On the Significance of Quasi-Probability in Quantum Mechanics

Kazuki Fukuda

Doctor of Philosophy

Department of Particle and Nuclear Physics
School of High Energy Accelerator Science
SOKENDAI (The Graduate University for Advanced Studies)

On the Significance of Quasi-Probability in Quantum Mechanics

Kazuki Fukuda

A dissertation submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

Department of Particle and Nuclear Physics
School of High Energy Accelerator Science
The Graduate University for Advanced Studies (SOKENDAI)

2015

Abstract

We present a theoretical study of quasi-probability based on the weak value and weak measurement with an aim to show its significance in quantum mechanics. First, we consider the general aspects of the quasi-probability underlying the weak value, which can be determined from the weak value up to a certain ambiguity parameterized by a complex number. We argue the legitimacy and the usefulness of the quasi-probability in quantum mechanics by showing that it has an analogous as the conventional probability. Next, we consider the post-selected measurement in operational probabilistic theory and show that the quasi-probability underlying the weak value has a epistemological feature. Finally, we consider a novel class of ontological models of quantum theory and uncover the conceptual utility of the quasi-probability. In particular, we show that, with the help of the quasi-probability, Bohmian mechanics can be regarded as an ontological model with a certain type of contextuality.

Acknowledgements

I would like to express my special appreciation and thanks to Dr. Izumi Tsutsui who during these five years has been a tolerant teacher. I greatly benefited from his clear thinking and from his originality. His insightful advice on writing, presentation and careers has proved invaluable. I would like to thank Tatsuya Morita, Takuya Mori, Jaeha Lee and Ryosuke Koganezawa for many useful discussions. Finally, I would like to thank my parents, Katsumi and Fumiko Fukuda, for their support throughout my education.

Contents

Chapter 1

Introduction

Quantum mechanics is a peculiar theory of physics. Ever since it was invented in early twentieth century, it has been verified to a very high accuracy and has never been questioned by any contradictions with experiments. Moreover, it is now widely regarded as the most fundamental theory that underlies almost all branches of modern physics except for gravity. Despite its success, there is still no consensus among physicists about what this theory really means. The existing several interpretations of quantum mechanics expresses this fact. There is no question that quantum mechanics works well as a tool for predicting what will occur in experiments. But once we try to understand this theory, we encounter the so-called quantum weirdness, the strange feature of quantum mechanics. The field of study which seeks to achieve the understanding of quantum weirdness is called the foundation of quantum mechanics or quantum foundations. The research of quantum foundations may also be essential for answering several basic questions in contemporary physics. For instance, since the early universe may be extremely small and should be described by quantum mechanics, the study of quantum foundations should be important. Moreover, quantum foundations may be crucial to find solutions for constructing a consistent quantum gravity theory.

Almost all quantum weirdness may be ascribed to the quantum feature of non-classicality, and we are still not sure whether the fundamental notions that are valid in classical mechanics can still be established in quantum mechanics or not. These notions are, for instance, physical reality, determinism, locality, time reversibility, and so on. Understanding the fundamental aspects of quantum theory may be achieved by learning whether or not these notions are valid within the current quantum theory, or whether we can discover a new theory which establishes firmly these notions and reproduces all predictions of quantum mechanics.

One of the important indications of quantum weirdness was uncovered by the so-called EPR paradox [1] proposed by Einstein, Podolsky and Rosen in 1935. They demonstrated that the wave function may not completely describe the elements of physical reality and stated that quantum mechanics may possibly be incomplete. The concluding sentence of their paper is;

> While we have shown that the wave function does not provide a complete description of the physical reality, we left open the question of whether or not such a description exists. We believe, however, the such a theory is possible.

Here, “such a theory” in the last sentence is sometimes called ‘hidden variable theory’ or in modern terms, ‘ontological model’, and Bohmian mechanics [2, 3] is a well-known example. Meanwhile, no-go theorems, which insist that a hidden variable theory satisfying conditions such as reality, locality or determinism contradicts quantum theory, had been put forward in various forms in several occasions. A notable example is Bell’s theorem introduced by Bell in 1964 [4]. This theorem states that a theory possessing some both of physical reality and locality cannot reproduce all predictions of quantum mechanics. Other famous examples are von-Neumann’s no-go theorem in 1932 [5] and Kochen-Specker theorem [6] in 1967.

Recently, along with the development of quantum information science and the improvement of experimental technology in microscopic regime, quantum foundations are being animated and developing with a renewed momentum. For example, the introduction of the ontological model by Spekkens in 2005 [7] and 2010 [8] is one of the recent developments in foundations of quantum mechanics. The ontological model is a novel framework of hidden variable theories of quantum mechanics, which incites several remarkable works including Pusey, Barrett and Rudolph (PBR) theorem [9] that shows that the quantum state must be ontic (i.e. a state of reality) if there exists an ontological model of quantum mechanics. There remains, however, a question whether their ontological model is completely general or not. Specifically, it has been pointed out that it cannot accommodate Bohmian mechanics despite its historical importance [10]. Another example of recent developments in quantum foundations is the weak value and the weak measurement introduced by Aharonov, Albert and Vaidman in 1988 [11]. The weak value is a new physical observable in quantum mechanics and a weak measurement is a method for detecting the weak value, both of which have a close connection with time symmetry in quantum mechanics. The application of the idea of the weak value has been studied extensively over the years, for instance, for solving Hardy's paradox [12, 13] and amplification for precision measurement [14, 15, 16]. Among them, however, the most fundamental is the finding that the weak value may motivate 'quasi-probability'.

Quasi-probability is an extended notion of probability. It does not satisfy all Kolmogorov's probability axioms [17] that the standard probability satisfies, and taking even a complex number, for instance. Historically, the first example of quasi-probability was introduced by Wigner [18] in 1932 in an attempt to construct a phase space distribution in quantum mechanics. This particular quasi-probability distributions is now called Wigner distribution. The well-known function called the Husimi Q function [19] proposed by Husimi in 1940 is another example. Unlike the Wigner distribution, it takes exclusively real positive values but may become greater than 1. Kirkwood-Dirac function [20, 21] is also an example. Recently, the topic of quasi-probability has become poplar along with the weak value and the weak measurement. Steinberg indicated its theoretical connection with the weak value in 1995 [22, 23], while Ozawa studied it from the perspective of simultaneous measurability [24] in 2011. The connection between time reversibility and quasi-probability is discussed by Hofmann [25] in 2012 and Morita *et al.* [26] in 2013, while the experimental realization of Kirkwood-Dirac distribution was demonstrated by Lundeen-Bamber [27] in 2014.

The central theme of this thesis is the study of the fundamental role of quasi-probability in quantum mechanics. We have presented theoretical studies of quasi-probability in quantum mechanics in topics selected from the weak value, operational theories, and the ontological model. We shall find through these studies that that quasi-probability has an inherent significance in quantum mechanics. It also provides us with a novel perspective in foundations of quantum mechanics. An overview of this thesis is as follows.

In chapter 2, we argue that legitimacy and usefulness of quasi-probability underlying the weak value in quantum mechanics. We will find a harmony between quasi-probability and the structure of quantum mechanics, in the sense that it is legitimate to accept the quasi-probability as a fundamental element of quantum mechanics and that quasi-probability is useful quantum mechanics. The quasi-probability underlying the weak value is found to be an extension of probability in quantum mechanics. We recall in subsection 2.2 that this extension is naturally brought by the generalized Gleason's theorem in quantum mechanics with some consistency condition. This extension admits an intrinsic ambiguity expressed by a complex valued parameter α. We investigate the relationship between the quasi-probability with the parameter α and the structure of quantum mechanics. In subsection 2.2.2, motivated by this α-extension, we introduce generalized products of quantum mechanical observables called α -products. It is seen that the joint quasi-probability distribution given by the α-products of observables can be regarded as a simultaneous quasi-probability distribution of incompatible observables, on account of the fact that its marginal probability gives the Born rule and the joint quasi-probability distribution of compatible observables coincide with the joint probability distribution for compatible observables. These are described in subsection 2.2.2 and 2.2.3. It will be noticed that the α -products include the well-known Jordan product for the special case (α =1/2), and from this we find that the physical significance of Jordan products can be given through the extension of probability. In section 2.3, we review the weak measurement procedure and thereby confirm the the measurability of the weak value-based quasi-probability theoretically.

In chapter 3, we attempt to reinforce the epistemological significance of the post-selected measurement from the operational probabilistic theories. Here, we review operational probabilistic theories for two reasons; first, for investigating the relationship between the ontological model of quantum mechanics and quasi-probability, we need operational probabilistic theories as a theoretical basis of the ontological model. This allows us to examine the physical reality from the epistemological point of view. Secondly, we need to analyze the epistemological significance of a post-selected measurement to find out a proper the quasi-probability in the weak measurement procedure. In section 3.1 we review the elements of operational probabilistic theories. In section 3.2, the definition of the simultaneous measurability and determinism is introduced and operational theories are classified with simultaneous measurability as the basic concept. This provides a clue as to what properties the general hidden variable theories should have. In section 3.3, we mention some examples of operational theories in physics. The concept of entanglement is introduced in section 3.4. At the end of this chapter, we show the epistemological significance of post-selected measurement, which is one of the key ideas of weak measurement, in terms of the time direction of inference in the probabilistic theory.

In chapter 4, we discuss the conceptual significance of quasi-probability in quantum mechanics. We show that by regarding the quasi-probability introduced in chapter 2 as a fundamental element of quantum mechanics, one can interpret the quantum mechanics realistically in a natural way. Also, it is shown that the weak value-based quasi-probability sheds a new light on Bohmian mechanics [2, 3]. We embed the notion of quasi-probability into Bohmian mechanics in section 4.2. This integration enables us to understand Bohmian mechanics as an ontological model with a proper extension, which is required to reconcile Bohmian mechanics with ontological model. In section 4.3, after reviewing the framework of the ontological model, we present a new type of ontological models which possess some kind of contextuality which we shall call 'synlogicality'. After this, we construct general ontological frameworks of quantum theory based on the quasi-probability and synlogicality. We shall then see that Bohmian mechanics can be regarded as a quasi-probabilistic synlogical model of quantum mechanics.

Finally, chapter 5 is devoted to our conclusion and the discussion of the future research which they suggest.

Chapter 2

Quasi-Probability, Weak Value and Weak Measurement

2.1 Quasi-Probability in Quantum Mechanics

In quantum mechanics, it is a major premise that a physical system cannot simultaneously have a well defined position and momentum. Due to this, one cannot give the operational definition of a joint probability that a system has a position and a momentum, i.e., one cannot define a true phase space probability distribution for a quantum mechanical system. Nonetheless, functions which have some similarity to classical phase space distribution, *"quasi-probability distribution"*, have been proposed in quantum mechanics. The term quasi-probability refers to the fact that these functions do not necessarily satisfy all Kolmogorov's probability axioms [17]. For this reason, some physicist calls it the *extended probability* [29].

First, we shall introduce the mathematical definition of standard probability and quasi-probability. The standard mathematical theory of probability is mostly based on *a probability space*, which is a triple $(\Omega, \mathcal{F}, \mu)$ where Ω is a set, $\mathcal{F} \subset 2^{\Omega}$ is a subset of power set 2^{Ω}, and $\mu : \mathcal{F} \to \mathbb{R}$ is a map that assigns probabilities to $F \in \mathcal{F}$ satisfying following conditions;

P1 (Positivity)

$$\mu(F) \in [0, 1] \tag{2.1}$$

for all $F \in \mathcal{F}$.

P2 (Countable additivity)

$$\mu(F) = \sum_i \mu(F_i) \tag{2.2}$$

for any mutually disjoint sequence $F_1, F_2, \ldots$ with $F = \cup_i F_i$.

P3 (Normalization condition)

$$\mu(\Omega) = 1. \tag{2.3}$$

The map $\mu : \mathcal{F} \to [0, 1]$ satisfying P1, P2, and P3 is called a *probability measure*. These assumptions are called Kolmogorov's probability [17] axioms.

Next, we give a mathematical definition of the quasi-probability. The quasi-probability is defined by eliminating the condition P1 of the standard Kolmogorov's axioms. Therefore, A *quasi-probability measure* is a map $\tilde{\mu} : \mathcal{F} \to \mathbb{C}$ of $\mathcal{F}$ into a complex plane $\mathbb{C}$ with P2 and P3 where $(\Omega, \mathcal{F})$ are same sets as standard probability space;

Q1 (Complex valued)

$$\tilde{\mu}(F) \in \mathbb{C} \tag{2.4}$$

for all $F \in \mathcal{F}$.

Q2 (Countable additivity)

$$\tilde{\mu}(F) = \sum_i \tilde{\mu}(F_i) \tag{2.5}$$

for any mutually disjoint sequence $F_1, F_2, \ldots$ with $F = \cup_i F_i$.

Q3 (Normalization condition)

$$\tilde{\mu}(\Omega) = 1. \tag{2.6}$$

From this definition, the quasi-probability measure is also referred as a normalized complex measure.

We shall now recall some historical examples of quasi-probability in quantum mechanics. The well-known Wigner function [18] was proposed to construct a quantum mechanical analogue to a classical phase space density. It should be noted that the motivation to understand the quantum mechanics in analogue to a classical mechanics is similar to it of the hidden variable theory. We shall consider the connection between the quasi-probability and the hidden variable theory in chapter 4.

Let us denote the position and momentum operators by $\mathbf{Q}$ and $\mathbf{P}$, respectively. These operators satisfy the canonical commutation relation;

$$[\mathbf{Q}, \mathbf{P}] := \mathbf{QP} - \mathbf{PQ} = i \tag{2.7}$$

where we put $\hbar = 1$. Wigner function [18] is defined by

$$\mu_\rho^{\mathrm{W}}(\mathbf{q}, \mathbf{p}) := \int_{\mathbb{R}} \langle \mathbf{q} - y | \rho | \mathbf{q} + y \rangle e^{2i\mathbf{p}y} \mathrm{d}y \tag{2.8}$$

where $\mathbf{q}$ and $\mathbf{p}$ are eigenvalues of $\mathbf{Q}$ and $\mathbf{P}$. This function can be rewritten as

$$\mu_\rho^{\mathrm{W}}(\mathbf{q}, \mathbf{p}) := \mathrm{Tr}\left[F^{\mathrm{W}}(\mathbf{q}, \mathbf{p})\, \rho\right] \tag{2.9}$$

where $F^{\mathrm{W}}(\mathbf{q}, \mathbf{p})$ is an operator defined by

$$F^{\mathrm{W}}(\mathbf{q}, \mathbf{p}) = \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{i\xi(\mathbf{Q}-\mathbf{q}) + i\eta(\mathbf{P}-\mathbf{p})} \mathrm{d}\xi \mathrm{d}\eta. \tag{2.10}$$

This function is both positive and negative in general. Another well-known example of the quasi-probability is the Husimi function [19] which is defined by

$$\mu_\rho^{\mathrm{H}}(\mathbf{q}, \mathbf{p}) := \mathrm{Tr}\left[F^{\mathrm{H}}(\mathbf{q}, \mathbf{p})\, \rho\right] \tag{2.11}$$

where

$$F^{\mathrm{H}}(\mathbf{q}, \mathbf{p}) = \frac{1}{(2\pi)^2} \int_{\mathbb{R}^2} e^{i\xi(\mathbf{Q}-\mathbf{q}) + i\eta(\mathbf{P}-\mathbf{p})} e^{-\frac{1}{4}\left(\xi^2 + \eta^2\right)} \mathrm{d}\xi \mathrm{d}\eta. \tag{2.12}$$

The Husimi function takes only positive value but may become greater than one. The important property of these quasi-probability distributions is that they reproduce quantum mechanical prediction of expectation value;

$$\begin{aligned}\langle A\rangle_{\rho} &= \int A(\mathbf{q},\mathbf{p})\,\mu_{\rho}^{\mathrm{W}}(\mathbf{q},\mathbf{p})\,\mathrm{d}\mathbf{q}\mathrm{d}\mathbf{p} \\ &= \int A(\mathbf{q},\mathbf{p})\,\mu_{\rho}^{\mathrm{H}}(\mathbf{q},\mathbf{p})\,\mathrm{d}\mathbf{q}\mathrm{d}\mathbf{p}\end{aligned}$$

where $A(\mathbf{q},\mathbf{p})$ is the classical function corresponding to operator A, i.e., the function that before quantization. The Dirac-Kirkwood function [20, 21] is also an important example of quasi-probability in quantum mechanics which has the simple form such that

$$\mu_{\rho}^{\mathrm{DK}}(\mathbf{q},\mathbf{p}) := \mathrm{Tr}\left[E^{\mathbf{Q}}(\mathbf{q})\,E^{\mathbf{P}}(\mathbf{p})\,\rho\right] \tag{2.13}$$

where $E^{\mathbf{Q}}(\mathbf{q})$ and $E^{\mathbf{P}}(\mathbf{p})$ are the spectral projection of position operators $\mathbf{Q}$ and $\mathbf{P}$ corresponding to eigenvalues $\mathbf{q}$ and $\mathbf{p}$. These can be written as $E^{\mathbf{Q}}(\mathbf{q}) = |\mathbf{q}\rangle\langle\mathbf{q}|$ in the Dirac bra-ket notation. It can be easily checked that these functions are quasi-probability distribution i.e., satisfying the conditions Q1, Q2 and Q3.

It may be an important question how the quasi-probability in quantum mechanics has to be directly interpreted. It may be then beneficial to cite some statements dealing with the interpretation of quasi-probability although we shall not address the problem of the direct interpretation of quasi-probability in this thesis. For instance, Shimony have stated that non-negativeness of probability is essential in any way of interpretation of probability and hence he is negative about negative probability ;

> *I am negative about negative probability. There are two generic concepts of probability. One is epistemic, and includes personal probability and logical probability. In both cases it is possible to give a fundamental proof for the non-negativeness of probability, once reasonable assumptions are made. The other is ontic, which includes the frequency and the propensity concepts — the latter postulating that probability makes sense in an individual case, but that the evidence for a propensity comes from an ensemble. Because of the role of frequencies for both versions of ontic probability, non-negativeness is essential. I do believe that quantum mechanics makes us change our fundamental concepts. It makes us introduce objective in-*

definiteness and entanglement. My main worry about negative probabilities is that it uses a rather formal device to shield us from facing the radical metaphysical consequences of quantum mechanics.

On the other hand, Feynman insists the importance of negative probability in quantum mechanics [28] :

The only difference between a probabilistic classical word and the equations of quantum world is that somehow or other it appears as if the probabilities would have to go negative, and that we do not know, as far as I know, how to simulate.

The other statements to quasi-probability by several physicists and another examples of quasi-probability (e.g., application of quasi-probability to EPR paradox) can be seen in the excellent review paper [29]. It is notable that the some type of quasi-probability, that is, the quasi-probability underlying the weak value, becomes to be measurable and, in fact, the measurement has been realized in experiments [27, 30].

2.2 Quasi-Probability and Weak Value

We have found that the usual measuring procedure for preselected and postselected ensembles of quantum systems gives unusual results. Under some natural conditions of weakness of the measurement, its result consistently defines a new kind of value for a quantum variable, which we call the weak value. (Yakir Aharonov, *et al.* 1988)

The concept of weak value first captured the attention of physics community with publication of the seminal paper of Aharonov, Albert and Vaidman [11], wherein it provides new kind of value for a quantum variable, *weak value*. They demonstrate how it is measured and how it takes strange values, a complex number. The weak value of the observable A is defined with two quantum states (ψ, ϕ) on $\mathcal{H}$. We shall denote the weak value of A with two states (ψ, ϕ) by $\langle A \rangle^{\mathrm{w}}_{\psi\phi}$. The explicit form of $\langle A \rangle^{\mathrm{w}}_{\psi\phi}$ is given by

$$\langle A \rangle^{\mathrm{w}}_{\psi\phi} = \frac{\langle \phi | A | \psi \rangle}{\langle \phi | \psi \rangle} \tag{2.14}$$

where the denominator is assumed to be non-vanishing; $\langle\phi|\psi\rangle \neq 0$. From this, it is obvious that this quantity in general takes a complex number. It is, however, measurable via a weak measurement procedure [11], which we give a brief review at the end of this chapter. The weak value is also under investigation on its own, especially in the conceptual aspect regarding its physical reality (For the detail, See Refs. [31, 32]). Concerning this, in what follows we shall focus on the implication of the weak value in the extension of probability in quantum mechanics.

Let us consider the spectrum decomposition of the observable $A = \int aE^A(a)\,\mathrm{d}a$, where $E^A(a) = |a\rangle\langle a|$ is the spectral projection with $|a\rangle$ is an eigenstate of $A|a\rangle = a|a\rangle$. With this decomposition, the weak value $\langle A\rangle^{\mathrm{w}}_{\psi\phi}$ may be written as

$$\langle A\rangle^{\mathrm{w}}_{\psi\phi} = \int a\frac{\langle\phi| E^A(a)|\psi\rangle}{\langle\phi|\psi\rangle}\mathrm{d}a. \tag{2.15}$$

Denoting the weak value of projection operator as

$$q(a\,|\psi,\phi) = \frac{\langle\phi| E^A(a)|\psi\rangle}{\langle\phi|\psi\rangle}, \tag{2.16}$$

the weak value becomes

$$\langle A\rangle^{\mathrm{w}}_{\psi\phi} = \int aq(a\,|\psi,\phi)\,\mathrm{d}a. \tag{2.17}$$

The expression (2.16) suggests that the the function $a \mapsto q(a\,|\psi,\phi)$ may be interpreted as an analogue of probability in that its average yields the weak value (2.17), even though the value $q(a\,|\psi,\phi)$ may go beyond the standard range of probability $[0,1]$ or even becomes complex. In the next subsection, we shall see that the function $a \mapsto q(a\,|\psi,\phi)$ is actually the quasi-probability measure introduced in the previous section. Before we mention it, we briefly sketch how it appears naturally in quantum mechanics [26] in next subsection. It should be noted that the quantity defined in equation (2.16) is applied to the explanation at the strange quantum phenomena such as [33, 35] which has become famous recently.

2.2.1 Conditional Quasi-Probability Distribution

Let $\nu_{(\psi,\phi)} : \mathcal{P} \to \mathbb{C}$ be a map of the collection $\mathcal{P}$ of projection operators on a Hilbert space $\mathcal{H}$ (finite dimension) into complex plane $\mathbb{C}$ for $\psi, \phi \in \mathcal{H}$ ($\langle\phi|\psi\rangle \neq 0$). We require that a map $\nu_{(\psi,\phi)}$ satisfies the following two conditions. First requirement is that a map $\nu_{(\psi,\phi)}$ is partially additive;

$$\nu_{(\psi,\phi)}\left(\sum_i P_i\right) = \sum_i \nu_{(\psi,\phi)}\left(P_i\right) \tag{2.18}$$

for $\{P_i\} \subset \mathcal{P}$ which are mutually orthogonal $P_i P_j = O$ (O is null operator) for $i \neq j$. This requirement is tantamount to the condition of the Gleason's theorem, which proves that the Born rule follows from the partially additive condition and the (one-state) normalization condition, except that our condition is for complex measure whereas Gleason's one is for standard probability measure. The generalization of Gleason's theorem to complex measure [34] guarantees that there exists the trace class operator W such that

$$\nu_{(\psi,\phi)}\left(P\right) = \mathrm{Tr}\left[PW\right] \tag{2.19}$$

for $\dim\left(\mathcal{H}\right) \geq 3$. Next, we require that a map $\nu_{(\psi,\phi)}$ satisfies two-state normalization conditions;

$$\nu_{(\psi,\phi)}\left(P_\psi\right) = 1, \tag{2.20}$$

$$\nu_{(\psi,\phi)}\left(P_{\psi^\perp}\right) = 0 \tag{2.21}$$

$$\nu_{(\psi,\phi)}\left(P_\phi\right) = 1, \tag{2.22}$$

$$\nu_{(\psi,\phi)}\left(P_{\phi^\perp}\right) = 0 \tag{2.23}$$

where $P_\psi = |\psi\rangle\langle\psi|$ and $\left|\psi^\perp\right\rangle$ and $\left|\phi^\perp\right\rangle$ are an arbitrary unit vectors on $\mathcal{H}$ such that $\left\langle\psi^\perp|\psi\right\rangle = 0$ and $\left\langle\phi^\perp|\phi\right\rangle = 0$.

We call equations (2.20), (2.21), (2.22) and (2.23) consistency conditions since these are the analogy of single state Gleason's theorem with standard probability. Morita, *et al.* proved the following theorem [26];

Theorem 2.1. *If a map $\nu_{(\psi,\phi)}$ for $\dim(\mathcal{H}) \geq 3$ satisfies the partial additivity (2.18) and the consistency condition, then it has the form;*

$$\nu_{(\psi,\phi)}(P) = \alpha \frac{\langle\phi| P |\psi\rangle}{\langle\phi|\psi\rangle} + (1-\alpha)\frac{\langle\psi| P |\phi\rangle}{\langle\psi|\phi\rangle} \tag{2.24}$$

for some $\alpha \in \mathbb{C}$.

A sketch of the proof runs as follows. Using equation (2.19), consistency conditions can be written as

$$\mathrm{Tr}\left[P_\psi W\right] = 1, \tag{2.25}$$

$$\mathrm{Tr}\left[P_{\psi^\perp} W\right] = 0, \tag{2.26}$$

$$\mathrm{Tr}\left[P_\phi W\right] = 1, \tag{2.27}$$

$$\mathrm{Tr}\left[P_{\phi^\perp} W\right] = 0. \tag{2.28}$$

Let $\{|e_i\rangle\}_{i=1}^d \subset \mathcal{H}$ be a complete orthonormal basis with $|e_1\rangle = |\psi\rangle$. In terms of this basis, the operator W is decomposed as

$$W = \sum_{i,j=1}^{d} \beta_{ij} |e_i\rangle\langle e_j|, \tag{2.29}$$

where $\beta_{ij} \in \mathbb{C}$. From (2.25) we have $\beta_{11} = 1$.

$$\begin{aligned}\mathrm{Tr}\left[P_{\psi^\perp} W\right] &= \mathrm{Tr}\left[(I - P_\psi) W\right] \\ &= \mathrm{Tr}W - 1 = 0\end{aligned} \tag{2.30}$$

The decomposition of $\psi^\perp$ in this basis is

$$\psi^\perp = \sum_{i=1}^{d} \gamma_i |e_i\rangle = \sum_{i=2}^{d} \gamma_i |e_i\rangle \tag{2.31}$$

where $\gamma_i = \langle e_i|\psi^\perp\rangle$ with $\gamma_1 = 0$. There is the vector $\psi^\perp \in \mathcal{H}$ such that $\gamma_i \neq 0$ for all i. Then we obtain $\beta_{ij} = 0$ for $i, j \geq 2$, since

$$0 = \left\langle \psi^{\perp} \middle| W \middle| \psi^{\perp} \right\rangle = \sum_{i,j \geq 2} \beta_{ij} \gamma_i^* \gamma_j. \tag{2.32}$$

Therefore, the operator W becomes

$$W = |\psi\rangle \langle\psi| + \beta_{12} |\psi\rangle \langle e_2| + \beta_{21} |e_2\rangle \langle\psi| . \tag{2.33}$$

Let $\{f_i\}_{i=1}^{d} \subset \mathcal{H}$ be a complete orthonormal basis with $|f_1\rangle = |\phi\rangle$. The operator W can be expanded in terms of this basis;

$$W = |\phi\rangle \langle\phi| + \gamma_{12} |\phi\rangle \langle f_2| + \gamma_{21} |f_2\rangle \langle\phi| . \tag{2.34}$$

Repeating same argument in terms of the basis $\{e_i\}_{i=1}^{d}$, we arrive at the equation (2.24) by using consistency conditions and condition $\langle\phi|\psi\rangle \neq 0$.

Putting $P = E^A(a)$ in equation (2.24) and let us denote $\nu_{(\psi,\phi)}\left(E^A(a)\right) = q^{\alpha}(a\,|\psi,\phi)$, i.e.,

$$q^{\alpha}(a\,|\psi,\phi) := \alpha \frac{\langle\phi| E^A(a) |\psi\rangle}{\langle\phi|\psi\rangle} + (1-\alpha) \frac{\langle\psi| E^A(a) |\phi\rangle}{\langle\psi|\phi\rangle}. \tag{2.35}$$

It may be easily seen that $q^{\alpha}(a\,|\psi,\phi)$ is a quasi-probability distribution for any α and any pair (ψ,ϕ) with $\langle\phi|\psi\rangle \neq 0$. We shall now generalize $q^{\alpha}(a\,|\psi,\phi)$ a little in order to see this mathematically. Suppose Δ is the interval on $\mathbb{R}$. Let us define $q^{\alpha}(a \in \Delta\,|\psi,\phi)$ as

$$q^{\alpha}(a \in \Delta\,|\psi,\phi) = \int_{\Delta} q^{\alpha}(a\,|\psi,\phi)\,\mathrm{d}a. \tag{2.36}$$

From equation (2.35), we observe that

$$q^{\alpha}(a \in \Delta\,|\psi,\phi) = \alpha \frac{\langle\phi| E^A(\Delta) |\psi\rangle}{\langle\phi|\psi\rangle} + (1-\alpha) \frac{\langle\psi| E^A(\Delta) |\phi\rangle}{\langle\psi|\phi\rangle}. \tag{2.37}$$

The α-parameterized conditional probability distribution $q^{\alpha}(a \in \Delta\,|\psi,\phi)$ satisfies following conditions (i) and (ii);

(i) (Countable additivity)

$$q^{\alpha}\left(a\in\Delta\,|\psi,\phi\,\right)=\sum_i q^{\alpha}\left(a\in\Delta_i\,|\psi,\phi\,\right) \tag{2.38}$$

for any mutually disjoint sequence of intervals $\Delta_1,\Delta_2,\dots$ with $\Delta=\cup_i\Delta_i$.

(ii) (Normalization condition)

$$q^{\alpha}\left(a\in\mathbb{R}\,|\psi,\phi\,\right)=1. \tag{2.39}$$

Since

$$\begin{aligned} q^{\alpha}\left(a\in\Delta\,|\psi,\phi\,\right) &= \alpha\frac{\langle\phi|\sum_i E^A\left(\Delta_i\right)|\psi\rangle}{\langle\phi|\psi\rangle}+\left(1-\alpha\right)\frac{\langle\psi|\sum_i E^A\left(\Delta_i\right)|\phi\rangle}{\langle\psi|\phi\rangle} \\ &= \sum_i\left\{\alpha\frac{\langle\phi|E^A\left(\Delta_i\right)|\psi\rangle}{\langle\phi|\psi\rangle}+\left(1-\alpha\right)\frac{\langle\psi|E^A\left(\Delta_i\right)|\phi\rangle}{\langle\psi|\phi\rangle}\right\} \\ &= \sum_i q^{\alpha}\left(a\in\Delta_i\,|\psi,\phi\,\right), \end{aligned} \tag{2.40}$$

and

$$\begin{aligned} q^{\alpha}\left(a\in\mathbb{R}\,|\psi,\phi\,\right) &= \alpha\frac{\langle\phi|E^A\left(\mathbb{R}\right)|\psi\rangle}{\langle\phi|\psi\rangle}+\left(1-\alpha\right)\frac{\langle\psi|E^A\left(\mathbb{R}\right)|\phi\rangle}{\langle\psi|\phi\rangle} \\ &= 1. \end{aligned} \tag{2.41}$$

One of the advantages of α parameter is that we can consider the real part and imaginary part of the quasi-probability (2.16) at once because the parameter α can be understood to represent the degree of mixture of real part and imaginary part of the quasi-probability (2.16). For example, the conditional α-parametrized quasi-probability distribution (2.35) with $\alpha=1$ becomes the equation (2.16);

$$q^{\alpha=1}\left(a\,|\psi,\phi\,\right)=\frac{\langle\phi|E^A\left(a\right)|\psi\rangle}{\langle\phi|\psi\rangle}. \tag{2.42}$$

If $\alpha = 0$,

$$q^{\alpha=0}\left(a\,|\psi,\phi\,\right) = \frac{\langle\psi|\,E^{A}\left(a\right)|\phi\rangle}{\langle\psi|\phi\rangle} = q^{\alpha=1}\left(a\,|\psi,\phi\,\right)^{*} \tag{2.43}$$

where $*$ represents the complex conjugate. The α-parametrized quasi-probability distribution with $\alpha = 1/2$ becomes real part of the equation (2.16);

$$\begin{aligned} q^{\alpha=1/2}\left(a\,|\psi,\phi\,\right): \quad &= \frac{1}{2}\left\{\frac{\langle\phi|\,E^{A}\left(a\right)|\psi\rangle}{\langle\phi|\psi\rangle} + \frac{\langle\psi|\,E^{A}\left(a\right)|\phi\rangle}{\langle\psi|\phi\rangle}\right\} \\ &= \mathrm{Re}\frac{\langle\phi|\,E^{A}\left(a\right)|\psi\rangle}{\langle\phi|\psi\rangle}. \end{aligned} \tag{2.44}$$

In particular, for the case that two states are same state, $\psi = \phi$, the α-parametrized conditional quasi-probability distribution becomes same form as Born rule of an observable A for any $\alpha \in \mathbb{C}$;

$$q^{\alpha}\left(a\,|\psi,\psi\,\right) := \langle\psi|\,E^{A}\left(a\right)|\psi\rangle = \left|\langle a|\psi\rangle\right|^{2}. \tag{2.45}$$

We thus learn that, under the the interpretation that the condition that two states are same is equivalent to adopting no further condition other than the one state condition, the α-parameterized conditional quasi-probability provides an extension of the Born rule.

2.2.2 Joint Quasi-Probability Distribution

Next, we proceed to define the joint quasi-probability and marginal quasi-probability from the conditional one introduced above. Let $\{b\}$ be the set of eigenvalues characterizing the eigenstates of an observable B, and let $p\left(b\,|\psi\,\right)$ be the probability that observable B takes a value b when the state of system is ψ. Then, supposing that α-parameterized conditional quasi-probability can be treated analogously to the standard quasi-probability, we define the *α-parametrized joint quasi-probability distribution* as;

$$q^{\alpha}\left(b,a\,|\psi\,\right) := q^{\alpha}\left(a\,|\psi,b\,\right)p\left(b\,|\psi\,\right). \tag{2.46}$$

using the conditional quasi-probability (2.35), *i.e.*,

$$q^{\alpha}\left(a\left|\psi,b\right.\right):=\alpha\frac{\left\langle b\right|E^{A}\left(a\right)\left|\psi\right\rangle}{\left\langle b|\psi\right\rangle}+\left(1-\alpha\right)\frac{\left\langle \psi\right|E^{A}\left(a\right)\left|b\right\rangle}{\left\langle \psi|b\right\rangle}. \tag{2.47}$$

If we admit the Born rule for the observable B, one obtains that

$$\begin{aligned} q^{\alpha}\left(a\left|\psi,b\right.\right)p\left(b\left|\psi\right.\right) &= \alpha\frac{\left\langle b\right|E^{A}\left(a\right)\left|\psi\right\rangle}{\left\langle b|\psi\right\rangle}\left|\left\langle b|\psi\right\rangle\right|^{2} \\ &\quad +\left(1-\alpha\right)\frac{\left\langle \psi\right|E^{A}\left(a\right)\left|b\right\rangle}{\left\langle \psi|b\right\rangle}\left|\left\langle b|\psi\right\rangle\right|^{2} \\ &= \alpha\left\langle \psi\right|E^{B}\left(b\right)E^{A}\left(a\right)\left|\psi\right\rangle \\ &\qquad +\left(1-\alpha\right)\left\langle \psi\right|E^{A}\left(a\right)E^{B}\left(b\right)\left|\psi\right\rangle \end{aligned} \tag{2.48}$$

where $E^{B}\left(b\right)$ is the spectrum decomposition of B (i.e. $B=\int bE^{B}\left(b\right)\mathrm{d}b$). This suggests that the α-parametrized joint quasi-probability distribution may be regarded as the generalized Kirkwood function[20] or the weak joint quasi-probability[24] argued by Ozawa. It should be noted that the joint quasi-probability distribution $q^{\alpha}\left(b,a\left|\psi\right.\right)$ is not invariant under the permutation of $E^{A}\left(a\right)$ and $E^{B}\left(b\right)$, that is $q^{\alpha}\left(b,a\left|\psi\right.\right)\neq q^{\alpha}\left(a,b\left|\psi\right.\right)$. Also, we emphasize that our formula of joint quasi-probability (2.48) is valid with the help of the Born rule applied for the observable B.

In order to investigate the properties of the α-parametrized joint quasi-probability distribution $q^{\alpha}\left(b,a\left|\psi\right.\right)$, it is convenient for us to introduce the $\circ_{\alpha}$-product for two operators X and Y on $\mathcal{H}$ as

$$X\circ_{\alpha}Y:=\alpha XY+\left(1-\alpha\right)YX. \tag{2.49}$$

With the $\circ_{\alpha}$-product, the joint quasi-probability distribution becomes much simpler;

$$q^{\alpha}\left(b,a\left|\psi\right.\right)=\left\langle \psi\right|E^{B}\left(b\right)\circ_{\alpha}E^{A}\left(a\right)\left|\psi\right\rangle. \tag{2.50}$$

For $\alpha=1$ and $\alpha=0$, the $\circ_{\alpha}$-product becomes the usual operator product;

$$X\circ_{\alpha=1}Y=XY, \tag{2.51}$$

$$X \circ_{\alpha=0} Y = YX. \tag{2.52}$$

For $\alpha = 1/2$, the $\circ_\alpha$-product becomes Jordan product[36, 72];

$$X \circ_{\alpha=\frac{1}{2}} Y = \frac{1}{2}(XY + YX). \tag{2.53}$$

It follows that the α-parametrized joint quasi-probability distribution with $\alpha = 1/2$ is symmetric under the permutation of $E^A(a)$ and $E^B(b)$;

$$\begin{aligned} q^{\alpha=1/2}(a, b\,|\psi) &= \mathrm{Re}\,\langle\psi|\,E^B(b)\,E^A(a)\,|\psi\rangle \\ &= \mathrm{Re}\,\langle\psi|\,E^A(a)\,E^B(b)\,|\psi\rangle = q^{\alpha=1/2}(b, a\,|\psi). \end{aligned} \tag{2.54}$$

If we put $\beta := \mathrm{Re}\alpha$ and $\gamma := \mathrm{Im}\alpha$, then,

$$\begin{aligned} X \circ_{\alpha=\beta+i\gamma} Y &= \beta XY + (1-\beta)\,YX + i\gamma\,[X, Y] \\ &= X \circ_\beta Y + i\gamma\,[X, Y] \end{aligned} \tag{2.55}$$

where $[X, Y] = XY - YX$. For $\alpha = \frac{1}{2}(1-i)$, the $\circ_\alpha$-product becomes

$$X \circ_{\alpha=\frac{1}{2}(1-i)} Y = \frac{1}{2}(XY + YX) + \frac{1}{2i}(XY - YX) \tag{2.56}$$

For every α, the $\circ_\alpha$-product has the following properties;

$$I \circ_\alpha X = X \circ_\alpha I = X, \tag{2.57}$$

$$X \circ_\alpha X = X^2, \tag{2.58}$$

$$X \circ_\alpha Y - Y \circ_\alpha X = (2\alpha - 1)\,[X, Y], \tag{2.59}$$

$$X \circ_\alpha Y + Y \circ_\alpha X = XY + YX. \tag{2.60}$$

$$[X, Y] = 0 \Rightarrow X \circ_\alpha Y = XY = YX \tag{2.61}$$

where I is the identity operator on $\mathcal{H}$ and $[X, Y] = XY - YX$. From the the third property (2.59), the difference between $q^{\alpha}(b, a\,|\psi\,)$ and $q^{\alpha}(a, b\,|\psi\,)$ represented by following form;

$$q^{\alpha}(b, a\,|\psi\,) - q^{\alpha}(a, b\,|\psi\,) = (2\alpha - 1)\,\langle\psi|\left[E^{B}(b)\,, E^{A}(a)\right]|\psi\rangle\,. \tag{2.62}$$

This equation asserts that if the observable A and B are commute then $q^{\alpha}(b, a\,|\psi\,) = q^{\alpha}(a, b\,|\psi\,)$. In this case,

$$q^{\alpha}(b, a\,|\psi\,) = \langle\psi|\,E^{B}(b)\,E^{A}(a)\,|\psi\rangle = \langle\psi|\,E^{A}(a)\,E^{B}(b)\,|\psi\rangle \tag{2.63}$$

for any $\alpha \in \mathbb{C}$. It should be noted again that our formula of joint quasi-probability is established by assuming the Born rule for the observable B.

2.2.3 Marginal Quasi-Probability Distribution

Finally, we shall define the marginal quasi-probability distribution of A on ψ by analogy with the usual probability;

$$q^{\alpha}(a\,|\psi\,) := \int q^{\alpha}(b, a\,|\psi\,)\,\mathrm{d}b = \int q^{\alpha}(a\,|\psi, b\,)\,p\,(b\,|\psi\,)\,\mathrm{d}b. \tag{2.64}$$

We then find from (2.48) and (2.57) that this marginal quasi-probability distribution is independent of the parameter α and yields the conventional probability of obtaining a particular value a of A,

$$\begin{aligned} q^{\alpha}(a\,|\psi\,) = \int q^{\alpha}(b, a\,|\psi\,)\,\mathrm{d}b &= \int q^{\alpha}(a\,|\psi, b\,)\,p\,(b\,|\psi\,)\,\mathrm{d}b \\ &= \int \langle\psi|\,E^{B}(b) \circ_{\alpha} E^{A}(a)\,|\psi\rangle\,\mathrm{d}b \\ &= \langle\psi|\,I \circ_{\alpha} E^{A}(a)\,|\psi\rangle \\ &= \langle\psi|\,E^{A}(a)\,|\psi\rangle \\ &= |\langle a|\psi\rangle|^{2}\,. \end{aligned} \tag{2.65}$$

Since the choice of observable A is arbitrary, this is in fact equivalent to the Born rule.

This result (2.65) shows that, whatever the interpretation one attaches to the α-parametrized quasi-probability distribution, one ends up with the conventional Born rule at the level of the marginal distribution that can be directly tested by measurement. At this point, we remark that equation (2.65) resembles the reproduction condition in the ontological model considered in [8], if the eigenstate $|b\rangle$ of the observable B is regarded as an *ontic state*. However, this is not quite the case since our quasi-probability is ψ-dependent in general whereas the counterpart in the ontological model is ψ-independent and standard probability. More on this comes later when we discuss the relevance of quasi-probability in the ontological model.

From the foregoing discussions, one may surmise that the α-parameterized conditional quasi-probability (2.35) provides a key ingredient of quantum theory. In fact, since our conditional quasi-probability is based upon the assumption of the observable B as a special reference observable satisfying the Born rule, one is alluded to the interpretation that the resultant theory is deterministic in the sense that

$$q^{\alpha}\left(b'\,|\psi,b\right)=\begin{cases}1 & b=b'\\ 0 & b\neq b'\end{cases}. \tag{2.66}$$

As we shall see in chapter 4, when the observable B is a position observable $\mathbf{X}$, this interpretation leads to the Bohmian mechanics [2, 3].

2.2.4 Quasi-Probability in Spin-1/2 system

Before closing this section, it is useful to discuss an example of simple physical system. We shall consider the Qbit (*i.e.,* the one particle with spin 1/2) system, in which case the corresponding Hilbert space is $\mathbb{C}^2$. It should be noted that our introduction of the α-parameterized quasi-probability is restricted to that dimension of Hilbert space is larger than dimension 3. It is, however, the quasi-probability which have the same form as the equation (2.35) can be achieved from another assumption. This argument is given in the appendix A.

Let $\sigma_x,\sigma_y,\sigma_z$ be the Pauli matrices. Put $\vec{\sigma}=(\sigma_x,\sigma_y,\sigma_z)$. The spectral projections of spin with $\vec{a}\in\mathbb{R}^3$ direction are given by

$$E^{\vec{a}}\left(+1\right)=\frac{1+\vec{a}\cdot\vec{\sigma}}{2},\qquad E^{\vec{a}}\left(-1\right)=\frac{1-\vec{a}\cdot\vec{\sigma}}{2}. \tag{2.67}$$

Then, we can write the spectral projections of $\vec{a}$-spin and $\vec{b}$-spin as

$$E^{\vec{a}}(r) = \frac{1 + r\vec{a}\cdot\vec{\sigma}}{2}, \qquad E^{\vec{b}}(s) = \frac{1 + s\vec{b}\cdot\vec{\sigma}}{2} \tag{2.68}$$

where $r \in \{-1, +1\}$ and $s \in \{-1, +1\}$. The operator product $E^{\vec{a}}(r)\, E^{\vec{b}}(s)$ becomes

$$\begin{aligned} E^{\vec{a}}(r)\, E^{\vec{b}}(s) &= \left(\frac{1 + r\vec{a}\cdot\vec{\sigma}}{2}\right)\left(\frac{1 + s\vec{b}\cdot\vec{\sigma}}{2}\right) \\ &= \frac{1}{4}\left\{\left(1 + rs\vec{a}\cdot\vec{b}\right) I + \left(r\vec{a} + s\vec{b} + rs\vec{a}\times\vec{b}\right)\cdot\vec{\sigma}\right\} \end{aligned} \tag{2.69}$$

where $\times$ represents a cross product in $\mathbb{R}^3$. We observe that the commutation relation of $E^{\vec{a}}(r)$ and $E^{\vec{b}}(s)$ is given by

$$\left[E^{\vec{a}}(r), E^{\vec{b}}(s)\right] = \frac{1}{2} rs\left(\vec{a}\times\vec{b}\right)\cdot\vec{\sigma}. \tag{2.70}$$

Let us calculate α-product $\circ_\alpha$ of $E^{\vec{a}}(r)$ and $E^{\vec{b}}(s)$;

$$E^{\vec{a}}(r)\circ_\alpha E^{\vec{b}}(s) = \frac{1}{4}\left\{\left(1 + rs\vec{a}\cdot\vec{b}\right) I + \left(r\vec{a} + s\vec{b} + (2\alpha - 1)\, rs\vec{a}\times\vec{b}\right)\cdot\vec{\sigma}\right\}, \tag{2.71}$$

$$E^{\vec{b}}(s)\circ_\alpha E^{\vec{a}}(r) = \frac{1}{4}\left\{\left(1 + rs\vec{a}\cdot\vec{b}\right) I + \left(r\vec{a} + s\vec{b} - (2\alpha - 1)\, rs\vec{a}\times\vec{b}\right)\cdot\vec{\sigma}\right\}. \tag{2.72}$$

Then the α-parametrized joint probability distribution of $\vec{a}$-spin and $\vec{b}$-spin on arbitrary state $\psi \in \mathbb{C}^2$ is given by

$$\begin{aligned} q^{\alpha}(s, r\,|\psi) &= \langle\psi|\, E^{\vec{b}}(s)\circ_\alpha E^{\vec{a}}(r)\,|\psi\rangle \\ &= \frac{1}{4}\left\{\left(1 + rs\vec{a}\cdot\vec{b}\right) + \left(r\vec{a} + s\vec{b} - (2\alpha - 1)\, rs\vec{a}\times\vec{b}\right)\cdot\langle\vec{\sigma}\rangle_\psi\right\}. \end{aligned} \tag{2.73}$$

where $\langle\vec{\sigma}\rangle_\psi = \left(\langle\sigma_x\rangle_\psi, \langle\sigma_y\rangle_\psi, \langle\sigma_z\rangle_\psi\right)$ and $\langle\sigma_i\rangle_\psi = \langle\psi|\,\sigma_i\,|\psi\rangle$, $(i = x, y, z)$. For $\alpha = 1/2$, the joint quasi-probability becomes

$$q^{1/2}(s, r\,|\psi) = \frac{1}{4}\left\{\left(1 + rs\vec{a}\cdot\vec{b}\right) + \left(r\vec{a} + s\vec{b}\right)\cdot\langle\vec{\sigma}\rangle_\psi\right\} = q^{1/2}(r, s\,|\psi). \tag{2.74}$$

The marginal of $q^{\alpha}(s, r|\psi)$ with respect to $\vec{b}$-spin provides Born's rule of $\vec{a}$-spin;

$$\begin{aligned}\sum_{s=+1,-1} q^{\alpha}(s, r|\psi) &= \frac{1}{4}\left\{\left(1 + r\vec{a}\cdot\vec{b}\right) + \left(r\vec{a} + \vec{b} - (2\alpha-1)\, r\vec{a}\times\vec{b}\right)\cdot\langle\vec{\sigma}\rangle_{\psi}\right\} \\ &\quad + \frac{1}{4}\left\{\left(1 - r\vec{a}\cdot\vec{b}\right) + \left(r\vec{a} - \vec{b} + (2\alpha-1)\, r\vec{a}\times\vec{b}\right)\cdot\langle\vec{\sigma}\rangle_{\psi}\right\} \\ &= \frac{1 + r\vec{a}\cdot\langle\vec{\sigma}\rangle_{\psi}}{2} = \langle\psi| E^{\vec{a}}(r)|\psi\rangle . \end{aligned} \tag{2.75}$$

The α-parametrized conditional quasi-probability is given by

$$q^{\alpha}(s|r, \psi) = \frac{q^{\alpha}(s, r|\psi)}{p(s|\psi)} = \frac{1}{2}\frac{\left\{\left(1 + rs\vec{a}\cdot\vec{b}\right) + \left(r\vec{a} + s\vec{b} - (2\alpha-1)\, rs\vec{a}\times\vec{b}\right)\cdot\langle\vec{\sigma}\rangle_{\psi}\right\}}{1 + s\vec{b}\cdot\langle\vec{\sigma}\rangle_{\psi}}. \tag{2.76}$$

2.3 Measurement of Quasi-Probability

In this section, we shall introduce the weak measurement procedure [11] and recall that the real part and imaginary part of weak value of projection

$$\frac{\langle\phi| E^{A}(a)|\psi\rangle}{\langle\phi|\psi\rangle} \tag{2.77}$$

can be measurable at least theoretically. Since the α-parametrized conditional quasi-probability can be rewritten as

$$\begin{aligned} q^{\alpha}(a|\psi, \phi) &= \alpha\frac{\langle\phi| E^{A}(a)|\psi\rangle}{\langle\phi|\psi\rangle} + (1-\alpha)\frac{\langle\psi| E^{A}(a)|\phi\rangle}{\langle\psi|\phi\rangle} \\ &= \mathrm{Re}\left[\frac{\langle\phi| E^{A}(a)|\psi\rangle}{\langle\phi|\psi\rangle}\right] + (2\alpha-1)\,\mathrm{Im}\left[\frac{\langle\phi| E^{A}(a)|\psi\rangle}{\langle\phi|\psi\rangle}\right], \end{aligned} \tag{2.78}$$

then, we can measure it expect of α.

The weak measurement procedure consists of the weak measurement whose measuring interaction is too weak, and the post-selection. The post-selection is the operation that fixes the state of the system after measurement. In usual measurement, the observer obtains the conditional probability distribution of some observable A on the same prepared state ψ, $p(a|\psi, A)$. In the post-selected measurement, the observer obtains the conditional probability distribution of some observable A on the process ϕ to ψ. Figure 2.3.1 gives the intuitive description of the post-selected measurement. We sometimes call pre-selection and post-selection the initial preparation of state and final choice of state and denote by the conditional probability $p(a\,|\psi, A, \phi)$ with pre-selection ψ and post-selection ϕ.

By using quantum measurement theory with the post-selection, it is seen that the post-selected expectation value in the weak measurement gives the weak value. Quantum measurement theory is originated by von Neumann [5] and developed by Ludwig [37, 38], Ozawa [39] and so on. Our description of quantum measurement theory that follows is [40] and we shall recollect the work of Aharonov *et al.* [11] by embedding the post-selection in quantum measurement theory.

The measurement of quantum system may be described by the indirect measurement model with probe system P in quantum measurement theory. Let us denote by $\mathcal{K}$ the Hilbert space of probe P. The indirect measurement model is described by the quadruple $(\mathcal{K}, \sigma, U, M)$ consisting of a Hilbert space $\mathcal{K}$, an initial state σ of P, an unitary operator U which represents the time evolution during the measurement interaction between the target system and the probe, and an observable M of P, which is the observable that the observer really measures. Let Π^X be the probability operator valued measure (POVM) corresponding to measurement of observable X of target system S. By this, we may regard the measurement of X as the measurement of A with some error. The probability distribution of measurement of X on ψ

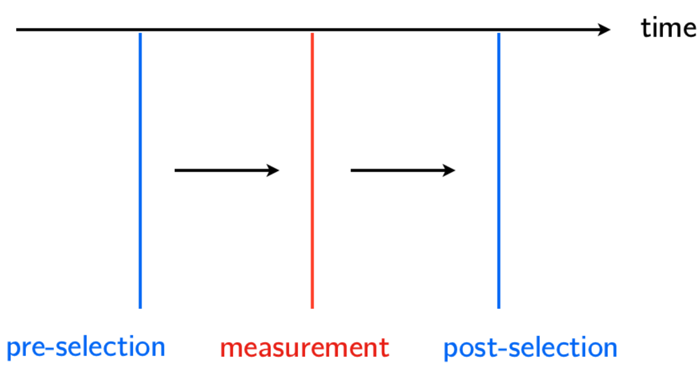

Figure 2.3.1: Post-selection

is given by

$$p\left(x\left|\psi, X\right.\right) = \mathrm{Tr}\left[\Pi^{X}\left(x\right)P_{\psi}\right]. \tag{2.79}$$

Let $(\mathcal{K}, \sigma, U, M)$ be the indirect measurement model of measurement of A, the probability distribution of X is given by

$$\begin{aligned} p\left(x\left|\psi, X\right.\right) &= \mathrm{Tr}\left[\left(I \otimes E^{M}\left(x\right)\right)U\left(P_{\psi}\otimes\sigma\right)U^{\dagger}\right] &(2.80)\\ &= \mathrm{Tr}\left[\mathrm{Tr}_{\mathcal{K}}\left[U^{\dagger}\left(I\otimes E^{M}\left(x\right)\right)U\left(I\otimes\sigma\right)U\right]P_{\psi}\right]. &(2.81)\end{aligned}$$

Then, we observe that

$$\Pi^{X}\left(x\right) = \mathrm{Tr}_{\mathcal{K}}\left[U^{\dagger}\left(I\otimes E^{M}\left(x\right)\right)U\left(I\otimes\sigma\right)U\right]. \tag{2.82}$$

Let us denote by ρ_x the state just after measurement conditional upon the outcome "$X = x$". From the indirect measurement model, it may be given by

$$\rho_x = \frac{\mathrm{Tr}_{\mathcal{K}}\left[U^{\dagger}\left(I\otimes E^{M}\left(x\right)\right)U\left(P_{\psi}\otimes\sigma\right)U\right]}{\mathrm{Tr}\left[U^{\dagger}\left(I\otimes E^{M}\left(x\right)\right)U\left(P_{\psi}\otimes\sigma\right)U\right]}. \tag{2.83}$$

The joint probability distribution, $p\left(x, y\left|\psi, X, Y\right.\right)$, of successive measurement A and B is given by

$$\begin{aligned} p\left(x, y\left|\psi, X, Y\right.\right) &= p\left(y\left|\rho_x, Y\right.\right)p\left(x\left|\psi, X\right.\right)\\ &= \mathrm{Tr}\left[\mathrm{Tr}_{\mathcal{K}}\left[U^{\dagger}\left(\Pi^{Y}\left(y\right)\otimes E^{M}\left(x\right)\right)U\left(I\otimes\sigma\right)\right]P_{\psi}\right]\\ &= \langle\psi|\,\mathrm{Tr}_{\mathcal{K}}\left[U^{\dagger}\left(\Pi^{Y}\left(y\right)\otimes E^{M}\left(x\right)\right)U\left(I\otimes\sigma\right)\right]|\psi\rangle. &(2.84)\end{aligned}$$

The post-selected probability can be represented by the form;

$$p\left(x\left|\psi, X, \phi\right.\right) = \frac{p\left(x, y\left|\psi, X, Y\right.\right)}{\int p\left(x, y\left|\psi, X, Y\right.\right)\mathrm{d}x} \tag{2.85}$$

where $|\phi\rangle = |y\rangle$. This becomes in the indirect measurement model that

$$p\left(x\left|\psi, X, \phi\right.\right) = \frac{\langle\psi|\operatorname{Tr}_{\mathcal{K}}\left[U^{\dagger}\left(\Pi^{Y}\left(y\right)\otimes E^{M}\left(x\right)\right)U\left(I\otimes\sigma\right)\right]|\psi\rangle}{\int\langle\psi|\operatorname{Tr}_{\mathcal{K}}\left[U^{\dagger}\left(\Pi^{Y}\left(y\right)\otimes E^{M}\left(x\right)\right)U\left(I\otimes\sigma\right)\right]|\psi\rangle\,\mathrm{d}x}. \tag{2.86}$$

The post-conditional expectation value of X, which we denote it by $\mathbb{E}_{\psi\phi}\left(X\right)$, is defined by

$$\mathbb{E}_{\psi\phi}\left(X\right) = \int xp\left(x\left|\psi, X, \phi\right.\right)\mathrm{d}x. \tag{2.87}$$

Next, we shall show that the post-conditional expectation value $\mathbb{E}_{\psi\phi}\left(X\right)$ in the typical indirect measurement model with weak measurement interaction gives the weak value. Suppose that the time evolution operator U is given by

$$U = U_g := e^{-igE^{A}(a)\otimes P} \tag{2.88}$$

where g is coupling constant of measurement interaction and P is the momentum observable of P, and the measurement observable is the position of probe Q. In this model $(\mathcal{K}, \sigma, A, Q)$, the successive joint probability (2.84) becomes for small g

$$\begin{aligned} p\left(x, \phi\left|\psi, X, Y\right.\right) &= \langle\psi|\operatorname{Tr}_{\mathcal{K}}\left[e^{igE^{A}(a)\otimes P}\left(P_{\phi}\otimes E^{Q}\left(x\right)\right)e^{-igE^{A}(a)\otimes P}\left(I\otimes\sigma\right)\right]|\psi\rangle \\ &= \langle\psi|P_{\phi}|\psi\rangle\operatorname{Tr}\left[E^{Q}\left(x\right)\sigma\right] \\ &\quad +g\langle\psi|\frac{1}{2}\left[P_{\phi}, E^{A}\left(a\right)\right]|\psi\rangle\operatorname{Tr}\left[\left\{P_{\phi}, E^{A}\left(a\right)\right\}\sigma\right] \\ &\quad -g\langle\psi|\frac{1}{2}\left\{P_{\phi}, E^{A}\left(a\right)\right\}|\psi\rangle\operatorname{Tr}\left[\left[P_{\phi}, E^{A}\left(a\right)\right]\sigma\right] + O\left(g^{2}\right) \end{aligned} \tag{2.89}$$

where $\{A, B\} = AB + BA$ and $[A, B] = AB - BA$. The post-conditional expectation value is given by

$$\begin{aligned} \mathbb{E}_{\psi\phi}\left(X\right) &= \operatorname{Tr}\left[Q\sigma\right] + g\frac{\langle\psi|\frac{1}{2}\left[P_{\phi}, E^{A}\left(a\right)\right]|\psi\rangle}{\langle\psi|P_{\phi}|\psi\rangle}\left\{\operatorname{Tr}\left[\left\{Q, P\right\}\sigma\right] - \operatorname{Tr}\left[Q\sigma\right]\operatorname{Tr}\left[P\sigma\right]\right\} \\ &\quad -ig\frac{\langle\psi|\frac{1}{2}\left\{P_{\phi}, E^{A}\left(a\right)\right\}|\psi\rangle}{\langle\psi|P_{\phi}|\psi\rangle}\operatorname{Tr}\left[\left[Q, P\right]\sigma\right] + O\left(g^{2}\right). \end{aligned} \tag{2.90}$$

It is easily seen that

$$\mathrm{Re}\left[\frac{\langle\phi|\,E^{A}\left(a\right)|\psi\rangle}{\langle\phi|\psi\rangle}\right]=\frac{\langle\psi|\,\frac{1}{2}\left\{E^{A}\left(a\right),P_{\phi}\right\}|\psi\rangle}{\langle\psi|\,P_{\phi}\,|\psi\rangle}, \tag{2.91}$$

$$\mathrm{Im}\left[\frac{\langle\phi|\,E^{A}\left(a\right)|\psi\rangle}{\langle\phi|\psi\rangle}\right]=\frac{\langle\psi|\,\frac{1}{2}\left[E^{A}\left(a\right),P_{\phi}\right]|\psi\rangle}{\langle\psi|\,P_{\phi}\,|\psi\rangle}. \tag{2.92}$$

Using the commutation relation $[Q,P]=i$, we observe that

$$\begin{aligned}\mathbb{E}_{\psi\phi}\left(X\right)\quad-\mathrm{Tr}\left[Q\sigma\right]=\quad & g\mathrm{Re}\left[\frac{\langle\phi|\,E^{A}\left(a\right)|\psi\rangle}{\langle\phi|\psi\rangle}\right] \\ & +g\mathrm{Im}\left[\frac{\langle\phi|\,E^{A}\left(a\right)|\psi\rangle}{\langle\phi|\psi\rangle}\right]\left\{\mathrm{Tr}\left[\left\{Q,P\right\}\sigma\right]-\mathrm{Tr}\left[Q\sigma\right]\mathrm{Tr}\left[P\sigma\right]\right\}+O\left(g^{2}\right).\end{aligned} \tag{2.93}$$

Let us choose the initial state of probe σ satisfying

$$\left\{\mathrm{Tr}\left[\left\{Q,P\right\}\sigma\right]-\mathrm{Tr}\left[Q\sigma\right]\mathrm{Tr}\left[P\sigma\right]\right\}=0, \tag{2.94}$$

then

$$\mathbb{E}_{\psi\phi}\left(X\right)-\mathrm{Tr}\left[Q\sigma\right]=g\mathrm{Re}\left[\frac{\langle\phi|\,E^{A}\left(a\right)|\psi\rangle}{\langle\phi|\psi\rangle}\right]+O\left(g^{2}\right). \tag{2.95}$$

The imaginary part of $\langle\phi|\,E^{A}\left(a\right)|\psi\rangle/\langle\phi|\psi\rangle$ can be measured by choosing the meter observable P instead of Q. In this case, the post-selected expectation value is given by

$$\mathbb{E}_{\psi\phi}\left(X\right)-\mathrm{Tr}\left[P\sigma\right]=g\mathrm{Im}\left[\frac{\langle\phi|\,E^{A}\left(a\right)|\psi\rangle}{\langle\phi|\psi\rangle}\right]\left\{\mathrm{Tr}\left[P^{2}\sigma\right]-\left(\mathrm{Tr}\left[P\sigma\right]\right)^{2}\right\}+O\left(g^{2}\right). \tag{2.96}$$

We can regard the quantity $\mathrm{Tr}\left[P^{2}\sigma\right]-\left(\mathrm{Tr}\left[P\sigma\right]\right)^{2}$ as some constant since it can be determined by other experience. Hence, we have recollected that the real and imaginary part of $\langle\phi|\,E^{A}\left(a\right)|\psi\rangle/\langle\phi|\psi\rangle$ is measurable. It should be noted that the real and imaginary part of quasi-probability are measured as the approximate value of the average of the post-selected probability $p\left(x\,|\psi,X,\phi\right)$. Using above technique, Bamber and Lundeen have realized the Kirkwood-Dirac distribution experimentally [27, 30].

Chapter 3

Operational Probabilistic Theories and Quasi-Probability

In this chapter, we shall introduce operational probabilistic theories which are also referred as generalized probabilistic theories. Operational probabilistic theories developed with adopting a purely operational description of physical experiments and a notion of probability. For this reason, operational probabilistic theories can provide unified theoretical framework in which both of classical and quantum theory appear as special cases. By "operational", we mean that concepts in the theory are meaningful only insofar as they correspond to physical operations. The idea of "operationalism" or "operationism" advocated by the American condensed matter physicist Bridgeman in 1927 [41]. The operational view, especially, got familiar to practicing physicists who almost are inspired by tradition of American pragmatism or the philosophy of logical positivism. The following Bridgeman's statement straightforwardly expresses the idea of operationalism [41] :

> *In general, we mean by any concept nothing more than a set of operations; the concept is synonymous with the corresponding set of operations. If the concept is physical, as of length, the operations are actual physical operations, namely, those by which length is measured; or if the concept is mental, as of mathematical continuity, the operations are mental operations, namely, those by which we determine whether a given aggregate of magnitudes is continuous [. . .] We must demand*

that the set of operations equivalent to any concept be a unique set, for otherwise there are possibilities of ambiguity in practical applications which we cannot admit.

The idea of operationalism is very useful to consider the quantum weirdness described in chapter 1. There has been a movement to reconstruct quantum mechanics from an operational viewpoint. This is, namely, an attempt to derive the Hilbert space structure of quantum theory from physically meaningful principles. Historically, early pioneers of these works are Birkoff-von Neumann [42]. Their works are so-called quantum logic approach. In 1957, American mathematician Mackey sketched a probabilistic framework for the mathematical foundations of quantum and classical mechanics.

In this chapter, we review operational probabilistic theories for two reasons; first, for investigating the relationship between the ontological model of quantum mechanics and quasi-probability, we need operational probabilistic theories as a theoretical basis of the ontological model. This allows us to examine the physical reality from the epistemological point of view. Secondly, we need to analyze the epistemological significance of a post-selected measurement to find out a proper quasi-probability in the weak measurement procedure. In section 3.1 we review the elements of operational probabilistic theories. In section 3.2, the definition of the simultaneous measurability and determinism is introduced and operational theories are classified with simultaneous measurability as the basic concept. This provides a clue as to what properties the general hidden variable theories should have. In section 3.3, we mention some examples of operational theories in physics. The concept of entanglement is introduced in section 3.4. At the end of this chapter, we show the epistemological significance of post-selected measurement, which is one of the key ideas of weak measurement, in terms of the time direction of inference in the probabilistic theory. The description of review part of operational probabilistic theories that follows is mainly based on Mackey [43] and Watanabe [44]. For the more details about operational theories, see Refs. [45, 46, 47, 49, 50].

3.1 Preparation, Measurement and Probability

In this section, we construct an operational framework of physical theories. For this purpose, we build a common picture of general physical experiments. First, it may be useful to suppose that there exists an 'object system' which is external to the observer. We denote by S an object system. The operations for an object system S may be made by *measurement apparatuses*. We use first capital letters of the alphabet, A, B, C, $\cdots$ to denote measurement apparatuses, and we use same letters to represent operations made apparatuses. If after the measurement is performed, the observer reads the *output value* of the measurement apparatus which he used in measurement. We use small letters , a, b, c, $\cdots$ to denote output values of corresponding measurement apparatuses A, B, C, $\cdots$. We assume that the output values of any apparatuses are real number. Let $\mathbb{K}_A$ be such a set that consists of all possible output values of the apparatus A. We can summarize the above by defining an "*observational proposition*" [44] which has the form:

> If one performs a certain well-defined operation A of observation on an object system, he obtain result $a \in \mathbb{K}_A$.

For simplicity, we write $A = a$ for this proposition. Broadly speaking, it is can be considered that measurement means such a action as one ascertain that an observational proposition $A = a$ is true or not. Let us denote by $\mathscr{M}$ the set of all considering measurement apparatuses[1]. We shall assume the following statement that seems to be almost inevitable for any physical theory;

Assumption (Experimental Proposition) [44]

> A theory in physics is supposed to produce an "experimental proposition" of the type: the probability that "$A = a$" will be true on "condition" P, $p\left(A = a \,|P\right)$, has such-and-such.

We call the conditional probability distribution $p\left(A = a \,|P\right)$ the *output probability distribution* of an operation A given a condition P. Hereafter, we mainly use the notation $p\left(a \,|P, A\right)$

[1] The set of all possible observational propositions in an operational framework is sometimes called *test space* [50].

for the output probability instead of $p\left(A = a\,|P\,\right)$ by considering that the choice of measurement apparatus is one of the conditions on probability. A condition P which appears in an experimental proposition can be considered as the conditioning operation or the preparatory measurement. For simplicity, we refer to a conditioning operation as a *preparation*. Let us denote by $\mathscr{P}$ the set of all possible preparation. It should be noted that there is a time order in P and A. The probability $p\left(a\,|P, A\,\right)$ can be regarded as the number accompanied by the inference of the value of A, a, from the a priori information P.

In summary, a general physical experiment can be characterized by a pair (A, P) consisting of the measurement A and the preparation P [2], and the observer obtains the output probability distribution $p\left(a\,|P, A\,\right)$ through the experiment. We shall define the probability weight function ω which assigns the $p\left(a\,|A, P\,\right)$ for a, A, and P:

$$\omega\left(a, A, P\right) = p\left(a\,|P, A\,\right). \tag{3.1}$$

This function ω describes a general physical experiment for given preparation and measurement. We shall assume that a map $a \mapsto \omega\left(a, A, P\right) = p\left(a\,|P, A\,\right)$ is a probability measure for any observable A.

We need not to distinguish the measurement apparatuses which observe same property in epistemological consideration. For example, a vernier caliper and a ruler, both of which are to measure the position. Then, let us define the following equivalent class. Two measurement A and A' are said to be operationally equivalent if they satisfy

$$p\left(a\,|P, A\,\right) = p\left(a'\,\middle|P, A'\right) \tag{3.2}$$

for all a, a' and P. In this case, we can regard that apparatuses A and A' measures same quantity. We shall denote $A = A'$ if A and A' are operationally equivalent. This relationship between A and A' defines the equivalence class in $\mathscr{M}$, because it satisfies following properties:

(i) $A = A$,

(ii) if $A \sim A'$, then $A' = A$,

[2] Leifer calls this pair a preparation-measurement fragment [51] .

(iii) if $A = A'$ and $A' = A''$, then $A = A''$.

We shall now call each equivalence class *observable*, respectively. Let $\mathscr{O}$ be the set of all considering observables. Similarly, we shall define the equivalent class in $\mathscr{P}$. Two preparation P and P' are operationally equivalent if these satisfy

$$p\left(a\left|P,A\right.\right) = p\left(a\left|P',A\right.\right) \tag{3.3}$$

for all a and A. We shall now give the same name preparation for each equivalent class respectively. Some physicist calls this equivalent class a *state*. In above words, It can be considered that a general physical experiment is modeled by a triple $(\mathscr{O},\mathscr{P},\omega)$. We shall now call a a triple $(\mathscr{O},\mathscr{P},\omega)$ operational probabilistic theory. Let OPT be a set of considerable operational probabilistic theories.

A preparation is, basically, measurement that made by a certain measurement apparatus, e.g., $B \in \mathscr{O}$. In this case, we shall denote $P =$ "$B = b$". We admit the probabilistic mixture of preparations such that $P = P_1 \,\mathrm{or}\, P_2$ where $P_1 =$ "$B = b_1$" and $P_2 =$ "$B = b_2$" with probabilities $p\left(P_1\left|P\right.\right)$ and $p\left(P_2\left|P\right.\right) = 1 - p\left(P_1\left|P\right.\right)$. In this case, the output probability $p\left(a\left|A,P\right.\right)$ becomes

$$p\left(a\left|P,A\right.\right) = p\left(a\left|P_1,A\right.\right)p\left(P_1\left|P\right.\right) + p\left(a\left|P_2,A\right.\right)p\left(P_2\left|P\right.\right). \tag{3.4}$$

We shall call the state P the mixed state of P_1 and P_2 if it satisfies the equation (3.4) for all a and A. The mixed state P is denoted as

$$P = \langle p_1, p_2; P_1, P_2\rangle, \quad p_i = p\left(P_i\left|P\right.\right) \quad (i = 1, 2) \tag{3.5}$$

The mixed state can be generalized to infinite number of probabilistic mixing:

$$p\left(a\left|P,A\right.\right) = \int_{\mathbb{K}} p\left(a\left|P_k,A\right.\right)p\left(P_k\left|P\right.\right)\mathrm{d}k \tag{3.6}$$

where k is some parameter which labels the preparations. By the nature of probability, $\int p\left(P_k\,|P\right)dk = 1$. In this case, we shall denote P as

$$P = \langle p_k : P_k \rangle_{k \in \mathbb{K}}\,, \quad p_k = p\left(P_k\,|P\right). \tag{3.7}$$

Notice that every preparation P satisfies the equation (3.4) in cases that

(i) $p\left(P_1\,|P\right) = 1$, (i.e., $P = P_1$),

(ii) $p\left(P_1\,|P\right) = 0$, (i.e., $P = P_2$),

(iii) $P_1 = P_2$, , (i.e., $P = P_1 = P_2$)

We shall call these cases (i), (ii) and (iii) *trivial mixing.* A preparation P is *pure* if it cannot be decomposed as the equation (3.5) except for trivial mixing cases. A preparation P is a mixed state if it is not a pure state. Mathematically, above discussion states that $\mathscr{P}$ has σ-convex structure. We shall not address the mathematics of convex theory in this thesis. We shall denote by $\mathscr{P}_{\text{pure}}$ the set of all pure preparation in $\mathscr{P}$. The mixed preparation corresponds to the case that observer's knowledge about the preparation is insufficient. For instance, the preparation in classical mechanics may correspond to the setting of initial condition of system. A pure preparation corresponds to the initial point on the phase space whereas a mixed preparation correspond to the initial distributions on the phase space. We shall see examples of concrete physical theories in detail later.

3.2 Determinism and Simultaneous Measurability

It is frequently said that quantum mechanics is not deterministic theory whereas classical mechanics is deterministic. This is because all predictions from standard quantum theory are only probability which is given by Born's rule. In this section, we shall define the determinism in our operational framework.

An operational theory $(\mathscr{O}, \mathscr{P}, \omega)$ is to be deterministic, if it satisfies following the two conditions;

(i) For all pure preparation P_{p} the output probability of any observational proposition is 1 or 0, i.e.,

$$p\left(a\left|P_{\mathrm{p}}, A\right.\right) = 1 \text{ or } 0 \tag{3.8}$$

for any observable $A \in \mathscr{O}$ and any output $a \in \mathbb{K}_A$ of A.

(ii)There exists a pure state $P\left(a\right)$ for arbitrary output a of arbitrary observable A such that $p\left(a\left|P, A\right.\right) = 1$

Let DET be a set of all deterministic theories. In DET theories, the output values of all observables are determined at the moment when we know the pure state. If the output values are fluctuated in DET theories, it is caused by that the preparation is probabilistically mixed (i.e., lack of knowledge of state-condition). It may be reasonable to require one more condition that all the pure preparation is determined by one special observable $Z \in \mathscr{O}$. Then, in this theory all output value of all observable are determined by the value of Z. We shall refer to this type theory as strong deterministic theory and write the set of strong deterministic theories as sDET.

One of the most important feature of quantum theory may be the existence of simultaneously immeasurable observables. In this section, we shall define the concept of simultaneous measurability in operational theories. It should be noted that the definition of simultaneous measurability we shall define in this section is equivalent to operator commutation in quantum theory. It also should be noted that we can define the concept of simultaneous measurability without using a probability.

First, we shall introduce the concept of function of observable. Suppose that A and C are observables of S. The observable A is a function of C, if there exists a surjective function $f : \mathbb{K}_C \to \mathbb{K}_A$ for observables A and C such that $f\left(c\right) = a$. If A is a function of C, we shall denote $A = f\left(C\right)$. By definition, the output value of A can be determined by output value of C. For this reason, we shall say that C implies A [44]and write

$$C \leqq A \tag{3.9}$$

if there exists a function f such that $A = f(C)$. Roughly speaking, if $C \leqq A$, C is a more detailed observation than A. It is easily seen that the the implication on $\mathscr{O}$ defined above defines partial order on $\mathscr{O}$;

(i) $A \leqq A$ for any A.

(ii) If $C \leqq A$ and $D \leqq C$, then $D \leqq A$.

(iii) $A \leqq B$ and $B \leqq A$, then $A = B$.

These properties are proved as follows: Let $\mathrm{id}_{\mathbb{R}}$ be an identity function $\mathbb{R}$. We can say that A implies A with $\mathrm{id}_{\mathbb{R}}$. Therefore, (i) is valid. If $C \to A$ and $D \to C$, then there are functions f_C^A and f_D^C such that $A = f_C^A(C)$ and $C = f_D^C(D)$. Putting $f_D^A := f_C^A \circ f_D^C$, then, $A = f_D^A(D)$. Therefore, $D \leqq A$.

It is notable that the output value of A is determined at the moment when the observable C is measured. In this case, we can associate *the joint probability distribution* $p(a, c\,|P, A, C)$ with the "joint observable" (A, C) for given P. This may be expressed as

$$p(a, c\,|P, A, C) = p(a\,|c, A)\, p(c\,|P, C) \tag{3.10}$$

where $p(a\,|c, A)$ is the probability that A takes a value a when C takes a value c.[3] Since $A = f(C)$, we have that

$$p(a\,|c) = \begin{cases} 1, & a = f(c) \\ 0, & a \neq f(c) \end{cases}. \tag{3.11}$$

The output probability distribution $p(a\,|A, P)$ of observable $A = f(C)$ given preparation P can be expressed as

$$\begin{aligned} p(a\,|P) &= \int_{\mathbb{K}_C} p(a, c\,|P, A, C)\, \mathrm{d}c \\ & \int_{\mathbb{K}_C} p(a\,|c, A)\, p(c\,|P, C)\, \mathrm{d}c. \end{aligned} \tag{3.12}$$

[3]It should be noted that we have now assumed that the form of the function f does not depend on the preparation P. We shall treat the case which f depends on P in chapter 5.

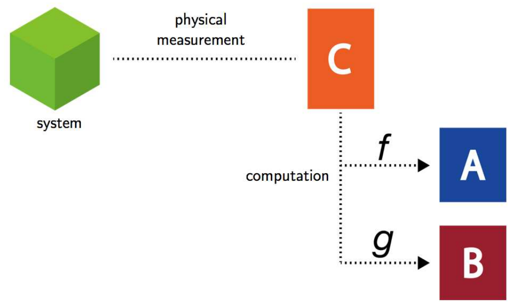

Figure 3.2.1: simultaneous observables

Now, we shall introduce the concept of simultaneous measurability. Suppose that A, B are the observables of S. A pair of observables A and B are *simultaneous measurable*, if there exists an observable C such that C implies A and B, i.e.,

$$C \leqq A \quad \wedge \quad C \leqq B. \tag{3.13}$$

where $\wedge$ represents the logical conjunction (i.e., "and"). This definition of simultaneous measurability can be paraphrased as follows: a pair of observables A and B are simultaneous measurable, if there exists an observable C, a surjective function $f : \mathbb{K}_C \to \mathbb{K}_A$ and a surjective function $g : \mathbb{K}_C \to \mathbb{K}_B$ such that

$$A = f(C), \tag{3.14}$$

$$B = g(C). \tag{3.15}$$

Since the output value of C determines the output values of A and B all at once, we shall call a pair A and B simultaneous measurable [43] (Figure 3.2.1).

Hereafter, "simultaneous measurable" is also referred to as "compatible". Let us denote by $A \rightleftharpoons B$ simultaneous measurable observables. We can easily generalize this definition of simultaneous measurability to arbitrary number of observables, e.g., $A_1 \rightleftharpoons A_2 \rightleftharpoons \ldots \rightleftharpoons A_n$. It should be noted that applying our definition of simultaneous measurability in quantum mechanics, it turns out to be equivalent to the operator commutativity. We shall see this in section 4.4.

It is remarkable that we cannot associate the joint probability distribution of observables A and B if they are not compatible since the theory does not bear the ability to predict and describe such a joint probability. This is one of motivations to consider quasi-probability in quantum mechanics.

Let $\mathscr{O}_{\mathrm{SM}} \subset \mathscr{O}$ be a subset of $\mathscr{O}$ such that any observable in $\mathscr{O}_{\mathrm{SM}}$ are mutually compatible. Let SM be a set of operational theories in whose observables are mutually compatible with any observable, i.e.,

$$\mathsf{SM} = \left\{ (\mathscr{O}, \mathscr{P}, \omega) \,|\, \mathscr{O} = \mathscr{O}_{\mathrm{SM}} \right\}. \tag{3.16}$$

The theories in SM are called *simple* if there exist an unique observable Z such that $Z \leqq A$ for any observable A. It may be natural to regard the physical SM theories as simple. In next section, we see that the classical mechanics is none other than the simple theory. We refer to Z as the *fundamental observable*. Let SIMP be a set of all simple SM theories. In simple theories, there is a surjective function $f^A : \mathbb{K}_Z \to \mathbb{K}_A$ for any observable A such that

$$A = f^A (Z). \tag{3.17}$$

Recall that a preparation must correspond to the observational proposition. It can be shown that the preparation of such a form as $P_z =$ "$Z = z$" is pure state. The proof is given as follows: Suppose P_z is mixed, then there exists pure preparations P_1 and P_2 ($P_1 \neq P_2$) such that

$$P_z = \langle k, 1-k; P_1, P_2 \rangle \tag{3.18}$$

where $k = p(P_1 \,|P_z) \in [0,1]$. The output probability distribution of any observable A on P_z is

$$p(a\,|P_z, A) = kp(a\,|P_1, A) + (1-k)\,p(a\,|P_2, A). \tag{3.19}$$

where $p(a\,|A, P_1) \neq p(a\,|A, P_2)$ since $P_1 \neq P_2$. If $P_1 =$ "$X = x$", there is a function f^X such that $X = f^X(Z)$ since Z is fundamental; $Z \leqq X$. Then,

$$k = p(x\,|P_z, X) = \begin{cases} 1, & x = f^X(z) \\ 0, & x \neq f^X(z) \end{cases}. \tag{3.20}$$

Therefore, there is no way to justify the equation (3.19) except for trivial mixing. We can also show that any pure preparation P in SIMP theories is $P =$ "$Z = z$" for some z. Suppose that $P_x =$ "$X = x$" is pure. The output probability distribution of any A on P_x is given by

$$p(a\,|P_x, A) = \int_{\mathbb{K}_Z} p(a\,|z, Z, A)\, p(z\,|Z, P_x)\, \mathrm{d}z \tag{3.21}$$

since Z is fundamental observable. From this equation, P_x turns out the mixed state;

$$P_x = \langle p(z\,|Z, P_x)\,; P_z \rangle_{z \in \mathbb{K}_z}\,. \tag{3.22}$$

It is clear that the case which P_x stays in pure is if and only if $X = Z$. Therefore, we can conclude that any pure preparation P in SIMP theories is $P =$ "$Z = z$" for some z. In summary we have shown that

$$\mathscr{P}^{\mathsf{SIMP}}_{\text{pure}} = \{\text{“}Z = z\text{”}\,|z \in \mathbb{K}_z\}\,. \tag{3.23}$$

This observation leads to the fact that theories in SIMP is intrinsically deterministic. Hence, we reached following theorem;

Theorem 3.1. *Theories in* SIMP *are strong deterministic;*

$$\mathsf{SIMP} = \mathsf{sDET}. \tag{3.24}$$

3.3 Examples of Operational Probabilistic Theory

In this section, we shall consider operational frameworks of theories in physics. We may see that the operational probabilistic theory, in general, gives us the panoramic view of physical theory.

3.3.1 Classical Mechanics

Although there are several mutually equivalent formalisms of classical mechanics, we shall focus on the phase space formalism developed by Hamilton, Boltzmann, Poincaré and Gibbs in late 19th century by briefly reconstructing from the operational viewpoint. The first axiom of operational classical mechanics is about the position and the momentum:

CM I. (position and momentum) The position observable $\mathbf{Q}$ and the momentum observable $\mathbf{P}$ are mutually simultaneously measurable.

This postulate CM I says that the position and the momentum of system, in principle, can be measured at once with any accuracy. Let us define the set Γ which consists of all possible values of position and momentum:

$$\Gamma = \{\mathbf{q}, \mathbf{p} \,|\, \mathbf{q} \in \mathbb{K}_{\mathbf{Q}},\, \mathbf{p} \in \mathbb{K}_{\mathbf{P}}\} = \mathbb{K}_{\mathbf{Q}} \times \mathbb{K}_{\mathbf{P}}. \tag{3.25}$$

The American physicist Gibbs called this **q-p** space the "*phase space*". We shall denote $\gamma = (\mathbf{q}, \mathbf{p})$ the element of phase space Γ [4].

CM II. (Observables) For any operationally equivalent observables of the classical mechanical system, e.g., $A \in \mathscr{O}$, there exists a function $A : \Gamma \to \mathbb{K}_A$ such that the value of observable A is given by $A(\gamma)$ if values of the position and the momentum is $\gamma = (\mathbf{q}, \mathbf{p})$.

It should be noted that we denote the function by the same symbol of observable, $A : \Gamma \to \mathbb{K}_A$. This notation would not generate any confusion because all values of observable A is completely determined by this function. If we rewrite the postulate CM II by using the

[4] It should be noted that there exists the classical mechanics whose phase space does not given by position and momentum such as equation (3.25). We shall not address this type exceptions.

notion of probability, it is that

$$p(a\,|\gamma, A) = \begin{cases} 1, & a = A(\gamma) \\ 0, & a \neq A(\gamma) \end{cases}. \tag{3.26}$$

We shall refer to probability (3.26) as an *indicator function*, considered as a function of phase space point γ. It is remarkable that all classical mechanical observables are simultaneously measurable since the function appeared in postulate CM II guarantees that if we know values of the position and the momentum, we can know the value of any observables in classical mechanics. For this reason, we shall call the position and the momentum the *fundamental observables* in classical mechanics.

The third axiom is about preparations.

CM III. (Preparations) Every operationally equivalent preparations on classical mechanical system correspond to the distribution on phase space Γ. We shall denote this phase space distribution given preparation P by $p(\gamma\,|\Gamma, P)$.

For example, a pure preparation corresponds to the assignment of point γ' in the phase space Γ. We shall denote this pure preparation by $P_{\gamma'}$, in which case the phase space distribution is given by

$$p(\gamma\,|P_{\gamma'}) = \begin{cases} 1, & \gamma = \gamma' \\ 0, & \gamma \neq \gamma' \end{cases}. \tag{3.27}$$

For the mixed preparation which mainly appears in statistical mechanics, the canonical distribution is one of the example;

$$p(\gamma\,|P) = \frac{1}{Z}e^{-\beta H(\gamma)}. \tag{3.28}$$

From the above postulates, the output probability distribution $p(a\,|A, P)$ of observable A

given preparation P in classical mechanics is given by

$$\begin{aligned} p\left(a\left|P,A\right.\right) &= \int_{\Gamma} p\left(a\left|\gamma,A\right.\right) p\left(\gamma\left|P\right.\right) \mathrm{d}\gamma \\ &= \int_{\Gamma} \delta\left(a - A\left(\gamma\right)\right) p\left(\gamma\left|P\right.\right) \mathrm{d}\gamma \end{aligned} \tag{3.29}$$

where δ is Dirac's delta. Let us define the subset of Γ as

$$\Gamma_a = \left\{\gamma \in \Gamma \middle| A\left(\gamma\right) = a\right\} \tag{3.30}$$

and call it Γ_a *constant-A surface* in phase space. Then, the equation (3.29) becomes the integral on the surface Γ_a :

$$p\left(a\left|P,A\right.\right) = \int_{\Gamma_a} p\left(\gamma\left|P\right.\right) \mathrm{d}\gamma. \tag{3.31}$$

The expectation value of the observable A given preparation P is defined as

$$\left\langle A\right\rangle_P := \int_{\mathbb{K}_A} ap\left(a\left|P,A\right.\right) \mathrm{d}a. \tag{3.32}$$

By substituting (3.29) into (3.32), the expectation value becomes average of $A\left(\gamma\right)$, with $p\left(\gamma\left|P\right.\right)$ being the weights;

$$\left\langle A\right\rangle_P = \int_{\mathbb{K}_A} \int_{\Gamma} a\delta\left(a - A\left(\gamma\right)\right) p\left(\gamma\left|P\right.\right) \mathrm{d}\gamma \mathrm{d}a \tag{3.33}$$

$$= \int_{\Gamma} A\left(\gamma\right) p\left(\gamma\left|P\right.\right) \mathrm{d}\gamma. \tag{3.34}$$

In summary, classical mechanics is the theory $(\mathscr{O}_{\mathrm{cl}}, \mathscr{P}_{\mathrm{cl}}, \omega)$ where

$$\mathscr{O}_{\mathrm{cl}} = \left\{\mathrm{functions\,on\,}\Gamma\right\}, \tag{3.35}$$

$$\Gamma_{\mathrm{cl}} = \left\{\mathrm{distributions\,on\,}\Gamma\right\} \tag{3.36}$$

and the probability weight is given by

$$\omega\left(P,A,a\right) = p\left(a\left|P,A\right.\right) = \int_{\Gamma} p\left(a\left|\gamma,A\right.\right) p\left(\gamma\left|P\right.\right) \mathrm{d}\gamma. \tag{3.37}$$

Therefore, classical mechanics is SIMP theory described in previous section.

Let us cite some typical examples. First, we shall consider the case that the target observable A is the position $\mathbf{Q}$. The indicator function is

$$p\left(\mathbf{q}\,|\gamma,\Gamma,\mathbf{Q}\,\right)=\begin{cases}1, & \pi_{\mathbf{Q}}\left(\gamma\right)=\mathbf{q}\\ 0, & \pi_{\mathbf{Q}}\left(\gamma\right)\neq\mathbf{q}\end{cases} \tag{3.38}$$

where $\pi_{\mathbf{Q}}:\Gamma=\mathbb{K}_{\mathbf{Q}}\times\mathbb{K}_{\mathbf{P}}\rightarrow\mathbb{K}_{\mathbf{Q}}$ is a $\mathbf{Q}$-projection on Γ. Therefore,

$$p\left(\mathbf{q}\,|P,\mathbf{Q}\,\right) = \int_{\Gamma}p\left(\mathbf{q}\,|\gamma,\mathbf{Q}\,\right)p\left(\gamma\,|P\,\right)\mathrm{d}\gamma \tag{3.39}$$

$$= \int_{\Gamma_{\mathbf{P}}}p\left(\mathbf{q},\mathbf{p}\,|P\,\right)\mathrm{d}\mathbf{p}. \tag{3.40}$$

The more trivial and important example is the case that the observable A is position and momentum $(\mathbf{Q},\mathbf{P})$,

$$p\left(\mathbf{q},\mathbf{p}\,|\gamma,\Gamma,(\mathbf{Q},\mathbf{P})\,\right)=\begin{cases}1, & \gamma=(\mathbf{q},\mathbf{p})\\ 0, & \gamma\neq(\mathbf{q},\mathbf{p})\end{cases}. \tag{3.41}$$

Then, the output probability distribution is the phase space distribution in this case;

$$p\left(\mathbf{q},\mathbf{p}\,|P,(\mathbf{Q},\mathbf{P})\,\right) = \int_{\Gamma}p\left(\mathbf{q},\mathbf{p}\,|\gamma,(\mathbf{Q},\mathbf{P})\,\right)p\left(\gamma\,|P\,\right)\mathrm{d}\gamma. \tag{3.42}$$

$$= p\left(\mathbf{q},\mathbf{p}\,|P\,\right). \tag{3.43}$$

Another easy case is, for instance, that the target observable is Hamiltonian of simple harmonic oscillator $H\left(\mathbf{q},\mathbf{p}\right)=\alpha\mathbf{p}^2+\beta\mathbf{q}^2$, $(\alpha,\beta\in\mathbb{R})$. The constant-energy surface is

$$\Gamma_h=\left\{\gamma=(\mathbf{q},\mathbf{p})\in\Gamma\middle|\alpha\mathbf{p}^2+\beta\mathbf{q}^2=h\right\}. \tag{3.44}$$

The indicator function is given by

$$p\left(h\,|\gamma,H\,\right)=\begin{cases}1, & h\in\Gamma_h\\ 0, & h\notin\Gamma_h\end{cases}. \tag{3.45}$$

The fiducial outcome probability distribution of energy given P is given by

$$p\left(h\left|H,P\right.\right)=\int_{\Gamma_h}p\left(\gamma\left|P\right.\right)\mathrm{d}\gamma. \tag{3.46}$$

3.3.2 Quantum Mechanics

Usual quantum mechanics is indeterministic theory and there exist simultaneous immeasurable observables. The mathematical structure of this theory is given by so-called von Neumann-Dirac axioms of quantum theory [5, 54]. In this subsection, we shall associate the vN-D axioms with operational probabilistic theory:

QM I. (Observables) Every operationally equivalent observables on quantum mechanical system corresponds to the self-adjoint operators in complex Hilbert space $\mathcal{H}$. We shall use same letter A for representing an observable $A\in\mathscr{O}$. An observational proposition "$A=a$" in quantum theory corresponds to spectral projection $E^A\left(a\right)=\left|a\right\rangle\left\langle a\right|$.

From this postulate, we can see the equivalence between the definition of simultaneous measurability in section 4.2 and operator commutation.

Theorem 3.2. *Two observables A and B simultaneous measurable in quantum mechanics is equivalent to requirement that $E^A\left(a\right)$ and $E^B\left(b\right)$ commutes; $\left[E^A\left(a\right),E^B\left(b\right)\right]=E^A\left(a\right)E^B\left(b\right)-E^B\left(b\right)E^A\left(a\right)=0$.*

Proof. From the definition of simultaneous measurability in section 4.2, if two observables A and B are compatible, there are another observable C and functions f^A and f^B such that $A=f^A\left(C\right)$ and $B=f^B\left(C\right)$. From the assumption QM I, an observational proposition in quantum theory represented by spectral projection. Then,

$$E^A\left(a\right)=E^C\left(f^A\left(c\right)\right),\quad E^B\left(b\right)=E^C\left(f^B\left(c\right)\right). \tag{3.47}$$

We immediately observe that

$$\left[E^C\left(f^A\left(c\right)\right),E^C\left(f^B\left(c\right)\right)\right]=0. \tag{3.48}$$

Conversely, suppose that $\left[E^A(a), E^B(b)\right] = 0$. Let us define

$$E^C(a,b) = E^A(a)\, E^B(b) = E^B(b)\, E^A(a)\,. \tag{3.49}$$

It is easily seen that $E^C(a,b)$ is the two-dimensional spectral projection. Let π^A and π^b are projections on $\mathbb{K}_A \times \mathbb{K}_B$; $\pi^A(a,b) = a$ and $\pi^B(a,b) = b$. We find that $A = \pi^A(C)$ and $B = \pi^B(C)$. □

In quantum mechanics, there are observables which do not commute. The most important typical example is position operator $\mathbf{Q}$ and momentum operator $\mathbf{P}$. These satisfies the canonical commutation relation;

$$[\mathbf{Q}, \mathbf{P}] := \mathbf{QP} - \mathbf{PQ} = i. \tag{3.50}$$

Another typical example is a spin observable. Let us denote Pauli matrices by σ_x, σ_y and σ_z. Spins with observable of spin 1/2 system is represented by Pauli matrices. These matrices satisfies

$$[\sigma_i, \sigma_j] = 2\epsilon_{ijk}\sigma_k$$

where ϵ_{ijk} is the Levi-Civita symbol and $i, j, k = x, y, z$.

QM II. (Preparations) Every operationally equivalent preparations on quantum mechanical system corresponds to the density operator in complex Hilbert space $\mathcal{H}$. We shall denote the density operator which corresponds to the preparation P by ρ_P.

QM III. (Born's rule) The output probability distribution of an observable A on the preparation P is given by Born's rule;

$$p(a\,|A, P) = \mathrm{Tr}\left[E^A(a)\, \rho_P\right]. \tag{3.51}$$

3.4 Composite System and Entanglement

In this section, we introduce concepts of composite system and entanglement. It is well known that entanglement is one of the important features of quantum theory. We define the entanglement in operational framework in this section and use it to classify theories.

Let S_1 and S_2 be physical systems which is described by the theories $(\mathscr{O}_1, \mathscr{P}_1, \omega_1)$ and $(\mathscr{O}_2, \mathscr{P}_2, \omega_2)$, respectively. A system S is called the composite system of S_1 and S_2 if it is completely characterized by S_1 and S_2. In this case, we call S_1 and S_2 partial systems of S. This means that all measurement apparatuses of S are constituted by the apparatuses of S_1 and S_2 only. In addition, we propose the assumption that every observation can be taken place simultaneously on S_1 and S_2. Let $(\mathscr{O}, \mathscr{P}, \omega)$ be the theory that describes the composite system S. The output value c of an observable $C \in \mathscr{O}$ of S can be written as

$$c = (a, b) \tag{3.52}$$

where a and b are the output values of $A \in \mathscr{O}_1$ and $B \in \mathscr{O}_2$ respectively. We shall denote such observable C as $C = (A, B)$. The output probability distribution of (A, B) on some preparation P is written as

$$p\left(a_1, a_2 \left| P, A_1, A_2\right.\right). \tag{3.53}$$

It may be most important feature of composite system that the preparation P of composite system S is not determined by the observational propositions of S_1 and S_2. This is because it is possible that there exists an interdependence between S_1 and S_2. Such situation happens in the case that the preparation procedure of composite system is made by interaction of S_1 and S_2. If the preparation P of the composite system S is made by preparations P_1 and P_2 of partial systems S_1 and S_2 independently, we call P product preparation and denote by $P = P_1 \otimes P_2$. In this case, the output probability distribution of the observable (A, B) of S may becomes

$$p\left(a_1, a_2 \left| P_1 \otimes P_2, A_1, A_2\right.\right) = p\left(a_1 \left| P_1, A_1\right.\right) p\left(a_2 \left| P_2, A_2\right.\right). \tag{3.54}$$

If the preparation P of S is not written as the equation (3.54), we shall refer to P as *entangled* preparation.

Let us introduce the marginal probability. Let A and B are observables of S_1 and S_2. The marginal probability distribution of A and B on P is respectively defined by

$$p\left(a\,|P,A,B\,\right):=\int p\left(a,b\,|P,A,B\,\right)\mathrm{d}b \tag{3.55}$$

$$p\left(b\,|P,A,B\,\right):=\int p\left(a,b\,|P,A,B\,\right)\mathrm{d}a \tag{3.56}$$

It is usually required the *no-signaling condition*, that is, if A and P are an arbitrary observable of S_1 and preparation on S, respectively, the marginal probability distribution of an arbitrary observable A on S_1 and arbitrary preparation P on S satisfies

$$p\left(a\,|P,A,B\,\right)=p\left(a\,\middle|P,A,B'\,\right) \tag{3.57}$$

for any A and P and for all $B,B'\in\mathscr{O}_2$. This means that the observation on S_2 never affect the probability distribution on S_1.

We shall introduce the Clauser-Horne-Shimony-Holt (CHSH) parameter [55] which statistically characterizes the existence of entanglement and hence it expresses that the theory bear the entanglement or not. We will calculate the CHSH parameter in several ontological model in next chapter.

Suppose that A and A' are observables of S_1 and t B and B' are observables of S_2. We assume that these observables are two-valued ones. Let us write the output values of A, B, A', and B' as a_i, b_j, a'_k and b'_l respectively. In addition, we require that for simplicity

$$a_i,b_j,a'_k,b'_l=\begin{cases}+1 & i,j,k,l=1\\ -1 & i,j,k,l=0\end{cases}. \tag{3.58}$$

Suppose that the theory provides the joint probability such as $p\left(a_i,b_j\,|P,A,B\,\right)$. The marginals are given by

$$p\left(a_i\left|P,A,B\right.\right)=\sum_j p\left(a_i,b_j\left|P,A,B\right.\right), \tag{3.59}$$

$$p\left(b_j\left|P,A,B\right.\right)=\sum_i p\left(a_i,b_j\left|P,A,B\right.\right). \tag{3.60}$$

The no-signaling condition for P is

$$p\left(a_i\left|A,B,P\right.\right)=p\left(a_i\left|A,C,P\right.\right) \tag{3.61}$$

For given preparation P, the observables A, B and C . In this case, we write the marginal probability as $p\left(a_i\left|A,P\right.\right)$. The expectation value of the observable A is given by

$$\begin{aligned}\left\langle A\right\rangle_P : &= \sum_i a_i p\left(a_i\left|A,P\right.\right)=p\left(+\left|A,P\right.\right)-\left(1-p\left(+\left|A,P\right.\right)\right) \\ &= 2p\left(+\left|A,P\right.\right)-1.\end{aligned} \tag{3.62}$$

Since $0\le p\left(+\left|A,P\right.\right)\le 1$, we have

$$-1\le\left\langle A\right\rangle_P\le 1 \tag{3.63}$$

for any P. Next, we shall consider the expectation of product of observables e.g. A and B;

$$\begin{aligned}C_P\left(A,B\right) &:= \sum_{i,j} a_i b_j p\left(a_i,b_j\left|A,B,P\right.\right) \\ &= p\left(a_1,b_1\left|A,B,P\right.\right)-p\left(a_1,b_0\left|A,B,P\right.\right) \\ &\quad -p\left(a_0,b_1\left|A,B,P\right.\right)+p\left(a_0,b_0\left|A,B,P\right.\right) \\ &= p\left(+,+\left|A,B,P\right.\right)-p\left(+,-\left|A,B,P\right.\right) \\ &\quad -p\left(-,+\left|A,B,P\right.\right)+p\left(-,-\left|A,B,P\right.\right)\end{aligned} \tag{3.64}$$

Since $\sum_{i.j} p\left(a_i,b_j\left|A,B,P\right.\right)=1$, we have

$$-1\le C_P\left(A,B\right)\le 1. \tag{3.65}$$

Let us define the following parameter of observables A, B, A', and B' for the preparation P;

$$\texttt{CHSH}_P := C_P(A,B) + C_P(A',B) + C_P(A',B') - C_P(A,B'). \tag{3.66}$$

If A and B are probabilistic independent observables on P;

$$p(a_i, b_j | A, B, P) = p(a_i | A, B, P)\, p(b_j | A, B, P) \tag{3.67}$$

then, we observe

$$\begin{aligned} C_{P_{\text{ind.}}}(A,B) &= \sum_{i,j} a_i b_j p(a_i, b_j | A, B, P_{\text{ind.}}) \\ &= \sum_{i,j} a_i p(a_i | A, P_{\text{ind.}})\, b_j p(b_j | B, P_{\text{ind.}}) \\ &= \langle A\rangle_{P_{\text{ind.}}} \langle B\rangle_{P_{\text{ind.}}}. \end{aligned} \tag{3.68}$$

If P is probabilistic independent preparation for all observables, the CHSH parameter becomes

$$\begin{aligned} \texttt{CHSH}_{P_{\text{ind.}}} &= \langle A\rangle_{P_{\text{ind.}}} \langle B\rangle_{P_{\text{ind.}}} + \langle A'\rangle_{P_{\text{ind.}}} \langle B\rangle_{P_{\text{ind.}}} \\ &\quad + \langle A'\rangle_{P_{\text{ind.}}} \langle B'\rangle_{P_{\text{ind.}}} - \langle A\rangle_{P_{\text{ind.}}} \langle B'\rangle_{P_{\text{ind.}}} \\ &= \langle A\rangle_{P_{\text{ind.}}} \left\{\langle B\rangle_{P_{\text{ind.}}} - \langle B'\rangle_{P_{\text{ind.}}}\right\} + \langle A'\rangle_{P_{\text{ind.}}} \left\{\langle B\rangle_{P_{\text{ind.}}} + \langle B'\rangle_{P_{\text{ind.}}}\right\} \end{aligned} \tag{3.69}$$

Using equation (3.63), we observe that

$$\begin{aligned} \left|\texttt{CHSH}_{P_{\text{ind.}}}(A,A',B,B')\right| &= \left|\langle A\rangle_{P_{\text{ind.}}} \left\{\langle B\rangle_{P_{\text{ind.}}} - \langle B'\rangle_{P_{\text{ind.}}}\right\} + \langle A'\rangle_{P_{\text{ind.}}} \left\{\langle B\rangle_{P_{\text{ind.}}} + \langle B'\rangle_{P_{\text{ind.}}}\right\}\right| \\ &\le \left|\langle A\rangle_{P_{\text{ind.}}} \left\{\langle B\rangle_{P_{\text{ind.}}} - \langle B'\rangle_{P_{\text{ind.}}}\right\}\right| + \left|\langle A'\rangle_{P_{\text{ind.}}} \left\{\langle B\rangle_{P_{\text{ind.}}} + \langle B'\rangle_{P_{\text{ind.}}}\right\}\right| \\ &= \left|\langle A\rangle_{P_{\text{ind.}}}\right| \left|\langle B\rangle_{P_{\text{ind.}}} - \langle B'\rangle_{P_{\text{ind.}}}\right| + \left|\langle A'\rangle_{P_{\text{ind.}}}\right| \left|\langle B\rangle_{P_{\text{ind.}}} + \langle B'\rangle_{P_{\text{ind.}}}\right| \\ &\le 2. \end{aligned} \tag{3.70}$$

Whereas, it is easy to see that the general no-signaling theories satisfies[73]

$$\left|\texttt{CHSH}_{P_{\text{NS}}}\right| \le 4. \tag{3.71}$$

The sketch of proof of (3.71) is as follows; The marginal probabilities is given by

$$p\left(a_1\left|A,P\right.\right)=p\left(a_1,b_1\left|A,B,P\right.\right)+p\left(a_1,b_0\left|A,B,P\right.\right)=p\left(a_1,b_1'\left|A,B,P\right.\right)+p\left(a_1,b_0'\left|A,B,P\right.\right), \tag{3.72}$$

$$p\left(a_0\left|A,P\right.\right)=p\left(a_0,b_1\left|A,B,P\right.\right)+p\left(a_0,b_0\left|A,B,P\right.\right)=p\left(a_0,b_1'\left|A,B,P\right.\right)+p\left(a_0,b_0'\left|A,B,P\right.\right), \tag{3.73}$$

and so on. The no-signaling conditions can be written explicitly as

$$p\left(a_1,b_1\left|A,B,P\right.\right)+p\left(a_0,b_1\left|A,B,P\right.\right)=p\left(a_1,b_1'\left|A,B,P\right.\right)+p\left(a_1,b_0'\left|A,B,P\right.\right), \tag{3.74}$$

$$p\left(a_1,b_1\left|A,B,P\right.\right)+p\left(a_1,b_0\left|A,B,P\right.\right)=p\left(a_1,b_1'\left|A,B,P\right.\right)+p\left(a_1,b_0'\left|A,B,P\right.\right), \tag{3.75}$$

etc. Then, if we can assign the probability to be valid the no-signaling condition (3.74) , and (3.75), and the equation (3.71). The tables 3.4 and 3.4 shows that it is possible.

$p(a_1,b_1\|A,B,P)$	0	$p(a_1',b_1\|A,B,P)$	1	$p(a_1,b_1'\|A,B,P)$	0	$p(a_1',b_1'\|A,B,P)$	1
$p(a_1,b_0\|A,B,P)$	0	$p(a_1',b_0\|A,B,P)$	0	$p(a_1,b_0'\|A,B,P)$	0	$p(a_1',b_0'\|A,B,P)$	0
$p(a_0,b_1\|A,B,P)$	1	$p(a_0',b_1\|A,B,P)$	0	$p(a_0,b_1'\|A,B,P)$	1	$p(a_0',b_1'\|A,B,P)$	0
$p(a_0,b_0\|A,B,P)$	0	$p(a_0',b_0\|A,B,P)$	0	$p(a_0,b_0'\|A,B,P)$	0	$p(a_0',b_0'\|A,B,P)$	0
$C_P(A,B)$	1	$C_P(A',B)$	1	$C_P(A,B')$	1	$C_P(A',B')$	1

Table 3.1: Joint probabilities and Correlations

$p(a_1\|A,P)$	0	$p(a_0\|A,P)$	1
$p(a_1,b_1\|A,B,P)+p(a_1,b_0\|A,B,P)$	0	$p(a_0,b_1\|A,B,P)+p(a_0,b_0\|A,B,P)$	1
$p(a_1,b_1'\|A,B,P)+p(a_1,b_0'\|A,B,P)$	0	$p(a_0,b_1'\|A,B,P)+p(a_0,b_0'\|A,B,P)$	1
$p(b_1\|B,P)$	1	$p(b_0\|B,P)$	0
$p(a_0,b_1\|A,B,P)+p(a_1,b_1\|A,B,P)$	1	$p(a_0,b_0\|A,B,P)+p(a_1,b_0\|A,B,P)$	0
$p(a_0',b_1\|A,B,P)+p(a_1',b_1\|A,B,P)$	1	$p(a_0',b_0\|A,B,P)+p(a_1',b_0\|A,B,P)$	0
$p(a_1'\|A,P)$	1	$p(a_0'\|A,P)$	0
$p(a_1',b_1\|A,B,P)+p(a_1',b_0\|A,B,P)$	1	$p(a_0',b_1\|A,B,P)+p(a_0',b_0\|A,B,P)$	0
$p(a_1',b_1'\|A,B,P)+p(a_1',b_0'\|A,B,P)$	1	$p(a_0',b_1'\|A,B,P)+p(a_0',b_0'\|A,B,P)$	0
$p(b_1'\|B,P)$	1	$p(b_0'\|B,P)$	0
$p(a_0,b_1'\|A,B,P)+p(a_1,b_1'\|A,B,P)$	1	$p(a_0,b_0'\|A,B,P)+p(a_1,b_0'\|A,B,P)$	0
$p(a_0',b_1'\|A,B,P)+p(a_1',b_1'\|A,B,P)$	1	$p(a_0',b_0'\|A,B,P)+p(a_1',b_0'\|A,B,P)$	0

Table 3.2: No-signaling conditions

In quantum mechanics, those dichotomic observables can be realized by spin observables of spin-1/2 particles S_1, S_2. Let A, A', B and B' are $\vec{a}$-directed spin of S_1, $\vec{a}$'-directed spin of S_1, $\vec{b}$-directed spin of S_2, and $\vec{b}'$-directed spin of S_2, respectively. We use the notation in section 2.2. The composite quantum system is given by following postulate:

QM IV. (Composite system) The Hilbert space of the composite system S of two systems S_1 associated with $\mathcal{H}_1$ and S_2 associated with $\mathcal{H}_2$ is the tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2$.

In our dichotomic case, the Hilbert space of S is $\mathbb{C}^2 \otimes \mathbb{C}^2$. The expectation value of $\vec{a}$-spin on the state of composite system $\Psi \in \mathbb{C}^2 \otimes \mathbb{C}^2$ is given by

$$\langle A \rangle_\Psi = \langle \Psi | \, \vec{a} \cdot \vec{\sigma} \otimes I \, | \Psi \rangle \,. \tag{3.76}$$

The correlation of $\vec{a}$-spin and $\vec{b}$-spin on Ψ defined in equation (3.64) is expressed as

$$C_\Psi \left(A, B \right) = \langle \Psi | \, \vec{a} \cdot \vec{\sigma} \otimes \vec{b} \cdot \vec{\sigma} \, | \Psi \rangle \,. \tag{3.77}$$

If Ψ is the product state, i.e., $\Psi = \psi_1 \otimes \psi_2$, the correlation becomes $C_\Psi (A, B) = \langle A \rangle_\Psi \langle B \rangle_\Psi$, and then, in this case, the CHSH parameter is 2. On the other hand, if Ψ is the entangled state, e.g., the singlet state;

$$|\Psi\rangle = \frac{1}{\sqrt{2}} \left(|+\rangle \otimes |-\rangle - |-\rangle \otimes |+\rangle \right) \tag{3.78}$$

where $|\pm\rangle$ is the eigenstate of spin of z-direction with eigenvalue ± 1. For simplicity, let us assign the vector $\vec{a}$ is z-axis and the vector $\vec{b}$ is on zx plane. It may not cause the loss of generality. Here, the correlation is given by

$$C_\Psi \left(A, B \right) = -\vec{a} \cdot \vec{b}. \tag{3.79}$$

Putting that $\vec{a}$ and $\vec{b}$ are same as z-axis and the vector $\vec{a}'$ and $\vec{b}'$ are on zx plane with

$\vec{a} \cdot \vec{a}' = -\cos\phi$, $\vec{a} \cdot \vec{b}' = -\cos\phi$ and $\vec{a}' \cdot \vec{b}' = -\cos 2\phi$. The CHSH parameter becomes

$$|\mathtt{CHSH}_{\Psi}| \leq 2\sqrt{2}. \tag{3.80}$$

Hence, for the quantum entanglement state (3.78), the CHSH parameter can take a value bigger than 2.

3.5 Post-selected Measurement

In this section, we shall see the epistemological significance of post-selected measurement that is one of the key features of weak measurement procedure. For this purpose, we shall introduce the concept of a prediction and a retrodiction. Prediction is an inference that consists of guessing the probability to the certain observational proposition introduced in section 3.1. On the other hand, retrodiction, which is originated by Watanabe [52], is the converse act, that is, an inference that consists of guessing the probability to the experimental proposition e.g., $A = a$ in past when we know that the proposition e.g., $B = b$ is true. In a word, prediction is an inference of a past data from a present ones.

First, let us see the retrodictive probability in classical mechanics. We can associate the joint probability distribution of the observables A and $(\mathbf{Q}, \mathbf{P})$ as the integrand in equation (3.29) ;

$$p\,(\gamma, a\,|P, A\,) := p\,(a\,|\gamma, A\,)\,p\,(\gamma\,|P\,)\,. \tag{3.81}$$

From this, we can associate the marginal and conditional probability as

$$p\,(\gamma\,|P, A\,) := \int_{\mathbb{K}_A} p\,(\gamma, a\,|P, A\,)\,\mathrm{d}a, \tag{3.82}$$

$$p\,(\gamma\,|P, a\,) := \frac{p\,(\gamma, a\,|P, A\,)}{p\,(a\,|P, A\,)}. \tag{3.83}$$

It should be noted that the conditional probability (3.83) is omitted form of $p\,(\gamma\,|P, A = a\,)$. This conditional probability can be regarded as the retrodictive probability since it expresses the inference of phase space point γ (past date) from the observation of A (present one). We can easily observe that

$$p(\gamma|a, P) = \frac{p(\gamma|P)\, p(a|\gamma, A)}{p(a|P, A)} = \frac{p(\gamma|P)\, p(a|\gamma, A)}{\int_\Gamma p(a|\gamma', A)\, p(\gamma'|P)\, \mathrm{d}\gamma'}. \tag{3.84}$$

This equation is known as Bayes' formula. It means that if the observer does not know the output value of A, then he associates probability $p(\gamma|P)$ with phase space point γ, but after knowing that A takes a value a, he associates a revised probability $p(\gamma|P, a)$. It should be noted that we cannot associate the retrodictive probability $p(\gamma|a, P)$ in almost all cases. The predictive probability $p(a|\gamma, A)$ can be given by the postulate CMII, $p(a|\gamma, A) = \delta\left(f^A(\gamma) - a\right)$. Then, unless we know the probability $p(\gamma|P)$ for all γ, we cannot obtain the retrodictive probability $p(\gamma|a, P)$ by Bayes' formula (3.84) except for the special case that the observable A is a pair of position and momentum. Let us consider the target observable be a pair of the position and the momentum $(\mathbf{Q}, \mathbf{P})$. In this case, it is easily seen that the predictive probability $p(\gamma'|\gamma, (\mathbf{Q}, \mathbf{P}))$ and the retrodictive probability $p(\gamma|P, \gamma')$ coincide for all preparation P. Putting the equation (3.43) into the Bayes' formula, we observe that

$$\begin{aligned} p\left(\gamma|P, \gamma'\right) &= \frac{p(\gamma|P)\, p(\gamma'|\gamma, (\mathbf{Q}, \mathbf{P}))}{p(\gamma'|(\mathbf{Q}, \mathbf{P}), P)} \\ &= \frac{p(\gamma|P)}{p(\gamma'|P)} p\left(\gamma'|\gamma, (\mathbf{Q}, \mathbf{P})\right). \end{aligned} \tag{3.85}$$

Recalling the equation (3.41), we found that

$$p\left(\gamma'|\gamma, (\mathbf{Q}, \mathbf{P})\right) = p\left(\gamma'|\gamma, (\mathbf{Q}, \mathbf{P}), P\right) = \begin{cases} 1, & \gamma = \gamma' \\ 0, & \gamma \neq \gamma' \end{cases}. \tag{3.86}$$

Hence, we observe that the coincidence of predictive and retrodictive probability is one of the feature of fundamental observable. This special case is called the bilaterally determinism [44]. It may be one of the epistemically important feature of the fundamental observable in the classical physics.

Next, we shall consider the post-selected probability introduced in section 2.3 in general operational theory. It can be written by

$$p\left(A_\tau = a \,|X_t = x, Y_{t'} = y\right) \tag{3.87}$$

where t, t' and τ are parameters expressing time at which observer takes place the measurement with $t < \tau < t'$. The post-selected probability $p\left(A_\tau = a \,|X_t = x, Y_{t'} = y\right)$ is consistent with one introduced in the previous chapter. It may be expressed in terms of predictive (sequential) probability;

$$p\left(A_\tau = a \,|X_t = x, Y_{t'} = y\right) = \frac{p\left(A_\tau = a, Y_{t'} = y \,|X_t = x\right)}{\int_{\mathbb{K}_A} p\left(A_\tau = a, Y_{t'} = y \,|X_t = x\right) \mathrm{d}a}. \tag{3.88}$$

On the other hand, it can be written as

$$p\left(A_\tau = a \,|X_t = x, Y_{t'} = y\right) = \frac{p\left(X_t = x, A_\tau = a \,|Y_{t'} = y\right)}{\int_{\mathbb{K}_A} p\left(X_t = x, A_\tau = a \,|Y_{t'} = y\right) \mathrm{d}a}. \tag{3.89}$$

It is remarkable that the probabilities in equation (3.89) are all retrodictive since $t < \tau < t'$.

In terms of the concept of ensemble, equations (3.88) and (3.89) may become clear. Let $n\left(x, a, y\right)$ be a number of objects which takes a value $X = x$ at t, $A = a$ at τ, and $Y = y$ at t' , $n\left(x\right)$ be a number of objects which takes a value $X = x$ at t and $n\left(x, y\right)$ be a number of objects which takes a value $X = x$ at t and $Y = y$ at t' while observation of A at τ is taken place without checking the value. The post-selected probability may be given by

$$p\left(A_\tau = a \,|X_t = x, Y_{t'} = y\right) \approx \frac{n\left(x, a, y\right)}{n\left(x, y\right)}. \tag{3.90}$$

Similarly, the probabilities appeared in equations (3.88) and (3.89) are

$$p\left(A_\tau = a, Y_{t'} = y \,|X_t = x\right) \approx \frac{n\left(x, a, y\right)}{n\left(x\right)}, \tag{3.91}$$

$$\int_{\mathbb{K}_A} p\left(A_\tau = a, Y_{t'} = y \,|X_t = x\right) \mathrm{d}a \approx \frac{n\left(x, y\right)}{n\left(x\right)}, \tag{3.92}$$

$$p\left(X_t = x, A_\tau = a \,|Y_{t'} = y\right) \approx \frac{n\left(x, a, y\right)}{n\left(y\right)}, \tag{3.93}$$

$$\int_{\mathbb{K}_A} p\left(X_t = x, A_\tau = a \left| Y_{t'} = y\right.\right) \mathrm{d}a \approx \frac{n\left(x,y\right)}{n\left(y\right)}. \tag{3.94}$$

Notice that

$$p\left(A_\tau = a \left| X_t = x, Y_{t'} = y\right.\right) \approx \frac{n\left(x,a,y\right)}{n\left(x,y\right)} = \frac{n\left(x,a,y\right)}{n\left(x\right)} \frac{n\left(x\right)}{n\left(x,y\right)} \tag{3.95}$$

$$p\left(A_\tau = a \left| X_t = x, Y_{t'} = y\right.\right) \approx \frac{n\left(x,a,y\right)}{n\left(x,y\right)} = \frac{n\left(x,a,y\right)}{n\left(y\right)} \frac{n\left(y\right)}{n\left(x,y\right)} \tag{3.96}$$

Therefore, we arrived at the equations (3.88) and (3.89). This consideration asserts that the post-selected probability is concerned with by the inference without being tied by the concepts of prediction and retrodiction, in which there is no time direction of inference.

Chapter 4

Quasi-Probabilistic Ontological Model

In this chapter, we clarify the conceptual significance of quasi-probability in quantum mechanics. We show that by regarding the quasi-probability introduced in chapter 2 as the fundamental element of quantum mechanics, one can interpret the quantum mechanics realistically in a natural way. The realistic interpretation of quantum theory, which is sometimes called an ontological interpretation, is the one that quantum mechanics possesses something physical reality such as classical mechanics has, whereas there may be no physical reality in an epistemological interpretation[1] (e.g., orthodox or Copenhagen interpretation or quantum bayesianism [57]). As seen in the previous chapter, one can imagine some physical realities in classical mechanics through its deterministic nature. Therefore, it may be no exaggeration to say that the ontological interpretation is achieved by constructing the classical mechanics like alternative theory to quantum mechanics. Such an alternative theory is called hidden variable theory or ontological model. There have been several attempts to interpret quantum mechanics realistically from the very beginning of quantum theory, and the most well-known hidden variable theory may be Bohmian mechanics advocated by Bohm in 1952 [2, 3] (The recent developments and applications of Bohmian mechanics can be seen, for example, in Refs. [63, 64, 65, 66, 67, 68].). It is recently animating to study which seeks to obtain the answer the

[1]For the detail of he interpretation of quantum theory, see Refs. [58, 59]

question, on the assumption that the ontological model of quantum theory exists, whether the quantum state ψ represents something about reality (ontic), or the observer's knowledge or belief (epistemic). Pusey, Barrett and Rudolph have proven a theorem that the quantum state must be ontic in the broad class of ontological model [9]. It is, however, pointed out [10] that the conventional framework of ontological model (i.e., the model proposed by Harrigan and Spekkens [8]) cannot accommodate Bohmian mechanics. Below we show that this is no longer the case if the framework of ontological model is extended properly. This section is organized as follows. In section 4.1, we review Bohmian mechanics briefly and in section 4.2 we embed the quasi-probability underlying the weak value into Bohmian mechanics. In section 4.3, after reviewing the conventional ontological model, we extend the framework of ontological models by introducing a certain contextuality. Finally, we propose a quasi-probabilistic ontological model and show the Bohmian mechanics can be regarded as it in section 4.4.

4.1 Bohmian Mechanics

We start by reviewing the Bohmian mechanics briefly on the based of the original Bohm's paper [2, 3]. The Bohmian mechanics is a well known interpretation of quantum mechanics in terms of hidden variables. We summarize first the basic postulates of this theory:

BM1 A quantum state ψ satisfies the Schrödinger equation:

$$i\hbar\frac{\partial}{\partial t}\psi = H\psi, \tag{4.1}$$

where H is Hamiltonian of a physical system (or *particle*) S.

BM2 If we write the wave function as $\psi(\mathbf{q}) = R(\mathbf{q})\, e^{iS(\mathbf{q})/\hbar}$ in the position representation where R and S are real functions, the particle momentum $\mathbf{p}$ is given by

$$\mathbf{p}_{\mathrm{B}}(\mathbf{q}) = \nabla S(\mathbf{q}). \tag{4.2}$$

BM3 The probability that a system on state ψ is at the point x, $p\left(\mathbf{q}\,|\psi,\mathbf{Q}\,\right)$, is given by

$$p\left(\mathbf{q}\,|\psi,\mathbf{Q}\,\right) = \left|\psi\left(\mathbf{q}\right)\right|^2. \tag{4.3}$$

Similarly to classical mechanics, the position and momentum introduced in the above postulates play an essential role in this theory. The precise values of the particle position and momentum must be regarded as "hidden" since we cannot measure them simultaneously. The wave function $\psi\left(\mathbf{q}\right)$ is regarded as "a mathematical representation of an objectively real field in this theory" [2]. It determines the dynamics of particle and probability distribution of position $\mathbf{Q}$ via the equation (4.2) in the BM2 and (4.3) in BM3 . For the reason that ψ determines dynamics, ψ is sometimes called the guiding wave or pilot wave. In fact, it is possible to draw a *trajectory* of particle as we see below. The particle S may be completely described by $(\mathbf{q},\psi)$, the particle position and the quantum state, since the particle momentum $\mathbf{p}$ is determined by the wave function. Since the particle moves along the trajectory in the Bohmian mechanics, this interpretation therefore eliminates the indeterminisim of the usual interpretation of quantum mechanics. In the view of Bohmian mechanics, it is considered that the indeterminism of usual interpretation is caused by our ignorance of the precise initial conditions of the particle. The postulate 3 states that the degree of our ignorance of the initial particle position $\mathbf{q}$ is given by $\left|\psi\left(\mathbf{q}\right)\right|^2$ and the postulate 2 restricts the initial particle momentum to $\mathbf{p}_{\mathrm{B}}\left(\mathbf{q}\right) = \nabla S\left(\mathbf{q}\right)$. when $\psi\left(\mathbf{q}\right) = R\left(\mathbf{q}\right)e^{iS(\mathbf{q})/\hbar}$ is the initial wave function. To summarize, the wave function ψ in Bohmian mechanics plays two roles as a guiding wave and as what assigns the distribution of particle's position. Since this interpretation does not admit the Born rule, it is an important question whether Bohmian mechanics reproduces all predictions of quantum theory. The answer to this question, in fact, was given by Bohm[2, 3] by considering the measurement procedure. It is, however, a little complicated argument.

The trajectory of particle can be drawn in Bohmian mechanics by integrating the following equations;

$$m\dot{\mathbf{q}} := \mathbf{p}_{\mathrm{B}}\left(\mathbf{q},t\right)\big|_{\mathbf{q}=\mathbf{q}(t)} = \nabla S\left(\mathbf{q},t\right)\big|_{\mathbf{q}=\mathbf{q}(t)} \tag{4.4}$$

where the dot $\dot{}$ represents the time derivative and m is the mass of particle. Suppose that the Hamiltonian of system is given by $H = \frac{1}{2m}\mathbf{P}^2 + V(\mathbf{Q})$, we have

$$\begin{aligned} \langle \mathbf{q} | H | \psi \rangle &= \left\{ \frac{(\nabla S(\mathbf{q},t))^2}{2m} - \frac{\nabla^2 R(\mathbf{q},t)}{2mR(\mathbf{q},t)} + V(\mathbf{Q}) \right\} \psi(\mathbf{q},t) \\ &\quad -i \left\{ \frac{(\nabla R(\mathbf{q},t))(\nabla S(\mathbf{q},t))}{mR(\mathbf{q},t)} + \frac{\left(\nabla^2 S(\mathbf{q},t)\right)}{2m} \right\} \psi(\mathbf{q},t). \end{aligned} \tag{4.5}$$

The Schrödinger equation becomes

$$i\frac{\partial}{\partial t}\psi(\mathbf{q},t) = -\frac{\partial S(\mathbf{q},t)}{\partial t}\psi(\mathbf{q},t) + i\frac{1}{R(\mathbf{q},t)}\frac{\partial R(\mathbf{q},t)}{\partial t}\psi(\mathbf{q},t). \tag{4.6}$$

Therefore, we obtain the following equation for R and S. Separating this equation into its real part and imaginary part, we obtain two differential equations;

$$\frac{\partial S(\mathbf{q},t)}{\partial t} = -\frac{(\nabla S(\mathbf{q},t))^2}{2m} + \frac{1}{2m}\frac{\nabla^2 R(\mathbf{q},t)}{R(\mathbf{q},t)} - V(\mathbf{Q}), \tag{4.7}$$

$$\frac{\partial R(\mathbf{q},t)}{\partial t} = -\frac{R(\mathbf{q},t)\left(\nabla^2 S(\mathbf{q},t)\right)}{2m} - \frac{(\nabla R(\mathbf{q},t))(\nabla S(\mathbf{q},t))}{m}. \tag{4.8}$$

The first equation is sometimes called the quantum Hamilton-Jacobi equation for S. It differs from classical Hamilton-Jacobi equation by the addition of a quantum potential term $V_{\mathrm{Q}}(\mathbf{q},t)$ which is defined as

$$V_{\mathrm{Q}}(\mathbf{q},t) := \frac{\nabla^2 R(\mathbf{q},t)}{R(\mathbf{q},t)}. \tag{4.9}$$

From the definition of Bohmian velocity $\dot{\mathbf{q}}$ in equation (4.4), we can draw the trajectory of particle by solving the quantum Hamilton-Jacobi equation (4.7). From this, one may infer that Bohmian mechanics is some kind of deterministic theory.

We can easily see that Bohmian mechanics is explicitly nonlocal from the definition of Bohmian velocity. Let us consider the composite system S which consists of two particles S_1 and S_2 and let Ψ be a wave function describing S. The Bohmian momentum of particle 1 may be given by

$$\mathbf{p}_{\mathrm{B}}^{1} = \mathrm{Im}\frac{\nabla_1 \Psi\left(\mathbf{q}_1,\mathbf{q}_2\right)}{\Psi\left(\mathbf{q}_1,\mathbf{q}_2\right)} \tag{4.10}$$

where $\mathbf{q}_1, \mathbf{q}_2$ are the positions of particles S_1 and S_2, respectively and ∇_1 is the derivative operator with respect to $\mathbf{q}_1$. From this, we see that Bohmian momentum of particle 1 explicitly depends on the position of particle 2. Therefore, we conclude that the Bohmian mechanics is explicitly nonlocal because of this dependence. Bell stated that this feature of Bohmian mechanics is a merit of it;

> *It is a merit of the de Broglie-Bohm version to bring this [nonlocality] out so explicitly that it cannot be ignored.* (Bell, 1980)

Holland defines the *local expectation value* of an arbitrary observable A on the position $\mathbf{x}$ as real number [69]

$$\langle A\rangle_{\psi}\left(\mathbf{q}\right) = \mathrm{Re}\frac{\psi^{*}\left(\mathbf{q}\right)\left(A\psi\right)\left(\mathbf{q}\right)}{\psi^{*}\left(\mathbf{q}\right)\psi\left(\mathbf{q}\right)}. \tag{4.11}$$

The local expectation value can be rewritten in Dirac's bra-ket notation as

$$\langle A\rangle_{\psi}\left(\mathbf{q}\right) = \mathrm{Re}\frac{\langle\mathbf{q}|\,A\,|\psi\rangle}{\langle\mathbf{q}|\psi\rangle} \tag{4.12}$$

where $|\mathbf{q}\rangle$ is an eigenstate of the position operator $\mathbf{Q}$, since

$$\frac{\psi^{*}\left(\mathbf{q}\right)\left(A\psi\right)\left(\mathbf{q}\right)}{\psi^{*}\left(\mathbf{q}\right)\psi\left(\mathbf{q}\right)} = \frac{\langle\psi|\mathbf{q}\rangle\,\langle\mathbf{q}|\,A\,|\psi\rangle}{\langle\psi|\mathbf{q}\rangle\,\langle\mathbf{q}|\psi\rangle} = \frac{\langle\mathbf{q}|\,A\,|\psi\rangle}{\langle\mathbf{q}|\psi\rangle}.$$

It is seen that the average of the local expectation value $\langle A\rangle_{\psi}\left(\mathbf{q}\right)$ with the weight $p\left(\mathbf{q}\,|\psi,\mathbf{Q}\,\right) = |\langle\mathbf{q}|\psi\rangle|^2$ (see BM3) is equals to the quantum mechanical expectation value;

$$\begin{aligned}\int \langle A\rangle_{\psi}\left(\mathbf{q}\right)p\left(\mathbf{q}\,|\psi\,\right)\mathrm{d}\mathbf{q} &= \int \mathrm{Re}\frac{\langle\mathbf{q}|\,A\,|\psi\rangle}{\langle\mathbf{q}|\psi\rangle}\,|\langle\mathbf{q}|\psi\rangle|^2\,\mathrm{d}\mathbf{q}. \\ &\int \mathrm{Re}\,\langle\psi\,|\mathbf{q}\rangle\,\langle\mathbf{q}|\,A\,|\psi\rangle\,\mathrm{d}\mathbf{q} \\ &= \langle\psi|\,A\,|\psi\rangle \end{aligned} \tag{4.13}$$

for any ψ and any A. In addition, the local expectation value can be regarded as the generalization of the equation (4.2) since the local expectation value of momentum operator $\mathbf{P}$ coincides with Bohmian momentum;

$$\begin{aligned} \langle \mathbf{P} \rangle_{\psi} (\mathbf{q}) &= \mathrm{Re} \frac{\langle \mathbf{q} | \mathbf{P} | \psi \rangle}{\langle \mathbf{q} | \psi \rangle} \qquad (4.14) \\ &= \mathrm{Re} \frac{-i\hbar \nabla \psi (\mathbf{q})}{\psi (\mathbf{q})} \\ &= \mathrm{Re} \left\{ \nabla S (\mathbf{q}) - i\hbar \frac{\nabla R (\mathbf{q})}{R (\mathbf{q})} \right\} \\ &= \mathrm{Im} \frac{\nabla \Psi (\mathbf{q})}{\Psi (\mathbf{q})} = \mathbf{p}_{\mathrm{B}} (\mathbf{q}) . \qquad (4.15) \end{aligned}$$

In virtue of equation (4.13) and (4.14), the local expectation value of arbitrary observable A may be understood as a hidden variable in terms of A or Bell's *beable*[74] of A. If we admit putting a local expectation value (4.12) as a postulate of de Bohmian mechanics, the postulate BM2 can be replaced as that the particle's general observable is given by a local expectation value.

4.2 Bohmian Mechanics with Quasi-Probability

We shall now embed the α-parametrized quasi-probability introduced in chapter 2 into Bohmian mechanics. Let $\mathbf{Q}$ be the position operator of S and $E^{\mathbf{Q}} (\mathbf{q}) = |\mathbf{q}\rangle \langle \mathbf{q}|$ be the spectrum projection of $\mathbf{Q}$ on the eigenstate $|\mathbf{q}\rangle$ of $\mathbf{Q}$. Then, $\mathbf{Q} = \int \mathbf{q} E^{\mathbf{Q}} (\mathbf{q})\, \mathrm{d}\mathbf{q}$. Consider the α-parametrized joint quasi-probability distribution (2.48) of the position $\mathbf{Q}$ and the arbitrary observable A on ψ ;

$$q^{\alpha} (\mathbf{q}, a \,|\psi\,) = \langle \psi | \, E^{\mathbf{Q}} (\mathbf{q}) \circ_{\alpha} E^{A} (a) \, |\psi\rangle \, . \qquad (4.16)$$

The advantage of considering the joint quasi-probability as a basic ingredient in Bohmian mechanics is that it leads to the marginal quasi-probability which ensures the Born rule, that is, the formula,

$$q \,(a \,|\psi\,) = \int q^{\alpha} (\mathbf{q}, a \,|\psi\,) \, \mathrm{d}\mathbf{q} = |\langle a|\psi\rangle|^{2} \, . \qquad (4.17)$$

for any observable A on the assumption of BM3. It should be noted that this formula is not included in Bohm's original postulates BM1 $\sim$ BM3. From equation (4.17), it can be seen that the Bohmian mechanics with the quasi-probability reproduces the quantum statistics straightforwardly.

We also find the α-parametrized conditional quasi-probability distribution (2.35) of an observable A on $(\mathbf{q}, \psi)$ reads

$$q^{\alpha}\left(a\left|\psi, \mathbf{q}\right.\right) = \alpha \frac{\langle \mathbf{q}| E^{A}(a) |\psi\rangle}{\langle \mathbf{q}|\psi\rangle} + (1-\alpha) \frac{\langle \psi| E^{A}(a) |\mathbf{q}\rangle}{\langle \psi|\mathbf{q}\rangle}. \tag{4.18}$$

In particular, the α-parametrized conditional quasi-probability distribution of position $\mathbf{Q}$ becomes

$$\begin{aligned} q^{\alpha}\left(\mathbf{q}'\left|\psi, \mathbf{q}\right.\right) &= \alpha \frac{\langle \mathbf{q}| |\mathbf{q}'\rangle \langle \mathbf{q}'| |\psi\rangle}{\langle \mathbf{q}|\psi\rangle} + (1-\alpha) \frac{\langle \psi| |\mathbf{q}'\rangle \langle \mathbf{q}'| |\mathbf{q}\rangle}{\langle \psi|\mathbf{q}\rangle} \\ &= \left\{ \alpha \frac{\langle \mathbf{q}'|\psi\rangle}{\langle \mathbf{q}|\psi\rangle} + (1-\alpha) \frac{\langle \psi|\mathbf{q}'\rangle}{\langle \psi|\mathbf{q}\rangle} \right\} \delta\left(\mathbf{q}-\mathbf{q}'\right) \end{aligned} \tag{4.19}$$

in which deterministic nature of Bohmian mechanics becomes manifest.

With the α-parametrized conditional quasi-probability, the conditional quasi-average of A is evaluated as

$$\begin{aligned} \langle A\rangle_{\psi}^{\alpha}(\mathbf{q}) : &= \int a q^{\alpha}\left(a\left|\mathbf{q}, \psi\right.\right) \mathrm{d}a \\ &= \alpha \frac{\langle \mathbf{q}| A |\psi\rangle}{\langle \mathbf{q}|\psi\rangle} + (1-\alpha) \frac{\langle \psi| A |\mathbf{q}\rangle}{\langle \psi|\mathbf{q}\rangle}. \end{aligned} \tag{4.20}$$

We thus observe that Holland's local expectation value $\langle A\rangle_{\psi}(\mathbf{q})$ in equation (4.12) arises as our expectation value at $\alpha = 1/2$;

$$\langle A\rangle_{\psi}(\mathbf{q}) = \langle A\rangle_{\psi}^{\alpha=1/2}(\mathbf{q}). \tag{4.21}$$

Note that the expectation value (4.20) is a linear combination of two weak values of a conjugate pair.

Now, since the definition of Bohmian velocity in postulate BM2 can be replaced with

(4.12), and since (4.12) allows for the extension (4.20), we may just replace BM2 with the new postulate:

BM2' Given a position $\mathbf{q}$ of the particle in the state ψ, the α-parametrized conditional quasi-probability distribution of an observable A , $q^{\alpha}\left(a\,|\psi,\mathbf{q}\,\right)$, is given by

$$q^{\alpha}\left(a\,|\psi,\mathbf{q}\,\right)=\alpha\frac{\langle\mathbf{q}|\,E^{A}\left(a\right)|\psi\rangle}{\langle\mathbf{q}|\psi\rangle}+\left(1-\alpha\right)\frac{\langle\psi|\,E^{A}\left(a\right)|\mathbf{q}\rangle}{\langle\psi|\mathbf{q}\rangle}. \tag{4.22}$$

At this point, it is important to recognize that adopting BM2' in place of BM2 does not modify the content of Bohmian mechanics, since the α-dependence appears only in the association of the quasi-probability and not in the probability which can be detected in experiment. Indeed, as in equation (4.13), the output probability has no α-dependence,

$$q^{\alpha}\left(a\,|\psi\,\right)=\int q^{\alpha}\left(a\,|\psi,\mathbf{q}\,\right)p\left(\mathbf{q}\,|\psi,\mathbf{Q}\,\right)\mathrm{d}\mathbf{q}=\langle\psi|\,E^{A}\left(a\right)|\psi\rangle\,. \tag{4.23}$$

It can be confirmed readily by use of equation (4.3) in BM3. We thus learn that our quasi-probability can be embedded in Bohmian mechanics without altering its physical content. As shown in (4.23), the α-dependence in equation (4.22), which exhibits the ambiguity in occurrence of the value of A at a position $\mathbf{q}$, disappears in the physical quantities after the average over all possible positions is performed.

4.3 Ontological Model

An ontological model introduced by Spekkens and Hariggan [7, 8] may be a novel general framework of hidden variable theories of quantum mechanics. Their main motivation of ontological model may be to examine the question "Is a quantum state ψ a physically real object (ontic) or is it an abstract entity of the observer's knowledge or information (epistemic)?". In order to consider this question, Spekkens made such a general model that it is epistemically equivalent to quantum theory and it bears some ontological entity. Then, they examine the question above by comparing ontological model with quantum theory. There remains, however, the question what model truly deserves to be called ontological model, or what is the really general ontological model. Indeed, it has been pointed out [10] that the conventional framework of the ontological model cannot accommodate Bohmian mechanics. In this section, we review the original ontological model [7, 8] briefly and extend it properly. It should be noted that there are excellent general review papers of ontological model, e.g., [51].

4.3.1 Conventional Model

Suppose that $(\mathscr{P}, \mathscr{O}, \omega)$ is a operational probabilistic theory with $\omega(a, A, P) = p(a\,|P, A)$ for $a \in \mathbb{K}_A$, $A \in \mathscr{O}$ and $P \in \mathscr{P}$ defined in chapter 3. Let Λ be a set. We shall call Λ an ontic state space and $\lambda \in \Lambda$ an ontic state. We assume that the set Λ plays a similar role as the phase space in classical mechanics. That is, the output probability distribution of observable A given P can be written as

$$p(a\,|P, A) = \int_\Lambda p(a\,|\lambda, A)\, p(\lambda\,|P)\, \mathrm{d}\lambda. \tag{4.24}$$

Here, two probabilities $p(a\,|\lambda, A)$ and $p(\lambda\,|P)$ are defined by the model. Then, the ontic state space Λ with probabilities $p(a\,|\lambda, A)$ and $p(\lambda\,|P)$ is called the ontological model of the operational theory $(\mathscr{P}, \mathscr{O}, \omega)$. The ontic state λ, therefore, may be regarded as the representation of some kind of physical reality of system or complete description of system. The probability $p(\lambda\,|P)$ represents the weight attached to the proposition that the preparation P really corresponds to ontic state λ and then $p(\lambda\,|P)$ is called the *epistemic state* since this probability can be regard as the degree of ignorance of our knowledge of system. The probability $p(a\,|\lambda\,, A)$ represents the weight attached to the proposition the the observable A of physical system takes a value a on ontic state λ. It is called *indicator function* by regarding it as a function of λ. Needless to say, we are mainly interested in the ontological model of quantum theory, that is $(\mathscr{P}, \mathscr{O}, \mu) = (\mathcal{S}(\mathcal{H}), \mathcal{L}_{\mathrm{SA}}(\mathcal{H}), \mu)$ with $\mu(a, A, P) = p(a\,|P, A) = \mathrm{Tr}\left[E^A(a)\,\rho_P\right]$. This is the conventional ontological model [7, 8]. More details and developments can be seen in their original papers [7, 8] and review paper e.g., [51] respectively.

	ontological model
ontic state space	Λ
epistemic state	$p(a\,|\lambda, A)$
indicator function	$p(\lambda\,|P)$

Table 4.1: Ontological model

It may be useful to calculate the CHSH parameter introduced in section 3.4 in this ontological model. Suppose that the target system is composite of S_1 and S_2. Let A and A' be the two-valued observables of S_1 and B and B' be the two-valued observables of S_2 as same as section 3.4. The output probability distribution is given by

$$p(a_i, b_j\,|P, A, B) = \int_\Lambda p(a_i, b_j\,|\lambda)\, p(\lambda\,|P)\,\mathrm{d}\lambda \tag{4.25}$$

Here, we require the following condition

$$p(a_i, b_j\,|\lambda) = p(a_i\,|\lambda)\, p(b_j\,|\lambda) \tag{4.26}$$

for any λ and any pair of observables. This means that the observations on S_1 and on S_2 are probabilistically independent. We shall call this condition the *Bell-locality* named after Bell's work and call the ontological model satisfying Bell-locality (Bell)-local ontological model. In such model, the expectation value of product of A and B becomes

$$\begin{aligned} C_P(A, B) &= \sum_{i,j=0,1} a_i b_j p(a_i, b_j\,|P, A, B) \\ &= \int_\Lambda \sum_{i,j=0,1} a_i b_j p(a_i\,|\lambda)\, p(b_j\,|\lambda)\, p(\lambda\,|P)\,\mathrm{d}\lambda \\ &= \int_\Lambda \langle A\rangle_\lambda \langle B\rangle_\lambda\, p(\lambda\,|P)\,\mathrm{d}\lambda \end{aligned} \tag{4.27}$$

where $\langle A\rangle_\lambda = \sum_i a_i p(a_i\,|\lambda)$ and $\langle B\rangle_\lambda = \sum_j b_j p(b_j\,|\lambda)$. Then, we obtain that the CHSH parameter in the local ontological model is given by

$$\mathtt{CHSH}_P^{\mathrm{L}} = \int_\Lambda \left\{\langle A\rangle_\lambda \langle B\rangle_\lambda + \langle A'\rangle_\lambda \langle B\rangle_\lambda + \langle A'\rangle_\lambda \langle B'\rangle_\lambda - \langle A\rangle_\lambda \langle B'\rangle_\lambda\right\} p(\lambda\,|P)\,\mathrm{d}\lambda \tag{4.28}$$

where L on _ means that this is in local ontological model. Since $\langle A\rangle_\lambda = \sum_i a_i p\left(a_i\,|\lambda\,\right) = 2p\left(+\,|\lambda\,\right) - 1$, we observe that

$$-1 \leq \langle A\rangle_\lambda \leq 1, \tag{4.29}$$

$$-1 \leq \langle B\rangle_\lambda \leq 1. \tag{4.30}$$

Therefore, the CHSH parameter satisfies

$$\left|\mathtt{CHSH}_P^{\mathrm{L}}\right| \leq 2 \tag{4.31}$$

for any P. In deterministic model, Bell locality is given by

$$p\left(a_i, b_j\,|\lambda\,\right) = \delta\left(a - v^A\left(\lambda\right)\right)\delta\left(a - v^B\left(\lambda\right)\right). \tag{4.32}$$

where v^A and v^B are called the value function of A and B respectively. By putting $\langle A\rangle_\lambda = v^A\left(\lambda\right) = \pm 1$ and $\langle B\rangle_\lambda = v^B\left(\lambda\right) = \pm 1$ in the equation (4.28), it is clear that the CHSH parameter in the local deterministic model satisfies the same equation as non deterministic one for any P;

$$\left|\mathtt{CHSH}_P^{\mathrm{LD}}\right| \leq 2 \tag{4.33}$$

where LD represents the local deterministic model.

4.3.2 Synlogical Model

In this subsection, we extend the ontological model by introducing a certain contextuality. Since our contextuality is slightly different from the Kochen-Specker's original type of contextuality [6] and the generalized contextuality discussed in [71], we employ the term 'synlogical' instead of 'contextual'.

We introduce an ontic state space Λ and an ontic state $\lambda \in \Lambda$ in the same way as conventional one. We shall first consider the conditional joint probability

$$p\left(a, \lambda\,|\Lambda, P, A\,\right) \tag{4.34}$$

of the outcome $a \in \mathbb{K}_A$ and ontic state $\lambda \in \Lambda$ given preparation P as the fundamental notion of the model. This represents the probability that the system S is in the ontic state λ and the value of the observable A of S is a given some preparation P. Here, we insert the letter Λ in the joint probability $p\,(a, \lambda\,|\Lambda, P, A\,)$ for stressing that this probability is given by the model with Λ. Let us define a map $\xi : \Lambda \times \mathbb{K}_A \times \mathscr{O} \times \mathscr{P} \to [0, 1]$ by $\xi\,(\lambda, a, A, P) = p\,(a, \lambda\,|\Lambda, A, P\,)$. From the joint probability $p\,(a, \lambda\,|\Lambda, P, A\,)$, we can associate the marginal probability in a natural way:

$$p\,(a\,|\Lambda, P, A\,) := \int_\Lambda p\,(\lambda, a\,|\Lambda, P, A\,)\,\mathrm{d}\lambda, \tag{4.35}$$

$$p\,(\lambda\,|\Lambda, P, A\,) := \int_{\mathbb{K}_A} p\,(\lambda, a\,|\Lambda, P, A\,)\,\mathrm{d}a. \tag{4.36}$$

Here, $p\,(a\,|\Lambda, P, A\,)$ represents the probability distribution of the observable A given an ontic state space Λ and preparation P. $p\,(\lambda\,|\Lambda, P, A\,)$ is the probability distribution of the ontic state λ given A and P. We shall propose the reproduction condition as

$$p\,(a\,|\Lambda, A, P\,) = \int_\Lambda p\,(\lambda, a\,|\Lambda, P, A\,)\,\mathrm{d}\lambda = p\,(a\,|P, A\,) \tag{4.37}$$

where $p\,(a\,|A, P\,)$ is the output probability of the operational theory $(\mathscr{O}, \mathscr{P}, \mu)$. Then, we shall define the model which satisfies the reproduction condition (4.37),

Definition 4.1. If a pair (Λ, ξ) satisfies the reproduction condition (4.37), (Λ, ξ) is *synlogical model* of the operational theory $(\mathscr{O}, \mathscr{P}, \mu)$.

It should be noted that the probability $p\,(\lambda\,|\Lambda, P, A\,)$ does not depend on the choice of observable A in conventional ontological models. This is one of the points of generalization on our constructing models. The probability (4.36) is the same notion as the "epistemic state" defined by Harrigan and Spekkens[8] besides our epistemic states depend on the choice of observable. The dependence on the choice of an observable A in the epistemic state (4.36) implies that there exists the interdependence between the ontic state space Λ and the observable A in our model. We shall call this property the observable-synlogicality.

It is useful to explicitly characterize the notion of observable-synlogicality. We call an synlogical model (A, B)-*nonsynlogical* if the marginal probability of λ given A and P, $p\,(\lambda\,|\Lambda, A, P\,)$, and given B and P , $p\,(\lambda\,|\Lambda, B, P\,)$ satisfies

$$p\,(\lambda\,|\Lambda, P, A\,) = p\,(\lambda\,|\Lambda, P, B\,) \tag{4.38}$$

for any preparation $P \in \mathscr{P}$ and any ontic state $\lambda \in \Lambda$. If the equation (4.38) is valid for any pair of observables $A, B \in \mathscr{O}$, the synlogical model is *observable-nonsynlogical*(O-NS). Otherwise, it is *observable-synlogical*(O-S).

If an synlogical model is O-NS, we shall write the marginal probability (4.36) of λ given $A \in \mathscr{O}$ and $P \in \mathscr{P}$ as $p\,(\lambda\,|\Lambda, P\,)$;

$$p\,(\lambda\,|\Lambda, P, A\,) = p\,(\lambda\,|\Lambda, P\,)\,. \tag{4.39}$$

Out of the joint probability $p\,(a, \lambda\,|\Lambda, P, A\,)$ and other marginal probabilities mentioned before, we can associate two types of conditional probabilities, those are, the conditional probability of output value a given λ and given $A \in \mathscr{O}$ and $P \in \mathscr{P}$ and the conditional probability of λ given a and given $A \in \mathscr{O}$ and $P \in \mathscr{P}$;

$$p\,(a\,|\lambda, \Lambda, P, A\,) := \frac{p\,(\lambda, a\,|\Lambda, P, A\,)}{p\,(\lambda\,|\Lambda, P, A\,)}, \tag{4.40}$$

$$p\,(\lambda\,|a, \Lambda, P, A\,) := \frac{p\,(\lambda, a\,|\Lambda P, A\,)}{p\,(a\,|\Lambda, P, A\,)}. \tag{4.41}$$

By construction, these functions fulfill Bayes' formula,

$$p\,(\lambda\,|a, \Lambda, P, A\,) = \frac{p\,(a\,|\lambda, \Lambda, P, A\,)\,p\,(\lambda\,|\Lambda, P, A\,)}{p\,(a\,|\Lambda, P, A\,)}. \tag{4.42}$$

Using these notions, the reproduction condition (4.37) can be rewritten as

$$p\,(a\,|P, A\,) = \int_\Lambda p\,(a\,|\lambda, \Lambda, P, A\,)\,p\,(\lambda\,|\Lambda, P, A\,)\,\mathrm{d}\lambda. \tag{4.43}$$

The conditional probability appeared in integrand of equation (4.43), that is the probability

(4.41), corresponds to the "indicator function" in conventional ontological models. We notice that the resultant conditional probabilities (4.41) depend on the preparation $P \in \mathscr{P}$ in general. This alludes us to call a model *prepration-synlogical*. It is useful to explicitly characterize this notion. If the conditional probability $p\left(a\,|\lambda, \Lambda, P, A\right)$ satisfies

$$p\left(a\,|\lambda, \Lambda, A, P\right) = p\left(a\,\middle|\lambda, \Lambda, A, P'\right) \tag{4.44}$$

for any $A \in \mathscr{O}$, any $\lambda \in \Lambda$, and any $a \in \mathbb{K}_A$, an ontological model (Λ, ξ) is (P, P')-*nonsynlogical*. If the equation (4.44) is valid for any pair of preparations $P, P' \in \mathscr{P}$, a synlogical model is *preparation-nonsynlogical* (P-NS). Otherwise, it is *preparation-synlogical* (P-C). If a synlogical model is P-NS, we write the conditional probability (4.40) of a given $A \in \mathscr{O}$ and $P \in \mathscr{P}$ as $p\left(a\,|\lambda, \Lambda, A\right)$;

$$p\left(a\,|\lambda, \Lambda, P, A\right) = p\left(a\,|\lambda, \Lambda, A\right) \tag{4.45}$$

If, in particular, the model is both O-NS and P-NS, we have

$$p\left(a\,|A, P\right) = \int_\Lambda p\left(a\,|\lambda, \Lambda, A\right) p\left(\lambda\,|\Lambda, P\right) \mathrm{d}\lambda. \tag{4.46}$$

If a model is of this type, let us call the model nonsynlogical, and otherwise we call it synlogical. The conventional framework of the ontological (or hidden variable) model [4, 8] is confined to the nonsynlogical case. Venn diagram on the synlogical model is illustrated in figure 4.3.1.

	ontological model	synlogical model
ontic state space	Λ	Λ
indicator function	$p\left(a\,|\Lambda, \lambda, A\right)$	$p\left(a\,|\lambda, \Lambda, P, A\right)$
epistemic state	$p\left(\lambda\,|\Lambda, P\right)$	$p\left(\lambda\,|\Lambda, P, A\right)$

Table 4.2: Ontological and synlogical model
Ontological model and synlogical model are compared. We put the Λ in the bracket of $p\left(\cdots\right)$ to stress that they depend on the model.

It may be useful to calculate the CHSH parameter in the synlogical model. Let S be a composite system of S_1 and S_2. Suppose that A and A' be the two-valued observables of S_1 and B and B' be the two-valued observables of S_2 as section 3.4. Recalling equation (4.37),

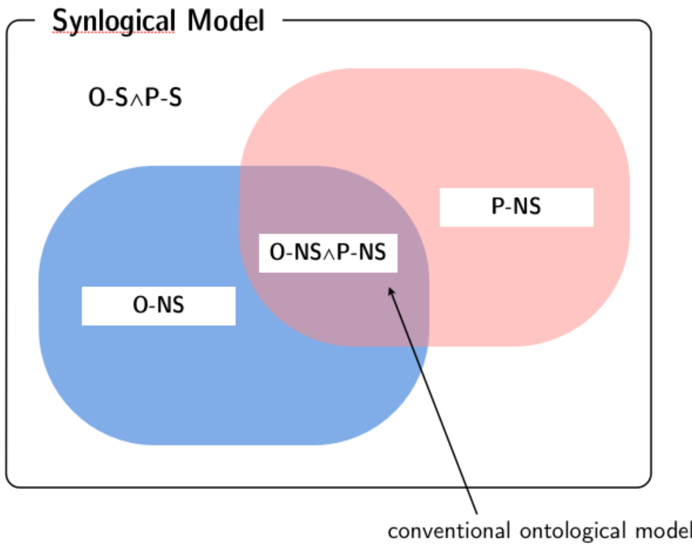

Figure 4.3.1: Venn diagram of synlogical model

the output probability distribution is given by

$$p\left(a_i, b_j \,|P, A, B\,\right) = \int_\Lambda p\left(\lambda, a_i, b_j \,|\Lambda, P, A, B\,\right) \mathrm{d}\lambda \tag{4.47}$$

$$= \int_\Lambda p\left(a_i, b_j \,|\Lambda, P, A, B\,\right) p\left(\lambda \,|\Lambda, P, A, B\,\right) \mathrm{d}\lambda. \tag{4.48}$$

Let us consider first the P-S ∧ O-NS model whose indicator function and epistemic state are given by $p\left(a_i, b_j \,|\Lambda, P, A, B\,\right)$ and $p\left(\lambda \,|\Lambda, P\,\right)$ respectively. If we require the locality (probabilistic independency) condition (4.26), the indicator function becomes

$$p\left(a_i, b_j \,|\Lambda, P, A, B\,\right) = p\left(a_i \,|\Lambda, P, A\,\right) p\left(b_j \,|\Lambda, P, B\,\right). \tag{4.49}$$

We can easily observe by the same calculation of previous section that

$$\left|\mathtt{CHSH}_P^{\mathrm{PS,\,L}}\right| \leq 2 \tag{4.50}$$

for any P.

Next, Let us consider the P-NS ∧ P-S model whose indicator function and epistemic state are given by $p\left(a_i, b_j \,|\Lambda, A, B\,\right)$ and $p\left(\lambda \,|\Lambda, P, A, B\,\right)$ respectively. From this, it is immediately observed that the no-signaling condition (3.57) introduced in chapter 3 is generally not valid in this model. Although it may be possible to be valid the no-signaling condition in this model by aligning the observable-dependence, the CHSH parameter is same as equation (4.50) if we assume the locality (probabilistic independency) condition (4.26);

$$\left|\mathtt{CHSH}_P^{\mathrm{OS,\,L}}\right| \leq 2. \tag{4.51}$$

4.4 Quasi-Probabilistic Ontological Model

We shall carry out the further extension of the ontological model by introducing a complex quasi-probability[2]. Let Λ be an ontic state space. We shall alter the joint probability (4.34) of λ and the value a of an observable A given P into quasi-probability;

$$q\,(\lambda, a\,|\Lambda, P, A\,) \in \mathbb{C} \tag{4.52}$$

We use the letter q to represent quasi-probability instead of p. We call a pair Λ and a map $\zeta : \lambda \mapsto q\,(\lambda, a\,|\Lambda, P, A\,)$ *quasi-probabilistic synlogical ontological model* of an operational theory if the joint quasi-probability $q\,(\lambda, a\,|\Lambda, A, P\,)$ satisfies

$$p\,(a\,|P, A\,) = \int_\Lambda q\,(\lambda, a\,|\Lambda, P, A\,)\,\mathrm{d}\lambda. \tag{4.53}$$

We can associate the indicator function and epistemic state from the joint quasi-probability distribution $q\,(\lambda, a\,|\Lambda, A, P\,)$. The *quasi-epistemic state* is given by

$$q\,(\lambda\,|P, A\,) := \int_{\mathbb{K}_A} q\,(\lambda, a\,|\Lambda, P, A\,)\,\mathrm{d}a, \tag{4.54}$$

and the *quasi-indicator function* is

$$q\,(a\,|\lambda, P\,) := \frac{q\,(\lambda, a\,|\Lambda, P, A\,)}{q\,(\lambda\,|\Lambda, P, A\,)}. \tag{4.55}$$

Synlogicalities in a quasi-probabilistic model can be defined in the same way as the onto-

[2] For the real-valued quasi-probabilistic ontological model, see Refs. [7, 75].

logical model with standard probability. Using these notions, the the reproduction condition (4.53) can be rewritten as

$$p\left(a\left|P,A\right.\right)=\int_{\Lambda}q\left(a\left|\lambda,\Lambda,P\right.\right)q\left(\lambda\left|\Lambda,P,A\right.\right)\mathrm{d}\lambda. \tag{4.56}$$

By construction, the quasi-probabilistic ontological model can be classified into three types. First type is the case that the epistemic state is quasi-probability but indicator function is standard probability, in which case the quasi-epistemic state can be regarded as the quasi-probability distribution on the phase space, and then quasi-epistemic state of this type may corresponds to the quasi-probability introduced in section 2.1., that is the Wigner function [18], Husimi function [19], etc. The second type is that epistemic state is standard probability but the indicator function is quasi-probability, which corresponds to Bohmian mechanics as we shall show below, thirdly, the case that both epistemic state and indicator function are quasi-probabilistic can be considered.

We shall now show that the Bohmian mechanics is the quasi-probabilistic P-S ontological model. To this end, we recall first that the ontic state space of Bohmian mechanics is just the position eigenspace,

$$\Lambda_{\mathbf{X}}=\left\{\left|\mathbf{x}\right\rangle\left|\mathbf{x}\in\mathbb{R}^{N}\right.\right\} \tag{4.57}$$

where $N=3n$ if n particles are present in the three dimensional space. Note that the postulate **BM2'** states that the indicator function on Bohmian mechanics is an (α-parameterized) conditional quasi-probability $q^{\alpha}\left(a\left|\psi,\mathbf{x}\right.\right)$ for preparation P_{ψ} which is associated with a quantum state ψ. Since the quasi-probability $q^{\alpha}\left(a\left|\psi,\mathbf{x}\right.\right)$ depends on ψ, this theory is obviously P-S, for which the reproduction condition is guaranteed by the equation (4.17). This shows that Bohmian mechanics is actually a quasi-probabilistic P-S ontological model defined by the ontic state space (4.57) and the foundational joint quasi-probability (4.52) of this model is given by

$$q^{\alpha}\left(\mathbf{x},a\left|P\right.\right)=\left\langle\psi\right|E^{X}\left(\mathbf{x}\right)\circ_{\alpha}E^{A}\left(a\right)\left|\psi\right\rangle. \tag{4.58}$$

In Bohmian mechanics, the quasi-epistemic state is given by

$$p(\mathbf{x}\,|\psi) = \int_{\mathbb{K}_A} q^{\alpha}(\mathbf{x}, a\,|\psi)\,\mathrm{d}a = |\langle\mathbf{x}|\psi\rangle|^2, \tag{4.59}$$

which ensures condition BM3. We also find that in Bohmian mechanics the quasi-indicator function (4.55) reads

$$q^{\alpha}(a\,|\psi, \mathbf{x}) = q^{\alpha}(\mathbf{x}, a\,|\psi)\,/p(\mathbf{x}\,|\psi)\,. \tag{4.60}$$

It should be noted that if the preparation P is associated with a mixed state ρ our result can be obtained by replacing $\langle\psi|\ldots|\psi\rangle$ with $\mathrm{Tr}\,[\ldots\rho]$. In addition, this model is ψ-epistemic because there are vectors $\psi, \phi \in \mathcal{H}$ such that $p(\mathbf{x}\,|\psi)\,p(\mathbf{x}\,|\phi) = |\langle\mathbf{x}|\psi\rangle|^2\,|\langle\mathbf{x}|\phi\rangle|^2 \neq 0$.

Chapter 5

Conclusions and Discussions

In this thesis, we have presented theoretical studies on the quasi-probability in quantum mechanics. These studies were done in three directions; structural, epistemological and conceptual directions. We now recollect our considerations and results, and discuss the future research which they suggest.

In chapter 2, we examined the structural significance of quasi-probability underlying the weak value in quantum mechanics. This was attained by showing the legitimacy and usefulness of quasi-probability in quantum mechanics. We found an internal consistency between the quasi-probability underlying the weak value and quantum mechanics. This result suggests that it is legitimate to accept the quasi-probability as a fundamental element of quantum mechanics and that quasi-probability is useful in quantum mechanics. The quasi-probability underlying the weak value is an extension of probability in quantum mechanics based on the weak value. We recalled that this extension is naturally brought by generalized Gleason's theorem in quantum mechanics with some consistency condition and that the quasi-probability is measurable by the weak measurement procedure. This extension possesses an intrinsic ambiguity expressed by a complex valued parameter α. This parameter prompts us to introduce the generalized product of quantum mechanical observables, α-product. The joint quasi-probability distribution which is expressed in terms of the α-product of observables may be regarded as the simultaneous quasi-probability distribution of incompatible observables since its marginal probability gives the Born rule and the joint quasi-probability distribution of compatible observables coincides with the joint probability distribution for compatible observables. It is notable that the α-product includes the well-known Jordan product for the special case (α=1/2), suggesting that the significance of Jordan products may be given through the extension of probability.

In chapter 3, we reinforced the epistemological significance of quasi-probability from the viewpoint of operational theories. This was done for two reasons; first, for investigating the relationship between the ontological model of quantum mechanics and quasi-probability, we need operational probabilistic theories as a theoretical basis of the ontological model. This allows us to examine the physical reality from the epistemological point of view. Secondly, we need to analyze the epistemological significance of a post-selected measurement to find out a proper quasi-probability in the weak measurement procedure. We investigated conditional probability in the view of the direction of inference and show that the post-selected probability does not depend on the direction of inference.

In chapter 4, we discussed the conceptual significance of the quasi-probability in quantum mechanics. We showed that the quasi-probability underlying the weak value gave a clue to interpret the quantum mechanics realistically. In particular, we showed that the quasi-probability sheds new light on the most familiar type of realistic interpretations of quantum mechanics, that is, the Bohmian mechanics. In addition, this brought a long-sought reconciliation between Bohmian mechanics and a properly extended framework of ontological model. For the extension, we introduced a complex quasi-probability and a certain contextually which we referred to as synlogicality, and thereby demonstrated that the Bohmian mechanics is a quasi-probabilistic synlogical ontological model. This way, we confirmed that quantum mechanics with quasi-probability naturally leads to a realistic interpretation of quantum mechanics.

In particular, two of our most important results are;

- **The legitimacy and the usefulness of quasi-probability underlying the weak value**: The quasi-probability introduced in this thesis is defined for two non-commuting observables A and B for which no joint probability is admitted in quantum mechanics due to the incompatibility of simultaneous measurements of the two observables. Moreover, the marginal of our joint quasi-probability naturally produces the Born rule. Compared to the quasi-probability mentioned in section 2.1, our quasi-probability possesses the α-dependence that disappears in physically testable situations.

- **Bohmian mechanics as a quasi-probabilistic ontological model:** The quasi-probability underlying the weak value can be embedded in Bohmian mechanics such that one of the premises of Bohmian mechanics is replaced by an alternative one which directly leads to the Born rule. This observation allows us to regard Bohmian mechanics as a quasi-probabilistic ontological model in a synlogical (contextual) type, clarifying its so far obscure status in the category of hidden variable models.

Our result suggests that the quasi-probability does form a basic ingredient in quantum mechanics, which means that, for instance, the arbitrariness of quasi-probability underlying the weak value expressed by α-parameter may provide a useful tool for the calculation of physical quantities such as the expectation value and the correlation of two physical observables by choosing the proper number α, or weak value based quasi-probabilistic ontological models may become helpful in tackling the conceptual problem of quantum mechanics.

There remain several important questions. For one, we have not determined whether the the complex parameter α which expresses the intrinsic ambiguity of the quasi-probability underlying the weak value has some kind of physical meaning. Although we have seen that this parameter α does not affect the verifiable quantities such as expectation values or correlations of observables in actual current experiments, it may be reasonable to expect that the parameter α may suggest a more fundamental physical theory than quantum theory. In addition, we have not yet obtained the exact answer to the question how the quasi-probability in quantum mechanics has to be interpreted directly. Our results clarified the legitimacy or the usefulness of quasi-probability in quantum mechanics, but we still do not know how to interpret the situation where, for instance, the energy of electron takes a certain value E_0 on the state ψ with probability -256% or $\frac{1}{2} + i\%$. This may be caused by the fact that the quasi-probability transcends far beyond the standard understanding of probability.

Despite this, we are quite sure that the quasi-probability is a convenient and useful concept in quantum mechanics. Our results suggest that the quasi-probability underlying with the weak value may become a powerful tool in tackling unsolved problems in physics such as quantum gravity. We believe that progress of the study of quasi-probability may help a deeper understanding of the conceptual, structural and practical aspects of quantum mechanics.

Appendix A

Another Derivation of Joint Quasi-Probability

In this appendix, we shall show that the α-parametrized joint quasi-probability defined in chapter 2 can be derived from another way. Although it needs some stringent assumptions, there are merits since the limitation of dimension of Hilbert space is excluded and the proof is clear.

Let us denote the joint quasi-probability measure as

$$\mu_\rho^{AB} : \left(\Delta, \Delta'\right) \mapsto q\left(a \in \Delta, b \in \Delta' \left|\rho\right.\right) \in \mathbb{C} \tag{A.1}$$

satisfying following properties

$$\mu_\rho^{AB}\left(\Delta, \Delta'\right) = \sum_i \mu_\rho^{AB}\left(\Delta_i, \Delta'\right) \tag{A.2}$$

$$\mu_\rho^{AB}\left(\Delta, \Delta'\right) = \sum_j \mu_\rho^{AB}\left(\Delta, \Delta'_j\right) \tag{A.3}$$

$$\mu_\rho^{AB}\left(\mathbb{R}, \mathbb{R}\right) = 1 \tag{A.4}$$

for any mutually disjoint sequences $\Delta_1, \Delta_2, \ldots$ with $\Delta = \cup_i \Delta_i$ and $\Delta'_1, \Delta'_2, \ldots$ with $\Delta' = \cup_i \Delta'_i$. Suppose that a map $\rho \in \mathcal{S}\left(\mathcal{H}\right) \mapsto \mu_\rho^{AB}\left(\Delta, \Delta'\right) \in \mathbb{C}$ is affine for all $\Delta,\ \Delta' \in \mathcal{B}^1$. This requirement is natural in an analogy of probability measure in quantum mechanics. The affineness leads us to following theorem;

Theorem A.1. *There exists the trace class operator* $T^{AB}(\Delta, \Delta')$ *such that*

$$\mu_\rho^{AB}\left(\Delta, \Delta'\right) = \mathrm{Tr}\left[T^{AB}\left(\Delta, \Delta'\right)\rho\right]. \tag{A.5}$$

The operator valued measure T^{AB}:$(\Delta, \Delta') \mapsto T^{AB}(\Delta, \Delta')$ is none other than complex valued version of POVM. Let us define some notions with $T^{AB}(\Delta, \Delta')$. The quasi-expectation value of A times B respect to μ_ρ^{AB} may be defined by

$$\begin{aligned}\mathbb{QE}\left(A, B\,|\rho\right): &= \iint ab\mu_\rho^{AB}(a,b)\,\mathrm{d}a\mathrm{d}b \\ &= \mathrm{Tr}\left[\tilde{T}^{AB}\rho\right]\end{aligned} \tag{A.6}$$

where

$$\tilde{T}(A,B) = \iint abT^{AB}(a,b)\,\mathrm{d}a\mathrm{d}b. \tag{A.7}$$

Let us assume that

$$\mu_\rho^{AA}(a,a) = \mathrm{Tr}\left[E^A(a)\rho\right] \tag{A.8}$$

for all ρ. Then, the operator T^{AA} is that $T^{AA}(a,a) = E^A(a)$. We have that

$$\tilde{T}(A,A) = \int a^2 E^A(a)\,\mathrm{d}a\mathrm{d}b = A^2. \tag{A.9}$$

We derive the α-parameterized joint quasi-probability as the special version of $\tilde{T}(A,B)$, i.e., $\tilde{T}\left(E^A(\Delta), E^B(\Delta')\right)$. This is equivalent to require the following postulates;

i) There exists a bilinear map $F: \mathcal{L}(\mathcal{H}) \times \mathcal{L}(\mathcal{H}) \to tc(\mathcal{H})$ for $T^{AB}(\Delta, \Delta')$ such that

$$T^{AB}\left(\Delta, \Delta'\right) = F\left(E^A(\Delta), E^B\left(\Delta'\right)\right) \tag{A.10}$$

where $E^A(a)$ and $E^B(b)$ is the spectral projection of A and B corresponding to eigenvalue a

and b.

ii) The map F satisfies following conditions

$$F(X, X) = X^2, \tag{A.11}$$

$$F(X, Y) - F(Y, X) = \theta [X, Y]. \tag{A.12}$$

where θ is an arbitrary complex parameter.

From the condition (A.11) and bilinearity, we observe

$$F(X + Y, X + Y) = (X + Y)^2 = X^2 + XY + YX + Y^2. \tag{A.13}$$

The left hand side of (A.13) becomes

$$\begin{aligned} F(X + Y, X + Y) &= F(X, X) + F(X, Y) + F(Y, X) + F(Y, Y) \\ &= X^2 + F(X, Y) + F(Y, X) + Y^2 \end{aligned} \tag{A.14}$$

Comparing (A.13) and (A.14),

$$F(X, Y) + F(Y, X) = XY + YX, \tag{A.15}$$

Using the condition (A.12), we observe that

$$F(X, Y) = \frac{1}{2}(XY + YX - \theta(XY - YX)) \tag{A.16}$$

$$= \frac{1}{2}(1 - \theta) XY + \frac{1}{2}(1 + \theta) YX. \tag{A.17}$$

If we put $\theta = 2\alpha - 1$, we have $F(X, Y) = \alpha XY + (1 - \alpha) YX$. Hence, we observe

$$T^{AB}(\Delta, \Delta') = \alpha E^A(\Delta) E^B(\Delta') + (1 - \alpha) E^B(\Delta') E^A(\Delta). \tag{A.18}$$

Bibliography

[1] Einstein, A., Podolsky, B. and Rosen, N., *Phys. Rev.*, **47**, 777 (1935).

[2] Bohm, D., *Phys. Rev.* **85**, 166 (1952).

[3] Bohm, D., *Phys. Rev.* **85**, 180 (1952).

[4] Bell, J.S., *Phys.*, **1(3)**, 195, (1964).

[5] Von Neumann, J. "*Mathematische Gründlagen der Quantenmechanik.*" Grundlehren der mathematischen Wissenschaften, volume Bd 38 (1932).

[6] Kochen, S. and Specker, E. P., *J. Math. Mech.* **17**, 59 (1967).

[7] Spekkens, R.W., *Phys. Rev. A* **71**, 052108 (2005).

[8] Harrigan, N. and Spekkens, R.W., *Found. Phys.* **40**, 125 (2010).

[9] Pusey, Matthew F., Barrett, J and Rudolph, T, *Nature Physics* **8**, 475 (2012).

[10] Feintzeig, B., *Studies in History and Philosophy of Modern Physics*, **48**, 59, (2014).

[11] Aharonov, Y., Albert, D. Z., and Vaidman, L., *Phys. Rev. Lett.*, **60**, 1351 (1988).

[12] Hardy, L., *Phys. Rev. Lett.*, **68**, 2981 (1992).

[13] Aharonov, Y., Botero, A., Popescu, S., Reznik, B., & Tollaksen, J., *Phys. Rev.* A, **301**, 130-138, (2002).

[14] Hosten, O. and Kwiat, P., Science **319**, 787 (2008).

[15] Dixon, P.B., Starling, D.J., Jordan, A, Howell, J.C., *Phys. Rev. Lett.,* **102**, 173601 (2009)

[16] Egan, P and Stone, J.A., *Opt. Lett.,* **37**, 4991 (2012).

[17] Kolmogorov, A.N, *"Grundbegriffe der Wahrscheinlichkeitrechnung"* Springer, Berlin, (1933), ["*Foundations of the Theory of Probability*", Second English Edition, Chelsea, New York, (1950)].

[18] Wigner, E.P., *Phys. Rev.*, **40**, 749 (1932).

[19] Husimi, K. *Proc. Phys. Math. Soc. Jpn,* **22**, 264 (1940).

[20] Kirkwood, J. G., *Phys. Rev.*, **44**, 31(1933).

[21] P. A. M. Dirac, *Rev. Mod. Phys.* **17**, 195 (1945).

[22] Steinberg, A. M., *Phys. Rev. A*, **52**, 32–42 (1995).

[23] Steinberg, A. M., *Phys. Rev. Lett.*, **74**, 2405–2409 (1995).

[24] Ozawa, M., *AIP Conf. Proc.* **1363**, 53 (2011).

[25] Hofmann, H.F., *New J. Phys.*, **14**, 043031 (2012).

[26] Morita, T., Sasaki, T. and Tsutsui, I., *Progress of Theoretical and Experimental Physics* **2013.5** 053A02, (2013).

[27] Lundeen, J. S. and Bamber, C., *Phys. Rev. Lett.* **112**, 070405 (2014).

[28] Feynman, R.P., *Int. J. Theor. Phys.*, **21**, 467, (1982).

[29] Muckenheim, W., Ludwig, G., Dewdney, C., Holland, P.R., Kyprianidis, A., Vigier, J.P.: Cufaro Petroni, N., Bartlett, M.S., Jaynes, E.T., *Phys. Rep.* **133**, 337 (1986).

[30] Denkmayr, T, Geppert, H, Sponar, S, Lemmel, H, Matzkin, A., Tollaksen, J. and Hasegawa, Y., *Nature Commn.,* DOI: 10.1038/ncomms5492 (2014).

[31] Shikano, Y., *"Measurements in Quantum Mechanics"*, edited by Pahlavani, M. R., In-Tech, Chapter 4 page 75., (2012).

[32] Svensson, B. E., *Quanta*, **2**, 18-49, (2013).

[33] Aharonov, Y., Popescu, S., Rohrlich, D., & Skrzypczyk, *New J. Phys.*, **15(11)**, 113015 (2013).

[34] T. Drisch, Int. J. *Theor. Phys.* **18** No.4 239. (1979).

[35] Danan, A., Farfurnik, D., Bar-Ad, S., & Vaidman, L., *Phys. Rev. Lett.*, **111(24)**, 240402 (2013).

[36] Jordan, P., *Zeits. f. Physik* **38**, 513 (1926).

[37] Ludwig, G., "*An Axiomatic Basis for Quantum Mechanics*", Springer, New York (1985)

[38] Ludwig, G., "*Foundation of Quantum Mechanics*", Springer, New York (1985).

[39] Ozawa, M, *J. Math. Phys.*, **25**, 79 (1984).

[40] Ozawa, M., *Ann. of Phys.,* **311.2** 350 (2004).

[41] Brigeman, P. W., *"The logic of modern physics.*" Macmillan, New York (1927).

[42] Brikoff, G., von Neumann, J. *The Annals of Mathematics,* 2nd Ser., **37**, No. 4. (1936).

[43] Mackey, G. W., "*Mathematical foundations of quantum mechanics*", New York–Amsterdam, W. A. Benjamin, Inc., (1963).

[44] Watanabe, S., *"Knowing and Guessing a Quantitative Study of Inference and Information.",* JOHN WILEY & SONS, Inc., New York (1969).

[45] Gudder, S. P. (2014). "*Stochastic methods in quantum mechanics*", Dover, New York, (1979).

[46] Busch, P., Grabowski, M., & Lahti, P. J., "*Operational quantum physics* (Vol. 31)", Springer Science & Business Media.(1995).

[47] Janotta, P and Hinrichsen, H., arXiv:1402.6562 (2014).

[48] Kraus, K., "*States, Effects and Operations: Fundamental Notions of Quantum Theory*", Springer Verlag (1983).

[49] Hardy, L., arXiv:1104.2066. (2011).

[50] Giulio, C. and Spekkens., R. W., "*Quantum Theory: Informational Foundations and Foils.*" Springer (2015).

[51] Leifer, M. S., "Is the quantum state real? An extended review of ψ-ontology theorems." *Quanta* **3.1**, 67, (2014).

[52] Watanabe, S., *Reviews of Modern Physics*, **27.2**, 179 (1955).

[53] Watanabe, S., *Progress of Theoretical Physics Supplement* **65**, 135 (1965).

[54] Dirac, P.A.M. *"The principles of Quantum Mechanics"*, Oxford University Press, (1930).

[55] Clauser, J., Horne, M., Shimony, A. and Holt, R., *Phys. Rev. Lett.,* **23** (1969).

[56] Hardy, L. and Spekkens, R., *Physics in Canada* **66** (2), 73-76 (2010)

[57] Fuchs, C. A., arXiv:1003.5209 (2010).

[58] d'Espagnat, B., "*Conceptual foundations of quantum mechanics*" Benjamin, Reading, MA, (1976).

[59] Jammer, M., "*The conceptual development of quantum mechanics.*" McGraw-Hill, (1966).

[60] Philippidis, C., Dewdney, C. and Hiley, B. J., *Nuovo Cimento,* **71B**, 75 (1979)

[61] Rockafellar, R., T., "*Convex Analysis*" Princeton University Press (1970).

[62] Nauenberg, M., *Quanta*, **3(1**), 43-46, (2014).

[63] Kendrick, BK., *J. Chem. Phys.*, **119**, 5805 (2003).

[64] Tumulka, R., *European Journal of Physics*, **26.3** L11 (2005).

[65] Nikolić, H., *American Journal of Physics*, **76.2** 143 (2008).

[66] Sanz, A. S., and Miret-Artés, S., *Journal of Physics A: Mathematical and Theoretical*, **44**.48 (2011).

[67] Sanz, A. S., and Miret-Artés, S., *American Journal of Physics,* **80.6**, 525 (2012).

[68] Norsen, T., *American Journal of Physics,* **81.4** 258 (2013).

[69] Holland, P. R., "*The quantum theory of motion: an account of the de Broglie-Bohm causal interpretation of quantum mechanics*" Cambridge university press, (1995).

[70] Bell, J. S., *Rev. Mod. Phys.*, **38(3)**, 447, (1966).

[71] Spekkens, R.W., *Phys. Rev. Lett.* **101.2** 020401 (2008).

[72] P. Jordan, J. v. Neumann, and E. Wigner, *Ann. Math.* **35**, 29 (1934).

[73] Popescu, S. and Rohrlich, D., *Found. Phys.* **24**, 379–385 (1994).

[74] Bell, J. S., *Epistemological Letters* (March 1976) (1975).

[75] Ferrie, C., *Rep. Prog. Phys.* **74**, 116001 (2011).

[76] Gudder, S. P., *J. of Math. Phys*, **11**, 431 (1970).