Contrasting classical and quantum theory in the context of quasi-probability

by

Christopher Scott Ferrie

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Applied Mathematics

Waterloo, Ontario, Canada, 2008

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Several finite dimensional quasi-probability representations of quantum states have been proposed to study various problems in quantum information theory and quantum foundations. These representations are often defined only on restricted dimensions and their physical significance in contexts such as drawing quantum-classical comparisons is limited by the non-uniqueness of the particular representation. In this thesis it is shown how the mathematical theory of frames provides a unified formalism which accommodates all known quasi-probability representations of finite dimensional quantum systems.

It is also shown that any quasi-probability representation is equivalent to a frame representation and it is proven that any such representation of quantum mechanics must exhibit either negativity or a deformed probability calculus.

Along the way, the connection between negativity and two other famous notions of non-classicality, namely contextuality and nonlocality, is clarified.

This thesis is an extension of work found in [16].

Acknowledgements

I thank my supervisor Joseph Emerson for the many spontaneous discussions we have had over the past two years. He has provided me with much advice and guidance not limited to my academic career and I thank him again for that. I also thank the other members of our research group: Cozmin Ududec for his clarity during our many debates; Easwar Magesan for his patience while answering my various technical question; Ryan Morris and Matt Leifer for many inspiring discussions.

Thanks go to John Watrous for technical assistance in the main proof of the thesis. I also thank Rob Spekkens for an influential discussion we had early in my graduate studies. I have also benefitted from the opportunities Chris Fuchs and Achim Kempf have provided by inviting me to present my work to their research groups and I thank them for it. Thanks also go to the members of the Institute for Quantum Computing, especially those who listened and provided useful feedback when I presented my work during the weekly lunch meetings. I thank Helen Warren and Ann Puncher for their support in many administrative tasks and the Natural Sciences and Engineering Research Council of Canada for financial support.

Lastly, I thank my family and friends for their moral support. In particular, I thank my wife Lindsay MacMillan for seeing me through.

Dedication

To Mom, for a lifetime of love and encouragement.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

At nearly a century old, quantum theory is humankind's most successful physical theory. However, its conceptual foundations are still debated today. Quantum theory is so counterintuitive that all other physical theories, when compared to it, are commonly referred to as "classical". The differences between the theories of classical and quantum physics are not hard to find. But, perhaps this is only because the mathematics of these theories are so different. And, perhaps if we are able to find a common mathematical language, many differences between quantum and classical theory will disappear. What is left may give insights into further understanding of both theories and how they contrast.

One approach in establishing a common ground between classical and quantum theory is phase space. Phase space is a natural concept in classical theory since it is equivalent to the state space. The idea of formulating quantum theory in phase space dates back to the early days of quantum theory when the so-called Wigner function was introduced [49]. The Wigner function is a *quasi-probability* distribution on a classical phase space which represents a quantum state[1]. The term quasi-probability refers to the fact that the function is not a true probability density as it takes on negative values for some quantum states. The Wigner function formalism can be lifted into a fully autonomous phase space theory which reproduces all the predictions of the standard quantum theory of infinite dimensional systems [3]. In other words, this phase space formulation of quantum theory is equivalent to the usual abstract formalism of quantum theory in the same sense that Heisenburg's matrix mechanics and Schrodinger's wave mechanics are equivalent to the abstract formalism.

In the abstract formulation of quantum theory there are many conceptual barriers to overcome in gaining an intuition for the behaviour of a quantum system. However, the phase space formulation allows for visualization and other analytical techniques that are already well understood and applied to classical probability

[1]Note that Wigner function is not the only such function. A review of the Wigner function and other choices appears in [34].

distributions. In this way, many conceptual problems are replaced by one: negative probability.

There is, however, one problem with the Wigner function approach: it is not valid when the quantum system it is describing has finitely many distinguishable states. Approaches to overcome this limitation have only been considered recently, probably motivated by experimental advances in the coherent control of finite dimensional quantum systems such as nuclear spins, superconducting circuits and trapped atomic systems. These systems cannot be described by the Wigner function phase space formulation discussed above, which is often called the *continuous phase space* approach.

In recent years various analogs of the Wigner function for finite dimensional quantum systems have been proposed. Again, here, the term *quasi-probability* is often used because of the appearance of negative values meant to represent probabilities. Such *discrete phase space* representations of quantum states have provided insight into fundamental structures for finite-dimensional quantum systems. For example, the representation proposed by Wootters identifies sets of mutually unbiased bases [50, 20]. Inspired by the discovery that quantum resources lead to algorithms that dramatically outperform their classical counterparts, there has also been growing interest in the application of discrete phase space formalism to analyze the quantum-classical contrast for finite-dimensional systems. Examples are quantum teleportation [41], the effect of decoherence on quantum walks [37], quantum Fourier transform and Grover's algorithm [38], conditions for exponential quantum computational speedup [18, 12], and quantum expanders [24].

Note that the term *phase space* is often given a meaning independent of the notion of state although it can be identified with the state space. The discrete phase space is an analogy to the classical phase space. Since this analogy is quite weak in some examples, the discrete phase space is more accurately described as a (discrete) state space. Where the analogy is weak, the terminology *phase space representation* is replaced with *quasi-probability representation*. A quasi-probability representation is a more general concept which includes phase space representations.

As noted above, central concept in studies of the quantum-classical contrast in the quasi-probability formalisms of quantum theory is the appearance of *negativity*. A non-negative quasi-probability function is a true probability distribution, prompting some authors to suggest that the presence of negativity in this function is a defining signature of non-classicality. Unfortunately the application of any one of these quasi-probability representations in the context of determining criteria for the non-classicality of a given quantum task is limited in significance by the non-uniqueness of that particular representation. Ideally one would like to determine whether the task can be expressed as a classical process in *any* quasi-probability representation. Indeed the sheer variety of proposed quasi-probability representations prompts the question of whether there is some shared underlying mathematical structure that might provide a means for identifying the full family of such representations.

Moreover, from an operational view, states alone are an incomplete description of an experimental arrangement. For example, Reference [12] proves that within a class of quasi-probability representations, the only positive pure states are the so-called stabilizer states. Hence, these states are "classical" from the point of view of allowing an efficient classical simulation via the stabilizer formalism [22]. However, this set of positive states includes the Bell state; and, Bell states maximally violate a Bell inequality [4]. Hence by a more conventional criteria of "classicality", namely *locality*, these states are maximally *non-classical*. The resolution of this paradox is that one must also consider a self-consistent representation of measurements in order to assess the classicality of *an entire experimental procedure*. Hence, it is important to elucidate the ways in which a quasi-probability representation *of states alone* can be lifted into an autonomous quasi-probability representation of *both the states and measurements* defining any set of experimental configurations.

The purpose of this thesis is to present a mathematical structure which underlies the known quasi-probability representation for finite dimensional quantum states. Also, this thesis outlines a general construction for lifting any representations of quantum states alone to a fully autonomous formulation of the whole quantum formalism. Along the way, it will be shown how non-locality, and more generally *contextuality* (another well-studied criterion for non-classicality), relates to presence of negativity in quasi-probability representations.

The outline of this thesis is as follows. Chapter 2 introduces the mathematical formulations of classical and quantum theory from an operational perspective. Chapter 3 offers a review of the existing quasi-probability representations found in the literature. In Chapter 4, the mathematical theory of *frames* is introduced and shown to be the sought after mathematical structure which unifies the quasi-probability representations of quantum states. Examples of *frame representations* are given in Chapter 5. In Chapter 6, it is shown, using frame theory, how to consistently incorporate quantum measurements into any quasi-probability formalism. It is shown in Chapter 7 that negativity in quasi-probability representations is necessary. The connection between negativity and another notion of non-classicality, *contextuality*, is studied in Chapter 8. The connection between negativity and, perhaps the most famous notion of non-classicality, *non-locality* is studied in Chapter 9. Conclusions and directions for future research is presented in Chapter 10. A introduction to some mathematical concepts and notations is provided in Appendix A. Some pictorial representations of Wigner functions are provided in Appendix B.

Chapter 2

Probability in classical and quantum theory

An introduction to the operational formulations of classical and quantum theory are presented in the following two sections. A standard treatment of classical physics heavily relies on dynamical equations of motion [21]. The same is true in quantum theory which is often introduced through the *quantization* of classical dynamical systems. Such a treatment is often omitted when considering an operational perspective of a given experiment. To be specific, an operational theory is one which specifies a set of instructions (called preparations and measurements) for an experiment. The role is to specify the outcomes (perhaps only probabilistically) of measurements performed when the preparation procedure is given.

2.1 Probabilistic structure of classical theory

Studying quantum theory gives a perspective on classical theories which one normally would not have. From this perspective one notices that many assumptions go unstated in an exposition of a classical theory. Two such assumptions are objectivity and determinism: a physical object that could be experimented on exists and possess properties whether a scientist is there to measure (or think about measuring) these properties or not and a scientist could in principal know with arbitrary accuracy the exact value of the properties possessed by the object. Of course technical constraints will always limit the accuracy of an experiment. However, in classical theories it is always assumed that a more accurate experiment could be devised.

On the other hand, as one requires more accuracy to predict the roll of dice for example, the experimental devices required to produce such accuracy are most likely to become unimaginably large and complex. Again, although it is assumed that such a device could be built, it is certain that a casino, for example, would not let you use such a device at its craps table. In the face of this necessary uncertainty, probability theory is used to gain as much constrained knowledge of the outcome of experiments as possible.

Consider a practical situation in which certainty in the knowledge (or acquisition of knowledge) of the properties of a classical system may or may not be obtained. In particular, consider a set of experimental configurations: a preparation device which produces a variety of classical systems and a measuring device which outputs a numerical value on a screen, audible "click" or some other sensory cue for a variety of detector settings. Such a situation is described mathematically as follows.

Let the set $\mathcal{S}$ along with the positive measure μ represent the properties of a classical system and the function $\rho(s) \geq 0$ represent the probabilistic knowledge of these properties. A measurement is a partitioning of the space $\mathcal{S}$ into disjoint subsets $\{\triangle_j\}$. The probability that the system has properties in $\triangle_j$ (called "outcome j") is

$$\Pr(j) = \int_{\triangle_j} d\mu(s)\ \rho(s) = \int_{\mathcal{S}} d\mu(s)\chi_j(s)\rho(s),$$

where $\chi_j(s) \in \{0,1\}$ is the indicator function of $\triangle_j$. The measurement is equivalently specified by the set $\{\chi_j(s)\}$. Each $\chi_j(s)$ is interpreted as the conditional probability of outcome j given the system is known to have the properties s. A measurement of this type is deterministic; it reveals with certainty the properties of the system. Consider now an indeterministic measurement specified by the conditional probabilities $\{M_k(s) \in [0,1]\}$. The above description is summarized with the following concise definition.

Definition 2.1.1. *Any model of a set of experimental configurations is a* classical probabilistic description *if all of the following properties hold.*

(a) There is a set of allowed properties $\mathcal{S}$ with a positive measure μ.

(b) A preparation (state) is represented by a probability density $\rho(s) \geq 0$ which satisfies the normalization condition $\int_{\mathcal{S}} d\mu(s)\rho(s) = 1$.

(c) A measurement is represented by a set $\{M_k(s) \in [0,1]\}$ which satisfies $\sum_k M_k(s) = 1$.

(d) For a system with probability density ρ subject to the measurement $\{M_k\}$, the probability of obtaining outcome k is given by the law of total probability

$$\Pr(k) = \int_{\mathcal{S}} d\mu(s)\rho(s)M_k(s). \tag{2.1}$$

To ensure the understanding of these concepts, they will now be applied to the simplest example possible: a coin toss. It is assumed that the results of a coin could be predicted with certainty if the experimenter had enough control over the tossing of the coin, the surface it lands on, and the vast number of air and dust particles striking it while in flight. Again, it would be painstakingly difficult to devise such a precise experiment. In a typical coin toss then, only two properties of the coin can be distinguished: heads and tails. The set of allowed properties or state space is then $\mathcal{S} = \{H, T\}$. Here the notion of measure is not needed. However, to conform to the definitions, the measure μ can be taken to be the counting measure. Then, any integration becomes summation over the state space. A preparation might be to choose a coin and take it in hand. Since the outcome of any measurement (toss) is uncertain, this preparation corresponds to the probability distribution $\rho(H) = \frac{1}{2}$ and $\rho(T) = \frac{1}{2}$. A measurement is the toss of the coin including its landing. This measurement is deterministic and is thus represented by the two indicator functions χ_H and χ_T. Verifying the law of total probability for the outcome H (heads) yields

$$\Pr(H) = \sum_{j=\{H,T\}} \rho(j)\chi_H(j) = \rho(H)\chi_H(H) + \rho(T)\chi_H(T) = \frac{1}{2}(1) + \frac{1}{2}(0) = \frac{1}{2},$$

as expected. This example is easily extrapolated to any of the usual applications of discrete and continuous probability densities: dice, cards, stock markets, etc.

A classical probabilistic description models many careful experiments done in physics labs as well. However, at the turn of the twentieth century, new experiments began to surface which violated the long standing physical laws of Newton. A new theory was slowly built out of necessity to describe the results of these experiments which seemed to suggest that uncertainty was an unavoidable fact of nature. This new theory came to be called quantum theory and it is described in the next section.

2.2 Probabilistic structure of quantum theory

As opposed to a classical theory, quantum theory is not derived from a set of physical axioms. It begins with a specification of the mathematical objects representing states, observables and values obtained in measurements. An introductory textbook [7] will present quantum theory as a set of laws or postulates assigning mathematical objects to the physical concepts of state, observable and measurement. Each state is assigned a vector in a Hilbert space. The observables are assigned self-adjoint operators and the values obtained in a measurement of an observable are the eigenvalues of the operator associated with that observable. If an observable is given a name, say A, then the operator assigned to that observable is denoted $\hat{A}$. Denote the eigenvalues and corresponding eigenvectors of $\hat{A}$ as $\{a_n, \psi_n\}$. If the state of the system is one of the eigenvectors ψ_n, the value of A obtained in a measurement is always a_n. However, there is no deterministic rule for the value obtained in a measurement of A when the state of the system is not an eigenvalue. The last postulate of quantum theory resolves this problem but at the expense of introducing uncertainty: the *probability* of obtaining the value a_n in a measurement of the observable A when the system is in state ψ is $|\langle \psi, \psi_n \rangle|^2$.

Ignoring foundational and interpretational issues, what has been described so far can be thought of as a traditional or orthodox approach to quantum theory. A more general and modern approach exists and its applications include fields such as quantum computing and quantum information theory. This modern approach often comes with the pragmatic interpretation that quantum theory is nothing more that a new theory of probability which may be necessary to describe experiments for which a classical probabilistic description does not exists. One variant of this approach is called *operationalism* [25]. This is a formulation of quantum theory which describes a set of experimental configurations as a preparation device which produces a variety of "physical systems" and a measuring device which outputs a numerical value on a screen, audible "click" or some other sensory cue for a variety of detector settings. Note that, despite the fact that the "physical system" is not assumed to be classical, this is the same experimental scenario that a classical probabilistic description is applied to. An operational set of axioms for quantum theory are as follows.

Definition 2.2.1. *Any model of a set of experimental configurations is a* quantum probabilistic description *if it can be modeled as follows.*

(i) There exists a Hilbert space $\mathcal{H}$ *with* $\dim(\mathcal{H}) = d$.

(ii) A preparation (state) is represented by a density operator $\hat{\rho}$ *satisfying* $\langle\psi, \hat{\rho}\psi\rangle \geq 0$*, for all* $\psi \in \mathcal{H}$*, and* $\mathrm{Tr}(\hat{\rho}) = 1$.

(iii) A measurement is represented by a set of effects $\{\hat{M}_k\}$ *satisfying* $0 \leq \langle\psi, \hat{M}_k\psi\rangle \leq 1$*, for all* $\psi \in \mathcal{H}$*, and* $\sum_k \hat{M}_k = \hat{\mathbb{1}}$.

(iv) For a system with density operator $\hat{\rho}$ *subject to the measurement* $\{\hat{M}_k\}$*, the probability of obtaining outcome* k *is given by the Born rule*

$$\Pr(k) = \mathrm{Tr}(\hat{M}_k\hat{\rho}). \tag{2.2}$$

The simplest example of such an experiment is the *Stern-Gerlach experiment* (see Figure 2.1 for a schematic diagram). A beam of silver atoms is directed at an inhomogeneous magnetic field. According to classical electrodynamics, the beam should be deflected in a continuum of directions. However, when the experiment is performed, the beam is deflected into two distinct directions. This happens no matter how the magnet is oriented.

In a operational sense, the silver atoms are prepared by the source and the measurement is made by the magnet (Stern-Gerlach device or S-G device) with two outcomes. Consider what happens when two identical S-G devices are placed in succession (Figure 2.2(a)). The first S-G device is aligned with the z-axis (of the lab frame) and by blocking one beam of particles it *prepares* "z-up" particles. As one would expect from classical intuition, the second S-G device *measures* the particles also to "be" (or have the "property") "z-up". One might think, then, this experiment is exactly analogous to a classical coin toss with "z-up" and "z-down" replacing "heads" and "tails". However, something peculiar happens when a sequence of Stern-Gerlach experiments are made. Consider the same situation

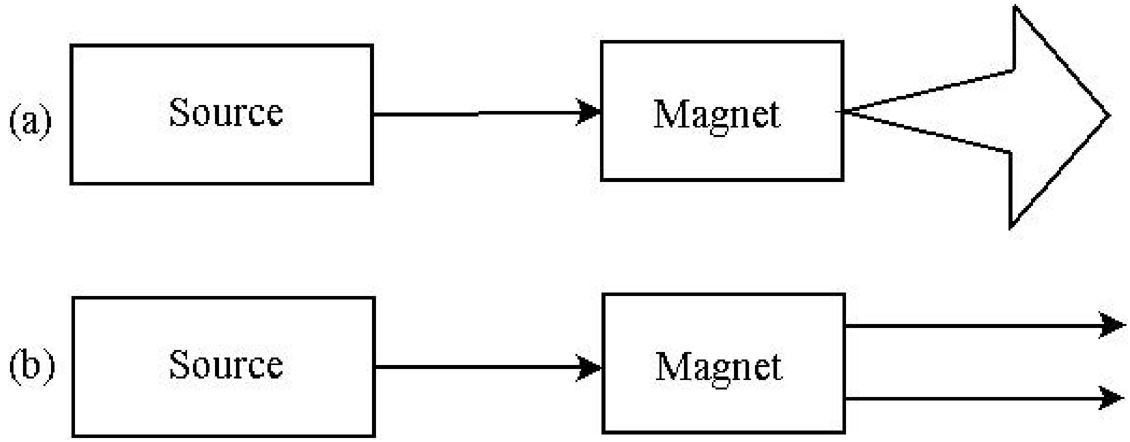

Figure 2.1: Schematic of the Stern-Gerlach experiment in (a) classical electrodynamics and (b) actual experiment.

but now the second S-G device is aligned with the x-axis (i.e. perpendicular to the first S-G device). This set-up is depicted in Figure 2.2(b). Again the first S-G device prepares "z-up" particles. The second S-G device measures whether the particles have the "property" "x-up" or "x-down". That is fine, each particle can have two "properties" (its z-direction and its x-direction) and this experiment has determined which two each particle "possesses", right? Wrong! Consider a third S-G device (aligned along z) placed after the "x-down" beam has been blocked in the previous example. This is depicted in Figure 2.2(c). Classical intuition suggests that the particles entering the third S-G device have the "z-up" "property" since they were prepared that way by the first S-G device. However, what happens in the experiment is a random mixture of "z-up" and "z-down" outcomes. So much for classical "properties"!

Certainly the Stern-Gerlach experiment is not definitive proof against objectivity and determinism although such experimental results (and those like them) have been difficult (if not impossible) to explain from any physical theory with intuitive classical assumptions. On the other hand, quantum theory has proven to be an incredibly accurate tool in predicting the probabilistic results of experiments such as the Stern-Gerlach experiment. This does not immediately suggest that a classical *probabilistic* description is impossible. Indeed, already some alignment can be seen in the classical (Definition 2.1.1) and quantum (Definition 6.2.1) description of an experiment. In fact, the only difference seems to be the mathematical language in which they are stated. There is no *a priori* reason to believe that the two are not (mathematically) equivalent. If this were the case, then the quantum density operator (state) could be interpreted as representing probabilistic knowledge over some classical state space. Building such an interpretation of quantum theory is the goal of the *hidden variables* program [19]. The hidden variables are classical states which are undetectable according to quantum theory. However, as discussed later in Chapters 7 through 9, these models have necessary features which

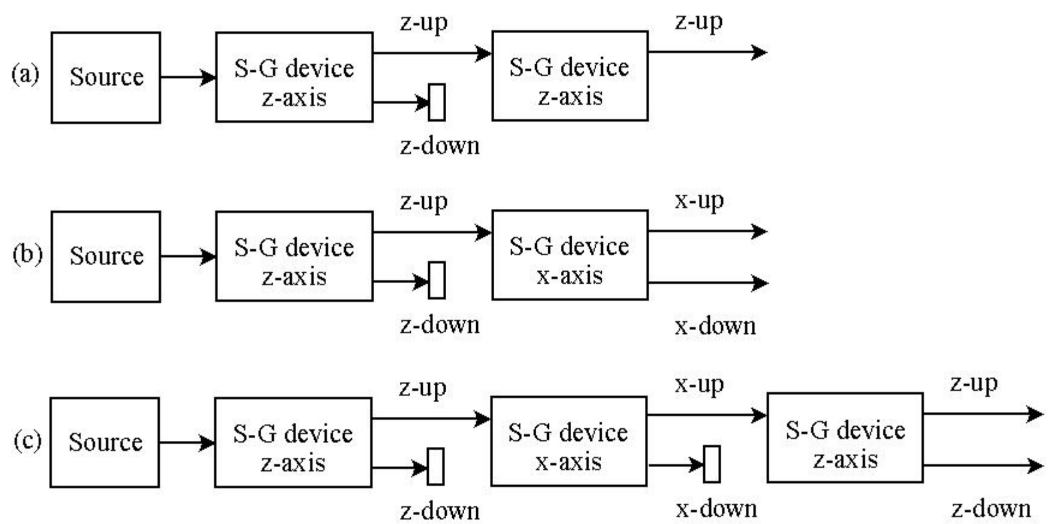

Figure 2.2: Sequences of Stern-Gerlach experiments.

are thought by most to be *non*-classical by criteria such as *negativity*, *non-locality*, and *contextuality*.

Chapter 3

Quasi-probability representations of quantum states

Reviewed in this chapter are the existing phase space formalisms of quantum states found in the literature. The original phase space representation put forth by Wigner and later realized as an alternative, and equivalent, formulation of the full quantum theory (of particles) by Moyal [39] and others [3] is reviewed first. This phase space picture is valid for infinite dimensional Hilbert spaces but it will be presented here as it has motivated all known phase space pictures for finite dimensions. Sections 3.1-3.6 of this chapter are devoted to reviewing a representative sample of the known quasi-probability representations of finite dimensional quantum states.

In Section 3.7 a summary of these and a few other quasi-probability representations is presented in a concise manner. The purpose of the chapter shifts from presentation of the known quasi-probability representations to unification in Section 3.8 where the precise mathematical definition of quasi-probability representation of quantum states is given.

3.1 Wigner phase space representation

The position operator, $\hat{Q}$ and momentum operator, $\hat{P}$, are the central objects in the abstract formalism of infinite dimensional quantum theory. The operators satisfy the canonical commutation relations

$$[\hat{Q}, \hat{P}] = i\hbar.$$

Since $\hat{Q}$ and $\hat{P}$ do not commute, the choice of the quantization map $(q, p) \mapsto (\hat{Q}, \hat{P})$ is not unique. This is the so-called "ordering problem". A class of solutions to this problem is the association $e^{i\xi q+i\eta p} \mapsto e^{i\xi\hat{Q}+i\eta\hat{P}} f(\xi, \eta)$ for some arbitrary function f (See Table 1 of [34] for a review of the traditional choices for f).

Consider the classical particle phase space $\mathbb{R}^2$ and the continuous set of operators

$$\left\{ \hat{F}(q,p) := \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} d\xi d\eta \; e^{i\xi(q-\hat{Q})+i\eta(p-\hat{P})} f(\xi,\eta) : (q,p) \in \mathbb{R}^2 \right\}. \tag{3.1}$$

When $f(\xi,\eta) = 1$, the distribution

$$\rho^{\text{Wigner}}(q,p) := \text{Tr}(\hat{\rho}\hat{F}(q,p)) \tag{3.2}$$

$$= \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} d\xi d\eta \; \text{Tr}\left[\hat{\rho} e^{i\xi(\hat{Q}-q)+i\eta(\hat{P}-p)}\right] \tag{3.3}$$

is the celebrated Wigner function [49]. The Wigner function is a member of a class of functions called *quasi-probability distributions*: functions on the phase space $\mathbb{R}^2$ which are both positive and negative. Some examples of Wigner functions are given in Appendix B.

The Wigner function is the unique quasi-probability distribution satisfying the properties [5]

Wig(1) For all $\hat{\rho}$, $\rho^{\text{Wigner}}(q,p)$ is real.

Wig(2) For all $\hat{\rho}_1$ and $\hat{\rho}_2$,

$$\text{Tr}(\hat{\rho}_1\hat{\rho}_2) = 2\pi \int_{\mathbb{R}^2} d\xi d\eta \; {\rho_1}^{\text{Wigner}}(\xi,\eta){\rho_2}^{\text{Wigner}}(\xi,\eta).$$

Wig(3) For all $\hat{\rho}$, integrating ρ^{Wigner} along the line $aq + bp = c$ in phase space yields the probability that a measurement of the observable $a\hat{Q} + b\hat{P} = c$ has the result c.

Notice from Equation (3.2) that Wigner function is obtained from the set of operators in Equation (3.1) (for $f = 1$) via the trace. Thus the properties Wig(1)-(3) can be transformed into properties on a set of operators $\hat{F}(q,p)$ which uniquely specify the set in Equation (3.1) for $f = 1$. These properties are

Wig(4) $\hat{F}(q,p)$ is Hermitian.

Wig(5) $2\pi\text{Tr}(\hat{F}(q,p)\hat{F}(q',p')) = \delta(q-q')\delta(p-p')$.

Wig(6) Let $\hat{P}_c$ be the projector onto the eigenstate of $a\hat{Q} + b\hat{P} = c$ with eigenvalue c. Then,

$$\int_{\mathbb{R}^2} dqdp \; \delta(aq + bp - c) = \hat{P}_c.$$

The Wigner functions has many properties and applications [28] which are not of concern here. However, it is important to note that wide variety of fruitful applications of the Wigner function is responsible for the interest in its generalization. The properties Wig(1)-(6) were presented here as most authors have aimed at a finite dimensional analogy of the Wigner function defined such that it satisfies properties analogous to Wig(1)-(6) for discrete phase spaces. The remainder of the chapter is devoted to generalizing the definition of the Wigner function to finite dimensional quantum systems.

3.2 Wootters discrete phase space representation

In [50], Wootters is interested in obtaining a discrete analog of the Wigner function. Associated with each Hilbert space $\mathcal{H}$ of finite dimension d is a *discrete phase space*. First assume d is prime. The *prime phase space*, Φ_d, is a $d \times d$ array of points $\alpha = (q, p) \in \mathbb{Z}_d \times \mathbb{Z}_d$. The simplest example of a discrete phase space[1] for a qubit (the common name for a quantum system with $d = 2$) is shown in Figure 3.1.

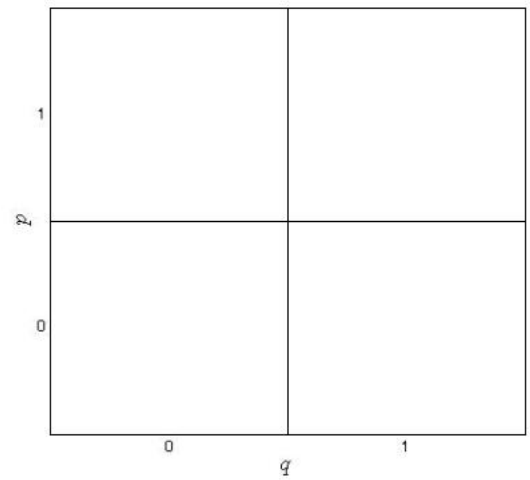

Figure 3.1: Discrete phase space of a qubit.

A *line*, λ, is the set of d points satisfying the linear equation $aq + bp = c$, where all arithmetic is modulo d. Two lines are *parallel* if their linear equations differ in the value of c. The prime phase space Φ_d contains $d + 1$ sets of d parallel lines called *striations*. The three sets of two parallel lines for the discrete phase space for $d = 2$ is depicted in Figure 3.2.

Assume the the Hilbert space $\mathcal{H}$ has composite dimension $d = d_1 d_2 \cdots d_k$. The discrete phase space of the entire d dimensional system is the Cartesian product of two-dimensional prime phase spaces of the subsystems. The phase space is thus a $d_1 \times d_1 \times d_2 \times d_2 \times \cdots \times d_k \times d_k$ array. Such as construction is formalized as follows. The *discrete phase space* is the multi-dimensional array $\Phi_d = \Phi_{d_1} \times \Phi_{d_2} \times \cdots \times \Phi_{d_k}$, where each Φ_{d_i} is a prime phase space. A *point* is the k-tuple $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k)$ of points $\alpha_i = (q_i, p_i)$ in the prime phase spaces. A *line* is the k-tuple $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$ of lines in the prime phase spaces. That is, a line is the set of d points satisfy the equation

$$(a_1 q_1 + b_1 p_1, a_2 q_2 + b_2 p_2, \ldots, a_k q_k + b_k p_k) = (c_1, c_2, \ldots, c_k),$$

[1] Although the phase space is defined to be a array of points, it is often easier to depict it as an array of boxes where each box represents a point.

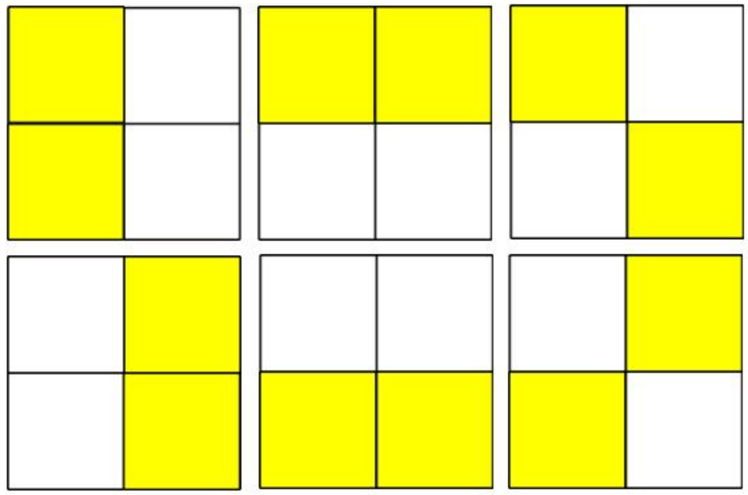

Figure 3.2: Complete set of lines in the qubit discrete phase space.

which is symbolically written $aq + bp = c$. Two lines are *parallel* if their equations differ in the value c. As was the case for the prime phase spaces, parallel lines can be partitioned into sets, again called striations; the discrete phase space Φ_d contains $(d_1 + 1)(d_2 + 1) \cdots (d_k + 1)$ sets of d parallel lines.

The construction of the discrete phase space has now been completed. To introduce Hilbert space into the discrete phase space formalism, Wootters chooses the following special basis for the space of Hermitian operators. The set of operators $\{\hat{A}_\alpha : \alpha \in \Phi_d\}$ acting on an d dimensional Hilbert space are called *phase point operators* if the operators satisfy

Woo(4) For each point α, $\hat{A}_\alpha$ is Hermitian.

Woo(5) For any two points α and β, $\mathrm{Tr}(\hat{A}_\alpha \hat{A}_\beta) = d\delta_{\alpha\beta}$.

Woo(6) For the lines λ in a given striation, the operators $\hat{P}_\lambda = \frac{1}{d} \sum_{\alpha \in \lambda} \hat{A}_\alpha$ form a projective valued measurement (PVM): a set of d orthogonal projectors which sum to identity.

Notice that these properties of the phase point operators Woo(4)-(6) are discrete analogs of the properties Wig(4)-(6) of the function $\hat{F}$ defining the original Wigner function. This definition suggests that the lines in the discrete phase space should be labeled with states of the Hilbert space. Since each striation is associated with a PVM, each of the d lines in a striation is labeled with an orthogonal state. For each Φ_d, there is a unique set of phase point operators up to unitary equivalence.

Although the sets of phase point operators are unitarily equivalent, the induced labeling of the lines associated to the chosen set of phase point operators are not equivalent. This is clear from the fact that unitarily equivalent PVMs do not project onto the same basis.

The choice of phase point operators in [50] will be adopted. For d prime, the phase point operators are

$$\hat{A}_\alpha = \frac{1}{d} \sum_{j,m=0}^{d-1} \omega^{pj-qm+\frac{jm}{2}} \hat{X}^j \hat{Z}^m, \tag{3.4}$$

where ω is a d'th root of unity and $\hat{X}$ and $\hat{Z}$ are the generalized Pauli operators (See Appendix A.1). For composite d, the phase point operator in Φ_d associated with the point $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k)$ is given by

$$\hat{A}_\alpha = \hat{A}_{\alpha_1} \otimes \hat{A}_{\alpha_2} \otimes \cdots \otimes \hat{A}_{\alpha_k}, \tag{3.5}$$

where each $\hat{A}_{\alpha_i}$ is the phase point operator of the point α_i in Φ_{d_i}.

The d^2 phase point operators are linearly independent and form a basis for the space of Hermitian operators acting on an d dimensional Hilbert space. Thus, any density operator $\hat{\rho}$ can be decomposed as

$$\hat{\rho} = \sum_{q,p} \rho^{\text{Wootters}}(q,p) \hat{A}(q,p),$$

where the real coefficients are explicitly given by

$$\rho^{\text{Wootters}}(q,p) = \frac{1}{d} \text{Tr}(\hat{\rho}\hat{A}(q,p)). \tag{3.6}$$

This discrete phase space function is the Wootters *discrete Wigner function*. This discrete quasi-probability function satisfies the following properties which are the discrete analogies of the properties Wig(1)-(3) the original continuous Wigner function satisfies.

Woo(1) For all $\hat{\rho}$, $\rho^{\text{Wootters}}(q,p)$ is real.

Woo(2) For all $\hat{\rho}_1$ and $\hat{\rho}_2$,

$$\text{Tr}(\hat{\rho}_1 \hat{\rho}_2) = d \sum_{q,p} {\rho_1}^{\text{Wootters}}(q,p) {\rho_2}^{\text{Wootters}}(q,p).$$

Woo(3) For all $\hat{\rho}$, sum ρ^{Wootters} along the line λ in phase space yields the probability that a measurement of the PVM associated with the striation which contains λ has the result associated with the outcome associated with λ.

3.3 Odd dimensional discrete Wigner functions

In [11], Cohendet *et al.* define a discrete analogue of the Wigner function which is valid for integer spin. That is, $\dim(\mathcal{H}) = d$ is assumed to be odd. Whereas Wootters builds up a discrete phase space before defining a Wigner function, the authors of [11] implicitly define a discrete phase space through the definition of their Wigner function.

The operators

$$\hat{W}_{mn}\phi_k = \omega^{2n(k-m)}\phi_{k-2m},$$

with $m, n \in \mathbb{Z}_d$ are the discrete analog of the Weyl operators. Then, the *discrete Wigner function* of a density operator $\hat{\rho}$ is

$$\rho^{\mathrm{odd}}(q,p) = \frac{1}{d}\mathrm{Tr}(\hat{\rho}\hat{W}_{qp}\hat{P}), \tag{3.7}$$

where $\hat{P}$ is the parity operator (see Appendix A.1).

The authors call the operators $\hat{\triangle}_{qp} = \hat{W}_{qp}\hat{P}$ *Fano operators* and note that they satisfy

$$\begin{aligned}
&\hat{\triangle}^{\dagger}_{qp} = \hat{\triangle}_{qp},\\
&\hat{\triangle}^{2}_{qp} = \hat{\mathbb{1}},\\
&\mathrm{Tr}(\hat{\triangle}_{qp}\hat{\triangle}_{q'p'}) = d\delta_{qq'}\delta_{pp'},\\
&\hat{W}^{\dagger}_{xk}\hat{\triangle}_{qp}\hat{W}_{xk} = \hat{\triangle}_{q-2x\;p-2k}.
\end{aligned}$$

The Fano operators play a role similar to Wootters' phase point operators; they form a complete basis of the space of Hermitian operators. The phase space implicitly defined through the definition of the discrete Wigner function (3.7) is $\mathbb{Z}_d \times \mathbb{Z}_d$. When d is an odd prime, this phase space is equivalent to Wootters discrete phase space. In this case the Fano operators are $\hat{\triangle}_{qp} = \hat{A}_{(-q,p)}$. This can seen by writing the Wootters phase point operators as

$$\hat{A}_{(q,p)} = \frac{1}{d}\hat{X}^{2q}\hat{Z}^{2p}\hat{P}\omega^{2qp}.$$

3.4 Even dimensional discrete Wigner functions

In [35], Leonhardt defines discrete analogues of the Wigner function for both odd and even dimensional Hilbert spaces. In a later paper [36], Leonhardt discusses the need for separate definitions for the odd and even dimension cases. Naively applying his definition, or that of Cohendet *et al.*, of the discrete Wigner function for odd dimensions to even dimensions yields unsatisfactory results. The reason for this is the discrete Wigner function carries redundant information for even dimensions which is insufficient to specify the state uniquely. The solution is to enlarge the phase space until the redundant information becomes sufficient to specify the state uniquely.

Suppose $\dim(\mathcal{H}) = d$ is odd. Leonhardt defines the discrete Wigner function as

$$\rho^{\text{Leonhardt}}(q,p) = \frac{1}{d}\text{Tr}(\hat{\rho}\hat{X}^{2q}\hat{Z}^{2p}\hat{P}\omega^{2qp}).$$

Leonhardt's definition of an odd dimensional discrete Wigner function is unitarily equivalent to the Cohendet *et al.* definition $\rho^{\text{Leonhardt}}(q,p) = \rho^{\text{odd}}(-q,p)$. To define a discrete Wigner function for even dimensions, Leonhardt takes half-integer values of q and p. This amounts to enlarging the phase space to $\mathbb{Z}_{2d} \times \mathbb{Z}_{2d}$. Thus the *even dimensional* discrete Wigner function is

$$\rho^{\text{even}}(q,p) = \frac{1}{2d}\text{Tr}(\hat{\rho}\hat{X}^{q}\hat{Z}^{p}\hat{P}\omega^{\frac{qp}{2}}),$$

where the operators

$$\triangle^{\text{even}}_{qp} = \frac{1}{2d}\hat{X}^{q}\hat{Z}^{p}\hat{P}\omega^{\frac{qp}{2}}$$

could be called the even dimensional Fano or phase point operators. Of course, these operators do not satisfy all the criteria which the Fano operators (in the case of Cohendet *et al.*) or the phase point operators (in the case of Wootters) satisfy; they are not orthogonal. Moreover, they are not even linearly independent which can easily be inferred since there are $4d^2$ of them and a set of linearly independent operators contains a maximum of d^2 operators.

3.5 Wigner functions on the sphere

In [27], Heiss and Weigert are concerned with a set of postulates put forth by Stratonovich [46]. The aim of Stratonovich was to find a Wigner function type mapping, analogous to that of a infinite dimensional system on $\mathbb{R}^2$, of a finite dimensional system (of dimension d of course) on the sphere $\mathbb{S}^2$. The first postulate is linearity and is always satisfied if the Wigner functions on the sphere satisfy

$$\rho^{\text{sphere}}(\mathbf{n}) = \text{Tr}(\hat{\rho}\hat{\triangle}(\mathbf{n})), \tag{3.8}$$

where $\mathbf{n}$ is a point on $\mathbb{S}^2$. The remaining postulates on this quasi-probability mapping are

$$\begin{aligned}
&\rho^{\text{sphere}}(\mathbf{n})^* = \rho^{\text{sphere}}(\mathbf{n}),\\
&\frac{d}{4\pi}\int_{\mathbb{S}^2} d\mathbf{n}\ \rho^{\text{sphere}}(\mathbf{n}) = 1,\\
&\frac{d}{4\pi}\int_{\mathbb{S}^2} d\mathbf{n}\ \rho_1^{\text{sphere}}(\mathbf{n})\rho_2^{\text{sphere}}(\mathbf{n}) = \text{Tr}(\hat{\rho}_1\hat{\rho}_2),\\
&(g\cdot\rho)^{\text{sphere}}(\mathbf{n}) = \rho^{\text{sphere}}(\mathbf{n})^g,\ g \in \text{SU}(2),
\end{aligned}$$

where $g \cdot \rho$ is the image of $\hat{U}_g \hat{\rho} \hat{U}_g^\dagger$ and $\hat{U} : \mathrm{SU}(2) \to \mathrm{U}(\mathcal{H})$ is an irreducible unitary representation of the group SU(2).

The continuous set of operators $\hat{\triangle}(\mathbf{n})$ is called a *kernel* and plays the role of the phase point and Fano operators of the previous sections. Requiring that Equation (3.8) hold changes the postulates to new conditions on the kernel

$$\hat{\triangle}(\mathbf{n})^\dagger = \hat{\triangle}(\mathbf{n}), \tag{3.9}$$

$$\frac{d}{4\pi} \int_{\mathbb{S}^2} d\mathbf{n} \, \hat{\triangle}(\mathbf{n}) = \hat{\mathbb{1}}, \tag{3.10}$$

$$\frac{d}{4\pi} \int_{\mathbb{S}^2} d\mathbf{n} \, \mathrm{Tr}(\hat{\triangle}(\mathbf{n})\hat{\triangle}(\mathbf{m}))\hat{\triangle}(\mathbf{n}) = \hat{\triangle}(\mathbf{m}), \tag{3.11}$$

$$\hat{\triangle}(g \cdot \mathbf{n}) = \hat{U}_g \hat{\triangle}(\mathbf{n}) \hat{U}_g^\dagger, \; g \in \mathrm{SU}(2). \tag{3.12}$$

Heiss and Weigert provide a derivation of 2^{2s}, where $s = \frac{d-1}{2}$ is the *spin*, unique kernels satisfying these postulates. They are

$$\hat{\triangle}(\mathbf{n}) = \sum_{m=-s}^{s} \sum_{l=0}^{2s} \epsilon_l \frac{2l+1}{2s+1} C^{s\,l\,s}_{m\,0\,m} \phi_m(\mathbf{n}) \phi_m^*(\mathbf{n}), \tag{3.13}$$

where C denotes the so-called *Clebsch-Gordon coefficients*; $\phi_m(\mathbf{n})$ are the eigenvectors of the operator $\hat{\mathbf{S}} \cdot \mathbf{n}$, where $\hat{\mathbf{S}} = (\hat{X}, \hat{Y}, \hat{Z})$; and $\epsilon_l = \pm 1$, for $l = 1 \ldots 2s$ and $\epsilon_0 = 1$.

Heiss and Weigert relax the postulates Equations (3.9)-(3.12) on the kernel $\hat{\triangle}(\mathbf{n})$ to allow for a pair of kernels $\hat{\triangle}^{\mathbf{n}}$ and $\hat{\triangle}_{\mathbf{m}}$. The pair individually satisfy Equation (3.9), while one of them satisfies Equation (3.10) and the other Equation (3.12). Together, the pair must satisfy the generalization of Equation (3.11)

$$\frac{d}{4\pi} \int_{\mathbb{S}^2} d\mathbf{n} \mathrm{Tr}(\hat{\triangle}^{\mathbf{n}} \hat{\triangle}_{\mathbf{m}}) \hat{\triangle}^{\mathbf{n}} = \hat{\triangle}_{\mathbf{m}}. \tag{3.14}$$

A pair of kernels, together satisfying Equation (3.14), is given by

$$\hat{\triangle}_{\mathbf{n}} = \sum_{m=-s}^{s} \sum_{l=0}^{2s} \gamma_l \frac{2l+1}{2s+1} C^{s\,l\,s}_{m\,0\,m} \phi_m(\mathbf{n}) \phi_m^*(\mathbf{n}),$$

$$\hat{\triangle}^{\mathbf{n}} = \sum_{m=-s}^{s} \sum_{l=0}^{2s} \gamma_l^{-1} \frac{2l+1}{2s+1} C^{s\,l\,s}_{m\,0\,m} \phi_m(\mathbf{n}) \phi_m^*(\mathbf{n}),$$

where $\gamma_l = \pm 1$ for $l = 1 \ldots 2s$ and $\gamma_0 = 1$. The original postulates are satisfied when $\gamma_l = \gamma_l^{-1} \equiv \epsilon_l$.

The major contribution of [27] is the derivation of a *discrete* kernel $\hat{\triangle}_\nu := \hat{\triangle}_{\mathbf{n}_\nu}$, for $\nu = 1 \dots d^2$ which satisfies the discretized postulates

$$\hat{\triangle}_\nu^\dagger = \hat{\triangle}_\nu, \tag{3.15}$$

$$\frac{1}{d}\sum_{\nu=1}^{d^2} \hat{\triangle}^\nu = \hat{\mathbb{1}}, \tag{3.16}$$

$$\frac{1}{d}\sum_{\nu=1}^{d^2} \mathrm{Tr}(\hat{\triangle}_\nu \hat{\triangle}^\mu)\hat{\triangle}_\nu = \hat{\triangle}^\mu, \tag{3.17}$$

$$\hat{\triangle}_{g\cdot\nu} = \hat{U}_g \hat{\triangle}_\nu \hat{U}_g^\dagger,\ g \in \mathrm{SU}(2). \tag{3.18}$$

The subset of points $\mathbf{n}_\nu$ is called a *constellation*. The linearity postulate is not explicitly stated since it is always satisfied under the assumption

$$\hat{\rho} \to \rho^{\text{constellation}}(\nu) = \mathrm{Tr}(\hat{\rho}\hat{\triangle}_\nu). \tag{3.19}$$

Equation (3.17) is called a *duality* condition. That is, it is only satisfied if $\hat{\triangle}_\nu$ and $\hat{\triangle}^\mu$ are *dual bases* for Herm($\mathcal{H}$). In particular,

$$\frac{1}{d}\mathrm{Tr}(\hat{\triangle}_\nu \hat{\triangle}^\mu) = \delta_{\nu\mu}.$$

Although the explicit construction of a pair of discrete kernels satisfying Equations (3.15)-(3.18) might be computationally hard, their existence is a trivial exercise in linear algebra. Indeed, so long as $\hat{\triangle}_\nu$ is a basis for Herm($\mathcal{H}$), its dual, $\hat{\triangle}^\mu$, is uniquely determined by

$$\hat{\triangle}^\mu = \sum_{\nu=1}^{d^2} \mathtt{G}_{\nu\mu}^{-1}\hat{\triangle}_\nu,$$

where the Gram matrix $\mathtt{G}$ is given by

$$\mathtt{G}_{\nu\mu} = \mathrm{Tr}(\hat{\triangle}_\nu \hat{\triangle}_\mu).$$

The authors of [27] note that almost any constellation leads to a discrete kernel $\hat{\triangle}_\nu$ forming a basis for Herm($\mathcal{H}$). The term *almost any* here means that a randomly selected discrete kernel will form, with probability 1, a basis for Herm($\mathcal{H}$).

3.6 Finite fields discrete phase space representation

Recall that when $\dim(\mathcal{H}) = d$ is prime, Wootters defines the discrete phase space as a $d \times d$ lattice indexed by the group $\mathbb{Z}_d$. In [51], Wootters generalizes his original construction of a discrete phase space to allow the $d \times d$ lattice to be indexed by a finite field $\mathbb{F}_d$ which exists only when $d = p^n$: an integer power of a prime number. This approach is discussed at length in the paper [20] authored by Gibbons, Hoffman and Wootters (GHW).

Similar to his earlier approach, Wootters defines the *phase space*, Φ_d, as a $d \times d$ array of points $\alpha = (q,p) \in \mathbb{F}_d \times \mathbb{F}_d$. A *line*, λ, is the set of d points satisfying the linear equation $aq + bp = c$, where all arithmetic is done in $\mathbb{F}_d$. Two lines are *parallel* if their linear equations differ in the value of c.

The mathematical structure of $\mathbb{F}_d$ is appealing because lines defined as above have the following useful properties: (i) given any two points, exactly one line contains both points, (ii) given a point α and a line λ not containing α, there is exactly one line parallel to λ that contains α, and (iii) two nonparallel lines intersect at exactly one point. Note that these are usual properties of lines in Euclidean space. As before, the d^2 points of the phase space Φ_d can be partitioned into $d+1$ sets of d parallel lines called *striations*. The line containing the point (q,p) and the origin $(0,0)$ is called a *ray* and consists of the points (sq, sp), where s is a parameter taking values in $\mathbb{F}_d$. We choose each ray, specified by the equation $aq + bp = 0$, to be the representative of the striation it belongs to.

A translation in phase space, $\mathcal{T}_{\alpha_0}$, adds a constant vector, $\alpha_0 = (q_0, p_0)$, to every phase space point: $\mathcal{T}_{\alpha_0}\alpha = \alpha + \alpha_0$. Each line, λ, in a striation is invariant under a translation by any point contained in its ray, parameterized by the points (sq, sp). That is,

$$\mathcal{T}_{(sq,sp)}\lambda = \lambda. \tag{3.20}$$

The discrete Wigner function is

$$\rho^{\text{field}}(q,p) = \frac{1}{d}\text{Tr}(\hat{\rho}\hat{A}_{(q,p)}),$$

where now the Hermitian *phase point operators* satisfy the following properties for a projector valued function $\hat{Q}$, called a *quantum net*, to be defined later.

GHW(4) For each point α, $\hat{A}$ is Hermitian.

GHW(5) For any two points α and β, $\text{Tr}(\hat{A}_\alpha \hat{A}_\beta) = d\delta_{\alpha\beta}$.

GHW(6) For any line λ, $\sum_{\alpha\in\lambda} \hat{A}_\alpha = d\hat{Q}(\lambda)$.

The projector valued function $\hat{Q}$ assigns quantum states to lines in phase space. This mapping is required to satisfy the special property of *translational covariance*, which is defined after a short, but necessary, mathematical digression. Notice first that properties GHW(4) and GHW(5) are identical to Woo(4) and Woo(4). Also note that if GHW(6) is to be analogous to Woo(6), the property of translation covariance must be such that the set $\{\hat{Q}(\lambda)\}$ when λ ranges over a striation forms a PVM.

The set of elements $E = \{e_0, ..., e_{n-1}\} \subset \mathbb{F}_d$ is called a *field basis* for $\mathbb{F}_d$ if any element, x, in $\mathbb{F}_d$ can be written

$$x = \sum_{i=0}^{n-1} x_i e_i, \tag{3.21}$$

where each x_i is an element of the prime field $\mathbb{Z}_p$. The *field trace*[2] of any field element is given by

$$\mathrm{tr}(x) = \sum_{i=0}^{n-1} x^{p^i}. \tag{3.22}$$

There exists a unique field basis, $\tilde{E} = \{\tilde{e}_0, ...\tilde{e}_{n-1}\}$, such that $\mathrm{tr}(\tilde{e}_i e_j) = \delta_{ij}$. We call $\tilde{E}$ the *dual* of E.

The construction presented in [20] is physically significant for a system of n objects (called *particles*) having a p dimensional Hilbert space. A translation operator, $\hat{T}_\alpha$ associated with a point in phase space $\alpha = (q, p)$ must act independently on each particle in order to preserve the tensor product structure of the composite system's Hilbert space. We expand each component of the point α into its field basis decomposition as in Equation (3.21)

$$q = \sum_{i=0}^{n-1} q_i e_i \tag{3.23}$$

and

$$p = \sum_{i=0}^{n-1} p_i f \tilde{e}_i, \tag{3.24}$$

with f any element of $\mathbb{F}_d$. Note that the basis we choose for p is a multiple of the dual of that chosen for q. Now, the translation operator associated with the point (q, p) is

$$\hat{T}_{(q,p)} = \bigotimes_{i=0}^{n-1} \hat{X}^{q_i} \hat{Z}^{p_i}, \tag{3.25}$$

Since $\hat{X}$ and $\hat{Z}$ are unitary, $\hat{T}_\alpha$ is unitary.

We assign with each line in phase space a pure quantum state. The quantum net $\hat{Q}$ is defined such that for each line, λ, $\hat{Q}(\lambda)$ is the operator which projects onto the pure state associated with λ. As a consequence of the choice of basis for p in Equation (3.24), the state assigned to the line $\tau_\alpha \lambda$ is obtained through

$$\hat{Q}(\tau_\alpha \lambda) = \hat{T}_\alpha \hat{Q}(\lambda) \hat{T}_\alpha^\dagger. \tag{3.26}$$

This is the condition of translational covariance and it implies that each striation is associated with an orthonormal basis of the Hilbert space. To see this, recall the property in Equation (3.20). From Equation (3.26), this implies that, for each $s \in \mathbb{F}_d$, $\hat{T}_{(sq,sp)}$ must commute with $\hat{Q}(\lambda)$, where the line λ is any line in the striation defined by the ray consisting of the points (sq, sp). That is, the states associated to the lines of the striation must be common eigenstates of the unitary translation operator $\hat{T}_{(sq,sp)}$, for each $s \in \mathbb{F}_d$. Thus, the states are orthogonal and form a basis

[2] Note that we will distinguish the field trace, $\mathrm{tr}(\cdot)$, from the usual trace of a Hilbert space operator, $\mathrm{Tr}(\cdot)$, by the case of the first letter.

for the Hilbert space. That is, there projectors form a PVM which makes GHW(6) identical to Woo(6) when d is prime.

In [20], the author's note that, although the association between states and vertical and horizontal lines is fixed, the quantum net is not unique. In fact, there are d^{d-1} quantum nets which satisfy Equation (3.20). When d is prime, one of these quantum nets corresponds exactly to the original discrete Wigner function defined by Wootters in Section 3.2.

3.7 Summary of existing quasi-probability functions

This section summarizes the phase space functions reviewed in Sections 3.2-3.6 which form only a subset of the literature on finite dimensional phase space functions. There are indeed several others (for a recent review see [47]). More generally, there exist what will be called *quasi-probability representations*, which are real-valued representations that do not necessarily reflect any preconceived classical phase space structure. For example in [25] Hardy shows that five axioms are sufficient to imply a special quasi-probabilistic representation which is equivalent to an operational form of quantum theory. In [26] Havel also proposes an kind of analog of the Wigner function called the "real density matrix".

A concise summary of all quasi-probability representations for finite dimensional quantum systems reviewed here (and a couple more) is presented in Table 3.1. The table gives the first author, reference and year of the publication. It also shows the phase space structure and mathematical field which indexes it (if applicable). The second to last column indicates whether or not the representation contains redundant information. The last column reveals the scope of quantum theory the paper aims to cover (notice that typically only states are considered).

Table 3.1: Finite quasi-probability representations

Author(s)	Year	Valid dimensions	Phase space	Index field	Redundancy	Quantum theory scope
Stratonovich [46]	1957	any	sphere	polar coordinates	continuous	states
Wootters [50]	1987	prime*	$d \times d$ lattice	$\mathbb{Z}_d$	no	standard
Cohendet *et al* [11]	1987	odd	$d \times d$ lattice	$\mathbb{Z}_d$	no	states
Leonhardt [35]	1995	even**	$2d \times 2d$ lattice	$\mathbb{Z}_{2d}$	four-fold	states
Heiss and Weigert [27]	2000	any	sphere***	arbitrary	no	states
Hardy [25]	2001	any	none	n/a	no	operational
Havel [26]	2003	any	none	n/a	no	states
Gibbons *et al* [20]	2004	power of prime	$d \times d$ lattice	$\mathbb{F}_d$	no	states
Chaturvedi *et al* [9]	2006	any	$d \times d$ lattice	$\mathbb{Z}_d$	no	states
Gross [23]	2006	odd	$d \times d$ lattice	$\mathbb{Z}_d$	no	states

Notes: *Wootters' original discrete Wigner function [50] is usually understood to be valid for prime dimension but, as discussed in Section 3.2, is easily extended to any dimension by combining prime dimensional phase spaces. **Leonhardt [35] also defines a discrete Wigner function valid for odd dimensional which is equivalent to the other odd dimensional cases [11]. ***The phase space of Heiss and Weigert is any subset of d points on the sphere which can indexed arbitrarily.

3.8 Unification of existing quasi-probability functions

Each of the quasi-probability functions discussed above are *linear* representations of the density operator. These representations are also *invertible* as the density operator can be obtained from any quasi-probability function. To ensure the finiteness of the formalism, each quasi-probability function is a member of the function space $L^2(\mathcal{S}, \mu)$, where $\mathcal{S}$ represents a classical state space (i.e. the phase space, where applicable). These three properties will constitute the following minimal definition of a quasi-probability representation of *quantum states alone*.

Definition 3.8.1. *A* quasi-probability representation of quantum states *is any map* $\mathrm{Herm}(\mathcal{H}) \to L^2(\mathcal{S}, \mu)$ *that is linear and invertible.*

One might object that the restrictions imposed on this map are too strong. Indeed there is no mention of linearity, invertibility, or L^2 spaces in the definition of classical probability. Recall however from Definition 2.1.1 that a classical probabilistic description describes an entire experimental arrangement. It could be argued that classical intuition is lost and mathematical descriptions become arbitrary if one begins to consider individual experimental procedures without specification of the entire scope of the experiment. How is one to describe a coin if it is not known whether it will be tossed or not? Definition 3.8.1 was chosen since all the known quasi-probability representations of states satisfy it. Moreover, later in Chapter 6 a definition of a quasi-probability representation of an entire quantum experiment (i.e. states and measurements) will be given that follows more closely a classical intuition *and* happens to reduce to Definition 2.1.1 on the part of the definition which represents states.

Given Definition 3.8.1, any phase space function is then a particular type of quasi-probability representation.

Definition 3.8.2. *If there exists symmetry group on* Γ*,* G*, carrying a unitary representation* $\hat{U} : G \to \mathcal{U}(\mathcal{H})$ *and a quasi-probability representation satisfying the covariance property* $\hat{U}_g \hat{A} \hat{U}_g^\dagger \mapsto \{A(g(\alpha))\}_{\alpha\in\Gamma}$ *for all* $\hat{A} \in \mathrm{Herm}(\mathcal{H})$ *and* $g \in G$*, then* $\hat{A} \mapsto A(\alpha)$ *is a* phase space representation of quantum states.

All phase space functions in the literature correspond to quasi-probability representations that satisfy this additional covariance condition.

Table 3.1 shows that the range of validity in the Hilbert space dimension of these functions are often disjoint. Moreover, the construction of the phase spaces use varying mathematical structures: integers, finite fields and points on a sphere. Coupled with the fact that at least two known representations required redundancy, it may seem at first that Definition 3.8.1 is as far as one can can go in unifying the quasi-probability functions. All is not lost however; in the next chapter it is shown that the mathematical theory of *frames* is both sufficient *and necessary* to describe any representation of quantum states satisfying Definition 3.8.1.

If one accepts negative probability, then these quasi-probability representations constitute a hidden variable theory. The variables here are "hidden" since negative probabilities have never been observed. Later, in Chapter 7, it will be shown that negativity is a necessary feature of quasi-probability representations.

Chapter 4

Frame representations of quantum states

This chapter first introduces the idea of a *frame* in the context of signal analysis, the branch of mathematics in which frames were discovered. Then, a precise definition of a frame is given in the context of quantum theory. Finally, it is shown that a *frame representation* (Definition 4.2.3) and a quasi-probability representation of quantum states (Definition 3.8.1) are equivalent.

4.1 Gentle introduction to frames

Frames are mathematical objects invented within the signal analysis community to rigorously deal with a common conceptual tool used in everyday life: redundancy. We all use redundancy; we repeat ourselves and double-check things to avoid errors. The same idea of reducing error is used in signal analysis [33]. A simple example will introduce the concept of a frame.

Consider the two dimensional Hilbert space $\mathcal{H}$. The canonical basis is $\{\phi_0, \phi_1\}$. The simplest frame is the Mercedes-Benz frame $\{\psi_k\}_{k=0}^{2} := \left\{\phi_0, -\frac{1}{2}\phi_0 + \frac{\sqrt{3}}{2}\phi_1, -\frac{1}{2}\phi_0 - \frac{\sqrt{3}}{2}\phi_1\right\}$. This is depicted in in Figure 4.1.

Any vector $\xi \in \mathcal{H}$ can be decomposed in the canonical basis as $\xi = \langle\xi, \phi_0\rangle\phi_0 + \langle\xi, \phi_1\rangle\phi_1$. In communications applications the coefficients $\alpha := \langle\xi, \phi_0\rangle$ and $\beta := \langle\xi, \phi_1\rangle$ are transmitted between parties. These coefficients will typically be subject to noise and what is received are $\tilde{\alpha}$ and $\tilde{\beta}$. The receiver can only reconstruct an estimate $\tilde{\xi} = \tilde{\alpha}\phi_0 + \tilde{\beta}\phi_1$ from the noisy coefficients. If the noise is white, with variation σ^2, the mean squared error (the average value of $\|\xi - \tilde{\xi}\|^2$) is also σ^2.

A short calculation will verify that the vector ξ can be decomposed in the Mercedes-Benz frame as $\xi = \frac{2}{3}\sum_{k=0}^{2}\langle\xi, \psi_k\rangle\psi_k$. Now, the three coefficients $\xi_k := \langle\xi, \psi_k\rangle$

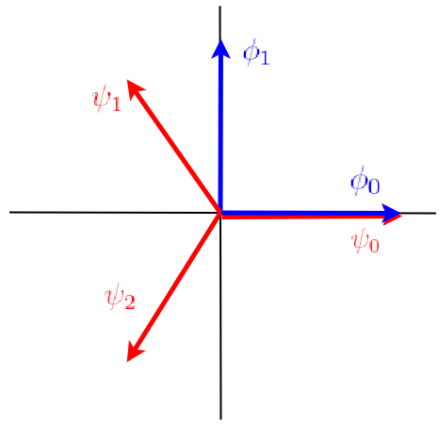

Figure 4.1: Mercedes-Benz frame $\{\psi_k\}$.

are transmitted, subjected to the same white noise and received as $\tilde{\xi}_k$. Again the receiver reconstructs the estimate $\tilde{\xi} = \frac{2}{3}\sum_{k=0}^{2} \tilde{\xi}_k \psi_k$ from the noisy coefficients. Now the mean square error can be verified to be $\frac{2}{3}\sigma^2$ which is lower than that for the same protocol using the basis coefficients.

As a second example, consider the case when one coefficient is lost in transmission. When the coefficients are taken from a basis, it is impossible to reconstruct the signal perfectly; the possible reconstructed vectors can never span the entire Hilbert space. However, if one of the Mercedes-Benz coefficients are lost, the remaining two can still be use to reconstruct the vector perfectly; any two elements of the frame still span $\mathcal{H}$.

The above two situations illustrate the utility of frames. This thesis is not so concerned with the communication applications of frames but it will be shown how various structures in quantum theory are equivalent to frames. In the next section a precise definition of a frame for a particular Hilbert space important in quantum theory will be presented.

4.2 Frame representations of quantum states

A *frame* can be thought of as a generalization of an orthonormal basis. However, the particular Hilbert space under consideration here is not $\mathcal{H}$. Considered here is a generalization of a basis for $\mathrm{Herm}(\mathcal{H})$, which is the set of Hermitian operators on an complex Hilbert space of dimension d. With the trace inner product (or Hilbert-Schmidt inner product) $\langle \hat{A}, \hat{B} \rangle := \mathrm{Tr}(\hat{A}\hat{B})$, $\mathrm{Herm}(\mathcal{H})$ forms a Hilbert space itself of dimension d^2. Let $\mathcal{S}$ be some set with positive measure μ.

Definition 4.2.1. *A* frame *for* $\mathrm{Herm}(\mathcal{H})$ *is a mapping* $\hat{F} : \mathcal{S} \to \mathrm{Herm}(\mathcal{H})$ *which satisfies*

$$a\|\hat{A}\|^2 \leq \int_{\mathcal{S}} d\mu(s)\ |\langle \hat{F}(s), \hat{A} \rangle|^2 \leq b\|\hat{A}\|^2, \tag{4.1}$$

for all $\hat{A} \in \mathrm{Herm}(\mathcal{H})$ *and some constants* $a, b > 0$.

The definition of a finite frame is of course subsumed by Definition 4.2.1. However, since the majority of the phase space representations are discrete, the definition of a finite frame will be given explicitly. When $|\mathcal{S}| < \infty$ and μ is the counting measure, a (finite) frame for $\mathrm{Herm}(\mathcal{H})$ is a set of operators $\mathcal{F} := \{\hat{F}(s) : s \in \mathcal{S}\} \subset \mathrm{Herm}(\mathcal{H})$ which satisfies

$$a\|\hat{A}\|^2 \leq \sum_{s\in\mathcal{S}} |\langle \hat{F}(s), \hat{A}\rangle|^2 \leq b\|\hat{A}\|^2, \tag{4.2}$$

for all $\hat{A} \in \mathrm{Herm}(\mathcal{H})$ and some constants $a, b > 0$. This definition generalizes a defining condition for an orthogonal basis $\{\hat{B}_k\}_{k=1}^{d^2}$

$$\sum_{k=1}^{d^2} |\langle \hat{B}_k, \hat{A}\rangle|^2 = \|\hat{A}\|^2, \tag{4.3}$$

for all $\hat{A} \in \mathrm{Herm}(\mathcal{H})$.

Definition 4.2.2. *A frame* $\hat{E} : \mathcal{S} \to \mathrm{Herm}(\mathcal{H})$ *which satisfies*

$$\hat{A} = \int_{\mathcal{S}} d\mu(s)\ \langle \hat{F}(s), \hat{A}\rangle \hat{E}(s), \tag{4.4}$$

for all $\hat{A} \in \mathrm{Herm}(\mathcal{H})$*, is a* dual frame *(to* $\hat{F}$*).*

The *frame operator* associated with the frame $\hat{F}$ is defined as

$$\tilde{S}(\hat{A}) := \int_{\mathcal{S}} d\mu(s)\ \hat{F}(s)\langle \hat{F}(s), \hat{A}\rangle.$$

If the frame operator satisfies $\tilde{S} = a\tilde{\mathbb{1}}$, the frame is called *tight*. The frame operator is invertible and thus every operator has a representation

$$\hat{A} = \tilde{S}^{-1}\tilde{S}\hat{A} = \int_{\mathcal{S}} d\mu(s)\ \langle \hat{F}(s), \hat{A}\rangle \tilde{S}^{-1}\hat{F}(s). \tag{4.5}$$

The map $\tilde{S}^{-1}\hat{F}$ is called the *canonical dual frame*. When $\mathcal{S}$ is finite and $|\mathcal{S}| = d^2$, the canonical dual frame is the unique dual, otherwise there are infinitely many choices for a dual.

A tight frame is ideal from the perspective that its canonical dual is proportional to the frame itself. Hence, the reconstruction is given by the convenient formula

$$\hat{A} = \tilde{S}^{-1}\tilde{S}\hat{A} = \frac{1}{a}\int_{\mathcal{S}} d\mu(s)\ \langle \hat{F}(s), \hat{A}\rangle \hat{F}(s).$$

The utility of this formula is emphasized when the frame is finite. It that case it becomes

$$\hat{A} = \frac{1}{a} \sum_{s \in \mathcal{S}} \langle \hat{F}(s), \hat{A} \rangle \hat{F}(s)$$

which is to be compared with

$$\hat{A} = \sum_{k=1}^{d^2} \langle \hat{B}_k, \hat{A} \rangle \hat{B}_k$$

which defines $\{\hat{B}_k\}_{k=1}^{d^2}$ as an orthonormal basis.

The mapping $\hat{A} \mapsto \langle \hat{F}(s), \hat{A} \rangle$ is usually called the *analysis operation* in the frame literature as it encodes the signal in terms of the frame. Here the notion of a signal not appropriate and a more suggestive name has been chosen and formalized in the following definition.

Definition 4.2.3. *A mapping* $\mathrm{Herm}(\mathcal{H}) \to L^2(\mathcal{S}, \mu)$ *of the form*

$$\hat{A} \mapsto A(s) := \langle \hat{F}(s), \hat{A} \rangle, \tag{4.6}$$

where $\hat{F}$ *is a frame, is a* frame representation *of* $\mathrm{Herm}(\mathcal{H})$*.*

4.3 Equivalence of the quasi-probability and frame representation of quantum states

Since each frame has at least a canonical dual, a frame representation (Definition 4.2.3) can always be inverted according to the reconstruction formula in Equation (4.5). A frame representation is defined such that it exists in $L^2(\mathcal{S}, \mu)$. It is clear that a frame representation is linear by virtue of the linearity of the inner product. Thus each frame representation is guaranteed to be a quasi-probability of quantum states (Definition 3.8.1) where $\mathcal{S}$ is interpreted as a classical state space (and possibly also a phase space). However, it is not clear that the converse is true, which is that every quasi-probability representation of quantum states is a frame representation. Indeed, the following lemma establishes the equivalence between frame representations and quasi-probability representations of quantum states.

Lemma 4.3.1. *A mapping* $R : \mathrm{Herm}(\mathcal{H}) \to L^2(\mathcal{S}, \mu)$ *is quasi-probability representation of quantum states (Definition 3.8.1) if and only if it is a frame representation for some unique frame* $\hat{F}$*.*

The proof of this lemma appears in [16]. For completeness it is reproduced here.

Proof. It is clear that a frame representation is a quasi-probability representation. Suppose a mapping $R : \mathrm{Herm}(\mathcal{H}) \to L^2(\mathcal{S}, \mu)$ is a quasi-probability representation. i.e. it is linear and invertible. The Riesz representation theorem implies that $R(\hat{A})(s) := \langle \hat{F}(s), \hat{A} \rangle$ for some unique mapping $\hat{F} : \mathcal{S} \to \mathrm{Herm}(\mathcal{H})$. Since $\mathrm{Herm}(\mathcal{H})$ is finite dimensional, the inverse R^{-1} is bounded. Thus R is bounded below by the bounded inverse theorem. That is, there exists a constant $a > 0$ such that

$$a\|\hat{A}\|^2 \leq \int_{\mathcal{S}} d\mu(s)\, |\langle \hat{F}(s), \hat{A} \rangle|^2.$$

Since $\langle \hat{F}(s), \hat{A} \rangle \in L^2(\mathcal{S})$, there exists a constant $b > 0$ such that

$$\int_{\mathcal{S}} d\mu(s)\, |\langle \hat{F}(s), \hat{A} \rangle|^2 \leq b\|\hat{A}\|^2.$$

Hence $\hat{F}$ is a frame. □

Thus there is a unique frame which defines each of the quasi-probability functions reviewed in Chapter 3. In the cases where the representation of the density operator is not redundant, the frame is just a basis. In the redundant cases, Leonhardt's even dimensional representation (Section 3.4) for example, the formalism of frame theory is necessary as a basis will not suffice. In the next chapter, the examples of quasi-probability functions will be analyzed using the frame formalism presented in this Chapter.

Chapter 5

Examples of frame representations of quantum states

Recall the examples of Chapter 3. Here it is demonstrated how each is a frame representation by identifying the frame which gives rise to each. After reading this chapter and going over the examples presented in Chapter 3 a second time, it becomes clear that the frame formalism presented in Chapter 4 provides a remarkably powerful tool in the unification of the known quasi-probability functions.

5.1 Wigner phase space representation

Let d be a prime number. Here $\mathcal{S} = \mathbb{Z}_d \times \mathbb{Z}_d$. Consider the frame $\mathcal{F}^{\text{Wootters}} = \{\hat{F}^{\text{Wootters}}(q,p) : (q,p) \in \mathcal{S}\}$, where

$$\hat{F}^{\text{Wootters}}(q,p) = \frac{1}{d^2}\hat{X}^{2q}\hat{Z}^{2p}\hat{P}\omega^{2qp}.$$

The quasi-probability function ρ^{Wootters} is a frame representation given by the frame $\mathcal{F}^{\text{Wootters}}$. The frame operator of $\mathcal{F}^{\text{Wootters}}$ is $\tilde{S} = d^{-1}\tilde{\mathbb{1}}$. The unique dual frame of $\mathcal{F}^{\text{Wootters}}$ is given by $\tilde{S}^{-1}\mathcal{F}^{\text{Wootters}}$, where here $\tilde{S}^{-1} = d\tilde{\mathbb{1}}$. Comparing this result to Equation (3.4), the dual frame to $\mathcal{F}^{\text{Wootters}}$ is a set of phase point operators.

Consider the group of translations on $\mathcal{S}$ with unitary representation $\hat{T}_{(q,p)} = \hat{X}^q\hat{Z}^p$. Then,

$$\begin{aligned}
\hat{T}_{(q,p)}\hat{F}^{\text{Wootters}}(q',p')\hat{T}^{\dagger}_{(q,p)} &= \frac{1}{d^2}\hat{X}^q\hat{Z}^p\hat{X}^{2q'}\hat{Z}^{2p'}\hat{P}\hat{Z}^{-p}\hat{X}^{-q}\omega^{2q'p'} \\
&= \frac{1}{d^2}\hat{X}^{2(q+q')}\hat{Z}^{2(p+p')}\hat{P}\omega^{2(q+q')(p+p')} \\
&= \hat{F}^{\text{Wootters}}(q+q',p+p').
\end{aligned}$$

Thus, by definition, the Wootters representation is a phase space representation.

Recall from Section 3.2 that Wootters also considered non-prime dimensions. In that case, the phase point operators (Equation (3.5)) were a tensor product of phase point operators (Equation (3.4)) for prime dimensions. The same is true here for the frame in composite dimensions. When d is composite with prime decomposition $d = d_1 d_2 \cdots d_k$. Let $\mathcal{S} = \mathcal{S}_1 \times \mathcal{S}_2 \times \cdots \mathcal{S}_k$ where each $\mathcal{S}_i = \mathbb{Z}_{d_i} \times \mathbb{Z}_{d_i}$. When d is composite the frame is $\mathcal{F}^{\text{Wootters}} = \{\hat{F}^{\text{Wootters}}(q_{(i)}, p_{(i)}) : (q_{(i)}, p_{(i)}) \in \mathcal{S}_{(i)}\}$ where

$$\hat{F}^{\text{Wootters}}(q_{(i)}, p_{(i)}) = \hat{F}^{\text{Wootters}}(q_{(1)}, p_{(1)}) \otimes \hat{F}^{\text{Wootters}}(q_{(2)}, p_{(2)}) \otimes \cdots \otimes \hat{F}^{\text{Wootters}}(q_{(k)}, p_{(k)})$$

and each $\hat{F}^{\text{Wootters}}(q_{(i)}, p_{(i)})$ is a frame as in Equation (5.1).

5.2 Odd dimensional discrete Wigner functions

Let d be an odd integer. Here $\mathcal{S} = \mathbb{Z}_d \times \mathbb{Z}_d$. Consider the frame $\mathcal{F}^{\text{odd}} = \{\hat{F}^{\text{odd}}(q,p) : (q,p) \in \mathcal{S}\}$, where

$$\hat{F}^{\text{odd}}(q,p) = \frac{1}{d^2} \hat{X}^{-2q} \hat{Z}^{2p} \hat{P} \omega^{-2qp}.$$

The quasi-probability function ρ^{odd} is a frame representation given by the frame $\mathcal{F}^{\text{odd}}$. The frame operator of $\mathcal{F}^{\text{odd}}$ is $\tilde{S} = d^{-1}\tilde{\mathbb{1}}$. The unique dual frame of $\mathcal{F}^{\text{odd}}$ is given by $\tilde{S}^{-1}\mathcal{F}^{\text{odd}}$, where here $\tilde{S}^{-1} = d\tilde{\mathbb{1}}$. The dual frame to $\mathcal{F}^{\text{odd}}$ is what Cohendet *et al.* call a set of Fano operators.

Consider the group of translations on $\mathcal{S}$ with unitary representation $\hat{T}_{(q,p)} = \hat{X}^{-q}\hat{Z}^{p}$. Then,

$$\hat{T}_{(q,p)} \hat{F}^{\text{odd}}(q',p') \hat{T}^{\dagger}_{(q,p)} = \hat{F}^{\text{odd}}(q+q', p+p').$$

By definition, this odd dimensional representation is a phase space representation.

5.3 Even dimensional discrete Wigner functions

Let d be an even integer. Then $\mathcal{S} = \mathbb{Z}_{2d} \times \mathbb{Z}_{2d}$. Consider the frame $\mathcal{F}^{\text{even}} = \{\hat{F}^{\text{even}}(q,p) : (q,p) \in \mathcal{S}\}$, where

$$\hat{F}^{\text{even}}(q,p) = \frac{1}{4d^2} \hat{X}^{q} \hat{Z}^{p} \hat{P} \omega^{\frac{qp}{2}}.$$

The quasi-probability function ρ^{even} is a frame representation given by the frame $\mathcal{F}^{\text{even}}$. The frame operator of $\mathcal{F}^{\text{even}}$ is $\tilde{S} = (2d)^{-1}\tilde{\mathbb{1}}$. However, this implies the frame is only tight; it is not a basis and the dual is not unique. The canonical dual frame of $\mathcal{F}^{\text{even}}$ is given by $\tilde{S}^{-1}\mathcal{F}^{\text{even}}$, where here $\tilde{S}^{-1} = 2d\tilde{\mathbb{1}}$. The canonical dual frame to $\mathcal{F}^{\text{even}}$ is what was called a set of Fano operators.

Consider the group of translations on $\mathcal{S}$ with unitary representation $\hat{T}_{(q,p)} = \hat{X}^{\frac{q}{2}}\hat{Z}^{\frac{p}{2}}$. Then,

$$\hat{T}_{(q,p)}\hat{F}^{\text{even}}(q',p')\hat{T}^{\dagger}_{(q,p)} = \hat{F}^{\text{even}}(q+q',p+p').$$

Thus, by definition, this odd dimensional representation is a phase space representation.

5.4 Wigner functions on the sphere

Let d be any integer. Here the phase space is $\mathcal{S} = \mathbb{S}^2$. Consider the (continuous) frame $\hat{F}^{\text{sphere}} : \mathbb{S}^2 \to \text{Herm}(\mathcal{H})$ given by

$$\hat{F}^{\text{sphere}}(\mathbf{n}) = \hat{\triangle}(\mathbf{n}),$$

where $\hat{\triangle}(\mathbf{n})$ is the same kernel given in Equation (3.13). The quasi-probability function ρ^{sphere} is a frame representation given by the frame $\hat{F}^{\text{sphere}}$. From Equation (3.11), it follows that the frame operator of $\hat{F}^{\text{sphere}}$ is $\tilde{S} = 4\pi d^{-1}\tilde{\mathbb{1}}$. Thus the frame is tight. Equation (3.12) is the group covariance property defining is a phase space representation for the group SU(2).

Now consider the discrete representation on sphere defined by Heiss and Weigert. Now the phase space $\mathcal{S}$ is a subset of points on the sphere which form a valid constellation. Consider the frame $\mathcal{F}^{\text{constellation}} = \{\hat{F}^{\text{constellation}}(s) : s \in \mathcal{S}\}$, where

$$F^{\text{constellation}}(s) = \hat{\triangle}_s,$$

where $\hat{\triangle}_s$ is a kernel satisfying the postulates (3.15)-(3.18). The quasi-probability function $\rho^{\text{constellation}}$ is a frame representation given by the frame $\mathcal{F}^{\text{constellation}}$. As was the case for the other discrete representation in the previous examples, the frame operator of $\mathcal{F}^{\text{constellation}}$ is $\tilde{S} = d^{-1}\tilde{\mathbb{1}}$. The unique dual frame of $\mathcal{F}^{\text{constellation}}$ is given by $\tilde{S}^{-1}\mathcal{F}^{\text{constellation}}$, where here $\tilde{S}^{-1} = d\tilde{\mathbb{1}}$. Thus, the dual frame to $\mathcal{F}^{\text{constellation}}$ is what Heiss and Weigert call a dual kernel. Again, from Equation (3.18), this representation satisfies definition of a phase space representation.

5.5 Finite fields discrete phase space representation

Let d be a power of a prime number. Here $\mathcal{S} = \mathbb{F}_d \times \mathbb{F}_d$. Consider the frame $\mathcal{F}^{\text{field}} = \{\hat{F}^{\text{field}}(q,p) : (q,p) \in \mathcal{S}\}$, where

$$\hat{F}^{\text{field}}(q,p) = \frac{1}{d}\left(\sum_{(q,p)\in\lambda} \hat{Q}(\lambda) - \hat{\mathbb{1}}\right).$$

The quasi-probability function ρ^{field} is a frame representation given by the frame $\mathcal{F}^{\text{field}}$. The frame operator of $\mathcal{F}^{\text{field}}$ is $\tilde{S} = d^{-1}\tilde{\mathbb{1}}$. The unique dual frame of $\mathcal{F}^{\text{field}}$ is given by $\tilde{S}^{-1}\mathcal{F}^{\text{field}}$, where here $\tilde{S}^{-1} = d\tilde{\mathbb{1}}$. The dual frame to $\mathcal{F}^{\text{field}}$ is a set of phase point operators.

This particular representation is constructed to be translationally covariant (recall Equation (3.26)) and is thus a phase space representation.

Chapter 6

Frame representations of quantum states and measurements

Table 3.1 shows that most proposed quasi-probability functions are representations of quantum states alone. In Sections 6.1 and 6.2 it is shown that there are two approaches within the frame formalism to lift any representation of states to a fully autonomous representation of finite dimensional quantum theory. In Section 6.3, a set of internal consistency conditions for each of the two approaches is given that allows one to view quantum theory independent of the standard operator theoretic formalism.

A short detour is taken in Section 6.4 to show that a more complete operational formulation of quantum theory (namely, one which includes *transformations*) is not outside the scope of the frame formalism. Finally in Section 6.5 a novel quasi-probability representation based on *SIC-POVMs* is presented which relates the frame formalism a recent research topic in quantum information.

6.1 Deformed probabilistic frame representations

The first frame representation approach consists of mapping both states and measurements to $L^2(\mathcal{S}, \mu)$ via a particular choice of frame $\hat{F}$. i.e. $\hat{\rho} \mapsto \rho(s) := \langle \hat{\rho}, \hat{F}(s) \rangle$ and $\hat{M}_k \mapsto M_k(s) := \langle \hat{M}_k, \hat{F}(s) \rangle$. The functions in the range of this frame representation when the domain is restricted to the density operators are called *quasi-probability densities*. Similarly, the functions in the range of the frame representation when the domain is restricted to the effects are called *conditional quasi-probabilities*. Together these mappings are called a *deformed probabilistic frame representation*. The reason for the qualifier *deformed* will become apparent. The deformed probabilistic frame representation achieves the following description of an experiment which is equivalent to Definition 2.2.1 of a quantum probabilistic description.

Definition 6.1.1. *Any model of a set of experimental configurations is a* deformed probabilistic frame description *if all of the following properties hold.*

(a) There is a set of allowed properties $\mathcal{S}$ with a positive measure μ.

(b) A preparation (state) is represented by a quasi-probability density $\rho(s) \in \mathbb{R}$ which satisfies the normalization condition $\int_{\mathcal{S}} d\mu(s)\rho(s) = 1$.

(c) A measurement is represented by a set of conditional quasi-probabilities $\{M_k(s) \in \mathbb{R}\}$ which satisfies $\sum_k M_k(s) = 1$ for all $s \in \mathcal{S}$.

(d) For a system with quasi-probability density ρ subject to the measurement $\{M_k\}$, the probability of obtaining outcome k is given by

$$\Pr(k) = \int_{\mathcal{S}} d\mu(s,r)\rho(s)M_k(r)\langle \hat{E}(s), \hat{E}(r)\rangle, \tag{6.1}$$

where $\hat{E}$ is any frame dual to $\hat{F}$.

Equation (6.1) is called the *deformed* law of total probability. Recall from Lemma 4.3.1 that all quasi-probability representations of states are frame representations. Given a quasi-probability representation (of states), one can identify the unique frame which gives rise to it. Then, using that frame to represent the measurement operators, one obtains a deformed probabilistic frame description of quantum experiment (Definition 6.1.1).

If the frame $\hat{F}$ is a positive operator, then $\rho(s) \geq 0$, and $M_k(s) \in [0,1]$ could be satisfied. Note that if this were the case and $\langle \hat{E}(s), \hat{E}(r)\rangle = \delta(s-r)$, then a deformed probabilistic frame description would be a classical probabilistic description (Definition 2.1.1).

6.2 Quasi-probabilistic frame representations

Notice that the deformed probability calculus in Equation (6.1) can be written

$$\Pr(k) = \int_{\mathcal{S}} d\mu(s)\ \rho(s)M'_k(s), \tag{6.2}$$

where

$$M'_k(s) = \int_{\mathcal{S}} d\mu(r)\ M_k(r)\langle \hat{E}(s), \hat{E}(r)\rangle. \tag{6.3}$$

Recall that M_k is the frame representation of $\hat{M}_k$ for the frame $\hat{F}$. Hence M'_k can be identified as the frame representation of $\hat{M}_k$ using a frame $\hat{E}$ that is dual to $\hat{F}$. The second frame representation approach consists of mapping states to $L^2(\Gamma,\mu)$ via a particular choice of frame $\hat{F}$ and measurements to $L^2(\Gamma,\mu)$ via a frame $\hat{E}$ that is dual to $\hat{F}$. i.e. $\hat{\rho} \mapsto \rho(\alpha) := \langle \hat{\rho}, \hat{F}(\alpha) \rangle$ and $\hat{M}_k \mapsto M_k(\alpha) := \langle \hat{M}_k, \hat{E}(\alpha) \rangle$. The term quasi-probability density has the same meaning as before. However, the functions in the range of the frame representation of the measurements (i.e. the frame representation defined via the dual $\hat{E}$) when the domain is restricted to the effects will still be called *conditional quasi-probabilities*. It should be clear which definition is being used from the context. Together these mappings are called a *quasi-probabilistic frame representation*. The quasi-probabilistic frame representation achieves the following description of an experiment which is equivalent to Definition 2.2.1 of a quantum probabilistic description.

Definition 6.2.1. *Any model of a set of experimental configurations is a* quasi-probabilistic frame description *if all of the following properties hold.*

(a) There is a set of allowed properties $\mathcal{S}$ with a positive measure μ.

(b) A preparation (state) is represented by a quasi-probability density $\rho(s) \in \mathbb{R}$ which satisfies the normalization condition $\int_{\mathcal{S}} d\mu(s)\rho(s) = 1$.

(c) A measurement is represented by a set of conditional quasi-probabilities $\{M'_k(s) \in \mathbb{R}\}$ which satisfies $\sum_k M'_k(s) = 1$ for all $s \in \mathcal{S}$.

(d) For a system with quasi-probability density ρ subject to the measurement $\{M'_k\}$, the probability of obtaining outcome k is given by the law of total probability Equation (6.2).

Equation (6.2), for true probabilities, is the law of total probability. Again, given a quasi-probability representation (of states), one can identify the unique frame which gives rise to it. Then, using that *dual frame* to represent the measurement operators, one obtains a quasi-probabilistic frame description of quantum experiment (Definition 6.1.1).

Forget, for the moment, about frames and consider the following definition.

Definition 6.2.2. *Any model of a set of experimental configurations is a* quasi-probabilistic description *if all of the following properties hold.*

(a) There is a set of allowed properties $\mathcal{S}$ with a positive measure μ.

(b) A preparation (state) is represented by a function $\rho(s) \in \mathbb{R}$ which satisfies the normalization condition $\int_{\mathcal{S}} d\mu(s)\rho(s) = 1$.

(c) A measurement is represented by a set of functions $\{M_k(s) \in \mathbb{R}\}$ which satisfies $\sum_k M_k(s) = 1$ for all $s \in \mathcal{S}$.

(d) For a system with preparation function ρ subject to the measurement $\{M_k\}$, the probability of obtaining outcome k is given by

$$\Pr(k) = \int_{\mathcal{S}} d\mu(s)\ \rho(s) M_k(s).$$

The differences in this definition and Definition 6.2.1 are subtle but important. Definition 6.2.2 does not refer to frames and is more general. If one has a quasi-probabilistic frame representation (a pair of mathematical mappings), then one has an effective description of a quantum experiment which satisfies Definition 6.2.1 and in turn also satisfies Definition 6.2.2. The converse is not true; if such a quasi-probabilistic description exists (Definition 6.2.2), it is not necessarily given by a quasi-probabilistic *frame* representation.

Note that, in Definitions 6.2.2 and 6.2.1 if $\rho(s) \geq 0$, and $M_k(s) \in [0,1]$, then a quasi-probabilistic description is a classical probabilistic description (Definition 2.1.1).

6.3 Internal consistency conditions

Recall that in a deformed probabilistic frame representation the definition of a quasi-probability density is a function in the range of a frame representation when the domain is restricted to the density operators. And, the conditional quasi-probabilities are the functions in the range of a frame representation when the domain is restricted to the effects. Of course, for a particular choice of frame, not every function in $L^2(\mathcal{S},\mu)$ will correspond to a valid quantum state or effect. Here a set of *internal* conditions is provided, independent of the standard axioms of quantum theory, which characterize the valid functions in $L^2(\mathcal{S},\mu)$. The conditions can be found by noting that the frame representation Equation (4.6) is an isometric and algebraic isomorphism from $\mathrm{Herm}(\mathcal{H})$ to $L^2(\mathcal{S},\mu)$ equipped with inner product

$$\langle A, B\rangle_{\mathtt{E}} := \int_{\mathcal{S}^2} d\mu(s,r)\ A(s)B(r)\mathtt{E}(s,r),$$

where $\mathtt{E}(s,r) := \langle \hat{E}(s), \hat{E}(r)\rangle$, and algebraic multiplication

$$(A \star_{\mathfrak{F}} B)(s) := \int_{\mathcal{S}^2} d\mu(r,t)\ A(r)B(\gamma)\mathfrak{F}(s,r,t),$$

where $\mathfrak{F}(s,r,t) = \langle \hat{F}(s), \hat{E}(r)\hat{E}(t)\rangle$.

Now the condition for a function in $L^2(\mathcal{S},\mu)$ to be a valid state or effect can be stated. A *pure state* is a function $\rho_{\mathrm{pure}} \in L^2(\mathcal{S},\mu)$ satisfying $\rho_{\mathrm{pure}} \star_{\mathfrak{F}} \rho_{\mathrm{pure}} = \rho_{\mathrm{pure}}$. A general *state* is a function $\rho \in L^2(\mathcal{S},\mu)$ satisfying $\langle \rho, \rho_{\mathrm{pure}}\rangle_{\mathtt{E}} \geq 0$ for all pure states and $\int_{\mathcal{S}} d\mu(s)\rho(s) = 1$. A *measurement* is represented by a set $\{M_k \in L^2(\mathcal{S},\mu)\}$ of *effects* which satisfies $\langle M_k, \rho_{\mathrm{pure}}\rangle_{\mathtt{E}} \geq 0$ for all pure states and for which $\sum_k M_k = \mathbb{1}$, where $\mathbb{1}$ is the identity element in $L^2(\mathcal{S},\mu)$ with respect to the algebra defined by $\star_{\mathfrak{F}}$. That is, $\mathbb{1}$ is the unique element satisfying $\mathbb{1} \star_{\mathfrak{F}} A = A \star_{\mathfrak{F}} \mathbb{1}$ for all $A \in L^2(\mathcal{S},\mu)$.

Recall for a quasi-probabilistic frame representation of quantum theory, the term quasi-probability density has the same meaning as in a deformed probability representation. And, the conditional quasi-probabilities are the functions in the range of the frame representation of the measurements (i.e. the frame representation defined via the dual $\hat{E}$) when the domain is restricted to the effects. Again for this approach, states and measurements in $L^2(\mathcal{S},\mu)$ must meet certain criteria to be valid. The conditions are similar to those in the deformed probability representation. Indeed the pure states and general states are equivalently characterized. However, a measurement is now represented by a set $\{M_k \in L^2(\mathcal{S},\mu)\}$ which satisfies $\langle M_k, \rho_{\text{pure}}\rangle \geq 0$ (now the usual pointwise inner product) for all pure states and for which $\sum_k M_k = \mathbb{1}$, where $\mathbb{1}$ is the identity element in $L^2(\mathcal{S},\mu)$ with respect to the algebra defined by $\star_{\mathfrak{E}}$ (which is defined in the same way as $\star_{\mathfrak{F}}$ with the roles of the frame and its dual reversed).

6.4 Transformations

A transformation is a superoperator (an operator acting on operators) $\tilde{\Phi} : \mathrm{D}(\mathcal{H}) \to \mathrm{D}(\mathcal{H})$. Operationally, an experiment consists of preparations followed by transformations and ending in a measurement. Note that, in a purely operational sense, the transformations are somewhat redundant as they could be bundled with either the preparations (to make new preparations) or measurements (to make new measurements).

A completely positive (CP) map is a linear superoperator $\tilde{\Phi}$ satisfying

$$\mathrm{Tr}[(\tilde{\Phi} \otimes \hat{\mathbb{1}})\hat{\rho}] \geq 0,$$

for every pure state $\hat{\rho}$ on extended system of arbitrary finite dimension. If in addition, the CP map satisfies $\mathrm{Tr}(\tilde{\Phi}(\hat{\rho})) = \mathrm{Tr}(\hat{\rho})$, it is called a completely positive trace-preserving (CPTP) map. The CPTP maps are the transformations which are admissible. Admissible means the mathematical objects which could conceivably describe the transitions a systems experiences between the preparation and measurement phase of an experiment.

In classical theories, transitions in probability are represented by matrices called *stochastic matrices*. It is natural to attempt a similar representation of transitions of quantum states here. Matrix representation are typical in quantum theory. A linear operator $\hat{A}$ is usually mapped to a matrix with entries a_{ij} given by $a_{ij} = \langle \phi_i, \hat{A}\phi_j\rangle$ where $\{\phi_i\}$ is an orthonormal basis for $\mathcal{H}$. Then the action of the operator is representation as the usual matrix multiplication. However, a slightly modified approach is required here when using frames (which reduces to usual matrix representations when the frame and an orthonormal basis coincide).

Let $A(s)$ be a frame representation of Hermitian operator $\hat{A}$ for a frame $\hat{F}$. Let $\hat{E}$ be a dual frame of $\hat{F}$ and consider the action of a superoperator

$$\tilde{\Phi}\hat{A} = \tilde{\Phi} \int_{\mathcal{S}} d\mu(s) A(s) \hat{E}(s). \tag{6.4}$$

Denote the frame representation of $\tilde{\Phi}\hat{A}$ as $A^{\Phi}(r) := \langle \hat{F}(r), \tilde{\Phi}\hat{A}\rangle$. As was the case for including measurements into a frame representations, two approaches can be identified.

The first approach follows directly from Equation (6.4)

$$A^{\Phi}(r) = \int_{\mathcal{S}} d\mu(s)\Phi^{\mathrm{qp}}(r,s)A(s), \tag{6.5}$$

where $\Phi^{\mathrm{qp}}(r,s) = \langle \hat{F}(r), \tilde{\Phi}\hat{E}(s)\rangle$ ("qp" is a label meant to abbreviate "quasi-probability"). Notice that Equation (6.5) is just the usual (perhaps infinite dimensional) matrix multiplication rule It is the same rule for transitioning probability distributions via stochastic matrices in classical theories. However, as opposed to stochastic matrices, Φ^{qp} could have negative entries.

Alternatively, consider the intermediate step

$$\hat{E}(s) = \int_{\mathcal{S}} d\mu(t)\langle \hat{E}(t), \hat{E}(s)\rangle \hat{F}(t). \tag{6.6}$$

Then Equation (6.4) becomes

$$A^{\Phi}(r) = \int_{\mathcal{S}^2} d\mu(s,t)\Phi^{\mathrm{def}}(r,t)\mathtt{E}(t,s)A(s), \tag{6.7}$$

where $\Phi^{\mathrm{def}}(r,t) = \langle \hat{F}(r), \tilde{\Phi}\hat{F}(t)\rangle$ ("def" is a label meant to abbreviate "deformed probability"). This second approach is analogous to the deformed probabilistic frame representation of Section 6.1.

6.5 Example: SIC-POVM representation

In [42], the authors conjecture[1] that the set $\{\phi_\alpha \in \mathcal{H} : \alpha \in \mathbb{Z}_d \times \mathbb{Z}_d\} = \{U_{(p,q)}\phi : (p,q) \in \mathbb{Z}_d \times \mathbb{Z}_d\}$ for some $\phi \in \mathcal{H}$ and

$$U_{(p,q)} = \omega^{\frac{pq}{2}} X^p Z^q \tag{6.8}$$

forms a *symmetric informationally complete positive operator valued measure* (SIC-POVM). The defining condition of a SIC-POVM is

$$|\langle \phi_\alpha, \phi_\beta\rangle|^2 = \frac{\delta_{\alpha\beta}d+1}{d+1}. \tag{6.9}$$

The set is called symmetric since the vectors have equal overlap. The POVM is formed by taking the projectors onto the one-dimensional subspaces spanned by the vectors. It is informationally complete since these d^2 projectors span $\mathrm{Herm}(\mathcal{H})$.

[1] Apparently this was conjectured earlier by Zauner in a Ph.D. thesis not available in english. See `http://www.imaph.tu-bs.de/qi/problems/23.html`.

As of writing, there is still no proof that SIC-POVMs exist in every dimension. However, it is still highly believed they do exist as there is numerical evidence for there existence for every dimension up to $d = 45$ [42]. Some authors have expressed urgency in determining whether SIC-POVMs exist or not by showing some desirable property they have *if* they were to exist [2].

Suppose then that for any dimension d, a SIC-POVM exists. Notice that a SIC-POVM forms a frame. Explicitly, let

$$\mathcal{F} = \left\{ \hat{F}_\alpha := \frac{1}{d}\phi_\alpha\phi_\alpha^* : \alpha \in \mathbb{Z}_d \times \mathbb{Z}_d \right\}$$

denote this frame. From the definition of the SIC-POVM, Equation (6.9),

$$\mathtt{F}_{\alpha\beta} := \langle \hat{F}_\alpha, \hat{F}_\beta \rangle = \frac{\delta_{\alpha\beta} d + 1}{d^2(d+1)}.$$

Since the frame forms a basis, the dual frame is unique and thus the inverse frame operator must be satisfy

$$\langle \hat{F}_\alpha, \tilde{S}^{-1}\hat{F}_\beta \rangle = \delta_{\alpha\beta}.$$

By inspection

$$\hat{E}_\beta = \tilde{S}^{-1}\hat{F}_\beta = d(d+1)\hat{F}_\beta - \hat{\mathbb{1}}.$$

Representing a quantum state via the frame or canonical dual yields the neat reconstruction formulae

$$\hat{\rho} = \sum_\alpha (d(d+1)\rho_\alpha - 1)\hat{F}_\alpha, \tag{6.10}$$

$$\hat{\rho} = \sum_\alpha \rho_\alpha (d(d+1)\hat{F}_\alpha - \hat{\mathbb{1}}), \tag{6.11}$$

where $\rho_\alpha := \langle \hat{F}_\alpha, \hat{\rho} \rangle$ is the frame representation of $\hat{\rho}$.

Equation (6.10) was given in [2]. This equation fits naturally into the deformed probabilistic frame representation formalism discussed in Section 6.1. Notice that the dual frame satisfies

$$\begin{aligned}
\mathtt{E}_{\alpha\beta} &= \langle \hat{E}_\alpha, \hat{E}_\beta \rangle \\
&= \langle d(d+1)\hat{F}_\alpha - \hat{\mathbb{1}}, d(d+1)\hat{F}_\beta - \hat{\mathbb{1}} \rangle \\
&= d^2(d+1)^2 \mathtt{F}_{\alpha\beta} - 2(d+1) + d \\
&= d(d+1)\delta_{\alpha\beta} - d + 1.
\end{aligned}$$

If an arbitrary measurement $\{\hat{M}_k\}$ is also represented via the SIC-POVM frame as $M_{k,\beta} := \langle \hat{F}_\beta, \hat{M}_k \rangle$, then Equation (6.10) is identical to the deformed law of total probability

$$\Pr(k) = \sum_{\alpha\beta} \rho_\alpha M_{k,\beta} \mathtt{E}_{\alpha\beta}.$$

Equation (6.11) fits more naturally into the quasi-probabilistic frame representation formalism discussed in Section 6.2. If an arbitrary measurement $\{\hat{M}_k\}$ is represented via the canonical dual to the SIC-POVM frame as $M'_{k,\alpha} := \langle \hat{E}_\alpha, \hat{M}_k \rangle$, then Equation (6.11) is identical to the deformed law of total probability

$$\Pr(k) = \sum_\alpha \rho_\alpha M'_{k,\alpha}.$$

Since the SIC-POVM frame is made of projectors, the frame representation of the density operator is a true probability distribution. However, the dual frame operators are not positive. Thus in a quasi-probabilistic frame representation (Section 6.2), the conditional quasi-probabilities will not be true probabilities as they must possess negative values. It will be shown in the next chapter than this is a general feature of quantum theory and applies to any quasi-probabilistic frame representation.

Chapter 7

Negativity and the non-classicality of quantum theory

It has been mentioned already in this thesis that negativity is a *necessary* feature of quasi-probability representations. This is significant because if it were false, a classical probabilistic description whole suffice for any physical experiment. A mapping from a quantum probabilistic description to a classical probabilistic description, called a *classical representation*, is defined in Section 7.1. A minimal generalization of a classical representation allows for a definition of a quasi-probability representation (of quantum theory) which makes no references to linearity, invertibility, or L^2 spaces as was the case for the quasi-probability representation for states alone (Definition 3.8.1). This definition of a quasi-probability representation for an entire quantum experiment is given in Section 7.2. It is then shown how frames are made necessary for this more general definition of a quasi-probability representation. Finally, in Section 7.3 the theorem that denies the existence of a classical representation and enforces negativity in quasi-probability representations is presented.

7.1 Classical representations of quantum theory

Suppose a quantum probabilistic description (Definition 2.2.1) is given for an experimental arrangement and an equivalent classical probabilistic description (Definition 2.1.1) is desired. Then a mapping from the set of density operators $\mathrm{D}(\mathcal{H})$ to probability densities $\hat{\rho} \mapsto \rho(s)$ and a mapping from the set of effects $\mathrm{E}(\mathcal{H})$ to the measurable functions $\hat{M} \mapsto M(s)$ is required. An important implicit assumption here is that the domain of these mappings are the sets $\mathrm{D}(\mathcal{H})$ and $\mathrm{E}(\mathcal{H})$. Operationally this means that the quantum probabilistic description is exhaustive in the sense that it is describing the most general quantum experiment. The most important requirement these mappings must satisfy is convex-linearity; they must preserve the convex structure of $\mathrm{E}(\mathcal{H})$. This can be motivated with the following counterexample. Consider a non-convex-linear representation in which the density operators $\hat{\rho}_1$ and $\hat{\rho}_2$ are represented by the functions $\rho_1(s)$ and $\rho_2(s)$. Now suppose a coin is flipped to decide which state to prepare; with probability p, $\hat{\rho}_1$ is prepared and with probability $1-p$ $\hat{\rho}_2$ is prepared. The quantum state that has been prepared is represented by $\hat{\rho} = p\hat{\rho}_1 + (1-p)\hat{\rho}_2$. However, the function $\rho(s)$, representing $\hat{\rho}$, is not equal to $p\rho_1(s) + (1-p)\rho_2(s)$. Thus the theory changes in a fundamental way depending on whether or not a simple coin is chosen to be tossed. Convexity is of such foundational importance in probability theory that is often taken as an obvious requirement and implicitly assumed.

In addition to convex-linearity, the mappings must also satisfy the conditions of the following definition.

Definition 7.1.1. *Let $\hat{\rho}$ be any arbitrary density operator and $\hat{M}$ an arbitrary effect. A* classical representation of quantum theory *is a pair of convex-linear mappings $\hat{\rho} \mapsto \rho(s)$ and $\hat{M} \mapsto M(s)$ for which*

$$\rho(s) \geq 0 \text{ and } \int_{\mathcal{S}} d\mu(s)\, \rho(s) = 1, \tag{7.1}$$

$$M(s) \in [0,1] \text{ and } \mathbb{1}(s) = 1, \tag{7.2}$$

where $\hat{\mathbb{1}} \mapsto \mathbb{1}(s)$ and the law of total probability holds

$$\mathrm{Tr}(\hat{M}\hat{\rho}) = \int_{\mathcal{S}} d\mu(s)\, M(s)\rho(s). \tag{7.3}$$

If such a classical representation exists, then a classical probabilistic description exists for any quantum experiment. However, as Theorem 7.3.1 will show, such a classical representation *does not* exist. It is important to note that this does not rule out a classical probabilistic description. Thus, a classical representation is still possible, but it is less *plausible* if the constraints on the mappings of a classical representation are held to be reasonable.

7.2 Equivalence of the quasi-probability and frame representation of quantum theory

Recall from Section 4.3 that a quasi-probability representation of quantum states is a map $\mathrm{Herm}(\mathcal{H}) \to L^2(\mathcal{S},\mu)$ which is linear and invertible. One might ask for a set of less restrictive conditions for a quasi-probability representation of quantum theory (i.e. of states *and* measurements). Consider Definition 7.1.1 of a classical representation. The minimal generalization of allowing negative values in the probability assignments will be made.

Definition 7.2.1. *Let $\hat{\rho}$ be any arbitrary density operator and $\hat{M}$ an arbitrary effect. A* quasi-probability representation of quantum theory *is a pair of convex-linear mappings $\hat{\rho} \to \rho(s)$ and $\hat{M} \to M(s)$ for which*

$$\rho(s) \in \mathbb{R} \textit{ and } \int_{\mathcal{S}} d\mu(s)\, \rho(s) = 1, \tag{7.4}$$

$$M(s) \in \mathbb{R} \textit{ and } \mathbb{1}(s) = 1, \tag{7.5}$$

where $\hat{\mathbb{1}} \to \mathbb{1}(s)$ and the law of total probability holds

$$\mathrm{Tr}(\hat{M}\hat{\rho}) = \int_{\mathcal{S}} d\mu(s)\, M(s)\rho(s). \tag{7.6}$$

If $\rho(s) \geq 0$ and $M(s) \in [0, 1]$, then indeed a quasi-probability representation is a classical representation of quantum theory. It can now be asked if this more general representation can be achieved. This is an easy question to answer because it is now known from Section 6.2 that a frame can achieve the desired result via a quasi-probabilistic frame representation. That is, frames are sufficient to describe this more general requirement of a quasi-probability representation of quantum theory. But are frames also necessary as was the case for only quantum states? The answer is no unless an additional assumption, which is discussed further along, is met. First a short lemma is needed.

Lemma 7.2.2. *A convex-linear mapping T from the set of effects to real valued functions on $\mathcal{S}$ can be uniquely extended to a linear function on the space of all Hermitian operators.*

Proof. Busch [8] has shown that a convex-linear map $R : \mathrm{E}(\mathcal{H}) \to [0, 1]$ can be extended uniquely to a linear map on the space of all Hermitian operators through the association $R(\hat{A}) = R(\hat{P}_1) - R(\hat{P}_2)$, where $\hat{A} = \hat{P}_1 - \hat{P}_2$ is any decomposition of $\hat{A} \in \mathrm{Herm}(\mathcal{H})$ in terms of positive operators. The same logic is applied to T. That is, too each convex-linear mapping T, there exists a unique extension to a linear map $T : \hat{A} \mapsto A(s)$ defined by $T(\hat{A}) = T(\hat{P}_1) - T(\hat{P}_2)$, where $\hat{A} = \hat{P}_1 - \hat{P}_2$ is any decomposition of $\hat{A} \in \mathrm{Herm}(\mathcal{H})$ in terms of positive operators. □

Unless noted otherwise, the mappings of a quasi-probability representation will be taken to be this unique extension to $\mathrm{Herm}(\mathcal{H})$. Recall that frames were not necessary to describe a quasi-probability representation of quantum theory (Definition 7.2.1). However, the following theorem shows that frame representations are in some sense "half-necessary".

Theorem 7.2.3. *The mapping $\hat{\rho} \mapsto \rho(s)$ in a quasi-probability representation is necessarily a frame representation of* $\mathrm{Herm}(\mathcal{H})$ *for some unique frame $\hat{F}$.*

Proof. Call the mapping T. Lemma 7.2.2 implies T has a unique linear extension on $\mathrm{Herm}(\mathcal{H})$. The Riesz representation theorem implies that $T(\hat{A})(s) = \langle \hat{F}(s), \hat{A} \rangle$

for some unique mapping $\hat{F} : \mathcal{S} \to \mathrm{Herm}(\mathcal{H})$ (not necessarily a frame). Condition (7.4) implies

$$\int_{\mathcal{S}} d\mu(s)\ \hat{F}(s) = \hat{\mathbb{1}}. \tag{7.7}$$

Application of the Cauchy-Schwarz inequality yields

$$\int_{\mathcal{S}} d\mu(s)\ |\langle \hat{F}(s), \hat{A}\rangle|^2 \le \int_{\mathcal{S}} d\mu(s)\ \|\hat{F}(s)\|^2 \|\hat{A}\|^2. \tag{7.8}$$

Now consider

$$\begin{aligned}\|\hat{F}(s)\|^2 &= \langle \hat{F}(s), \hat{F}(s)\rangle \\ &\le \max_r \langle \hat{F}(s), \hat{F}(r)\rangle \\ &=: \langle \hat{F}(s), \hat{B}\rangle,\end{aligned}$$

where $\hat{B}$ has be implicitly defined as the operator which achieves this maximum. Define $b := \mathrm{Tr}(\hat{B}) < \infty$. From (7.8) and using (7.7)

$$\begin{aligned}\int_{\mathcal{S}} d\mu(s)\ |\langle \hat{F}(s), \hat{A}\rangle|^2 &\le \int_{\mathcal{S}} d\mu(s)\ \langle \hat{F}(s), \hat{B}\rangle \|\hat{A}\|^2 \\ &= \langle \hat{\mathbb{1}}, \hat{B}\rangle \|\hat{A}\|^2 \\ &= b\|\hat{A}\|^2.\end{aligned}$$

Thus $T(\hat{A}) \in L^2(\mathcal{S})$ for all $\hat{A} \in \mathrm{Herm}(\mathcal{H})$. Now, Equation (7.6) must hold for all $\hat{\rho} \in \mathrm{D}(\mathcal{H})$ which also spans $\mathrm{Herm}(\mathcal{H})$. Thus

$$\hat{M} = \int_{\mathcal{S}} d\mu(s)\ M(s)\hat{F}(s). \tag{7.9}$$

Since this holds for all $\hat{M} \in \mathrm{E}(\mathcal{H})$, $\hat{F}$ must also span $\mathrm{Herm}(\mathcal{H})$. Suppose $T(\hat{A}) = T(\hat{B})$ for all $\hat{A}, \hat{B} \in \mathrm{Herm}(\mathcal{H})$. In particular, $\langle \hat{F}(s), \hat{A}\rangle = \langle \hat{F}(s), \hat{B}\rangle$. Linearity and the fact that $\hat{F}$ spans $\mathrm{Herm}(\mathcal{H})$ implies $\hat{A} = \hat{B}$ and T is therefore one-to-one. Applying Lemma 4.3.1 yields the desired result. □

Note that sufficiency holds if the frame satisfies Equation (7.7). Also note that the theorem does not imply that the mapping $\hat{M} \mapsto M(s)$ is a frame representation. Although there is a unique operator valued mapping $\hat{E} : \mathcal{S} \to \mathrm{Herm}(\mathcal{H})$ such that $M(s) = \langle \hat{E}(s), \hat{M}\rangle$, it need not be a frame. This is due to the fact that there is no condition on $\hat{E}$ analogous to Equation (7.7) for the frame $\hat{F}$. The following theorem provides a necessary and sufficient condition on the mapping $\hat{M} \mapsto M(s)$ being a frame representation.

Theorem 7.2.4. *The mapping $\hat{M} \mapsto M(s)$ in a quasi-probability representation is a frame representation of* $\mathrm{Herm}(\mathcal{H})$ *for some unique frame $\hat{E}$, dual to $\hat{F}$ if and only if the measure space $(\mathcal{S}, \mu)$ is finite.*

Proof. Assume the measure space is finite. That is, $\mu(\mathcal{S}) < \infty$. Through Lemma 7.2.2 and the Riesz representation theorem there is a unique operator valued mapping $\hat{E} : \mathcal{S} \to \mathrm{Herm}(\mathcal{H})$ such that $M(s) = \langle \hat{E}(s), \hat{M} \rangle$ which need not be a frame. Application of the Cauchy-Schwarz inequality yields

$$\int_{\mathcal{S}} d\mu(s)\ |\langle \hat{E}(s), \hat{A} \rangle|^2 \leq \int_{\mathcal{S}} d\mu(s)\ \|\hat{E}(s)\|^2 \|\hat{A}\|^2. \tag{7.10}$$

Since $\mathrm{Herm}(\mathcal{H})$ is finite, $\|\hat{E}(s)\|^2 \leq e$ for some finite e. Then Equation (7.2) becomes

$$\int_{\mathcal{S}} d\mu(s)\ |\langle \hat{E}(s), \hat{A} \rangle|^2 \leq e\|\hat{A}\|^2 \int_{\mathcal{S}} d\mu(s) = e\|\hat{A}\|^2 \mu(\mathcal{S}) < \infty.$$

Now, Equation (7.6) must hold for all $\hat{M} \in \mathrm{E}(\mathcal{H})$ which also spans $\mathrm{Herm}(\mathcal{H})$. Thus

$$\hat{\rho} = \int_{\mathcal{S}} d\mu(s)\ \rho(s)\hat{E}(s). \tag{7.11}$$

Since this holds for all $\hat{\rho} \in \mathrm{D}(\mathcal{H})$, $\hat{E}$ must also span $\mathrm{Herm}(\mathcal{H})$. Suppose $\langle \hat{E}(s), \hat{A} \rangle = \langle \hat{E}(s), \hat{B} \rangle$. Linearity and the fact that $\hat{E}$ spans $\mathrm{Herm}(\mathcal{H})$ implies $\hat{A} = \hat{B}$ and the mapping is therefore one-to-one. Lemma 4.3.1 implies that $\hat{E}$ is a frame. Rewriting Equation (8.4) as

$$\hat{\rho} = \int_{\mathcal{S}} d\mu(s)\ \langle \hat{F}(s), \hat{\rho} \rangle \hat{E}(s)$$

shows that $\hat{E}$ is dual to $\hat{F}$.

Now assume the converse: $\hat{E}$ is a frame (being dual to $\hat{F}$ does not matter here). By definition

$$\int_{\mathcal{S}} d\mu(s)\ |\langle \hat{E}(s), \hat{M} \rangle|^2 < \infty. \tag{7.12}$$

Consider the particular effect $\hat{\mathbb{1}}$. From Equation (7.5) in the definition of a quasi-probability representation, $\hat{\mathbb{1}} \mapsto \mathbb{1}(s) = 1$. Applying this particular choice to Equation (7.12) yields

$$\mu(\mathcal{S}) = \int_{\mathcal{S}} d\mu(s) < \infty.$$

Thus $(\mathcal{S}, \mu)$ is a finite measure space. □

7.3 Non-classicality of quantum theory

Theorem 7.3.1. *A classical representation of quantum theory (Definition 7.1.1) does not exist.*

Proof. The statement of the theorem is equivalent to the following statement. A quasi-probability representation of quantum theory satisfying $\rho(s) \geq 0$ and $M(s) \in [0,1]$ does not exists. Recall that for a quasi-probability representation, Equation (7.9) must hold. Now although $\hat{M} \mapsto M(s)$ is not necessary a frame representation, it must be given by $M(s) = \langle \hat{E}(s), \hat{M} \rangle$ for some unique mapping $\hat{E} : \mathcal{S} \to \mathrm{Herm}(\mathcal{H})$ (again, not necessarily a frame). Thus

$$\hat{M} = \int_{\mathcal{S}} d\mu(s)\ \langle \hat{E}(s), \hat{M} \rangle \hat{F}(s). \tag{7.13}$$

Consider the mapping

$$\tilde{\Phi}(\hat{M}) = \int_{\mathcal{S}} d\mu(s)\ \langle \hat{E}(s), \hat{M} \rangle \hat{F}(s), \tag{7.14}$$

If $\tilde{\Phi}$ were the identity super-operator, then $\hat{F}$ and $\hat{E}$ would satisfy the constraints of a quasi-probability representation in Equation (7.13). It is now shown this is not possible when both $\hat{F}(s)$ and $\hat{E}(s)$ are positive operators. Let $\{\phi_i \phi_j^* : i,j \in \mathbb{Z}_d\}$ be the standard basis for $\mathrm{L}(\mathcal{H})$. Then the Choi-Jamiolkowski representation of $\tilde{\Phi}$ is

$$\begin{aligned}
J(\tilde{\Phi}) &= \sum_{i,j\in\mathbb{Z}_d} \tilde{\Phi}(\phi_i \phi_j^*) \otimes \phi_i \phi_j^* \\
&= \int_{\mathcal{S}} d\mu(s)\ \left(\sum_{i,j\in\mathbb{Z}_d} \langle \phi_j, \hat{E}(s)\phi_i \rangle \hat{F}(s) \otimes \phi_i \phi_j^* \right) \\
&= \int_{\mathcal{S}} d\mu(s)\ \left(\hat{F}(s) \otimes \sum_{i,j\in\mathbb{Z}_d} \langle \phi_j, \hat{E}(s)\phi_i \rangle \phi_i \phi_j^* \right) \\
&= \int_{\mathcal{S}} d\mu(s)\ \left(\hat{F}(s) \otimes \hat{E}(s) \right),
\end{aligned}$$

which is a separable operator [48] on $\mathcal{H} \otimes \mathcal{H}$ when both $\hat{F}(s)$ and $\hat{E}(s)$ are positive operators. However, $J(\tilde{\mathbb{1}})$ is not a separable operator on $\mathcal{H} \otimes \mathcal{H}$ and thus $\tilde{\Phi}$ cannot be the identity super-operator. Hence there does not exist a quasi-probability representation in which $\rho(s) \geq 0$ and $M(s) \in [0,1]$. □

This theorem can also be proven using the results of Reference [29]. Theorem 2 of that paper shows that the channel $\tilde{\Phi}$ defined by Equation (7.14) for positive operators $\hat{F}$ and $\hat{E}$ is so-called *entanglement breaking*. However, Theorem 6 of Reference [29] states that if $\tilde{\Phi}$ has fewer than d Kraus operators, it is *not* entanglement breaking. Since the identity superoperator has fewer than d Kraus operators, $\tilde{\Phi}$ is not entanglement breaking and therefore $\hat{E}$ is not the dual of $\hat{F}$.

The following corollary, of purely mathematical interest, was proven in [16] using the same proof technique.

Corollary 7.3.2. *There does not exist a dual frame of positive operators for a frame of positive operators.*

Theorem 7.3.1 establishes the necessity of negativity in quasi-probability representations of quantum theory. In the next chapter another, older, notion of non-classicality, *contextuality*, will be compared with negativity.

Chapter 8

Connection with contextuality

Within the quantum formalism there are many notions of non-classicality. In the previous chapter, negativity was proven to be such a notion. In this chapter the idea of *contextuality* is considered as an alternative candidate. In Section 8.1 the traditional notion of contextual is reviewed. A generalization due to Spekkens is presented in Section 8.2. In Section 8.3 the argument of Spekkens that "negativity and contextuality are equivalent" is studied. It is shown in Section 8.4 that a generalization of the traditional notion of contextuality can be stated in a concise mathematical definition which makes the relationship between classicality and contextuality clear (as is case for the definition of quasi-probability). The most famous example of a contextual hidden variable model (the de Broglie Bohm model) is discussed in Section 8.5.

8.1 Traditional definition of contextuality

The traditional definition of contextuality evolved from a theorem which appears in a paper by Kochen and Specker [32]. The Kochen-Specker theorem concerns the standard quantum formalism: physical systems are assigned states in a complex Hilbert space $\mathcal{H}$ and measurements are made of observables represented by Hermitian operators (the standard quantum formalism was discussed in Section 2.2). The theorem establishes a contradiction between a set of plausible assumptions which would imply quantum systems possess values for observable quantities in the classical sense. Let $\mathcal{H}$ be the Hilbert space associated with a quantum system and $\hat{A} \in \text{Herm}(\mathcal{H})$ be the operator associated with an observable A . The function $f_\psi(A)$ represents the value of the observable A when the system is in state ψ. The added assumption used to derive the contradiction is for any function F, $f_\psi(F(A)) = F(f_\psi(A))$. This is plausible because, for example, we would expect that the value of A^2 could be obtained in this way from the value of A.

Assuming that physical systems do possess values which can be revealed via measurements, the Kochen-Specker theorem leads to the following counterintuitive example [31]. Suppose three operators $\hat{A}$, $\hat{B}$, and $\hat{C}$ satisfy $[\hat{A}, \hat{B}] = 0 = [\hat{A}, \hat{C}]$, but $[\hat{B}, \hat{C}] \neq 0$. The the value of the observable A will depend on whether observable B or C is chosen to be measured as well. That is, the value of A depends on the *context* of the measurement.

What the Kochen-Specker theorem establishes then is the mathematical framework of quantum theory does not allow for a *noncontextual* model. This fact is often shortened to the terminology "quantum theory is contextual".

8.2 Generalized definition of contextuality

The original notion of contextuality in lacking in the sense that it only applies to the standard form of quantum theory and does not apply to general operational models. This problem was addressed by Spekkens in Reference [45] and the results of that paper will be discussed here.

Give the preparation procedures the label $\mathcal{P}$ and measurement procedures the label $\mathcal{M}$. In a general operational model (including classical and quantum theory), the role is to specify the probabilities $p(k|\mathcal{P}, \mathcal{M})$ for the outcomes of a measurement procedure $\mathcal{M}$ given preparation procedure $\mathcal{P}$. Each $\mathcal{P}$ belongs to an equivalence class $e(\mathcal{P})$ in which any two preparations, $\mathcal{P}$ and $\mathcal{P}'$ are equivalent if $p(k|\mathcal{P}, \mathcal{M}) = p(k|\mathcal{P}', \mathcal{M})$ for all $\mathcal{M}$. Each $\mathcal{M}$ defines an equivalence class is a similar manner. The features of an experimental configuration which are not specified by the equivalence class of the procedure are called the *context* of the experiment.

One may supplement the operational theory with a classical state space (or ontological space) $(\mathcal{S}, \mu)$. Then the preparation procedures become probability densities $\rho_{\mathcal{P}}(s)$ while the measurement procedures become conditional probabilities $M_{\mathcal{M},k}(s)$. The probabilities of the outcomes of the measurements is required to satisfy the law of total probability

$$p(k|\mathcal{P}, \mathcal{M}) = \int_{\mathcal{S}} d\mu(s)\ \rho_{\mathcal{P}}(s) M_{\mathcal{M},k}(s) \tag{8.1}$$

Such a supplemented operation model is called an ontological model. The ontological model is *preparation noncontextual* if

$$\rho_{\mathcal{P}}(s) = \rho_{e(\mathcal{P})}(s). \tag{8.2}$$

That is, the representation of the preparation procedure is independent of context. Similarly the ontological model is *measurement noncontextual* if

$$M_{\mathcal{M},k}(s) = M_{e(\mathcal{M}),k}(s). \tag{8.3}$$

The terminology "contextual" is again shorthand for the inability of an operational theory to admit a (preparation or measurement) noncontextual ontological model.

However, the term "contextual" is also used to describe specific ontological model which do not satisfy Equations (8.2) and (8.3).

At first sight, this might not seem like a generalization of the traditional notion initiated by Kochen and Specker. However, the standard quantum formalism is an instance of an operational model. Moreover, considering only measurements, one can see that Spekkens generalizes the notion of *noncontextuality* from outcomes of individual measurements being independent of the measurement context to *probabilities* from outcomes of measurements being independent of the measurement context. In Reference [45], it was proven that quantum theory is both preparation and measurement contextual.

8.3 Equivalence of negativity and contextuality

Recall from Definition 2.2.1 that a preparation is specified by a density operator $\hat{\rho}$ and a measurement outcome by an effect $\hat{M}$. Thus a (preparation and measurement) noncontextual ontological model of quantum theory would require

$$\begin{aligned} \rho_{\mathcal{P}}(s) &= \rho_{\hat{\rho}}(s), \\ M_{\mathcal{M}}(s) &= M_{\hat{M}}(s), \end{aligned}$$

and the law of total probability

$$\mathrm{Tr}(\hat{M}\hat{\rho}) = \int_{\mathcal{S}} d\mu(s)\ M_{\hat{M}}(s)\rho_{\hat{\rho}}(s).$$

Assuming the probabilities satisfy the usual normalization conditions these equations are equivalent to those in Definition 7.1.1 of a classical representation. Spekkens first noticed this equivalence [44]. He has therefore independently obtained a connection between negativity and non-classicality. Similarly, our direct proof of the non-existence of a positive dual frame to a positive frame gives a new independent proof of this generalized notion of contextuality.

In addition to this indirect proof of the necessity of negativity, Spekkens provides the following independent direct proof [44]. Assume a classical representation exists (Definition 7.1.1). Theorem 7.2.3 and the comment following the proof imply

$$\begin{aligned} \rho(s) &= \langle \hat{F}(s), \hat{\rho} \rangle, \\ M(s) &= \langle \hat{E}(s), \hat{M} \rangle, \end{aligned}$$

where $\hat{F}$ is a frame and $\hat{E}$ may not be a frame (although it is an injective linear mapping $\mathcal{S} \to \mathrm{Herm}(\mathcal{H})$). Equation (7.6) must hold for for all effects $\hat{M} \in \mathrm{E}(\mathcal{H})$ and thus

$$\hat{\rho} = \int_{\mathcal{S}} d\mu(s)\ \langle \hat{F}(s), \hat{\rho} \rangle \hat{E}(s). \tag{8.4}$$

Since this must hold for all states $\hat{\rho}$, it must hold in particular for pure states. Let $\hat{\rho}$ be pure. Since $\rho(s) \geq 0$, Equation (8.4) implies $\hat{\rho}$ is in the convex hull of the range of $\hat{E}$. However, pure states are extremal and admit only the trivial decomposition. Thus $\rho(s)$ must be zero for all $s \in \mathcal{S}$ for which $\hat{F}(s) \neq \hat{\rho}$. That is, $\hat{F}$ is a POVM which discriminates the entire set of pure states. This is known to be impossible and hence the contradiction is derived.

Note that the terminology "negativity and contextuality are equivalent" is somewhat misleading. A quasi-probability model of a quantum experiment is *not* equivalent to a contextual ontological model. Within the formalism of quasi-probability representations of quantum theory, a classical representation (Definition 7.1.1) is a particular instance. A classical representation of quantum theory is also a particular instance of a general operational model, a non-contextual ontological model to be specific. The "equivalence" is obtained through the fact one can establish the non-existence of a classical representation (Theorem 7.3.1) starting from either formalism. In other words, a classical representation does not exist but one can relax the constraints on Definition 7.1.1 to achieve a representation which may contain negativity, contextuality, both, or something entirely different which has not yet been considered.

8.4 Concise definition of contextuality

Recall Definition 7.1.1 of a classical representation. As shown in Theorem 7.3.1 and the previous section, the requirements of such a representation are inconsistent. However, relaxing the requirement of positivity, one can obtain a quasi-probability representation (Definition 7.2.1). In other words, a classical representation is a special case of a quasi-probability representation when $\rho(s) > 0$ and $M(s) \in [0, 1]$. Presented in this concise mathematical fashion, the difference between classical and quasi-probability is clear. The notion of contextuality provided by Spekkens is the inability of an operational theory to admit a non-contextual ontological model. This notion does not make it clear how a classical representation is an instance of an ontological model which may be contextual. A definition of contextuality which makes the relationship clear is as follows.

Definition 8.4.1. *Let $\hat{\rho}$ be any arbitrary density operator and $\hat{M}$ an arbitrary effect. Let $C_{\mathcal{P}}$ and $C_{\mathcal{M}}$ be sets (called the contexts) and denote their elements by $c_{\mathcal{P}}$ and $c_{\mathcal{M}}$, respectively. A* contextual representation of quantum theory *is a pair of convex-linear (for each fixed context) mappings $(\hat{\rho}, c_{\mathcal{P}}) \mapsto \rho_{c_{\mathcal{P}}}(s)$ and $(\hat{M}, c_{\mathcal{M}}) \mapsto M_{c_{\mathcal{P}}}(s)$ for which*

$$\rho_{c_{\mathcal{P}}}(s) \geq 0 \text{ and } \int_{\mathcal{S}} d\mu(s)\, \rho_{c_{\mathcal{P}}}(s) = 1, \tag{8.5}$$

$$M_{c_{\mathcal{M}}}(s) \in [0, 1] \text{ and } \mathbb{1}_{c_{\mathcal{M}}}(s) = 1, \tag{8.6}$$

where $(\hat{\mathbb{1}}, c_{\mathcal{M}}) \to \mathbb{1}_{c_{\mathcal{M}}}(s)$ *and the law of total probability holds*

$$\mathrm{Tr}(\hat{M}_{c_{\mathcal{M}}}\hat{\rho}_{c_{\mathcal{P}}}) = \int_{\mathcal{S}} d\mu(s)\ M_{c_{\mathcal{M}}}(s)\rho_{c_{\mathcal{P}}}(s). \tag{8.7}$$

The sets $C_{\mathcal{P}}$ and $C_{\mathcal{M}}$ form the context of the preparation and measurement. They could be a list of instructions on what to do in the lab for example. Thus a classical representation is a contextual representation for which $\rho_{c_{\mathcal{P}}}(s) = \rho_{d_{\mathcal{P}}}(s)$ for all $c_{\mathcal{P}}, d_{\mathcal{P}} \in C_{\mathcal{P}}$ and $M_{c_{\mathcal{M}}}(s) = M_{d_{\mathcal{M}}}(s)$ for all $c_{\mathcal{M}}, d_{\mathcal{M}} \in C_{\mathcal{M}}$ or, without loss of generality, $|C_{\mathcal{P}}| = |C_{\mathcal{M}}| = 1$. It is now explicitly clear that negativity and contextuality are *not* equivalent if one takes these notions to describe specific representations which satisfy a more relaxed set of constraints than Definition 7.1.1 imposes on a classical representation.

8.5 de Broglie-Bohm model

A hidden variable theory originally formulated by de Broglie [13] and later by Bohm [6] is perhaps the most famous example of an ontological model of quantum theory. The model assumes that for a given experimental configuration, there exists particles with well defined trajectories and a quantum state ψ. The hidden variables are the positions of the particles in real spaces. That is, the classical state space is $\mathcal{S} = \mathbb{R}^3 \times \mathcal{H}$. The Hilbert space is included in the state space as its serves as a wave which guides the particle. The equation of motion of the particles is such that the quantum probability distribution $|\psi|^2$ is invariant. Thus, so long as it is assumed that the particles are prepared according to this distribution, the model provides the same predictions as the standard formulation of quantum theory.

Note that this model does not fit into the framework of quasi-probability representations or ontological models for two reasons. First, the model applies to infinite dimensional Hilbert spaces. The theory can be accommodated by extending the quasi-probability framework to infinite dimensions. Indeed the results of negativity and contextuality are conjectured to hold in infinite dimensions; it is assumed that the subtleties of infinite dimensions can be accounted for upon deeper mathematical analysis. The second reason the de Broglie-Bohm model does not fit into the framework considered in this thesis is not a deficiency of the framework considered in this thesis. The reason is the de Broglie-Bohm model does not consider the entire range of possible quantum states. Where a classical representation (Definition 7.1.1) contains a convex-linear mapping $\hat{\rho} \mapsto \rho(s)$ for all $\hat{\rho} \in \mathrm{D}(\mathcal{H})$, the de Broglie-Bohm model considers only a mapping $\psi \mapsto \rho_\psi(s)$ for all $\psi \in \mathcal{H}$. Bell notes that [4] "in the de Broglie-Bohm theory a fundamental significance is given to the wavefunction, and it cannot be transferred to the density matrix."

Bell *does not* claim that the situation is such that the de Broglie-Bohm model *cannot* be extended to include density operators. The key words in his comment are "fundament significance". Indeed, the de Broglie-Bohm model *can* be extended to include density operators provided this extension is either contextual or contains negativity. In either case, the pure states (wavefunctions) retain their significance while the density operators possess non-classical features.

As an example, the de Broglie-Bohm model could be such that $(\hat{\rho}, c_{\mathcal{P}}) \mapsto \rho_{c_{\mathcal{P}}}(s)$ where each preparation consists of a density operator $\hat{\rho}$ supplied with a context $c_{\mathcal{P}}$ which specifies a particular convex decomposition of $\hat{\rho}$ into pure states. Such a model would be preparation contextual.

Chapter 9

Connection with non-locality

Up to now no mention of composite systems has been made in this thesis. In order to prove the various theorems connecting quantum theory with non-locality, one requires knowledge of the quantum formalism of composite systems. However, understanding the statements and consequences of these results does not. References will be made to the *tensor product* and *collapse postulate*. These are concepts which appear in the axioms of quantum theory when composite systems are considered and can be found in any introductory textbook on quantum theory [7]. They are only mentioned to assist the keen reader.

In Section 9.1 *locality* is defined. Also discussed is the famous argument of Einstein and his colleagues that sparked the debate about the curious *nonlocal* features possessed by quantum theory which could be avoided if there existed hidden variables. Bell's theorem claimed to show, as recalled in Section 9.2, that even a hidden variable theory must nevertheless possess non-locality. This conclusion, when extended to the broadest class of hidden variable models, is questioned and its relation to negativity is presented in Section 9.3.

9.1 EPR incompleteness argument

In a paper by Einstein, Podolsky and Rosen (EPR) [15], it was argued that quantum mechanics is *incomplete* (each element of physical reality does not have a counterpart in quantum theory) if special relativity remains valid. The latter means physical causation must be local or events cannot have causes outside of their past light cones. Using a particular spatially separated quantum system, and some standard quantum theory, EPR concluded that quantum mechanics is either incomplete or nonlocal (or both!). Locality was such a desired property of any theory that quantum mechanics was concluded to be incomplete. That is, there must be elements of physical reality (hidden variables) which quantum mechanics does not account for.

The argument of EPR was reformulated by Bohm [6] for the simplest bipartite system possible: two qubits. The argument is built around the following hypothetical experiment. Two parties, Alice and Bob, are at distant locations with a source midway between them creating quantum systems described by the quantum state

$$\psi = \frac{1}{\sqrt{2}}(\phi_1 \otimes \phi_2 - \phi_2 \otimes \phi_1), \tag{9.1}$$

where $\{\phi_1, \phi_2\}$ is an orthonormal basis for a qubit. One particle is sent to Alice and the other to Bob. Alice performs the projective two-outcome measurement $\{\hat{P}_1, \hat{P}_2\}$ on the particle which was sent to her. The state in Equation (9.1) is such that Alice, once she performs her measurement, she can predict with certainty the outcome Bob receives when he performs the same measurement at his side of the experiment *regardless of whether or not the measurement events are spacelike separated (i.e. nonlocal)*. For example, Alice could perform the measurement $\{\phi_1\phi_1^*, \phi_2\phi_2^*\}$. According to the collapse postulate, if Alice registers the first outcome, Bob particle will immediately collapse to ϕ_2 and he is certain to obtain the second outcome if he were to make the same measurement. Therefore, unless there exists hidden variables which pre-determine the possible outcomes when the particles are created, quantum theory is *nonlocal*. Out of these arguments came the notion of entanglement and what Einstein referred to as “spooky action-at-a-distance”.

9.2 Bell’s theorem

Bell later investigated the possibility of finding the hidden variables Einstein thought to exist [4]. He noted immediately that the de Broglie-Bohm theory was such a theory yet in contained an astonishingly nonlocal character. He soon was able to prove that any hidden variable theory of quantum phenomena must possess nonlocal features. This is now called Bell’s theorem.

The proof is by contraction and follows the general line of reasoning which lead to the results in this thesis: build a mathematical model with assumptions that can be identified with (or motivated by) some notion of classicality then prove that quantum theory does not satisfy these assumptions. Consider the EPR experimental setup where the particles sent to Alice and Bob are not assumed to be quantum systems. Alice and Bob can each perform a two-outcome measurement with outcomes labeled A and B, respectively. Without loss of generality, the outcomes can be assigned numerical values $A, B = \pm 1$.

Suppose there exist a classical state space $(\mathcal{S}, \mu)$ (i.e. a set of hidden variables) which serve to determine the outcomes A and B. Probabilistic knowledge of the state is represented by a density $\rho(s) \geq 0$ which is normalized

$$\int_{\mathcal{S}} d\mu(s)\rho(s) = 1.$$

The different measurements Alice and Bob can perform are parameterized by detector settings a and b, respectively. Locality is enforced by assuming that the outcomes A and B depend only the local detector settings and the global state. That is $A = A(a, s)$ is allowed but $A = A(a, b, s)$ is not. Define the correlation function

$$C(a, b) = \int_{\mathcal{S}} d\mu(s) A(a, s) B(b, s) \rho(s). \tag{9.2}$$

Bell's theorem states that the correlations obtained in the EPR experiment (i.e. a particular quantum experiment) cannot satisfy this equation. The proof follows by deriving an inequality from Equation (9.2) such as

$$|C(a, b) - C(a, c)| \leq 1 + C(b, c). \tag{9.3}$$

This inequality holds for any hidden variable model which satisfies the locality assumption. For the quantum state in Equation (9.1), the inequality is violated. This is the contradiction between the quantum theory and a local hidden variable model which proves Bell's theorem.

It was noted that the assumptions which go into the hidden variable models first considered by Bell imply those models are *deterministic*. That is, the theorem did not exclude models which suggested quantum theory only provided *stochastic* (or probabilistic) information of the possible outcomes of measurements. Bell later extended the theorem to include such models. For the EPR experimental setup, let the conditional probability of outcome $A = 1$, for Alice, given the state (hidden variable) is $s \in \mathcal{S}$ be denoted $M_A(s)$ and similarly define $M_B(s)$ for Bob. Now denote the conditional joint probability of the simultaneous outcomes $A, B = 1$ by $M_{AB}(s)$. Fine [17] defines a *stochastic hidden variable model* as one which satisfies

$$\Pr(A = 1) = \int_{\mathcal{S}} d\mu(s) M_A(s) \rho(s) \tag{9.4}$$

and

$$\Pr(A = 1, B = 1) = \int_{\mathcal{S}} d\mu(s) M_{AB}(s) \rho(s). \tag{9.5}$$

If $M_{AB}(s) = M_A(s) M_B(s)$, then the model is *factorizable*. Bell claimed this also encoded the assumption of locality. Again, it can be shown that quantum theory is in contradiction with an inequality derived from these assumptions. The proof is quite simple. Fine [17] showed that a factorizable stochastic hidden variable model exists for the EPR-type correlation experiment if and only if a deterministic hidden variable model exists for the experiment. Since the latter is ruled out, the former is also ruled out.

It is often stated that *the* consequence of Bell's theorem is "quantum theory is non-local". However the theorem only states that quantum theory does not satisfy the assumptions which go into a classical model *Bell defines as local*. It is not necessary that the locality assumption is violated. Nor is it unanimously agreed that the mathematical condition Bell refers to as locality in the stochastic hidden variable models reflects any physical significance [17]. In the next section it will be shown how such a claim can be supported by appealing to the notions of negativity.

9.3 Negativity connection

Notice that a (non-factorizable) stochastic hidden variable model is exactly a classical representation. Suppose that for the EPR experiment such a classical probabilistic description is possible. Then, factorizability (Bell locality) is an additional requirement. In Section 7.3, a classical probabilistic description was deemed questionable due to Theorem 7.3.1 which ruled out a classical representation independently of any assumption of locality for the most general quantum experiment.

Theorem 7.3.1 implies that one (or more) of the constraints which go into Definition 7.1.1 are false. Indeed, relaxing the assumptions of positive probability yields a quasi-probability representation which is not in conflict with quantum theory. This poses a problem for those who hold that Bell's theorem implies locality is violated in quantum theory. The existence of *positive* probability distributions in the hidden variables Bell is trying to rule out is considered an unquestionable assumption. However, if it is the case that negative probability encodes something about Nature that is independent of locality, then it is not necessary that Bell's theorem implies locality is violated.

If Bell's theorem turned out not to have any significance on the concept of locality, this would come as a surprise to many students of quantum theory who have been taught otherwise. The current situation, however, bears a remarkable similarity to the situation Bell faced at the time he introduced his theorem. Indeed, an early argument of von Neumann claimed that no hidden variable model (local or otherwise) of quantum theory could exist. Bell was able to show that an assumption von Neumann made of hidden variable theories was too strong stating [4] "the formal proof of von Neumann does not justify his informal conclusion."

Chapter 10

Conclusions and future directions

In this thesis it has been shown that frame theory provides a formalism that unifies the known quasi-probability representations of quantum *states*. It was then shown that there exists two different ways (the *deformed* and *quasi-probability* approach) to lift a quasi-probability representation of states to a consistent and equivalent formulation of quantum theory. These quasi-probability representations of quantum states and *measurements* were then shown to require either negativity or a deformation of the rule for calculating probabilities. Thus a mathematically rigorous set of criteria have been that establish the (long suspected) connection between negativity and non-classicality.

10.1 Directions for future research

10.1.1 Infinite dimensions

While the results of this paper have been proven only for finite dimensional Hilbert spaces (while allowing for continuous representation spaces), it is conjectured that the results continue to hold also for infinite dimensional quantum systems (i.e., all separable Hilbert spaces).

For separable Hilbert spaces the definition of a frame is essentially the same [10]. However, a frame is not simply a spanning set as in the finite dimensional case. The first immediate road block in applying the methods used in this thesis to all separable Hilbert spaces is the unboundedness of operators on infinite dimensional spaces. In particular, the Hilbert-Schmidt (trace) inner product is undefined for unbounded operators such as the familiar position and momentum operators. One could however restrict the class of operators to the bounded ones as an unbounded operator can be approximated arbitrarily well by a sequence of bounded operators [14]. In any case, the set of mathematical tools [1] used to analyze frames in infinite dimensional Hilbert space is different from those used in this thesis.

10.1.2 Quantum mechanics

The scope of quantum theory that has been consider in this thesis can be thought of as *kinematical*; only the description of experimental configurations is of concern. There exist many *dynamical* approaches to quantum theory where the interest shifts to how and why quantum systems change in time. This theory is often called quantum *mechanics*.

Finite dimensional quantum mechanics differs from what is presented in this thesis (namely Definition 2.2.1) by the addition of *equations of evolution*. A significant reason why adding evolution to the frame representation formalism could be fruitful is the Wigner function analogy. Recall the original Wigner function (described in Section 3.1) for infinite dimensional Hilbert spaces. Using the Wigner function formalism to describe the dynamical transformations predicted by quantum mechanics yields the dynamical law

$$\frac{\partial \rho}{\partial t} = \{H, \rho\} + \sum_{n=1}^{\infty} \frac{1}{2^{2n}(2n+1)!} \frac{\partial^{2n+1} H}{\partial q^{2n+1}} \frac{\partial^{2n+1} \rho}{\partial p^{2n+1}}, \tag{10.1}$$

where H is the classical Hamiltonian and $\{H, \rho\}$ is the classical Poisson bracket. Notice then that Equation (10.1) is of the form "classical evolution" + "quantum correction terms". Using this formalism, one can then do more than discuss which experimental procedures are classical. Now one can discuss the *transitions* between quantum and classical descriptions, a process known as *decoherence* [52].

It would be of interest to see if an equation analogous to Equation (10.1) can be found for a frame representation of finite dimensional quantum mechanics. This is not a simple task as finite dimensional quantum systems have no classical analogue. Whereas, infinite dimensional quantum system are usually *quantized* systems of classical mechanical particles. The classical analogy is then obvious and its mathematical description is well known. The first step for the finite dimensional case is to develop appropriate classical analogies to finite dimensional quantum systems and derive the dynamical laws these classical system must satisfy. Then it can be determined if the quantum dynamical laws are represented in the frame representation as "classical evolution" + "quantum correction terms".

10.1.3 "Quantumness" of information processing tasks

The "physical system" prepared according to a set of experimental configurations may not be classical in the sense that it obeys the assumptions of objectivity and determinism, yet it is conceivable that technical constraints allows us to interpret these "physical systems" as classical. In other words, a experiment on a truly classical system can achieve the same mathematical description. It is thus of interest to known when a particular experiment with a quantum probabilistic description also has a classical probabilistic description. This question is important from a quantum information theoretic prospective since it is of great interest to know when a given communication protocol or algorithm can be simulated with classical systems. In other words, one would like to know whether a given communication protocol possess "quantumness" in the sense that it requires truly quantum resources.

The proof by contradiction of Theorem 7.3.1 establishes that a classical representation of quantum theory does not exist. This fact on its own is not very insightful. A future direction of research is to analyze quantum information tasks within the frame representation formalism with the goal of identifying the true quantumness of such tasks. Consider a given quantum information task. Such a task will not require the full state space nor the full set of possible measurements allowed by quantum theory. It is then conceivable that one can identify a frame and a dual which represents the *restricted set* of states and measurements used in a particular protocol as positive functions hence satisfying the constraints of a classical model.

More generally, it would be of interest to characterize the various combinations of subsets of $\mathrm{D}(\mathcal{H})$ and $\mathrm{E}(\mathcal{H})$ which require negativity. For example, supposes a set of experimental configuration is such that it can prepare any density operator. This leads to (Equation (7.9) in the proof of Theorem 7.2.3)

$$\hat{M} = \int_{\mathcal{S}} d\mu(s)\ M(s)\hat{F}(s).$$

The question becomes which subsets of $\mathrm{E}(\mathcal{H})$ must the experimental procedure be able to perform before $M(s)$ must have negative values? One obvious answer is $\mathrm{E}(\mathcal{H})$ itself, this is of course the statement of Theorem 7.3.1. Suppose only the trivial measurement $\hat{\mathbb{1}}$ can be performed. Then,

$$\hat{\mathbb{1}} = \int_{\mathcal{S}} d\mu(s)\ \mathbb{1}(s)\hat{F}(s).$$

From Equation (7.7), $\mathbb{1}(s) = 1$ and so long as $\hat{F}(s)$ is positive, this constitutes a classical description of the experiment. In some sense this is obvious; there is no "quantumness" in an experiment which asks "is something there". In other words, a quantum system on its own is classical; it is what one can "do" with it that is truly quantum.

Appendix A

Mathematical Background

Provided in this appendix are some brief comments to serve as an introduction to the mathematical concepts and notations used throughout this thesis. Some knowledge of the reader will be assumed. Namely, the knowledge that a typical student (say, the author of this thesis) would acquire during an undergraduate degree in mathematical physics. The mathematical concepts used in quantum information are presented in detail in Nielsen and Chaung [40].

The term Hilbert space is ubiquitous in any branch of quantum theory. In quantum information theory, where only finite dimensions are concerned, a Hilbert space is better understood as an inner product space, a concept which is assumed to be known to the reader. An attempt will be made to explain some of the notation which is commonplace in the quantum information community. The Hilbert space is a vector space (which can be over the real or complex field) along with an inner product. The Hilbert space and the underlying vector space are always given the same label, say $\mathcal{H}$. The inner product itself is always denoted $\langle\cdot,\cdot\rangle_L$, where the label L is to distinguish the underlying vector space where ambiguity is possible. More often than not, the vector space can be inferred immediately from the vectors which appear in the inner product. In such a case, the label is removed for brevity.

All finite dimensional Hilbert spaces are *isomorphic*; a bijective mapping preserving the linear and inner product structure exists between any two Hilbert spaces of the same dimension. It is assumed in quantum theory that the "quantum system" is associate with an *abstract* complex Hilbert space of a given dimension. Since all Hilbert spaces of a given dimension are isomorphic, any *concrete* Hilbert space will do. For dimension d, this concrete Hilbert space is most often taken to be $\mathbb{C}^d$ with the usual dot product (or pointwise product) as the inner product.

A.1 Linear operators

Linear operators on a complex Hilbert space $\mathcal{H}$, with dimension d, are always denoted with a "hat" $\hat{A}$. The unique operator $\hat{A}^\dagger$ implicitly defined via $\langle \hat{A}^\dagger\psi,\psi\rangle =$

$\langle \psi, \hat{A}\psi \rangle$ for all $\psi \in \mathcal{H}$ is called the adjoint of $\hat{A}$. The subset of all linear operators which satisfy $\hat{A} = \hat{A}^\dagger$ are called *Hermitian* and denoted Herm($\mathcal{H}$). The Hilbert-Schmidt or trace inner product is defined as $\langle \hat{A}, \hat{B} \rangle := \mathrm{Tr}(\hat{A}\hat{B})$. The space of Hermitian operators together with the trace inner product is a *real* Hilbert space. The dimension of this space is d^2.

Here are a few examples of a Hermitian operators important to quantum theory. Consider the operator $\hat{Z}$ whose spectrum is $\mathrm{spec}(\hat{Z}) = \{\omega^k : k \in \mathbb{Z}_d\}$, where $\omega^k = e^{\frac{i2\pi}{d}k}$. The eigenvectors form a basis for $\mathcal{H}$ and are denoted $\{\phi_k : k \in \mathbb{Z}_d\}$. Consider also the operator defined by $\hat{X}\phi_k = \phi_{k+1}$, where all arithmetic is modulo d. Define $\hat{Y}$ implicitly through $[\hat{X}, \hat{Z}] = 2i\hat{Y}$. The operators $\hat{Z}, \hat{X}$ and $\hat{Y}$ are often called *generalized Pauli operators* since they are indeed the usual Pauli operators when $d = 2$. The *parity operator* is defined by $\hat{P}\phi_k = \phi_{-k}$.

Three important subsets of the Hermitian operators are used throughout this thesis: the *projectors*, the *density operators* (also known as density matrices) and the *effects*. All three are sets of *positive* operators. A positive operators $\hat{A}$ is one that satisfies $\langle \psi, \hat{A}\psi \rangle \geq 0$ for all $\psi \in \mathcal{H}$. The short-hand $\hat{A} \geq 0$ is often used to denote a positive operator with $\hat{A} \geq \hat{B}$ meaning $\hat{A} - \hat{B} \geq 0$. A projector is an operator $\hat{P}$, such that $\hat{P}^2 = \hat{P}$. The set of projectors is denoted P($\mathcal{H}$). A density operator $\hat{\rho}$ is a positive operator with trace one. That is $\hat{\rho} \geq 0$ and $\mathrm{Tr}(\hat{\rho}) = 1$. The set of all density operators is denoted D($\mathcal{H}$). An effect $\hat{M}$ is a positive operator which is "less than" the identity. That is $0 \leq \hat{M} \leq \hat{\mathbb{1}}$. The set of all effects is denoted E($\mathcal{H}$). Note that P($\mathcal{H}$) $\subset$ D($\mathcal{H}$) $\subset$ E($\mathcal{H}$).

A.2 Convexity

The notion of convexity is vital to both classical and quantum probability. It will be defined and discussed here. For a reference see [43].

A subset C of a vector space X is *convex* if $(1-p)x + py \in C$ whenever $x \in C, y \in C$ and $0 < p < 1$. For example, the set of all classical probability vectors is convex. A *convex combination* is a sum

$$p_1 x_1 + \cdots + p_n x_n$$

in which each $p_j > 0$ and $p_1 + \cdots + p_n = 1$. An equivalent definition of a convex set is a set which contains all the convex combinations of its elements. A point x in a convex set C is called an *extreme point* if there is no way to express x as a convex combination $(1-p)y + pz$ such that $y \in C, z \in C$ and $0 < p < 1$ unless $x = y = z$. A mapping $f : C \to C'$ of one convex set into another is *convex-linear* if $f((1-p)x + py) = (1-p)f(x) + pf(y)$ for all $x \in C, y \in C$ and $0 < p < 1$.

The important examples of these notions are the projectors, density operators and effects. The sets D($\mathcal{H}$) and E($\mathcal{H}$) are convex. The projectors P($\mathcal{H}$) are extreme points of both sets D($\mathcal{H}$) and E($\mathcal{H}$).

Appendix B

Gallery of Wigner functions

If $\hat{\rho} = \psi\psi^*$ is a pure state, the the definition of the Wigner function in Equation (3.2) reduces to

$$\rho^{\text{Wigner}}(x,p) = \frac{1}{\pi}\int_{\mathbb{R}} dy\ \overline{\psi(x+y)}\psi(x-y)e^{2ipy}. \tag{B.1}$$

Some examples of Wigner functions are now given. Consider the quantum state

$$\psi(x) = \frac{1}{\sqrt{2b\sqrt{\pi}}}e^{\frac{-(x-a)^2}{2b^2}}. \tag{B.2}$$

The Wigner function of this state is depicted in Figure B.1. The Wigner function is a Gaussian distribution and quantum states of form in Equation B.2 have come to be known as Gaussian states. Notice that no negative values appear in the Wigner function. It is tempting to conclude that such a state is classical. In some sense it is; Gaussian states saturate the lower bound of the Heisenburg uncertainty principle. The state is often interpreted as a particle which is a well localized as quantumly possible. Now consider a superposition between two Gaussian states

$$\psi(x) = \frac{1}{\sqrt{2b\sqrt{\pi}(1+e^{-a^2})}}\left(e^{\frac{-(x-a)^2}{2b^2}} + e^{\frac{-(x+a)^2}{2b^2}}\right). \tag{B.3}$$

The Wigner function of this state is depicted in Figure B.2. This state is often called a *cat state* as it is interpreted as the superposition of two macroscopically distinct states. Notice the negative values. This has prompted some authors to suggest that negativity in the Wigner function is the signature of "quantumness". A final example is the $n = 4$ eigenstate of a harmonic oscillator

$$\psi(x) = \left(\frac{m\omega}{576\pi}\right)^{\frac{1}{4}} e^{\frac{-m\omega x^2}{2}}(4(\sqrt{m\omega}x)^4 - 12(\sqrt{m\omega}x)^2 + 3). \tag{B.4}$$

The Wigner function of this state is depicted in Figure B.3. Notice again the negative values. It was shown that negativity is quite ubiquitous; the only quantum states which have positive Wigner function are Gaussian states [30]. For the interested reader, Figures B.1-B.3 can be reproduced using the following MATLAB script.

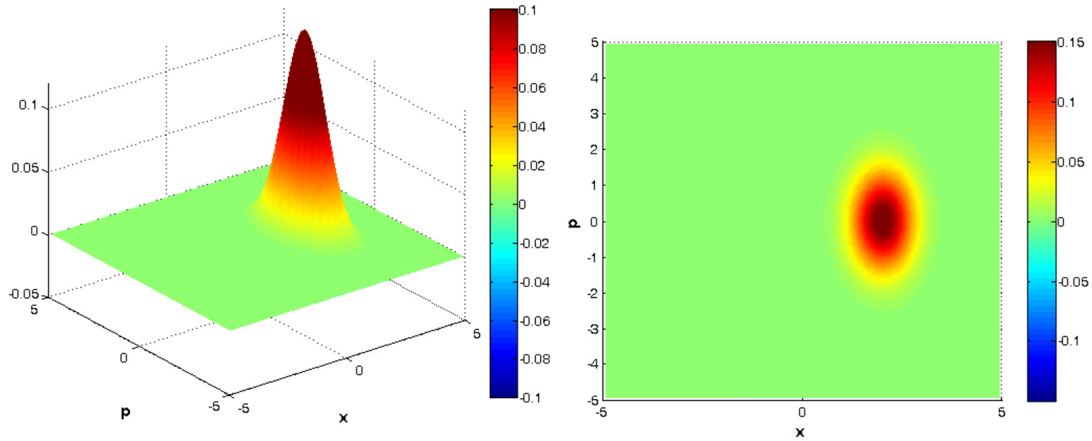

Figure B.1: Wigner function of the Gaussian state (B.2) for $a = 2$ and $b = 1$.

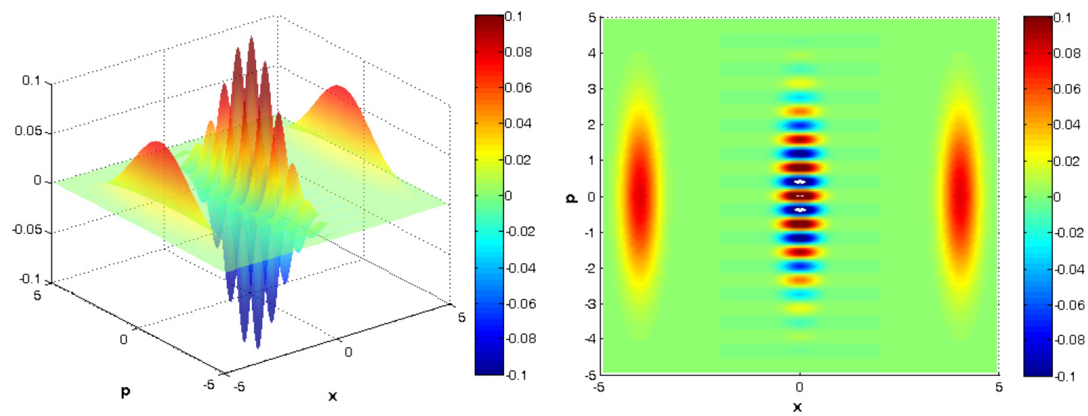

Figure B.2: Wigner function of the cat state (B.3) for $a = 4$ and $b = 0.5$.

```
%physical constants
m=1; omega=1; hbar=1;

%constants defining the Gaussian IC
a=4; b=1;

% number of grid points (total is N+1) and the grid
N=128; z=linspace(-10,10,N+1); [y,x,q]=meshgrid(z,z,z);

%initial wave functions (choose one!)

%classical gaussian
psi=@(q)1/sqrt(2*b*sqrt(pi))*(exp(-(q-a).^2/(2*b^2)));

%cat state
```

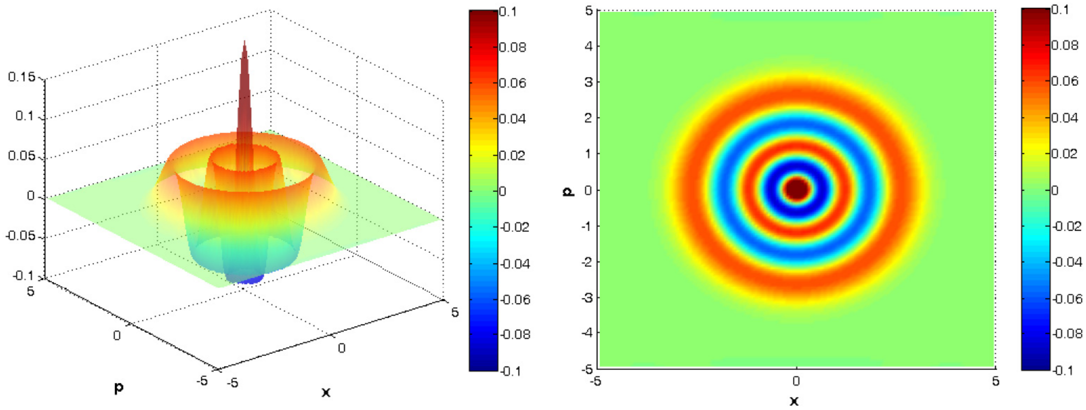

Figure B.3: Wigner function of a harmonic oscillator eigenstate ($n = 4$) (B.4) for $m = \omega = 1$.

```
% psi=@(q)1/sqrt(2*b*sqrt(pi)*(1+exp(-a^2)))*(exp(-(q-a).^2/(2*b^2))+...
%      exp(-(q+a).^2/(2*b^2)));

%harmonic oscillator eigenstate n=4
% psi=@(q)(m*omega/hbar/pi/576)^(1/4)*exp(-m*omega*q.^2/(2*hbar)).*...
%      (4*(sqrt(m*omega/hbar)*q).^4-12*(sqrt(m*omega/hbar)*q).^2+3);

%Wigner function of initial wave function
u0 =1/2/pi/hbar *
trapz(z,(exp(-i*y.*q)/hbar).*conj(psi(x+q/2)).*psi(x-q/2),3);

%plot
figure(1)
clf
caxis([-0.1,0.1]) axis([-10 10 -10 10 -0.05 0.1])
colorbar;
xlabel('x','fontweight','bold','fontsize',12);
ylabel('p','fontweight','bold','fontsize',12);
surf(x,y,u0);
shading interp;alpha(.4);
```

References

[1] S. T. Ali, J.-P. Antoine, and J.-P. Gazeau. Continuous frames in hilbert space. *Ann. Phy.*, 222:1, 1993. 59

[2] D. Appleby, H. B. Dang, and C. A. Fuchs. Physical significance of symmetric informationally-complete sets of quantum states, 2007. http://arxiv.org/abs/quant-ph/0707.2071v1. 40

[3] G. A. Baker. Formulation of quantum mechanics based on the quasi-probability distribution induced on phase space. *Phys. Rev.*, 109:2198, 1958. 1, 10

[4] J. Bell. *Speakable and Unspeakable in Quantum Mechanics*. Cambridge University Press, Cambridge, 1987. 3, 53, 56, 58

[5] J. Bertrand and P. Bertrand. A tomographic approach to wigner's function. *Foundations of Phyiscs*, 17:397, 1987. 11

[6] D. Bohm. *Quantum Theory*. Prentice-Hall, New York, 1951. 53, 56

[7] B. Bransden and C. Joachain. *Quantum Mechanics*. Pearson Education Limited, London, 2000. 6, 55

[8] P. Busch. Quantum states and generalized observables: A simple proof of gleasons theorem. *Phys. Rev. Lett.*, 91:120403, 2003. 44

[9] S. Chaturvedi, E. Ercolessi, G. Marmo, G. Morandi, and R. Simon. Wigner-weyl correspondence in quantum mechanics for continuous and discrete systems - a dirac-inspired view. *J. Phys. A: Math. Gen.*, 39:1405, 2006. 22

[10] O. Christensen. *Introduction to Frames and Riesz Bases*. Birkhäuser, Boston, 2003. 59

[11] O. Cohendet, P. Combe, M. Sirugue, and M. Sirugue-Collin. A stochastic treatment of the dynamics of an integer spin. *J. Phys. A: Math. Gen.*, 21:2875, 1987. 15, 22

[12] C. Cormick, E. F. Galvao, D. Gottesman, J. P. Paz, and A. O. Pittenger. Classicality in discrete wigner functions. *Phys. Rev. A*, 73:012301, 2006. 2, 3

[13] L. de Broglie. *J. de Phys.*, 8:225, 1927. 53

[14] D. A. Dubin, M. A. Hennings, and T. B. Smith. *Mathematical aspects of Weyl quantization and phase*. World Scientific, Singapore, 2000. 59

[15] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777, 1935. 55

[16] C. Ferrie and J. Emerson. Frame representations of quantum mechanics and the necessity of negativity in quasi-probability representations. *J. Phys. A: Math. Theor.*, 41:352001, 2008. iii, 28, 47

[17] A. Fine. Hidden variables, joint probabilities and the bell inequalities. *Phys. Rev. Lett.*, 48:291, 1982. 57

[18] E. F. Galvao. Discrete wigner functions and quantum computational speedup. *Phys. Rev. A*, 71:042302, 2005. http://arxiv.org/abs/quant-ph/0405070v2. 2

[19] M. Genovese. Research on hidden variable theories: a review of recent progesses. *Physics Reports*, 413:319, 2005. 8

[20] K. S. Gibbons, M. J. Hoffman, and W. K. Wootters. Discete phase space based on finite fields. *Phys. Rev. A*, 70:062101, 2004. http://arxiv.org/abs/quant-ph/0401155v6. 2, 18, 20, 21, 22

[21] H. Goldstein, C. Poole, and J. Safko. *Classical Mechanics (3rd edition)*. Addison Wesley, San Francisco, 2001. 4

[22] D. Gottesman. *Stabilizer Codes and Quantum Error Correction*. PhD thesis, California Institute of Technology, 1997. 3

[23] D. Gross. Hudson's theorem for finite-dimensional quantum systems. *J. Math. Phys.*, 47:122107, 2006. http://arxiv.org/abs/quant-ph/0602001v3. 22

[24] D. Gross and J. Eisert. Quantum margulis expanders, 2007. http://arxiv.org/abs/quant-ph/0710.0651v1. 2

[25] L. Hardy. Quantum theory from five reasonable axioms, 2001. http://arxiv.org/abs/quant-ph/0101012v4. 7, 21, 22

[26] T. F. Havel. The real density matrix. *Quantum Information Processing*, 1:511, 2003. http://arxiv.org/abs/quant-ph/0302176v5. 21, 22

[27] S. Heiss and S. Weigert. Discrete moyal-type representations for a spin. *Phys. Rev. A*, 63:012105, 2000. 16, 17, 18, 22

[28] M. Hillery, R. O'Connell, M. Scully, and E. Wigner. Distrubution functions in physics: fundamentals. *Physics Reports*, 106:121, 1984. 11

[29] M. Horodecki, P. W. Shor, and M. B. Ruskai. Entanglement breaking channels. *Rev. Math. Phys.*, 15:269, 2003. http://arxiv.org/abs/quant-ph/0302031v2. 47

[30] R. Hudson. When is the wigner quasi-probability density non-negative. *Rep. Math. Phys.*, 6:249, 1974. 64

[31] C. J. Isham. *Lectures on Quantum Theory: Mathematical and Structural Foundations.* Imperial College Press, London, 1995. 50

[32] S. Kochen and E. Specker. The problem of hidden variables in quantum mechanics. *J. Math. Mech.*, 17:59, 1967. 49

[33] J. Kovacevic and A. Chebira. Life beyond bases: the advent of frames (part i). *IEEE Signal Processing Mag.*, 24:86, 2007. 25

[34] H.-W. Lee. Theory and application of the qunatum phase-space distrbution functions. *Phys. Rep.*, 259:147, 1994. 1, 10

[35] U. Leonhardt. Quantum state tomography and discrete wigner functions. *Phys. Rev. Lett.*, 74:4101, 1995. 15, 22

[36] U. Leonhardt. Discrete wigner functions and quantum state tomography. *Phys. Rev. A*, 53:2998, 1996. 15

[37] C. C. López and J. P. Paz. Phaes-space approach to the study of decoherence in quantum walks. *Phys. Rev. A*, 68:0052305, 2003. http://arxiv.org/abs/quant-ph/0308104v3. 2

[38] C. Miquel, J. P. Paz, and M. Saraceno. Quantum computers in phase space. *Phys. Rev. A*, 65:062309, 2002. 2

[39] J. E. Moyal. Quantum mechanics as a statistical theory. *Proc. Camb. Phil. Soc.*, 45:99, 1949. 10

[40] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information.* Cambridge University Press, Cambridge, 2000. 62

[41] J. P. Paz. Discrete wigner functions and the phase-space representation of quantum teleportation. *Phys. Rev. A*, 65:062311, 2002. http://arxiv.org/abs/quant-ph/0204150v1. 2

[42] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves. Symmetric informationally complete quantum measurements. *J. Math. Phys.*, 45:2171, 2004. http://arxiv.org/abs/quant-ph/0310075v1. 39, 40

[43] R. T. Rockafellar. *Convex Analysis.* Princeton University Press, Princeton, 1970. 63

[44] R. Spekkens. Negativity and contextuality are equivalent notions of nonclassicality, 2007. http://arxiv.org/abs/quant-ph/0710.5549v1. 51

[45] R. W. Spekkens. Contextuality for preparations, transformations, and unsharp measurements. *PRA*, 71:052108, 2005. 50, 51

[46] R. L. Stratonovich. *JETP*, 4:891, 1957. 16, 22

[47] A. Vourdas. Quantum systems with finite hilbert space. *Rep. Prog. Phys.*, 67:267, 2004. 21

[48] R. F. Werner. Quantum states with einstein-podolsky-rosen correleations admitting a hidden-variable model. *Phys. Rev. A*, 40:4277, 1989. 47

[49] E. Wigner. On the quantum correction for thermodynamic equilibrium. *Phys. Rev.*, 40:0749, 1932. 1, 11

[50] W. K. Wootters. A wigner-function formulation of finite-state quantum mechanics. *Ann. Phy.*, 176:1, 1987. 2, 12, 14, 22

[51] W. K. Wootters. Picturing qubits in phase space. *IBM J. Rev. & Dev.*, 48:99, 2004. 18

[52] W. H. Zurek. Decoherence, einselection, and the quantum origins of the classical. *Rev. Mod. Phys.*, 75:715, 2003. http://arxiv.org/abs/quant-ph/0105127v3. 60