to itself[185] by using the caps  and cups  .

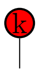
In **FHilb**, orthonormal bases are in a one to one correspondence with †-special commutative Frobenius algebras[86]. This definition for observable structures has been shown[85] to be equivalent to the spider laws depicted below.

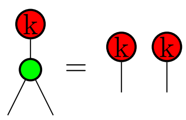


We can then illustrate spiders with n inputs and m outputs by describing them, in terms of a given orthonormal basis $\{|0\rangle, |1\rangle\}$, as:

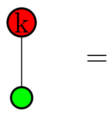


$$= \begin{cases} |00...0\rangle \longmapsto |00...0\rangle \\ |11...1\rangle \longmapsto e^{i\alpha} |11...1\rangle \\ others \longmapsto 0 \end{cases} \tag{2.40}$$

We define a classical point for an observable structure (A, δ, ϵ) as a self conjugate morphism k: $I \to A$  obeying:



and



This means that classical points are those which get copied by the copying map and deleted by the deleting map. In **FHilb**, for example, they are the basis states corresponding to the observable structure.

Symmetric monoidal categories and observable structures will play a key role in our analysis of operational physical theories.

Chapter 3

Background II: Quantum theory

3.1 Operational theories

Our scientific theories aim to accurately describe every phenomenon that can possibly occur in the world we live in. Moreover, 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 any theoretical construction is flawed. Nevertheless, in order to make predictions about physical events, it is necessary to provide a mathematical formalism, a common language used to describe physical systems and processes.

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. We take the following processes as primitive concepts for any operational physical theory: *preparations*, *transformations* and *measurements*. First of all, the preparation of a physical system consists of a repeatable procedure which outputs a valid state (Figure 3.1).

Figure 3.1: A preparation process.

Next, transformations are processes which convert valid physical systems of the theory into other valid systems (Figure 3.2).

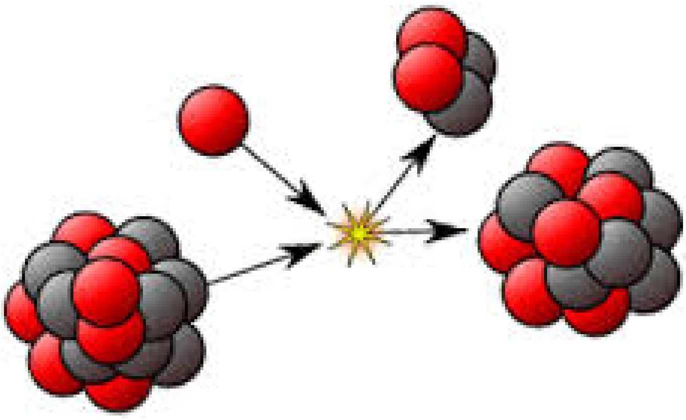


Figure 3.2: A transformation process.

Finally, measurements are repeatable procedures that receive a physical system and then produce a macroscopically distinguishable outcome from a set of possible outcomes (Figure 3.3).

Figure 3.3: A measurement process.

Each operational physical theory associates these three physical processes with mathematical objects. This provides an unambiguous description of an **operational physical theory**.

3.2 Quantum mechanics introduced

3.2.1 Orthodox postulates

A natural starting point for an analysis of the foundations of quantum theory is to present the postulates of quantum mechanics:

Axiom 1

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

Axiom 2

The evolution of a closed system is a unitary transformation:

$$|\psi(t)\rangle = U(t, t_0)\,|\psi(t_0)\rangle \tag{3.1}$$

(such that $U^{-1} = U^{\dagger}$) depending only on the initial time t_0 and the final time t.

Axiom 3

Associated with each observable property of a system is a Hermitian operator M, which therefore satisfies $M = M^{\dagger}$, has real eigenvalues and has orthogonal eigenvectors.
Hence, $M = \sum_m mP_m$, 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:

$$p(m) = \langle\psi|\,P_m\,|\psi\rangle \tag{3.2}$$

Axiom 4

Given that outcome m occurred, the state of the system changes discontinuously as:

$$|\psi\rangle \rightarrow \frac{P_m\,|\psi\rangle}{p(m)} \tag{3.3}$$

Axiom 5

If two systems $|\psi_1\rangle$ and $|\psi_2\rangle$ have state spaces H_1 and H_2 respectively and if we treat these two systems as one single compound system $|\psi_1\rangle \otimes |\psi_2\rangle$, then 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. The 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 and where the division lies between observer and observed.

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.

3.2.2 Operational axioms

Quantum theory is well suited for an operational presentation providing a minimal explanation of observable phenomena. This can be achieved by giving a description of physical preparation (P), transformation (T) and measurement (M) procedures which yields correct statistics for experiments that can be performed. In such a setting, the axioms of quantum theory can be reformulated as:

Axiom 1: Preparation

A preparation P is associated to a trace one positive operator ρ, known as the density operator, acting on a 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 a single compound preparation corresponding to 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 with a partial trace: $\rho_1 = Tr_2(\rho_{12})$.

Axiom 2: Transformation

A transformation T is associated to a completely positive trace non-decreasing map:

$$\mathcal{E} : \rho \to \mathcal{E}(\rho) \tag{3.4}$$

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).

Note that (i), (ii) and (iii) are formally equivalent to either of the following [225]:

(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 followed by a projective measurement P (that has some chance of failure) then trace out the environment.

Axiom 3: Measurement

Measurements are now a special case of Axiom 2 where each 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 mathematically unambiguous. They provide a clear target which alternative interpretations of quantum theory must reproduce.

3.3 Quantum computation

A recent approach to studying quantum theory has been to present physical processes from the viewpoint of computer science. In the last few decades, this outlook has provided an insightful new perspective. Given their direct relevance for the rest of this thesis, we will now introduce some basic ideas from Quantum Computation.

3.3.1 Quantum circuits

The **quantum circuit** model is a fundamental model of quantum computation, where finite dimensional quantum processes can be described through their linear algebraic representation. In this way, we can introduce *quantum gates* in order to depict the unitary matrices representing quantum transformations. In quantum computation, the basic concept of a classical bit, which can be in state 0 or 1, is extended to the notion of a qubit $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$, where $\alpha, \beta \in \mathbb{C}$ and $\{|0\rangle, |1\rangle\}$ is the computational basis, an orthonormal basis for the state. This allows for the *superposition* of quantum states, which can be linearly combined to form new states. Single qubit states can then be visualized on the surface of the Bloch sphere [55].

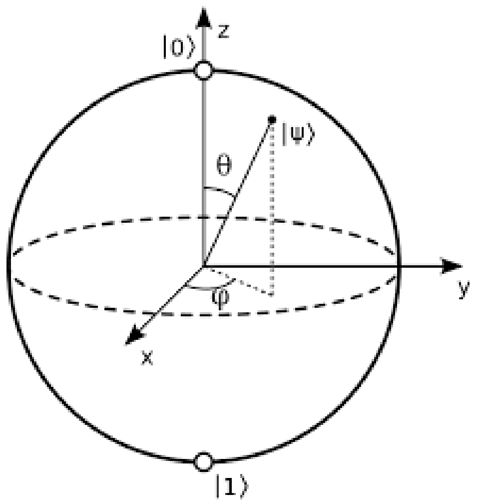


Figure 3.4: The Bloch sphere representation of a qubit.

State preparations in the circuit model consist of tensor products of qubits and quantum gates act on multiple qubit states. We introduce several examples of gates in Figure 3.5.

Measurements are represented as projection operators in the computational basis. Using the Neumark extension theorem [234,142], it is then possible to perform an arbitrary quantum (POVM) measurement by adding ancillary states to enlarge the system Hilbert space, and then performing a projective quantum measurement in the enlarged space.

We now will introduce some useful notation. The qubit Pauli operators, which are examples of single qubit gates, are defined as:

$$I := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad ; \qquad Z := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{3.5}$$

$$X := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad ; \qquad Y := iZX = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \tag{3.6}$$

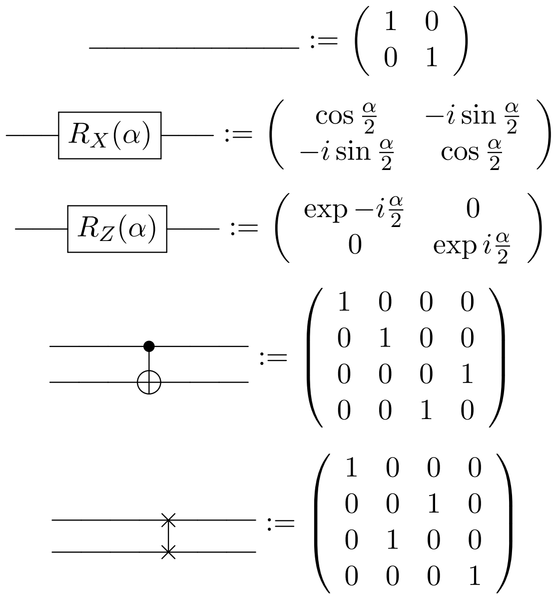


Figure 3.5: Examples of basic quantum gates.
From top to bottom: a simple wire, X and Z rotation gates, the CNOT gate and the SWAP gate.

Note that the four Pauli matrices form an orthogonal basis for the complex Hilbert space of 2×2 matrices. We denote the eigenvectors of the Pauli matrices as:

$$|0\rangle := \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad ; \qquad |1\rangle := \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{3.7}$$

$$|\pm\rangle := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix} \quad ; \qquad |\pm i\rangle := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm i \end{pmatrix} \tag{3.8}$$

Other interesting single qubit gates include the Hadamard gate H, the phase gate S and the $\frac{\pi}{8}$ gate T:

$$H := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \;;\; S := \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \;;\; T := \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix} \tag{3.9}$$

We can also define the controlled-NOT gate CNOT and the controlled phase gate CZ as:

$$CNOT\,|i\rangle\,|j\rangle := |i\rangle\,|i+j\rangle \quad ; \qquad CZ\,|i\rangle\,|j\rangle := (-1)^{ij}\,|i\rangle\,|j\rangle \tag{3.10}$$

where addition and multiplication are modulo 2.

An important question which arises is whether one can find a finite set of quantum gates which are *universal* for quantum computation, in the sense that any unitary operation can be approximated to arbitrary accuracy by a quantum circuit using only these gates. It has been shown[225,103] that the set of quantum gates: {CNOT, S, T, H} is universal for quantum computation.

There are a handful of **quantum algorithms** which currently outperform the best known classical algorithms, including Grover's algorithm for searching an unstructured database[159], Shor's factoring algorithm[272] and algorithms for solving the hidden subgroup problem[131].

3.3.2 Other quantum computation models

An alternative framework to the quantum circuit model is **measurement based quantum computation**[252,65]. In this formalism, quantum computation is performed by starting with a fixed entangled state and then performing computation by applying a sequence of measurements, in designated bases, to this initial state. Earlier measurement outcomes may affect the basis chosen for later measurements and the final result of the computation can be determined from the classical data of all the measurement outcomes. Measurement based quantum computation is universal for quantum computation, meaning that any quantum unitary transformation can be reproduced within this model.

To illustrate the idea, we will describe an example of cluster state quantum computation. A cluster state is prepared by forming a two-dimensional rectangular grid of $|+\rangle$ states and then applying a CZ gate to each nearest neighbor pair. Computation then proceeds by performing single qubit measurements, either in the computational basis $\{|0\rangle\,, |1\rangle\}$ or in a basis: $M(\theta) = \{|0\rangle \pm e^{i\theta}\,|1\rangle\}$. The computation is one-way since the initial entangled cluster state is irreversibly degraded as the computation proceeds through layers of measurements. Given a cluster state of sufficient size, this process allows the implementation of any quantum gate array[65,176].

Topological quantum computation is a framework where quantum computation is implemented by using the fusion and braiding properties of anyons (quasi-particles in topological systems)[228]. Anyonic computation can be illustrated through the Kitaev toric code[187] (in Figure 3.6) and the Kitaev honeycomb lattice model[188]. When measurement based quantum computation is implemented on a periodic three-dimensional lattice cluster state, then it can be used to implement topological quantum error correction[253].

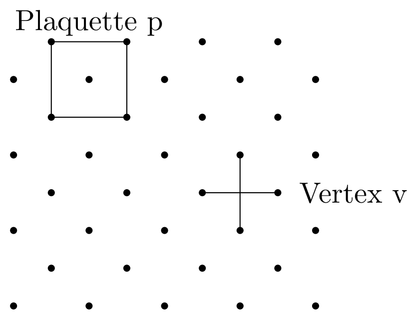


Figure 3.6: Plaquette and vertex operators on a section of the toric code.

Another quantum computational model is **adiabatic quantum computation**[133]. In this paradigm, we take a Hamiltonian (quantum operator corresponding to the total energy of a quantum system) acting on a set of particles (encoding qubits), with a non-degenerate ground state and finite energy gap above the ground state at all times. Adiabaticity (when energy is transfered only as work) ensures that the kinetic energy corresponding to the speed at which the Hamiltonian parameters change over time is considerably smaller than the energy gap above the ground state. This means that transitions away from the ground state are suppressed. Therefore, quantum algorithms are implemented by an adiabatic process where the initial ground state is easily prepared and the final ground state is the solution of the quantum computation.

The following is an example of adiabatic quantum computation. Take the Hamiltonian:

$$H(\epsilon) = (1-\epsilon)(Z \otimes I - I \otimes Z) + \epsilon(Z \otimes Z - X \otimes X) \tag{3.11}$$

acting on a two-qubit system. As we adiabatically change ϵ from 0 to 1, there is always a non-zero energy gap between the ground state and the first excited state. This adiabatic computation takes the initial ground state: $|00\rangle$ to the final ground state: $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$.

It has been shown[11] that adiabatic quantum computation is equivalent to the circuit model. Topological quantum computation closely resembles a constant energy gap adiabatic quantum computation[228].

3.4 Non locality and Contextuality

3.4.1 Realism and quantum theory

Practicing physicists usually take the philosophical view that there is a conjectured state of things as they actually exist and that our theoretical models are only the approximation of this underlying reality. Scientific progress can then be understood as an ongoing effort to improve our mind's correspondence to this reality, and every new observation brings us closer to understanding an aspect of this underlying reality. This physical reality includes everything that is and has been, whether or not it is observable or comprehensible by human beings, and is ontologically independent of our beliefs, language or theoretical constructions.

This philosophical realism has shaped our current physical theories and defined the aim of Physics as a discipline of human thought. In particular, a realist approach to quantum theory must aim to go further than just giving an account of all the results of physical 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 leads to unexpected consequences.

3.4.2 EPR

In their 1935 paper[128], 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.

3.4.3 Bohr response

Not long after the publication of the EPR paper, Bohr published a response[61] 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".

Schrödinger[263] 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, meaning that $|\psi\rangle \neq |\alpha\rangle \otimes |\beta\rangle$ for any states $|\alpha\rangle$ and $|\beta\rangle$. This leads Schrödinger to study quantum steering, or the influence of the measuring procedure of one subsystem on the other subsystem, as described by Bohr.

Bohr also 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.

3.4.4 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[127] 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[294] 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 $\langle A\rangle^2 \neq \langle A^2\rangle$ 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 [47], 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:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.12}$$

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 simulates the quantum mechanical probability $\frac{(1-\langle\psi|A|\psi\rangle/m)}{2}$

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

The average result is then:

$$\langle A\rangle = \langle\psi|\, A\, |\psi\rangle = \frac{m(1+\langle\psi|\, A\, |\psi\rangle /m)}{2} - \frac{m(1-\langle\psi|\, A\, |\psi\rangle /m)}{2} \tag{3.13}$$

which perfectly agrees with the quantum mechanical prediction since experiments yield a uniform distribution of λ between -1 and 1. Measurement of B gives 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, $\langle A + B\rangle = \langle A\rangle + \langle B\rangle$ 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.

3.4.5 Bell's theorem and the CHSH inequality

Bell derived a quantitative criterion for the existence of a realistic interpretation of any local theory [46]. Consider as an example a system of two spin half particles. Note that we could reformulate this example in terms of boxes with switches and lights flashing such that the inequality obtained is purely about operational correlations. Suppose that both '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:

$$P(a, b) = \int d\lambda \rho(\lambda) \bar{A}(a, \lambda) \bar{B}(b, \lambda) \tag{3.14}$$

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:

$$|P(a,b) - P(a,c)| \leq \int d\lambda\rho(\lambda)[1 \pm \bar{A}(c,\lambda)\bar{B}(d,\lambda)] + \int d\lambda\rho(\lambda)[1 \pm \bar{A}(c,\lambda)\bar{B}(b,\lambda)] \quad (3.15)$$

This then yields an inequality that cannot be violated by a local hidden variable theory first derived by Clauser, Holt, Shimony and Horne [79] (CHSH inequality):

$$|C| = |P(a,b) - P(a,d)| + |P(c,d) + P(c,b)| \leq 2 \quad (3.16)$$

The original form of the result, given in Bell's original paper [46], can be derived using c=d and $P(d,d) = -1$ such that the CHSH inequality becomes:

$$|P(a,b) - P(a,d)| \leq 1 + P(d,b) \quad (3.17)$$

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 [21] verifying this violation of the CHSH inequality using pairs of photons. Several loopholes [48] 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 [140], 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 that everything was once in the same light cone).

3.4.6 Cirelson bound

Cirelson asked whether quantum theory enforces an upper limit on non-local correlations [78], 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 bound. 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.

3.4.7 Popescu Rohrlich boxes

In a 1994 article, Popescu and Rohrlich asked the question of whether non-locality can be used as an axiom for quantum theory [243]. 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 [242], 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 corresponds 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:

$$|C| = |P(a, b) - P(a, d)| + |P(c, d) + P(c, b)| = 4 \tag{3.18}$$

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 [286].

Simulation of entangled states would be a bit too easy and communication complexity would become trivial if PR boxes existed in nature [289]. 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 [260]. Studying extra features that we expect the world to satisfy can yield valuable potential physical axioms for quantum theory, or even theories going beyond quantum mechanics.

3.4.8 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 [149].

The CHSH inequality can also be easily generalized [236] 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 the directions: $a_1, a_3, ..., a_{2n-1}$ 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:

$$|\langle A_1 B_2\rangle + \langle B_2 A_3\rangle + ... + \langle A_{2n-1} B_{2n}\rangle - \langle B_{2n} A_1\rangle| \leq 2n - 2 \tag{3.19}$$

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 [161,237], with more observers, measuring settings and measurement results. Some of these generalized Bell-type inequalities may be undiscovered and have novel features and applications. Brunner et al. [67] wrote an excellent review of recent developments in the study of non-locality. In the next section, we will describe a generalization to three observers which is particularly elegant and interesting.

3.4.9 Mermin non-locality

Based on an argument by Greenberger, Horne and Zeilinger, Mermin described a new test of non-locality [218] which does not depend on an inequality based upon the statistics of data accumulated in many runs, but relies instead on the outcome of a single experimental run.

Let a source emit a trio of particles which go to three far-away detectors. These detectors have two switch settings 1 and 2 and emit either a red or green light, as in the operational description of the CHSH experiment.

Einstein locality would then lead us to conclude that all the information concerning 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).

Every one of these sets of instructions also 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.

However, we can create a set-up where 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.

This can be achieved using quantum mechanics. Indeed, let one prepare a three particle GHZ state:

$$|GHZ\rangle = \frac{1}{\sqrt{2}}(|000\rangle - |111\rangle) \tag{3.20}$$

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:

$$\sigma_x \otimes \sigma_x \otimes \sigma_x = -(\sigma_x \otimes \sigma_y \otimes \sigma_y)(\sigma_y \otimes \sigma_x \otimes \sigma_y)(\sigma_y \otimes \sigma_y \otimes \sigma_x) \tag{3.21}$$

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 space-like 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.

3.4.10 The over-protective seer

In order to illustrate his early thoughts on the limitations of non-contextuality, Specker introduced a mathematical parable[274]. The story is that of an overprotective seer who does not wish for his daughter to marry any of her suitors. If they hope to claim the hand of the seer's daughter, then the suitors had to overcome the following trial.

They were each given three boxes, which may or may not contain a gem, and told to pick any two boxes and state whether they expect both boxes to be empty or both boxes to be full. After each suitor had made his prediction, he was ordered by the father to open the two boxes which he had predicted to be both empty/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' associated to the 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 observed (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).

3.4.11 Gleason's theorem

Gleason[151] was interested in reformulating quantum theory using a weaker set of axioms than Von Neumann's[294]. 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 a function which associates a non-negative real number to each closed subspace, such that for any countable collection of mutually orthogonal subspaces A_i having closed linear span B, we get:

$$\mu(B) = \sum_i \mu(A_i) \tag{3.22}$$

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: $\langle P(uv)\rangle = \langle P(u)\rangle + \langle P(v)\rangle$.

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 addressing hidden variables, Gleason's work was an important source of inspiration for the no-go theorems of Bell and Kochen-Specker.

3.4.12 Bell corollary of Gleason's theorem

In a paper written before his famous non-locality article, Bell derived an important corollary[47] of Gleason's work in the form of a no-go theorem against non-contextual hidden variable theories.

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

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

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

Now, let u be a normalized vector such that, for a given state, $\langle P(u)\rangle = 1$ and let v be a vector such that $\langle P(v)\rangle = 0$. 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: $\langle P(u')\rangle = \langle P(u'')\rangle = 0$.

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

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

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

$\langle P(u)\rangle = \langle P(u + u'')\rangle + \langle P(u - u'')\rangle = 0$, 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 $\langle P(u)\rangle \neq \langle P(v)\rangle$. 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, then 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 the one Bohr made against Einstein's criterion of reality when he introduced the notion of complementarity [61].

Bell elegantly explains 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.

3.4.13 Kochen Specker theorem

The Kochen Specker theorem [190] 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. This means that the result of a measurement depends on the context of the measurement.

A more precise statement of the Kochen-Specker theorem is that in a Hilbert space of dimension N 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 carefully 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[190] used a set of 117 vectors in real three dimensional space but a number of proofs involve fewer vectors. Conway and Kochen found a proof using 31 vectors[236] and Peres came up with two particularly elegant proofs[235] using 33 rays in $\mathbb{R}^3$ and 20 rays in $\mathbb{R}^4$. In higher dimensions, the theorem can usually be proven using fewer vectors[238], particularly if we restrict the analysis to a known state[184].

Similarly to the Bell theorem, the Kochen Specker theorem does not only 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 measurement.

A recent analysis which is worthy of mentioning here is the Cabello Severini Winter graph-theoretic approach to contextuality[69]. The Kochen Specker result can also be recast in logical terms as a result about partial Boolean algebras within a category-theoretic framework[45] and Abramsky and Hardy introduced logical Bell inequalities[4] based on logical non-contextual consistency conditions. In general, Abramsky, Brandenburger and co-workers have used sheaf theory to give a unified treatment of non-locality and contextuality[2,6].

3.4.14 Mermin magic square

We will now conclude our discussion of contextuality by presenting an elegant result by Mermin[217].

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:

$$\begin{array}{ccc} I \otimes \sigma_z & \sigma_z \otimes I & \sigma_z \otimes \sigma_z \\ \sigma_x \otimes I & I \otimes \sigma_x & \sigma_x \otimes \sigma_x \\ \sigma_x \otimes \sigma_z & \sigma_z \otimes \sigma_x & \sigma_y \otimes \sigma_y \end{array} \tag{3.23}$$

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, since each row and column of the square must multiply to one except the last row, which gives -1. This contradiction leads us to conclude that observables cannot 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 [218], using a five-pointed star instead of a square.

Contextuality is a central and recurring topic in the foundations of quantum theory.

3.4.15 Leggett-Garg inequality

Having opened this section with the discussion of the EPR article, it is fitting to conclude by introducing the Leggett-Garg inequality. It has been shown that the predictions of quantum mechanics are incompatible with the following postulates [203]:
(i) Macroscopic realism: "A macroscopic object, which has available to it two or more macroscopically distinct states, is at any given time in a definite one of those states."
(ii) Noninvasive measurability: "It is possible in principle to determine which of these states the system is in without any effect on the state itself, or on the subsequent system dynamics."

Indeed, by assuming (i) and (ii), we can define a physical quantity Q which can take on two distinct values Q $= \pm 1$, as well as the correlation functions $K_{ij} := \langle Q(t_i)Q(t_j)\rangle$ (where i $<$ j) for three times $t_1 < t_2 < t_3$. The assumptions (i) and (ii) then impose the inequality [203]:

$$K_{12} + K_{23} - K_{13} \leq 1 \tag{3.24}$$

Quantum mechanics, on the other hand, violates this inequality with a maximal value of $K_{12}+K_{23}-K_{13} = \frac{3}{2}$. As with the Bell inequalities, there are a range of different Leggett-Garg inequalities, whose violation has been demonstrated in a wide array of physical systems [130]. In essence, these are the analogue of Bell inequalities but with space-like separation of observers replaced by separation in time.

3.5 Ontological models for 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. 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.

One can make this quest for a realist interpretation of quantum theory more formal by introducing ontological models[165,167]. 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 distribution over Λ:

$$|\psi\rangle \in H^{(d)} \leftrightarrow (\mu_{P,|\psi\rangle}(\lambda)) \tag{3.25}$$

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:

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

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:

$$p(k|P,M) = \langle \xi_{M,k}(\lambda)\mu_{P,|\psi\rangle}(\lambda)\rangle_\Lambda := \int d\lambda \xi_{M,k}(\lambda)\mu_{P,|\psi\rangle}(\lambda) \tag{3.26}$$

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. Also, measurements can 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 more subtle meta-ontological model as we discuss in Chapter 5).

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. We will discuss possible alternative mathematical formulations of the ontic space Λ in Chapter 5.

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

3.5.1 Examples of ontological models

As an illustration, we shall now study several examples of simple ontological models [257,165].

(A) The first of these is the Beltrametti-Bugajski model [49]. 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:

$$\xi(k|\lambda, M) = Tr(|\lambda\rangle\langle\lambda| M_k) \tag{3.27}$$

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:

$$pr(k|M, \psi) = \int d\lambda \xi(k|\lambda, M)\mu(\lambda|\psi) = Tr(|\psi\rangle\langle\psi| M_k) \tag{3.28}$$

(B) The next model, which is for two dimensional Hilbert spaces, is due to Kochen and Specker[190]. The ontic states are vectors λ on the unit sphere Λ and the quantum state ψ is associated with the probability distribution:

$$\mu(\lambda|\psi) = \frac{1}{\pi}\Theta(\psi \cdot \lambda)\psi \cdot \lambda \tag{3.29}$$

where Θ is the Heaviside function, defined by $\Theta(x) = 1$ 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:

$$p(\phi|\psi) = \int d\lambda \xi(\phi|\lambda)\mu(\lambda|\psi) = \frac{1}{2}(1 + \psi \cdot \phi) = |\langle\psi|\phi\rangle|^2 \tag{3.30}$$

Note that Bell's hidden variable model[47], which we previously described as a counter-example of 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[257]. The ontic state in this case consists of all the rank one projectors in $GL(3, \mathbb{C})$, which is the general linear group of degree 3, or the set of all 3×3 invertible complex matrices.

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

$$\mu(\lambda|\psi) = N(Tr(\lambda\lambda_\psi) - \Delta) \text{ if } Tr(\lambda\lambda_\psi) - \Delta \geq 0$$
$$\text{or } \mu(\lambda|\psi) = 0 \text{ otherwise.}$$

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

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 extremely 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.

3.5.2 Spekkens toy theory

In defense of ψ-epistemic interpretations of quantum theory, Spekkens introduced a toy theory [276] 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:

$$1 \vee 2 \leftrightarrow |0\rangle,\ 1 \vee 3 \leftrightarrow |+\rangle,\ 1 \vee 4 \leftrightarrow |-i\rangle,\ 2 \vee 3 \leftrightarrow |+i\rangle,$$
$$2 \vee 4 \leftrightarrow |-\rangle,\ 3 \vee 4 \leftrightarrow |1\rangle \text{ and } 1 \vee 2 \vee 3 \vee 4 \leftrightarrow \frac{I}{2}$$

where $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $|\pm i\rangle = \frac{1}{\sqrt{2}}(|0\rangle + i\,|1\rangle)$.

Transformations on the ontic states $IV \to IV$ 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 the 24 permutation elements of the symmetric group S_4 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:

$$1 \vee 2 \text{ or } 3 \vee 4\ ? \quad ; \quad 1 \vee 3 \text{ or } 2 \vee 4\ ? \quad ; \quad 1 \vee 4 \text{ or } 2 \vee 3\ ?$$

Note that measurement would be deterministic if the ontic state were 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 which were 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:

(i) 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.

(ii) Applying any allowed operation to a state must yield an epistemic state which satisfies 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 [197]. Recently, Spekkens and Schreiber have been working on generalizing the theory to higher dimensional systems [262]. They have shown that an epistemic restriction based on a discrete version of the uncertainty principle allows us to extend the toy model to higher dimensions. This will be discussed in more detail in Chapter 5.

3.5.3 Contextuality for ontological models

Spekkens has introduced an operational definition of contextuality which applies to ontological models and to arbitrary operational theories [275]. This generalized notion of noncontextuality 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:

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

Similarly, transformation (or measurement) noncontextuality is the feature 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:

$$p(k|P,M) = p(k|P,M')(forallP,k) \implies \xi_{M,k}(\lambda) = \xi_{M',k}(\lambda) \tag{3.31}$$

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[275]. 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 [277] [69] [207], 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.

3.5.4 PBR theorem

So far, we have seen that ontological models for quantum mechanics 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 [163], 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 [221], assuming that the dynamics of the ontic states is Markovian. Moreover, Colbeck and Renner [94,95] have demonstrated that an extensive class of hidden-variable extensions of quantum theory cannot give any more predictive information about the outcomes of future measurements than quantum theory itself.

Pusey, Barrett and Rudolph, in an attempt to clarify what a quantum state represents, introduced another no-go theorem for ontological models [248]. 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:

$$\int ... \int p(k|\lambda_1, ..., \lambda_n)\mu_{x_1}(\lambda_1)...\mu_{x_n}(\lambda_n)d\lambda_1...d\lambda_n \tag{3.32}$$

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[248] 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 between 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[206], 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.

Another interesting no-go result similar to PBR[36] provides an upper bound on the extent to which the probability distributions in ψ-epistemic models can overlap if they are to be consistently reproduce quantum predictions.

We could alternatively take the approach of quantum Bayesianism [71,72,139] and argue for a ψ-epistemic interpretation of quantum theory, where the quantum state represents information about possible measurement outcomes (regardless of any underlying ontology), 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.

3.6 Ontological interpretations 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.

3.6.1 Bohmian mechanics

The ontic state in Bohmian mechanics [58] [59] 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 Schrödinger equation:

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi + V(r)\psi \tag{3.33}$$

where $\psi(r,t) = R(r,t)\exp\frac{iS(r,t)}{\hbar}$, along with the guidance equation:

$$\frac{d\xi(t)}{dt} = \frac{1}{m}[\nabla S(r,t)]_{r=\xi(t)} \tag{3.34}$$

which is a first order equation. Note that we choose a spacetime frame [x,t] and that this is not a fundamentally Lorentz invariant theory.

The Hamilton-Jacobi equation (real part of the Schrödinger equation):

$$\frac{\partial S}{\partial t} + \frac{(\nabla S)^2}{2m} + V + Q = 0 \tag{3.35}$$

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 Schrödinger equation):

$$\frac{\partial R^2}{\partial t} + \nabla \cdot (R^2 \frac{\nabla S}{m}) = 0 \tag{3.36}$$

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, meaning that it is not affected by ξ. This field provides active information to the particle: very little energy directs a much greater energy.

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[190], 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. We shall not delve further into the details of this theory but note that they are well described in Bohm and Hiley's book[60]. We will not go through objections of de-Broglie Bohm theory here, but will instead move on to a description of many-worlds theory.

3.6.2 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 [132] and DeWitt [111].

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 [278]. 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 [182] 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[108], 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[295]).

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.

3.6.3 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 Schrödinger'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.

These collapse models will play a role in the final chapter of the thesis where a novel collapse model is introduced, providing an interesting quantum-like theory that will be discussed at some length. In that chapter, we will present collapse models in some detail so we will keep this section brief and only give a feeling for spontaneous quantum collapse.

Let us now briefly look at an example of a dynamic collapse model due to Ghirardi, Rimini and Weber[145]. The wave function for N particles is assumed to evolve according to the Schrödinger equation: $i\hbar\frac{\partial}{\partial t}\left|\psi(t)\right\rangle = H\left|\psi(t)\right\rangle$ 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):

$$\left|\psi(t+dt)\right\rangle = \frac{1}{\sqrt{p(q_k)}}\sqrt{E^{(k)}(q_k)}\left|\psi(t)\right\rangle \tag{3.37}$$

where

$$E^{(k)}(q_k) = \int dr_k K \exp\frac{-(r_k-q_k)^2}{\sigma^2}\left|r_k\right\rangle\left\langle r_k\right| \tag{3.38}$$

is a positive operator which has expectation values:

$$p_k = \left\langle\psi(t)\right| E^{(k)}(q_k)\left|\psi(t)\right\rangle \tag{3.39}$$

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 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:

$$\psi = \sum_n C_n \psi_n(x) \phi_n(y_1, ..., y_R, Y) \tag{3.40}$$

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 very rapidly leave the single result $\phi_m(y_1, ..., y_R, Y)$ with a well defined pointer reading.

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.

We will return to a more detailed survey of quantum collapse models in Chapter 6.

To conclude this section, we note that there are many other interpretations of quantum theory, such as the two-state vector formalism [12], the consistent histories approach [114], quantum measure theory [273], quantum causal sets [115], the transactional interpretation [100], modal interpretations [292], and quantum logic [54].

3.7 Generalized probabilistic theories

Whether a physical theory is aiming to reproduce natural phenomena 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 through the eyes of 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. We will now present a larger space of hypothetical theories, containing quantum theory, namely generalized probabilistic theories.

3.7.1 Hardy's operational framework

Based on previous work by Ludwig [212] and collaborators, Hardy [162] introduced a simple framework for convex operational theories, where 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 which 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 K×K 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:

$$r_m \cdot p_n = \delta_{mn} \tag{3.41}$$

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 [162], 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.

3.7.2 Information theoretic constraints for quantum theory

Two years after Hardy's paper, Clifton, Bub and Halvorson [80] 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.

3.7.3 Information processing in generalized probabilistic theories

Barrett introduced a framework[35] 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[162]. The Barrett assumptions, however, take the degrees of freedom of the state as internal degrees of freedom, which requires a closer analysis of the role of spacetime and treats transformations and measurements in a unified way. The latter assumptions deal with how systems combine to make other systems. These allow us to derive a theorem that systems combine according to a specific 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[138], the multiple decompositions of a mixed state into pure states, the no-deleting theorem[229] and the no-cloning theorem[301]. The paper also describes in detail how classical and quantum theory fit into the framework and introduces a general non-signaling theory (box-world), containing states giving rise to PR-box correlations[242], as well as a generally local theory (GLT), 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[34], time and causal structure[164,13], phase[141] and computation[201,202] in these generalized probabilistic theories.

Chapter 4

The logic of Stabilizer Quantum Mechanics

There is often a consensus among scientists that disciplines such as Physics, Chemistry and Biology are concerned with empirical knowledge. This would mean that our knowledge of physics can only be acquired through the experience acquired from our sense data, in conjunction with a process of induction, abstraction and synthesis. Logic, on the other hand, is concerned only with a priori knowledge, which need not be justified by experience. But what is the nature of this a priori knowledge; should it be interpreted as a special insight into physical reality, a profound understanding of our own minds or a simply a disambiguation of language?

How can one make any statements about the universality of a physical law based upon an inductive generalization coming from observation? Indeed, from our empirical experience we can only ever make a finite number of imperfect observations, so can we then assumingly proceed to attribute philosophical truth to physical theories? Moreover, it is not only impossible to deny the possibility of exceptions to physical laws but a closer examination shows that we can never expect to exactly verify any law. Yet, the fact that a physical measurement of the hypotenuse of a right triangle with sides of length one will never yield $\sqrt{2}$ does not lead us to question the validity of any mathematical axioms in physics.

Mathematical theories define objects and use a language that are detached from our sensory experience. The efficiency of Euclidean geometry as a tool for building bridges does not mean that we can give any physical meaning to concepts like a *point* which has no part or an infinitely extended *straight line*. Are physical experiments just self-fulfilling prophecies, where we are defining abstract objects and processes that trivially obey the rules we set out because they are defined in that way? Should our physical theories be understood as a priori knowledge that cannot be strengthened at all by observing additional evidence towards our claim?

In addition, what can we make of the existence of numerous conflicting axiomatic systems which may all be logically consistent? Can arguments of parsimony and elegance such as Ockham's razor really give us a reason to prefer some theoretical constructs rather than others? The danger in the mathematical foundations of physics is that we may be inclined to feel that certain axioms and theorems are true and therefore our attention might be taken away from the logical interrelations between them. It is often tempting to believe that apparently obvious theorems follow from premises which do not rigorously entail them through a deductive system of reasoning. Conversely, belief that certain results are absurdly false or not in line with "physical intuition" may lead one to discard consistent physical theories.

Therefore, formal logical systems must play a prevalent role in Physics. In particular, studying whether logical systems describing physical processes obey properties such as universality, soundness or completeness can yield important information concerning physical theories. In the logical foundations of Physics, we can define an abstract process language whose primitive terms are physical processes. Equivalence of physical processes then provides a high-level axiomatic system and theorems follow from this system by pure deductive logic. Physical meaning no longer plays a fundamental role, and physical objects have no more meaning than that defined by the mathematical rules of the formal system.

Given the numerous foundational accounts of Quantum Theory, it is essential to provide a coherent narrative, where physical theories follow deductively from consistent sets of axioms. One must start with a careful examination of the mathematical terms in a physical theory: the context and interpretation of the abstract terms are essential. We can then follow Tarski[281] is requiring that truth can only be understood relative to a given language $\mathcal{L}$ and can only be expressible in a meta-language $\mathcal{M}$, which contains $\mathcal{L}$ and can be used to analyze the syntax of $\mathcal{L}$. Truth can be defined via a physical meta-language $\mathcal{M}$, which provides a framework to study a variety of alternative theories. Can we define such a meta-language of physics?

4.1 Introduction

Studying quantum theory from a computational and information-theoretic point of view has provided important no-go theorems [46,190,301,33,248], a description of new physical phenomena [52,50,51,303] and a better understanding of the importance of quantum resources, like entanglement [169]. The development of quantum computation as a sub-discipline of computer science in its own right, moreover, leads us to ask important new questions that would not normally occur to physicists.

There are several natural logical properties that are important with regard to quantum computation. The first one of these is universality. It has been shown [110,32] that a universal set of gates for any quantum computation consists of single qubit gates and the controlled-not gate. This means that any valid quantum circuit can be built up using composition and tensor products of gates in this universal set.

Two other essential properties for logical systems are soundness and completeness. Previous work has focused on whether abstract diagrammatical systems are sound or complete for quantum mechanics [268,82,24]. Here, we wish to present soundness and completeness in a more concrete setting by describing them in terms of familiar quantum circuits. This should clarify the importance of these logical properties from the viewpoint of quantum computation, in analogy with the work done on the role of universality in quantum computation.

Assume that we are given a set of equations between quantum circuits. New circuit equations can be obtained by locally substituting parts of circuits by equal quantum circuits.

Soundness guarantees that any equation between quantum circuits that can be deduced from an original set of equations is in agreement with quantum theory. A set of circuit equations is sound if each quantum circuit equation in the original set of equations agrees with quantum mechanics and if any equation built from this original set is also in agreement with quantum theory.

Completeness ensures that any equation between quantum circuits that is true in quantum theory can be deduced from the original set of equations. A complete set of circuit equations for quantum mechanics is one from which the equality of any two quantum circuits corresponding to the same physical process can be deduced. Although constructing a set of circuit equations that is sound for quantum theory is simple, finding a complete set of circuit equations is far from trivial. Such a set, if it exists, would provide a logical set of axioms from which one could formally derive whether or not any two quantum processes are equivalent.

In this chapter, based on joint work with Bob Coecke[251], we restrict the search for a complete set of circuit equations to a subclass of quantum mechanics, namely stabilizer quantum theory. A stabilizer quantum mechanics process consists of tensor products and compositions of computational basis state preparations, Clifford unitaries and measurements of observables in the Pauli group (or at least one of these three). Two such physical processes are equivalent if they can be described by exactly the same quantum circuit.

This naturally leads us to ask the following question:

Can one find a sound and complete set of quantum circuit equations from which one can deduce the equivalence of any two stabilizer processes?

We answer this question in the affirmative. The crux of the proof draws from converting an abstract graphical calculus into quantum circuits.

In the following, we construct a logical circuit calculus whose elements correspond to physical stabilizer processes. We show that this calculus is equivalent to an abstract graphical calculus called the ZX network[82].

This demonstrates that familiar quantum circuits can always be used instead of the algebraic calculus to study stabilizer theory. However, since the ZX network diagrams are not restricted to the structure of circuits, the ZX network is a more flexible and convenient tool for calculation. The abstract calculus relies on reasoning with diagram elements which have no explicit physical interpretation.

The elements of the circuit calculus, on the other hand, correspond directly to physical systems and processes. Therefore, we can use this graphical language to study the physical theory of stabilizer quantum mechanics from a logical point of view. This allows us to explicitly present a **complete set of quantum circuit equations for stabilizer quantum mechanics**.

This is an important result towards understanding the logic of stabilizer quantum mechanics: this complete set of circuit equations is a set of axioms from which any two stabilizer quantum circuits which are identical can be proven to be the same. Note that the existence of such a definable complete set of *circuit equations* cannot be deduced from only studying the abstract ZX network.

4.2 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:

$$P_n := \{\alpha g_1 \otimes ... \otimes g_n : \alpha \in \{\pm 1, \pm i\},\ \text{with: } g_k \in \{I, \sigma_x, \sigma_y, \sigma_z\}, \forall k \in \{1, ..., n\}\} \tag{4.1}$$

The Clifford group is the group of unitary operations:

$$C_n := \{U : UgU^\dagger \in P_n, \forall g \in P_n\} \tag{4.2}$$

It is generated by the phase, Hadamard and C-NOT gates.
Stabilizer quantum mechanics[155] includes state preparations in the computational basis, Clifford unitaries and measurements of observables in the Pauli group. This non-universal subclass of quantum mechanics is particularly important for a large number of quantum protocols, including quantum teleportation[50], super-dense coding[52] and quantum key distribution[150]. It also underlies the current theory of quantum error correction.

By the Gottesman-Knill theorem[154], stabilizer quantum mechanics can be efficiently simulated by a classical computer. It has been shown[91] [247] that there is a close relationship between the stabilizer formalism and Spekkens' toy theory[276].

Independently from work presented here, a recent result by Selinger[269] presents a rewrite system by which any Clifford operator can be reduced to a unique normal form. This is similar in spirit to the logical analysis of stabilizer quantum mechanics we consider here.

4.3 ZX network

We will now describe the ZX network[82,83], which is a two-colored pictorial calculus aiming to reproduce certain aspects of quantum theory. This calculus directly allowed us to find the complete set of circuit equations for stabilizer quantum mechanics presented below.

General network diagrams are built out of parallel (tensor product) and downward compositions of generating diagrams from Figure 4.1.

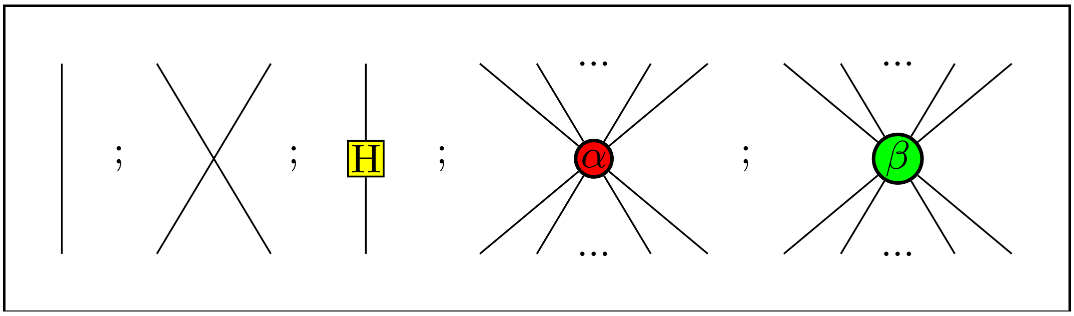


Figure 4.1: Generating diagrams for the ZX network.

The axioms of the ZX network are summarized in Figure 4.2. The (T) rule means that after identifying the inputs and outputs of any part of a ZX network, any topological deformation of the internal structure does not matter. The (H) rule was introduced in[118].

Two network diagrams can be shown to be equal by locally replacing some part of a diagram with a diagram equal to it.

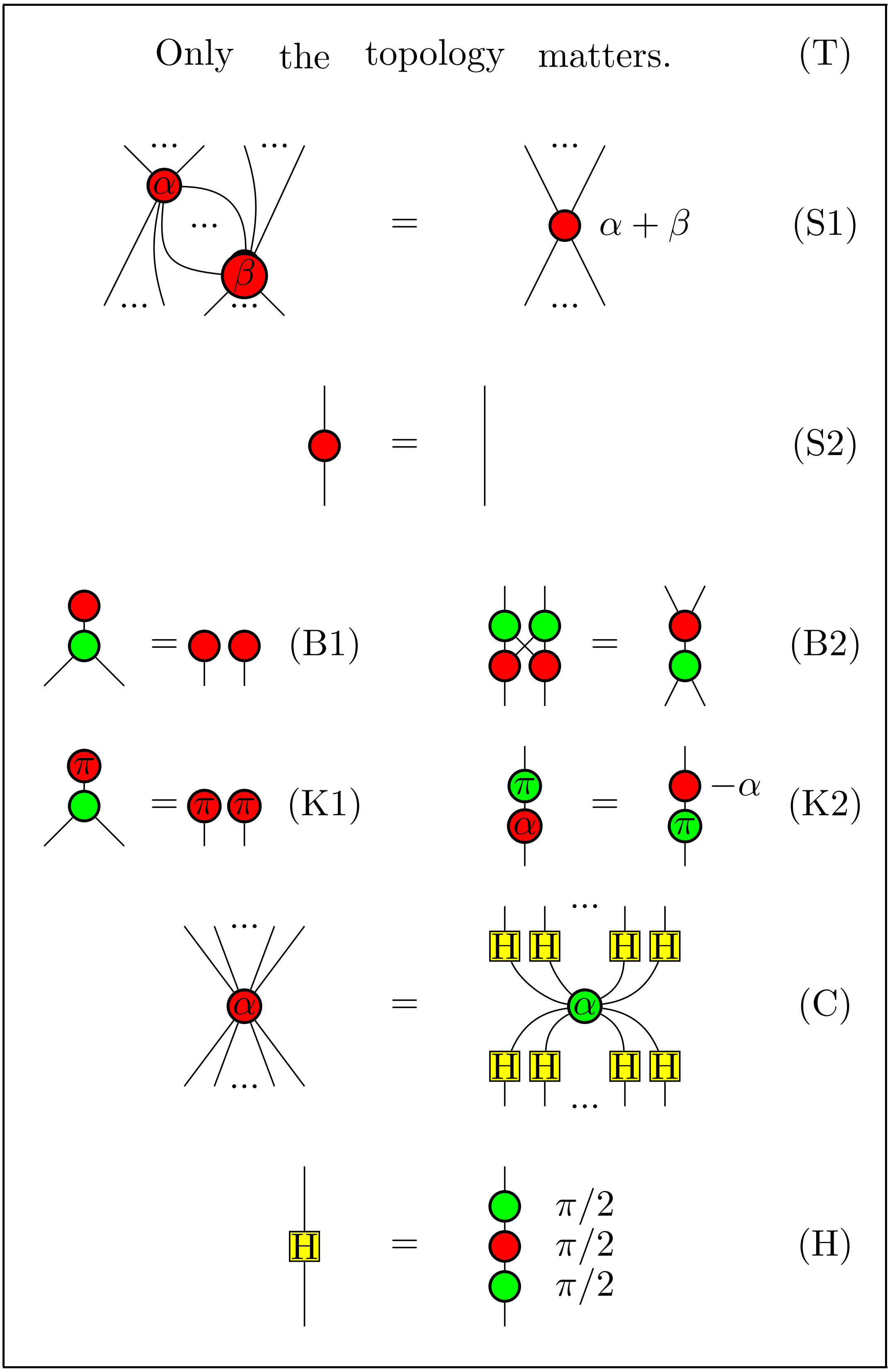


Figure 4.2: Diagrammatic rules for the ZX network.

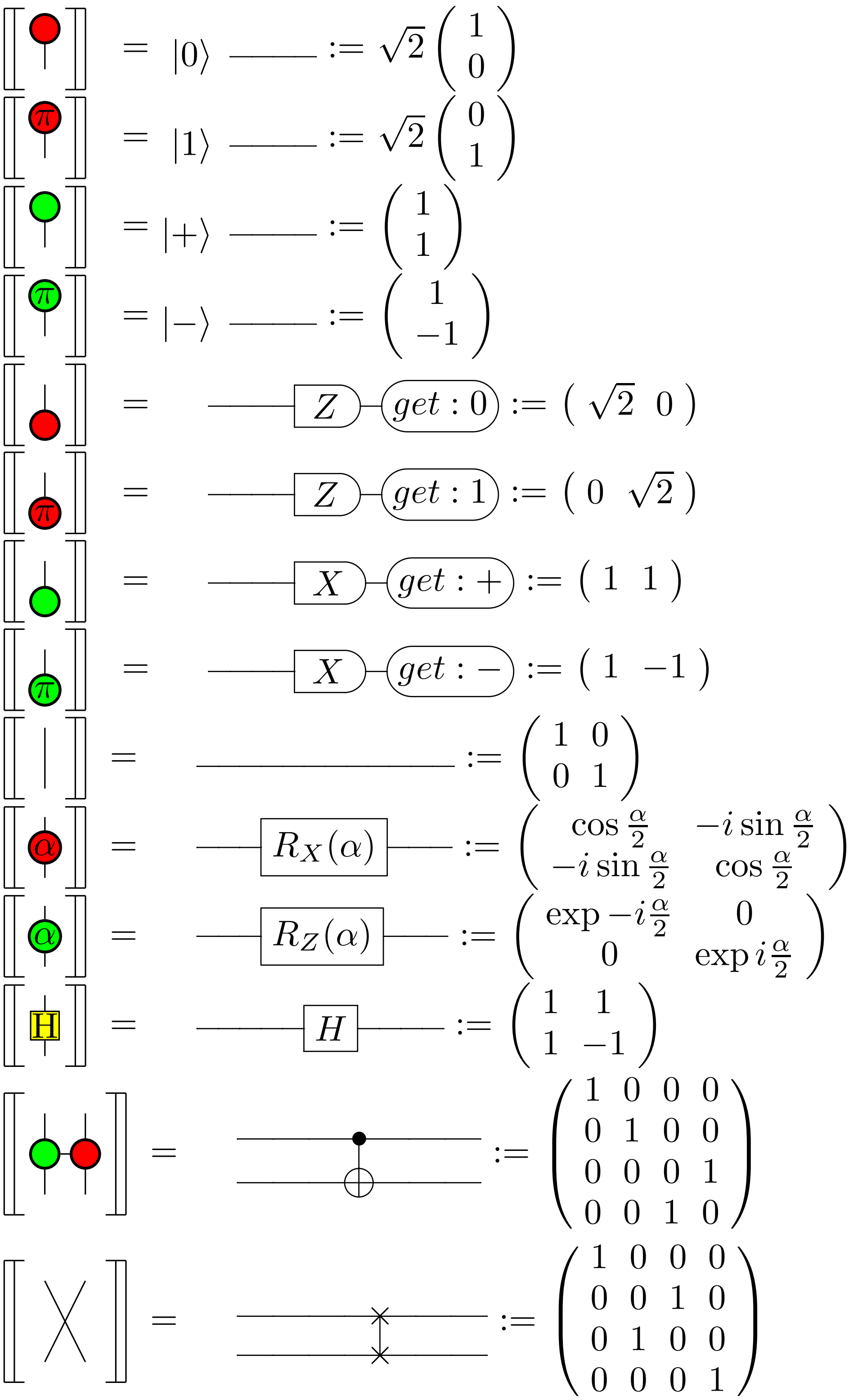


Figure 4.3: Quantum circuit interpretation of the ZX network elements.

ZX network diagrams are logical elements which have no explicit physical meaning and can be modeled in many different ways. Indeed, there are structures that appear in ZX networks but don't have a circuit interpretation. A particular interpretation in terms of quantum circuits can be constructed from the diagrams of the ZX network as shown in Figure 4.3. The ZX network is universal for quantum computation since any quantum circuit can be built in this way.

We know that the ZX network is sound for quantum mechanics: if two diagrams are equal according to the rules of the ZX network then their corresponding quantum circuits are equivalent[82]. Note that the converse is not true: it can be impossible, from the axioms, to show the equality of two ZX network diagrams whose corresponding quantum circuits are equivalent. The ZX network 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[119,170,87].

4.4 Completeness of the ZX calculus

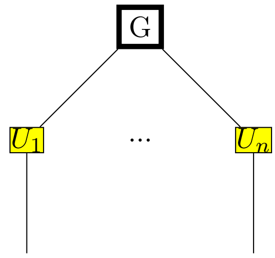
Theorem (Backens)[24]: The ZX network is complete for stabilizer quantum mechanics.

This means that any equation between two ZX network diagrams (put into matrix mechanics) which can be shown to be true using stabilizer quantum mechanics is derivable using the rules of the ZX network. Note that this completeness result only requires the axioms in Figure 4.2 to hold with phases α and β in the set $\{-\pi/2, 0, \pi/2, \pi\}$.

We will present an outline of the proof[24], which uses results on quantum graph states and local Clifford operations[290,129] to bring diagrams into a normal form.

Recall that a graph state $|G\rangle$, where G=(E,V) is a graph of order n with adjacency matrix A, is defined as the eigenstate of all the operators $X_v \otimes \bigotimes_{u \in V} Z_u^{A_{uv}}$ $(\forall v \in V)$. In ZX network diagrams[82,118], this graph state can be represented by a green node with one output for each vertex $v \in V$ and a Hadamard node connected to the green nodes for vertices u,v for each edge $\{u, v\} \in E$.

Definition: A GS-LC (*graph state- local Clifford*) diagram consists of a ZX network graph state representation with arbitrary single-qubit Clifford operators (called *vertex operators*) applied to each output:



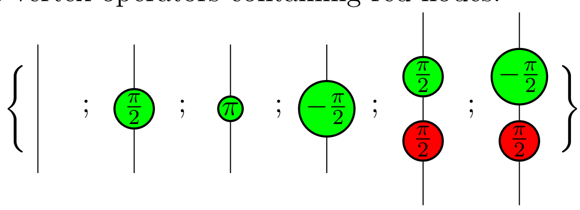
Lemma: Any stabilizer state diagram is equal to some GS-LC diagram within the ZX network.

The proof of this lemma is inspired from a paper[14] showing that stabilizer quantum mechanics can be simulated efficiently on classical computers using a GS-LC representation. It uses the fact that any stabilizer ZX network diagram can be written as a combination of the four green spider diagrams with: (i) a single input, (ii) a single output, (iii) two inputs and an output, (iv) one input and two outputs, as well as the 24 single-qubit Clifford unitaries (depicted using their Euler decompositions). The proof proceeds by induction[24], demonstrating that applying each of the basic components to a GS-LC diagram yields another GS-LC diagram.

In fact, one can strengthen this lemma and show that[24]:

Lemma: Any stabilizer state diagram is equal to some **reduced GS-LC diagram** within the ZX network, where a reduced GS-LC diagram is a GS-LC diagram where:

(i) Two adjacent vertices cannot both have vertex operators containing red nodes.

(ii) All vertex operators belong to the set: 

This proves that there is a non-unique normal form for stabilizer state ZX network diagrams consisting of a graph state diagram and local Clifford operators.

Even though this reduced GS-LC normal form is not unique, there is a straightforward algorithm for testing equality of diagrams given in this form, based on a result for graph states[129].

Definition: A pair of reduced GS-LC diagrams is simplified if there are no pairs (p,q) of qubits, adjacent in at least one of the diagrams, such that p has a red node in its vertex operator in the first diagram but not the second and q has a red node in the second diagram but not the first.

<u>Lemma:</u> Two diagrams making up a *simplified pair of reduced GS-LC diagrams* correspond to the same quantum state if and only if they are identical. Moreover, any pair of reduced GS-LC diagrams can be simplified.

Since the Choi-Jamiolkowski isomorphism preserves equalities, this result extends to diagrams which represent operators and not states. Indeed, we can always use map-state duality to turn pairs of operators into states and then transform these states into simplified pairs of reduced GS-LC diagrams and then apply map-state duality to transform these states back into operators.

This shows that the ZX network is complete for stabilizer quantum mechanics. Note that any unitary single-qubit operator can be approximated to arbitrary accuracy using only Clifford operators and the $T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix}$ operator and that the ZX network for single qubits remains complete upon adding the operator T to the single-qubit stabilizer operations[25].

4.5 Quantum circuits for the ZX network axioms

This section and the next present the formal proof of the result stated in the introduction.

In light of Backens' theorem, the quantum circuit equations corresponding to the axioms of the ZX network will be complete for stabilizer quantum mechanics. First of all, note that directly using Figure 4.3 to convert the ZX network axioms into equations between linear operators does not yield a complete set of equations between quantum circuits since some of the resulting equations between linear operators cannot be expressed as quantum circuit equalities.

Therefore, in order to obtain the desired set of sound and complete circuit equations for stabilizer theory, we need to clarify the relationship between the ZX network and quantum stabilizer circuits. In order to do this formally, we introduce a symmetric monoidal category of stabilizer quantum circuits and show that it is equivalent to the symmetric monoidal category of the ZX network:

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

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

$\mathcal{F}_{SMC}(Circ)$ is a free symmetric monoidal category over the monoidal signature[186]:

$$\mathcal{S} := \{CNOT; SWAP; prepare\,|0\rangle\,; prepare\,|+\rangle\,; postselect\,|0\rangle\,, postselect\,|+\rangle\,, R_x(\alpha); R_z(\beta)\} \tag{4.3}$$

These are the consistuent 'gates' of the symmetric monoidal category, which can be combined using composition and the tensor product.

The axioms for the category $\mathcal{F}_{SMC}(Circ)$, which are quantum circuit equations corresponding directly to the axioms of the ZX network ($\mathcal{F}_{SMC}(ZX)$), are given in Figure 4.5. This gives us a new insight into the structure of the ZX network, namely an understanding of what the axioms of the network mean, in terms of familiar quantum circuits.

This equivalence of categories means that there exists a full, faithful, essentially surjective functor $[[\cdot]] : \mathcal{F}_{SMC}(ZX)/\equiv_{ZX} \to \mathcal{F}_{SMC}(Circ)/\equiv_{Circ}$. For the constructive proof of the existence of this functor, we use the functor $[[\cdot]]$ in Figure 4.3 and check that it is full, faithful and essentially surjective.

In practice, this requires us to find a set of ZX network equations which are equivalent to the axioms of the ZX network ($\equiv_{ZX}$) and are in a form that can be directly related to quantum circuits using Figure 4.3. Such a set of ZX network circuit-like equations is shown in Figure 4.4, in the following section. If we use the quantum circuit equations obtained by applying the functor in Figure 4.3 to the network equations in Figure 4.4 as the axioms $\equiv_{Circ}$ for the category $\mathcal{F}_{SMC}(Circ)$, then $[[\cdot]] : \mathcal{F}_{SMC}(ZX)/\equiv_{ZX} \to \mathcal{F}_{SMC}(Circ)/\equiv_{Circ}$ is full, faithful and essentially surjective by construction.

The next section proves that the set of equations in Figure 4.4 are equivalent to the ZX network axioms. These ZX network equations can be directly related to the axioms $\equiv_{Circ}$ for the category $\mathcal{F}_{SMC}(Circ)$ in Figure 4.5, using the functor in Figure 4.3. Note that the equivalence in this lemma holds for arbitrary phases α and β in the ZX network axioms.

4.6 Proof of the Equivalence Lemma

We will now prove that the set of ZX network equations given in Figure 4.4, which are in a form that can be directly related to quantum circuits using Figure 4.3, are equivalent to the axioms of the ZX network. Note that normalization is not relevant for the proof of completeness so we ignore scalar factors.

Note first of all that the rule (T) of the ZX network states that after enumerating the inputs and outputs of a diagram, any topological deformation of the internal structure will give an equal diagram. A version of the (T) rule can be used as part of the new set of ZX axioms in the form resembling circuit equations. The topological rigidity of quantum circuits, however, means that the complete set of quantum circuit equations will contain several equations for each ZX network rule, one for each possible choice of assignments of inputs and outputs.

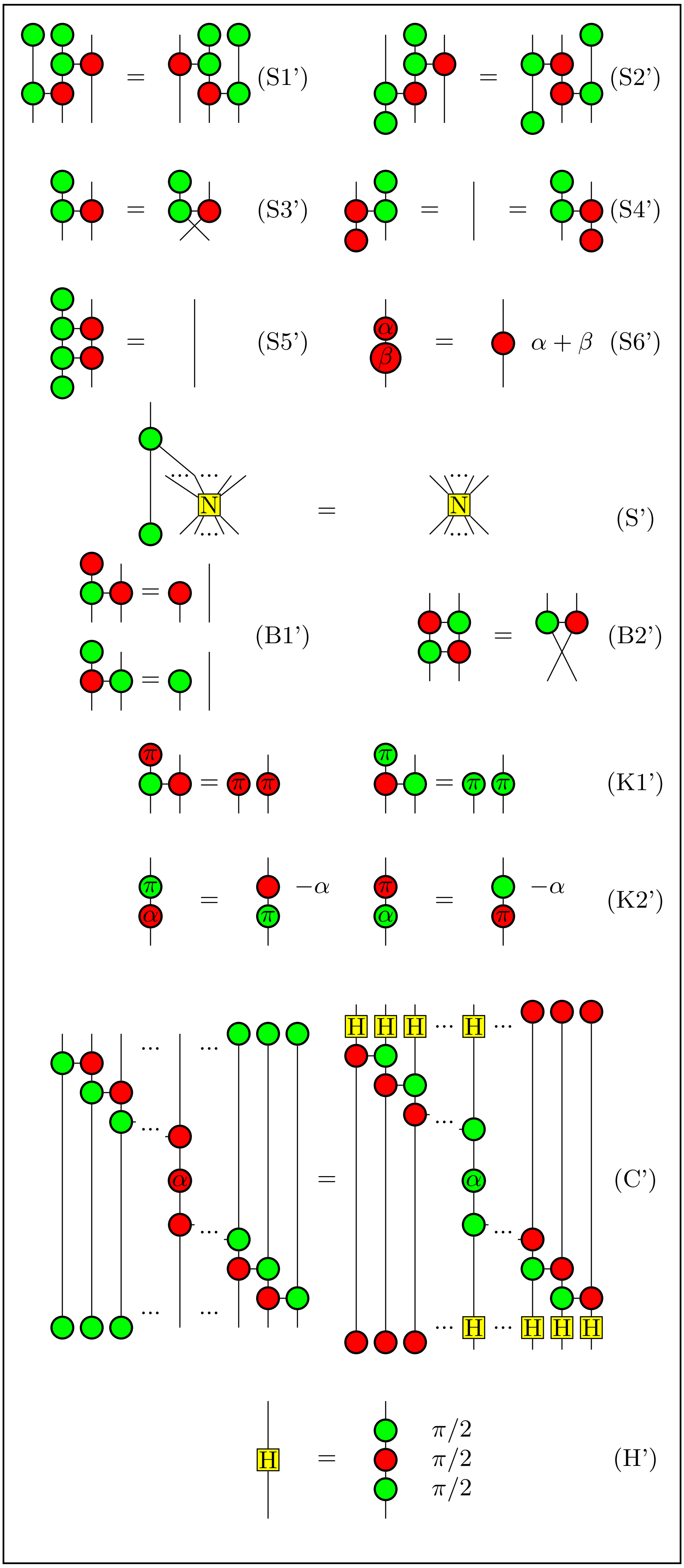
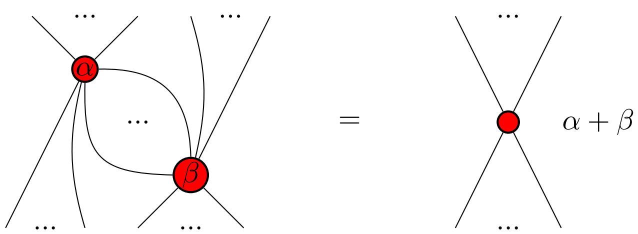
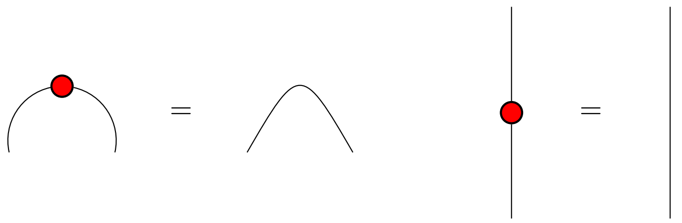


Figure 4.4: Alternative ZX axioms in a form resembling quantum circuit equations.

<u>Lemma A1</u>: The ZX network rules (S1'), (S2'), (S3'), (S4'), (S5'), (S6') and (S') taken together are equivalent to the (S) rules of the ZX network:

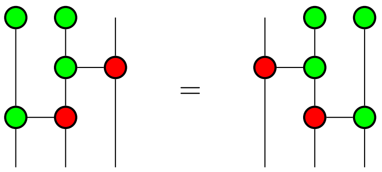


(S1)

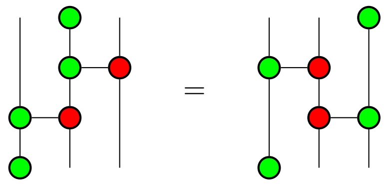


(S2)

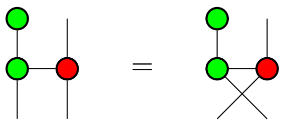
⇔



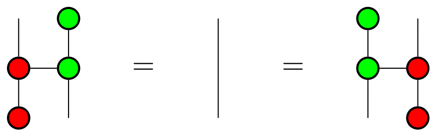
(S1')



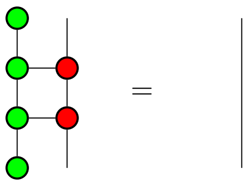
(S2')



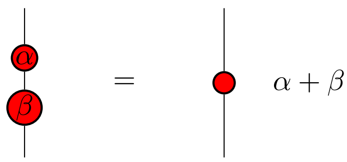
(S3')



(S4')



(S5')

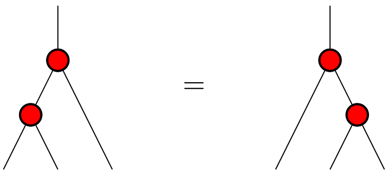


(S6')

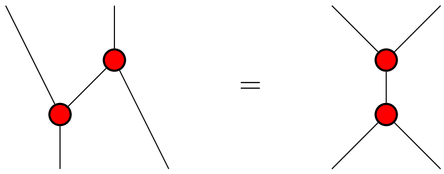
(S')

This equivalence assumes that the (T) rule holds and that the (C) rule holds in one direction.

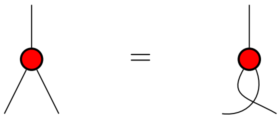
Proof: By theorems 6.11 and 6.12 of[82], we know that (S1) and (S2) are equivalent to:



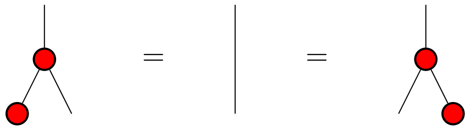
(S1o')



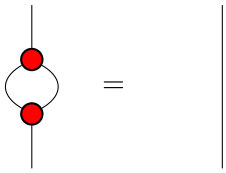
(S2o')



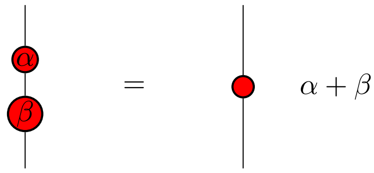
(S3o')



(S4o')

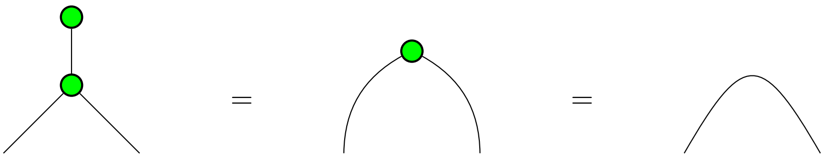


(S5o')



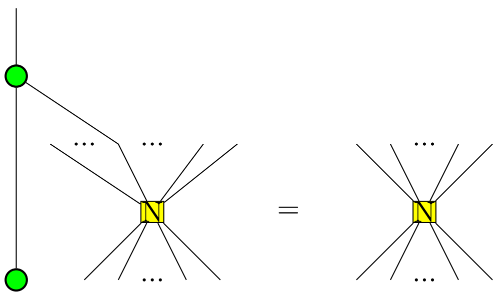
(S6o')

In particular, these equations, together with (T) and (C), imply:



(So')

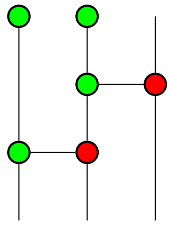
therefore we can assume that (So') holds in one direction of the proof. We now add a rule (S') to the new set of circuit equations which is trivially equivalent to (So'):



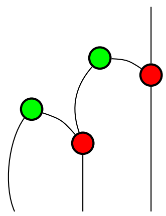
(S')

where the N box is an arbitrary ZX network. Adding (So') to the new set of network equations means that we can now assume that (So') holds in both directions of the proof. Note that we only assume that (C) holds in the proof that:{(S1), (S2)} ⇒ {(S1'), (S2'), (S3'), (S4'), (S5'), (S6') } and not in the other direction.

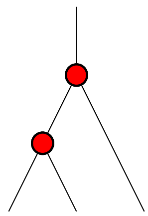
The equation (S6o') is the same as the equation (S6'). If we assume that (So') and (T) hold, then each of the individual equations (S1o'), (S2o'), (S3o') and (S5o'), is equivalent to (S1'), (S2'), (S3'), (S4') and (S5') respectively. For example:



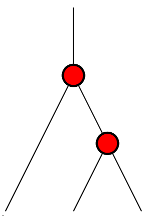
= (T) (So')



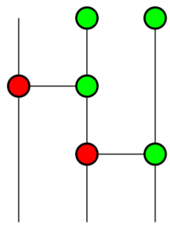
= (T) (So')



= (S1o')



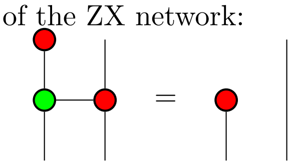
=



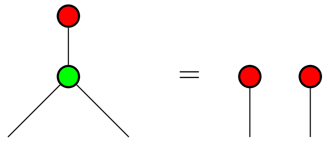
shows that (S1o') is equivalent to (S1'). The other four equivalences follow in the same way, by repeatedly using (So').

The proof of Lemma A1 is the most delicate stage in proving the Equivalence Lemma as it is not trivial to express the diagrammatic spider laws in terms of the rigid structure of quantum circuits. The other three lemmas are more straightforward to prove once the circuit equivalent of the (S) laws are in place.

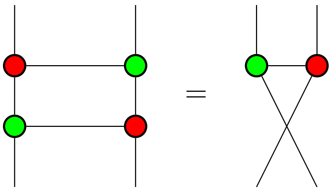
Lemma A2: The ZX network equations (B1') and (B2') are equivalent to the (B) rules of the ZX network:



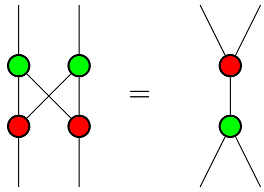
(B1') iff



(B1)



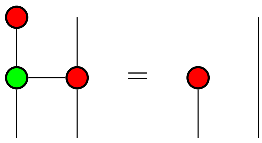
(B2') iff



(B2)

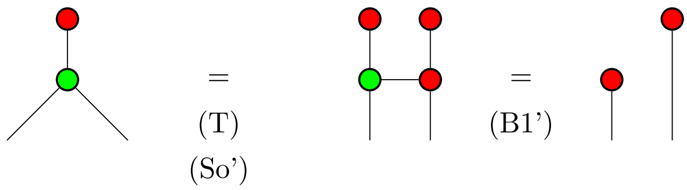
Proof: Note that we assume that the rules (T) and (S) hold, which is not a problem since our goal is to prove the equivalence of the whole set of ZX network equations given in Figure 4.4 with the ZX axioms from Figure 4.2. The proof consists of four steps:

(i) (B1') ⇒ (B1):



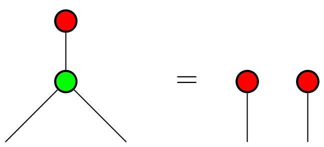
(B1')

⇒



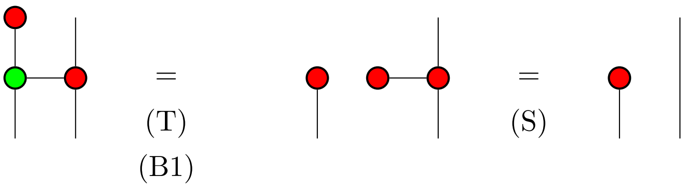
(B1)

(ii) (B1') ⇐ (B1):



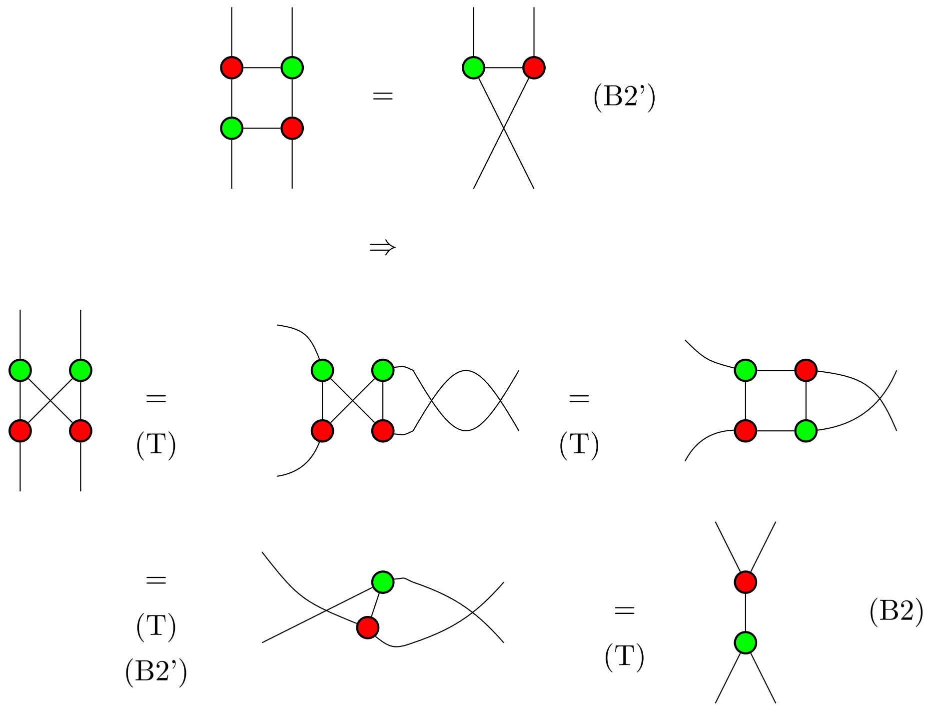
(B1)

⇒

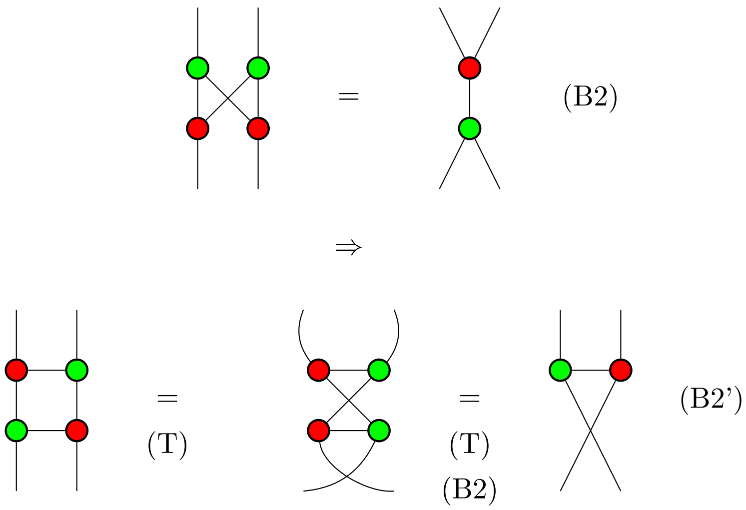


(B1')

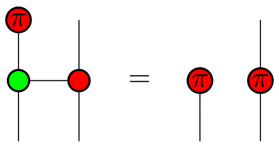
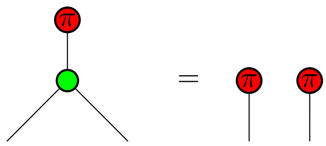
(iii) (B2') $\Rightarrow$ (B2):



(iv) (B2') $\Leftarrow$ (B2):



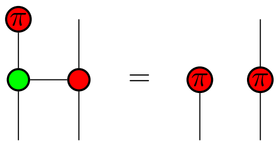
Lemma A3: The ZX network equations (K1') and (K2') are equivalent to the (K) rules of the ZX network:


(K1') iff

(K1)

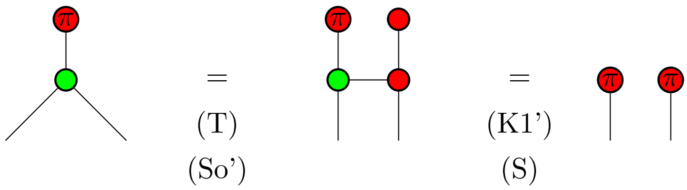
and (K2') is the same as (K2).

Proof: Once again, we assume that the (S) and (T) rules hold. We show the equivalence in two steps:

(i) (K1') (⇒) (K1):

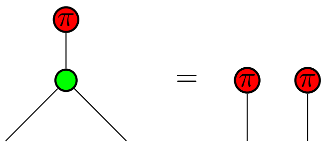

(K1')

⇒

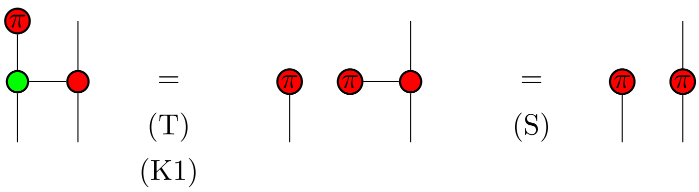


(K1)

(ii) (K1')(⇐) (K1):

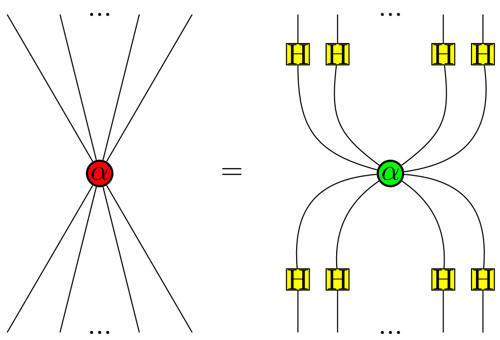

(K1)

⇒



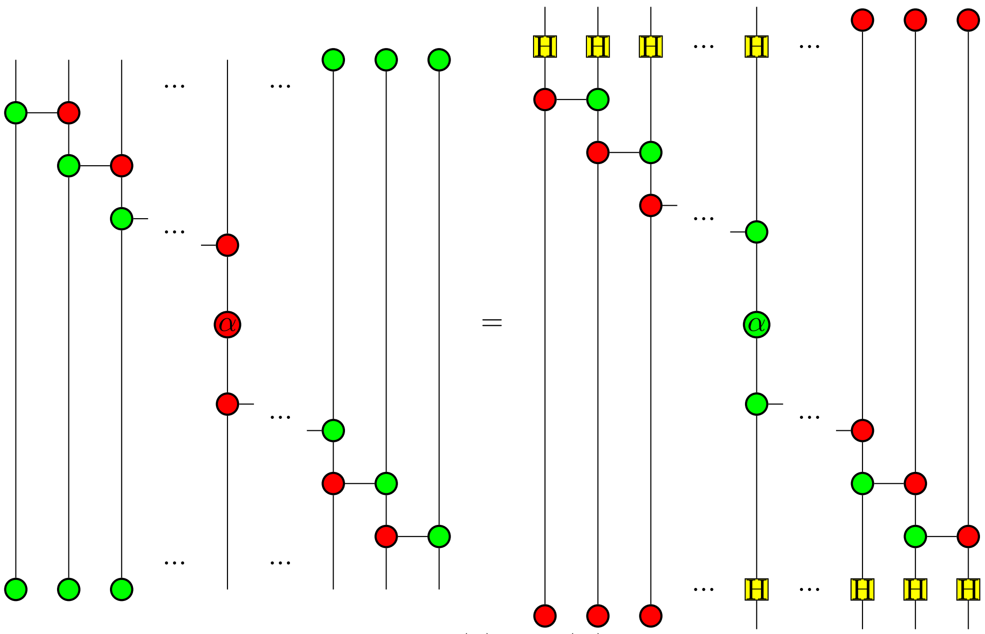
(K1')

Lemma A4: The ZX network equation (C') is equivalent to the (C) rule of the ZX network:



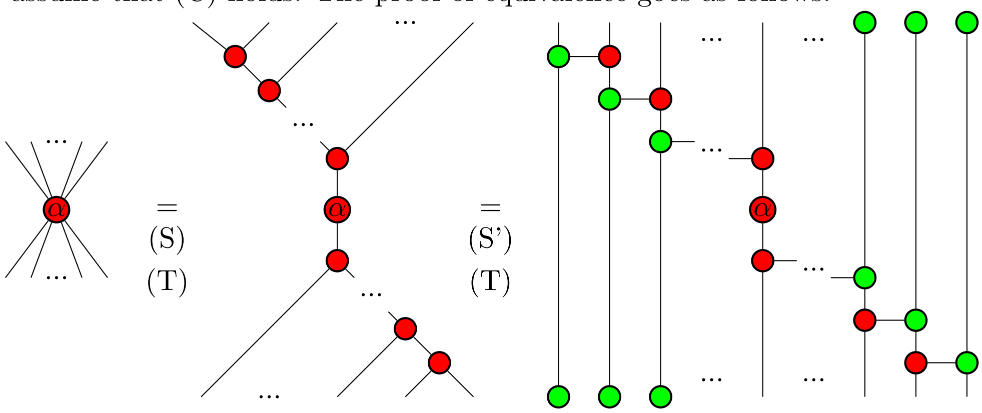
(C)

iff

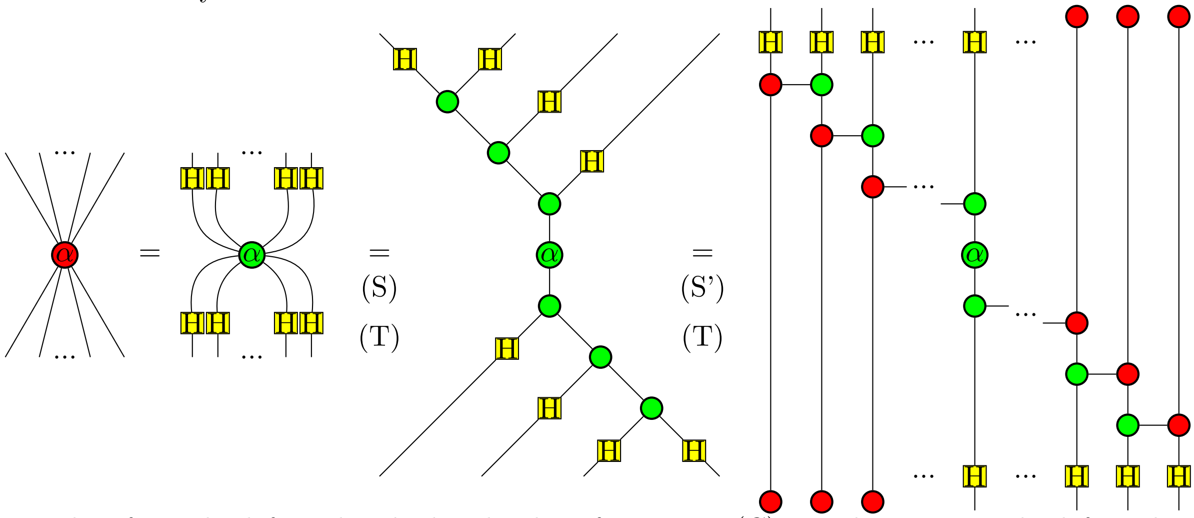


(C')

Proof: Again, we assume that the (S) and (T) rules hold. This is not a problem since the proof that {(S1'), (S2'), (S3'), (S4'), (S5'), (S6') } ⇒ {(S1), (S2)} in Lemma A1 does not assume that (C) holds. The proof of equivalence goes as follows:



and similarly:

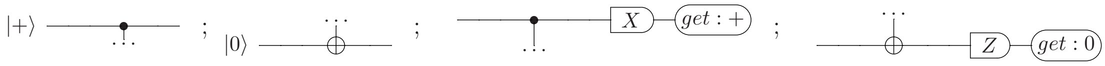


Therefore, the left and right hand sides of equation (C) are the same as the left and right hand sides respectively of equation (C'), which shows that (C) and (C') are equivalent. Note that both (C) and (C') rules include the case where there are no inputs or no outputs. Note that (H') is the same as (H). Lemmas A1-A4 taken together show that the set of ZX network equations given in Figure 4.4, are equivalent to the axioms of the ZX network.

Note that the transition from the alternative ZX network equations we presented here to the circuit equations from the next section can be understood by fixing a grid-like structure for the quantum circuits, enumerating each circuit input and output and considering all the circuit equations that arise from each ZX network equation (including when the colours are reversed).

We expect that the relationship between each ZX network equation from Figure 4.4 and its corresponding set of quantum circuit equations in Figure 4.5 is clear, except for the case of the (S') rule, which we will now explain further. By fixing equation (S') in a grid-like 'circuit' structure and enumerating all of the inputs and outputs, we can see that the (S') rule, when interpreted in the circuit calculus is equivalent to the following rule:

Let us associate a number to each input and output of a quantum circuit Q. If we can obtain a valid quantum circuit Q', whose inputs and outputs (which do not include truncated CNOT lines) are numbered in the same way as Q, by replacing a finite number of times the following quantum circuit fragments:



by wires with the same number as the corresponding input or output (regardless of topological structure), then the circuits Q and Q' are equivalent. The CNOT vertex attached to one of these circuit elements in circuit Q is included in circuit Q'. (Scirc)

By considering all the cases when this rule can arise, we can enumerate all the instances when the quantum circuit fragments in circuits Q being replaced by wires in circuits Q' leads to a valid quantum circuit equation. This leads to the quantum circuit equations which are presented in the (Scirc) rule in Figure 4.5. Note that due to composition and repitition with the other circuit equations, a small number of circuit equations are sufficient.

4.7 A complete set of circuit equations for stabilizer quantum mechanics

The Equivalence Lemma from section 4.5 shows that any quantum circuit equation which, when written in the ZX network, can be shown to be true using the ZX axioms from Figure 4.2, can be shown to be true using the equivalent circuit equations in Figure 4.5.

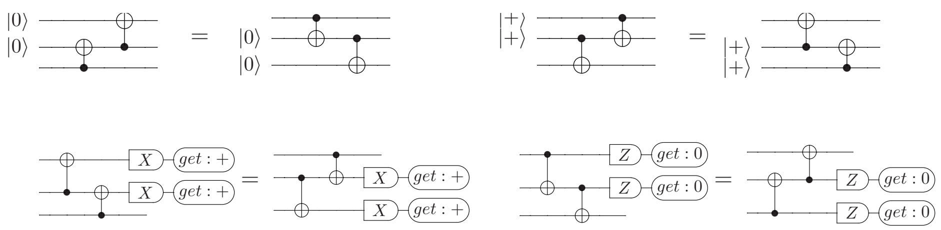
Backens' theorem states that any quantum circuit equation which can be shown to be true using stabilizer quantum mechanics is derivable using the ZX axioms when written as an equation between two ZX network diagrams.

Combining the Equivalence Lemma with the fact that the ZX network is sound for stabilizer quantum mechanics shows that any equation between quantum circuits which can be derived from the circuit equations in Figure 4.5 is in agreement with stabilizer quantum mechanics.

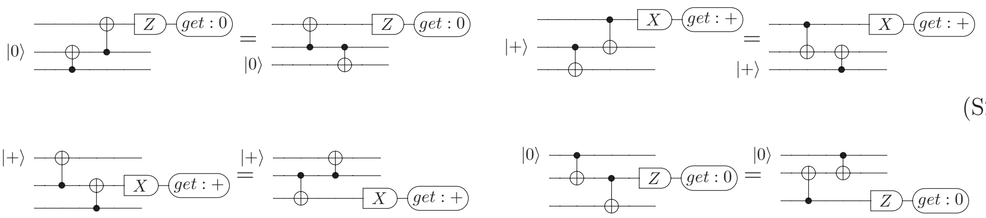
Synthesizing these results yields the main result of this chapter:

Theorem: The set of quantum circuit equations in Figure 4.5 with phases α and β in the set $\{-\pi/2, 0, \pi/2, \pi\}$ is both sound and complete for stabilizer quantum mechanics.

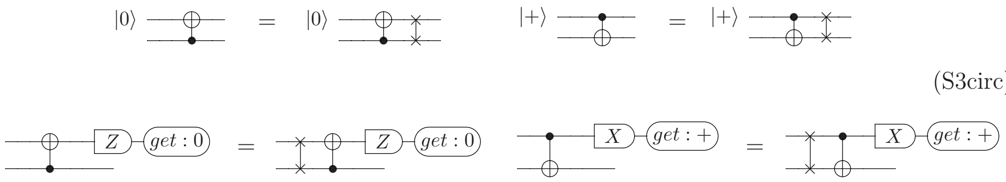
We now present this sound and complete set of quantum circuit equations:



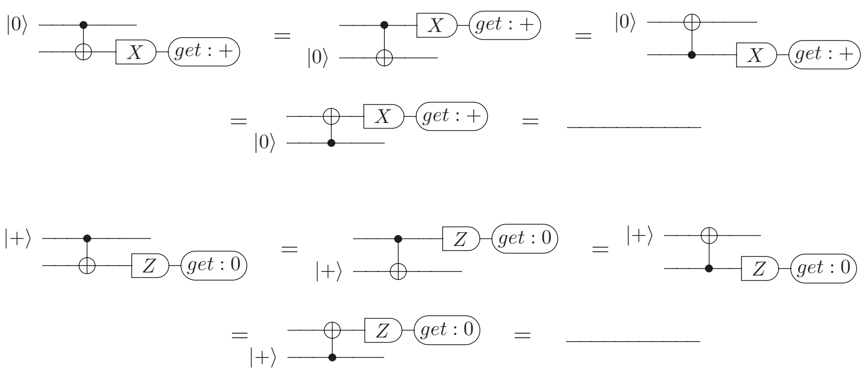
(S1circ)



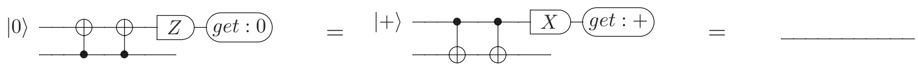
(S2circ)



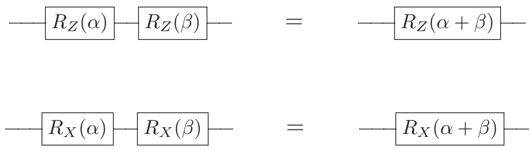
(S3circ)



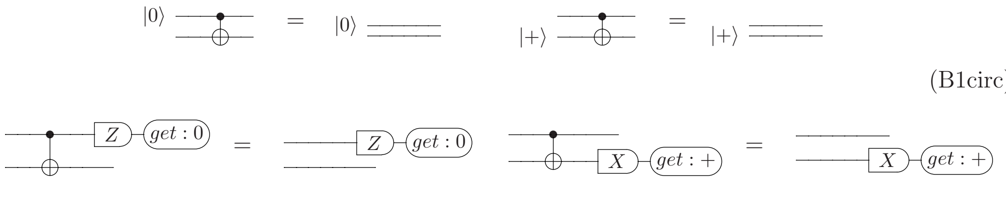
(S4circ)



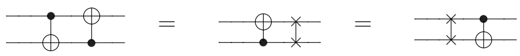
(S5circ)



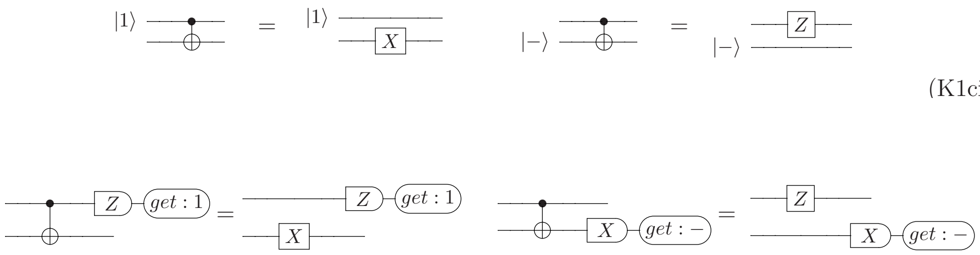
(S6circ)



(B1circ)

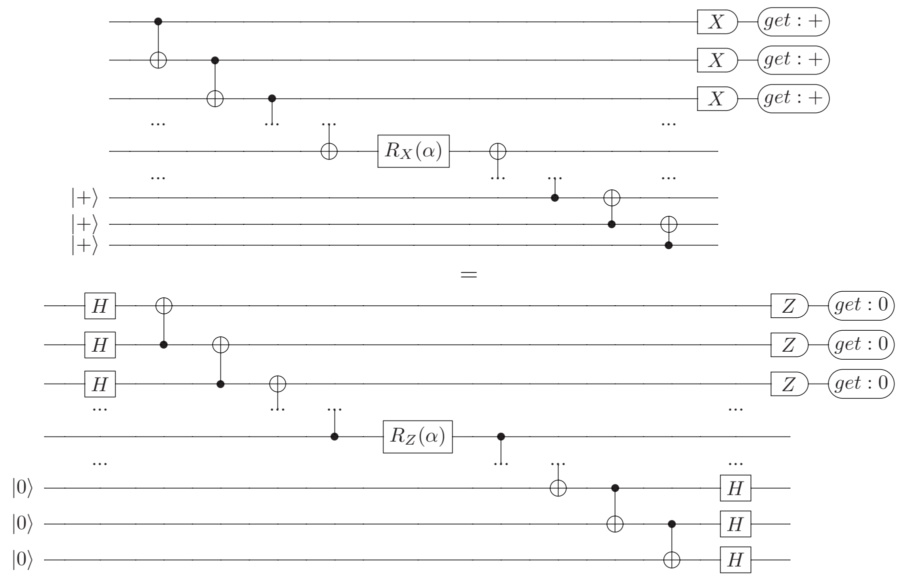


(B2circ)



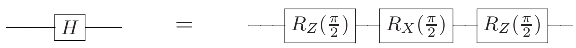
(K1circ)

(K2circ)

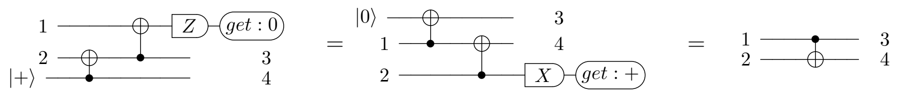


(Ccirc)

Note that this rule also holds if both sides of the (Ccirc) equation above only contain the top/bottom half of the quantum circuit (corresponding to the (C) rule with no inputs/outputs respectively).



(Hcirc)



(Scirc)

Figure 4.5: Sound and complete set of circuit equations for stabilizer quantum mechanics.

Therefore, we have found a **complete set of quantum circuit equations for stabilizer quantum mechanics**. Any circuit equation which can be shown to be true using stabilizer theory—in the sense that both quantum circuits in the equation correspond to equivalent processes in stabilizer quantum mechanics—can be derived from this set.

This provides a novel insight into the logical foundation of the stabilizer formalism.

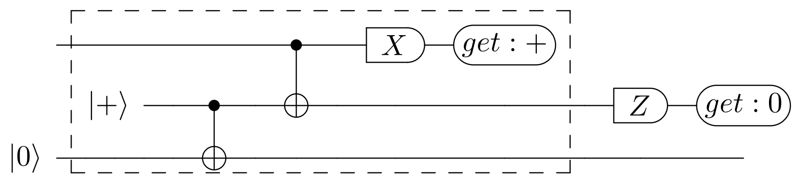
4.8 Derivation of an equation between stabilizer quantum circuits from the complete set

The proof of the result relies heavily upon categorical quantum mechanics. It would have been difficult to find this set of circuits without the flexibility of the ZX network and the theorem may have been difficult to prove without appealing to category theory.

The theorem itself, however, is purely a result about quantum circuits and stabilizer quantum mechanics, which can readily be understood without any knowledge of category theory or formal logic.

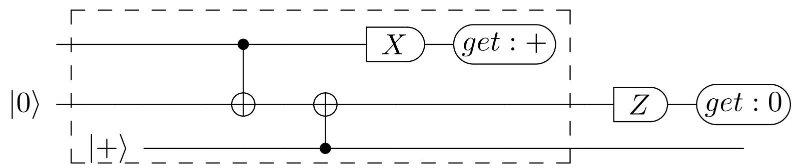
In order to make this clear and provide an illustration of the general result, we now give an example of using the complete set of circuit equations to formally derive a well known equation between stabilizer quantum circuits.

The first quantum circuit of the equation below corresponds to the standard quantum teleportation protocol[50], where a Bell state $|00\rangle + |11\rangle$ is prepared on the second and third qubits and the Bell basis is measured on the first two qubits (the result corresponding to $|00\rangle + |11\rangle$ is post-selected). We use the complete set of circuit equations from Figure 4.5 to show that this is the same quantum process as taking the first qubit to the third qubit:



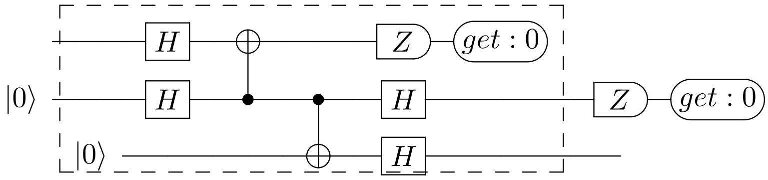
$=$

(S2circ)



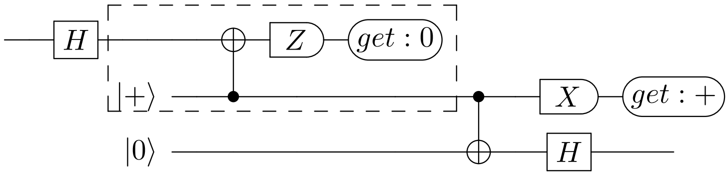
$=$

(Ccirc)



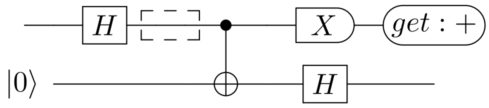
$=$

(Ccirc)



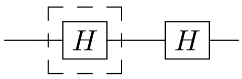
$=$

(S4circ)



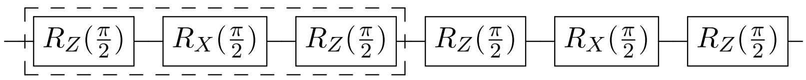
$=$

(S4circ)



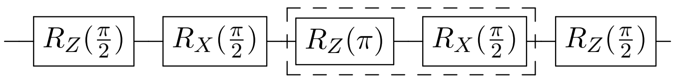
$=$

(Hcirc)



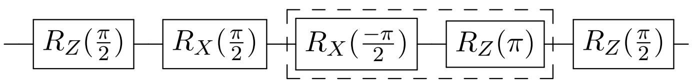
$=$

(S6circ)



$=$

(K2circ)



$=$

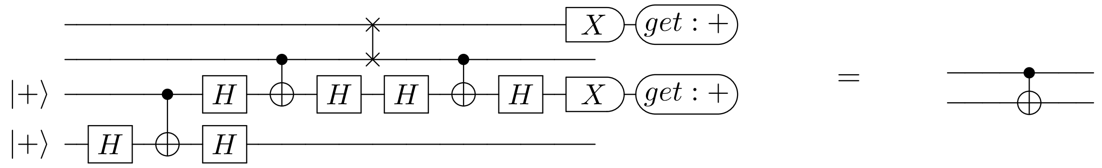
(S6circ),(K2circ)

This is a proof of the validity of quantum teleportation from a set of axioms for quantum stabilizer theory. The dotted boxes indicate a circuit substitution using a circuit equation from Figure 4.5. Any equivalence between two quantum circuits corresponding to the same stabilizer process can be formally shown from the complete set of circuit equations by using this reasoning by substitution.

4.9 Reasoning with the ZX network is easier than using the quantum circuit calculus

A quick comparison of the ZX network axioms from Figure 4.2 with the set of quantum circuit axioms from Figure 4.5 makes it clear that demonstrating the equivalence of quantum processes with the quantum circuit calculus will be far more cumbersome than using the ZX network. For instance, in the previous section, the circuit calculus takes more than 10 steps to prove the validity of the post-selected teleportation protocol, whereas the ZX network can verify validity in a single step.

Now, let us briefly present another example of a derivation which is less trivial using the ZX network. This demonstrates how the flexibility of the spider law allows the ZX network to show validity of a quantum circuit equation far more intuitively and efficiently than the quantum circuit calculus. Both the ZX network and the quantum circuit calculus can prove that the following measurement based quantum computing program computes a CNOT gate:



This only requires a straightforward repeated application of the (S) law and 2 applications of the (C) law using the ZX network[82]. The circuit calculus, however, requires applications of the (Hcirc), (S6circ), (K2circ), (Ccirc), (S2circ), (S3circ) and (Scirc) rules to demonstrate the validity of the previous equation. Therefore, using the circuit calculus to check correctness not only requires a larger total number of axioms to be used but also uses more distinct axioms, whose application is far less intuitive than in the ZX network case.

The examples presented above are circuit equations whose validity can be shown in a small number of steps. For larger circuit equations, we expect the use of the circuit calculus to be unviable. The skeptical reader is challenged to verify the correctness of the 7 qubit Steane code[117] using the circuit calculus instead of the ZX network.

We conclude this section by stressing once again that the elements of the ZX network have no explicit physical meaning. Indeed, the network elements are not restricted to the circuit structure of quantum processes. This mathematical flexibility is at the core of the calculational power of the network calculus relative to the circuit calculus. For example, a primitive circuit element like the CNOT gate is broken down into two abstract elements in the ZX network, corresponding to red and green nodes. These elements obey algebraic rules, some of which have no evident physical interpretation, but which appear to play a fundamental logical role. In contrast, every rule in the circuit calculus has an explicit physical interpretation.

Note that we could find similar completeness results for other process theories by using a similar method to the one presented here. Using a recent completeness result for Spekkens' toy theory[26], for example, we could give a complete set of toy theory process equations, by finding the equations corresponding to the ZX network axioms.

4.10 Conclusion

Studying quantum theory from a logical, computer science perspective has provided an insight into the foundations of stabilizer quantum mechanics. The axiomatic approach presented here yields a representation of the systems and processes of an operational physical theory, together with all the equational laws they obey.

Describing physical processes directly using a logical language may dispense with the need of a more elaborate mathematical description which would require a more refined language and further axioms. Some of this extra structure may be unnecessary and undesirable to fully model an operational physical theory. The introduction of a formal logical system describing physical processes provides a framework which is both perspicuous and parsimonious.

Furthermore, such a formalization of the foundations of physics allows one to rigorously ask certain questions about consistency, soundness and completeness of physical theories. Is it possible to find a consistent, sound and complete set of quantum circuit equations which can prove the validity of any true quantum circuit equation? Are there fundamental incompleteness theorems for the foundations of physics?

In any case, the study of the logical foundation of physical theories is an essential method of testing their validity, especially in realms of nature in which experiments are very difficult or impossible to perform. Logic seems to be the most suited tool to rigorously study the foundations of mathematical theories of nature from a human perspective.

Chapter 5

A periodic table of quantum-like theories

The analysis of physical processes hinges on the use of a synthetic and elegant conceptual framework. The extent to which an abstract theory is considered parsimonious and powerful often relies upon **symmetry**. As Hermann Weyl said:
"Symmetry denotes that sort of concordance of several parts by which they integrate into a whole. Beauty is bound up with symmetry". Symmetry is ubiquitous, both in nature and in human activities such as art, music and architecture[99]. Given our desire to find patterns, it is natural that symmetrical considerations also play a key role in our scientific frameworks.

Figure 5.1: Examples of symmetry in nature: the snowflake, honeycomb lattice and aloe polyphylla.

Even in Ancient Greece, fundamental physical theories were strongly influenced by a desire to emphasize symmetry. Following the discovery that there exist exactly five convex regular polyhedra, later called Platonic Solids, the theory was put forward[239] that these symmetrical shapes can be associated to the classical natural elements (air, water, fire, earth) which combine to form all physical matter. Euclid[168] placed a strong emphasis on constructing the five Platonic Solids, shown in Figure 5.2 and deriving their properties from his geometric axioms.

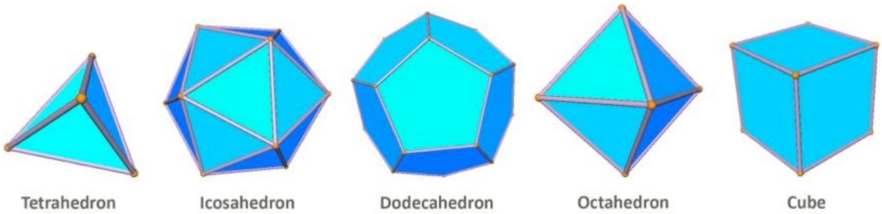


Figure 5.2: The five Platonic Solids.

Symmetry was also a central concern for Kepler when he introduced his laws of planetary motion[183], which were the product of imposing notions of symmetry to the motion of planets around the sun. From these symmetry relations, Newton derived equations of motion[224] which moreover embodied the additional principle of equivalence of inertial frames. The work of Einstein[126] and Noether[227] in the foundations of physics, most notably the derivation of conservation laws and dynamical equations from symmetry principles, further brought symmetry at the forefront. Fundamental symmetries have become the center-piece of modern theoretical physics. The standard model of particle physics, for example, arises from the requirement that physical laws are reference-frame and gauge invariant, meaning that they satisfy global Poincaré symmetry, and local internal SU(3) $\times$ SU(2)$\times$ U(1) gauge symmetry[284].

Furthermore, the language of symmetry provides an excellent tool for efficient **classification**. The search for regularity often leads to a thorough analysis of all possible patterns. For instance, the observation that one can find a number of distinct tessellations, or periodic tiling of a plane using geometric shapes, played an important role in Islamic art. The Alhambra palace in Granada, shown in Figure 5.3, serves as a testimony to the human desire to discover new patterns and contains 17 distinct types of tessellation.

Figure 5.3: Hall of the Abencerrajes in the Alhambra palace.

Formal analysis of the plane symmetry (wallpaper) groups later revealed that these 17 tessellations, depicted in Figure 5.4, fully exhaust all possible periodic tilings of the plane[134,241].

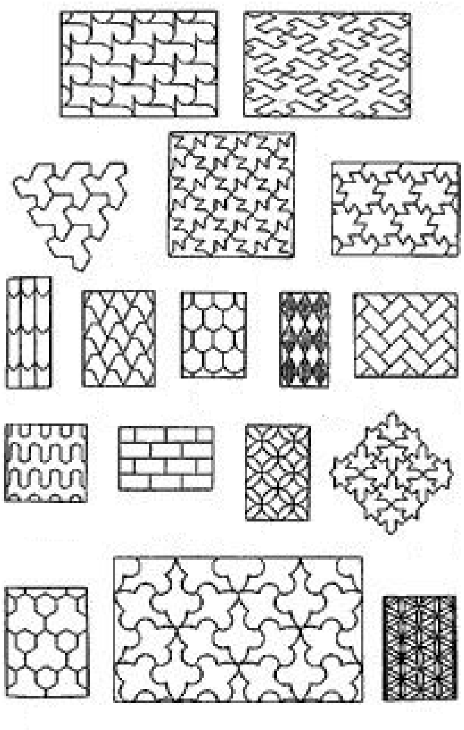


Figure 5.4: Polya's representation of the 17 plane symmetry groups.

A remarkable example of classification arising from the analysis of symmetry is the classification theorem for finite simple groups[98,153], which we presented in Chapter 2. Indeed, this impressive result provides a tangible decomposition of the abstract notion of symmetry, through a classification of the different types of **group**.

In the foundations of physics, we should embrace our desire for elegant theoretical parsimony and ensure that the central role of symmetry is made explicit. In this regard, it is essential to analyze the interplay between group theory and physics, particularly in the study of **alternative physical theories**. We will now focus on this fascinating relationship and study how symmetry can be utilized to extract a classification of physical theories.

5.1 Introduction

An interesting approach to understanding the foundations of quantum mechanics is to study sets of alternative theories which exhibit similar structural or physical features as quantum theory. Several mathematical formalisms for operational physical theories have been proposed[3,35,77] which encompass quantum mechanics as one possible theory within a space of different potential theories. These provide a setting in which we can determine which features are truly particular to quantum theory and which ones are more generic. This approach can pave the way towards novel axiomatizations of quantum mechanics and could yield precious clues about future physical theories which may supersede quantum theory, such as a theory of quantum gravity. As Lewis Carroll aptly put it: "If you don't know where you are going, any road will get you there".

In the previous chapter, we saw that symmetric monoidal categories (SMCs) provide a general framework for physical theories, since they contain two interacting modes, $\otimes$ and $\circ$, of composing systems and processes. Previous work has investigated which additional structure must be imposed on a SMC in order to recover the structure of quantum theory[3]. This approach has yielded the **ZX calculus**, an intuitive graphical language which we introduced in the previous chapter[83]. As we described, the calculus is sound and universal for quantum mechanics and is complete for stabilizer quantum mechanics, given a certain choice of phases[24]. The ZX calculus has proven useful in the study of quantum foundations[92], quantum computation[119] and quantum error-correction[170].

In this chapter, we will sketch a theoretical formalism for analyzing and classifying physical theories that resemble quantum theory. At the core of this framework lies a concern to understand the role of **symmetry** in physics and to use group theory as a tool for classification. We shall build on the description of operational theories through symmetric monoidal categories and isolate a key ingredient, called the **phase group**[83,91]. This allows for the introduction of a **Periodic Table of quantum-like theories**.

The methodology we propose follows five main stages, which will each be presented in some detail. Note that each one of the five levels of analysis of quantum-like theories can be studied independently and that certain physical theories may not admit a description within a given level.

(A) The first stage of analysis provides an explicit presentation of a model for an operational theory. This requires a mathematical representation of preparations, transformations and measurements, as we discussed in Chapter 3. In addition to quantum theory, we also explicitly define two important groups of quantum-like operational theories, *stabilizer quantum theory for qudits*[156] and *Spekkens-Schreiber's toy theory for dits*[262]. This initial level of description is the most familiar to physicists.

(B) The second stage involves a category theoretic description of operational physical theories. This requires us to define symmetric monoidal categories, which furnish an abstract and unified definition of preparations, transformations and measurements. For this purpose, we generalize the ZX calculus to qudit systems and show that the resulting calculus is universal for quantum mechanics. We utilize this calculus as a pictorial tool to depict quantum-like theories and we define the notion of a **mutually unbiased qudit theory** (MUQT), which can be represented by a symmetric monoidal category whose observable structures are all mutually unbiased.

(C) The third stage of analysis involves classifying MUQTs in terms of a particular Abelian group, called the **phase group**. This approach aims to give symmetry a central role in the study of physical theories. Previous work has shown that in the case of qubits[91], there are essentially two MUQTs: stabilizer quantum mechanics[155], which has phase group $\mathbb{Z}_4$, and Spekken's toy theory for bits[276], which has phase group $\mathbb{Z}_2 \times \mathbb{Z}_2$. Furthermore, the phase groups of these theories determine whether or not they admit a local hidden variable model. We aim to generalize this work to higher dimensional systems. In particular, we focus on two interesting families of MUQTs, corresponding to *stabilizer quantum theory for qudits* [156] and *Spekkens-Schreiber's toy theory for dits* [262] and provide a novel proof that these theories are operationally equivalent in three dimensions. This is a first step towards a **Periodic Table of quantum-like theories**, where physical theories can be classified according to their phase groups.

(D) The final stage of analysis briefly outlines a way to generalize the ontological models of quantum mechanics, which were described in Chapter 3, to ontological models for operational theories. We allow ontic spaces which are no longer restricted to measure spaces but can be more intricate mathematical objects. We discuss the idea of topological ontic models and categorical ontic models.

5.2 Explicit models of theories

The standard operational presentation of a physical theory involves associating separate mathematical objects to preparation, transformation and measurement procedures and describing how these mathematical objects relate to each other. The typical example of such an explicit model is the operational presentation of quantum theory. As we discussed in Chapter 3, quantum preparation, transformation and measurement processes are associated with trace one positive density operators acting on Hilbert spaces, completely positive trace non-decreasing maps and positive operator valued measures respectively. The axioms of quantum mechanics then aim to make the relationship between these three mathematical objects explicit.

Note that it is not necessarily possible to always describe operational physical theories in terms of mathematical models which are as concrete and clear-cut as this presentation of quantum theory. Other examples of explicit models of physical theories consist of Spekkens' toy theory, presented in Chapter 3, and stabilizer quantum mechanics, described in Chapter 4. We will now introduce explicit models for two families of quantum-like theories.

5.2.1 Qudit stabilizer quantum mechanics

We describe the generalization of qubit stabilizer quantum mechanics [155] to quantum systems of dimension D, where D can be higher than 2 [156]. Stabilizer states are eigenstates with eigenvalue 1 of each operator in a subgroup of the generalized Pauli group of operators acting on the Hilbert space of n qudits:

$$\mathcal{P}_{D,n} := \{\sqrt{\eta}^{\lambda} g_1 \otimes ... \otimes g_n : \eta = e^{\frac{2\pi i}{D}} \wedge \lambda \in \mathbb{Z}_{2D}\} \tag{5.1}$$

with: $g_k = X^{x_k} Z^{z_k}$ and $x_k, z_k \in \mathbb{Z}_D; \forall k \in \{1, ..., n\}$. Note that sums and multiplication are all modulo D and $\mathbb{Z}_D$ are integers modulo D.

The single qudit Z and X operators are:

$$Z = \sum_{j=0}^{D-1} \eta^j \left|j\right\rangle \left\langle j\right| \quad \text{and} \quad X = \sum_{j=0}^{D-1} \left|j\right\rangle \left\langle j+1\right| \tag{5.2}$$

One can easily see that: $XZ = \eta ZX$ and $Z^D = X^D = \mathbb{I}$.

The generalized Clifford group on n qudits consists of the unitary operations that leave Pauli operators invariant under conjugation:

$$C_n := \{U : UgU^{\dagger} \in \mathcal{P}_{D,n}, \forall g \in \mathcal{P}_{D,n}\} \tag{5.3}$$

The following gates are generalizations of standard qubit gates to higher dimensions [144].

The generalization of the Hadamard gate is the Fourier gate: $F := \frac{1}{\sqrt{D}} \sum_{j,k=0}^{D-1} \eta^{jk} \left|j\right\rangle \left\langle k\right|$. Another important set of qudit gates are the multiplicative gates: $S_q := \sum_{j=0}^{D-1} \left|j\right\rangle \left\langle jq\right|$, where $q \in \mathbb{Z}_D$ such that $\exists \bar{q} \in \mathbb{Z}_D$ with $q\bar{q} = 1$.

We define the qudit controlled NOT and controlled phase gates between control qudit **a** and target qudit **b** as:

$$CNOT_{a,b} := \sum_{j,k=0}^{D-1} \left|k\right\rangle \left\langle j\right|_a \otimes \left|k\right\rangle \left\langle k+j\right|_b \quad \text{and} \quad CP_{a,b} := \sum_{j,k=0}^{D-1} \eta^{jk} \left|j\right\rangle \left\langle j\right|_a \otimes \left|k\right\rangle \left\langle k\right|_b \tag{5.4}$$

The swap gate is: $SWAP_{a,b} := \sum_{j,k=0}^{D-1} |k\rangle \langle j|_a \otimes |j\rangle \langle k|_b$. Note that the SWAP gate can be decomposed as:

$$SWAP_{a,b} = CNOT_{a,b} CNOT_{b,a}^{\dagger} CNOT_{a,b}(F_a^2 \otimes \mathbb{I}_b) \tag{5.5}$$

Similarly to the qubit case, the controlled phase gate can be decomposed as:

$$CP_{a,b} = (\mathbb{I}_a \otimes F_b)^{\dagger} CNOT_{a,b}(\mathbb{I}_a \otimes F_b) \tag{5.6}$$

The generalized Clifford group is generated[144,171] by the set of three gates: $\{F, S_q, CNOT_{a,b}\}$.

Stabilizer quantum mechanics for qudits[156] includes state preparations in the computational basis $\{|0\rangle, |1\rangle, |2\rangle, ...\}$, generalized Clifford unitaries and measurements of observables in the generalized Pauli group. In addition to its foundational importance, the theory of qudit stabilizer quantum mechanics plays a key role in quantum information theory, in quantum key distribution and in quantum error correction.

Extending the Gottesman-Knill theorem shows that qudit stabilizer quantum mechanics can be efficiently simulated by a classical computer. Indeed, a group of order K has at most log(K) generators therefore the qudit stabilizer group can be compactly described using the group generators. One can show that if D is prime then any n-dimensional stabilizer group can be described using at most n generators[156]. In composite dimensions one can have more than n generators but no more than 2n[144].

5.2.2 Spekkens toy theory in higher dimensions

Previous work in quantum foundations[276,37,262] has shown that considering a classical statistical theory together with a fundamental restriction on the allowed statistical distributions over phase space allows one to reproduce a large part of operational quantum mechanics. We will now introduce some of this work for physical systems with discrete degrees of freedom[262]. We call the theory described here **Spekkens-Schreiber toy theory for dits**.

Let phase space $\Omega = (\mathbb{Z}_d)^{2n}$ consist of a set of points (ontic states):

$$m \equiv (x_1, p_1, ..., x_n, p_n) \in \Omega \tag{5.7}$$

We can then define functionals on phase space $F : \Omega \to \mathbb{Z}_d$ and a Poisson bracket of functionals:

$$\{F,G\}(m) := \sum_{j=1}^{n}(F[m+e_{x_j}]-F[m])(G[m+e_{p_j}]-G[m])-(F[m+e_{p_j}]-F[m])(G[m+e_{x_j}]-G[m]) \tag{5.8}$$

where e_{x_j} and e_{p_j} have a 1 in position x_j and p_j respectively and zeros everywhere else. We define canonical variables as the linear functionals:

$$\begin{aligned} F =& a_1X_1 + b_1P_1 + ... + a_nX_n + b_nP_n \\ G =& c_1X_1 + d_1P_1 + ... + c_nX_n + d_nP_n \end{aligned} \tag{5.9}$$

where $X_k(m) = x_k$, $P_k(m) = p_k$ and $a_j, b_j, c_j, d_j \in \mathbb{Z}_d$, $\forall j \in \{1, ..., n\}$.

These form the dual space $\Omega^\star \equiv (\mathbb{Z}_d)^{2n}$ such that: $F = (a_1, b_1, ..., a_n, b_n), G = (c_1, d_1, ..., c_n, d_n) \in \Omega^\star$. We can then write the Poisson bracket of canonical variables as a symplectic inner product of vectors:

$$\{F,G\}(m) = \sum_{j=1}^{n}(a_jd_j - b_jc_j) = F^TJG \tag{5.10}$$

where:

$$J = \bigoplus_{k=1}^{n} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{5.11}$$

We then define the principle of **classical complementarity** in the following way: an observer can only have knowledge of the values of a commuting set of canonical variables (whose Poisson brackets all vanish) and is maximally ignorant otherwise.

The Spekkens-Schreiber toy theory for dits can then be described in the following way:

(a) *Valid epistemic states* are specified by isotropic subspaces $V \subseteq \Omega^\star$, such that $\{F,G\} = 0$; $\forall F, G \in V$, together with a valuation vector $v : V \to \mathbb{Z}_d$ ($v \in V^\star$) such that: $v(F) = F^Tv$; $\forall F \in V$. Therefore, V specifies which set of canonical variables are known and v describes what is known about them. Note the analogy with the commuting set of eigen-operators of the quantum state, together with their eigenvalues.

Epistemic states can also be characterized by a probability distribution over phase space

Ω. We can define the orthogonal complement of V as:

$$V^{\perp} := \{m \in \Omega | P_V m = 0\} \tag{5.12}$$

where P_V is the projector onto V. Note that the phase space points $m \in \Omega$ which are consistent with an epistemic state associated to the isotropic subspace V and valuation vector v are those which satisfy:

$$F^T m = F^T v, \; \forall F \in V \tag{5.13}$$

Therefore, the probability distribution for the epistemic state associated to the isotropic subspace V and valuation vector v is: $p_{V,v} : \Omega \to [0,1]$ such that:

$$p_{V,v}(m) = \frac{1}{|V^{\perp}|}\delta_{V^{\perp}+v}(m) \tag{5.14}$$

where $|V^{\perp}|$ is the cardinality of $V^{\perp}$ and $\delta_{V^{\perp}+v}(m)$ is 1 if $m \in V^{\perp} + v$ and zero otherwise.

(b) *Valid reversible transformations* correspond to all the symplectic, affine transformations (analogues of the Clifford operations). These are the phase space maps $C : \Omega \to \Omega$ such that: $C(m) = Sm + a$ where $a \in \Omega$ and $\{Su, Sv\} = \{u, v\}, \forall u, v \in (\mathbb{Z}_d)^{2n}$.

(c) *Valid measurements* are described by sets of indicator functions $\xi_k : \Omega \to [0,1]$ such that $\sum_k \xi_k = u$ (where u is a function mapping every point of phase space to 1) which correspond to some choice of a set of non-conjugate variables. The outcome probability can then be obtained by:

$$p_k = \sum_{\lambda} v(\lambda)\xi_k(\lambda) \tag{5.15}$$

where $v(\lambda)$ is the epistemic state.

The Spekkens-Schreiber theory, for any number of dits of any dimension, can be represented using matrices to describe the valid epistemic states, transformations and measurements. This corresponds to the subcategory of **FRel** which we will describe below.

Note that Spekkens toy model for bits [276] is a special instance of Spekkens-Schreiber theory for dimension 2 and that the 'knowledge balance principle' is superseded by the principle of classical complementarity described above.

5.3 Depicting qudit quantum mechanics and toy models

The development of categorical quantum mechanics has introduced the idea of describing operational theories by symmetrical monoidal categories representing preparation, transformation and measurement processes. As we saw in Chapter 2, we can use a dagger compact symmetric monoidal category $\mathcal{C}$ to define:

(i) Processes as arrows $\psi : I \to A$, where $A, I \in OBJ(\mathcal{C})$ and I is an initial object

(ii) Transformations as arrows $T : A \to B$ where $A, B \in OBJ(\mathcal{C})$

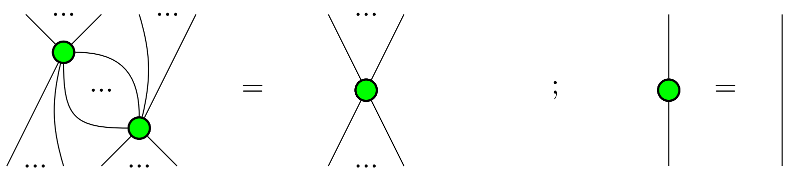
(iii) Measurements using observable structures which generalize linear algebraic measurement bases.

This abstract categorical characterization and the corresponding diagrammatic representation is at the heart of the ZX calculus[82,24,251], that we described in the previous chapter and provides a second level of analysis of quantum-like theories. We will now present a generalization of the ZX calculus to higher dimensional systems.

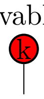
5.3.1 Derivation of the qudit ZX calculus

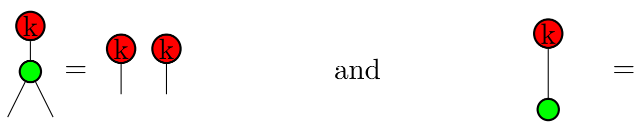
Chapter 2 introduced dagger compact symmetric monoidal categories and how these can be depicted using a formal graphical calculus[181].

Recall that **observable structures**, which are †-special commutative Frobenius algebras, can be defined through the spider laws[85] depicted below.



In **FHilb**, the category of finite dimensional Hilbert spaces, orthonormal bases are in a one to one correspondence with observable structures[86].

Let (A, δ, ϵ) be an observable structure. We can define a classical point as a self-conjugate morphism k: $I \to A$  obeying:



This means that classical points are those which get copied by the copying map and deleted by the deleting map. In **FHilb**, for example, they are the basis states corresponding to the observable structure.

We will now introduce a notion of phase relative to a given basis[82] which allows us to study unbiasedness and the interplay between several bases.

Let (A, δ, ϵ) be an observable structure. For any two points α, β: $I \to A$, we define a multiplication operation:

$$\alpha \odot \beta = \delta^\dagger \circ (\alpha \otimes \beta) \circ \lambda_I \tag{5.16}$$

Note that this multiplication on points is commutative, associative and $\epsilon^\dagger \odot \alpha = \alpha$ for any point α.

A point $\alpha : I \to A$ is called *unbiased* relative to an observable structure (A, δ, ϵ) if there exists a scalar s: $I \to I$ such that: $s.\alpha \odot \alpha^\star = \epsilon^\dagger$. This is a generalization of the usual definition of an unbiased vector with respect to a basis.

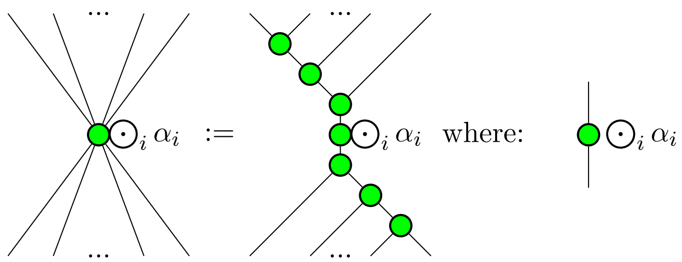
For each state and observable structure (A, δ, ϵ), we introduce a *phase map* Λ which takes each point α: $I \to A$, to the morphism:

$$\Lambda(\alpha) = \delta^\dagger \circ (\alpha \otimes 1_A) : A \to A \tag{5.17}$$

The phase map satisfies several properties:
(i) $\Lambda(\alpha \odot \beta) = \Lambda(\alpha) \circ \Lambda(\beta)$ (ii) $\Lambda(\epsilon^\dagger) = 1_A$ (iii) $\Lambda(\alpha^\star) = \Lambda(\alpha)^\dagger$ (iv) Phase maps commute freely with the observable structure since: $\Lambda(\alpha) \circ \delta^\dagger = \delta^\dagger \circ (1_A \otimes \Lambda(\alpha)) = \delta^\dagger \circ (\Lambda(\alpha) \otimes 1_A)$.

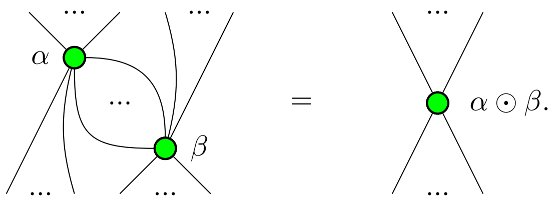
We can extend the spider laws to account for phases relative to an observable structure.

<u>Theorem 3.1:</u> Any morphism $A^{\otimes n} \to A^{\otimes m}$ generated from an observable structure (A, δ, ϵ), together with one occurrence of each unbiased point $\alpha_i : I \to A$ can be written in the form[82]:



is the phase map $\Lambda(\bigodot_i \alpha_i)$.

These spider maps compose according to the generalized spider law:



This theorem follows from the spider laws together with the fact that the phase maps commute freely with the observable structure (property (iv) given above).

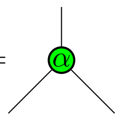
Let α: $I \to A$ be a point satisfying $\alpha^\dagger \circ \alpha = dim(A)$. Then α is unbiased iff $\Lambda(\alpha)$ is unitary.

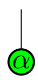
Note that the choice of $\alpha^\dagger \circ \alpha = dim(A)$ is taken for unbiased points from this point onwards.

All the points which are unbiased with respect to the basis corresponding to the observable structure (A, δ, ϵ) form an Abelian group $\mathcal{U}$ with respect to the multiplication $\odot$. This is clear since $\odot$ is closed for unbiased points, commutative, associative, admits the unique identity point $\epsilon^\dagger$ and each point has a unique inverse, its conjugate.

The phase maps, restricted to act on unbiased points relative to the observable structure, form an abelian group with map composition as the group operation, which is isomorphic to $\mathcal{U}$. We call this the ***phase group*** Π.

Note that, for each unbiased point α we can define a new observable structure (A, δ_α, ϵ_α) where:

$\delta_\alpha=$ 

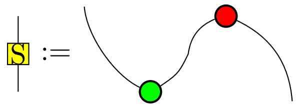
$:= (\Lambda(\alpha) \otimes \Lambda(\alpha)) \circ \delta \circ \Lambda(\alpha)^\dagger$ and $\epsilon_\alpha=$ 

$:= \epsilon \circ \Lambda(\alpha)^\dagger$.

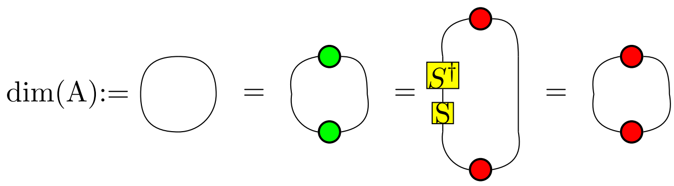
We use the phase group for observable structures as a tool to study physical theories from an abstract algebraic perspective.

We now proceed to study how two complementary observable structures interact[82]. In general, we cannot assume that the dagger compact structures of two distinct observable structures coincide[89].

Therefore, we define the *dualizer* of observable structures $(A, \delta_Z, \epsilon_Z)$ and $(A, \delta_X, \epsilon_X)$ as:



By the spider laws for red and green, we can see that the dualizer is unitary. This shows that the dimension of a dagger symmetric monoidal category does not depend on the choice of observable structure[82] since:



In a Hilbert space of D dimensions, two orthonormal bases $\{u_1, u_2, ..., u_D\}$ and $\{v_1, v_2, ..., v_D\}$ are called *unbiased* if:

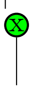
$$D\,|\langle v_i, u_j\rangle|^2 = 1,\ \ \forall i, j \in \{1, 2, ..., D\} \tag{5.18}$$

If a quantum system is prepared in a state corresponding to a vector in one of these bases, then all the outcomes of a measurement, with respect to the other mutually unbiased basis, will occur with equal probabilities. No information can be retrieved by performing such a measurement. In this sense, two mutually unbiased bases corresponding to eigenstates of two non-degenerate quantum observables describe mutually exclusive physical measurement procedures.

This provides a mathematical expression for Bohr's principle of complementarity 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 exhausts the possible information about the objects".

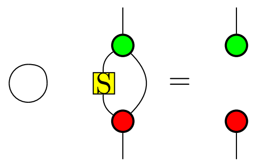
Two observable structures $(A, \delta_Z, \epsilon_Z)$ and $(A, \delta_X, \epsilon_X)$ are called ***complementary*** if:

(i) whenever a point z: $I \to A$  is classical for (δ_Z, ϵ_Z), it is unbiased for (δ_X, ϵ_X)).

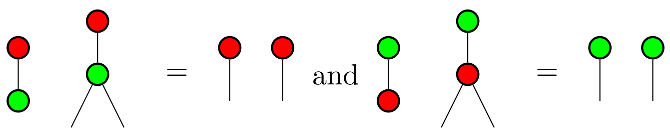
(ii) whenever a point x: $I \to A$  is classical for (δ_X, ϵ_X), it is unbiased for (δ_Z, ϵ_Z)).

This definition of complementarity could easily be generalized to more than two observable structures by requiring that whenever a point is classical for one observable structure, it must be unbiased for all the other observable structures. One can show the following theorem[82], assuming that at least one of the observable structures describes a basis:

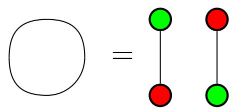
Theorem 3.2: Two observable structures are complementary iff they obey:



Two observable structures $(A, \delta_Z, \epsilon_Z)$ and $(A, \delta_X, \epsilon_X)$ are called *coherent* if the erasing point for each observable structure is a classical point for the other observable structure. This can be pictured as:



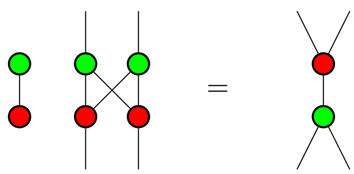
In line with qudit quantum theory, states and erasing points are defined such that:



The classical points K_Z of an observable structure $\{A, \delta_Z, \epsilon_Z\}$ are called *closed* for an observable structure $\{A, \delta_X, \epsilon_X\}$ if, for all $k, k' \in K_Z$, we have $k \odot_X k' = \delta_X \circ (k \otimes k') \in K_Z$.

One can easily show[82] that for every Hilbert space we can find a pair of coherent, closed, complementary observable structures. In fact, the observable structures corresponding to the Z and X qudit operators are closed, coherent and complementary.

Two observables structures $\{A, \delta_Z, \epsilon_Z\}$ and $\{A, \delta_X, \epsilon_X\}$ are said to be ***strongly complementary*** if:



This condition is called strong complementarity since[82]:

Theorem 3.3: A pair of coherent, strongly complementary observable structures are complementary.

One can show that[82]:

Theorem 3.4: If $\{A, \delta_Z, \epsilon_Z\}$ and $\{A, \delta_X, \epsilon_X\}$ are coherent strongly complementary observable structures and the set K_X of classical points for $\{A, \delta_X, \epsilon_X\}$ is finite, then K_X is a subgroup of the group $(U_Z, \odot_Z)$ of unbiased points for $\{A, \delta_Z, \epsilon_Z\}$.

The ZX calculus for qubits is restricted to two dimensions. However, since this algebraic characterization of bases applies to arbitrary dimensions, we can generalize the pictorial calculus to higher dimensional quantum systems.

As with the qubit ZX calculus, the use of graphical notation is justified since **FHilb** is a †-CSMC. We let all the edges be implicitly labeled by $\mathbb{C}^D$ and focus on a pair of observable structures corresponding to the Z and X observables from qudit quantum mechanics.

The green observable structure, corresponding to the qudit observable $Z = \sum_{j=0}^{D-1} \eta^j |j\rangle \langle j|$, is defined via the copying and deleting maps:

$$\delta_Z = \begin{pmatrix} e_1| & e_2| & ... & |e_D \end{pmatrix} \text{ and } \epsilon_Z = \frac{1}{D}(1, 1, ..., 1).$$

Where δ_Z is a $D^2 \times D$ matrix with D columns e_i which have one 1 in row $D \times (i-1)+i$ and zeros in all the other rows.

Unbiased points for the green observable structure are in the form: $|\{\alpha_1, \alpha_2, ..., \alpha_{D-1}\}_Z\rangle = |0\rangle + \sum_{j=1}^{D-1} e^{i\alpha_j} |j\rangle$ and therefore, the phase group consists of matrices of the form:

$$\Lambda_Z(\alpha_1, \alpha_2, ..., \alpha_{D-1}) = \begin{pmatrix} 1 & 0 & 0 & ... & 0 \\ 0 & e^{i\alpha_1} & 0 & ... & 0 \\ 0 & 0 & e^{i\alpha_1} & ... & ... \\ ... & ... & ... & ... & 0 \\ 0 & 0 & ... & 0 & e^{i\alpha_{D-1}} \end{pmatrix} \tag{5.19}$$

Therefore the phase group for the green observable Π_Z and the group of unbiased points $(U_Z, \odot_Z)$ are both the ***D-torus group***, corresponding to the direct product of D circle groups $S^1 \times ... \times S^1$.

The green part of the ZX calculus for qudits follows from the generalized green spider law.

The red observable structure, corresponding to the qudit observable $X = \sum_{j=0}^{D-1} |j\rangle \langle j+1|$, is defined via the copying and deleting maps:

$$\delta_X = \begin{pmatrix} \mathbb{I}_{D\times D} \\ P_1(\mathbb{I}_{D\times D}) \\ ... \\ P_{D-1}(\mathbb{I}_{D\times D}) \end{pmatrix} \text{ and } \epsilon_Z = \frac{1}{\sqrt{D}}(1,0,...,0).$$

Where δ_X is a $D^2 \times D$ matrix composed of D matrix blocks $\mathbb{I}_{D\times D}$ and $P_j(\mathbb{I}_{D\times D})$ (j=1,2, ... D-1), which are $D \times D$ matrices corresponding to the identity matrix $\mathbb{I}_{D\times D}$, with all its rows permuted to the right by j.

Unbiased points for the red observable structure are in the form: $|\{\alpha_1, \alpha_2, ..., \alpha_{D-1}\}_X\rangle = |+_0\rangle + \sum_{k=1}^{D-1} e^{i\alpha_k} |+_k\rangle = \frac{1}{\sqrt{D}}(c_0 |0\rangle + \sum_{j=1}^{D-1} c_j |j\rangle)$, where $|+_k\rangle$ are the D eigenvectors of the X matrix (and the c_j are the computation basis decomposition coefficients). Therefore, the phase group consists of matrices of the form:

$$\Lambda_X(\alpha_1, \alpha_2, ..., \alpha_{D-1}) = \frac{1}{D} \begin{pmatrix} c_0 & c_{D-1} & c_{D-2} & ... & c_2 & c_1 \\ c_1 & c_0 & c_{D-1} & ... & c_3 & c_2 \\ c_2 & c_1 & c_0 & ... & c_4 & c_3 \\ ... & ... & ... & ... & ... & ... \\ c_{D-1} & c_{D-2} & c_{D-3} & ... & c_1 & c_0 \end{pmatrix} \tag{5.20}$$

Which can be shown to be unitary and which satisfy:

$$\Lambda_X(\beta_1, \beta_2, ..., \beta_{D-1}) \circ \Lambda_X(\alpha_1, \alpha_2, ..., \alpha_{D-1}) = \Lambda_X(\alpha_1 + \beta_1, \alpha_2 + \beta_2, ..., \alpha_{D-1} + \beta_{D-1}) \tag{5.21}$$

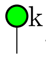
Therefore the phase group for the red observable Π_X and the group of unbiased points $(U_X, \odot_X)$ are both the ***D-torus group***, corresponding to the direct product of D circle groups $S^1 \times ... \times S^1$.

The red part of the ZX calculus for qudits follows from the generalized red spider law.

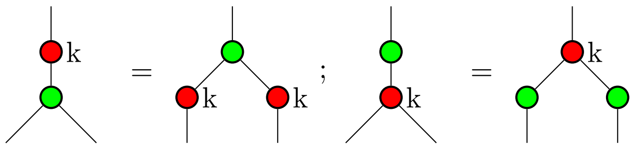
Note that the red and green observable structures do not induce the same compact structure since:

$$\eta_Z = \delta_Z \circ \epsilon_Z^\dagger \neq \delta_X \circ \epsilon_X^\dagger = \eta_X \tag{5.22}$$

The classical points of the green Z observable are , where k corresponds to phase values $\{\alpha_1, ..., \alpha_{D-1}\}$, $\{\beta_1, ..., \beta_{D-1}\}$, ... such that the red unbiased points $|\{\alpha_1, ..., \alpha_{D-1}\}_X\rangle$, $|\{\beta_1, ..., \beta_{D-1}\}_X\}\rangle$, ... are the classical points (eigenvectors) $|0\rangle, |1\rangle, ..., |D-1\rangle$ of the green observable Z. By theorem 3.4, these D points form an abelian subgroup of the D-torus group $(U_X, \odot_X)$, where $|0\rangle$ is the identity.

Similarly, the classical points of the red X observable are , where k corresponds to phase values $\{\alpha_1, ..., \alpha_{D-1}\}$, $\{\beta_1, ..., \beta_{D-1}\}, ...$ such that the green unbiased points $|\{\alpha_1, ..., \alpha_{D-1}\}_Z\rangle$,$|\{\beta_1, ..., \beta_{D-1}\}_Z\rangle, ...$ correspond to the classical points (eigenvectors) $|+_0\rangle, |+_1\rangle, ..., |+_{D-1}\rangle$ of the red observable X. By theorem 3.4, these D points form an abelian subgroup of the D-torus group $(U_Z, \odot_Z)$, where $|+_0\rangle$ is the identity.

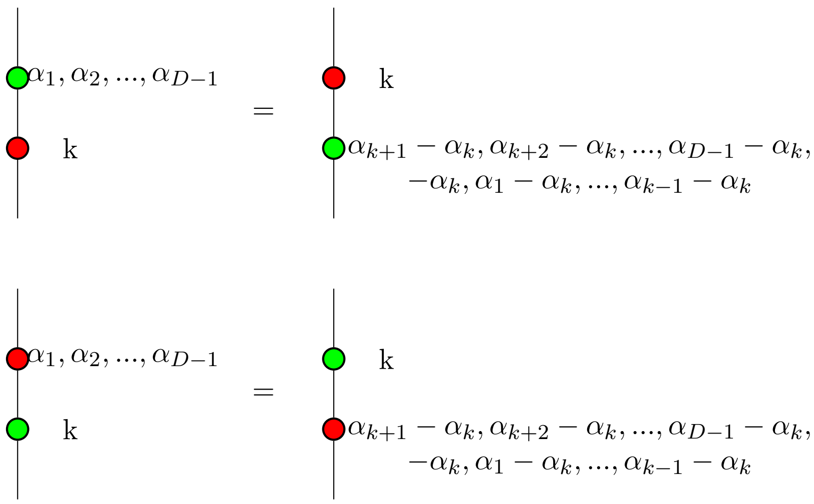
Therefore, the red and green observable structures are a closed pair of coherent observable structures. The (D), (B1) and (B2) rules (presented in the next section) then follow as before. The (K1) rule becomes:



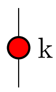
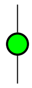
(K1)

There are $2D - 2$ equations in (K1), one equation for each of the D classical points of each colour (except the (0,0,..., 0), phaseless points of each colour). Note that if you add to (K1) the case where k corresponds to the (0,0,..., 0), phaseless points of each colour then the rule (B1) follow as a special case of (K1).

We obtain the (K2) rule, by calculating the action of the red (or green) phase maps of either colour, corresponding to classical points in K_Z (or K_X), on the unbiased green (or red) points in U_Z (or U_X). One can then see that the (K2) rule is:



(K2)

There are 2D-2 equations in (K2) corresponding to the the D phase maps  associated to the D classical points for Z and the D phase maps  k associated to the D classical points for X (except the (0,0,..., 0) phaseless maps for each colour).

For clarity, we illustrate this rule for the case of qudits of dimension four. This requires us to calculate the action of K_Z on U_Z:

$$\begin{aligned}\Lambda^X(\left|\{\frac{\pi}{2},\pi,\frac{3\pi}{2}\}_X\right\rangle)(|\{\alpha_1,\alpha_2,\alpha_3\}_Z\rangle) &= \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ e^{i\alpha_1} \\ e^{i\alpha_2} \\ e^{i\alpha_3} \end{pmatrix} \\ &= e^{i\alpha_1}\begin{pmatrix} 1 \\ e^{i(\alpha_2-\alpha_1)} \\ e^{i(\alpha_3-\alpha_1)} \\ e^{i(-\alpha_1)} \end{pmatrix} = (|\{\alpha_2-\alpha_1,\alpha_3-\alpha_1,-\alpha_1\}_Z\rangle)\end{aligned} \tag{5.23}$$

$$\begin{aligned}\Lambda^X(|\{\pi,0,\pi\}_X\rangle)(|\{\alpha_1,\alpha_2,\alpha_3\}_Z\rangle) &= \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ e^{i\alpha_1} \\ e^{i\alpha_2} \\ e^{i\alpha_3} \end{pmatrix} \\ &= e^{i\alpha_2}\begin{pmatrix} 1 \\ e^{i(\alpha_3-\alpha_2)} \\ e^{i(-\alpha_2)} \\ e^{i(\alpha_1-\alpha_2)} \end{pmatrix} = (|\{\alpha_3-\alpha_2,-\alpha_2,\alpha_1-\alpha_2\}_Z\rangle)\end{aligned} \tag{5.24}$$

$$\begin{aligned}\Lambda^X(\left|\{\frac{3\pi}{2},\pi,\frac{\pi}{2}\}_X\right\rangle)(|\{\alpha_1,\alpha_2,\alpha_3\}_Z\rangle) &= \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ e^{i\alpha_1} \\ e^{i\alpha_2} \\ e^{i\alpha_3} \end{pmatrix} \\ &= e^{i\alpha_3}\begin{pmatrix} 1 \\ e^{i(-\alpha_3)} \\ e^{i(\alpha_1-\alpha_3)} \\ e^{i(\alpha_2-\alpha_3)} \end{pmatrix} = (|\{-\alpha_3,\alpha_1-\alpha_3,\alpha_2-\alpha_3\}_Z\rangle)\end{aligned} \tag{5.25}$$

The action of K_X on U_X is exactly dual to this.

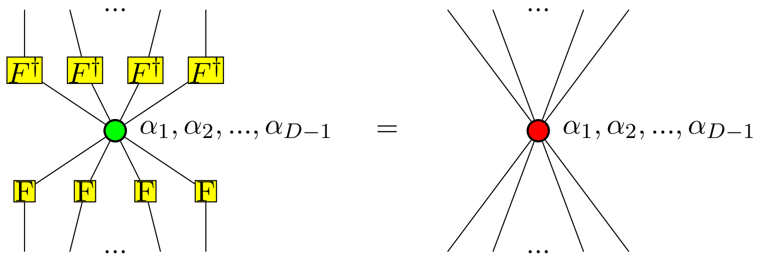
The last rule will correspond to the definition of the Fourier gate $F = \frac{1}{\sqrt{D}}\sum_{j,k=0}^{D-1}\eta^{jk}\,|j\rangle\,\langle k|$ in the calculus. In general, one can show that:

$$(F\otimes F)\circ\delta_Z\circ F^{\dagger} = \delta_X \tag{5.26}$$

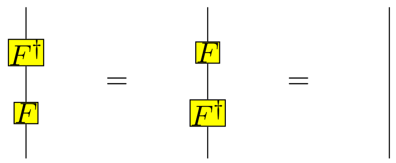
and:

$$F(|\{\alpha_1, \alpha_2, ...\}_Z\rangle) = |\{\alpha_1, \alpha_2, ...\}_X\rangle \tag{5.27}$$

where F is the unitary Fourier matrix. This holds for all dimensions and allows us to introduce the Fourier gate in the qudit ZX calculus in much the same way as the Hadamard matrix was introduced in the qubit ZX calculus[82], except that the Fourier gate corresponds to box with a vertical (involutive) asymmetry. This gives us the (F) rules of the qudit ZX calculus:



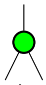
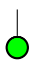
(F1)



(F2)

Therefore, we have justified all the rules of the qudit ZX calculus from the algebraic properties of the Z and X observables and of the Fourier map. This construction, together with Theorem 5.2 in Chapter 2, shows that qudit ZX calculus is **sound** for quantum mechanics.

5.3.2 The ZX calculus for qudit quantum mechanics:

We now present the ZX calculus for qudit quantum mechanics[250]. This is a generalization of the standard qubit ZX calculus[82]. Recall that an observable structure, which is a generalization of the Hilbert space concept of an orthonormal basis, consists of a copying map δ :  and a deleting map ϵ :  satisfying certain algebraic conditions. A state (or point) ψ is *classical* (or an eigenstate) for an observable structure if it is copied by the copying map and deleted by the deleting map. ψ is *unbiased* with respect to an observable structure if: $s(\delta^\dagger \circ (\psi \otimes \psi^\star)) = \epsilon^\dagger$ for some scalar s.

Given an observable structure, each state ψ has a corresponding *phase map*: $\Lambda(\psi) := \delta^\dagger \circ (\psi \otimes \mathbb{I})$. The set of all phase maps corresponding to unbiased states for an observable structure, together with map composition, form a group called the **phase group**. We will now present the rules of the calculus and its relationship to quantum theory.

General network diagrams are built out of parallel (tensor product) and downward compositions of generating diagrams from Figure 5.5.

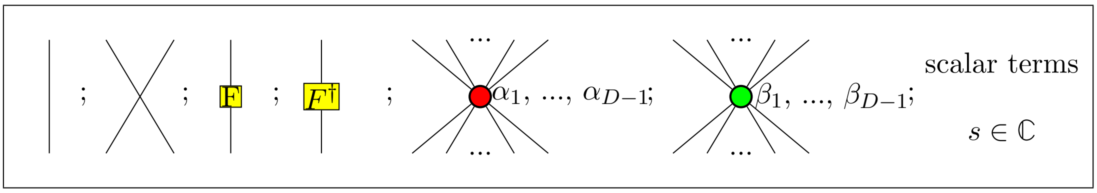
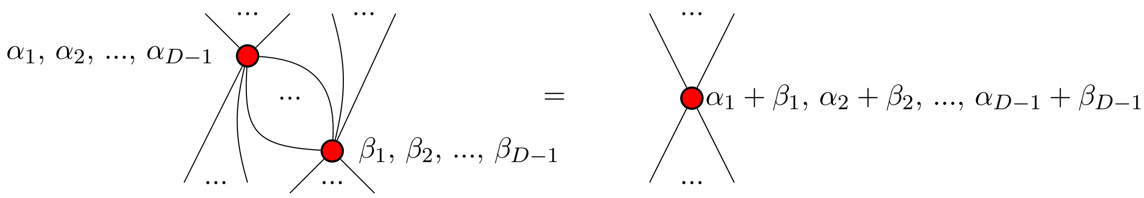
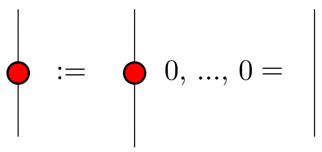


Figure 5.5: Generating diagrams for the qudit ZX calculus.

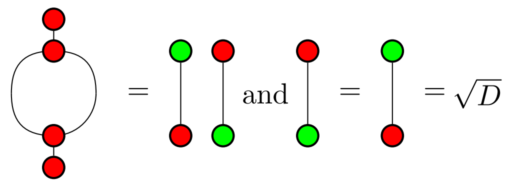
The rules of the ***qudit ZX calculus*** are the (S), (D), (B), (K), and (F) rules below (and their reversed colour counterparts), together with a (T) rule which states that after identifying the inputs and outputs of any part of a ZX network, any topological deformation of the internal structure does not matter.



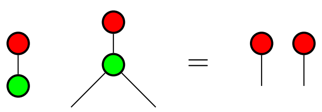
(S1)



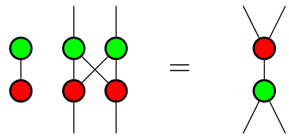
(S2)



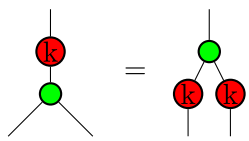
(D)



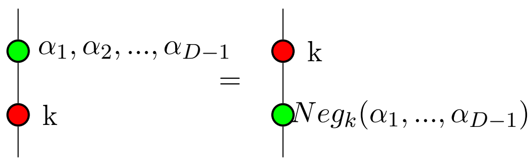
(B1)



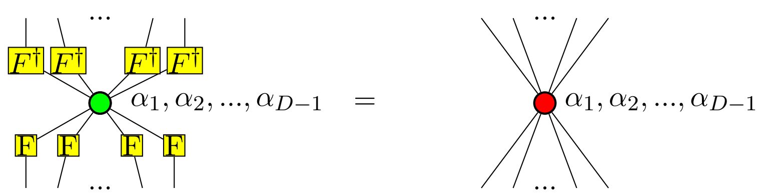
(B2)



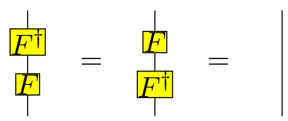
(K1)



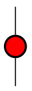
(K2)



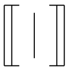
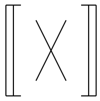
(F1)

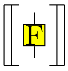
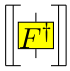


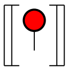
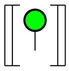
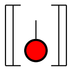
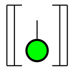
(F2)

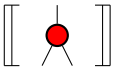
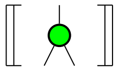
where $Neg_k(\alpha_1, ..., \alpha_{D-1}) := \alpha_{k+1} - \alpha_k, \alpha_{k+2} - \alpha_k, ..., \alpha_{D-1} - \alpha_k, -\alpha_k, \alpha_1 - \alpha_k, ..., \alpha_{k-1} - \alpha_k$, and where there are D-1 different red k vertices which have phases $\alpha_1, ..., \alpha_{D-1}$ such that  k are the phase maps corresponding to the D-1 classical points for Z whose phases are not all zero. In higher dimensions, the (K) rules give rise to more intricate interference phenomena, since the D classical points of an observable structure each permute the phase group elements.

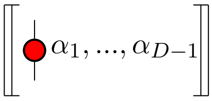
Diagrammatic reasoning in the qudit calculus is identical to reasoning in the qubit

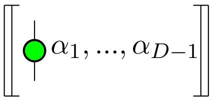
 $= \mathbb{I}_{D\times D} := \sum_{k=0}^{D-1} |k\rangle\langle k|$;  $= SWAP_{a,b} := \sum_{j,k=0}^{D-1} |k\rangle\langle j|_a \otimes |j\rangle\langle k|_b$

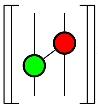
 $= Fourier := \frac{1}{\sqrt{D}} \sum_{j,k=0}^{D-1} \eta^{jk} |j\rangle\langle k|$;  $= Fourier^{\dagger}.$

 $= |0\rangle := \sqrt{D}\begin{pmatrix} 1 \\ 0 \\ \dots \\ 0 \end{pmatrix}$;  $= |+\rangle := \begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix}$;  $= \epsilon_X := \langle 0|$;  $=$ $\epsilon_Z := \langle +|$

 $= \delta_X := \begin{pmatrix} \mathbb{I}_{D\times D} \\ P_1(\mathbb{I}_{D\times D}) \\ \dots \\ P_{D-1}(\mathbb{I}_{D\times D}) \end{pmatrix}$;  $= \delta_Z := \begin{pmatrix} e_1| & e_2| & \dots & |e_D \end{pmatrix}$

 $= \Lambda_X(\alpha_1, \alpha_2, \dots, \alpha_{D-1}) := \frac{1}{D}\begin{pmatrix} c_0 & c_{D-1} & c_{D-2} & \dots & c_2 & c_1 \\ c_1 & c_0 & c_{D-1} & \dots & c_3 & c_2 \\ c_2 & c_1 & c_0 & \dots & c_4 & c_3 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ c_{D-1} & c_{D-2} & c_{D-3} & \dots & c_1 & c_0 \end{pmatrix}$

 $= \Lambda_Z(\alpha_1, \alpha_2, \dots, \alpha_{D-1}) := \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & e^{i\alpha_1} & 0 & \dots & 0 \\ 0 & 0 & e^{i\alpha_2} & \dots & \dots \\ \dots & \dots & \dots & \dots & 0 \\ 0 & 0 & \dots & 0 & e^{i\alpha_{D-1}} \end{pmatrix}$

 $= CNOT_{a,b} := \sum_{j,k=0}^{D-1} |k\rangle\langle j|_a \otimes |k\rangle\langle k+j|_b$

$$= \begin{pmatrix} \mathbb{I}_{D\times D} & 0 & 0 & \dots & 0 \\ 0 & P_1(\mathbb{I}_{D\times D}) & 0 & \dots & 0 \\ 0 & 0 & P_2(\mathbb{I}_{D\times D}) & \dots & \dots \\ \dots & \dots & \dots & \dots & 0 \\ 0 & 0 & \dots & 0 & P_{D-1}(\mathbb{I}_{D\times D}) \end{pmatrix}$$

Figure 5.6: Hilbert space interpretation of the qudit ZX calculus elements.

calculus. As before, two network diagrams can be shown to be equal by locally replacing some part of a diagram with a diagram equal to it.

Note that the restricted case of the ZX calculus for qutrits has been studied independently by Quanlong Wang and co-workers[296].

As with the qubit case, we can model the calculus in Hilbert space. We interpret all diagram edges by $\mathbb{C}^D$ and elements of the qudit calculus correspond to the following Hilbert space elements:

where $P_j(\mathbb{I}_{D\times D})$ (j=1,2, ... D-1) are $D \times D$ matrices corresponding to the identity matrix

$\mathbb{I}_{D\times D}$, with all its rows permuted to the right by j and where e_i are the vectors which have one 1 in row $D \times (i-1)+i$ and D-1 zeros in all the other rows.

The c_j elements of the Λ_X matrix are defined by: $|+\rangle + \sum_{k=1}^{D-1} e^{i\alpha_k} |+_k\rangle = \frac{1}{\sqrt{D}}(c_0 |0\rangle + \sum_{j=1}^{D-1} c_j |j\rangle)$ where $|+_k\rangle$ are the D eigenvectors of the $X = \sum_{j=0}^{D-1} |j\rangle \langle j+1|$ matrix.

We now proceed to show the universality of the qudit ZX calculus for quantum mechanics. Muthukrishnan and Stroud [222] have shown that the families of gates Z_j and X_j, which we define below, are are sufficient to simulate all single qudit transformations. Moreover, Brylinski [68] has proven that the collection of all one-qudit gates together with a single imprimitive two-qudit gate (as defined below) produces all n-qudit gates up to arbitrary precision.

We combine these two important results into a theorem:

<u>Theorem 3.5:</u> The following set of qudit quantum gates are universal for quantum computing: (a) The two following families of D-dimensional transforms (which are universal for single qudit quantum mechanics) [222]:

$$Z_j(b_0, b_1..., b_{D-1}) : b_0 |0\rangle + b_1 |1\rangle + ... + b_{D-1} |D-1\rangle \mapsto |j\rangle \tag{5.28}$$

for $j \in \{0, 1, ..., D-1\}$, where the b_j are complex coefficients normalized to unity. Note that this equation does not determine the map Z_j uniquely.

$$X_j(\phi) : b_0 |0\rangle + b_1 |1\rangle + ... + b_{D-1} |D-1\rangle \mapsto b_0 |0\rangle + ... + e^{i\phi} b_j |j\rangle + b_{D-1} |D-1\rangle \tag{5.29}$$

for $j \in \{0, 1, ..., D-1\}$.

(b) Any imprimitive 2 qudit gate [68], where a 2 qudit gate is called imprimitive if there exist no single qudit gates S and T such that: $V = S \otimes T$ or $V = (S \otimes T)SWAP$ where $SWAP |xy\rangle = |yx\rangle$.

<u>Theorem 3.6:</u> The qudit ZX calculus is universal for quantum mechanics.

Proof. (a) Single qudit universality:

We can show that all 2D maps Z_0, Z_1, ..., Z_{D-1}, $X_0(\phi_0)$, $X_1(\phi_1)$, ..., and $X_{D-1}(\phi_{D-1})$

are included in the qudit ZX calculus.

One can easily check that (up to global phase): $X_0(\phi_0) = \Lambda_Z(-\phi_0, ..., -\phi_0)$, $X_j(\phi_j) = \Lambda_Z(0, ..., 0, \alpha_j = \phi_j, 0, ..., 0)$ for $j \neq 0$.

Next, we show that the d maps $Z_j(b_0, b_1, ..., b_{D-1})$ can be written (up to global phase) as $\Lambda_X(\alpha_1, \alpha_2, ..., \alpha_{D-1})$ for some set $\{\alpha_1, \alpha_2, ..., \alpha_D\}$.

First of all, note that $\Lambda_Z(\alpha_1, \alpha_2, ..., \alpha_{D-1})$ has determinant one and that, since $\Lambda_X(\alpha_1, \alpha_2, ..., \alpha_{D-1})$ is unitary and is obtained by applying the Fourier gate and its transpose to $\Lambda_Z(\alpha_1, \alpha_2, ..., \alpha_{D-1})$, we have that:

$$det(\Lambda_X(\alpha_1, \alpha_2, ..., \alpha_{D-1})) = 1 \neq 0 \tag{5.30}$$

for any values of $\{\alpha_1, \alpha_2, ..., \alpha_{D-1}\}$.

This means that there exists a unique solution set $\{b_0, b_1, ..., b_{D-1}\}$ to the equation:

$$\Lambda_X(\alpha_1, \alpha_2, ..., \alpha_{D-1}).(b_0, b_1, ..., b_{D-1})^T = e_j \tag{5.31}$$

for each vector $e_j = (0, ..., 0, 1, 0, ..., 0)^T$ with a single 1 in the j^{th} column.

Using the definition of Λ_X, this means that there exists a unique set of $\{b_0, b_1, ..., b_{D-1}\}$ such that:

$$c_j b_0 + c_{j-1} b_1 + ... + c_0 b_j + c_{D-1} b_{j+1} + ... + c_{j+1} b_{D-1} = D \tag{5.32}$$

$$c_k b_0 + c_{k-1} b_1 + ... + c_0 b_k + c_{D-1} b_{k+1} + ... + c_{k+1} b_{D-1} = 0; \forall k \neq j \tag{5.33}$$

Recall that the c_k are defined as: $c_k = 1 + \sum_{l=1}^{D-1} \eta^{r_k(l)} e^{i\alpha_l}$ where r_k permutes the entries l (there is one r_k for each k). We know that: $\sum_{k=0}^{D-1} \eta^k = 0$, where $\eta = e^{\frac{2\pi i}{D}}$, and that all the b_k ($k = 0, 1, ..., D-1$) are complex numbers of unit norm so that: $\sum_{k=0}^{D-1} b_k^\star b_k = 1$. This means that there exists a solution to this set of equations. Therefore, up to global phase, it is always possible to find values for $\{\alpha_1, ..., \alpha_{D-1}\}$ such that (5.32,5.33) is satisfied.

But and this means that each map $Z_D(b_0, ..., b_{D-1})$ (each one corresponding to a value of j in (5.31)) can be written in the form $\Lambda_X(\alpha_1, \alpha_2, ..., \alpha_{D-1})$ for some set $\{\alpha_1, \alpha_2, ..., \alpha_D\}$.

Using theorem 3.5, this shows that the qudit ZX calculus contains all single qudit unitary transformations.

(b) The qudit CNOT gate is an imprimitive 2 qudit gate which is contained within the qudit calculus.

Note also that any map from n qudits to m qudits can be constructed by using diagrammatic map-state duality [3]. Therefore, any qudit quantum state and (post-selected) measurement and any quantum gate can be written in the qudit ZX calculus and therefore it is **universal** for quantum mechanics. □

To illustrate the proof for single qudit universality, note that in the qutrit case, we can explicitly find an assignment of the α values such that (up to global phase):

$$\begin{aligned} Z_0(b_0,b_1,b_2) = \Lambda_X(&-i\log\left(\frac{(b_0+b_1+b_2)(b_0\eta-b_1(\eta+1)+b_2)}{\eta\,(b_0^2(-\eta)(\eta+1)-b_0(b_1+b_2)+b_1^2+b_1b_2\eta(\eta+1)+b_2^2)}\right),\\ &-i\log\left(\frac{(b_0+b_1+b_2)(b_0\eta+b_1-b_2(\eta+1))}{\eta\,(b_0^2(-\eta)(\eta+1)-b_0(b_1+b_2)+b_1^2+b_1b_2\eta(\eta+1)+b_2^2)}\right)) \end{aligned} \tag{5.34}$$

$$\begin{aligned} Z_1(b_0,b_1,b_2) = \Lambda_X(&-i\log\left(\frac{(b_0+b_1+b_2)(b_0+b_1\eta-b_2(\eta+1))}{\eta\,(b_0^2+b_0(b_2\eta(\eta+1)-b_1)-(b_1\eta+b_1-b_2)(b_1\eta+b_2))}\right),\\ &-i\log\left(-\frac{(b_0+b_1+b_2)(b_0\eta+b_0-b_1\eta-b_2)}{\eta\,(b_0^2+b_0(b_2\eta(\eta+1)-b_1)-(b_1\eta+b_1-b_2)(b_1\eta+b_2))}\right)) \end{aligned} \tag{5.35}$$

$$\begin{aligned} Z_2(b_0,b_1,b_2) = \Lambda_X(&-i\log\left(-\frac{(b_0+b_1+b_2)(b_0\eta+b_0-b_1-b_2\eta)}{\eta\,(b_0^2+b_0(b_1\eta(\eta+1)-b_2)+(b_1+b_2\eta)(b_1-b_2(\eta+1)))}\right),\\ &-i\log\left(\frac{(b_0+b_1+b_2)(b_0-b_1(\eta+1)+b_2\eta)}{\eta\,(b_0^2+b_0(b_1\eta(\eta+1)-b_2)+(b_1+b_2\eta)(b_1-b_2(\eta+1)))}\right)) \end{aligned} \tag{5.36}$$

where $\eta = e^{\frac{2\pi i}{3}}$.

By construction, any equation which can be shown to be true using the qudit ZX calculus is true in quantum mechanics so the qudit ZX calculus is **sound** for quantum mechanics. Moreover, extending the qudit ZX calculus to account for mixed states and general quantum evolution described by completely positive maps can be achieved by using the same standard constructions [266,87,119] as in the qubit case.

We know that [24] the qubit ZX calculus is **complete** for qubit stabilizer quantum

theory, in the sense that any two equivalent qubit stabilizer processes can be shown to be equal by using the qubit ZX calculus. Backens' proof of this result [24], however, relies on results for qubit graph states and it is unclear whether it can be generalized to show completeness of the qudit ZX calculus for qudit stabilizer theory. Therefore, we leave this as an open question:
Is the qudit ZX calculus, with additional rules analogous to the Euler decomposition of the Hadamard vertex, complete for qudit stabilizer quantum mechanics? If it is not, then which other rules need to be added for completeness?

Another important question is how the qudit and qubit ZX calculi are related. More generally, it would be interesting to understand exactly how the ZX calculus for qudits of dimension m is related to the ZX calculus of dimension $n > m$. Perhaps, we could introduce maps which "*create*" and "*annihilate*" dimensions. This could lead to an interesting structure and provide insight into the relationship between qubit and qudit quantum mechanics.

We anticipate that the new calculus will provide a practical tool to study quantum information and computation from a high-level point of view. For example, the qudit calculus for dimensions higher than two should be well suited to understanding structural properties of quantum algorithms, quantum key distribution and quantum error-correction. Moreover, as the complexity of the quantum systems we study will grow, computer software such as Quantomatic [185], which allows automated reasoning within the calculus, may play an important role in the design of future quantum networks.

The framework presented thus far is limited to pure states and measurements as post-selected projections. We now briefly describe three possible ways to extend the qudit ZX calculus to account for mixed states, measurements, decoherence and general quantum evolution described by completely positive maps.

The first method of augmenting the graphical language is to use the **Selinger CPM construction** [266] which, for each dagger compact category C of pure states and maps produces a new dagger compact category CPM(C) of mixed states and completely positive maps. This new category is constructed in the following way:

The objects of CPM(C) are the same as the objects of C. The morphisms of CPM(C) are the morphisms of C that can be written in the form $\mathcal{E}= (1\otimes\eta^{\dagger}\otimes 1)\circ(f_{\star}\otimes f)$: $A^{\star}\otimes A \to B^{\star}\otimes B$

(**completely positive maps**), where $f : A \to D \otimes B$ is a morphism in C (Kraus morphism) and D is an object in C (ancillary system). Identities and tensor products are inherited from C and composition is as in C. If $\mathcal{E}$: $A^\star \otimes A \to B^\star \otimes B$ then the adjoint in CPM(C) is given by: $\mathcal{E}$: $B^\star \otimes B \to A^\star \otimes A$. The construction preserves the dagger compact structure of C.

One can then study probability distributions, quantum measurements, decoherence, etc. in the resulting category[90]. Quantum operations are then encoded using two wires and classical information can be described using a single wire.

The second formalism which allows one to go from a pure state qudit ZX calculus to a general theory with mixed states and completely positive maps is to introduce an environment structure[87,84].

An ***environment structure*** for a dagger symmetric monoidal category is defined as a dagger symmetric monoidal "supercategory" $\hat{C}$ with the same objects as C but where each object A has a morphism $\top_A$: $A \to I$ which satisfies the following axioms:
(i) $\top_I = 1_I$ and for all objects A and B: $\top_A \otimes \top_B = \top_{A \otimes B}$ in $\hat{C}$.
(ii) For all $f, g \in C(A, C \otimes B)$ we have: $f^\dagger \circ f = g^\dagger \circ g$ in C iff $\top_C \circ f = \top_C \circ g$ in $\hat{C}$.
(iii) For each $\hat{f} \in \hat{C}(A, B)$, there is $f \in C(A, C \otimes B)$ st: $\hat{f} = \top_C \circ f$ in $\hat{C}$ (purification).

If C consist of pure states then the category $\hat{C}$ consists of mixed states and the $\top$ is an environment map. This approach can in fact be related to the CPM construction[87,84]. Adding the environment to the ZX calculus allows one to account for decoherence and measurement in a general way.

The third possible generalization of the graphical calculus is to introduce a set of variables that encode the outcome of measurements. This allows one to study determinism and information flow using *conditional diagrams*[119].

We define a set of variables V and valuation functions $V \to \{0, 1\}$. A **conditional diagram** is a ZX calculus diagram $\mathcal{D}$ where each Z or X vertex v has an associated variable subset $U_v \subseteq V$. From each conditional diagram and valuation function pair $\{\mathcal{D}, f\}$, we can obtain an evaluated diagram $\mathcal{D}_f$ by modifying the phase at each vertex (according to the product of the valuation functions f(u) for all $u \in U_v$) then forgetting U_v. In this way, valuations correspond to possible sets of measurement outcomes for measurements corresponding to variables. The evaluated diagram $\mathcal{D}_f$ depicts the measurement process when a particular outcome corresponding to f is observed.

We can then construct a CP map: $\mathcal{E} : \rho \mapsto \sum_f D_f \rho D_f^\dagger$ by taking the Kraus linear maps associated with the evaluated diagrams $\mathcal{D}_f$ and summing over all possible valuations.

Therefore, we have seen that the ZX calculus for qudits can be extended to describe mixed states, completely positive maps and measurement theory.

5.3.3 Mutually unbiased qudit theories

One of the main goals of this chapter is to use the abstract structures we introduced to study the foundation of quantum theory. In this respect, we aim to define a class of theories which exhibit many key features of quantum mechanics, within a single mathematical framework.

Therefore, we will generalize the previous approach of studying mutually unbiased qubit theories using dagger compact symmetric monoidal categories [91,123] to the case of qudits.

<u>Definition:</u> A **mutually unbiased qudit theory** is a dagger symmetric monoidal category with observable structures such that:

(i) The objects of the category are the unit and finite tensor products of qudit-like systems Q.

(ii) The observables on a given object are all mutually unbiased, have the same number of eigenstates and have the same phase groups.

(iii) All states of Q are eigenstates of some observable.

We will study mutually unbiased qudit theories for dimensions higher than two. In the following two sections, we analyze in detail two key examples of mutually unbiased qudit theories: *qudit stabilizer quantum mechanics* and *Spekkens-Screiber theory for dits.* Note that we are interested in the diagrams from the ZX calculus which can be directly related to physical processes and thereby we use the calculus to depict MUQTs.

5.3.4 Picturing stabilizer quantum mechanics

We define the process category **DStab** as the †-compact symmetric monoidal subcategory of **FHilb** corresponding to qudit stabilizer quantum mechanics, which is generated by the unit, n-fold tensor products of $\mathbb{C}^D$, single qudit Clifford operations and the quantum copying and deleting maps. **DStab** can be depicted using the qudit ZX diagrams, where the allowed phases are restricted according to the phase group.

In the case of the standard qubit stabilizer quantum mechanics, the phase group is the cyclic group $\mathbb{Z}_4$, which is a finite subgroup of the quantum qubit phase group S^1 (the circle group). Since the unbiased circles for the Z and X observables coincide on the points corresponding to $|+i\rangle$ and $|-i\rangle$, we can completely picture single qubit stabilizer quantum theory using the Bloch sphere.

Can one find an analogous picture for qutrit quantum mechanics?
Let $\{|0\rangle, |1\rangle, |2\rangle\}$ and $\{|+\rangle, |\top\rangle, |\bot\rangle\}$ be the eigenbases for the qutrit Z and X observables respectively. Then the unbiased states for the Z and X observable:

$$|\{\alpha_1, \alpha_2\}_Z\rangle = |0\rangle + e^{i\alpha}|1\rangle + e^{i\alpha_2}|2\rangle \;\; ; \;\; |\{\alpha_1, \alpha_2\}_X\rangle = |+\rangle + e^{i\alpha}|\top\rangle + e^{i\alpha_2}|\bot\rangle \tag{5.37}$$

under pairwise addition of phases form a torus group $S^1 \times S^1$.

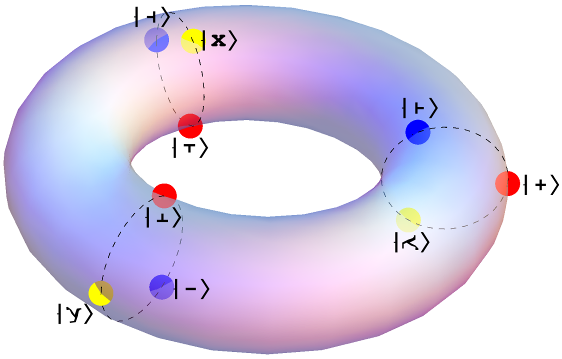
All the single qutrit stabilizer states, corresponding to the eigenstates of the qutrit X, Z, XZ and XZ^2 operators, can be written as unbiased states for either the Z basis or the X basis since:

$$\begin{aligned}
&|0\rangle = |\{0,0\}_X\rangle, |1\rangle = \left|\{\frac{4\pi}{3},\frac{2\pi}{3}\}_X\right\rangle, |2\rangle = \left|\{\frac{2\pi}{3},\frac{4\pi}{3}\}_X\right\rangle;\\
&|+\rangle = |\{0,0\}_Z\rangle, |\top\rangle = \left|\{\frac{2\pi}{3},\frac{4\pi}{3}\}_Z\right\rangle, |\bot\rangle = \left|\{\frac{4\pi}{3},\frac{2\pi}{3}\}_Z\right\rangle;\\
&|-\rangle = \left|\{\frac{4\pi}{3},\frac{4\pi}{3}\}_Z\right\rangle = \left|\{\frac{2\pi}{3},\frac{2\pi}{3}\}_X\right\rangle, |\vdash\rangle = \left|\{0,\frac{2\pi}{3}\}_Z\right\rangle = \left|\{\frac{4\pi}{3},0\}_X\right\rangle, |\dashv\rangle = \left|\{\frac{2\pi}{3},0\}_Z\right\rangle = \left|\{0,\frac{4\pi}{3}\}_X\right\rangle;\\
&|\times\rangle = \left|\{\frac{2\pi}{3},\frac{2\pi}{3}\}_Z\right\rangle = \left|\{\frac{4\pi}{3},\frac{4\pi}{3}\}_X\right\rangle, |\curlywedge\rangle = \left|\{\frac{4\pi}{3},0\}_Z\right\rangle = \left|\{\frac{2\pi}{3},0\}_X\right\rangle, |\prec\rangle = \left|\{0,\frac{4\pi}{3}\}_Z\right\rangle = \left|\{0,\frac{2\pi}{3}\}_X\right\rangle
\end{aligned} \tag{5.38}$$

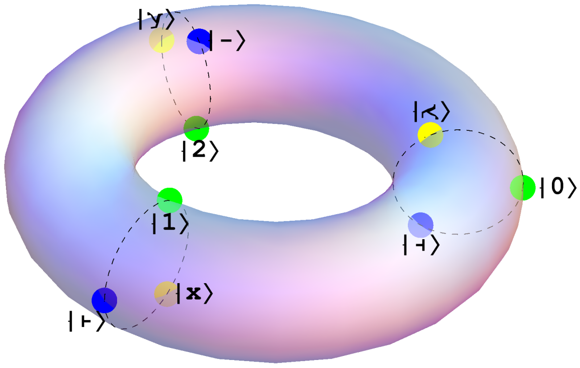
Single qutrit stabilizer operations take subsets of these 12 states to other subsets of these 12 states. This shows that the phase group for qutrit stabilizer quantum mechanics is $\mathbb{Z}_3 \times \mathbb{Z}_3$.

Therefore, single qutrit stabilizer quantum theory can be pictured using 12 points on two toruses, which is a direct generalization of the Bloch sphere case, where the 4 elements on each of the two unbiased circles (coinciding on two elements) visualized in three dimensions are replaced by **9 elements on each of two unbiased toruses coinciding on six points** (the blue and yellow points in Figures (5.7a, 5.7b)).

In fact, this picture can easily be generalized to higher dimensional qudit stabilizer the-



(a) Qutrit stabilizer states on the unbiased torus for the Z observable.



(b) Qutrit stabilizer states on the unbiased torus for the X observable.

Figure 5.7: Depicting qutrit stabilizer theory on two tori.

ories for prime dimensions. In that case, the single qudit states of qudit stabilizer quantum theory correspond to the vectors in the D+1 mutually unbiased eigenbases of the single qudit operators: $X, Z, XZ, XZ^2, ..., XZ^{D-1}$. The mutually unbiased points with respect to each of these bases forms a D-torus. If we chose an observable structure, whose eigenstates are a basis, then all the other stabilizer states are on the unbiased D-torus of the chosen basis. In this way, qudit stabilizer theory for prime dimension D can be pictured using D^2 **points on each of two D-toruses** (unbiased toruses for the Z and X operators for example), which coincide on $D^2 - D$ points and can be visualized in D+1 dimensions.

In general, the phase group for qudit stabilizer quantum theory of dimension $D > 3$ is an Abelian subgroup of the group $\mathbb{Z}_D \times \mathbb{Z}_D \times ... \times \mathbb{Z}_D$ (D-1 times). In fact, every finite dimensional closed subgroup of the torus group is isomorphic to a product of finite cyclic groups[258]. Therefore, the phase group for mutually unbiased qudit theories which are also sub-theories of quantum mechanics must be of the form $Z_{n_1} \times Z_{n_2} \times ... \times Z_{n_k}$ for positive integers $n_1, ..., n_k$. In further work, we will study how these integers $n_1, ..., n_k$ for stabilizer phase groups depend on the dimension D. In general, we would like a physical classification of all the mutually unbiased qudit sub-theories of quantum mechanics in terms of $n_1, ..., n_k$. Once we have determined their **phase group**, the qudit ZX calculus allows us to fully describe these physical theories.

As an example of the ZX calculus for stabilizer quantum theory in non-prime dimensions, we study the four-dimensional case.

In four dimensions, we can write three coherent mutually unbiased bases corresponding to the eigenbases of the operators Z, X and $Y = XZ^2$. The eigenbasis of Z can be written as:

$$\{|0\rangle = (1,0,0,0)^T,\ |1\rangle = (1,0,0,0)^T,\ |2\rangle = (1,0,0,0)^T,\ |3\rangle = (1,0,0,0)^T\} \tag{5.39}$$

The Z and Y eigenbases can respectively be written as:

$$\begin{aligned} \{\,|+\rangle &= \frac{1}{2}(|0\rangle + |1\rangle + |2\rangle + |3\rangle),\ \ |\times\rangle = \frac{1}{2}(|0\rangle + i\,|1\rangle - |2\rangle - i\,|3\rangle), \\ |-\rangle &= \frac{1}{2}(|0\rangle - |1\rangle + |2\rangle - |3\rangle),\ \ |\div\rangle = \frac{1}{2}(|0\rangle - i\,|1\rangle - |2\rangle + i\,|3\rangle)\} \end{aligned} \tag{5.40}$$

$$\begin{aligned} \{\,|\top\rangle &= \frac{1}{2}(|0\rangle - |1\rangle - |2\rangle - |3\rangle),\ \ |\bot\rangle = \frac{1}{2}(|0\rangle - i\,|1\rangle + |2\rangle + i\,|3\rangle), \\ |\vdash\rangle &= \frac{1}{2}(|0\rangle + i\,|1\rangle + |2\rangle - i\,|3\rangle),\ \ |\dashv\rangle = \frac{1}{2}(|0\rangle + |1\rangle - |2\rangle + |3\rangle)\}. \end{aligned} \tag{5.41}$$

Note that we can always construct D+1 mutually unbiased bases when D is an integer power of a prime [31,189] but that in the case when D is not an integer power of a prime then the maximal number of mutually unbiased bases is not known [66,175].

The set of unbiased states (up to global phase) for the $|0\rangle$, $|1\rangle$, $|2\rangle$, $|3\rangle$ eigenbasis of the qudit Z observable can be written as:

$$|\{\alpha_1, \alpha_2, \alpha_3\}_Z\rangle = |0\rangle + e^{i\alpha}\,|1\rangle + e^{i\alpha_2}\,|2\rangle + e^{i\alpha_3}\,|3\rangle\,. \tag{5.42}$$

These unbiased states for the Z observable form a 3-torus group under the operation:

$$|\{\alpha_1, \alpha_2, \alpha_3\}_Z\rangle \odot_Z |\{\beta_1, \beta_2, \beta_3\}_Z\rangle = |\{\alpha_1+\beta_1, \alpha_2+\beta_2, \alpha_3+\beta_3\}_Z\rangle\,. \tag{5.43}$$

Similarly, the set of unbiased states (up to global phase) for the $|+\rangle$, $|\times\rangle$, $|-\rangle$, $|\div\rangle$

eigenbasis of the ququit X observable can be written as:

$$|\{\alpha_1, \alpha_2, \alpha_3\}_X\rangle = |+\rangle + e^{i\alpha} |\times\rangle + e^{i\alpha_2} |-\rangle + e^{i\alpha_3} |\div\rangle \tag{5.44}$$

and these unbiased states for the X observable form a 3-torus group under the operation:

$$|\{\alpha_1, \alpha_2, \alpha_3\}_X\rangle \odot_X |\{\beta_1, \beta_2, \beta_3\}_X\rangle = |\{\alpha_1 + \beta_1, \alpha_2 + \beta_2, \alpha_3 + \beta_3\}_X\rangle . \tag{5.45}$$

All the single four-dimensional stabilizer states, correspond to the eigenstates of the X, Z and XZ^2 operators and these can be written as unbiased states for either the Z basis or the X basis. Indeed, the 8 stabilizer states unbiased for the Z observable can be written as:

$$\begin{aligned}
&|+\rangle = |\{0,0,0\}_Z\rangle\,, \quad |\times\rangle = \left|\{\frac{\pi}{2}, \pi, \frac{3\pi}{2}\}_Z\right\rangle, \\
&|-\rangle = |\{\pi,0,\pi\}_Z\rangle\,, \quad |\div\rangle = \left|\frac{3\pi}{2}, \pi, \frac{\pi}{2}\}_Z\right\rangle; \\
&|\top\rangle = |\{\pi,\pi,\pi\}_Z\rangle\,, \quad |\bot\rangle = \left|\frac{3\pi}{2}, 0, \frac{\pi}{2}\}_Z\right\rangle, \\
&|\vdash\rangle = \left|\frac{\pi}{2}, 0, \frac{3\pi}{2}\}_Z\right\rangle, \quad |\dashv\rangle = |\{0,\pi,0\}_Z\rangle
\end{aligned} \tag{5.46}$$

Similarly, the 8 stabilizer states unbiased for the X observable can be written as:

$$\begin{aligned}
&|0\rangle = |\{0,0,0\}_X\rangle\,, \quad |1\rangle = \left|\{\frac{3\pi}{2}, \pi, \frac{\pi}{2}\}_X\right\rangle, \\
&|2\rangle = |\{\pi,0,\pi\}_X\rangle\,, \quad |3\rangle = \left|\{\frac{\pi}{2}, \pi, \frac{3\pi}{2}\}_X\right\rangle; \\
&|\top\rangle = |\{\pi,\pi,\pi\}_X\rangle\,, \quad |\bot\rangle = |\{\pi,0,0\}_X\rangle\,, \\
&|\vdash\rangle = |\{0,0,\pi\}_X\rangle\,, \quad |\dashv\rangle = |\{0,\pi,0\}_X\rangle
\end{aligned} \tag{5.47}$$

Single stabilizer operations take subsets of these 12 states to other subsets of these 12 states. The group of unbiased points for a basis in four-dimensional quantum stabilizer theory is a proper abelian subgroup of $\mathbb{Z}^4 \times \mathbb{Z}^4 \times \mathbb{Z}^4$ with eight elements which has the group multiplication table given in Figure 5.8 below.

Figure 5.8: Group table for the four-dimensional stabilizer phase group.

+	id	a	b	c	d	e	f	g
id	id	a	b	c	d	e	f	g
a	a	g	id	f	e	c	d	b
b	b	id	g	e	f	d	g	a
c	c	f	e	id	g	b	a	d
d	d	e	f	g	id	a	b	c
e	e	c	d	b	a	g	id	f
f	f	d	g	a	b	id	g	e
g	g	b	a	d	c	f	e	id

Note that if we use addition modulo 4 and modulo 2 we can take:

$$\begin{aligned} id = (0,0), \quad a = (1,1), \quad b = (3,1), \quad c = (0,1) \\ d = (2,1), \quad e = (3,0), \quad f = (1,0), \quad g = (2,0) \end{aligned} \tag{5.48}$$

Therefore the phase group is: $\mathbb{Z}_4 \times \mathbb{Z}_2$.

It seems odd that stabilizer quantum mechanics in four dimensions only uses three of the five possible mutually unbiased bases. Indeed, this means that single qudit four dimensional stabilizer theory has exactly the same number of states as three dimensional stabilizer theory. Perhaps, it would be interesting to extend four-dimensional stabilizer quantum mechanics to a theory which has all the 20 vectors from all five mutually unbiased bases as single qudit states. We would then expect the phase group to be a larger subgroup of $\mathbb{Z}^4 \times \mathbb{Z}^4 \times \mathbb{Z}^4$ than the one above. In either case, we can picture qudit stabilizer quantum mechanics using two 3-toruses as we described before.

In general, qudit stabilizer theory for non-prime dimension D will only have $3 \times D$ states corresponding to three mutually unbiased bases. It is still an open question whether there exist sets with more than three mutually unbiased bases in non prime power dimensions, such as D=6,10,... .

Thus, we have shown how qudit Stabilizer theory can be described as a †-compact symmetric mondonoidal theory of processes using the qudit ZX calculus, where the choice of the **phase group** determines which state preperations, effects and single qudit maps $\Lambda_X(\alpha_1, ..., \alpha_D)$ and $\Lambda_Z(\alpha_1, ..., \alpha_D)$ are included in the pictorial calculus. The CNOT and SWAP gates are always included in the calculus and together with single qudit gates, they provide arbitrary Clifford operations.

5.3.5 Depicting Spekkens-Schreiber toy theory for dits

We define the category **FRel** whose objects are finite sets and whose morphisms are relations. By taking the Cartesian product of sets as the tensor product, the single element set $\{\star\}$ as the identity object and the relational converse as the dagger, **FRel** can be viewed as a SMC with dagger structure.

We can then define the category **DSpek** as a subcategory of **FRel** whose objects are the single element set $I = \{\star\}$ and n-fold Cartesian products of the D^2-element set: $\mathcal{D} := \{1, 2, ..., D^2\}$.

The morphisms of **DSpek** are those generated by composition, Cartesian product and relational converse from the following relations:

(a) All $(D^2)!$ permutations $\sigma_i : \mathcal{D} \to \mathcal{D}$ of the D^2- element set.

(b) The copying relation: $\delta_Z : \mathcal{D} \to \mathcal{D} \times \mathcal{D}$ defined as:

1	2	…	D												
D	1	…	D-1												
…	…	…	…												
2	3	…	1												
				D+1	D+2	…	2D								
				2D	D+1	…	2D-1								
				…	…	…	…								
				D+2	D+3	…	D+1								
								…	…	…	…				
								…	…	…	…				
								…	…	…	…				
								…	…	…	…				
												D(D-1)+1	D(D-1)+2	…	D^2
												D^2	D(D-1)+1	…	D^2-1
												…	…	…	…
												D(D-1)+2	D(D-1)+3	…	D(D-1)+1

where there is x in the (y,z) location of the grid iff $\delta_Z : x \sim (y, z)$.

(c) The deleting relation: $\epsilon_Z : \mathcal{D} \to I$ defined as: $\{1, D+1, 2D+1, ..., D(D-1)+1\} \sim \star$.

(d) The relevant unit, associativity and symmetry natural isomorphisms.

If we interpret relations from I to n-fold tensor products of $\mathcal{D}$ as epistemic states on phase space then this category corresponds to Spekkens-Schreiber theory for dits with only states of maximal knowledge. Adding the maximally mixed state $\perp_D :: \{\star\} \sim \{1, 2, ..., D^2\}$ to **DSpek** yields the category **MDSpek**, corresponding to Spekkens-Schreiber theory for dits of dimension D.

DSpek and **MDSpek** inherit symmetric monoidal and †-compact structure from **FRel** since we can define Bell states (corresponding to compact structures) as:

$$\mu_D := \delta_Z \circ \epsilon_Z^\dagger : I \to \mathcal{D} \times \mathcal{D} :: \star \sim \{(1,1), (2,2), ..., (D^2, D^2)\} \tag{5.49}$$

We can define the other copying map as:

$$\delta_X = (\Pi_{k=1}^{D-1}\sigma_{(k+1,(kD)+1)} \otimes \Pi_{k=1}^{D-1}\sigma_{(k+1,(kD)+1)}) \circ \delta_Z \circ (\Pi_{k=1}^{D-1}\sigma_{(k+1,(kD)+1)}) \tag{5.50}$$

where $\sigma_{(j,k)}$ permutes entries j and k of the input D^2-element set (epistemic state). This map is explicitly: $\delta_X : \mathcal{D} \to \mathcal{D} \times \mathcal{D}$ such that: $\delta_X : x \sim (y, z)$ iff there is x in the (y,z) location of the following grid:

1				D+1				...	...	...	...	D(D-1)+1			
	2				D+2			...	...	...	...		D(D-1)+2		
		...				...		...	...	...	...			...	
			D				2D	...	...	...	...				D^2
D(D-1)+1				1				...	...	...	...	D(D-2)+1			
	D(D-1)+2				2			...	...	...	...		D(D-2)+2		
		...				...		...	...	...	...			...	
			D^2				D	...	...	...	...				D(D-1)
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
D+1				2D+1				...	...	...	...	1			
	D+2				2D+2			...	...	...	...		2		
		...				...		...	...	...	...			...	
			2D				3D	...	...	...	...				D

and the other erasing map as: $\epsilon_X : \mathcal{D} \to I$ such that: $\{1, 2, 3, ..., D\} \sim \star$. It is easy to check that this then gives us two strongly complementary observable structures, analogous to the Z and X observable structures in quantum theory.

In fact, we can use the fact that **3Spek** can be depicted in the qutrit ZX calculus to provide a novel proof of the following known result:

Theorem 3.7[158,262]: Spekkens-Schreiber theory for trits is operationally equivalent (meaning equivalence of preparation, transformation and measurement processes) to stabilizer theory for qutrits.

Proof. We can define the Z and X observable structures $(\mathcal{D}, \delta_Z^{(trit)}, \epsilon_Z^{(trit)})$ and $(\mathcal{D}, \delta_X^{(trit)}, \epsilon_X^{(trit)})$ for **3Spek** as described above.

The Z observable structure is: $(\mathcal{D} = \{1, 2, ..., 9\}, \delta_Z : \mathcal{D} \times \mathcal{D} \to \mathcal{D}, \epsilon_Z : \mathcal{D} \to I)$ where: $\delta_Z :: x \sim (y, z)$ iff there is x in the (y,z) location of the following grid:

1	2	3						
3	1	2						
2	3	1						
			4	5	6			
			6	4	5			
			5	6	4			
						7	8	9
						9	7	8
						8	9	7

and $\epsilon_Z : \{1, 4, 7\} \sim \star$.

The X observable structure is: $(\mathcal{D} = \{1, 2, ..., 9\}, \delta_X : \mathcal{D} \times \mathcal{D} \to \mathcal{D}, \epsilon_X) : \mathcal{D} \to I$

where: $\delta_X = ((\sigma_{(2,4)} \circ \sigma_{(3,7)}) \otimes (\sigma_{(2,4)} \circ \sigma_{(3,7)})) \circ \delta_Z \circ (\sigma_{(2,4)} \circ \sigma_{(3,7)})$ defined as usual by the table:

1			4			7		
	2			5			8	
		3			6			9
7			1			4		
	8			2			5	
		9			3			6
4			7			1		
	5			8			2	
		6			9			3

and $\epsilon_X : \{1,2,3\} \sim \star$.

The Z observable structure has three classical states:

$$z_0 :: \star \sim \{1,2,3\} \ ; \ z_1 :: \star \sim \{4,5,6\} \ ; \ z_2 :: \star \sim \{7,8,9\} \tag{5.51}$$

and nine unbiased states:

$$x_0 :: \star \sim \{1,4,7\} \ ; \ x_1 :: \star \sim \{2,5,8\} \ ; \ x_2 :: \star \sim \{3,6,9\}$$

$$(xz)_0 :: \star \sim \{1,6,8\} \ ; \ (xz)_1 :: \star \sim \{2,4,9\} \ ; \ (xz)_2 :: \star \sim \{3,5,7\}$$

$$(xz^2)_0 :: \star \sim \{1,5,9\} \ ; \ (xz^2)_1 :: \star \sim \{2,6,7\} \ ; \ (xz^2)_2 :: \star \sim \{3,4,6\} \tag{5.52}$$

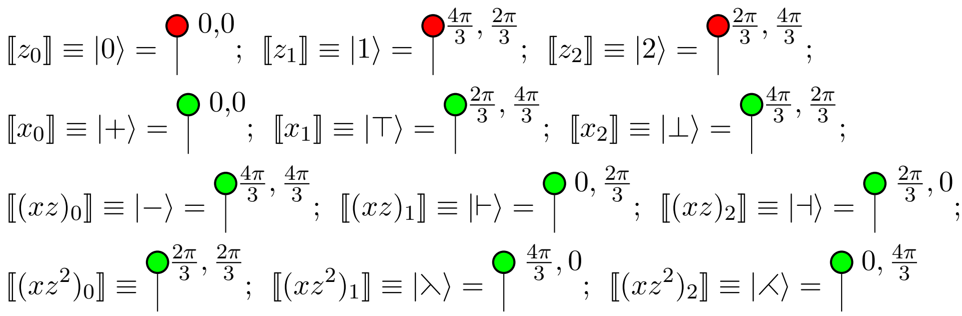
The X observable also has 3 classical states x_0, x_1, x_2 and nine unbiased states z_0, z_1, z_2, $(xz)_0, (xz)_1, (xz)_2$, $(xz^2)_0, (xz^2)_1$ and $(xz^2)_2$.

Similarly, one could define two other observable structures corresponding to "XZ" and "XZ^2" which each have three classical states and nine unbiased states. Therefore, **3Spek** contains 4 observable structures and 12 single trit states. Also, there are 12 (measurement) effects, corresponding to taking the converse relations of the 12 states.

If we define phase maps: $\Lambda_Z(\psi) := (\delta_Z^{(trit)})^\dagger(\psi \otimes 1_\mathcal{D}) : \mathcal{D} \to \mathcal{D}$ and $\Lambda_X(\psi) := (\delta_X^{(trit)})^\dagger(\psi \otimes 1_\mathcal{D}) : \mathcal{D} \to \mathcal{D}$, then it is clear that the *phase group* of **3Spek** is $\mathbb{Z}_3 \times \mathbb{Z}_3$.

Therefore, both **3Spek** and **3Stab** can be expressed in the qutrit ZX calculus as mutually unbiased qutrit theories with twelve states and phase group $\mathbb{Z}_3 \times \mathbb{Z}_3$. This uniquely determines all the allowable preperations, measurements and transformations (compositions of spiders with phases adding according to $\mathbb{Z}_3 \times \mathbb{Z}_3$).

Explicitly, we can associate the twelve single trit states of **3Spek** with those of **3Stab** according to:



(5.53)

The post-selected measurements (indicator functions) are the adjoints of these and the reversible transformations in **3Spek** are mapped to the reversible of **3Stab** corresponding to the same spider diagram. This means that our previous discussion of qutrit stabilizer quantum theory carries through to **3Spek**. □

We will now study **DSpek** in the four-dimensional case.

Now, the Z observable structure has a copying map δ_Z given by:

1	2	3	4												
4	1	2	3												
3	4	1	2												
2	3	4	1												
				5	6	7	8								
				8	5	6	7								
				7	8	5	6								
				6	7	8	5								
								9	10	11	12				
								12	9	10	11				
								11	12	9	10				
								10	11	12	9				
												13	14	15	16
												16	13	14	15
												15	16	13	14
												14	15	16	13

and a deleting map: $\epsilon_Z :: \{1, 5, 9, 13\} \sim \star$. This has classical points:

$$z_0 :: \star \sim \{1, 2, 3, 4\} \tag{5.54}$$

$$z_1 :: \star \sim \{5, 6, 7, 8\} \tag{5.55}$$

$$z_2 :: \star \sim \{9, 10, 11, 12\} \tag{5.56}$$

$$z_3 :: \star \sim \{13, 14, 15, 16\} \tag{5.57}$$

.

And similarly, the X observable has a copying map δ_X described by:

1				5				9				13			
	2				6				10				14		
		3				7				11				15	
			4				8				12				16
13				1				5				9			
	14				2				6				10		
		15				3				7				11	
			16				4				8				12
9				13				1				5			
	10				14				2				6		
		11				15				3				7	
			12				16				4				8
5				9				13				1			
	6				10				14				2		
		7				11				15				3	
			8				12				16				4

and a deleting map $\epsilon_X :: \{1, 2, 3, 4\} \sim \star$. The classical points of the X observable structure are:

$$x_0 :: \star \sim \{1, 5, 9, 13\} \equiv |\{0, 0\}_Z\rangle \tag{5.58}$$

$$x_1 :: \star \sim \{2, 6, 10, 14\} \equiv \left|\{\frac{\pi}{2}, 0\}_Z\right\rangle \tag{5.59}$$

$$x_2 :: \star \sim \{3, 7, 11, 15\} \equiv |\{\pi, 0\}_Z\rangle \tag{5.60}$$

$$x_3 :: \star \sim \{4, 8, 12, 16\} \equiv \left|\{\frac{3\pi}{2}, 0\}_Z\right\rangle \tag{5.61}$$

The Bell state is given by:

$$BELL := \{(k,k)|k = 1, 2, ...16\} = \delta_Z \circ \epsilon_Z^\dagger = \delta_X \circ \epsilon_X^\dagger \tag{5.62}$$

There are now 12 unbiased points for both the Z and X observables, such that $\delta_Z^\dagger \circ (p_i \times p_i)\lambda_I = \epsilon_Z$ and $\delta_X^\dagger \circ (p_i \times p_i)\lambda_I = \epsilon_X$, corresponding to distinct epistemic states. These are:

$$u_1 :: \star \sim \{1,8,11,14\} \equiv \left|\{0, \frac{\pi}{2}\}_Z\right\rangle;\ u_2 :: \star \sim \{2,5,12,15\} \equiv \left|\{\frac{\pi}{2}, \frac{\pi}{2}\}_Z\right\rangle;$$
$$u_3 :: \star \sim \{3,6,9,16\} \equiv \left|\{\pi, \frac{\pi}{2}\}_Z\right\rangle;\ u_4 :: \star \sim \{4,7,10,13\} \equiv \left|\{\frac{3\pi}{2}, \frac{\pi}{2}\}_Z\right\rangle;$$
$$u_5 :: \star \sim \{1,6,11,16\} \equiv |\{0, \pi\}_Z\rangle;\ u_6 :: \star \sim \{2,7,12,13\} \equiv \left|\{\frac{\pi}{2}, \pi\}_Z\right\rangle;$$
$$u_7 :: \star \sim \{3,8,9,14\} \equiv |\{\pi, \pi\}_Z\rangle;\ u_8 :: \star \sim \{4,5,10,15\} \equiv \left|\{\frac{3\pi}{2}, \pi\}_Z\right\rangle;$$

$$u_9 :: \star \sim \{1,7,10,16\} \equiv \left|\{0, \frac{3\pi}{2}\}_Z\right\rangle;\ u_{10} :: \star \sim \{2,8,11,13\} \equiv \left|\{\frac{\pi}{2}, \frac{3\pi}{2}\}_Z\right\rangle; \tag{5.63}$$

$$u_{11} :: \star \sim \{3,5,12,14\} \equiv \left|\{\pi, \frac{3\pi}{2}\}_Z\right\rangle;\ u_{12} :: \star \sim \{4,6,9,15\} \equiv \left|\{\frac{3\pi}{2}, \frac{3\pi}{2}\}_Z\right\rangle \tag{5.64}$$

Note the redundancy in the unbiased points (similar to a choice of global phase) that leads us to keep only half of the 24=4! relations that satisfy the unbiasedness relation for Z and X. So we can see that the group of unbiased points for Z (or X) in **4Spek** can be interpreted as the direct product of the $\mathbb{Z}_4$ group corresponding to the 'position' part of the phase space in Spekkens-Schreiber theory and a $\mathbb{Z}_4$ group corresponding to the 'momentum' part of the phase space in Spekkens-Schreiber theory. Note that theorem 3.4 is satisfied since the group $\mathbb{Z}_4$ of classical points for the X observable (or Z observable) is a subgroup of the unbiased group $\mathbb{Z}_4 \times \mathbb{Z}_4$ for the Z observable (or X observable).

Therefore, the phase group of **4Spek** is $\mathbb{Z}_4 \times \mathbb{Z}_4$.

In general, **DSpek** contains D classical points of the Z (or X) observable and D^2 unbiased points for the Z (or X) observable structures. Note that: $D!+D = D^2$ iff $D = 2, 3$, otherwise there is a redundancy in the relations satisfying the unbiasedness condition which means that $D! + D - D^2$ of them must be discarded since they correspond to a repeated epistemic state (with the same isotropic subspace and valuation vector). Therefore, the phase group for **DSpek** in general is $\mathbb{Z}_D \times \mathbb{Z}_D$.

This should allow us to depict these theories for any dimension using (a version of) the qudit ZX calculus. We can then study the relationship between Spekkens-Schreiber theory for dits and qudit stabilizer theory in the general case.

5.4 A periodic table of quantum-like theories

We have shown how to study operational physical theories using symmetric monoidal categories and diagrammatic calculi. The key ingredient in our analysis has been the **phase group**. Isolating this particular feature provides a method for classification and yields a periodic table of quantum-like operational theories, described by the Phase Group.

Definition 4.1: Let Π be an Abelian group. We can interpret this group as a category $\mathcal{P}$ with a single object X and arrows from X to X, corresponding to the underlying set of Π. Let $\mathcal{F}_{SMC}(ZX_D)/\equiv_{ZX_D}$ be the free symmetric monoidal category of the ZX calculus for qudits in dimension D, quotient to the axioms of the qudit ZX calculus.

We can map the phase group to a symmetric monoidal category defined using the qudit calculus corresponding to a MUQT, which allows us to classify alternative operational theories by using their phase group. This yields the following Periodic Table of Quantum-like theories:

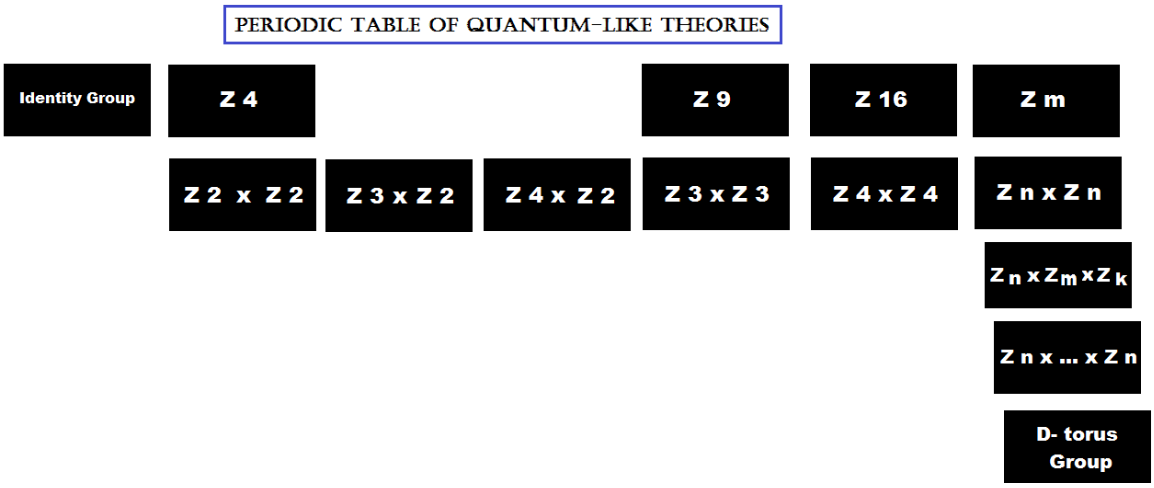


Figure 5.9: Periodic table of quantum-like theories.

This provides a classification of physical theories arising from fundamental symmetry within the framework, as illustrated in Figure 5.9. Note that the horizontal axis represents the order of the phase group and the vertical axis represents the number of direct products of component cyclic groups. We can summarize by recalling the phase groups corresponding to the theories we have analyzed:

In two dimensions– Spekkens' theory: $\mathbb{Z}_2 \times \mathbb{Z}_2$ and Stabilizer theory: $\mathbb{Z}_4$

In three dimensions– Spekkens' theory and Stabilizer theory: $\mathbb{Z}_3 \times \mathbb{Z}_3$

In four dimensions– Spekkens' theory: $\mathbb{Z}_4 \times \mathbb{Z}_4$ and Stabilizer theory: $\mathbb{Z}_4 \times \mathbb{Z}_2$

Quantum theory– Torus group $S^1 \times ... \times S^1$.

A natural question involves whether this periodic table can be extended to include more groups, such as non-Abelian groups and Lie groups.

Note that it could be interesting to interpret the phase group as a Galois group. An operational theory can then be identified in terms of a field extension (of the rational numbers $\mathbb{Q}$, for instance).

Each physical theory can be associated with a collection of polynomials, corresponding to a specific field extension of the rational numbers $\mathbb{Q}$. The analysis of operational theories through the phase group then follows from the application of Galois theory. The phase group arises from a **fundamental polynomial** of a physical theory, by the fundamental theorem of Galois theory.

Example 5.1:

Consider the trivial field extension of the rationals $\mathbb{Q}/\mathbb{Q}$. The phase field is in correspondence with any polynomial which has only rational roots, for example (x-2)2, or (x-2)(x-1). The Galois group is then the trivial group and therefore corresponds to a trivial operational theory, where the only physical process is the identity map.

Example 5.2:

Consider the field extension of the rationals $\mathbb{Q}(\sqrt{2}, \sqrt{3})/\mathbb{Q}$.

This has the fundamental polynomial: p(x) $= x^4 - 10x^2 + 1$, shown in Figure 5.10.

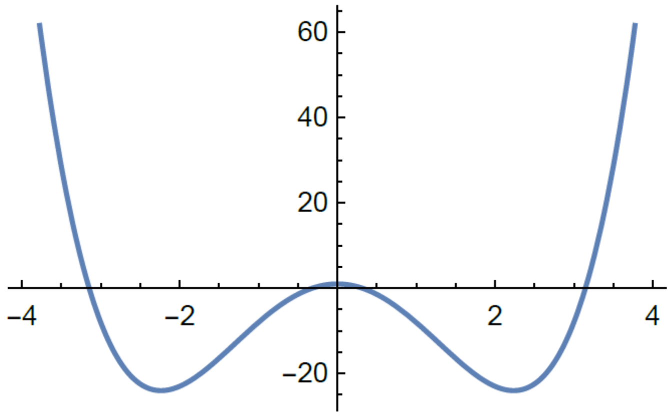


Figure 5.10: Plot of the fundamental polynomial for Spekkens toy theory.

Therefore, the corresponding Galois group is the phase group $Gal(p) = \mathbb{Z}_2 \times \mathbb{Z}_2$ so this theory is Spekkens toy theory (in two dimensions).

Example 5.3:

Consider the field extension of the rationals $\mathbb{Q}(e^{\frac{2\pi i}{5}}, e^{\frac{4\pi i}{5}})/\mathbb{Q}$. This field admits the fundamental polynomial: p(x) $= x^4 + x^3 + x^2 + x + 1$ shown in Figure 5.11.

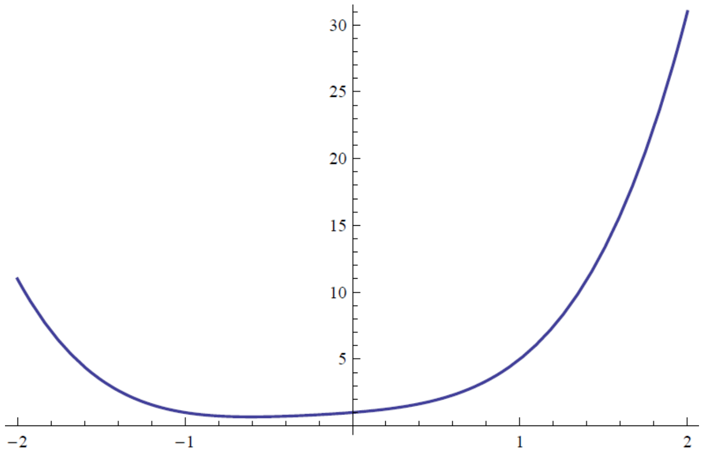


Figure 5.11: Plot of the fundamental polynomial for stabilizer quantum mechanics.

Therefore, the corresponding Galois group is the phase group $Gal(p) = \mathbb{Z}_4$ so this theory is stabilizer quantum theory (in two dimensions).

In fact, we can consider a quartic fundamental polynomial

$$f(x) = x^4 + ax^2 + b \tag{5.65}$$

with $a, b \in \mathbb{Z}$, which has roots $\{\pm\alpha, \pm\beta\}$ and take $\alpha^2, \alpha \pm \beta \in \mathbb{Q}$ so that f(x) is irreducible. Then the phase group $\Pi = Gal(\mathbb{Q}(\alpha, \beta)/\mathbb{Q})$ corresponding to the fundamental polynomial f is isomorphic to[96]:

(i) $\mathbb{Z}_2 \times \mathbb{Z}_2$ iff $\alpha\beta \in \mathbb{Q}$.

(ii) $\mathbb{Z}_4$ iff $\mathbb{Q}(\alpha, \beta) = \mathbb{Q}(\alpha^2)$.

(iii) $\mathbb{Z}_4 \times \mathbb{Z}_2$ iff $\alpha\beta \notin \mathbb{Q}(\alpha^2)$.

Cases (i), (ii) and (iii) respectively correspond to Spekkens toy theory, stabilizer quantum mechanics in two dimensions and stabilizer quantum mechanics in four dimensions.

As we can see from this example, this method provides an efficient way of classifying quantum-like theories, through the features of fundamental polynomials.

5.5 Topological Ontological models

The previous three levels of analysis of quantum-like theories have solely focused on an operational interpretation of these theories, without seeking any ontological significance for the theoretical constructs used to define physical theories. In this section, we aim to provide a realist ontic level of analysis of alternative physical theories, based upon an extension of the usual measure space ontic model of Bell [48], Harrigan, Spekkens and Rudolph [165,167,166].

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 a parameter λ to reproduce quantum statistical predictions. If we 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. Stressing this point is the main goal of this section.

Thus far in the study of ontological models, several tacit mathematical assumptions have been made with regard to the nature of the ontic space. The main assumption we shall question here is that the ontic space must be a measure space. It is clear that the capacity to define integration and thereby associate a number to subspaces of the ontic space is a valuable and desirable feature to retrieve the operational theories from our posited underlying reality. Without the measure space structure it is difficult to account for probabilities and measurement structure.

Nevertheless, it feels over-simplified to assume that a mathematical object aiming to describe something called "underlying reality" should pander to our desire to associate numbers to physical objects and procedures. Moreover, if we seek to define ontological models for alternative operational theories then we should allow for greater generality. The aim to directly reproduce quantum theory from underlying ontic assumptions is then no longer the prime concern. This leads to the notion of meta-ontological models, where the ontic space can be any mathematical object and all transformations are general abstractions of those for standard ontological models.

In the following, we question the assumption that the ontic space must always be a measure space. This leads to the introduction of **topological ontological models**, where the ontic space is a topological space. We also discuss how these models relate to the current measure space ontic models.

We shall now restrict the mathematical form of ontic spaces to topological spaces and introduce the notion of a **topological ontic space**. This will lead to an alternative framework for ontological models where topology is at the heart.

Let us first of all take the ontic space Λ to be a topological space with a topology τ. We can then define a topological ontic model in the following way:

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

(ii) An operational preparation procedure within a physical theory can be obtained from an incomplete description of the underlying reality. This is defined by introducing a measure μ, constructed from the Borel sigma-algebra $\mathrm{B}(\Lambda)$ generated by all the open sets in the topological ontic space Λ. Preparation procedures can then be obtained from the measure μ by defining a distribution:

$$|\psi\rangle \leftrightarrow (\mu(\lambda)) \tag{5.66}$$

(iii) Measurements correspond to introducing an ensemble of **separated sets**, corresponding to subsets of the ontic topological space Λ that are neither overlapping nor touching. The exact notion of separation to be used is related to the Trennungsaxiom Hierarchy which we introduced in Chapter 2. Indeed, depending on which separation axiom applies to the topological ontic state, we can call subsets L_1, L_2 of Λ separated if one of the following holds:

(1) L_1 and L_2 are disjoint, meaning that their intersection is empty.

(2) L_1 and L_2 are disjoint from each other's closure.

(3) L_1 and L_2 are separated by neighborhoods, meaning that there are neighborhoods U_1 of L_1 and U_2 of L_2 such that U_1 and U_2 are disjoint.

(4) L_1 and L_2 are separated by a function, meaning that there exists a continuous function $f : \Lambda \to \mathbb{R}$ such that $f(L_1) = 0$ and $f(L_2) = 1$.

Measurement results are obtained by testing for the inclusion of the ontic state $\lambda \in \Lambda$ into one of the separated subsets.

As before, we can construct measures ξ_i constructed on the Borel sigma-algebra $\mathrm{B}(L_i)$, generated by all the open sets, in each separated topological subspace L_i. Measurement procedures can then be obtained from the measures $\{\xi_i\}$, which can be used to define distributions $\{\xi_i(\lambda)\}$ in the usual way. These satisfy:

$$0 \leq \xi_i(\lambda) \leq 1 \text{ and } \textstyle\sum_i \xi_i(\lambda) = 1, \text{ for all } \lambda.$$

(iv) The probability of getting outcome k for a measurement M given preparation P is then given by 'averaging ' over the measure space obtained via the ontic space through the use of the measures μ and ξ_i, which we previously defined.

$$p(i|\mu, M) = \langle \xi_i(\lambda)\mu(\lambda)\rangle_\Lambda := \int d\lambda \xi_i(\lambda)\mu(\lambda) \tag{5.67}$$

This allows us to compare the predictions of the ontological model with the operational framework we wish to consider, as in the case of the standard ontological models. One could also, however, decide that the transition from the topological ontic model formalism to the measure space framework which allows us to make statistical predictions requires an excessive loss of information and that predictive power weakens the model's aptitude to approximate "underlying reality".

(v) Transformation of the topological ontic space Λ correspond to continuous maps. Also, measurements can disturb the space Λ and the model should account for this by defining continuous measurement maps.

Borel measure spaces, which are the mathematical object used to define random variables and probability spaces, arise as a special case of topological spaces. Therefore, we can recover the usual structure of ontological models of quantum mechanics as a special case of the topological ontic model formalism. Naturally, this may require restrictions on the allowable topological ontic spaces and we expect a trade-off between abstraction and the reproduction of predictions of operational theories. For example, practical considerations may dictate that the ontic topological space Λ should be restricted to a metrizable space, and obey the conditions from the metrization theorems from Chapter 2. In general, what types of topological spaces should we use as topological ontic spaces for quantum-like theories?

Furthermore, we stress that the theoretical analysis of topological ontic models could be conducted independently from the retrieval of familiar measure-theoretic notions. In the present section, an effort was made to ensure that we can associate values with measurement results and reproduce the predictions of operational physical theories, even if this means that a process of approximation is inevitable. In future work, it will be important to consider methods for obtaining real numbers as the results of physical measurements, which come directly from topological methods, perhaps through the use of *sheaf theory*.

Could we define notions of psi-ontic, psi-epistemic and psi-calculational topological ontic models, independently of measure-theoretic structure? Is it possible that the use of abstract mathematical objects to describe physical reality might provide a new light on no-go results such as the Bell, Kochen-Specker and PBR theorems? Could mathematical intricacy and abstraction provide a novel defense of psi-epistemic interpretations of quantum theory?

Another interesting direction is to add a manifold structure to topological ontic spaces. The key question would then be to understand the meaning of these ontic manifolds and whether they may be related to our notions of space-time. Naturally, imposing additional mathematical structure to the ontic space reduces the likelihood that our abstract domain of discourse can claim any ontological significance.

Finally, we can also consider the idea of using category theory to describe the ontological space which underlies our operational physical theories. This leads to the notion of a **categorical ontic model**, where the ontic space is modelled by a category.

A possible method of comparing predictions of the ontological model with operational frameworks is by relating the categorical ontic space to the category **Meas** of measure spaces and measure preserving maps. measure preserving maps. In future work it would be desirable to bypass the use of measure spaces and rely on a more direct method to relate the underlying ontology with operational predictions.

5.6 Further work

We will conclude this chapter with a brief outline of possible avenues of research which follow from the work presented here.

The framework we have presented is rather incomplete in a number of respects. We have hardly touched on the relationship between the five levels of analysis. A particular point which requires further analysis is the ontological significance of the phase group and of Galois particles.

As we mentioned earlier, understanding how the qubit calculus fits into the general qudit calculus and proving the completeness of the generalized qudit ZX calculus for stabilizer quantum mechanics would certainly provide new insights into qudit stabilizer quantum mechanics. This might lead to modifications of the qudit ZX calculus before it reaches its final form[27].

For example, can the qudit ZX calculus be expressed without angles by adding axioms relating to graph structure[120] or use multiple edges between vertices? This approach could simplify proofs of completeness or provide a graphical depiction of non-locality. Another possible mathematical framework for studying the qudit ZX calculus would be to use product and permutation categories (PROPs)[62]. This approach may yield an elegant synthetic axiomatization of numerous physical process theories and could provide new completeness theorems for corresponding graphical calculi.

On a more practical note, the calculus for qudit stabilizer quantum theory can help generalize qubit protocols to qudits and understand new features of familiar quantum processes. For example, the formalism could be used to give a general description of error correction and fault tolerance for qudits, such that links can be made between error correction in various dimensions. Furthermore, getting new insights into the abstract structure of qudit quantum mechanics could play a pivotal role in the development of new quantum algorithms.

There are also a number of quantum foundations questions which could be addressed next. It would be interesting to develop the periodic table of quantum-like theories and include more explicit examples. For instance, we know that the single qudit stabilizer theory is operationally equivalent to Spekkens-Schreiber theory for dits for *finite odd dimensions* and therefore admits a non-contextual, local hidden variable model in those cases. But what is the relationship between qudit stabilizer theory and Spekkens-Schreiber toy theory in general? We could also study van Enk's toy model [291] as a MUQT and find its phase group.

More generally, it would be useful to classify all the mutually unbiased qudit theories and determine which physical features each one exhibits. For example, we can build on previous work aiming to elucidate the relationship between a theory's phase group and whether it admits a local hidden variable interpretation [91,152]. The study of the qudit ZX calculus with different Abelian phase groups should produce a large class of interesting toy models. In the future, we could also consider theories where distinct observable structures have phase groups that are non-Abelian or Lie groups.

Moreover, the qudit ZX calculus provides the ideal framework to study other similar foundational questions related to complementarity. We could, for example, use the categorical framework to study how various notions of complementarity arise in different dimensions. Can one find a pictorial calculus which captures complementarity of more than two observables?

Finally, it would be interesting to understand the interpretation of the D-torus phase groups for qudit quantum mechanics observables from a physical point of view. Perhaps studying the operational interpretation of phase [141] in physical theories could help us find the physical reason for each phase group taking the form it does. The study of phase and complementarity from an operational point of view may also provide insight into the relationship between categorical quantum mechanics and generalized probabilistic theories.

Chapter 6

Quantum collapse theories and Quantum Integrated Information

Throughout this thesis, we have analyzed possible formulations of quantum theory and alternative theories in quantum foundations. In this final chapter, we will pursue this same objective from a different angle, through the study of **quantum collapse theories**.

As we have discussed, quantum theory admits the delicate coexistence of two radically different dynamics. Unobserved systems undergo linear, deterministic, unitary evolution whereas observation causes a non-linear, probabilistic, non-unitary "collapse" of the quantum state. In addition, the ontological significance of the quantum state is unclear. Moreover, the quantum superposition of distinguishable states and the arising of probabilities seem to contradict the behavior we observe in macroscopic systems. Is there a classical/quantum divide and if so, where does it lie?

These issues are inextricably related to the impossibility of separating the physical system under examination from the observer acquiring knowledge about the system. If we admit that measuring devices should be described by the same dynamical equations as the systems under consideration, then why does the measurement process break the superposition of states? This leads us to follow Bell[48] in asking:

"What exactly qualifies some physical systems to play the role of 'measurer'?"

In a joint project with Kobi Kremnizer[192], we aim to provide a potential answer to this question. We postulate that physical systems act more or less as measuring devices depending on how much they exhibit a property called **quantum Integrated Information** (QII). This leads us to outline a novel, experimentally falsifiable theory with a universal dynamics depending on the levels of QII of physical systems.

There have been numerous proposals to replace both unitary and measurement dynamics by a single, universal dynamics governing all physical processes [230,146,147,113,232,7]. Such a dynamical theory could be described using a non-linear, stochastic differential equation which does not allow superluminal signaling. This equation is expected to both reduce to Schrödinger's equation in the quantum regime and also provide an accurate description of the classical behavior of macroscopic objects.

We stress that such a model aims to describe the physical world from an ontological perspective, whether or not any act of observation takes place. Knowledge about physical systems plays no fundamental role.

An important question which naturally arises is the basis which should be chosen for the localization of the wavefunction. From our experience of macroscopic superpositions rapidly collapsing into localized states, it may seem that position should be considered as a privileged basis for collapse. We will discuss the role that relevant properties of the physical state could play in determining the basis on which the wavefunction is localized.

From a phenomenological point of view, all space collapse models are equivalent: they induce a collapse of the wavefunction in space, such that the collapse rate depends on the size of the system. The assumption that the speed of localization of the system in space depends only on the size of the system but on none of its other properties seems rather ad hoc and naive.

The key idea we explore here is that the relevant property of a physical system affecting the rate of collapse of the state might not be its size (or mass distribution) but should rather be related to its informational complexity.

This naturally follows from the idea that quantum mechanical observers are expected to exhibit some form of 'consciousness' which induces the wavefunction collapse. We take the view that consciousness plays a crucial role in quantum collapse and that conscious perceptions do not obey the linear laws of quantum mechanics. This leads to the difficult problem of finding a physicalist measure of consciousness. In the present work, we make no claims of having resolved this intricate philosophical issue but instead we take a working approach to this problem.

For the purpose of the present theory, we use a modified version of an existing 'measure of consciousness', called Integrated Information (II) [287,288]. The II of a physical system is defined as the information of the whole system above and beyond the information contained in its parts.

We introduce a quantum version of this measure, called Quantum Integrated Information (QII), which enables us to explicitly present a novel ***Integrated Information-induced collapse theory***.

This theory may be interpreted as a modification of existing collapse models, where the rate of collapse of states is determined by a specific feature of their informational complexity: the QII. We believe that this already provides an important conceptual shift, even if QII is completely unrelated to consciousness.

This chapter will spend some time presenting the philosophy of consciousness and previous quantum collapse models. We will then introduce quantum Integrated Information and present the universal theory of Integrated Information-induced collapse. We shall also describe potential experimental tests of the new theory in realms where it might not agree with quantum mechanics. Finally, we will discuss some of the modifications we might expect this collapse theory to undergo and sketch some issues that may arise.

6.1 The philosophy of consciousness

6.1.1 History

Despite the ubiquity of doubt in human experience, Descartes [107] encounters certitude through the process of thought: "cogito ergo sum". This can be interpreted as an inductive definition of existence through consciousness. Indeed, Descartes [106] states that: "By the word 'thought', I understand all that of which we are conscious as operating in us" and "ainsi l'activité de l'esprit et la conscience me caractérisent : la conscience est l'essence de la pensée".

Similarly, Locke uses consciousness as a cornerstone of his theory of personal identity [211]: "[A person] can consider it self as it self, the same thinking thing in different times and places; which it does only by that consciousness, which is inseparable from thinking, and as it seems to me essential to it: It being impossible for any one to perceive, without perceiving that he does perceive".

Aiming to understand the essence of this key concept, Leibniz [204] presented the analogy of the mill:

"Supposing that there were a machine whose structure produced thought, sensation, and perception, we could conceive of it as increased in size with the same proportions until one was able to enter into its interior, as he would into a mill. Now, on going into it he would find only pieces working upon one another, but never would he find anything to explain perception. It is accordingly in the simple substance, and not in the compound nor in a machine that the perception is to be sought. Furthermore, there is nothing besides perceptions and their changes to be found in the simple substance. And it is in these alone that all the internal activities of the simple substance can consist".

This denial that consciousness and perception are constricted to the physical world of matter is a recurrent argument in the philosophy of consciousness. Indeed the 19th century biologist Huxley colorfully asked: "How it is that anything so remarkable as a state of consciousness comes about as a result of irritating nervous tissue, is just as unaccountable as the appearance of the Djin, when Aladdin rubbed his lamp"?

To Kant, the unity of consciousness [178] is an essential feature of the human mind:

"The experiences must have a single common subject [...] The consciousness that this subject has of represented objects and/or representations must be unified".

The manner in which our experience is tied together through consciousness is an essential Kantian justification for truth in mathematics and physics and reflects the way that physical objects in the world must be tied together [178]: "If, therefore, there exist any pure a priori concepts, they cannot indeed contain anything empirical; they must, nevertheless, all be a priori conditions of a possible experience, for on this ground alone can their objective reality rest".

This integration of conscious thought led William James [174] to describe the stream of consciousness: "Consciousness, then, does not appear to itself chopped up in bits [...] it is nothing jointed; it flows. A 'river' or a 'stream' are the metaphors by which it is most naturally described".

Analyzing the unity of the conscious mind has played an important role in historical debates on consciousness[172,64,254,174]. For example, investigating the limitations in the range of psychological phenomena over which unified consciousness ranges and whether most of what goes on in our mind is due to conscious thought led Freud to popularize the idea of the subconscious: "The conscious mind may be compared to a fountain playing in the sun and falling back into the great subterranean pool of subconscious from which it rises".

We will end this section by mentioning that considerable scientific progress has recently been made in understanding the neural basis for consciousness[101,265,38].

6.1.2 Philosophical positions

To define the term 'consciousness', we can borrow one of the multifarious definitions given by the Oxford English Dictionary: "The state or faculty of being conscious, as a condition and concomitant of all thought, feeling and volition".

We can distinguish three broad positions[57] concerning the nature of consciousness. The first of these states that consciousness cannot be understood in a materialist ontology but requires an immaterial explanation. This interpretation is an extension of Cartesian dualism, with the realm of res cogitans[106], or of Karl Popper's World 2 of mental objects and events[245]. We will not thoroughly investigate the denial of a physical basis for consciousness but it is important to remember the progress made by avenues of human inquiry that are not seeking a scientist's reductionist materialist ontology.

A second position doubts that consciousness is a coherent philosophical concept and/or denies that human beings have the mental capabilities to comprehend their own state of consciousness. It can be argued that it is impossible to bridge the "explanatory gap"[205] between the material brain and the lived world of conscious experience. Are we even capable of understanding[216] how "the water of the physical brain is turned into the wine of consciousness"? Some philosophers place the concept of consciousness on the same footing as ghosts and ether, concepts that, according to Churchland[15]: "under the suasion of a variety of empirical-cum-theoretical forces [...] lose their integrity and fall apart".

The third position states that consciousness is a natural physical phenomenon, intricate and complex but not beyond analysis using an advanced scientific framework. The goal is then to figure out how the diverse fields of philosophy, psychology, neuroscience, computer science, physics, physiology can work together to provide a physicalist framework for consciousness.

Using methodology from the study of animal behavior, we can attempt a scientific analysis of consciousness by asking Tinbergen's four questions [285].

(i) Function: Why does consciousness exist? Does it have a function, and if so what is it? Does it affect the operation of the environment which contains it, and if so how and why?

(ii) Evolution: How did consciousness come to exist? Did evolution through natural selection play a crucial role? Can consciousness arise from nonconscious entities and processes?

(iii) Development: How does consciousness arise in individuals? What is the process explaining its genesis through reproduction and embryonic growth? What genetic and environmental factors play a key role in the development period?

(iv) Causation: What is consciousness? What are its physical features and how can these be modeled? Does it act causally and if so with what types of effects and mechanisms? What defines a conscious being and where is the locus of consciousness?

6.1.3 Problems

In 'What is it like to be a bat?' Thomas Nagel argues that the essential component of consciousness is that there is something that it is (or feels) like to be a particular conscious thing [223]: "But fundamentally an organism has conscious mental states if and only if there is something it is like to be that organism – something it is like for the organism".

Nagel reasons that:

"[...] if the facts of experience– facts about what it is like for the experiencing organism– are accessible only from one point of view, then it is a mystery how the true character of experiences could be revealed in the physical operation of that organism. [...] A Martian scientist with no understanding of visual perception could understand the rainbow, or lightning, or clouds as physical phenomena, though he would never be able to understand the human concepts of rainbow, lightning, or cloud".

There is an asymmetry between our understanding and access to our own consciousness compared with that of other beings: this is the first person versus third person problem.

Does this mean that one cannot comprehend the consciousness of others and moreover our own self consciousness is incapable of understanding itself, since[220]: “Turning a tool on itself may be as futile as trying to soar off the ground by a tug at one’s bootstraps”?

Block[56] has introduced a distinction between **access** consciousness and **phenomenal** consciousness: “Phenomenal consciousness is experience; the phenomenally conscious aspect of a state is what it is like to be in that state. The mark of access-consciousness, by contrast, is availability for use in reasoning and rationally guiding speech and action”.

This leads to a contrast between representational access consciousness (such as thoughts beliefs and desires) used in reasoning and experiential phenomenal consciousness (resulting from sensory experiences) corresponding to ‘what is is like’ to be in a state.

One can then introduce **qualia**, or instances of subjective conscious experiences, which are at the heart of the philosophy of consciousness. Qualia cannot be reduced to physical information or communicated but they are private and immediately apprehensible to the subject of a phenomenal experience.

This notion is well captured by Schrödinger’s statement that[264]: “The sensation of color cannot be accounted for by the physicist’s objective picture of light-waves. Could the physiologist account for it, if he had fuller knowledge than he has of the processes in the retina and the nervous processes set up by them in the optical nerve bundles and in the brain? I do not think so”.

The idea that qualia do not affect the course of physical events has led to interesting philosophical inquiry. The inverted spectrum thought experiment, first introduced by Locke[211], asks whether it is conceivable that we could wake up one day to find that two colours have been inverted, whilst no physical change has occurred that would explain the phenomenon.

Similarly, one can define philosophical zombies[193,74] which are beings whose behavior, functional, and physical structure are identical to those of normal human beings but who lack any conscious experience.

Of course, the notion of qualia is not universally accepted. Dennett[104], for example, has argued that: “conscious experience has no properties that are special in any of the ways qualia have been supposed to be special”.

We mention in passing Wittgenstein’s denial of the existence of a private language[299,300] where: "The words of this language are to refer to what can be known only to the speaker; to his immediate, private, sensations". He argues[299] that such a language must be unintelligible to its supposed originator and that another cannot understand the language.

David Chalmers[75] has introduced a distinction between the easy and hard problems of consciousness. Chalmers lists some of the easy problems, which could readily be understood through computational and neural mechanisms:

"• the ability to discriminate, categorize, and react to environmental stimuli;

- the integration of information by a cognitive system;
- the reportability of mental states;
- the ability of a system to access its own internal states;
- the focus of attention;
- the deliberate control of behavior;
- the difference between wakefulness and sleep".

The problems of conscious experience, phenomenal consciousness and qualia, on the other hand, are described as hard in the sense that they may elude any scientific explanation.

6.2 Consciousness and Integrated Information

It has been suggested that physical systems exhibiting consciousness must satisfy two fundamental properties[23,219,287,288]. Firstly, *differentiation of information* states that consciousness should allow discrimination of a single possibility amongst a vast repertoire of possible states, leading to the acquisition of information. Secondly, *integration* is the feature that this differentiation should be performed by a unified physical system, not decomposable into a collection of independent parts.

These concepts can be illustrated[287] by considering two unconscious physical systems. On the one hand, a digital camera with a million photodiodes exhibits a high level of differentiation but very little integration since it can enter a large number of distinct states but each photodiode acts independently. On the other hand, a million Christmas lights connected to a single switch exhibit a large amount of integration but almost no differentiation since either all the lights are on or they are all off. Both of these examples are in contrast with the neural networks associated with consciousness in the human brain, since such physical systems are known to exhibit high levels of both differentiation of information and integration [215,29].

This observation hints that the amount of 'consciousness' a physical system may manifest can be related to how much it exhibits a property called Integrated Information [287,288].

For our purpose, we define quantum Integrated Information (QII) as a general property of a quantum system, which corresponds to how much information the parts of a physical system contain above and beyond the information generated by the system as a whole. Therefore QII embodies this particular definition of consciousness as the capacity to process information in an integrated way.

<u>Definition:</u> Given a quantum system in a Hilbert space $\mathcal{H}$ described by a density matrix ρ, we define the system's quantum Integrated Information as:

$$\Phi(\rho) = \inf S(\rho || \bigotimes_{i=1}^{N} Tr_{\bar{i}}(\rho)) : \mathcal{H} \cong_{\phi} \mathcal{H}_1 \otimes ... \otimes \mathcal{H}_N \tag{6.1}$$

where we take the infimum over decompositions of the Hilbert space into subsystem Hilbert spaces $\mathcal{H}_i$ (by the isomorphism ϕ). Note that we fix the basis used for the decomposition of the total Hilbert space $\mathcal{H}$ (as the position basis for example) and fix N. The trace over $\bar{i}$ denotes the trace taken over all the subspaces other than the i subspace. Following terminology used in the definition of Integrated Information [288] we call the Hilbert space partition which minimizes the QII the minimum information partition (MIP).

S is the quantum relative entropy:

$$S(\sigma_1 || \sigma_2) := Tr(\sigma_1 \log \sigma_1) - Tr(\sigma_1 \log \sigma_2) \tag{6.2}$$

between the state of the system and the tensor product of the states obtained by tracing out each subsystem i in the MIP.

Note that we can extend this definition to the case where the Hilbert space is decomposed into an infinite number of subspaces such that: $\mathcal{H} \cong \bigotimes_{i \in I} \mathcal{H}_i$, where the index set I is no longer the finite set $\{1, ..., N\}$.

An interesting question is whether the MIP always splits the Hilbert space into two subsystems. We expect that finding the MIP and calculating the QII of realistic physical systems will rely on the use of approximations and numerical techniques.

6.3 Calculating the Quantum Integrated Information

We will now explicitly calculate the QII of two simple tripartite systems: the GHZ [157] and W [121] states.

The density matrices for these pure states are:

$$GHZ = \frac{1}{2}(|000\rangle + |111\rangle)(\langle 000| + \langle 111|) \tag{6.3}$$

$$W = \frac{1}{3}(|001\rangle + |010\rangle + |100\rangle)(\langle 001| + \langle 010| + \langle 100|) \tag{6.4}$$

Since both of these states are symmetrical, we only need to consider two candidate splittings for the MIP, namely separating the Hilbert space into three subsystems A, B and C or into two subsystems A and BC. Calculating the relevant reduced density matrices yields:

$$GHZ_A \otimes GHZ_{BC} = \frac{1}{4}\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{6.5}$$

$$GHZ_A \otimes GHZ_B \otimes GHZ_C = \frac{\mathbb{I}}{8} \tag{6.6}$$

$$W_A \otimes W_{BC} = \frac{1}{9}\begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{6.7}$$

$$W_A \otimes W_B \otimes W_C = \frac{1}{27}\begin{pmatrix} 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{6.8}$$

Matrix diagonalization gives us:

$$\log{(GHZ)} = \log{(W)} = \mathbf{0} \tag{6.9}$$

Therefore:

$$\begin{aligned} S(GHZ||GHZ_A \otimes GHZ_{BC}) &= -Tr(GHZ \log{(GHZ_A \otimes GHZ_{BC})}) \\ &= Tr\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = 2 \end{aligned} \tag{6.10}$$

$$S(GHZ||GHZ_A \otimes GHZ_B \otimes GHZ_C) = 3 \tag{6.11}$$

$$S(GHZ||W_A \otimes W_{BC}) = 2\log{(\frac{3}{2})} \approx 1.17 \tag{6.12}$$

$$S(GHZ||W_A \otimes W_B \otimes W_C) = \frac{1}{3}(\log\frac{27}{2} + 2\log\frac{27}{4}) \approx 3.09 \tag{6.13}$$

Hence, we get that the QII of these states are: $\Phi(\text{GHZ}) = 2$ and $\Phi(\text{W}) = 2\log\frac{3}{2} \approx 1.17$.

6.4 A review of existing quantum collapse theories

Quantum mechanics admits a clash between the linear deterministic evolution of an unobserved system and the nonlinear stochastic collapse of observed systems [294,40]. This dichotomy is at the heart of the difficulty in interpreting quantum theory and leads to the impossibility of attributing definite properties to physical systems independently of measurement.

We will now review some of the main quantum collapse theories [44] which aim to provide a unified dynamical model describing both observed and unobserved physical systems.

6.4.1 Pearle's collapse equation

The seminal article investigating the possibility of using a stochastic nonlinear modification of the Schrödinger equation to explain quantum measurement is due to Pearle [230]. He postulates that a non-linear term can be added to the Schrödinger equation which, upon measurement, rapidly drives the amplitude of one of the state vectors in a superposition to one and the other amplitudes to zero.

Pearle proposes the following non-linear equation describing his collapse model in terms of the probability amplitudes:

$$i\hbar\frac{da_n}{dt} = \hbar\omega_n a_n + \sum_{m\neq 1}^{N}\langle\phi_n(t)|H_I|\phi_m(t)\rangle a_m + \lambda\hbar\frac{a_n^r}{a_n^*}\sum_{m=1}^{N}(a_m^*)^r\alpha_{nm}\exp ir\beta_{nm} \tag{6.14}$$

where $A_{nm} := \alpha_{nm}\exp ir\beta_{nm}$ are elements of a Hermitian matrix (such that $\alpha_{nm} = \alpha_{mn}$ and $\beta_{nm} = -\beta_{mn}$), λ is a real coupling constant, H_I is the usual interaction Hamiltonian and $a_n := \langle\phi_n(t)|\psi(t)\rangle$ is the interaction picture probability amplitude for the n^{th} state.

Given the non-linear collapse equation, one can derive (using a weak coupling approximation) a *diffusion equation*, describing the reduction of an ensemble of state vectors (described by a density matrix ρ):

$$\frac{\partial\rho(\vec{x},t)}{\partial t} = \lambda^2 \sum_{n<m}^{N} [(\frac{\partial}{\partial x_n} - \frac{\partial}{\partial x_m})^2 \alpha_{nm} x_n^r x_m^r] \rho(\vec{x},t) \tag{6.15}$$

Note that the rate of collapse depends on the Hermitian matrix A (through the elements α_{nm}) and the coupling constant λ. Experimental verification is expected to constrain the allowable values of the constants in equation (6.14).

Two possible shortcomings of Pearle's model are the lack of an explicit description of the preferred collapse basis on which reductions take place and a missing description of the amplification mechanism reducing superpositions when moving from the microscopic to the macroscopic level.

6.4.2 GRW Model

Both of these limitations were overcome ten years later when Ghirardi, Rimini and Weber[145] presented their GRW collapse model. In their theory, the basis on which reductions take place is chosen such that macroscopic objects have a definite position in space and there is an amplification mechanism such that objects composed of more particles undergo a higher rate of collapse.

GRW consider a system of N particles represented by a wavefunction $\psi(t)$ which evolves according to the Schrödinger equation:

$$i\hbar\frac{\partial}{\partial t} |\psi(t)\rangle = H |\psi(t)\rangle \tag{6.16}$$

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):

$$|\psi(t+dt)\rangle = \frac{1}{\sqrt{p(q_k)}} \sqrt{E^{(k)}(q_k)} |\psi(t)\rangle \tag{6.17}$$

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 $\tau := \lambda_{GRW}^{-1} \simeq 10^{16}s$, and the localization width of each particle $\sigma \simeq 10^{-7}m$. Jumps are assumed to be distributed in time similarly to a Poissonian process with frequency λ_{GRW}. 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.

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:

$$\psi = \sum_n C_n \psi_n(x) \phi_n(y_1, ..., y_R, Y) \tag{6.18}$$

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.

We can write a GRW master equation:

$$\frac{d\rho(t)}{dt} = -\frac{i}{\hbar}[H, \rho(t)] - \sum_{i=1}^{N} \mathcal{T}_i[\rho(t)] \tag{6.19}$$

where there are N non-linear $\mathcal{T}_i$ operators (one for each particle) such that:

$$\langle \mathbf{x} | \mathcal{T}_i[\rho(t)] | \mathbf{y} \rangle = \tau^{-1}[1 - \exp \frac{-(\mathbf{x} - \mathbf{y})^2}{4\sigma^2}] \langle \mathbf{x} | \rho(t) | \mathbf{y} \rangle \tag{6.20}$$

in the position representation. It is then clear that the collapse amplification mechanics depends directly on the number of particles (or the size of the system).

6.4.3 QMULP Model

Diósi[113] introduced a gravity-based version of the GRW model where unwanted macroscopic superpositions of quantum states become destroyed in very short times for massive objects due to gravitational measures for reducing macroscopic fluctuations of the mass density. This led to the introduction of the QMULP collapse model, which we shall describe now.

Quantum mechanics with universal position localization (QMULP) is an alternative collapse model which admits a more streamlined mathematical form. The dynamics can be described by using a stochastic non-linear dynamical equation:

$$d\Psi_t = [-\frac{i}{\hbar}Hdt + \sum_{i=1}^{N}\sqrt{\lambda_i}(q_i - \langle q_i\rangle)dW_t^{(i)} - \frac{1}{2}\sum_{i=1}^{N}\lambda_i(q_i - \langle q_i\rangle)^2 dt]\Psi_t \tag{6.21}$$

where H is the quantum Hamiltonian, q_i are the position operators of the N particles and $\langle q_i\rangle := \langle\Psi_t|q_i|\Psi_t\rangle$. The λ_i are collapse coefficients for each particle which can be taken as: $\lambda_i = \frac{m_i}{m_{nuc}}\lambda_{QMULP}$, where m_i is the particle mass, m_{nuc} is the nucleon mass and λ_{QMULP} is a universal collapse constant. There are N independent Wiener processes $W_t^{(i)}$, which are continuous-time stochastic processes for $t \geq 0$ with $W_0^{(i)} = 0$, such that each increment $W_s^{(i)} - W_u^{(i)}$ is Gaussian with mean 0 and variance s-u for any $0 \leq u < s$ and increments for non overlapping time intervals are independent.

In contrast with the GRW model described above, the QMULP model does not have a parameter corresponding to the localization width σ of each particle since stochastic fluctuations only take place in time. Also, the model is built for systems of distinguishable particles. Note that work has been done to extend the QMULP model to include more realistic noise[39] and dissipative effects[42].

6.4.4 Continuous Spontaneous Localization Model

The Continuous Spontaneous Localization (CSL) model[147] is the most current space collapse model. It can be defined by using the following stochastic differential equation:

$$d\Psi_t = [-\frac{i}{\hbar}Hdt + \frac{\sqrt{\gamma}}{m_{nuc}}\int d\mathbf{x}(M(\mathbf{x}) - \langle M(\mathbf{x})\rangle)dW_t(\mathbf{x}) - \frac{\gamma}{2(m_{nuc})^2}\int d\mathbf{x}(M(\mathbf{x}) - \langle M(\mathbf{x})\rangle)^2 dt]\Psi_t \tag{6.22}$$

where H is the quantum Hamiltonian, m_{nuc} is the nucleon mass, γ is a positive coupling constant and the $W_t(\mathbf{x})$ are an ensemble of independent Wiener processes (one for each point in space).

$M(\mathbf{x})$ is a mass density operator:

$$M(\mathbf{x}) = \sum_j m_j \int d\mathbf{y}(\sqrt{2\pi}\sigma)^{-3} \exp -\frac{(\mathbf{y}-\mathbf{x})^2}{2\sigma^2} a_j^\dagger(\mathbf{y})a_j(\mathbf{y}) \tag{6.23}$$

where $a_j^\dagger(\mathbf{y})$ and $a_j(\mathbf{y})$ are the creation and annihilation operators of a particle of mass m_j at position $\mathbf{y}$ and σ is the particle localization width, which is a fundamental constant of the model.

We can define the collapse rate for the model as:

$$\lambda_{CSL} := \frac{\gamma}{(4\pi\sigma^2)^{\frac{3}{2}}} \approx 2.2 \times 10^{-17} s^{-1} \tag{6.24}$$

The CSL model can be generalized by including more elaborate (non-white) noise[10] but the dynamical equations then become non-Markovian.

6.4.5 Gravity-induced collapse models

An alternative class of collapse models puts forward the idea that spontaneous collapse might be related to the curvature of spacetime produced by material bodies. Gravity would then play the key role in wave-function reduction for macroscopic objects, whilst leaving the microscopic domain unaffected. In that light, gravity may provide a fundamental underpinning for spontaneous collapse models and explicate the new parameters for rate of collapse and localization width.

The idea of reconciling quantum theory and general relativity led Karolyhazy [179] to combine the Heisenberg's uncertainty relations with gravitation and derive a quantitative limit on the 'sharpness' of the structure of space-time.

According to the Karolyhazy uncertainty relation [179], the distance s in Minkowski space-time cannot be known to a better accuracy than:

$$\Delta s = (\frac{G\hbar}{2c^3})^{\frac{1}{3}} s^{\frac{1}{3}} \tag{6.25}$$

Including this relation into the dynamical (Klein-Gordon) equation for the propagation of the quantum wavefunction gives a novel theory where pure wavefunctions evolve into mixtures and a single pure wavefunction survives only as long as it corresponds to a sufficiently small spread in the position of any massive part of the system under investigation. In this K model, which can be related to the GRW model [137], the reduction time decreases with increasing mass and there are no new free parameters.

Diósi [112] followed Karolyhazy and introduced an explicit gravity-induced collapse model described by the master equation:

$$\frac{d}{dt}\rho(t) = -\frac{i}{\hbar}[H, \rho(t)] - \frac{G}{2\hbar}\int\int \frac{d\mathbf{r}d\mathbf{r'}}{|\mathbf{r}-\mathbf{r'}|}[f(\mathbf{r})[f(\mathbf{r'}), \rho(t)]] \tag{6.26}$$

where $f(\mathbf{r})$ is the local mass density operator at the point $\mathbf{r}$. The collapse rate can then be calculated in terms of the local mass density operator so that the collapse rate free parameter is replaced by the gravitational constant G. Interpreting this collapse model by using a stochastic Schrödinger equation gives a model called QMUDL (quantum mechanics with universal density localization) [113], which is analogous to the QMULP model but with the mass density operator $f(\mathbf{r})$ playing the role of the position operators q_i. At present there is no gravity-induced collapse model corresponding to the CSL model.

Penrose [232,233] has argued in favor of gravity-induced quantum collapse by noting that time translation and the operator $\frac{\partial}{\partial t}$ are not well defined in the presence of gravitation. This leads to a fundamental uncertainty in the energy of states in a quantum superposition [232] due to the fact that there must be two different spacetimes (one for each one of the two superposed quantum states) which cannot be identified with each other because of the general covariance principle. Quantum states in a superposition then have a finite lifetime, with a collapse rate $\tau \approx \frac{\hbar}{E_\Delta}$ inversely proportional to the energy uncertainty. This provides an interpretation where, in the presence of gravity, spontaneous collapse of the wavefunction arises naturally from the laws of General Relativity.

Finally, an important difficulty for gravity-induced collapse is to specify which quantum states are to be regarded as the stable basic states which are not considered as superpositions and do not decay by spontaneous state reduction. Penrose argued that [233] these basic stationary states can be taken as stationary solutions of the Schrödinger-Newton system of partial differential equations [255], where a nonlinear modification corresponding to a Newtonian gravitational potential is added to the Schrödinger equation. Taken together with the Diósi master equation described above, this gives the Diósi-Penrose collapse model.

6.4.6 Adler trace dynamics

We shall conclude our presentation of collapse models by mentioning that Adler [7] has proposed that quantum field theory emerges from a matrix theory where particles (bosons and fermions) are represented by Grassmannian non-commuting matrices and the Lagrangian is constructed by taking the trace of a function of these matrices. Quantum theory is then treated as a thermodynamic approximation to a general statistical mechanics of the matrix models and Brownian motion around the thermodynamic approximation naturally yields non-linear stochastic modifications of the Schrödinger equation.

This trace dynamics method gives an underlying framework for spontaneous collapse theories, where fluctuations about the equilibrium lead away from quantum theory. Adler's work, however, does not provide any understanding of the arising of fundamental parameters such as collapse rate or localization width.

6.5 Integrated Information and state-vector reduction

As we have seen, collapse theories are alternatives to standard quantum mechanics, which aim to resolve its issues by presenting a universal non deterministic, nonlinear evolution law such that microprocesses and macroprocesses are governed by a single dynamics [230,146,147,113,232,7].

We expect a universal dynamical equation to satisfy the following constraints, which

strongly restrict the allowed form of the non-linear modification to Schrödinger's equation:
(i) It must be almost identical to Schrödinger's equation in the quantum regime but should break the superposition principle at the macroscopic level.
(ii) It must be stochastic and should explain why measurement situations yield results distributed according to the Born rule.
(iii) It must not allow for superluminal signaling[148] in order to preserve relativistic causal structure.

Previous work on collapse models has shown that a universal equation of the form:

$$\frac{d}{dt}\rho(t) = -\frac{i}{\hbar}[H, \rho(t)] - \mathcal{I}[\rho(t)] \tag{6.27}$$

where $\mathcal{I}$ is a non-linear operator representing the effect of the spontaneous collapse, can satisfy all three constraints.

Standard space collapse models (such as GRW or CSL) are astutely set up such that each particle undergoes random collapse leading to larger systems collapsing faster than small systems. In the dynamical equations, the rate of collapse is directly dependent on the number of particles or size of the physical system under study.

In our model, however, particles no longer undergo random collapse at random times but instead we consider that the spontaneous collapse follows from a type of group behavior. We expect that a physical system exhibiting a certain amount of informational complexity has an increased chance of spontaneous collapse. In that sense, we expect collapse to be less random than in other space collapse models: physical systems which have a high QII should naturally collapse faster.

We believe that a physical system's capacity to act as an observer should not depend on its size but on other physical properties instead. Indeed, localization follows from the process of observation which occurs in a measurement. This observation process taking place should require the observer in question to exhibit consciousness. This leads us to postulate that the main physical property determining whether or not a system can act as an observer is directly related to a key aspect of its informational complexity, namely its capacity to process information in an integrated way.

The idea that a physical description of consciousness could be at the heart of resolving fundamental issues in quantum theory is not new [298,279]. In the present chapter we make no claims of presenting such a description, but assume that quantum Integrated Information determines how much a system acts like an observer and exhibits spontaneous collapse.

We introduce a novel collapse model where the rate of collapse does not depend on a system's size but on how much QII it exhibits. The general evolution equation we propose is of the form:

$$\frac{d}{dt}\rho(t) = -\frac{i}{\hbar}[H,\rho(t)] + \sum_{n,m=1}^{N^2-1} h_{n,m}(\Phi(\rho(t)))(L_n\rho(t)L_m^\dagger - \frac{1}{2}(\rho(t)L_m^\dagger L_n + L_m^\dagger L_n\rho(t))) \quad (6.28)$$

where the Hermitian matrix elements $h_{n,m}$ are continuous functions of the QII of ρ (which are all zero when $\Phi(\rho) = 0$) and $\{L_k\}$ is a basis of operators on the N dimensional system Hilbert space, which determines the basis in which the state collapses.

Note that this is a highly non-linear Markovian collapse equation [43]. It has been argued [93,41] that macro-objectification must take place in space and time and that position must therefore play the preferred role in collapse theories. Since space collapse models appear to be the only ones which explain the classical behavior of macroscopic objects, we must choose the $\{L_k\}$ basis such that the wavefunction localizes in the position basis.

Hence, our model's objective description of how macroscopic reality arises is rather similar to the one resulting from the standard space collapse theories [146,147], but where the mechanism causing the collapse onto the position basis depends on the QII. An underlying equation for wave function dynamics, whose general form would resemble that of standard space collapse models [43] but with parameters related to QII, could also provide an alternative description of our model.

We can produce a large class of Integrated Information collapse models by replacing this evolution equation by equation (6.27), with a more general non-linear operator $\mathcal{I}$ describing how the collapse rate depends on the system's QII.

In the future, we expect a slightly modified version of the QII dynamical reduction equation to be compatible with relativity. This universal dynamics may emerge from a fundamental underlying theory in the spirit of trace dynamics [7,9] or of quantum theory without spacetime[210].

It could also turn out that the level of QII of a physical system is not the optimal measure of its capacity to encompass various distinguishable states and process information in a cohesive, integrated manner. Therefore, QII may have to be replaced by a more astute measure or one which is more convenient to calculate. We stress that the key idea of this article is that informational complexity, and more precisely the capacity to process information in an integrated manner, should replace size as the property of a physical system which determines its rate of collapse. Further details will require more fine tuning and input from experiments.

6.6 Experimental tests of Integrated Information-induced collapse

The Integrated Information collapse model we have presented here is an experimentally verifiable theory which is expected to yield some physical predictions which are in conflict with quantum mechanics. We will briefly discuss potential experiments which could serve to validate, reject or at least refine the new theory.

The predictions of the new theory almost coincide with those of standard quantum mechanics at the microscopic level. Most current collapse models become significantly different from quantum theory when the size of the system under study increases. This leads to numerous experimental challenges due to the fact that environmental influences become more and more difficult to eliminate for larger systems.

Typical experiments testing collapse models aim to set bounds on model parameters by studying the collapse of sizable physical systems in a large superposition[135,226,18]. The aim of most superposition experiments is to observe spontaneous collapse of the wavefunction at a mesoscopic scale, after reducing the interaction with the environment. Tests of superposition include diffraction experiments with large molecules[19,143,124], optomechanical systems[214], microsphere interferometers[256] and indirect tests using cosmological data[8,102].

Testing Integrated Information collapse is different from previous work on verifying the validity of collapse models. It is no longer sufficient to study large systems in order to increase the predicted rate of collapse. Indeed, we expect novel behavior in conflict with quantum theory to arise in situations where physical systems with a high level of QII exhibit non-linear collapse and cause a breakdown of the quantum principle of linear superposition.

Therefore, the first step in verifying QII collapse consists of calculating the quantum Integrated Information of various interesting physical systems. This may require some numerical approximations and clever optimization in order to determine the minimum information partition (MIP) for each system.

The next step would then be to compare the collapse rate of various physical systems with very different QII. We expect these experiments testing quantum superposition to be similar in nature to current collapse model tests. They would require an extremely precise control of the environment since the effects of decoherence need to be accounted for to a high precision. Note that one would expect conscious beings to clearly exhibit high levels of QII and therefore physical systems including such beings would undergo spontaneous collapse. It may be the case, however, that certain complex inanimate objects may have a high QII and therefore also behave as observers, in the sense that their presence within a larger physical system leads to collapse.

In some respects, the experimental tests of QII collapse models may be simpler to implement than those for standard spontaneous collapse since the systems under examination might not have to be as large. Indeed, several relatively small mesoscopic systems of similar size may exhibit very different levels of QII and have observably different spontaneous collapse rates.

These experiments should help us refine the collapse model dynamics and determine the $h_{n,m}(\Phi)$ matrix elements in equation (6.28). They will also lead to a better understanding of whether QII is indeed the best measure of a physical system's capacity to spontaneously collapse.

6.7 Conclusion

We have presented a novel theory which is in conflict with quantum mechanics. Even if it turns out that QII spontaneous collapse does not agree with future experiments, we feel that the theoretical implications of the new collapse theory are of interest for their own sake and may shed some light on various features of quantum theory.

First of all, it may be interesting to study computational properties of the new collapse model. How would the spontaneous collapse of systems with high QII affect the possibility of performing large ‘quantum’ computations. Can one define a modified version of many-worlds theory which can be related to the QII collapse model?

Moreover, we believe that the basis on which wavefunction localization takes place should not always be position. The relationship between another physical definition of Integrated Information and the so-called *quantum factorization problem* has been addressed in[282].

In general, we expect that the collapse basis for each system may depend on properties of a quantum version of qualia space[30], corresponding to the quality of consciousness of the system in question. In this sense, dynamics would not just be governed by the QII of a physical system but also by the set of all the informational relationships that causally link its elements.

In the model we are currently proposing, the collapse mechanism is universal and not related to specific systems since position plays a fundamental role, similarly to the current spontaneous collapse models. Further work, however, could redefine equation (6.28) and the operator basis $\{L_k\}$ such that the collapse basis is different for each physical system in a way which explains the apparent fundamental role of the position basis. Space-time would then emerge from the fact that we cannot extract ourselves from the physical systems we examine.

This may lead to alternative versions of quantum field theory, where space-time does not play a fundamental role. We expect new particles – *complexetrons*– to arise due to the spontaneous collapse term in equation (6.27).

We look forward to revealing the physical world described by Integrated Information-induced collapse.

Chapter 7

Conclusion

"There must be some way out of here,

Said the joker to the thief,

There's too much confusion,

I can't get no relief."

Bob Dylan

Our brief foray through the foundations of alternative quantum theories was only a succinct introduction. The use of mathematical abstraction and symmetry as a tool for classification and clarification in the foundations of physics is a promising avenue of inquiry. The thesis outlined a research program whose formal development is still in the initial stages.

In the preceding chapters, we used several different approaches to explore the world of alternative theories that share features with quantum mechanics. Chapter 4 focused on the use of a quantum circuit calculus to analyze the logical features of stabilizer quantum theory, a sub-theory of quantum mechanics. Chapter 5 emphasized the role of symmetry and of the phase group in the study of mutually unbiased qudit theories. We presented five different levels of analysis for physical theories: using an explicit operational representation, a categorical representation, a group-theoretic representation, a finite field representation and finally a generalized ontological model representation. Chapter 6 introduced another class of quantum-like theories, called collapse models, and defined a new type of quantum collapse model.

We took a step towards reaching the goal of presenting and examining a diverse range of physical theories by using an elegant and concise abstract framework. Could it be, however, that the aim of reducing confusion through the use of synthetic mathematical analysis might be undesirable and destined to ineluctable failure from the start? Perhaps we should simply accept that:

"Science is essentially an anarchic enterprise: theoretical anarchism is more humanitarian and more likely to encourage progress than its law-and-order alternatives."

Bibliography

[1] N. H. Abel. Mémoire sur les équations algébriques, où l'on démontre l'impossibilité de la résolution de l'équation générale du cinquième degré. 1824.

[2] S. Abramsky and A. Brandenburger. The sheaf-theoretic structure of non-locality and contextuality. *New Journal of Physics*, 13(11):113036, Nov. 2011. doi: 10.1088/1367-2630/13/11/113036.

[3] S. Abramsky and B. Coecke. A categorical semantics of quantum protocols. *University Computing*, 415:21, 2004. doi: 10.1109/LICS.2004.1.

[4] S. Abramsky and L. Hardy. Logical Bell inequalities. *Phys. Rev. A*, 85:062114, Jun 2012. doi: 10.1103/PhysRevA.85.062114.

[5] S. Abramsky and N. Tzevelekos. Introduction to categories and categorical logic. *University Computing*, 813:96, 2011.

[6] S. Abramsky, S. Mansfield, and R. Soares Barbosa. The Cohomology of Non-Locality and Contextuality. *Electronic Proceedings in Theoretical Computer Science*, 95:1–14, 2011. doi: 10.4204/EPTCS.95.1.

[7] S. L. Adler. *Quantum Theory as an Emergent Phenomenon: The Statistical Mechanics of Matrix Models as the Precursor of Quantum Field Theory*. Cambridge University Press, Oct. 2004. ISBN 0521831946.

[8] S. L. Adler. Lower and upper bounds on CSL parameters from latent image formation and IGM heating. *Journal of Physics A Mathematical General*, 40:2935–2957, Mar. 2007. doi: 10.1088/1751-8113/40/12/S03.

[9] S. L. Adler. Incorporating gravity into trace dynamics: the induced gravitational action. *Classical and Quantum Gravity*, 30(19):195015, Oct. 2013. doi: 10.1088/0264-9381/30/19/195015.

[10] S. L. Adler and A. Bassi. Collapse models with non-white noises. *Journal of Physics A Mathematical General*, 40:15083–15098, Dec. 2007. doi: 10.1088/1751-8113/40/50/012.

[11] D. Aharonov, W. van Dam, J. Kempe, Z. Landau, S. Lloyd, and O. Regev. Adiabatic Quantum Computation is Equivalent to Standard Quantum Computation. *SIAM Journal of Computing*, 37:166–194, May 2007.

[12] Y. Aharonov and L. Vaidman. The two-state vector formalism: an updated review. In *Time in quantum mechanics*, pages 399–447. Springer, 2008.

[13] S. W. Al-Safi. *Quantum theory from the perspective of general probabilistic theories.* PhD thesis, University of Cambridge, 2015.

[14] S. Anders and H. J. Briegel. Fast simulation of stabilizer circuits using a graph-state representation. *Phys. Rev. A*, 73:022334, Feb 2006. doi: 10.1103/PhysRevA.73.022334.

[15] D. Archard and P. S. Churchland. Neurophilosophy: Toward a unified science of the mind/brain. *Radical Philosophy*, 49:41, 1988.

[16] G. B. Arfken and H. J. Weber. *Mathematical Methods for Physicists*, chapter 9.3. Academic Press, 2005. ISBN 9780123846556.

[17] Aristotle. *Prior Analytics and Posterior Analytics.* Digireads, 2010. ISBN 9781420935646.

[18] M. Arndt and K. Hornberger. Testing the limits of quantum mechanical superpositions. *Nat Phys*, 10(4):271–277, Apr. 2014. doi: 10.1038/nphys2863.

[19] M. Arndt, O. Nairz, J. Vos-Andreae, C. Keller, G. van der Zouw, and A. Zeilinger. Waveparticle duality of C60 molecules. *Nature*, 401(6754):680–682, Oct. 1999. doi: 10.1038/44348.

[20] R. Ash and C. Doléans-Dade. *Probability and Measure Theory.* Harcourt/Academic Press, 2000. ISBN 9780120652020.

[21] A. Aspect, P. Grangier, and G. Roger. Experimental Tests of Realistic Local Theories via Bell's Theorem. *Phys. Rev. Lett.*, 47:460–463, Aug 1981. doi: 10.1103/PhysRevLett.47.460.

[22] S. Awodey. *Category Theory.* Oxford Logic Guides. Ebsco Publishing, 2006. ISBN 9780191513824.

[23] B. J. Baars. *A Cognitive Theory of Consciousness.* Cambridge University Press, 1988. ISBN 0521427436.

[24] M. Backens. The ZX-calculus is complete for stabilizer quantum mechanics. *New Journal of Physics*, 16(9):093021, 2014. doi: 10.1088/1367-2630/16/9/093021.

[25] M. Backens. The ZX-calculus is complete for the single-qubit Clifford+T group. In B. Coecke, I. Hasuo, and P. Panangaden, editors, Proceedings 11th workshop on *Quantum Physics and Logic*, volume 172 of *Electronic Proceedings in Theoretical Computer Science*, pages 293–303, 2014. doi: 10.4204/EPTCS.172.21.

[26] M. Backens and A. N. Duman. A Complete Graphical Calculus for Spekkens' Toy Bit Theory. *Foundations of Physics*, 46:70–103, Jan. 2016. doi: 10.1007/s10701-015-9957-7.

[27] M. Backens, S. Perdrix, and Q. Wang. A Simplified Stabilizer ZX-calculus. *ArXiv e-prints*, Feb 2016.

[28] J. C. Baez. An introduction to n-categories. In *Category theory and computer science*, pages 1–33. Springer, 1997.

[29] D. Balduzzi and G. Tononi. Integrated information in discrete dynamical systems: motivation and theoretical framework, Jun 2008.

[30] D. Balduzzi and G. Tononi. Qualia: The geometry of integrated information. *PLoS Computational Biology*, 5(8), Aug 2009. doi: 10.1371/journal.pcbi.1000462.

[31] S. Bandyopadhyay, P. O. Boykin, V. Roychowdhury, and F. Vatan. A new proof for the existence of mutually unbiased bases. *eprint arXiv:quant-ph/0103162*, Mar 2001.

[32] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, and H. Weinfurter. Elementary gates for quantum computation. *Phys. Rev. A*, 52:3457–3467, Nov 1995. doi: 10.1103/PhysRevA.52.3457.

[33] H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa, and B. Schumacher. Noncommuting Mixed States Cannot Be Broadcast. *Physical Review Letters*, 76:2818–2821, Apr. 1996. doi: 10.1103/PhysRevLett.76.2818.

[34] H. Barnum, J. Barrett, L. O. Clark, M. Leifer, R. Spekkens, N. Stepanik, A. Wilce, and R. Wilke. Entropy and information causality in general probabilistic theories. *New Journal of Physics*, 12(3):033024, 2010. doi: 10.1088/1367-2630/12/3/033024.

[35] J. Barrett. Information processing in generalized probabilistic theories. *Phys. Rev. A*, 75:032304, Mar 2007. doi: 10.1103/PhysRevA.75.032304.

[36] J. Barrett, E. G. Cavalcanti, R. Lal, and O. J. E. Maroney. No ψ-Epistemic Model Can Fully Explain the Indistinguishability of Quantum States. *Physical Review Letters*, 112(25):250403, Jun 2014. doi: 10.1103/PhysRevLett.112.250403.

[37] S. D. Bartlett, T. Rudolph, and R. W. Spekkens. Reconstruction of Gaussian quantum mechanics from Liouville mechanics with an epistemic restriction. *Phys. Rev. A*, 86(1):012103, July 2012. doi: 10.1103/PhysRevA.86.012103.

[38] D. S. Bassett and M. S. Gazzaniga. Understanding complexity in the human brain. *Trends in Cognitive Sciences*, 15(5):200–209, May 2011. doi: 10.1016/j.tics.2011.03.006.

[39] A. Bassi and L. Ferialdi. Non-Markovian Quantum Trajectories: An Exact Result. *Phys. Rev. Lett.*, 103:050403, Jul 2009. doi: 10.1103/PhysRevLett.103.050403.

[40] A. Bassi and G. Ghirardi. A general argument against the universal validity of the superposition principle. *Physics Letters A*, 275:373–381, Oct. 2000. doi: 10.1016/S0375-9601(00)00612-5.

[41] A. Bassi and G. Ghirardi. Dynamical reduction models. *Physics Reports*, 379:257–426, jun 2003. doi: 10.1016/S0370-1573(03)00103-0.

[42] A. Bassi, E. Ippoliti, and B. Vacchini. On the energy increase in space-collapse models. *Journal of Physics A Mathematical General*, 38:8017–8038, Sept. 2005. doi: 10.1088/0305-4470/38/37/007.

[43] A. Bassi, D. Dürr, and G. Hinrichs. Uniqueness of the Equation for Quantum State Vector Collapse. *Physical Review Letters*, 111(21):210401, Nov. 2013. doi: 10.1103/PhysRevLett.111.210401.

[44] A. Bassi, K. Lochan, S. Satin, T. P. Singh, and H. Ulbricht. Models of wave-function collapse, underlying theories, and experimental tests. *Reviews of Modern Physics*, 85:471–527, Apr. 2013. doi: 10.1103/RevModPhys.85.471.

[45] J. L. Bell. *Logical Reflections on the Kochen-Specker Theorem*, pages 227–235. Springer Netherlands, 1996. ISBN 978-94-015-8656-6. doi: 10.1007/978-94-015-8656-6_17.

[46] J. S. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1(3):195–200, 1964.

[47] J. S. Bell. On the problem of hidden variables in quantum mechanics. *Rev. Mod. Phys.*, 38:447–452, Jul 1966. doi: 10.1103/RevModPhys.38.447.

[48] J. S. Bell. *Speakable and unspeakable in quantum mechanics.* Collected papers on quantum philosophy. Cambridge Univ. Press, Cambridge, 2004. doi: 10.1017/CBO9780511815676.

[49] E. G. Beltrametti and S. Bugajski. A classical extension of quantum mechanics. *J. Phys. A: Math. Gen.*, 28:3329–3343, 1995. doi: 10.1088/0305-4470/28/12/007.

[50] C. Bennett, G. Brassard, C. Crepeau, R. Jozsa, A. Peres, and W. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Physical Review Letters*, 70(13):1895–1899, 1993. doi: 10.1103/PhysRevLett.70.1895.

[51] C. H. Bennett. Quantum cryptography using any two nonorthogonal states. *Phys. Rev. Lett.*, 68:3121–3124, May 1992. doi: 10.1103/PhysRevLett.68.3121.

[52] C. H. Bennett and S. J. Wiesner. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Phys. Rev. Lett.*, 69:2881–2884, Nov 1992. doi: 10.1103/PhysRevLett.69.2881.

[53] R. Bing, S. Singh, S. Armentrout, and R. Daverman. *The collected papers of R.H. Bing.* American Mathematical Society, 1988. ISBN 9780821801178.

[54] G. Birkhoff and J. von Neumann. The logic of quantum mechanics. *Annals of Mathematics*, 37:823843, 1936.

[55] F. Bloch. Nuclear induction. *Phys. Rev.*, 70:460–474, Oct 1946. doi: 10.1103/PhysRev.70.460.

[56] N. Block. On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, 18:227–287, 1995. doi: 10.1017/S0140525X00038188.

[57] N. Block, O. Flanagan, and G. Güzeldere. *The Nature of Consciousness: Philosophical Debates*. Bradford Books, 1997. ISBN 9780262522106.

[58] D. Bohm. A suggested interpretation of the quantum theory in terms of "hidden" variables. i. *Phys. Rev.*, 85:166–179, Jan 1952. doi: 10.1103/PhysRev.85.166.

[59] D. Bohm. A suggested interpretation of the quantum theory in terms of "hidden" variables. ii. *Phys. Rev.*, 85:180–193, Jan 1952. doi: 10.1103/PhysRev.85.180.

[60] D. Bohm and B. Hiley. *The undivided universe: an ontological interpretation of quantum theory*. Taylor & Francis, 1995. ISBN 9780415121859.

[61] N. Bohr. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 48:696–702, Oct 1935. doi: 10.1103/PhysRev.48.696.

[62] F. Bonchi, F. Gadducci, A. Kissinger, P. Sobociński, and F. Zanasi. Rewriting modulo symmetric monoidal structure. In *Thirty-first annual ACM/IEEE symposium on Logic and Computer Science (LiCS '16)*, 2016.

[63] N. Bourbaki, H. Eggleston, and S. Madan. *Topological vector spaces*. Éléments de mathématique. Springer-Verlag, 1987. ISBN 9783540136279.

[64] F. Brentano. *Psychology from An Empirical Standpoint*. Routledge Classics. Taylor & Francis, 1874. ISBN 0415106613.

[65] H. J. Briegel, D. E. Browne, W. Dür, R. Raussendorf, and M. Van den Nest. Measurement-based quantum computation. *Nature Physics*, 5.

[66] S. Brierley and S. Weigert. Constructing mutually unbiased bases in dimension six. *Phys. Rev. A*, 79:052316, May 2009. doi: 10.1103/PhysRevA.79.052316.

[67] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner. Bell nonlocality. *Reviews of Modern Physics*, 86:419–478, Apr. 2014. doi: 10.1103/RevModPhys.86.419.

[68] J.-L. Brylinski and R. Brylinski. Universal quantum gates. *eprint arXiv:quant-ph/0108062*, Aug 2001.

[69] A. Cabello, S. Severini, and A. Winter. (Non-)Contextuality of Physical Theories as an Axiom. *ArXiv e-prints*, Oct 2010.

[70] R. Carter. *Simple Groups of Lie Type*. Wiley Classics Library. Wiley, 1989. ISBN 9780471506836.

[71] C. M. Caves, C. A. Fuchs, and R. Schack. Quantum probabilities as bayesian probabilities. *Phys. Rev. A*, 65:022305, Jan 2002. doi: 10.1103/PhysRevA.65.022305.

[72] C. M. Caves, C. A. Fuchs, and R. Schack. Subjective probability and quantum certainty. *Studies in History and Philosophy of Science Part B*, 38(2):255–274, 2007. doi: 10.1016/j.shpsb.2006.10.007.

[73] J. Chabert and E. Barbin. *A History of Algorithms: From the Pebble to the Microchip.* chabert: A History of Algorithms. Springer Berlin Heidelberg, 1999. ISBN 9783540633693.

[74] D. Chalmers. *The Conscious Mind: In Search of a Fundamental Theory.* Oxford paperbacks. OUP USA, 1997. ISBN 9780195117899.

[75] D. J. Chalmers. Facing up to the problem of consciousness. *Journal of Consciousness Studies*, 2:200–219, 1995.

[76] E. Cheng and A. Lauda. Higher-dimensional categories: an illustrated guide book. 2004.

[77] G. Chiribella, G. M. D'Ariano, and P. Perinotti. Probabilistic theories with purification. *Phys. Rev. A*, 81:062348, Jun 2010. doi: 10.1103/PhysRevA.81.062348.

[78] B. S. Cirel'son. Quantum generalizations of Bell's inequality. *Letters in Mathematical Physics*, 4:93–100, 1980. doi: 10.1007/BF00417500.

[79] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, 23:880–884, Oct 1969. doi: 10.1103/PhysRevLett.23.880.

[80] R. Clifton, J. Bub, and H. Halvorson. Characterizing quantum theory in terms of information-theoretic constraints. *Foundations of Physics*, 33:1561–1591, 2003. doi: 10.1023/A:1026056716397.

[81] B. Coecke. Introducing categories to the practicing physicist. *What is category theory*, 1:45–76, 2006.

[82] B. Coecke and R. Duncan. Interacting quantum observables. 2008. doi: 10.1007/978-3-540-70583-3/25. Extended version: arXiv:quant-ph/09064725.

[83] B. Coecke and R. Duncan. Interacting quantum observables: categorical algebra and diagrammatics. *New Journal of Physics*, 13(4):043016, 2011. doi: 10.1088/1367-2630/13/4/043016.

[84] B. Coecke and C. Heunen. Pictures of complete positivity in arbitrary dimension. *Electronic Proceedings in Theoretical Computer Science*, 95:27–35, 2011. doi: 10.4204/EPTCS.95.4.

[85] B. Coecke and E. O. Paquette. POVMs and Naimarks theorem without sums. *Electronic Notes in Theoretical Computer Science*, 210:15–31, 2006. doi: 10.1016/j.entcs.2008.04.015.

[86] B. Coecke and D. Pavlovic. Quantum measurements without sums. In G. Chen, L. Kauffman, and S. Lamonaco, editors, *Mathematics of Quantum Computing and Technology*, pages 567–604. Taylor and Francis, 2007. doi: 10.1201/9781584889007.ch16.

[87] B. Coecke and S. Perdrix. Environment and classical channels in categorical quantum mechanics. *ArXiv e-prints*, Apr 2010. doi: 10.2168/LMCS-8(4:14)2012.

[88] B. Coecke, D. Pavlovic, and J. Vicary. A new description of orthogonal bases. *Electronic Notes in Theoretical Computer Science*, 2008. doi: 10.1017/S0960129512000047.

[89] B. Coecke, S. Perdrix, and E. O. Paquette. Bases in Diagrammatic Quantum Protocols. *Electronic Notes in Theoretical Computer Science*, 218(0):131 – 152, 2008. doi: 10.1016/j.entcs.2008.10.009.

[90] B. Coecke, E. O. Paquette, and D. Pavlovic. Classical and quantum structuralism. *ArXiv e-prints*, Apr 2009.

[91] B. Coecke, B. Edwards, and R. W. Spekkens. Phase groups and the origin of non-locality for qubits. *Electronic Notes in Theoretical Computer Science*, 270(2):15–36, 2011. doi: 10.1016/j.entcs.2011.01.021.

[92] B. Coecke, R. Duncan, A. Kissinger, and Q. Wang. Strong Complementarity and Non-locality in Categorical Quantum Mechanics. *Logic in Computer Science (LICS), 27th Annual IEEE Symposium*, pages 245–254, Jun 2012. doi: 10.1109/LICS.2012.35.

[93] R. Cohen and J. Stachel. *Potentiality, Entanglement and Passion-at-a-Distance: Quantum Mechanical Studies for Abner Shimony, Volume Two*. Boston Studies in the Philosophy and History of Science. Springer Netherlands, 1997. ISBN 9780792344537.

[94] R. Colbeck and R. Renner. No extension of quantum theory can have improved predictive power. *Nature Communications*, 2:411, Aug 2011. doi: 10.1038/ncomms1416.

[95] R. Colbeck and R. Renner. Is a system's wave function in one-to-one correspondence with its elements of reality? *Phys. Rev. Lett.*, 108:150402, Apr 2012. doi: 10.1103/PhysRevLett.108.150402.

[96] K. Conrad. Galois groups of cubics and quartics (not char. 2), 2012.

[97] K. Conrad. Galois theory at work: concrete examples. 2013.

[98] J. Conway. *Atlas of Finite Groups: Maximal Subgroups and Ordinary Characters for Simple Groups*. Clarendon Press, 1985. ISBN 9780198531999.

[99] J. Conway, H. Burgiel, and C. Goodman-Strauss. *The Symmetries of Things*. Ak Peters Series. Taylor & Francis, 2008. ISBN 9781568812205.

[100] J. G. Cramer. The transactional interpretation of quantum mechanics. *Rev. Mod. Phys.*, 58:647–687, Jul 1986. doi: 10.1103/RevModPhys.58.647.

[101] F. Crick and C. Koch. Toward a neurobiological theory of consciousness. *Seminars in the Neurosciences*, 2:263–275, 1990.

[102] S. Das, K. Lochan, and A. Bassi. Bounds on Spontaneous Collapse model of Quantum Mechanics from formation of CMBR and Standard Cosmology. *ArXiv e-prints*, Feb. 2013.

[103] C. M. Dawson and M. A. Nielsen. The Solovay-Kitaev algorithm. *eprint arXiv:quant-ph/0505030*, May 2005.

[104] D. Dennett. Quining Qualia. In A. Marcel and E. Bisiach, editors, *Consciousness in Modern Science*. Oxford University Press, 1988.

[105] J. Derrida. *Writing and Difference*. University of Chicago Press, 1978. ISBN 9780226143293.

[106] R. Descartes:. *Principles of Philosophy*. Synthese Historical Library. 1640. ISBN 9789027717542.

[107] R. Descartes. *A Discourse on Method - (1637)*. Read Books, 2008. ISBN 9781443733748.

[108] D. Deutsch. Quantum theory of probability and decisions. *Royal Society of London Proceedings Series A*, 455:3129, Aug. 1999. doi: 10.1098/rspa.1999.0443.

[109] D. Deutsch. *The Fabric of Reality*. Penguin Books Limited, 2011. ISBN 9780141969619.

[110] D. Deutsch, A. Barenco, and A. Ekert. Universality in quantum computation. *Computer Bulletin*, 449:669–677, 1995. doi: 10.1098/rspa.1995.0065.

[111] B. DeWitt and N. Graham. *The Many-Worlds Interpretation of Quantum Mechanics*. Princeton Series in Physics. Books on Demand, 1973. ISBN 9780783719429.

[112] L. Diósi. A universal master equation for the gravitational violation of quantum mechanics. *Physics Letters A*, 120(8):377 – 381, 1987. doi: 10.1016/0375-9601(87)90681-5.

[113] L. Diósi. Models for universal reduction of macroscopic quantum fluctuations. *Physical Review A*, 40(3):1165–1174, Aug. 1989. doi: 10.1103/physreva.40.1165.

[114] F. Dowker and A. Kent. Properties of consistent histories. *Phys. Rev. Lett.*, 75:3038–3041, Oct 1995. doi: 10.1103/PhysRevLett.75.3038.

[115] F. Dowker, J. Henson, and R. D. Sorkin. Quantum Gravity Phenomenology, Lorentz Invariance and Discreteness. *Modern Physics Letters A*, 19:1829–1840, 2004. doi: 10.1142/S0217732304015026.

[116] P. Duhem. *The Aim and Structure of Physical Theory*. Atheneum paperbacks. Princeton University Press, 1914. ISBN 9780691025247.

[117] R. Duncan and M. Lucas. Verifying the Steane code with Quantomatic. *ArXiv e-prints*, Jun 2013.

[118] R. Duncan and S. Perdrix. Graphs States and the necessity of Euler Decomposition. *ArXiv e-prints*, Feb 2009.

[119] R. Duncan and S. Perdrix. Rewriting measurement-based quantum computations with generalised flow. *Proceedings of the 37th international colloquium conference on Automata, languages and programming*, pages 285–296, 2010. doi: 10.1007/978-3-642-14162-1/24.

[120] R. Duncan and S. Perdrix. Pivoting makes the ZX-calculus complete for real stabilizers. *ArXiv e-prints*, Jul 2013.

[121] W. Dür, G. Vidal, and J. I. Cirac. Three qubits can be entangled in two inequivalent ways. *Physical Review A*, 62(6):062314, Dec. 2000. doi: 10.1103/PhysRevA.62.062314.

[122] J. R. Durbin. *Modern algebra: an introduction*. Wiley, 2009. ISBN 0470384433.

[123] B. Edwards. Phase groups and local hidden variables. Technical Report RR-10-15, September 2010.

[124] S. Eibenberger, S. Gerlich, M. Arndt, M. Mayor, and J. Tuxen. Matter-wave interference of particles selected from a molecular library with masses exceeding 10 000 amu. *Phys. Chem. Chem. Phys.*, 15(35):14696–14700, 2013. doi: 10.1039/c3cp51500a.

[125] S. Eilenberg and S. MacLane. General theory of natural equivalences. *Transactions of the American Mathematical Society*, 58(2):231–294, 1945. doi: 10.2307/1990284.

[126] A. Einstein. *Die Grundlage der allgemeinen Relativitätstheorie*. Annalen der Physik und Chemie. Barth, 1916.

[127] A. Einstein. Physics and reality. *Journal of the Franklin Institute*, 221(3):349 – 382, 1936. doi: 10.1016/S0016-0032(36)91047-5.

[128] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777–780, May 1935. doi: 10.1103/PhysRev.47.777.

[129] M. B. Elliott, B. Eastin, and C. M. Caves. Graphical description of the action of Clifford operators on stabilizer states. *Phys. Rev. A*, 77:042307, Apr 2008. doi: 10.1103/PhysRevA.77.042307.

[130] C. Emary, N. Lambert, and F. Nori. Leggett-Garg inequalities. *Reports on Progress in Physics*, 77(1):016001, Jan. 2014. doi: 10.1088/0034-4885/77/1/016001.

[131] M. Ettinger, P. Høyer, and E. Knill. The quantum query complexity of the hidden subgroup problem is polynomial. *Information Processing Letters*, 91(1):43–48, 2004.

[132] H. Everett. "Relative State" Formulation of Quantum Mechanics. *Rev. Mod. Phys.*, 29:454–462, Jul 1957. doi: 10.1103/RevModPhys.29.454.

[133] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda. A Quantum Adiabatic Evolution Algorithm Applied to Random Instances of an NP-Complete Problem. *Science*, 292(5516):472–475, 2001. doi: 10.1126/science.1057726.

[134] E. Fedorov. Simmetrija na ploskosti [symmetry in the plane]. *Zapiski Imperatorskogo Sant-Petersburgskogo Mineralogicheskogo Obshchestva [Proceedings of the Imperial St. Petersburg Mineralogical Society]*, 28(2):245–291, 1891.

[135] W. Feldmann and R. Tumulka. Parameter diagrams of the GRW and CSL theories of wavefunction collapse. *Journal of Physics A Mathematical General*, 45(6):065304, Feb. 2012. doi: 10.1088/1751-8113/45/6/065304.

[136] P. Feyerabend. *Against Method*. Verso, 1975. ISBN 9780860916468.

[137] A. Frenkel. Spontaneous localizations of the wave function and classical behavior. *Foundations of Physics*, 20(2):159–188, 1990. doi: 10.1007/BF00731645.

[138] C. A. Fuchs and A. Peres. Quantum-state disturbance versus information gain: Uncertainty relations for quantum information. *Phys. Rev. A*, 53:2038–2045, Apr 1996. doi: 10.1103/PhysRevA.53.2038.

[139] C. A. Fuchs and R. Schack. Quantum-bayesian coherence. *Reviews of Modern Physics*, 85(4):1693, 2013. doi: 10.1103/RevModPhys.85.1693.

[140] A. Garg and N. D. Mermin. Detector inefficiencies in the einstein-podolsky-rosen experiment. *Phys. Rev. D*, 35:3831–3835, Jun 1987. doi: 10.1103/PhysRevD.35.3831.

[141] A. J. P. Garner, O. C. O. Dahlsten, Y. Nakata, M. Murao, and V. Vedral. A framework for phase and interference in generalized probabilistic theories. *New Journal of Physics*, 15(9):093044, 2013. doi: 10.1088/1367-2630/15/9/093044.

[142] I. Gelfand and M. Neumark. On the imbedding of normed rings into the ring of operators in Hilbert space. *Rec. Math. [Mat. Sbornik] N.S.*, 12(54):197–217, 1943.

[143] S. Gerlich, S. Eibenberger, M. Tomandl, S. Nimmrichter, K. Hornberger, P. J. Fagan, J. T. Ãuxen, M. Mayor, and M. Arndt. Quantum interference of large organic molecules. *Nature Communications*, 2:263, Apr. 2011. doi: 10.1038/ncomms1263.

[144] V. Gheorghiu. Standard Form of Qudit Stabilizer Groups. *ArXiv e-prints*, jan 2011. doi: 10.1016/j.physleta.2013.12.009.

[145] G. C. Ghirardi, A. Rimini, and T. Weber. Unified dynamics for microscopic and macroscopic systems. *Phys. Rev. D*, 34:470–491, Jul 1986. doi: 10.1103/PhysRevD.34.470.

[146] G. C. Ghirardi, A. Rimini, and T. Weber. Unified dynamics for microscopic and macroscopic systems. *Phys. Rev. D*, 34:470–491, Jul 1986. doi: 10.1103/PhysRevD.34.470.

[147] G. C. Ghirardi, P. Pearle, and A. Rimini. Markov processes in Hilbert space and continuous spontaneous localization of systems of identical particles. *Physical Review A*, 42:78–89, July 1990. doi: 10.1103/physreva.42.78.

[148] N. Gisin. Stochastic quantum dynamics and relativity. *Helvetica Physica Acta*, 62(4):363–371, 1989.

[149] N. Gisin and A. Peres. Maximal violation of bell's inequality for arbitrarily large spin. *Physics Letters A*, 162(1):15 – 17, 1992. doi: 10.1016/0375-9601(92)90949-M.

[150] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden. Quantum cryptography. *Rev. Mod. Phys.*, 74:145–195, Mar 2002. doi: 10.1103/RevModPhys.74.145.

[151] A. Gleason. Measures on the closed subspaces of a hilbert space. *Indiana Univ. Math. J.*, 6:885–893, 1957. doi: 10.1512/iumj.1957.6.56050.

[152] S. Gogioso and W. Zeng. Mermin Non-Locality in Abstract Process Theories. *QPL 2015, Electronic Proceedings in Theoretical Computer Science*, Jun 2015.

[153] D. Gorenstein, R. Lyons, and R. Solomon. *The Classification of the Finite Simple Groups*. Number 3 in Mathematical surveys and monographs. American Mathematical Society, 1998. ISBN 9780821803912.

[154] D. Gottesman. The Heisenberg Representation of Quantum Computers. *Proceedings of the XXII International Colloquium on Group Theoretical Methods in Physics*.

[155] D. Gottesman. Stabilizer codes and quantum error correction. *Energy*, 2008:114, 1997.

[156] D. Gottesman. Fault tolerant quantum computation with higher dimensional systems. *Chaos Solitons Fractals*, 10:1749–1758, 1999. doi: 10.1016/S0960-0779(98)00218-5.

[157] D. M. Greenberger, M. A. Horne, and A. Zeilinger. Going beyond Bells theorem. In *Bells theorem, quantum theory and conceptions of the universe*, pages 69–72. Springer, 1989.

[158] D. Gross. Hudson's theorem for finite-dimensional quantum systems. *Journal of Mathematical Physics*, 47(12):122107, Dec. 2006. doi: 10.1063/1.2393152.

[159] L. K. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the Twenty-eighth Annual ACM Symposium on Theory of Computing*, STOC '96, pages 212–219. ACM, 1996. ISBN 0-89791-785-5. doi: 10.1145/237814.237866.

[160] P. Halmos. *A Hilbert Space Problem Book*. Graduate Texts in Mathematics. Springer New York, 1982. ISBN 9780387906850.

[161] L. Hardy. Nonlocality for two particles without inequalities for almost all entangled states. *Phys. Rev. Lett.*, 71:1665–1668, Sep 1993. doi: 10.1103/PhysRevLett.71.1665.

[162] L. Hardy. Quantum Theory From Five Reasonable Axioms. *eprint arXiv:quant-ph/0101012*, jan 2001.

[163] L. Hardy. Quantum ontological excess baggage. *Studies in History and Philosophy of Science Part B*, 35(2):267–276, 2004. doi: 10.1016/j.shpsb.2003.12.001.

[164] L. Hardy. Towards quantum gravity: a framework for probabilistic theories with non-fixed causal structure. *Journal of Physics A: Mathematical and Theoretical*, 40(12): 3081, 2007. doi: 10.1088/1751-8113/40/12/S12.

[165] N. Harrigan and T. Rudolph. Ontological models and the interpretation of contextuality. *ArXiv e-prints*, Sept. 2007.

[166] N. Harrigan and R. Spekkens. Einstein, Incompleteness, and the Epistemic View of Quantum States. *Foundations of Physics*, 40:125–157, Feb. 2010. doi: 10.1007/s10701-009-9347-0.

[167] N. Harrigan, T. Rudolph, and S. Aaronson. Representing probabilistic data via ontological models. *ArXiv e-prints*, Sep 2007.

[168] T. L. Heath and Euclid. *The Thirteen Books of Euclid's Elements*. Dover Publications, Incorporated, 1956. ISBN 0486600882.

[169] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81:865–942, Jun 2009. doi: 10.1103/RevModPhys.81.865.

[170] C. Horsman. Quantum picturalism for topological cluster-state computing. *New Journal of Physics*, 13(9):18, 2011. doi: 10.1088/1367-2630/13/9/095011.

[171] E. Hostens, J. Dehaene, and B. De Moor. Stabilizer states and clifford operations for systems of arbitrary dimensions and modular arithmetic. *Phys. Rev. A*, 71:042315, Apr 2005. doi: 10.1103/PhysRevA.71.042315.

[172] D. Hume. *A Treatise of Human Nature*. Dover Publications, 1739. ISBN 0198751729.

[173] I. Isaacs. *Algebra: A Graduate Course*. Graduate studies in mathematics. American Mathematical Society, 1994. ISBN 9780821847992.

[174] W. James. *The Principles of Psychology*. American science series: advanced course. H. Holt, 1890. ISBN 0486203816.

[175] P. Jaming, M. Matolcsi, and P. Mora. The problem of mutually unbiased bases in dimension 6. *ArXiv e-prints*, Jan 2012.

[176] R. Jozsa. An introduction to measurement based quantum computation. *eprint arXiv:quant-ph/0508124*, Aug 2005.

[177] R. Kadison and J. Ringrose. *Fundamentals of the Theory of Operator Algebras: Advanced theory*. Number v. 2 in Fundamentals of the Theory of Operator Algebras. American Mathematical Society, 1997. ISBN 9780821808207.

[178] I. Kant. *Critique of Pure Reason*. The Cambridge Edition of the Works of Immanuel Kant. Cambridge University Press, 1781. ISBN 9781107467057. Translated by Paul Guyer and Allen W. Wood.

[179] F. Karolyhazy. Gravitation and quantum mechanics of macroscopic objects. *Nuovo Cimento A Serie*, 42:390–402, Mar. 1966. doi: 10.1007/BF02717926.

[180] J. Kelley. *General Topology*. Graduate Texts in Mathematics. Springer, 1975. ISBN 9780387901251.

[181] G. M. Kelly and M. L. Laplaza. Coherence for compact closed categories. *Journal of Pure and Applied Algebra*, 19(1):193–213, 1980. doi: 10.1016/0022-4049(80)90101-2.

[182] A. Kent. Against Many-Worlds Interpretations. *International Journal of Modern Physics A*, 5:1745–1762, 1990. doi: 10.1142/S0217751X90000805.

[183] J. Kepler and W. Donahue. *Selections from Kepler's Astronomia Nova*. Science classics for humanities studies. Green Lion Press, 1609. ISBN 9781888009286.

[184] M. Kernaghan and A. Peres. Kochen-specker theorem for eight-dimensional space. *Physics Letters A*, 198(1):1 – 5, 1995. doi: 10.1016/0375-9601(95)00012-R.

[185] A. Kissinger. Exploring a Quantum Theory with Graph Rewriting and Computer Algebra. *Lecture Notes in Artificial Intelligence (LNCS/LNAI)*, Jul 2009.

[186] A. Kissinger. Pictures of processes: Automated graph rewriting for monoidal categories and applications to quantum computing. *ArXiv e-prints*, mar 2012.

[187] A. Kitaev. Fault-tolerant quantum computation by anyons. *Annals of Physics*, 303 (1):2–30, 2003. doi: 10.1016/S0003-4916(02)00018-0.

[188] A. Kitaev. Anyons in an exactly solved model and beyond. *Annals of Physics*, 321: 2–111, Jan. 2006. doi: 10.1016/j.aop.2005.10.005.

[189] A. Klappenecker and M. Roetteler. Constructions of Mutually Unbiased Bases. *eprint arXiv:quant-ph/0309120*, sep 2003.

[190] S. Kochen and E. Specker. The problem of hidden variables in quantum mechanics. *Journal of Mathematics and Mechanics*, 17(3):59–87, 1967.

[191] T. Körner. Metric and topological spaces. 2010.

[192] K. Kremnizer and A. Ranchin. Integrated Information-Induced Quantum Collapse. *Foundations of Physics*, 45(8):889–899, 2015. doi: 10.1007/s10701-015-9905-6.

[193] S. A. Kripke. Identity and necessity. In M. K. Munitz, editor, *Identity and Individuation*, pages 135–164. New York University Press, 1971. ISBN 0814753523.

[194] T. S. Kuhn. *The Structure of Scientific Revolutions*. University of Chicago Press, 1970. ISBN 0226458083.

[195] S. Kwapien. Isomorphic characterizations of inner product spaces by orthogonal series with vector valued coefficients. *Studia Mathematica*, 44(6):583–595, 1972.

[196] I. Lakatos, J. Worrall, and G. Currie. *The Methodology of Scientific Research Programmes: Volume 1: Philosophical Papers*. Cambridge paperback library. Cambridge University Press, 1980. ISBN 9780521280310.

[197] J. A. Larsson. A contextual extension of Spekkens' toy model. 1424:211–220, Mar 2012. doi: 10.1063/1.3688973.

[198] L. Laudan. *Progress and Its Problems: Towards a Theory of Scientific Growth.* Campus (Berkeley). University of California Press, 1978. ISBN 9780520037212.

[199] L. Laudan. A confutation of convergent realism. *Philosophy of Science*, 48(1):19–49, 1981.

[200] F. Lawvere. Functorial semantics of algebraic theories and some algebraic problems in the context of functorial semantics of algebraic theories. *Repr. Theory Appl. Categ.*, (5):1–121, 2004.

[201] C. M. Lee and J. Barrett. Computation in generalised probabilisitic theories. *New Journal of Physics*, 17(8):083001, 2015. doi: 10.1088/1367-2630/17/8/083001.

[202] C. M. Lee and J. H. Selby. Generalised phase kick-back: the structure of computational algorithms from physical principles. *New Journal of Physics*, 18(3):033023, 2016. doi: 10.1088/1367-2630/18/3/033023.

[203] A. J. Leggett and A. Garg. Quantum mechanics versus macroscopic realism: Is the flux there when nobody looks? *Phys. Rev. Lett.*, 54:857–860, Mar 1985. doi: 10.1103/PhysRevLett.54.857.

[204] G. Leibniz. *Monadology.* edition tablie par E. Boutroux, 1714. ISBN 1514389002.

[205] J. Levine. Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly*, 64(4):354361, 1983.

[206] P. G. Lewis, D. Jennings, J. Barrett, and T. Rudolph. Distinct quantum states can be compatible with a single state of reality. *Physical review letters*, 109(15):150404, 2012. doi: 10.1103/PhysRevLett.109.150404.

[207] Y.-C. Liang, R. W. Spekkens, and H. M. Wiseman. Speckers parable of the overprotective seer: A road to contextuality, nonlocality and complementarity. *Physics Reports*, 506(12):1 – 39, 2011. doi: 10.1016/j.physrep.2011.05.001.

[208] S. Lie. *Influence de Galois sur le développement des mathématiques.* Le centenaire de l'Ecole Normale 17951895. Hachette, 1895.

[209] E. Lieb and M. Loss. *Analysis.* Crm Proceedings & Lecture Notes. American Mathematical Society, 2001. ISBN 9780821827833.

[210] K. Lochan, S. Satin, and T. P. Singh. Statistical Thermodynamics for a Noncommutative Special Relativity: Emergence of a Generalized Quantum Dynamics. *Foundations of Physics*, 42:1556–1572, Dec 2012. doi: 10.1007/s10701-012-9683-3.

[211] J. Locke. *An essay concerning human understanding.* Prometheus Books, 1841. ISBN 0879759178.

[212] G. Ludwig. Attempt of an axiomatic foundation of quantum mechanics and more general theories, ii. *Communications in Mathematical Physics*, 4(5):331–348, 1967.

[213] S. Mac Lane. *Categories for the working mathematician.* Springer, 1998. ISBN 1441931236.

[214] W. Marshall, C. Simon, R. Penrose, and D. Bouwmeester. Towards quantum superpositions of a mirror. *Phys. Rev. Lett.*, 91:130401, Sep 2003. doi: 10.1103/PhysRevLett.91.130401.

[215] M. Massimini, F. Ferrarelli, S. K. Esser, B. A. Riedner, R. Huber, M. Murphy, M. J. Peterson, and G. Tononi. Triggering sleep slow waves by transcranial magnetic stimulation. *Proc Natl Acad Sci U S A*, 104(20):8496–8501, May 2007. doi: 10.1073/pnas.0702495104.

[216] C. McGinn. *The problem of consciousness: Essays toward a resolution.* 1991. ISBN 0631188037.

[217] N. D. Mermin. Simple unified form for the major no-hidden-variables theorems. *Phys. Rev. Lett.*, 65:3373–3376, Dec 1990. doi: 10.1103/PhysRevLett.65.3373.

[218] N. D. Mermin. Quantum mysteries revisited. *American Journal of Physics*, 58(8): 731–734, 1990. doi: 10.1119/1.16503.

[219] T. Metzinger. *Being No One: The Self-Model Theory of Subjectivity.* MIT Press, 2003. ISBN 0262633086.

[220] G. Miller. *Psychology: the science of mental life.* Harper & Row, 1962. ISBN 0140134891.

[221] A. Montina. Exponential complexity and ontological theories of quantum mechanics. *Phys. Rev. A*, 77:022104, Feb 2008. doi: 10.1103/PhysRevA.77.022104.

[222] A. Muthukrishnan and C. R. Stroud, Jr. Multivalued logic gates for quantum computation. *Phys. Rev. A*, 62(5):052309, Nov. 2000. doi: 10.1103/PhysRevA.62.052309.

[223] T. Nagel. What is it like to be a bat? *The philosophical review*, 83(4):435–450, 1974.

[224] I. Newton. *Philosophiae naturalis principia mathematica.* J. Societatis Regiae ac Typis J. Streater, 1687. ISBN 0520088174.

[225] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*, volume 70. Cambridge University Press, 2000.

[226] S. Nimmrichter and K. Hornberger. Macroscopicity of Mechanical Quantum Superposition States. *Physical Review Letters*, 110(16):160403, Apr. 2013. doi: 10.1103/PhysRevLett.110.160403.

[227] E. Noether. Invariant variation problems. *Transport Theory and Statistical Physics*, 1:186–207, 1918. doi: 10.1080/00411457108231446.

[228] J. Pachos. *Introduction to Topological Quantum Computation.* Introduction to Topological Quantum Computation. Cambridge University Press, 2012. ISBN 9781107005044.

[229] A. K. Pati and S. L. Braunstein. Impossibility of deleting an unknown quantum state. *Nature*, 404(6774):164–165, 2000. doi: 10.1038/404130b0.

[230] P. Pearle. Reduction of the state vector by a nonlinear Schrödinger equation. *Physical Review D*, 13(4):857–868, Feb. 1976. doi: 10.1103/physrevd.13.857.

[231] B. Peirce. Linear associative algebra. *American Journal of Mathematics*, 4(1):97–229, 1881. doi: 10.2307/2369153.

[232] R. Penrose. On Gravity's role in Quantum State Reduction. *General Relativity and Gravitation*, 28(5):581–600, May 1996. doi: 10.1007/bf02105068.

[233] R. Penrose. Quantum computation, entanglement and state reduction. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 356(1743):1927–1939, 1998. doi: 10.1098/rsta.1998.0256.

[234] A. Peres. Neumark's theorem and quantum inseparability. *Foundations of Physics*, 20(12):1441–1453, 1990. doi: 10.1007/BF01883517.

[235] A. Peres. Two simple proofs of the Kochen-Specker theorem. *Journal of Physics A: Mathematical and General*, 24(4):L175, 1991. doi: 10.1088/0305-4470/24/4/003.

[236] A. Peres. *Quantum Theory: Concepts and Methods*. Fundamental Theories of Physics. Springer, 1995. ISBN 9780792336327.

[237] A. Peres. All the Bell inequalities. *Found.Phys.*, 29:589–614, 1999. doi: 10.1023/A:1018816310000.

[238] M. Planat. On small proofs of the Bell-Kochen-Specker theorem for two, three and four qubits. *European Physical Journal Plus*, 127:86, Aug. 2012. doi: 10.1140/epjp/i2012-12086-x.

[239] Plato. *Timaeus and Critias*. Classics Series. Penguin Books, 1971. ISBN 9780140442618.

[240] Plato. *The Republic of Plato*. Basic Books, September 1991. ISBN 0465069347.

[241] G. Pólya. *Ü*ber die analogie der kristallsymmetrie in der ebene. *Zeitschrift fr Kristallographie*, 60:278282, 1924.

[242] S. Popescu and D. Rohrlich. Which states violate Bell's inequality maximally? *Physics Letters A*, 169(6):411 – 414, 1992. doi: 10.1016/0375-9601(92)90819-8.

[243] S. Popescu and D. Rohrlich. Nonlocality as an axiom. *Foundations of Physics*, 24 (379), 1994. doi: 10.1007/BF02058098.

[244] K. Popper. *The Logic of Scientific Discovery*. Classics Series. Routledge, 1959. ISBN 9780415278447.

[245] K. Popper. *Three worlds - The Tanner Lecture on Human Values*. Delivered by Karl Popper at The University of Michigan. 1979.

[246] T. U. F. Program. *Homotopy Type Theory: Univalent Foundations of Mathematics*. 2013.

[247] M. Pusey. Stabilizer Notation for Spekkens Toy Theory. *Foundations of Physics*, 42: 688–708, 2012. doi: 10.1007/s10701-012-9639-7.

[248] M. F. Pusey, J. Barrett, and T. Rudolph. On the reality of the quantum state. *Nature Physics*, 8:476–479, June 2012. doi: 10.1038/nphys2309.

[249] W. V. Quine. Two Dogmas of Empiricism. *The Philosophical Review*, 60(1):20–43, Jan. 1951. doi: 10.2307/2181906.

[250] A. Ranchin. Depicting qudit quantum mechanics and mutually unbiased qudit theories. *QPL 2014, Electronic Proceedings in Theoretical Computer Science*, 172:68–91, Apr 2014. doi: 10.4204/EPTCS.172.6.

[251] A. Ranchin and B. Coecke. Complete set of circuit equations for stabilizer quantum mechanics. *Phys. Rev. A*, 90:012109, Jul 2014. doi: 10.1103/PhysRevA.90.012109.

[252] R. Raussendorf and H. J. Briegel. A one-way quantum computer. *Phys. Rev. Lett.*, 86:5188–5191, May 2001. doi: 10.1103/PhysRevLett.86.5188.

[253] R. Raussendorf, J. Harrington, and K. Goyal. Topological fault-tolerance in cluster state quantum computation. *New Journal of Physics*, 9(6):199, 2007. doi: 10.1088/1367-2630/9/6/199.

[254] T. Reid. *Essays on the Intellectual Powers of Man*. 1785. ISBN 0748611894.

[255] O. Robertshaw and P. Tod. Lie point symmetries and an approximate solution for the Schrodinger–Newton equations. *Nonlinearity*, 19(7):1507, 2006. doi: 10.1088/0951-7715/19/7/002.

[256] O. Romero-Isart, L. Clemente, C. Navau, A. Sanchez, and J. I. Cirac. Quantum magnetomechanics with levitating superconducting microspheres. *Phys. Rev. Lett.*, 109:147205, Oct 2012. doi: 10.1103/PhysRevLett.109.147205.

[257] T. Rudolph. Ontological Models for Quantum Mechanics and the Kochen-Specker theorem. *eprint arXiv:quant-ph/0608120*, Aug 2006.

[258] H. Salzmann. *The Classical Fields: Structural Features of the Real and Rational Numbers*. Encyclopedia of Mathematics. Cambridge University Press, 2007. ISBN 9780521865166.

[259] F. Saussure, C. Bally, A. Sechehaye, and A. Riedlinger. *Course in General Linguistics*. Open Court classics. 1983. ISBN 9780812690231.

[260] V. Scarani. Feats, Features and Failures of the PR-box. *AIP Conference Proceedings (Melville, New York)*, 844:309–320, 2006. doi: 10.1063/1.2219371.

[261] U. Schoning. Graph isomorphism is in the low hierarchy. *Journal of Computer and System Sciences*, 37(3):312 – 323, 1988. doi: 10.1016/0022-0000(88)90010-4.

[262] O. Schreiber and R. W. Spekkens. Reconstruction of the stabilizer formalism for qutrits from a statistical theory of trits with an epistemic restriction. *to be published*, 2012.

[263] E. Schrödinger. Discussion of probability relations between separated systems. *Mathematical Proceedings of the Cambridge Philosophical Society*, 31(04):555–563, 1935.

[264] E. Schrödinger. *What is life? The physical aspect of the living cell.* Cambridge, 1944. ISBN 1107604664.

[265] J. R. Searle. Consciousness. *Annual Review of Neuroscience*, 23(1):557–578, 2000.

[266] P. Selinger. Dagger compact closed categories and completely positive maps (extended abstract). *Electronic Notes in Theoretical Computer Science*, 170:139–163, 2007. doi: 10.1016/j.entcs.2006.12.018.

[267] P. Selinger. A survey of graphical languages for monoidal categories. *New Structures for Physics*, pages 1–63, 2009. doi: 10.1007/978-3-642-12821-9/4.

[268] P. Selinger. Finite dimensional Hilbert spaces are complete for dagger compact closed categories. *ArXiv e-prints*, Jul 2012.

[269] P. Selinger. Generators and relations for n-qubit Clifford operators. *ArXiv e-prints*, Oct 2013.

[270] G. Shafer. *A Mathematical Theory of Evidence.* Princeton University Press, 1976. ISBN 069110042X.

[271] W. Shakespeare. *Antony and Cleopatra.* Giunti classics. Giunti Editore, 1606. ISBN 9788809020856.

[272] P. W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Comput.*, 26(5):1484–1509, Oct. 1997. doi: 10.1137/S0097539795293172.

[273] R. D. Sorkin. Quantum Measure Theory and its Interpretation. *ArXiv General Relativity and Quantum Cosmology e-prints*, July 1995.

[274] E. Specker. Die logik nicht gleichzeitig entscheidbarer aussagen. *Dialectica*, 1(14): 239–246, 1960. doi: 10.1111/j.1746-8361.1960.tb00422.x.

[275] R. W. Spekkens. Contextuality for preparations, transformations, and unsharp measurements. *Phys. Rev. A*, 71:052108, May 2005. doi: 10.1103/PhysRevA.71.052108.

[276] R. W. Spekkens. Evidence for the epistemic view of quantum states: A toy theory. *Physical Review A*, 75(3):032110, 2007. doi: 10.1103/PhysRevA.75.032110.

[277] R. W. Spekkens, D. H. Buzacott, A. J. Keehn, B. Toner, and G. J. Pryde. Preparation contextuality powers parity-oblivious multiplexing. *Phys. Rev. Lett.*, 102:010401, Jan 2009. doi: 10.1103/PhysRevLett.102.010401.

[278] H. Stapp. The basis problem in many-worlds theories. *Canadian Journal of Physics*, 80(9):1043–1052, 2002. doi: 10.1139/p02-068.

[279] H. Stapp. *Mindful Universe: Quantum Mechanics and the Participating Observer.* The Frontiers Collection. Springer, 2011. ISBN 9783642180767.

[280] M. H. Stone. The Theory of Representation for Boolean Algebras. *Transactions of the American Mathematical Society*, 40(1):37–111, 1936. doi: 10.2307/1989664.

[281] A. Tarski. The concept of truth in the languages of the deductive sciences. *Zygmunt*, 34:13172, 1933.

[282] M. Tegmark. Consciousness as a State of Matter. *ArXiv e-prints*, Jan. 2014.

[283] P. R. Thagard. Why astrology is a pseudoscience. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 1978:223–234, 1978.

[284] M. Thomson. *Modern Particle Physics*. Modern Particle Physics. Cambridge University Press, 2013. ISBN 9781107034266.

[285] N. Tinbergen. On aims and methods of ethology. *Zeitschrift fr Tierpsychologie*, 20: 410–433, 1963.

[286] B. F. Toner and D. Bacon. Communication cost of simulating bell correlations. *Phys. Rev. Lett.*, 91:187904, Oct 2003. doi: 10.1103/PhysRevLett.91.187904.

[287] G. Tononi. An information integration theory of consciousness. *BMC Neuroscience*, 5(1):42, 2004. doi: 10.1186/1471-2202-5-42.

[288] G. Tononi. Consciousness as integrated information: a provisional manifesto. *The Biological Bulletin*, 215(3):216–242, 2008.

[289] W. van Dam. Implausible Consequences of Superstrong Nonlocality. *eprint arXiv:quant-ph/0501159*, Jan. 2005.

[290] M. Van den Nest, J. Dehaene, and B. De Moor. Graphical description of the action of local clifford transformations on graph states. *Phys. Rev. A*, 69:022316, Feb 2004.

[291] S. J. van Enk. A Toy Model for Quantum Mechanics. *Foundations of Physics*, 37: 1447–1460, Oct. 2007. doi: 10.1007/s10701-007-9171-3.

[292] B. C. Van Fraassen. The Semantic Approach to Scientific Theories. pages 105–124, 1987. doi: 10.1007/978-94-009-3519-8_6.

[293] J. Van Heijenoort. *From Frege to Godel: A Source Book in Mathematical Logic, 1879-1931*. Source books in the history of the sciences. Harvard University Press, 1967. ISBN 9780674324497.

[294] J. von Neumann. *The Mathematical Foundations of Quantum Mechanics*. Princeton Univ. Press, Princeton NJ, 1st edition, 1932.

[295] D. Wallace. Quantum Probability and Decision Theory, Revisited. *eprint arXiv:quant-ph/0211104*, Nov. 2002.

[296] Q. Wang and X. Bian. Qutrit Dichromatic Calculus and its Universality. *ArXiv e-prints*, Jun 2014.

[297] I. Weiss. A note on the metrizability of spaces. *Algebra Universalis*, 73(2):179–182, 2015. doi: 10.1007/s00012-015-0319-2.

[298] E. P. Wigner. Remarks on the mind-body question. pages 284–304. 1962.

[299] L. Wittgenstein. *Philosophical Investigations*. Basil Blackwell, 1953. ISBN 0631146709.

[300] L. Wittgenstein. Notes for lectures on private experience and sense-data. *The Philosophical Review*, 77:275320, 1968.

[301] W. K. Wootters and W. H. Zurek. A single quantum cannot be cloned. *Nature*, 299 (5886):802–803, Oct. 1982. doi: 10.1038/299802a0.

[302] E. Zermelo. Untersuchungen uber die grundlagen der mengenlehre. i. *Mathematische Annalen*, 65(2):261–281, 1908. doi: 10.1007/BF01449999.

[303] M. Ziegler. Computational power of infinite quantum parallelism. *International Journal of Theoretical Physics*, 44(11):2059–2071, 2005. doi: 10.1007/s10773-005-8984-0.