Research Report

KSTS/RR-16/001
January 8, 2016

Linguistic interpretation of quantum mechanics: Quantum Language [Ver. 2]

by

Shiro Ishikawa

Shiro Ishikawa
Department of Mathematics
Keio University

Department of Mathematics
Faculty of Science and Technology
Keio University

3-14-1 Hiyoshi, Kohoku-ku, Yokohama, 223-8522 Japan

KSTS/RR-16/001
January 8, 2016

Linguistic interpretation of quantum mechanics: Quantum Language [Ver. 2]

Shiro ISHIKAWA (ishikawa@math.keio.ac.jp)

Department of mathematics, Faculty of science and Technology, Keio University, 3-14-1, Hiyoshi, Kouhokuku, in Yokohama, 223-8522, Japan

Preface

This is the lecture note for graduate students. This lecture has been continued, with gradually improvement, for about 15 years in the faculty of science and technology of Keio university[1]. In this lecture, I explain "quantum language"(="measurement theory"), which was proposed as

the fundamental theory of quantum information science

by myself. Quantum language is a language that is inspired by the Copenhagen interpretation of quantum mechanics, but it has a great power to describe classical systems as well as quantum systems. In this lecture, I assert that quantum language, roughly speaking, has the three aspects as follows.

The three aspects of quantum language

①: the standard interpretation of quantum mechanics
(i.e., the true colors of the Copenhagen interpretation)
②: the final goal of the dualistic idealism (Descartes=Kant philosophy)
③: theoretical statistics of the future

And therefore, I assert that

The main assertion of this lecture

Quantum language is the most fundamental language in science.

---

[1] This preprint is the second version of

Ref. [47]: S. Ishikawa, "Linguistic interpretation of quantum mechanics: Quantum Language", Dept. Math. Keio University, 2015, KSTS/RR-15/001, 416 pages, [Ver. 1]
(`http://www.math.keio.ac.jp/en/academic/research.html`)
or, (`http://www.math.keio.ac.jp/academic/research_pdf/report/2015/15001.pdf`)

Roughly speaking, we say that [Ver. 2]="[Ver. 1]+ Sec.11.2(Wave function collapse)", though there are many small improved points.
Also, for my recent results, see my homepage (`http://www.math.keio.ac.jp/~ishikawa/indexe.html`)

The purpose of this lecture is to explain these assertions. Also, this lecture note may be regarded as the revised edition of the following two:

- [29]: S. Ishikawa, *Mathematical Foundations of Measurement Theory,* Keio University Press Inc. 2006, (335 pages) .
- [38]: S. Ishikawa, *Measurement Theory in the Philosophy of Science,* arXiv:1209.3483 [physics.hist-ph] 2012, (177 pages)

KSTS/RR-16/001
January 8, 2016

Contents

Chapter 1

My answer to Feynman's question

Dr. R. P. Feynman (one of the founders of quantum electrodynamics) said the following wise words:($\sharp_1$) and ($\sharp_2$):[1]

($\sharp_1$) There was a time when the newspapers said that only twelve men understood the theory of relativity. I do not believe there ever was such a time. There might have been a time when only one man did, because he was the only guy who caught on, before he wrote his paper. But after people read the paper a lot of people understood the theory of relativity in some way or other, certainly more than twelve. On the other hand, I think I can safely say that nobody understands quantum mechanics.

and

($\sharp_2$) We have always had a great deal of difficulty understanding the world view that quantum mechanics represents. $\cdots\cdots$ I cannot define the real problem, therefore I suspect there's no real problem, but I'm not sure there's no real problem.

In this lecture, I will answer Feynman's question ($\sharp_1$) and ($\sharp_2$) as follows.

($\flat$) I am sure there's no real problem. Therefore, since there is no problem that should be understood, it is a matter of course that nobody understands quantum mechanics.

This answer may not be uniquely determined, however, I am convinced that the above ($\flat$) is one of the best answers to Feynman's question ($\sharp_1$) and ($\sharp_2$).

The purpose of this lecture is to explain the answer ($\flat$). That is, I show that

If we start from the answer ($\flat$),
we can double the scope of quantum mechanics.

And further, I assert that

Metaphysics (which might not be liked by Feynman)
is located in the center of science.

In this lecture, I will show the above.

[1] The importance of the two ($\sharp_1$) and ($\sharp_2$) was emphasized in Mermin's book [61]

1.1 Quantum language (= measurement theory)

1.1.1 Introduction

In this lecture, I will explain "quantum language (= measurement theory (=MT))", which is located as illustrated in the following figure:

Figure 1.1. [The location of quantum language in the history of world-description (*cf.* ref.[31])]

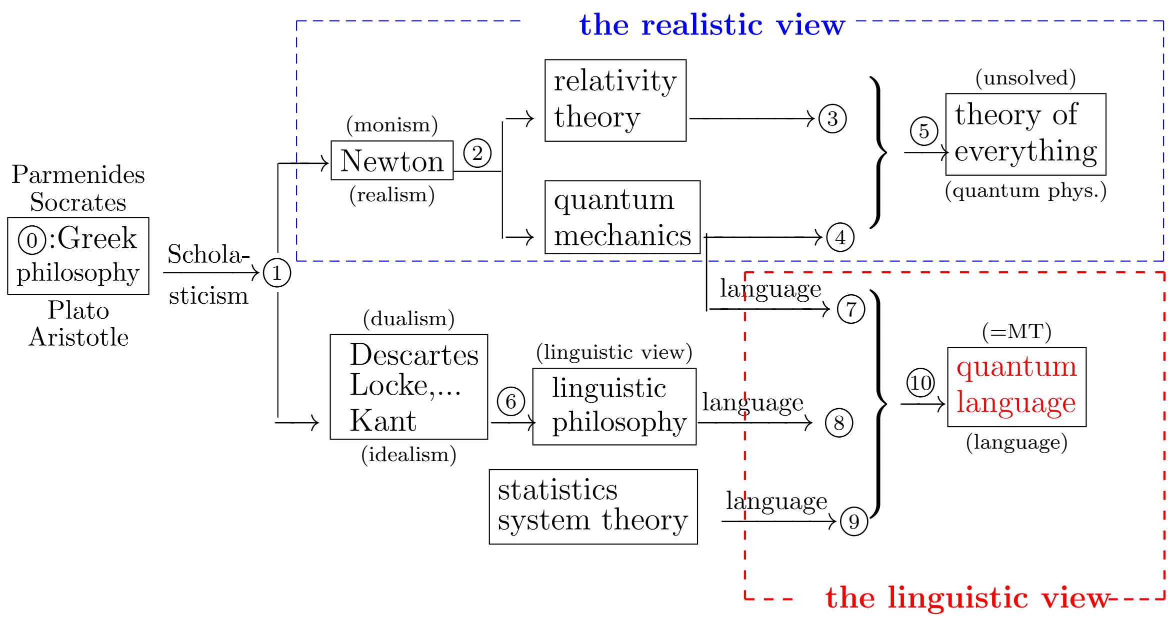


Figure 1.1: The history of the world-view

It should be noted that the above figure automatically gives answers to the following questions

- **⑦: What should be the standard interpretation of quantum mechanics?**
- **⑧: What did Descartes-Kant philosophy want to do?** [2]
- **⑨: How will theoretical statistics evolve?**

Therefore,

Figure 1.1 is all in this lecture.

[2] In this note, philosophy is always used as a metaphor.

♠**Note 1.1.** If most physicists feel something like metaphysics in quantum mechanics, the reason is due to Figure 1.1. That is, we consider that there are two "quantum mechanics", that is, "(realistic) quantum mechanics" in ⑤ and "(metaphysical) quantum mechanics" in ⑩. Namely,

- quantum mechanics $\begin{cases} \text{"(realistic) quantum mechanics" in ⑤} \\ \\ \text{"(metaphysical) quantum mechanics" in ⑩} \end{cases}$

The former is not completed yet. The latter is "the usual quantum mechanics" studied in undergraduate course of university. In this lecture, we are not concerned with the former.

♠**Note 1.2.** If readers are familiar with quantum mechanics, it may be recommended to read the following short papers before reading this lecture text.

(a) Ref. [30]: S. Ishikawa, *A New Interpretation of Quantum Mechanics:* Journal of quantum information science: Vol.1(2), pp.35-42, 2011

(b) Ref. [31]:S. Ishikawa, *Quantum Mechanics and the Philosophy of Language: Reconsideration of traditional philosophies,* Journal of quantum information science, Vol. 2(1), pp.2-9, 2012

(c) Ref. [49] S. Ishikawa, *Linguistic interpretation of quantum mechanics; Projection Postulate,* Journal of quantum information science, Vol. 5, No.4 , 150-155, 2015, DOI: 10.4236/jqis.2015.54017
(`http://www.scirp.org/Journal/PaperInformation.aspx?PaperID=62464`)

The similarities and differences between the linguistic interpretation and so called Copenhagen interpretation have been clarified in the above (c).

1.1.2 From Heisenberg's uncertainty principle to the linguistic interpretation

As explained in §4.2,

(A) In 1991(*cf.* ref. [22])[3], I found the mathematical formulation of Heisenberg's uncertainty principle (i.e., $\Delta_x \cdot \Delta_p \ge \hbar/2$ in (4.36)), which clarified that

- under what kind of condition does Heisenberg's uncertainty principle hold?

I thought that this result is interesting. However, from immediately after the discovery (A), the interpretation of quantum mechanics began to worry me. There are many interpretations

[3]Ref.[22]:S. Ishikawa, "Uncertainty relation in simultaneous measurements for arbitrary observables" Rep. Math. Phys. Vol.29(3), pp.257–273, 1991,

of quantum mechanics, for example, "the Copenhagen interpretation", "the many world interpretation", "the probabilistic interpretation", etc. In the applied field of quantum mechanics, we can expect that the same conclusion is derived from different interpretations. In this sense, the problem of "the interpretation of quantum mechanics" is not serious.

However, concerning Heisenberg's uncertainty principle, this problem is important. That is because the meaning of "errors" in Heisenberg's uncertainty principle depend on the interpretation of quantum mechanics $\Big($for example, the meaning of "errors (Δ_x and Δ_p)" depends on the acceptance of "the collapse of wave function" or not$\Big)$. Thus,

- I want to establish the "standard" interpretation of quantum mechanics.

In what follows, let me mention my idea (i.e., the linguistic interpretation of quantum mechanics):

Recalling that quantum mechanics was called "matrix mechanics" (when quantum mechanics was proposed (i.e., 1920s), I consider that

(B_1) **from the mathematical point of view, quantum mechanics is the theory of "square matrix"**

On the other hand,

(B_2) **from the mathematical point of view, classical mechanics is the theory of "diagonal matrix"**

Thus, we have the following problem:

(C) What is the interpretation which is common to both quantum system (B_1) and classical system (B_2)?

And we conclude that

(D) **the answer to the question (C) is uniquely determined as "quantum language",**

where quantum language can describe classical systems as well as quantum systems.

Since quantum language is not physics but language (= metaphysics), quantum language (= the linguistic interpretation of quantum mechanics) is completely different from other quantum interpretations. In this sense, we are convinced that

(E) **quantum language (= the linguistic interpretation of quantum mechanics) is forever**,
even if some propose the "final" interpretation of quantum mechanics in the realistic view (i.e., ⑤ in Figure 1.1)

KSTS/RR-16/001
January 8, 2016

1.2 The outline of quantum language

1.2.1 The classification of quantum language (=measurement theory)

Quantum language (= measurement theory) is classified as follows.

$$
\text{(A)} \quad \underset{\text{(=quantum language)}}{\text{measurement theory}} \begin{cases} \underset{(\mathrm{A}_1)}{\text{pure type}} \begin{cases} \text{classical system : Fisher statistics} \\ \text{quantum system : usual quantum mechanics} \end{cases} \\ \\ \underset{(\mathrm{A}_2)}{\text{mixed type}} \begin{cases} \text{classical system : including Bayesian statistics, Kalman filter} \\ \text{quantum system : quantum decoherence} \end{cases} \end{cases}
$$

Therefore, we have two kinds of quantum language, i.e., pure measurement theory and mixed measurement theory. The former is formulated as follows.

$$
(\mathrm{A}_1) \quad \underset{\text{(=quantum language)}}{\boxed{\text{pure measurement theory}}} := \underbrace{\overset{\text{[(pure)Axiom 1]}}{\underset{(\textit{cf.}\ \S 2.7)}{\boxed{\text{pure measurement}}}} + \overset{\text{[Axiom 2]}}{\underset{(\textit{cf.}\ \S 10.3)}{\boxed{\text{Causality}}}}}_{\text{a kind of spell(a priori judgment)}} + \underbrace{\overset{\text{[quantum linguistic interpretation]}}{\underset{(\textit{cf.}\ \S 3.1)}{\boxed{\text{Linguistic interpretation}}}}}_{\text{the manual to use spells}}
$$

And the mixed measurement theory (or, statistical measurement theory) is formulated as follows.

$$
(\mathrm{A}_2) \quad \underset{\text{(=quantum language)}}{\boxed{\text{mixed measurement theory}}} := \underbrace{\overset{\text{[(mixed)Axiom}^{(m)}\text{ 1]}}{\underset{(\textit{cf.}\ \S 9.1)}{\boxed{\text{mixed measurement}}}} + \overset{\text{[Axiom 2]}}{\underset{(\textit{cf.}\ \S 10.3)}{\boxed{\text{Causality}}}}}_{\text{a kind of spell(a priori judgment)}} + \underbrace{\overset{\text{[quantum linguistic interpretation]}}{\underset{(\textit{cf.}\ \S 3.1)}{\boxed{\text{Linguistic interpretation}}}}}_{\text{the manual to use spells}}
$$

1.2.2 Axiom 1 (measurement) and Axiom 2 (causality)

Since the pure measurement theory is the most fundamental, we mainly devote ourselves to pure measurement theory. Although it is impossible to read Axiom 1 (measurement: §2.7) and Axiom 2 (causality; §10.3) at the present time, we present them as follows.

(B):Axiom 1 (measurement) pure type

(This will be able to be read in §2.7)

With any system S, a basic structure $[\mathcal{A} \subseteq \overline{\mathcal{A}}]_{B(H)}$ can be associated in which measurement theory of that system can be formulated. In $[\mathcal{A} \subseteq \overline{\mathcal{A}}]_{B(H)}$, consider a **$W^*$-measurement** $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}{=}(X, \mathcal{F}, F), S_{[\rho]})$ $\Big($ or, **C^*-measurement** $\mathsf{M}_{\mathcal{A}}(\mathsf{O}{=}(X, \mathcal{F}, F), S_{[\rho]})$ $\Big)$. That is, consider

- a W^*-measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}, S_{[\rho]})$ $\Big($ or, C^*-measurement $\mathsf{M}_{\mathcal{A}}(\mathsf{O}{=}(X, \mathcal{F}, F), S_{[\rho]})$ $\Big)$ of an **observable** $\mathsf{O}{=}(X, \mathcal{F}, F)$ for a **state** $\rho (\in \mathfrak{S}^p(\mathcal{A}^*)$: state space)

Then, the probability that a measured value x $(\in X)$ obtained by the W^*-measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}, S_{[\rho]})$ $\Big($ or, C^*-measurement $\mathsf{M}_{\mathcal{A}}(\mathsf{O}{=}(X, \mathcal{F}, F), S_{[\rho]})$ $\Big)$ belongs to Ξ $(\in \mathcal{F})$ is given by

$$\rho(F(\Xi)) (\equiv {}_{\mathcal{A}^*}(\rho, F(\Xi))_{\overline{\mathcal{A}}}) \tag{1.1}$$

(if $F(\Xi)$ is essentially continuous at ρ, or see Definition 2.14).

And

(C): Axiom 2 (causality)

(This will be able to be read in §10.3)

Let T be a **tree** (i.e., semi-ordered tree structure). For each $t (\in T)$, a basic structure $[\mathcal{A}_t \subseteq \overline{\mathcal{A}}_t]_{B(H_t)}$ is associated. Then, the **causal chain** is represented by a **W^*- sequential causal operator** $\{\Phi_{t_1,t_2} : \overline{\mathcal{A}}_{t_2} \to \overline{\mathcal{A}}_{t_1}\}_{(t_1,t_2) \in T^2_{\leqq}}$ $\Big($ or, **C^*- sequential causal operator** $\{\Phi_{t_1,t_2} : \mathcal{A}_{t_2} \to \mathcal{A}_{t_1}\}_{(t_1,t_2) \in T^2_{\leqq}}$ $\Big)$

Here, note that

(D$_1$) **the above two axioms are kinds of spells (i.e., incantation, magic words, metaphysical statements), and thus, it is impossible to verify them experimentally.**

In this sense, the above two axioms correspond to "a priori synthetic judgment" in Kant's philosophy (*cf.* [53]). Therefore,

(D$_2$) **what we should do is not to understand the two, but to learn the spells (i.e., Axioms 1 and 2) by rote.**

Of course, the "learning by rote" means that we have to understand the mathematical definitions of followings:

- basic structure $[\mathcal{A} \subseteq \overline{\mathcal{A}}]_{B(H)}$, state space $\mathfrak{S}^p(\mathcal{A}^*)$, observable $\mathsf{O}{=}(X, \mathcal{F}, F)$, etc.

♠**Note 1.3.** If metaphysics did something wrong in the history of science, it is because metaphysics attempted to answer the following questions seriously in ordinary language:

($\sharp_1$) What is the meaning of the keywords (e.g., measurement, probability, causality) ?

Although the question ($\sharp_1$) looks attractive, it is not productive. What is important is *to create a language* to deal with the keywords. So we replace ($\sharp_1$) by

($\sharp_2$) How are the keywords (e.g., measurement, probability, causality) used in quantum language ?

The problem ($\sharp_1$) will now be solved in the sense of ($\sharp_2$).

♠**Note 1.4.** *Metaphysics* is an academic discipline concerning propositions in which empirical validation is impossible. Lord Kelvin (1824–1907) said

Mathematics is the only good metaphysics.

Here we step forward:

($\sharp$) *Quantum language is another good metaphysics.*

Lord Kelvin might think that Kant philosophy (Critique of Pure Reason [53]) is not good metaphysics. However, I consider that *a priori synthetic judgment* (i.e., axiom which cannot be examined by experiment) corresponds to [Axiom 1 and Axiom 2]. That is,

$$\underset{\text{(Kant philosophy)}}{\boxed{\text{a priori synthetic judgment}}} \quad \underset{\text{(correspondence)}}{\longleftrightarrow} \quad \underset{\text{(quantum language)}}{\boxed{\text{Axiom 1 and Axiom 2}}}$$

See ref. [31]:S. Ishikawa, *Quantum Mechanics and the Philosophy of Language: Reconsideration of traditional philosophies,* Journal of quantum information science, Vol. 2(1), pp.2-9, 2012

KSTS/RR-16/001
January 8, 2016

1.2.3 The linguistic interpretation

Axioms 1 and 2 are all of quantum language. Therefore,

($\sharp$) after learning Axioms 1 and 2 by rote, we need to brush up our skills to use them through trial and error.

Here, let us recall a wise saying

- *Experience is the best teacher*, or *custom makes all things*

and our experience

- A manual helps us to master the rules quickly.

Thus, we understand

$$
\begin{aligned}
& \text{to master the linguistic interpretation of quantum mechanics} \\
= \;& \text{to make practice with a manual to use Axioms 1 and 2}
\end{aligned}
$$

Although the linguistic interpretation (= the linguistic Copenhagen interpretation) is composed of many statements, the simplest and best representation may be as follows.

(E):The linguistic interpretation)

(This will be explained in §3.1)

Only one measurement is permitted.

We can also choose apparently opposite viewpoints concerning the linguistic interpretation, though they look a bit too extreme.

(E_1) Through trial and error, we can do well without the linguistic interpretation.

(E_2) All that are written in this note are a part of the linguistic interpretation.

They are viewpoints obtained from the opposite standpoints. In this sense, there is a reason to regard this lecture note as something like a cookbook.

♠**Note 1.5.** Kolmogorov's probability theory (*cf.* [54]) starts from the following spell:

(♯) Let $(X, \mathcal{F}, P)$ be a probability space. Then, the probability that a event $\Xi (\in \mathcal{F})$ happens is given by $P(\Xi)$

And, through trial and error, Kolmogorov found his extension theorem, which says that

(♯) **Only one probability space is permitted.**

This surely corresponds to the linguistic interpretation "Only one measurement is permitted." That is,

$$\underset{\text{(Only one probability space is permitted)}}{\overset{\text{(the most fundamental theorem)}}{\boxed{\text{Probability theory}}}} \quad \xleftrightarrow{\text{(correspondence)}} \quad \underset{\text{(Only one measurement is permitted)}}{\overset{\text{(the linguistic interpretation)}}{\boxed{\text{Quantum language}}}}$$

In this sense, we want to assert that

(♯) **Kolmogorov is one of the main discoverers of the linguistic interpretation.**

Therefore, we are optimistic to believe that the linguistic interpretation "Only one measurement is permitted" can be, after trial and error, acquired if we start from Axioms 1 and 2. That is, we consider, as mentioned in (H_1), that we can theoretically do well without the linguistic interpretation.

1.2.4 Summary

Summing up the above arguments, we see:

(F): Summary (All of quantum language)

Quantum language (= measurement theory) is formulated as follows.

$$\underset{\text{(=quantum language)}}{\boxed{\text{measurement theory}}} := \underbrace{\overset{\text{[Axiom 1]}}{\underset{(\textit{cf.}\ \S2.7)}{\boxed{\text{Measurement}}}} + \overset{\text{[Axiom 2]}}{\underset{(\textit{cf.}\ \S10.3)}{\boxed{\text{Causality}}}}}_{\text{a kind of spell(a priori judgment)}} + \underbrace{\overset{\text{[quantum linguistic interpretation]}}{\underset{(\textit{cf.}\ \S3.1)}{\boxed{\text{Linguistic interpretation}}}}}_{\text{manual to use spells}} \tag{1.2}$$

[Axioms]. Here

(F1) Axioms 1 and 2 are kinds of spells, (i.e., incantation, magic words, metaphysical statements), and thus, it is impossible to verify them experimentally. In this sense, I consider that

$$\underset{\text{(Kant philosophy)}}{\boxed{\text{a priori synthetic judgment}}} \xrightarrow[\text{quantization}]{} \underset{\text{(quantum language)}}{\boxed{\text{Axioms 1 and 2}}}$$

Therefore, what we should do is not "to understand" but "to use". After learning Axioms 1 and 2 by rote, we have to improve our skills to use them through trial and error.

[The linguistic interpretation]. From a pure theoretical point of view, we do well without the interpretation. However,

(F2) it is better to know the linguistic interpretation of quantum mechanics (= the manual to use Axioms 1 and 2), if we want to make quick progress in using quantum language.

The most important statement in the linguistic interpretation (§3.1) is

Only one measurement is permitted.

After all, we think that

$$\underset{\text{[dualistic idealism]}}{\text{Descartes philosophy}} \longrightarrow \left\{ \begin{array}{c} \underset{\text{[Axioms]}}{\text{Continental Rationalism}} \\ \\ \underset{\text{[Linguistic interpretation]}}{\text{British empiricism}} \end{array} \right\} \longrightarrow \underset{\text{[quantum language]}}{\text{Kant philosophy}}$$

1.3 Example: "Cold" or "Hot"

Axioms 1 and 2 (mentioned in the previous section) are too abstract. And thus, I am afraid that the readers feel that it is too hard to use quantum language. Hence, let us add a simple example in this section.

It is sufficient for the readers to consider that our purpose in the next chapters is

- to bury the gap between Axiom 1 and the following simple example (i.e., "Cold" or "Hot").

Example 1.2. [The measurement of "Cold or Hot" for the water in a cup] Let testees drink water with various temperature ω °C ($0 \leqq \omega \leqq 100$). And assume: you ask them "Cold or Hot ?" alternatively. Gather the data, (for example, $g_c(\omega)$ persons say "Cold", $g_h(\omega)$ persons say "Hot") and normalize them, that is, get the polygonal lines such that

$$
\begin{aligned}
f_c(\omega) &= \frac{g_c(\omega)}{\text{the numbers of testees}} \\
f_h(\omega) &= \frac{g_h(\omega)}{\text{the numbers of testees}}
\end{aligned}
\tag{1.3}
$$

And

$$
f_c(\omega) = \begin{cases} 1 & (0 \leqq \omega \leqq 10) \\ \frac{70-\omega}{60} & (10 \leqq \omega \leqq 70) \\ 0 & (70 \leqq \omega \leqq 100) \end{cases} , \qquad f_h(\omega) = 1 - f_c(\omega)
$$

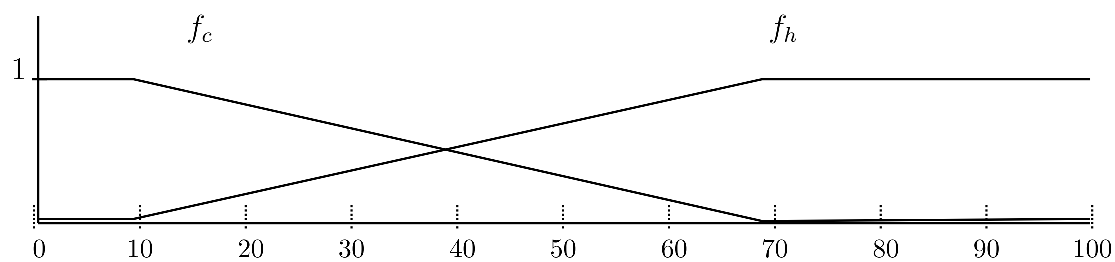


Figure 1.2: Cold or hot?

Therefore, for example,

(A$_1$) You choose one person from the testees, and you ask him/her whether the water (with 55 °C) is "cold" or "hot" ?. Then the probability that he/she says $\begin{bmatrix} \text{"cold"} \\ \text{"hot"} \end{bmatrix}$ is given by $\begin{bmatrix} f_c(55) = 0.25 \\ f_h(55) = 0.75 \end{bmatrix}$

In what follows, let us describe the statement (A_1) in terms of quantum language (i.e., Axiom 1).

Define the state space Ω such that Ω = interval $[0, 100] (\subset \mathbb{R}(=$ the set of all real numbers$))$ and measured value space $X = \{c, h\}$ (where "c" and "h" respectively means "cold" and "hot"). Here, consider the "[C-H]-thermometer" such that

(A_2) for water with ω °C, [C-H]-thermometer presents $\begin{bmatrix} \mathrm{c} \\ \mathrm{h} \end{bmatrix}$ with probability $\begin{bmatrix} f_c(\omega) \\ f_h(\omega) \end{bmatrix}$. This [C-H]-thermometer is denoted by $\mathsf{O} = (f_c, f_h)$

Note that this [C-H]-thermometer can be easily realized by "random number generator".

Here, we have the following identification:

$$(\mathrm{A}_3) \qquad (\mathrm{A}_1) \Longleftrightarrow (\mathrm{A}_2)$$

Therefore, the statement (A_1) in ordinary language can be represented in terms of measurement theory as follows.

(A_4) When an **observer** takes a measurement by $\underset{\textbf{measuring instrument}\mathsf{O}=(f_c, f_h)}{\text{[[C-H]-instrument]}}$ for $\underset{\textbf{(System (measuring object))}}{\text{[water]}}$ with $\underset{(\textbf{state}(=\omega \in \Omega)\ )}{[55\ ^\circ\mathrm{C}]}$, the probability that **measured value** $\begin{bmatrix} \mathrm{c} \\ \mathrm{h} \end{bmatrix}$ is obtained is given by $\begin{bmatrix} f_c(55) = 0.25 \\ f_h(55) = 0.75 \end{bmatrix}$

This example will be again discussed in the following chapter(Example 2.31).

Chapter 2

Axiom 1 — measurement

Quantum language (= measurement theory) is formulated as follows.

- $\underset{\text{(=quantum language)}}{\boxed{\text{measurement theory}}} := \underbrace{\overset{\text{[Axiom 1]}}{\underset{(cf.\ \S 2.7)}{\boxed{\text{Measurement}}}} + \overset{\text{[Axiom 2]}}{\underset{(cf.\ \S 10.3)}{\boxed{\text{Causality}}}}}_{\text{a kind of spell(a priori judgment)}} + \underbrace{\overset{\text{[quantum linguistic interpretation]}}{\underset{(cf.\ \S 3.1)}{\boxed{\text{Linguistic interpretation}}}}}_{\text{manual to use spells}}$

Measurement theory asserts that

- Describe every phenomenon modeled on Axioms 1 and 2 (by a hint of the linguistic interpretation)!

In this chapter, we introduce Axiom 1 (measurement). Axiom 2 concerning causality will be explained in Chapter 10.

2.1 The basic structure[$\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)$]; General theory

The Hilbert space formulation of quantum mechanics is due to von Neumann. I cannot emphasize too much the importance of his work (*cf.* [70]).

2.1.1 Hilbert space and operator algebra

Let H be a complex Hilbert space with a inner product $\langle \cdot, \cdot \rangle$, where it is assumed that $\langle u, \alpha v \rangle = \alpha \langle u, v \rangle$ $(\forall u, v \in H, \alpha \in \mathbb{C}$(= the set of all complex numbers)). And define the norm $\|u\| = |\langle u, u \rangle|^{1/2}$. Define $B(H)$ by

$$B(H) = \{T : H \to H \mid T \text{ is a continuous linear operator}\} \tag{2.1}$$

$B(H)$ is regarded as the Banach space with the operator norm $\| \cdot \|_{B(H)}$, where

$$\|T\|_{B(H)} = \sup_{\|x\|_H = 1} \|Tx\|_H \quad (\forall T \in B(H)) \tag{2.2}$$

Let $T \in B(H)$. The dual operator $T^* \in B(H)$ of T is defined by

$$\langle T^* u, v \rangle = \langle u, Tv \rangle \quad (\forall u, v \in H)$$

The followings are clear.

$$(T^*)^* = T, \quad (T_1 T_2)^* = T_2^* T_1^*$$

Further, the following equality (called the "C^*-condition") holds:

$$\|T^*T\| = \|TT^*\| = \|T\|^2 = \|T^*\|^2 \quad (\forall T \in B(H)) \tag{2.3}$$

When $T = T^*$ holds, T is called a self-adjoint operator (or, Hermitian operator). Let $T_n (n \in \mathbb{N} = \{1, 2, \cdots\}), T \in B(H)$. The sequence $\{T_n\}_{n=1}^\infty$ is said to converge weakly to T (that is, $w - \lim_{n \to \infty} T_n = T$), if

$$\lim_{n \to \infty} \langle u, (T_n - T)u \rangle = 0 \qquad (\forall u \in H) \tag{2.4}$$

Thus, we have two convergences (i.e., norm convergence and weakly convergence) in $B(H)$[1].

Definition 2.1. [C^*-algebra and W^*-algebra] $\mathcal{A}(\subseteq B(H))$ is called a C^*-algebra, if it satisfies that

(A$_1$) $\mathcal{A}(\subseteq B(H))$ is the closed linear space in the sense of the operator norm $\| \cdot \|_{B(H)}$.

(A$_2$) $\mathcal{A}$ is $*$-algebra, that is, $\mathcal{A}(\subseteq B(H))$ satisfies that

$$F_1, F_2 \in \mathcal{A} \Rightarrow F_1 \cdot F_2 \in \mathcal{A}, \qquad F \in \mathcal{A} \Rightarrow F^* \in \mathcal{A}$$

Also, a C^*-algebra$\mathcal{A}(\subseteq B(H))$ is called a W^*-algebra, if it is weak closed in $B(H)$.

2.1.2 Basic structure$[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]$; general theory

Definition 2.2. Consider the basic structure $[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]$ $\Big($ or, denoted by $[\mathcal{A} \subseteq \overline{\mathcal{A}}]_{B(H)}$ $\Big)$. That is,

- $\mathcal{A}(\subseteq B(H))$ is a C^*-algebra, and $\overline{\mathcal{A}}(\subseteq B(H))$ is the weak closure of $\mathcal{A}$.

Note that W^*-algebra $\overline{\mathcal{A}}$ has the pre-dual Banach space $\overline{\mathcal{A}}_*$(that is, $(\overline{\mathcal{A}}_*)^* = \overline{\mathcal{A}}$) uniquely. Therefore, the basic structure$[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]$ is represented as follows.

(B): General basic structure:$[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]$

$$\begin{array}{ccccc} \mathcal{A}^* & & & & \\ \uparrow{\scriptstyle \text{dual}} & & & & \\ \boxed{\mathcal{A}} & \xrightarrow[\text{subalgebra}\cdot\text{weak-closure}]{\subseteq} & \boxed{\overline{\mathcal{A}}} & \xrightarrow[\text{subalgebra}]{\subseteq} & \boxed{B(H)} \\ & & \downarrow{\scriptstyle \text{pre-dual}} & & \\ & & \overline{\mathcal{A}}_* & & \end{array} \tag{2.5}$$

[1]Although there are many convergences in $B(H)$, in this paper we devote ourselves to the two.

2.1.3 Basic structure$[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]$ and state space; General theory

The concept of "state space" is fundamental in quantum language. This is formulated in the dual space $\mathcal{A}^*$ of C^*-algebra $\mathcal{A}$ (or, in the pre-dual space $\overline{\mathcal{A}}_*$ of W^*-algebra $\overline{\mathcal{A}}$).

Let us explain it as follows.

Definition 2.3. [State space, mixed state space] Consider the basic structure:

$$[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]$$

Let $\mathcal{A}^*$ be the dual space of the C^*-algebra$\mathcal{A}$. The mixed state space $\mathfrak{S}^m(\mathcal{A}^*)$ and the pure state space $\mathfrak{S}^p(\mathcal{A}^*)$ is respectively defined by

(a) $\mathfrak{S}^m(\mathcal{A}^*) = \{\rho \in \mathcal{A}^* \mid \|\rho\|_{\mathcal{A}^*} = 1, \rho \geq 0 \text{ (i.e., } \rho(T^*T) \geq 0 (\forall T \in \mathcal{A}))\}$

(b) $\mathfrak{S}^p(\mathcal{A}^*) = \{\rho \in \mathfrak{S}^m(\mathcal{A}^*) \mid \rho \text{ is a pure state}\}$. Here, $\rho(\in \mathfrak{S}^m(\mathcal{A}^*))$ is a pure state if and only if

$$\rho = \alpha\rho_1 + (1-\alpha)\rho_2, \quad \rho_1, \rho_2 \in \mathfrak{S}^m(\mathcal{A}^*), 0 < \alpha < 1 \Longrightarrow \rho = \rho_1 = \rho_2$$

The mixed state space $\mathfrak{S}^m(\mathcal{A}^*)$ and the pure state space $\mathfrak{S}^p(\mathcal{A}^*)$ are locally compact spaces (*cf.* ref.[74]).

Assume that $\overline{\mathcal{A}}_*$ is the pre-dual space of $\overline{\mathcal{A}}$. Then, another mixed state space $\overline{\mathfrak{S}}^m(\overline{\mathcal{A}}_*)$ is defined by

(c) $\overline{\mathfrak{S}}^m(\overline{\mathcal{A}}_*) = \{\rho \in \overline{\mathcal{A}}_* \mid \|\rho\|_{\overline{\mathcal{A}}_*} = 1, \rho \geq 0 \text{ (i.e., } \rho(T^*T) \geq 0 (\forall T \in \overline{\mathcal{A}}))\}$

That is, we have two "mixed state spaces", that is, C^*-mixed state space $\mathfrak{S}^m(\mathcal{A}^*)$ and W^*-mixed state space $\overline{\mathfrak{S}}^m(\overline{\mathcal{A}}_*)$.

The above arguments are summarized in the following figure:

(C): General basic structure and State spaces

$$
\begin{tikzcd}
\underset{C^*\text{-pure state}}{\mathfrak{S}^p(\mathcal{A}^*)} \subset \underset{C^*\text{-mixed state}}{\mathfrak{S}^m(\mathcal{A}^*)} \subset \mathcal{A}^* & & \\
\boxed{\mathcal{A}} \arrow[u, "\text{dual}"'] \arrow[r, "\subseteq", "\text{subalgebra·weak-closure}"'] & \boxed{\overline{\mathcal{A}}} \arrow[r, "\subseteq", "\text{subalgebra}"'] \arrow[d, "\text{pre-dual}"] & \boxed{B(H)} \\
& \underset{W^*\text{-mixed state}}{\overline{\mathfrak{S}}^m(\overline{\mathcal{A}}_*)} \subset \overline{\mathcal{A}}_* &
\end{tikzcd}
\tag{2.6}
$$

Remark 2.4. In order to avoid the confusions, three "state spaces" should be explained in what follows.

$$
\text{(D)}\quad \text{“state spaces”}\begin{cases} \text{Fisher statistics} & \cdots \text{pure state space:} \mathfrak{S}^p(\mathcal{A}^*)\text{: most fundamental} \\ \text{Bayes statistics} & \cdots \begin{cases} C^*\text{-mixed state space:} \mathfrak{S}^m(\mathcal{A}^*) : \text{easy} \\ W^*\text{-mixed state space:} \overline{\mathfrak{S}}^m(\overline{\mathcal{A}}_*)\text{: natural, useful} \end{cases} \end{cases}
$$

In this note, we mainly devote ourselves to the W^*-mixed state$\overline{\mathfrak{S}}^m(\overline{\mathcal{A}}_*)$ rather than the C^*-mixed state$\mathfrak{S}^m(\mathcal{A}^*)$, though the two play the similar roles in quantum language.

2.2 Quantum basic structure$[\mathcal{C}(H) \subseteq B(H) \subseteq B(H)]$ and State space

If a conclusion is said previously, we say the following classification of (i.e., quantum state space and classical state space):

(A)

$\boxed{\text{General basic structure}[\mathcal{A} \subseteq \overline{\mathcal{A}}]_{B(H)}}$

pure state space $\mathfrak{S}^p(\mathcal{A}^*)$
C^*-mixed state space $\mathfrak{S}^m(\mathcal{A}^*)$
W^*-mixed state *space* $\overline{\mathfrak{S}}^m(\overline{\mathcal{A}}_*)$

$\Longrightarrow$

$\boxed{\text{(A}_1\text{):Quantum basic structure}[\mathcal{C}(H) \subseteq B(H)]_{B(H)}}$

pure state space $\mathfrak{S}^p(\mathcal{T}r(H)(\approx H))$
C^*-mixed state space $\mathfrak{S}^m(\mathcal{T}r(H))(=\mathcal{T}r_{+1}(H))$
W^*-mixed state space $\mathfrak{S}^m(\mathcal{T}r(H))(=\mathcal{T}r_{+1}(H))$

$\boxed{\text{(A}_2\text{):Classical basic structure}[C_0(\Omega) \subseteq L^\infty(\Omega, \nu)]_{B(L^2(\Omega,\nu))}}$

pure state space Ω
C^*-mixed state space $\mathcal{M}_{+1}(\Omega)$
W^*-mixed state space $L^1_{+1}(\Omega,\nu)$

In what follows, we shall explain the above classification (A):

2.2.1 Quantum basic structure$[\mathcal{C}(H) \subseteq B(H) \subseteq B(H)]$;

In quantum system, the basic structure$[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]$ is characterized as

$$[\mathcal{C}(H) \subseteq B(H) \subseteq B(H)] \tag{2.7}$$

That is, we see:

(B): Quantum basic structure:$[\mathcal{C}(H) \subseteq B(H) \subseteq B(H)]$

$$\begin{tikzcd} \mathcal{T}r(H) & & \\ \boxed{\mathcal{C}(H)} \arrow[u, "\text{dual}"'] \arrow[r, "\subseteq", "\text{subalgebra}\cdot\text{weak-closure}"'] & \boxed{B(H)} \arrow[d, "\text{pre-dual}"] \arrow[r, "\subseteq", "\text{subalgebra}"'] & \boxed{B(H)} \\ & \mathcal{T}r(H) & \end{tikzcd} \tag{2.8}$$

Before we explain "compact operators class $\mathcal{C}(H)$" and "trace class $\mathcal{F}(H)$", we have to prepare "Dirac notation" and "CONS" as follows.

Definition 2.5. [(i):Dirac notation] Let H be a Hilbert space. For any $u, v \in H$, define $|u\rangle\langle v| \in B(H)$ such that

$$(|u\rangle\langle v|)w = \langle v, w\rangle u \quad (\forall w \in H) \tag{2.9}$$

Here, $\langle v|$ $\big[$ resp. $|u\rangle$ $\big]$ is called the "Bra-vector" $\big[$ resp. "Ket-vector"$\big]$.

[(ii):ONS(orthonormal system), CONS(complete orthonormal system)] The sequence $\{e_k\}_{k=1}^\infty$ in a Hilbert space H is called an orthonormal system (i.e., ONS), if it satisfies

$(\sharp_1)$ $\langle e_k, e_j\rangle = \begin{cases} 1 & (k=j) \\ 0 & (k \neq j) \end{cases}$

In addition, an ONS $\{e_k\}_{k=1}^\infty$ is called a complete orthonormal system (i.e., CONS), if it satisfies

$(\sharp_2)$ $\langle x, e_k\rangle = 0$ $(\forall k = 1, 2, ...)$ implies that $x = 0$.

Theorem 2.6. [The properties of compact operators class $\mathcal{C}(H)$] Let $\mathcal{C}(H)(\subseteq B(H))$ be the compact operators class. Then, we see the following (C_1)-(C_4) $\Big($ particularly, "$(C_1) \leftrightarrow (C_2)$" may be regarded as the definition of the compact operators class $\mathcal{C}(H)(\subseteq B(H))$ $\Big)$.

(C_1) $T \in \mathcal{C}(H)$. That is,

- for any bounded sequence $\{u_n\}_{n=1}^\infty$ in Hilbert space H, $\{Tu_n\}_{n=1}^\infty$ has the subsequence which converges in the sense of the norm topology.

(C_2) There exist two ONSs $\{e_k\}_{k=1}^\infty$ and $\{f_k\}_{k=1}^\infty$ in the Hilbert space H and a positive real sequence $\{\lambda_k\}_{k=1}^\infty$ (where, $\lim_{k\to\infty} \lambda_k = 0$) such that

$$T = \sum_{k=1}^{\infty} \lambda_k |e_k\rangle\langle f_k| \qquad \text{(in the sense of weak topology)} \tag{2.10}$$

(C_3) $\mathcal{C}(H)(\subseteq B(H))$ is a C^*-algebra. When $T(\in \mathcal{C}(H))$ is represented as in (C_2), the following equality holds

$$\|T\|_{B(H)} = \max_{k=1,2,\cdots} \lambda_k \tag{2.11}$$

(C_4) The weak closure of $\mathcal{C}(H)$ is equal to $B(H)$. That is,

$$\overline{\mathcal{C}(H)} = B(H) \tag{2.12}$$

Theorem 2.7. [The properties of trace class $\mathcal{T}r(H)$] Let $\mathcal{T}r(H)(\subseteq B(H))$ be the trace class. Then, we see the following (3D$_1$)-(D$_4$)(particularly, "(D$_1$)$\leftrightarrow$ (D$_2$)" may be regarded as the definition of the trace class $\mathcal{T}r(H)(\subseteq B(H))$).

(D$_1$) $T \in \mathcal{T}r(H)(\subseteq \mathcal{C}(H) \subseteq B(H))$.

(D$_2$) There exist two ONSs $\{e_k\}_{k=1}^{\infty}$ and $\{f_k\}_{k=1}^{\infty}$ in the Hilbert space H and a positive real sequence $\{\lambda_k\}_{k=1}^{\infty}$ (where, $\sum_{k=1}^{\infty} \lambda_k < \infty$) such that

$$T = \sum_{k=1}^{\infty} \lambda_k |e_k\rangle\langle f_k| \qquad \text{(in the sense of weak topology)}$$

(D$_3$) It holds that

$$\mathcal{C}(H)^* = \mathcal{T}r(H) \tag{2.13}$$

Here, the dual norm $\|\cdot\|_{\mathcal{C}(H)^*}$ is characterized as the trace norm $\|\cdot\|_{Tr}$ such as

$$\|T\|_{Tr} = \sum_{k=1}^{\infty} \lambda_k \tag{2.14}$$

when $T(\in \mathcal{T}r(H))$ is represented as in (D$_2$),

(D$_4$) Also, it holds that

$$\mathcal{T}r(H)^* = B(H) \qquad \text{in the same sense,} \qquad \mathcal{T}r(H) = B(H)_* \tag{2.15}$$

Remark 2.8. Assume that a Hilbert space H is finite dimensional, i.e., $H = \mathbb{C}^n$, i.e., $\mathbb{C}^n = \{z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ x_n \end{bmatrix} \mid z_k \in \mathbb{C}, k = 1, 2, ..., n\}$. Put

$$M(\mathbb{C}, n) = \text{The set of all } (n \times n)\text{-complex matrices}$$

and thus,

$$\mathcal{A} = \overline{\mathcal{A}} = B(\mathbb{C}^n) = \mathcal{C}(H) = \mathcal{T}r(H) = M(\mathbb{C}, n) \tag{2.16}$$

However, it should be noted that the norms are different as mentioned in (C$_3$) and (D$_3$).

2.2.2 Quantum basic structure$[\mathcal{C}(H) \subseteq B(H) \subseteq B(H)]$ and State space;

Consider the quantum basic structure:

$$[\mathcal{C}(H) \subseteq B(H) \subseteq B(H)]$$

and see the following diagram:

(E): Quantum basic structure and State space

$$\begin{tikzcd}
\underset{C^*\text{-pure state}}{\mathfrak{S}^p(\mathcal{T}r(H))} \subset \underset{C^*\text{-mixed state}}{\mathfrak{S}^m(\mathcal{T}r(H))} \subset \mathcal{T}r(H) & & \\
\boxed{\mathcal{C}(H)} \arrow[u, "\text{dual}"'] \arrow[r, "\subseteq", "\text{subalgebra}\cdot\text{weak-closure}"'] & \boxed{B(H)} \arrow[d, "\text{pre-dual}"] \arrow[r, "\subseteq", "\text{subalgebra}"'] & \boxed{B(H)} \\
& \underset{W^*\text{-mixed state}}{\overline{\mathfrak{S}}^m(\mathcal{T}r(H))} \subset \mathcal{T}r(H) &
\end{tikzcd} \tag{2.17}$$

In what follows, we shall explain the above diagram.

Firstly, we note that

$$\mathcal{C}(H)^* = \mathcal{T}r(H), \qquad \mathcal{T}r(H)^* = B(H) \tag{2.18}$$

and

$$\begin{aligned}
&\mathfrak{S}^m(\mathcal{T}r(H)) = \overline{\mathfrak{S}}^m(\mathcal{T}r(H)) \\
&=\{\rho = \sum_{n=1}^\infty \lambda_n |e_n\rangle\langle e_n| \;\;:\;\; \{e_n\}_{n=1}^\infty \text{ is ONS },\; \sum_{n=1}^\infty \lambda_n = 1, \lambda_n > 0\} \\
&=: \mathcal{T}r_{+1}(H)
\end{aligned} \tag{2.19}$$

Also, concerning the pure state space, we see:

$$\begin{aligned}
&\mathfrak{S}^p(\mathcal{T}r(H)) \\
&=\{\rho = |e\rangle\langle e| \;\;:\;\; \|e\|_H = 1\} =: \mathcal{T}r_{+1}^p(H)
\end{aligned} \tag{2.20}$$

Therefore, under the following identification:

$$\mathfrak{S}^p(\mathcal{T}r(H)) \ni |u\rangle\langle u| \underset{\text{identification}}{\longleftrightarrow} u \in H \qquad (\|u\| = 1) \tag{2.21}$$

we see,

$$\mathfrak{S}^p(\mathcal{T}r(H)) = \{u \in H \;:\; \|u\| = 1\} \tag{2.22}$$

where we assume the equivalence: $u \approx e^{i\theta} u$ $(\theta \in \mathbb{R})$.

Definition 2.9. Define the trace $\mathrm{Tr} : \mathcal{T}r(H) \to \mathbb{C}$ such that

$$\mathrm{Tr}(T) = \sum_{n=1}^{\infty} \langle e_n, Te_n \rangle \qquad (\forall T \in \mathcal{T}r(H)) \tag{2.23}$$

where $\{e_n\}_{n=1}^{\infty}$ is a CONS in H. It is well known that the $\mathrm{Tr}(T)$ does not depend on the choice of CONS $\{e_n\}_{n=1}^{\infty}$. Thus, clearly we see that

$$_{\mathcal{T}rH}\Big(|u\rangle\langle u|, F\Big)_{B(H)} = \mathrm{Tr}(|u\rangle\langle u| \cdot F) = \langle uFu \rangle \quad (\forall ||u||_H = 1, F \in B(H)) \tag{2.24}$$

Remark 2.10. Assume that a Hilbert space H is finite dimensional, i.e., $H = \mathbb{C}^n$. Then,

$$M(\mathbb{C}, n) = \text{The set of all } (n \times n)\text{-complex matrices}$$

That is,

$$F = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ f_{n1} & f_{n2} & \cdots & f_{nn} \end{bmatrix} \in M(\mathbb{C}, n) \tag{2.25}$$

As mentioned before, we see

$$\mathcal{A} = \overline{\mathcal{A}} = B(\mathbb{C}^n) = \mathcal{C}(H) = \mathcal{T}r(H) = M(\mathbb{C}, n) \tag{2.26}$$

and further, under the following notations:

$$\mathcal{T}r_{+1}^{D}(\mathbb{C}^n) = \Big\{ \text{diagonal matrix} F = \begin{bmatrix} f_{11} & 0 & \cdots & 0 \\ 0 & f_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f_{nn} \end{bmatrix} \;\Big|\; f_{kk} \geq 0, \; \sum_{k=1}^{n} f_{kk} = 1 \Big\}$$

$$\mathcal{T}r_{+1}^{DP}(\mathbb{C}^n) = \Big\{ F = \begin{bmatrix} f_{11} & 0 & \cdots & 0 \\ 0 & f_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f_{nn} \end{bmatrix} \in \mathcal{T}r_{+1}^{D}(\mathbb{C}^n) \;\Big|\; f_{kk} = 1 \text{ (for some } k = j), = 0 \ (k \neq j) \Big\}$$

We see,

$$\text{mixed state space: } \mathcal{T}r_{+1}(\mathbb{C}^n) = \Big\{ UFU^* \ : \ F \in \mathcal{T}r_{+1}^{D}(\mathbb{C}^n), \ U \text{ is a unitary matrix} \Big\} \tag{2.27}$$

$$\text{pure state space: } \mathcal{T}r_{+1}^{p}(\mathbb{C}^n) = \Big\{ UFU^* \ : \ F \in \mathcal{T}r_{+1}^{DP}(\mathbb{C}^n), \ U \text{ is a unitary matrix} \Big\} \tag{2.28}$$

2.3 Classical basic structure$[C_0(\Omega) \subseteq L^\infty(\Omega,\nu) \subseteq B(L^2(\Omega,\nu))]$

2.3.1 Classical basic structure$[C_0(\Omega) \subseteq L^\infty(\Omega,\nu) \subseteq B(L^2(\Omega,\nu))]$

In classical systems, the basic structure$[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]$ is restricted to the classical basic structure:

$$[C_0(\Omega) \subseteq L^\infty(\Omega,\nu) \subseteq B(L^2(\Omega,\nu))]$$

And we get the following diagram:

(A): Classical basic structure: $[C_0(\Omega) \subseteq L^\infty(\Omega,\nu) \subseteq B(L^2(\Omega,\nu))]$

$$\begin{array}{ccccc} \mathcal{M}(\Omega) & & & & \\ \uparrow{\scriptstyle \text{dual}} & & & & \\ \boxed{C_0(\Omega)} & \xrightarrow[\text{subalgebra·weak-closure}]{\subseteq} & \boxed{L^\infty(\Omega,\nu)} & \xrightarrow[\text{subalgebra}]{\subseteq} & \boxed{B(L^2(\Omega,\nu))} \\ & & \downarrow{\scriptstyle \text{pre-dual}} & & \\ & & L^1(\Omega,\nu) & & \end{array} \tag{2.29}$$

In what follows, we shall explain this diagram.

2.3.1.1 Commutative C^*-algebra $C_0(\Omega)$ and Commutative W^*-algebra $L^\infty(\Omega,\nu)$

Let Ω a locally compact space, for example, it suffices to image Ω as follows.

$$\mathbb{R}(= \text{the real line}), \quad \mathbb{R}^2(= \text{plane}), \quad \mathbb{R}^n(= n\text{-dimensional Euclidean space}),$$

$$[a,b](= \text{interval}), \quad \underset{(\text{with discrete metric } d_D)}{\text{finite set}\Omega(= \{\omega_1,...,\omega_n\})}$$

where the discrete metric d_D is defined by $d_D(\omega,\omega') = 1$ $(\omega \neq \omega'), = 0$ $(\omega = \omega')$.
Define the continuous functions space $C_0(\Omega)$ such that

$$C_0(\Omega) = \{f : \Omega \to \mathbb{C} \mid f \text{ is complex-valued continuous on } \Omega, \ \lim_{\omega\to\infty} f(\omega) = 0\} \tag{2.30}$$

where "$\lim_{\omega\to\infty} f(\omega) = 0$" means

(B) for any positive real $\epsilon > 0$, there exists a compact set $K(\subseteq \Omega)$ such that

$$\{\omega \mid \omega \in \Omega \setminus K, |f(\omega)| > \epsilon\} = \emptyset$$

Therefore, if Ω is compact, the, the condition "$\lim_{\omega\to\infty} f(\omega) = 0$" is not needed, and thus, $C_0(\Omega)$ is usually denoted by $C(\Omega)$. In this note, even if Ω is compact, we often denote $C(\Omega)$ by $C_0(\Omega)$.

Defining the norm $\|\cdot\|_{C_0(\Omega)}$ in a complex vector space $C_0(\Omega)$ such that

$$\|f\|_{C_0(\Omega)} = \max_{\omega\in\Omega} |f(\omega)| \tag{2.31}$$

we get the Banach space $\Big(C_0(\Omega), \|\cdot\|_{C_0(\Omega)}\Big)$.

Let Ω be a locally compact space, and consider the σ-finite measure space $(\Omega, \mathcal{B}_\Omega, \nu)$, where, $\mathcal{B}_\Omega$ is the Borel field, i.e., the smallest σ-field that contains all open sets. Further, assume that

(C) for any open set $U \subseteq \Omega$, it holds that $0 < \nu(U) \leqq \infty$

♠**Note 2.1.** Without loss of generality, we can assume that Ω is compact by the Stone-Čech compactification. Also, we can assume that $\nu(\Omega) = 1$.

Define the Banach space $L^r(\Omega,\nu)$ (where, $r = 1, 2, \infty$) by the all complex-valued measurable functions $f : \Omega \to \mathbb{C}$ such that

$$\|f\|_{L^r(\Omega,\nu)} < \infty$$

The norm $\|f\|_{L^r(\Omega,\nu)}$ is defined by

$$\|f\|_{L^r(\Omega,\nu)} = \begin{cases} \Big[\int_\Omega |f(\omega)|^r \,\nu(d\omega)\Big]^{1/r} & (\text{when } r = 1, 2) \\ \\ \underset{\omega\in\Omega}{\text{ess.sup}}|f(\omega)| & (\text{when } r = \infty) \end{cases} \tag{2.32}$$

where

$$\text{ess.sup}_{\omega\in\Omega}|f(\omega)| = \sup\{a \in \mathbb{R} \mid \nu(\{\omega \in \Omega \;:\; |f(\omega)| \geqq a\;\}) > 0\}$$

$L^r(\Omega,\nu)$ is often denoted by $L^r(\Omega)$ or $L^r(\Omega, \mathcal{B}_\Omega, \nu)$.

Remark 2.11. $[C_0(\Omega) \subseteq L^\infty(\Omega,\nu) \subseteq B(L^2(\Omega,\nu))]$ Consider a Hilbert space H such that

$$H = L^2(\Omega,\nu)$$

For each $f \in L^\infty(\Omega)$, define $T_f \in B(L^2(\Omega,\nu))$ such that

$$L^2(\Omega,\nu) \ni \phi \longrightarrow T_f(\phi) = f \cdot \phi \in L^2(\Omega,\nu)$$

Then, under the identification:

$$L^\infty(\Omega) \ni f \underset{\text{identification}}{\longleftrightarrow} T_f \in B(L^2(\Omega,\nu)) \tag{2.33}$$

we see that

$$f \in L^\infty(\Omega) \subseteq B(L^2(\Omega,\nu))$$

and further, we have the classical basic structure:

$$[C_0(\Omega) \subseteq L^\infty(\Omega) \subseteq B(L^2(\Omega,\nu))] \tag{2.34}$$

This will be shown in what follows.

Riese theorem (*cf.* [74]) says that

$$C_0(\Omega)^* = \mathcal{M}(\Omega)(= \text{the set of all complex-valued measures on } \Omega\ ) \tag{2.35}$$

Therefore, for any $F \in C_0(\Omega)$, $\rho \in C_0(\Omega)^* = \mathcal{M}(\Omega)$, we have the bi-linear form which is written by the several ways such as

$$\rho(F) = {}_{C_0(\Omega)^*}\Big(\rho, F\Big)_{C_0(\Omega)} = {}_{\mathcal{M}(\Omega)}\Big(\rho, F\Big)_{C_0(\Omega)} = \int_\Omega F(\omega)\rho(d\omega) \tag{2.36}$$

Also, the dual norm is calculated as follows.

$$\begin{aligned} \|\rho\|_{C_0(\Omega)^*} &= \sup\{|\rho(F)| \mid \|F\|_{C_0(\Omega)} = 1\} = \sup_{\|F\|_{C_0(\Omega)}=1} |\int_\Omega F(\omega)\rho(d\omega)| \\ =& \sup_{\Xi,\Gamma\in\mathcal{B}_\Omega} \Big(|Re(\rho(\Xi)) - Re(\rho(\Xi^c))|^2 + |Im(\rho(\Gamma)) - Im(\rho(\Gamma^c))|^2 \Big)^{1/2} \\ =& \|\rho\|_{\mathcal{M}(\Omega)} \end{aligned} \tag{2.37}$$

where, Ξ^c is the complement of Ξ, and $Re(z)$="the real part of the complex number z", $Im(z)$="the imaginary part of the complex number z".

Further, we see that

$$L^1(\Omega,\nu)^* = L^\infty(\Omega,\nu) \qquad \text{in the same sense,} \qquad L^1(\Omega,\nu) = L^\infty(\Omega,\nu)_*$$

Also, it is clear that

$$C_0(\Omega) \subseteq L^\infty(\Omega,\nu)$$

For any $f \in L^\infty(\Omega,\nu)$, there exist $f_n \in C_0(\Omega), n = 1,2,..$ such that

$$\begin{cases} \nu(\{\omega \in \Omega \mid \lim_{n\to\infty} f_n(\omega) \neq f(\omega)\} = 0 \\ \\ |f_n(\omega)| \leq \|f\|_{L^\infty(\Omega,\nu)} \quad (\forall\omega \in \Omega, \forall n = 1,2,3,...) \end{cases}$$

Therefore, we see

$$\lim_{n\to\infty} |\Big\langle \phi, (f - f_n)\phi \Big\rangle_{L^2(\Omega,\nu)}| \le \lim_{n\to\infty} \int_\Omega |f_n(\omega) - f(\omega)| \cdot |\phi(\omega)|^2 \nu(d\omega) = 0 \quad (\forall \phi \in L^2(\Omega,\nu))$$

Hence,

the weak closure of $C_0(\Omega)$ is equal to $L^\infty(\Omega,\nu)$

Then, we have the classical basic structure:

$$[C_0(\Omega) \subseteq L^\infty(\Omega) \subseteq B(L^2(\Omega,\nu))] \tag{2.38}$$

Theorem 2.12. [Gelfand theorem (*cf.* [67])] Consider a general basic structure:

$$[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]$$

where it is assumed that $\mathcal{A}$ is commutative. Then, there exists a measure space $(\Omega, \mathcal{B}_\Omega, \nu)$ (where Ω is a locally compact space) such that

$$\mathcal{A} = C_0(\Omega), \quad \overline{\mathcal{A}} = L^\infty(\Omega,\nu), \quad B(H) = B(L^2(\Omega,\nu))$$

where Ω is called a spectrum.

2.3.2 Classical basic structure$[C_0(\Omega) \subseteq L^\infty(\Omega,\nu) \subseteq B(L^2(\Omega,\nu))]$ and State space

Consider the classical basic structure $[C_0(\Omega) \subseteq L^\infty(\Omega,\nu) \subseteq B(L^2(\Omega,\nu))]$. Then, we see the following diagram:

(D): Classical basic structure and State space

$$\begin{array}{ccccccc}
\underset{\substack{(\approx\Omega) \\ C^*\text{-pure state}}}{\mathcal{M}_{+1}^p(\Omega)} \subset \underset{\substack{\text{(probability measure)} \\ C^*\text{-mixed state}}}{\mathcal{M}_{+1}(\Omega)} \subset & \mathcal{M}(\Omega) & & & & & \\
& \Big\uparrow \text{dual} & & & & & \\
& \boxed{C_0(\Omega)} & \xrightarrow[\substack{\text{subalgebra} \\ \text{weak-closure}}]{\subseteq} & \boxed{L^\infty(\Omega)} & \xrightarrow[\text{subalgebra}]{\subseteq} & \boxed{B(L^2(\Omega))} & \\
& & & \Big\downarrow \text{pre-dual} & & & \\
& & \underset{\substack{\text{(probability density function)} \\ W^*\text{-mixed state}}}{L_{+1}^1(\Omega,\nu)} & \subset L^1(\Omega,\nu) & & &
\end{array} \tag{2.39}$$

In the above, the mixed state space $\mathfrak{S}^m(C_0(\Omega)^*)$ is characterized as

$$\begin{aligned}\mathfrak{S}^m(C_0(\Omega)^*) =&\{\rho \in \mathcal{M}(\Omega) \ : \ \rho \geq 0, ||\rho||_{\mathcal{M}(\Omega)} = 1\} \\ =&\{\rho \in \mathcal{M}(\Omega) \ : \ \rho \text{ is a probability measure on } \Omega \ \} \\ =:&\mathcal{M}_{+1}(\Omega)\end{aligned} \tag{2.40}$$

Also, the pure state space $\mathfrak{S}^p(C_0(\Omega)^*)$ is

$$\begin{aligned}&\mathfrak{S}^p(C_0(\Omega)^*) \\ =&\{\rho = \delta_{\omega_0} \in \mathfrak{S}^p(C_0(\Omega)^*) \ : \ \delta_{\omega_0} \text{ is the point measure at } \omega_0(\in \Omega), \omega_0 \in \Omega\} \\ \equiv&\mathcal{M}_{+1}^p(\Omega)\end{aligned} \tag{2.41}$$

Here, the **point measure** $\delta_{\omega_0} \in \mathcal{M}(\Omega)$ is defined by

$$\int_\Omega f(\omega)\delta_{\omega_0}(d\omega) = f(\omega_0) \quad (\forall f \in C_0(\Omega))$$

Therefore,

$$\mathcal{M}_{+1}^p(\Omega) = \mathfrak{S}^p(C_0(\Omega)^*) \ni \delta_\omega \underset{\text{identification}}{\longleftrightarrow} \omega \in \Omega \tag{2.42}$$

Under this identification, we consider that

$$\mathfrak{S}^p(C_0(\Omega)^*) = \Omega$$

Also, it is well known that

$$L^1(\Omega, \nu)^* = L^\infty(\Omega, \nu)$$

Therefore, the W^*-mixed state space is characterized by

$$\begin{aligned}L_{+1}^1(\Omega, \nu) =& \{f \in L^1(\Omega, \nu) \ : \ f \geq 0, \ \int_\Omega f(\omega)\nu(d\omega) = 1\} \\ =& \text{ the set of all probability density functions on } \Omega\end{aligned} \tag{2.43}$$

Remark 2.13. [The case that Ω is finite: $C_0(\Omega) = L^\infty(\Omega, \nu)$, $\mathcal{M}(\Omega) = L^1(\Omega, \nu)$] Let Ω be a finite set $\{\omega_1, \omega_2, ..., \omega_n\}$ with the discrete metric d_D and the counting measure ν. Here, the counting measure ν is defined by

$$\nu(D) = \sharp[D](= \text{"the number of the elements of } D\text{"})$$

Then, we see that

$$C_0(\Omega) = \{F : \Omega \to \mathbb{C} \mid \ F \text{ is a complex valued function on } \Omega\} = L^\infty(\Omega,\nu)$$

And thus, we see that

$$\rho \in \mathcal{M}_{+1}(\Omega) \iff \rho = \sum_{k=1}^n p_k \delta_{\omega_k} \ \ (\sum_{k=1}^n p_k = 1, p_k \geq 0)$$

and

$$f \in L^1_{+1}(\Omega,\nu) \iff \sum_{k=1}^n f(\omega_k) = 1. \ \ f(\omega_k) \geq 0$$

In this sense, we have the following identifications:

$$\mathcal{M}_{+1}(\Omega) = L^1_{+1}(\Omega,\nu) \quad (\text{ or, } \mathcal{M}(\Omega) = L^1(\Omega,\nu))$$

After all, we have the following identification:

$$C_0(\Omega) = L^\infty(\Omega) = \mathbb{C}^n \qquad \mathcal{M}(\Omega) = L^1(\Omega) = \mathbb{C}^n \tag{2.44}$$

where the norm $\|\cdot\|_{C_0(\Omega)}$ in the former is defined by

$$\|z\|_{C_0(\Omega)} = \max_{k=1,2,\ldots,n} |z_k| \qquad \forall z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{C}^n \tag{2.45}$$

and the norm $\|\cdot\|_{\mathcal{M}(\Omega)}$ in the latter is defined by

$$\|z\|_{\mathcal{M}(\Omega)} = \sum_{k=1}^n |z_k| \qquad \forall z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{C}^n \tag{2.46}$$

2.4 State and Observable—the primary quality and the secondary quality—

2.4.1 In the beginning

Our present purpose is to learn the following spell (= Axiom 1) by rote.

(A): Axiom 1(pure measurement)(*cf.* This will be able to be read in §2.7)

With any system S, a basic structure $[\mathcal{A} \subseteq \overline{\mathcal{A}}]_{B(H)}$ can be associated in which measurement theory of that system can be formulated. In $[\mathcal{A} \subseteq \overline{\mathcal{A}}]_{B(H)}$, consider a **$W^*$-measurement** $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}{=}(X, \mathcal{F}, F), S_{[\rho]})$ $\Big($ or, **C^*-measurement** $\mathsf{M}_{\mathcal{A}}(\mathsf{O}{=}(X, \mathcal{F}, F), S_{[\rho]})$ $\Big)$. That is, consider

- a W^*-measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}, S_{[\rho]})$ $\Big($ or, C^*-measurement $\mathsf{M}_{\mathcal{A}}(\mathsf{O}{=}(X, \mathcal{F}, F), S_{[\rho]})$ $\Big)$ of an **observable** $\mathsf{O}{=}(X, \mathcal{F}, F)$ for a **state** $\rho(\in \mathfrak{S}^p(\mathcal{A}^*)$: state space)

Then, the probability that a measured value x $(\in X)$ obtained by the W^*-measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}, S_{[\rho]})$ $\Big($ or, C^*-measurement $\mathsf{M}_{\mathcal{A}}(\mathsf{O}{=}(X, \mathcal{F}, F), S_{[\rho]})$ $\Big)$ belongs to Ξ $(\in \mathcal{F})$ is given by

$$\rho(F(\Xi))(\equiv {}_{\mathcal{A}^*}(\rho, F(\Xi))_{\overline{\mathcal{A}}})$$

(if $F(\Xi)$ is essentially continuous at ρ, or see Definition 2.14).

The "learning by rote" urges us to understand the mathematical definitions of

($\sharp_1$) Basic structure$[\mathcal{A} \subseteq \overline{\mathcal{A}}]_{B(H)}$, state space $\mathfrak{S}^p(\mathcal{A}^*)$

($\sharp_2$) observable $\mathsf{O}{=}(X, \mathcal{F}, F)$, etc.

In the previous section, we studied the above ($\sharp_1$), that is, we discussed the following classification:

(B) General basic structure$\underset{\text{state space } [\mathfrak{S}^p(\mathcal{A}^*), \mathfrak{S}^m(\mathcal{A}^*), \overline{\mathfrak{S}}^p(\overline{\mathcal{A}}_*)]}{[\mathcal{A} \subseteq \overline{\mathcal{A}}]_{B(H)}}$

$$\Longrightarrow \begin{cases} \text{Quantum basic structure}\underset{\text{state space } [\mathfrak{S}^p(\mathcal{T}r(H)), \mathfrak{S}^m(\mathcal{T}r(H)) = \overline{\mathfrak{S}}^m(\mathcal{T}r(H))]}{[\mathcal{C}(H) \subseteq B(H)]_{B(H)}} \\ \\ \text{Classical basic structure}\underset{\text{state space } [\Omega, \mathcal{M}_{+1}(\Omega), L^\infty(\Omega,\nu)]}{[C_0(\Omega) \subseteq L^\infty(\Omega, \nu)]_{B(L^2(\Omega,\nu))}} \end{cases}$$

In this section, we shall study the above ($\sharp_2$), i.e.,

"Observable"

Recall the famous words: **"the primary quality"** and **"the secondary quality"** due to John Locke, an English philosopher and physician regarded as one of the most influential of Enlightenment thinkers and known as the "Father of Classical Liberalism". We think the following correspondence:

$$\begin{cases} \text{[state]} & \longleftrightarrow \text{[the primary quality]} \\ \text{[observable]} & \longleftrightarrow \text{[the secondary quality]} \end{cases} \tag{2.47}$$

And thus, we think

- These (i.e., "state" and "observable") are the concepts which form the basis of dualism.

Also, the following table (which may include my fiction) promotes the better understanding of quantum language as well as the other world-views(i.e., the conventional philosophies).

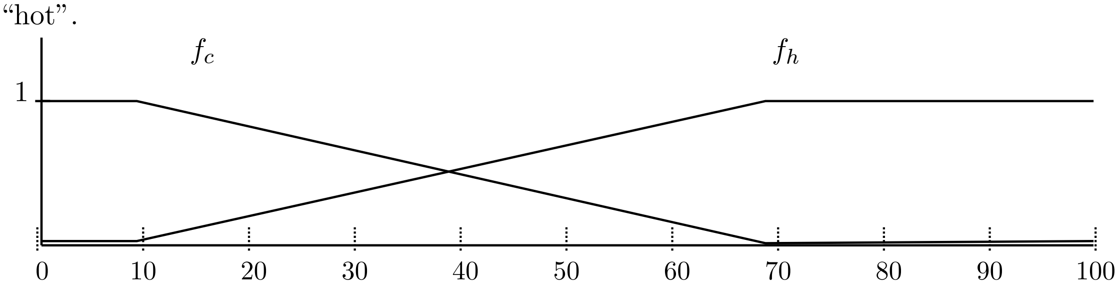
Table 2.1: Observable · State · System in world-views (*cf.* Table 3.1)

World description╲Quantum language	observable	state	system
Plato	idea	/	/
Aristotle	/	eidos	hyle
Locke	secondary quality	primary quality	/
Newton	/	state	point mass
statistics	/	parameter	population
quantum mechanics	observable	state(≈ wave function)	particle

♠**Note 2.2.** It may be understandable to consider

$$\text{"observable"} = \text{"the partition of word"} = \text{"the secondary quality"} \tag{2.48}$$

For example, Chapter 1 (Figure 1.2) says that (f_c, f_h) is the partition between "cold" and "hot".



Chapter 1 (Figure 1.2): Cold or hot?

Also, "measuring instrument" is the instrument that choose a word among words. In this sense, we consider that "observable"= "measurement instrument". Also, The reason that John Locke's sayings *"primary quality* (e.g., length, weight, etc.)*"* and *"secondary quality* (e.g., sweet, dark, cold, etc.)*"* is that these words form the basis of dualism.

2.4.2 Dualism (in philosophy) and duality (in mathematics)

The following question may be significant:

(C_1) Why did philosophers continue persisting in dualism?

As the typical answer, we may consider that

(C_2) "I" is the special existence, and thus, we would like to draw a line between "I" and "matter".

But, we think that this is only quibbling. We want to connect the question (C_1) with the following mathematical question:

(C_3) Why do mathematicians investigate "dual space"?

Of course, the question "why?" is non-sense in mathematics. If we have to answer this, we have no answer except the following (D):

(D) If we consider the dual space $\mathcal{A}^*$, calculation progresses deeply.

Thus, we want to consider the relation between the dualism and the dual space such as

$$
\begin{cases}
\text{[the primary quality]} & \longleftrightarrow \text{the state in the dual space } \mathcal{A}^* \\
\text{[the secondary quality]} & \longleftrightarrow \text{the observable in } C^* \text{ algebra } \mathcal{A} \text{ (or, } W^*\text{-algebra } \overline{\mathcal{A}})
\end{cases}
\tag{2.49}
$$

Thus, we consider that the answer to the (C_1) is also "calculation progresses deeply".

2.4.3 Essentially continuous

In §2.1.2, we introduced the following diagram:

(E):General basic structure and state space

$$
\begin{tikzcd}
\underset{C^*-purestate}{\mathfrak{S}^p(\mathcal{A}^*)} \subset \underset{C^*\text{-mixed state}}{\mathfrak{S}^m(\mathcal{A}^*)} \subset \mathcal{A}^* & & & \\
\boxed{\mathcal{A}} \arrow[u, "\text{dual}"] \arrow[r, "\subseteq", "\text{subalgebra}\cdot\text{weak-closure}"'] & \boxed{\overline{\mathcal{A}}} \arrow[d, "\text{pre-dual}"] \arrow[r, "\subseteq", "\text{subalgebra}"'] & \boxed{B(H)} \\
\underset{W^*\text{-mixed state}}{\overline{\mathfrak{S}}^m(\overline{\mathcal{A}}_*)} \subset \overline{\mathcal{A}}_* & &
\end{tikzcd}
\tag{2.50}
$$

In the above diagram, we introduce the following definition.

Definition 2.14. [Essentially continuous (*cf.* ref. [30])] An element $F(\in \overline{\mathcal{A}})$ is said to be **essentially continuous** at $\rho_0(\in \mathfrak{S}^m(\mathcal{A}^*))$, if there uniquely exists a complex number α such that

(F$_1$) if ρ_n $(\in \overline{\mathfrak{S}}^m(\overline{\mathcal{A}}_*))$ weakly converges to $\rho_0(\in \mathfrak{S}^m(\mathcal{A}^*))$ (That is, $\lim_{n\to\infty} {}_{\overline{\mathcal{A}}_*}\Big(\rho_n, G\Big)_{\mathcal{A}} = {}_{\mathcal{A}^*}\Big(\rho_0, G\Big)_{\mathcal{A}}$ $(\forall G \in \mathcal{A}(\subseteq \overline{\mathcal{A}})$), then $\lim_{n\to\infty} {}_{\overline{\mathcal{A}}_*}\Big(\rho_n, F\Big)_{\overline{\mathcal{A}}} = \alpha$

Then, the value $\rho_0(F)$ $(= {}_{\mathcal{A}^*}\Big(\rho_0, F\Big)_{\overline{\mathcal{A}}})$ is defined by the α
Of course, for any $\rho_0(\in \mathfrak{S}^m(\mathcal{A}^*))$, $F(\in \mathcal{A})$ is essentially continuous at ρ_0.
This "essentially continuous" is chiefly used in th case that $\rho_0(\in \mathfrak{S}^p(\mathcal{A}^*))$.

Remark 2.15. [Essentially continuous in quantum system and classical system]
[I]: Consider the quantum basic structure $[\mathcal{C}(H) \subseteq B(H)]_{B(H)}$. Then, we see

$$(\mathcal{C}(H))^* = \mathcal{T}(H) = B(H)_*$$

Thus, we have $\rho \in \mathfrak{S}^p(\mathcal{C}(H)^*) \subseteq \mathcal{T}r(H)$, $F \in \overline{\mathcal{C}(H)} = B(H)$, which implies that

$$\rho(G) = {}_{\mathcal{C}(H)^*}\Big(\rho, F\Big)_{B(H)} = {}_{\mathcal{T}r(H)}\Big(\rho, F\Big)_{B(H)} \tag{2.51}$$

Thus, we see that "essentially continuous" ⇔ "continuous" in quantum case.
[II]: Next, consider the classical basic structure $[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))]$. A function F $(\in L^\infty(\Omega, \nu))$ is essentially continuous at ω_0 $(\in \Omega = \mathfrak{S}^p(C_0(\Omega)^*))$, if and only if it holds that

(F$_2$) if $\rho_n(\in L^1_{+1}(\Omega, \nu)$ satisfies that

$$\lim_{n\to\infty} \int_\Omega G(\omega)\rho_n(\omega)\nu(d\omega) = G(\omega_0) \qquad (\forall G \in C_0(\Omega))$$

then there uniquely exists a complex number α such that

$$\lim_{n\to\infty} \int_\Omega F(\omega)\rho_n(\omega)\nu(d\omega) = \alpha \tag{2.52}$$

Then, the value of $F(\omega)$ is defined by α, that is, $F(\omega_0) = \alpha$.

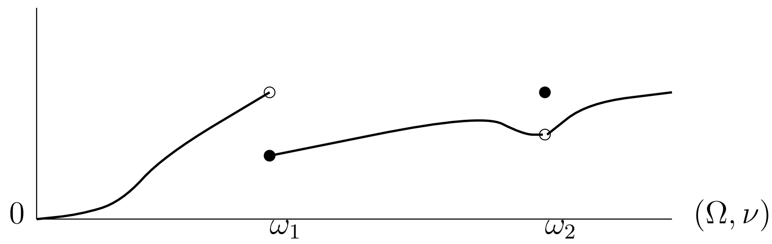


Figure 2.1: not essentially continuous at ω_1, essentially continuous at ω_2

2.4.4 The definition of "observable (=measuring instrument)"

In this section, we introduce "observable", which is also said to be "measuring instrument" or "POVM (=positive operator valued measure space)".

Definition 2.16. [Set ring, set field, σ-field] Let X be a set (or locally compact space). The $\mathcal{F}\Big(\subseteq 2^X = \mathcal{P}(X) = \{A \mid A \subseteq X\}$, the power set of $X\Big)$ (or, the pair $(X, \mathcal{F})$) is called a **ring (of sets)**, if it satisfies that

$$
\begin{aligned}
&(\mathrm{a}) : \emptyset(=\text{"empty set"}) \in \mathcal{F}, \\
&(\mathrm{b}) : \Xi_i \in \mathcal{F} \quad (i = 1, 2, \ldots) \Longrightarrow \bigcup_{i=1}^{n} \Xi_i \in \mathcal{F}, \quad \bigcap_{i=1}^{n} \Xi_i \in \mathcal{F} \\
&(\mathrm{c}) : \Xi_1, \Xi_2 \in \mathcal{F} \Longrightarrow \Xi_1 \setminus \Xi_2 \in \mathcal{F} \quad (\text{ where, } \Xi_1 \setminus \Xi_2 = \{x \mid x \in \Xi_1, x \notin \Xi_2\})
\end{aligned}
$$

Also, if $X \in \mathcal{F}$ holds, the ring $\mathcal{F}$(or, the pair $(X, \mathcal{F})$) is called a **field (of sets)**.
And further,

(d) if the formula (b) holds in the case that $n = \infty$, a field $\mathcal{F}$ is said to be **σ-field**. And the pair $(X, \mathcal{F})$ is called a **measurable space**.

The following definition is most important. In this note, we mainly devote ourselves to the W^*-observable.

Definition 2.17. [Observable,measured value space] Consider the basic structure

$$[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]$$

(G_1):C^*- observable

A triplet $\mathsf{O}=(X, \mathcal{R}, F)$ is called a **C^*-observable (or, C^*-measuring instrument)** in $\mathcal{A}$, if it satisfies as follows.

(i) $(X, \mathcal{R})$ is a ring of sets.

(ii) a map $F : \mathcal{R} \to \mathcal{A}$ satisfies that

(a) $0 \leqq F(\Xi) \leq I \quad (\forall \Xi \in \mathcal{R}),\ F(\emptyset) = 0,$

(b) for any $\rho(\in \mathfrak{S}^p(\mathcal{A}^*))$, there exists a probability space $(X, \overline{\mathcal{R}}, P_\rho)$ such that (where, $\overline{\mathcal{R}}$ is the smallest σ-field such that $\mathcal{R} \subseteq \overline{\mathcal{R}}$) such that

$$ {}_{\mathcal{A}_*}\Big(\rho, F(\Xi)\Big)_{\mathcal{A}} = P_\rho(\Xi) \qquad (\forall \Xi \in \mathcal{R}) \tag{2.53} $$

Also, X [resp. $(X, \mathcal{F}, P_\rho)$] is called a **measured value space** [resp. **sample probability space**].

(G$_2$):W^*- observable

A triplet O=$(X, \mathcal{F}, F)$ is called a **W^*-observable (or, W^*-measuring instrument)** in $\overline{\mathcal{A}}$, if it satisfies as follows.

(i) $(X, \mathcal{F})$ is a σ-field.

(ii) a map $F : \mathcal{F} \to \overline{\mathcal{A}}$ satisfies that

(a) $0 \leqq F(\Xi) \quad (\forall \Xi \in \mathcal{F}),\ F(\emptyset) = 0,\ F(X) = I$

(b) for any $\overline{\rho}(\in \overline{\mathfrak{S}}^m(\overline{\mathcal{A}}_*))$, there exists a probability space $(X, \mathcal{F}, P_{\overline{\rho}})$ such that

$$ {}_{\overline{\mathcal{A}}_*}\Big(\overline{\rho}, F(\Xi)\Big)_{\overline{\mathcal{A}}} = P_{\overline{\rho}}(\Xi) \qquad (\forall \Xi \in \mathcal{F}) \tag{2.54} $$

The observable O=$(X, \mathcal{F}, F)$ is called a **projective observable**, if it holds that

$$ F(\Xi)^2 = F(\Xi) \qquad (\forall \Xi \in \mathcal{F}). $$

In this note, we aways assume Hypothesis 2.19 below:

Definition 2.18. Let $\rho \in \mathfrak{S}^m(\mathcal{A}^*)$, and $(X, \mathcal{F}, F)$ be a W^*-observable in $\overline{\mathcal{A}}$. $\mathcal{F}_\rho = \{\Xi \in \mathcal{F} \mid F(\Xi)$ is essentially continuous at ρ $\}$. The probability space $(X, \mathcal{F}, P_\rho)$ is called its sample probability space, if it holds that

($\sharp_1$) $\mathcal{F}$ is the smallest σ-field that contains $\mathcal{F}_\rho$.

($\sharp_2$)

$$ {}_{\mathcal{A}^*}\Big(\rho, F(\Xi)\Big)_{\overline{\mathcal{A}}} = P_\rho(\Xi) \qquad (\forall \Xi \in \mathcal{F}_\rho) \tag{2.55} $$

Concerning the C^*-observable, the sample probability space clearly exists. On the other hand, concerning the W^*-observable, we have to say something as follows. As mentioned in Remark 2.15, in quantum cases (thus, $\mathcal{A}^* = \mathcal{T}r(H) = \overline{\mathcal{A}}_*$), the ($\sharp_1$) and ($\sharp_2$) clearly hold.

KSTS/RR-16/001
January 8, 2016

However, in the classical cases, we do not know whether the existence of the sample probability space follows from the definition of the W^*-observable. Thus, in this note, we do not add the condition ($\sharp$) in the definition of the W^*-observable.

Hypothesis 2.19. [Sample probability space]. In the above situation, the existence of the sample probability space is always assumed.

2.5 Examples of observables

We shall mention several examples of observables. The observables introduced in Example 2.20-Example 2.23 are characterized as a C^*- observable as well as a W^*- observable.

In what follows (except Example 2.20), consider the classical basic structure:

$$[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))]$$

Example 2.20. [Existence observable] Consider the basic structure:

$$[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]$$

Define the observable $\mathsf{O}^{(\text{exi})} \equiv (X, \{\emptyset, X\}, F^{(\text{exi})})$ in W^*-algebra $\overline{\mathcal{A}}$ such that:

$$F^{(\text{exi})}(\emptyset) \equiv 0, \quad F^{(\text{exi})}(X) \equiv I \tag{2.56}$$

which is called the *existence observable* (or, *null observable*).
Consider any observable $\mathsf{O} = (X, \mathcal{F}, F)$ in $\overline{\mathcal{A}}$. Note that $\{\emptyset, X\} \subseteq \mathcal{F}$. And we see that

$$F(\emptyset) = 0, \quad F(X) = I$$

Thus, we see that $(X, \{\emptyset, X\}, F^{(\text{exi})}) = (X, \{\emptyset, X\}, F)$, and therefore, we say that any observable $\mathsf{O} = (X, \mathcal{F}, F)$ includes the existence observable $\mathsf{O}^{(\text{exi})}$.

♠**Note 2.3.** The above is associated with Berkley's words:

($\sharp_1$) **To be is to be perceived (by George Berkeley(1685-1753))**

which is peculiar to dualism: This is opposite to Einstein's saying in monism :

($\sharp_2$) The moon is there whether one looks at it or not. (i.e., Physics holds without observers.)

in Einstein and Tagore's conversation. (*cf.* Note 12.2)。

Example 2.21. [The resolution of the identity I; The word's partition] Let $[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))]$ be the classical basic structure. We find the similarity between an observable O and *the resolution of the identity* I in what follows. Consider an observable $\mathsf{O} \equiv (X, \mathcal{F}, F)$ in $L^\infty(\Omega)$ such that X is a countable set (i.e., $X \equiv \{x_1, x_2, ...\}$) and $\mathcal{F} = \mathcal{P}(X) = \{\Xi \mid \Xi \subseteq X\}$, i.e., the power set of X. Then, it is clear that

(i) $F(\{x_k\}) \geq 0$ for all $k = 1, 2, ...$

(ii) $\sum_{k=1}^{\infty}[F(\{x_k\})](\omega) = 1 \quad (\forall \omega \in \Omega)$,

which imply that the $[F(\{x_k\}) \ : \ k = 1, 2, ...]$ can be regarded as *the resolution of the identity element I.* Thus we say that

- **An observable O $\big(\equiv (X, \mathcal{F}, F) \big)$ in $L^\infty(\Omega)$ can be regarded as**

" the resolution of the identity I

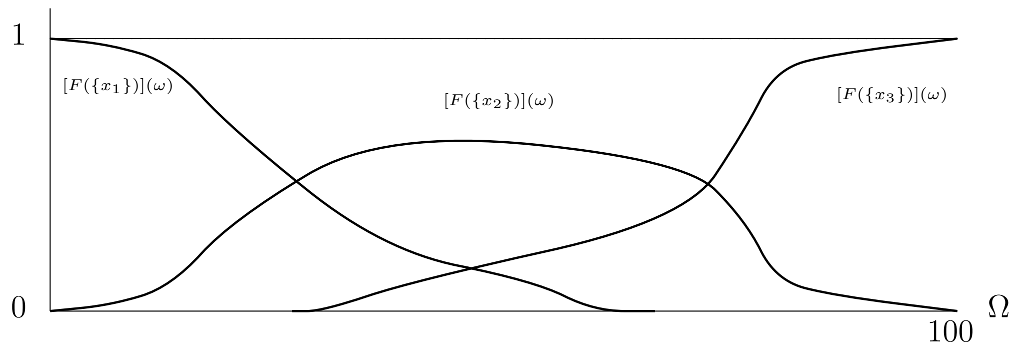


Figure 2.2: $\mathsf{O} \equiv (\{x_1, x_2, x_3\},\ 2^{\{x_1,x_2,x_3\}},\ F)$

In Figure 2.2, assume that $\Omega = [0, 100]$ is the axis of temperatures (°C), and put $X = \{$C(="cold"), L (="lukewarm" = "not hot enough"), H(="hot") $\}$. And further, put $f_{x_1} = f_{\rm C}$, $f_{x_2} = f_{\rm L}$, $f_{x_3} = f_{\rm H}$. Then, the resolution $\{f_{x_1}, f_{x_2}, f_{x_3}\}$ can be regarded as the word's partition C(="cold"), L(="lukewarm"="not hot enough"), H(="hot") .
Also, putting

$$\mathcal{F}(= 2^X) = \{\emptyset, \{x_1\}, \{x_2\}, \{x_3\}, \{x_1, x_2\}, \{x_2, x_3\}, \{x_1, x_3\}, X\}$$

and

$$\begin{aligned}
&[F(\emptyset)](\omega) = 0, \ \ [F(X)](\omega) = f_{x_1}(\omega) + f_{x_2}(\omega) + f_{x_3}(\omega) = 1 \\
&[F(\{x_1\})](\omega) = f_{x_1}(\omega), \ \ [F(\{x_2\})](\omega) = f_{x_2}(\omega), \ \ [F(\{x_3\})](\omega) = f_{x_3}(\omega) \\
&[F(\{x_1, x_2\})](\omega) = f_{x_1}(\omega) + f_{x_2}(\omega), \ \ [F(\{x_2, x_3\})](\omega) = f_{x_2}(\omega) + f_{x_3}(\omega) \\
&[F(\{x_1, x_3\})](\omega) = f_{x_1}(\omega) + f_{x_3}(\omega)
\end{aligned}$$

then, we have the observable $(X, \mathcal{F}(= 2^X), F)$ in $L^\infty([0, 100])$.

Example 2.22. [Triangle observable] Let $[C_0(\Omega) \subseteq L^\infty(\Omega,\nu) \subseteq B(L^2(\Omega,\nu))]$ be the classical basic structure. For example, define the state space Ω by the closed interval $[0,100]$ $(\subseteq \mathbb{R})$. For each $n \in \mathbb{N}_{10}^{100} = \{0, 10, 20, \ldots, 100\}$, define the (triangle) continuous function $g_n : \Omega \to \mathbb{R}$ by

$$g_n(\omega) = \begin{cases} 0 & (0 \leqq \omega \leqq n-10) \\ \dfrac{\omega - n - 10}{10} & (n-10 \leqq \omega \leqq n) \\ -\dfrac{\omega - n + 10}{10} & (n \leqq \omega \leqq n+10) \\ 0 & (n+10 \leqq \omega \leqq 100) \end{cases} \tag{2.57}$$

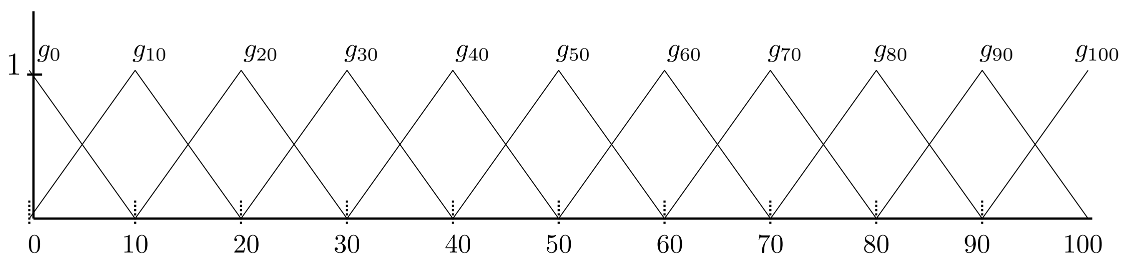


Figure 2.3: Triangle observable

Putting $Y = \mathbb{N}_{10}^{100}$ and define the triangle observable $\mathsf{O}^{\triangle} = (Y, 2^Y, F^{\triangle})$ such that

$$[F^{\triangle}(\emptyset)](\omega) = 0, \qquad [F^{\triangle}(Y)](\omega) = 1$$
$$[F^{\triangle}(\Gamma)](\omega) = \sum_{n \in \Gamma} g_n(\omega) \quad (\forall \Gamma \in 2^{\mathbb{N}_{10}^{100}})$$

Then, we have the triangle observable $\mathsf{O}^{\triangle} = (Y(= \mathbb{N}_{10}^{100}), 2^Y, F^{\triangle})$ in $L^\infty([0,100])$.

Example 2.23. [Normal observable]

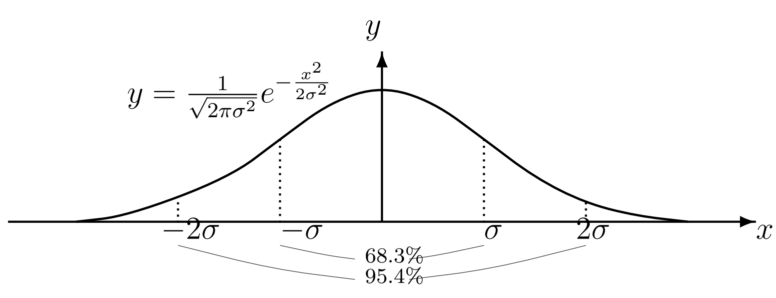


Figure 2.4: Error function

Consider a classical basic structure $[C_0(\Omega) \subseteq L^\infty(\Omega,\nu) \subseteq B(L^2(\Omega,\nu))]$. Here, $\Omega = \mathbb{R}$(= the real line) or, $\Omega =$ interval $[a,b]$ $(\subseteq \mathbb{R})$, which is assumed to have Lebesgue measure $\nu(d\omega)$(=

$d\omega$). Let $\sigma > 0$, which is call a standard deviation. The **normal observable** O_{G_σ}=$(\mathbb{R}, \mathcal{B}_{\mathbb{R}}, G_\sigma)$ in $L^\infty(\Omega,\nu)$ is defined by

$$[G_\sigma(\Xi)](\omega) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_\Xi e^{-\frac{(x-\omega)^2}{2\sigma^2}} dx \quad (\forall \Xi \in \mathcal{B}_{\mathbb{R}}(\text{Borel field}), \forall\omega \in \Omega(= \mathbb{R} \text{ or } [a,b]))$$

This is the most fundamental observable in statistics.

The following examples introduced in Example 2.24 and Example 2.25 are not C^*- observables but W^*- observables. This implies that the W^*-algebraic approach is more powerful than the C^*-algebraic approach. Although the C^*-observable is easy, it is more narrow than the W^*-observable. Thus, throughout this note, we mainly devote ourselves to W^*-algebraic approach.

Example 2.24. [Exact observable] Consider the classical basic structure: $[C_0(\Omega) \subseteq L^\infty(\Omega,\nu) \subseteq B(L^2(\Omega,\nu))]$. Let $\mathcal{B}_\Omega$ be the Borel field in Ω, i.e., the smallest σ-field that contains all open sets. For each $\Xi \in \mathcal{B}_\Omega$, define the definition function $\chi_\Xi : \Omega \to \mathbb{R}$ such that

$$\chi_\Xi(\omega) = \begin{cases} 1 & (\omega \in \Xi) \\ \\ 0 & (\omega \notin \Xi) \end{cases} \tag{2.58}$$

Put $[F^{(\text{exa})}(\Xi)](\omega) = \chi_\Xi(\omega)$ $(\Xi \in \mathcal{B}_\Omega, \omega \in \Omega)$. The triplet $\mathsf{O}^{(\text{exa})} = (\Omega, \mathcal{B}_\Omega, F^{(\text{exa})})$ is called the *exact observable* in $L^\infty(\Omega,\nu)$. This is the W^*-observable and not C^*-observable, since $[F^{(\text{exa})}(\Xi)](\omega)$ is not always continuous. For the argument about the sample probability space (*cf.* Definition 2.18), see Example 2.33.

Example 2.25. [Rounding observable] Define the state space Ω by $\Omega = [0, 100]$. For each $n \in \mathbb{N}_{10}^{100}$=$\{0, 10, 20, \ldots, 100\}$, define the discontinuous function $g_n : \Omega \to [0,1]$ such that

$$g_n(\omega) = \begin{cases} 0 & (0 \leqq \omega \leqq n-5) \\ 1 & (n-5 < \omega \leqq n+5) \\ 0 & (n+5 < \omega \leqq 100) \end{cases}$$

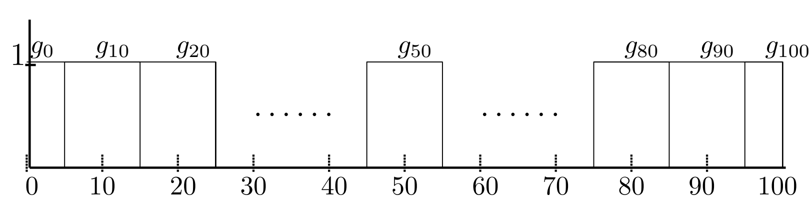


Figure 2.5: Round observable

KSTS/RR-16/001
January 8, 2016

Define the observable $\mathsf{O}_{\text{RND}} = (Y(=\mathbb{N}_{10}^{100}), 2^Y, G_{\text{RND}})$ in $L^\infty(\Omega, \nu)$ such that

$$
\begin{aligned}
&[G_{\text{RND}}(\emptyset)](\omega) = 0, \quad [G_{\text{RND}}(Y)](\omega) = 1 \\
&[G_{\text{RND}}(\Gamma)](\omega) = \sum_{n \in \Gamma} g_n(\omega) \quad (\forall \Gamma \in 2^Y = 2^{\mathbb{N}_{10}^{100}})
\end{aligned}
$$

Recall that g_n is not continuous. Thus, this is not C^*-observable but W^*-observable.

2.6 System quantity — The origin of observable

In classical mechanics, the term "observable" usually means the continuous real valued function on a state space (that is, physical quantity). An observable in measurement theory (= quantum language) is characterized as the natural generalization of the physical quantity. This will be explained in the following examples.

Example 2.26. [System quantity] Let $[C_0(\Omega) \subseteq L^\infty(\Omega,\nu) \subseteq B(L^2(\Omega,\nu))]$ be the classical basic structure. A continuous real valued function $\widetilde{f} : \Omega \to \mathbb{R}$ (or generally, a measurable $\mathbb{R}^n$-valued function $\widetilde{f} : \Omega \to \mathbb{R}^n$) is called a system quantity (or in short, quantity) on Ω. Define the projective observable $\mathsf{O} = (\mathbb{R}, \mathcal{B}_\mathbb{R}, F)$ in $L^\infty(\Omega,\nu)$ such that

$$[F(\Xi)](\omega) = \begin{cases} 1 & \text{when } \omega \in \widetilde{f}^{-1}(\Xi) \\ \\ 0 & \text{when } \omega \notin \widetilde{f}^{-1}(\Xi) \end{cases} \qquad (\forall \Xi \in \mathcal{B}_\mathbb{R})$$

Here, note that

$$\widetilde{f}(\omega) = \lim_{N\to\infty} \sum_{n=-N^2}^{N^2} \frac{n}{N} \Big[F\ \big([\frac{n}{N}, \frac{n+1}{N}) \ \big) \Big] (\omega) = \int_\mathbb{R} \lambda [F(d\lambda)](\omega) \tag{2.59}$$

Thus, we have the following identification:

$$\underset{\text{(system quantity on }\Omega)}{\widetilde{f}} \quad \longleftrightarrow \quad \underset{\text{(projective observable in } L^\infty(\Omega,\nu))}{\mathsf{O} = (\mathbb{R}, \mathcal{B}_\mathbb{R}, F)} \tag{2.60}$$

This O is called the observable representation of a system quantity $\widetilde{f}$. Therefore, we say that

(a) An observable in measurement theory is characterized as the natural generalization of the physical quantity.

Example 2.27. [Position observable , momentum observable , energy observable] Consider Newtonian mechanics in the classical basic algebra $[C_0(\Omega) \subseteq L^\infty(\Omega,\nu) \subseteq B(L^\infty(\Omega,\nu))]$. For simplicity, consider the two dimensional space

$$\Omega = \mathbb{R}_q \times \mathbb{R}_p = \{(q,p) = (\text{position}, \text{momentum}) \mid q,p \in \mathbb{R}\}$$

The following quantities are fundamental:

$$(\sharp_1) : \widetilde{q} : \Omega \to \mathbb{R}, \qquad \widetilde{q}(q,p) = q \quad (\forall (q,p) \in \Omega)$$

$$(\sharp_2)\ :\widetilde{p}:\Omega\to\mathbb{R},\qquad \widetilde{p}(q,p)=p\quad(\forall(q,p)\in\Omega)$$

$$(\sharp_3)\ :\widetilde{e}:\Omega\to\mathbb{R},\qquad \widetilde{e}(q,p)=[\text{potential energy }]+[\text{kinetic energy }] \\ =\underset{\text{(Hamiltonian)}}{U(q)+\frac{p^2}{2m}}\qquad(\forall(q,p)\in\Omega)$$

where, m is the mass of a particle. Under the identification (2.60), the above $(\sharp_1)$, $(\sharp_2)$ and $(\sharp_3)$ is respectively called a position observable, a momentum observable and an energy observable.

Example 2.28. [Hermitian matrix is projective observable] Consider the quantum basic structure in the case that $H=\mathbb{C}^n$, that is,

$$[B(\mathbb{C}^n)\subseteq B(\mathbb{C}^n)\subseteq B(\mathbb{C}^n)]$$

Now, we shall show that an Hermitian matrix $A(\in B(\mathbb{C}^n))$ can be regarded as a projective observable. For simplicity, this is shown in the case that $n=3$. We see (for simplicity, assume that $x_j\neq x_k$(if $j\neq k$))

$$A=U^*\begin{bmatrix}x_1&0&0\\0&x_2&0\\0&0&x_3\end{bmatrix}U \tag{2.61}$$

where U $(\in B(\mathbb{C}^3))$ is the unitary matrix and $x_k\in\mathbb{R}$. Put

$$F_A(\{x_1\})=U^*\begin{bmatrix}1&0&0\\0&0&0\\0&0&0\end{bmatrix}U,\quad F_A(\{x_2\})=U^*\begin{bmatrix}0&0&0\\0&1&0\\0&0&0\end{bmatrix}U,$$
$$F_A(\{x_3\})=U^*\begin{bmatrix}0&0&0\\0&0&0\\0&0&1\end{bmatrix}U\quad F_A(\mathbb{R}\setminus\{x_1,x_2,x_3\})=\begin{bmatrix}0&0&0\\0&0&0\\0&0&0\end{bmatrix},$$

Thus, we get the projective observable $\mathsf{O}_A=(\mathbb{R},\mathcal{B}_\mathbb{R},F_A)$ in $B(\mathbb{C}^3)$. Hence, we have the following identification[2]:

$$\underset{\text{(Hermitian matrix)}}{A}\longleftrightarrow\underset{\text{(projective observable)}}{\mathsf{O}_A=(\mathbb{R},\mathcal{B}_\mathbb{R},F_A)} \tag{2.62}$$

---

[2] For example, in the case that $x_1=x_2$, it suffices to define

$$F_A(\{x_1\})=U^*\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix}U,\ \ F_A(\{x_3\})=U^*\begin{bmatrix}0&0&0\\0&0&0\\0&0&1\end{bmatrix}U\ \ F_A(\mathbb{R}\setminus\{x_1,x_3\})=\begin{bmatrix}0&0&0\\0&0&0\\0&0&1\end{bmatrix}$$

And, we have the projection observable $\mathsf{O}_A=(\mathbb{R},\mathcal{B}_\mathbb{R},F_A)$.

Let $A(\in B(\mathbb{C}^n))$ be an Hermitian matrix. Under this identification, we have the quantum measurement $\mathsf{M}_{B(\mathbb{C}^n)}(\mathsf{O}_A,\, S_{[\rho]})$, where

$$\rho = |\omega\rangle\langle\omega|, \quad \omega = \begin{bmatrix} \omega_1 \\ \omega_2 \\ \vdots \\ \omega_n \end{bmatrix} \in \mathbb{C}^n, \|\omega\| = 1$$

Born's quantum measurement theory (or, Axiom 1 (§2.7)) says that

(♯) The probability that a measured value $x(\in \mathbb{R})$ is obtained by the quantum measurement $\mathsf{M}_{B(\mathbb{C}^n)}(\mathsf{O}_A, S_{[\rho]})$ is given by $\mathrm{Tr}(\rho \cdot F_A(\{x\}))$ ($= \langle \omega, F_A(\{x\})\omega \rangle$).

(for the trace: "Tr", recall Definition 2.9).
Therefore, the expectation of a measured value is given by

$$\int_{\mathbb{R}} x\langle \omega, F_A(dx)\omega\rangle = \langle \omega, A\omega\rangle \tag{2.63}$$

Also, its variance $(\delta_A^\omega)^2$ is given by

$$\begin{aligned}(\delta_A^\omega)^2 &= \int_{\mathbb{R}} (x - \langle \omega, A\omega\rangle)^2 \langle \omega, F_A(dx)\omega\rangle = \langle A\omega, A\omega\rangle - |\langle \omega, A\omega\rangle|^2 \\ &= ||(A - \langle \omega, A\omega\rangle)\omega||^2 \end{aligned} \tag{2.64}$$

Example 2.29. [Spectrum decomposition] Let H be a Hilbert space. Consider the quantum basic structure

$$[\mathcal{C}(H) \subseteq B(H) \subseteq B(H)].$$

The spectral theorem (*cf.* [74]) asserts the following equivalence: ((a)⇔(b)), that is,

(a) T is a self-adjoint operator on Hilbert space H

(b) There exists a projective observable $\mathsf{O} = (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, F)$ in $B(H)$ such that

$$T = \int_{-\infty}^{\infty} \lambda F(d\lambda) \tag{2.65}$$

Since the definition of "unbounded self-adjoint operator" is not easy, in this note we regard the (b) as the definition. In the sense of the (b), we consider the identification:

$$\text{self-adjoint operator } T \underset{\text{identification}}{\longleftrightarrow} \text{spectrum decomposition } \mathsf{O} = (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, F) \tag{2.66}$$

This quantum identification should be compared to the classical identification (2.60).

The above argument can be extended as follows. That is, we have the following equivalence: ((c)⇔(d)), that is,

(c) T_1, T_2 are commutative self-adjoint operators on Hilbert space H

(d) There exists a projective observable $\widehat{\mathsf{O}} = (\mathbb{R}^2, \mathcal{B}_{\mathbb{R}^2}, G)$ in $B(H)$ such that

$$T_1 = \int_{\mathbb{R}^2} \lambda_1 G(d\lambda_1 d\lambda_2), \qquad T_2 = \int_{\mathbb{R}^2} \lambda_2 G(d\lambda_1 d\lambda_2) \tag{2.67}$$

2.7 Axiom 1 — No science without measurement

Measurement theory (= quantum language) is formulated as follows.

$$\bullet \underset{\text{(=quantum language)}}{\boxed{\text{measurement theory}}} := \underbrace{\overset{\text{[Axiom 1]}}{\underset{(cf.\ \S 2.7)}{\boxed{\text{Measurement}}}} + \overset{\text{[Axiom 2]}}{\underset{(cf.\ \S 10.3)}{\boxed{\text{Causality}}}}}_{\text{a kind of spells (a priori judgment)}} + \underbrace{\overset{\text{[quantum linguistic interpretation]}}{\underset{(cf.\ \S 3.1)}{\boxed{\text{Linguistic interpretation}}}}}_{\text{manual to use spells}}$$

Now we can explain Axiom 1 (measurement).

2.7.1 Axiom 1 for measurement

With any system S, a basic structure $[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]$ can be associated in which measurement theory of the system can be formulated. A *state* (or precisely, *pure state*) of the system$S$ is represented by an element of *state space* $\mathfrak{S}^p(\mathcal{A}^*)$. An *observable* (= *measuring instrument*) is represented by a C^*-observable $\mathsf{O} = (X, \mathcal{F}, F)$ in $\mathcal{A}$ (or, W^*-observable $\mathsf{O} = (X, \mathcal{F}, F)$ in $\overline{\mathcal{A}}$).

(A$_1$) An *observer* takes a measurement of an observable $[\mathsf{O}]$ for a state ρ, and gets a measured value $x (\in X)$.

In a basic structure $[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]$, consider a W^**-measurement* $\mathsf{M}_{\overline{\mathcal{A}}}\big(\mathsf{O}{=}(X, \mathcal{F}, F), S_{[\rho]}\big)$ $\Big($ or, C^**-measurement* $\mathsf{M}_{\mathcal{A}}\big(\mathsf{O}{=}(X, \mathcal{F}, F), S_{[\rho]}\big)$ $\Big)$.

Preparation 2.30. Consider

- a W^*-measurement $\mathsf{M}_{\overline{\mathcal{A}}}\big(\mathsf{O}, S_{[\rho]}\big)$ $\Big($ or, C^*-measurement $\mathsf{M}_{\mathcal{A}}\big(\mathsf{O}{=}(X, \mathcal{F}, F), S_{[\rho]}\big)$ $\Big)$ of an *observable* $\mathsf{O}{=}(X, \mathcal{F}, F)$ for a *state* $\rho (\in \mathfrak{S}^p(\mathcal{A}^*)$: state space)

Note that

$$(\mathrm{A}_2) \begin{cases} W^*\text{-}measurement \ \ \mathsf{M}_{\overline{\mathcal{A}}}\big(\mathsf{O}, S_{[\rho]}\big) & \cdots \ \mathsf{O} \text{ is } W^*\text{- observable }, \rho \in \mathfrak{S}^p(\mathcal{A}^*) \\ C^*\text{-}measurement \ \ \mathsf{M}_{\mathcal{A}}\big(\mathsf{O}, S_{[\rho]}\big) & \cdots \ \mathsf{O} \text{ is } C^*\text{- observable }, \rho \in \mathfrak{S}^p(\mathcal{A}^*) \end{cases}$$

In this lecture, we mainly devote ourselves to W^*-measurements.

(B): Axiom 1(measurement) pure type

(This can be read under the preparation to this section)

With any system S, a basic structure $[\mathcal{A} \subseteq \overline{\mathcal{A}}]_{B(H)}$ can be associated in which measurement theory of that system can be formulated. In $[\mathcal{A} \subseteq \overline{\mathcal{A}}]_{B(H)}$, consider a **$W^*$-measurement** $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}{=}(X, \mathcal{F}, F), S_{[\rho]})$ $\Big($ or, **C^*-measurement** $\mathsf{M}_{\mathcal{A}}(\mathsf{O}{=}(X, \mathcal{F}, F), S_{[\rho]})$ $\Big)$. That is, consider

- a W^*-measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}, S_{[\rho]})$ $\Big($ or, C^*-measurement $\mathsf{M}_{\mathcal{A}}(\mathsf{O}{=}(X, \mathcal{F}, F), S_{[\rho]})$ $\Big)$ of an **observable** $\mathsf{O}{=}(X, \mathcal{F}, F)$ for a **state** $\rho (\in \mathfrak{S}^p(\mathcal{A}^*)$: state space)

Then, the probability that a measured value x $(\in X)$ obtained by the W^*-measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}, S_{[\rho]})$ $\Big($ or, C^*-measurement $\mathsf{M}_{\mathcal{A}}(\mathsf{O}{=}(X, \mathcal{F}, F), S_{[\rho]})$ $\Big)$ belongs to Ξ $(\in \mathcal{F})$ is given by

$$\rho(F(\Xi)) (\equiv {}_{\mathcal{A}^*}(\rho, F(\Xi))_{\overline{\mathcal{A}}})$$

(if $F(\Xi)$ is essentially continuous at ρ, or see Definition 2.14).

This axiom is a kind of generalization (or, a linguistic turn) of Born's probabilistic interpretation of quantum mechanics. [3] That is,

(the law proposed by Born)
$\boxed{\text{quantum mechanics (Born's quantum measurement)}}$ $\xrightarrow[\text{linguistic turn}]{}$
(physics)

(a kind of spell)
$\boxed{\text{measurement theory(Axiom 1)}}$ (2.68)
(metaphysics, language)

♠**Note 2.4.** The above axiom is due to Max Born (1926). There are many opinions for the term "probability". For example, Einstein sent Born the following letter (1926):

($\sharp_1$) Quantum mechanics is certainly imposing. But an inner voice tells me that it is not yet the real thing. The theory says a lot, but does not really bring us any closer to the secret of the "old one." I, at any rate, am convinced that He does not throw dice.

From a viewpoint of quantum mechanics, I want to believe that both Born and Einstein are right. That is because I assert that quantum mechanics is not physics.

2.7.2 A simplest example

Now we shall describe Example1.2 (Cold or hot?) in terms of quantum language (i.e., Axiom 1).

[3] Ref. [6]: Born, M. "Zur Quantenmechanik der Stoßprozesse (Vorläufige Mitteilung)", Z. Phys. (37) pp.863–867 (1926).

Example 2.31. [(continued from Example1.2) The measurement of "cold or hot" for water in a cup] Consider the classical basic structure:

$$[C_0(\Omega) \subseteq L^\infty(\Omega,\nu) \subseteq B(L^2(\Omega,\nu))]$$

Here, Ω = the closed interval $[0,100](\subset \mathbb{R})$ with Lebesgue measure ν. The state space $\mathfrak{S}^p(C_0(\Omega)^*)$ is characterized as

$$\mathfrak{S}^p(C_0(\Omega)^*) = \{\delta_\omega \in \mathcal{M}(\Omega) \mid \omega \in \Omega\} \approx \Omega = [0,100]$$

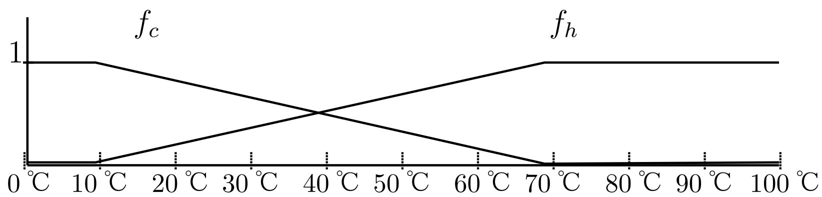


Figure 2.6: Cold? Hot?

In Example 1.2, we consider this [C-H]-thermometer $\mathsf{O} = (f_c, f_h)$, where the state space $\Omega = [0,100]$, the measured value space $X = \{c, h\}$. That is,

$$f_c(\omega) = \begin{cases} 1 & (0 \leqq \omega \leqq 10) \\ \frac{70-\omega}{60} & (10 \leqq \omega \leqq 70) \\ 0 & (70 \leqq \omega \leqq 100) \end{cases}, \qquad f_h(\omega) = 1 - f_c(\omega)$$

Then, we have the (cold-hot) observable $\mathsf{O}_{ch} = (X, 2^X, F_{ch})$ in $L^\infty(\Omega)$ such that

$$[F_{ch}(\emptyset)](\omega) = 0, \qquad [F_{ch}(X)](\omega) = 1$$
$$[F_{ch}(\{c\})](\omega) = f_c(\omega), \qquad [F_{ch}(\{h\})](\omega) = f_h(\omega)$$

Thus, we get a measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{ch}, S_{[\delta_\omega]})$ (or in short, $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{ch}, S_{[\omega]})$. Therefore, for example, putting $\omega = 55$ °C, we can, by Axiom 1 (§2.7), represent the statement (A_1) in Example 1.2 as follows.

(a) the probability that a measured value$x(\in X{=}\{c,h\})$ obtained by measurement

$$\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{ch}, S_{[\omega(=55)]}) \text{ belongs to set } \begin{bmatrix} \emptyset \\ \{c\} \\ \{h\} \\ \{c,h\} \end{bmatrix} \text{ is given by } \begin{bmatrix} [F_{ch}(\emptyset)](55) = 0 \\ [F_{ch}(\{c\})](55) = 0.25 \\ [F_{ch}(\{h\})](55) = 0.75 \\ [F_{ch}(\{c,h\})](55) = 1 \end{bmatrix}$$

Or more precisely,

(b) When an **observer** takes a measurement by $\underset{\textbf{measuring instrument}\mathsf{O}_{ch}=(X,2^X,F_{ch})}{\text{[[C-H]-instrument]}}$

for $\underset{\textbf{(system(measuring object))}}{\text{[water in cup]}}$ with $\underset{\textbf{(state}(=\omega\in\Omega)\textbf{)}}{[55\ °\mathrm{C}]}$, the probability that **measured value**

$$\begin{bmatrix} \mathrm{c} \\ \mathrm{h} \end{bmatrix} \text{ is obtained is given by } \begin{bmatrix} f_c(55) = 0.25 \\ f_h(55) = 0.75 \end{bmatrix}$$

2.8 Classical simple examples (urn problem, etc.)

2.8.1 linguistic world-view — Wonder of man's linguistic competence

The applied scope of physics physics (realistic world-description method) is rather clear. But the applied scope of measurement theory is ambiguous.

What we can do in measurement theory (= quantum language) is

(a) $\begin{cases} (a_1)\text{: Use the language defined by Axiom 1 (§2.7)} \\ \\ (a_2)\text{: Trust in man's linguistic competence} \end{cases}$

Thus, some readers may doubt that

(b) Is it science?

However, it should be noted that the spirit of measurement theory is different from that of physics.

2.8.2 Elementary examples—urn problem, etc.

Since measurement theory is a language, we can not master it without exercise. Thus, we present simple examples in what follows.

Example 2.32. [The measurement of the approximate temperature of water in a cup (continued from Example2.22 [triangle observable])] Consider the classical basic structure:

$$[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))]$$

where $\Omega =$ "the closed interval $[0, 100]$" with the Lebesgue measure ν.
Let testees drink water with various temperature ω °C ($0 \leqq \omega \leqq 100$). And you ask them "How many degrees(°C) is roughly this water?" Gather the data, (for example, $h_n(\omega)$ persons say n °C ($n = 0, 10, 20, \ldots, 90, 100$). and normalize them, that is, get the polygonal lines. For example, define the state space Ω by the closed interval $[0, 100]$ ($\subseteq \mathbb{R}$) with the Lebesgue measure. For each $n \in \mathbb{N}_{10}^{100} = \{0, 10, 20, \ldots, 100\}$, define the (triangle) continuous function $g_n : \Omega \to [0, 1]$ by

$$g_n(\omega) = \begin{cases} 0 & (0 \leqq \omega \leqq n - 10) \\ \dfrac{\omega - n - 10}{10} & (n - 10 \leqq \omega \leqq n) \\ -\dfrac{\omega - n + 10}{10} & (n \leqq \omega \leqq n + 10) \\ 0 & (n + 10 \leqq \omega \leqq 100) \end{cases}$$

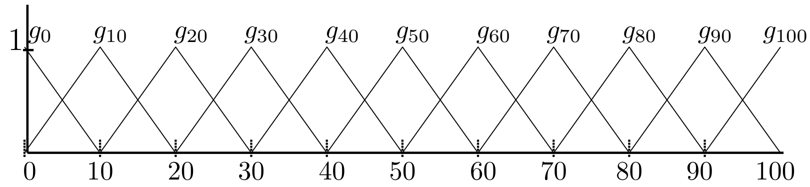


Figure 2.7: Triangle observable

(a) You choose one person from the testees, and you ask him/her "How many degrees(°C) is roughly this water?". Then the probability that he/she says $\begin{bmatrix} \text{"about 40 °C"} \\ \text{"about 50 °C"} \end{bmatrix}$ is given by $\begin{bmatrix} g_{40}(47) = 0.25 \\ f_{50}(47) = 0.75 \end{bmatrix}$

This is described in terms of Axiom 1 (§2.7) in what follows.
Putting $Y = \mathbb{N}_{10}^{100}$, define the triangle observable $\mathsf{O}^{\triangle} = (Y, 2^Y, G^{\triangle})$ in $L^\infty(\Omega)$ such that

$$[G^{\triangle}(\emptyset)](\omega) = 0, \qquad [G^{\triangle}(Y)](\omega) = 1$$
$$[G^{\triangle}(\Gamma)](\omega) = \sum_{n \in \Gamma} g_n(\omega) \qquad (\forall \Gamma \in 2^{\mathbb{N}_{10}^{100}}, \forall \omega \in \Omega = [0, 100])$$

Then, we have the triangle observable $\mathsf{O}^{\triangle} = (Y(= \mathbb{N}_{10}^{100}), 2^Y, G^{\triangle})$ in $L^\infty([0, 100])$. And we get a measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}^{\triangle}, S_{[\delta_\omega]})$. For example, putting ω=47 °C, we see, by Axiom 1 (§2.7), that

(b) the probability that a measured value obtained by the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}^{\triangle}, S_{[\omega(=47)]})$ is $\begin{bmatrix} \text{about 40 °C} \\ \text{about 50 °C} \end{bmatrix}$ is given by $\begin{bmatrix} [G^{\triangle}(\{40\})](47) = 0.3 \\ [G^{\triangle}(\{50\})](47) = 0.7 \end{bmatrix}$

Therefore, we see:

$$\underset{\text{(ordinary language)}}{\boxed{\text{statement (a)}}} \xrightarrow[\text{translation}]{} \underset{\text{(quantum language)}}{\boxed{\text{statement (b)}}} \tag{2.69}$$

///

Example 2.33. [Exact measurement] Consider the classical basic structure:

$$[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))]$$

Let $\mathcal{B}_\Omega$ be the Borel field. Then, define the exact observable $\mathsf{O}^{(\mathrm{exa})} = (X(= \Omega), \mathcal{F}(= \mathcal{B}_\Omega), F^{(\mathrm{exa})})$ in $L^\infty(\Omega, \nu)$ such that

$$[F^{(\mathrm{exa})}(\Xi)](\omega) = \chi_\Xi(\omega) = \begin{cases} 1 & (\omega \in \Xi) \\ \\ 0 & (\omega \notin \Xi) \end{cases} \qquad (\forall \Xi \in \mathcal{B}_\Omega)$$

Let $\delta_{\omega_0} \approx \omega_0 (\in \Omega)$. Consider the exact measurement $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O}^{(\mathrm{exa})}, S_{[\delta_{\omega_0}]})$. Here, Axiom 1 (§2.7) says:

(a) Let $D(\subseteq \Omega)$ be arbitrary open set such that $\omega_0 \in D$. Then, the probability that a measured value obtained by the exact measurement $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O}^{(\mathrm{exa})}, S_{[\delta_{\omega_0}]})$ belongs to D is given by

$$
{}_{C_0(\Omega)^*}\Big(\delta_{\omega_0}, \chi_D\Big)_{L^\infty(\Omega,\nu)} = 1
$$

From the arbitrariness of D, we conclude that

(b) a measured value ω_0 is, with the probability 1, obtained by the exact measurement $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O}^{(\mathrm{exa})}, S_{[\delta_{\omega_0}]})$.

Further, put

$$
\mathcal{F}_{\omega_0} = \{\Xi \in \mathcal{F} \;:\; \omega_0 \notin \text{“the closure of } \Xi\text{”} \setminus \text{“the interior of } \Xi\text{”}\}
$$

Then, when $\Xi \in \mathcal{F}_{\omega_0}$, $F(\Xi)$ is continuous at ω_0. And, $\mathcal{F}$ is the smallest σ-field that contains $\mathcal{F}_{\omega_0}$. Therefore, we have the probability space $(X, \mathcal{F}, P_{\delta_{\omega_0}})$ such that

$$
P_{\delta_{\omega_0}}(\Xi) = [F(\Xi)](\omega_0) \qquad (\forall \Xi \in \mathcal{F}_{\omega_0})
$$

that is,

(c) the exact measurement $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O}^{(\mathrm{exa})}, S_{[\delta_{\omega_0}]})$ has the sample space $(X, \mathcal{F}, P_{\delta_{\omega_0}})$ $(= (\Omega, \mathcal{B}_\Omega, P_{\delta_{\omega_0}}))$

Example 2.34. [Urn problem] There are two urns U_1 and U_2. The urn U_1 [resp. U_2] contains 8 white and 2 black balls [resp. 4 white and 6 black balls] (*cf.* Table 2.2, Figure 2.7).

Table 2.2: urn problem

Urn＼ w·b	white ball	black ball
Urn U_1	8	2
Urn U_2	4	6

Here, consider the following statement (a):

(a) When one ball is picked up from the urn U_2, the probability that the ball is white is 0.4.

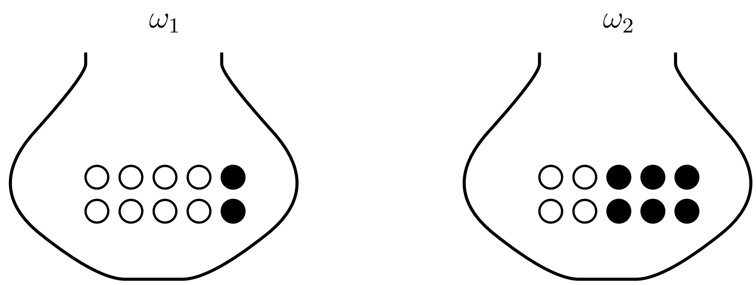


Figure 2.8: Urn problem

In measurement theory, the statement (a) is formulated as follows: Assuming

$$U_1 \quad \cdots \quad \text{“the urn with the state } \omega_1\text{”}$$
$$U_2 \quad \cdots \quad \text{“the urn with the state } \omega_2\text{”}$$

define the state space Ω by $\Omega = \{\omega_1, \omega_2\}$ with the discrete metric and the counting measure ν (i.e., $\nu(\{\omega_1\}) = \nu(\{\omega_2\}) = 1$). That is, we assume the identification;

$$U_1 \approx \omega_1, \quad U_2 \approx \omega_2,$$

Thus, consider the classical basic structure:

$$[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))]$$

Put “w” = “white”, “b” = “black”, and put $X = \{w, b\}$. And define the observable $\mathsf{O}\big(\equiv (X \equiv \{w, b\}, 2^{\{w,b\}}, F)\big)$ in $L^\infty(\Omega)$ by

$$[F(\{w\})](\omega_1) = 0.8, \qquad [F(\{b\})](\omega_1) = 0.2,$$
$$[F(\{w\})](\omega_2) = 0.4, \qquad [F(\{b\})](\omega_2) = 0.6.$$

Thus, we get the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[\delta_{\omega_2}]})$. Here, Axiom 1 (§2.7) says that

(b) the probability that a measured value w is obtained by $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[\delta_{\omega_2}]})$ is given by

$$F(\{b\})(\omega_2) = 0.4$$

Therefore, we see:

$$\underset{\text{(ordinary language)}}{\boxed{\text{statement (a)}}} \xrightarrow[\text{translation}]{} \underset{\text{(quantum language)}}{\boxed{\text{statement (b)}}} \tag{2.70}$$

♠**Note 2.5.** [$L^\infty(\Omega,\nu)$, or in short, $L^\infty(\Omega)$] In the above example, the counting measure ν (i.e., $\nu(\{\omega_1\}) = \nu(\{\omega_2\}) = 1$) is not absolutely indispensable. For example, even if we assume that $\nu(\{\omega_1\}) = 2$ and $\nu(\{\omega_2\}) = 1/3$, we can assert the same conclusion. Thus, in this note,

$L^\infty(\Omega,\nu)$ is often abbreviated to $L^\infty(\Omega)$.

♠**Note 2.6.** The statement (a) in Example 2.34 is not necessarily guaranteed, that is,

> When one ball is picked up from the urn U_2, the probability that the ball is white is 0.4.

is not guaranteed. What we say is that

> the statement (a) in ordinary language should be written by the measurement theoretical statement (b)

It is a matter of course that "probability" can not be derived from mathematics itself. For example, the following ($\sharp_1$) and ($\sharp_2$) are not guaranteed.

($\sharp_1$) From the set $\{1,2,3,4,5\}$, choose one number. Then, the probability that the number is even is given by $2/5$

($\sharp_2$) From the closed interval $[0,1]$, choose one number x. Then, the probability that $x \in [a,b] \subseteq [0,1]$ is given by $|b-a|$

The common sense — "probability" can not be derived from mathematics itself — is well known as Bertrand's paradox (cf. §9.11). Thus, it is usual to add the term "at random" to the above ($\sharp_1$) and ($\sharp_2$). In this note, this term "at random" is usually omitted.

Example 2.35. [Blood type system] The ABO blood group system is the most important blood type system (or blood group system) in human blood transfusion. Let U_1 be the whole Japanese's set and let U_2 be the whole Indian's set. Also, assume that the distribution of the ABO blood group system [O:A:B:AB] concerning Japanese and Indians is determined in (Table 2.3).

Table 2.3: The ratio of the ABO blood group system

J or I╲ABO blood group	O	A	B	AB
Japanese U_1	30%	40%	20%	10%
Indian U_2	30%	20%	40%	10%

Consider the following phenomenon:

(a) Choose one person from the the whole Indian's set U_2 at random. Then the probability that the person's blood type is $\begin{bmatrix} O \\ A \\ B \\ AB \end{bmatrix}$ is given by $\begin{bmatrix} 0.3 \\ 0.2 \\ 0.4 \\ 0.1 \end{bmatrix}$

In what follows, we shall translate the statement (a) described in ordinary language to quantum language. Put $\Omega = \{\omega_1, \omega_2\}$ and consider the discrete metric (Ω, d_D). We get consider the classical basic structure:

$$[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))]$$

Therefore, the pure state space is defined by

$$\mathfrak{S}^p(C_0(\Omega)^*) = \{\delta_{\omega_1}, \delta_{\omega_2}\}$$

Here, consider

$$\begin{aligned} \delta_{\omega_1} \quad & \cdots \quad \text{"the state of the whole Japanese's set } U_1 \text{(i.e., population)"}^4 \\ \delta_{\omega_2} \quad & \cdots \quad \text{"the state of the whole India's set } U_1 \text{(i.e., population)"}, \end{aligned}$$

That is, we consider the following identification: (Therefore, image Figure 2.9):

$$U_1 \approx \delta_{\omega_1}, \qquad U_2 \approx \delta_{\omega_2}$$

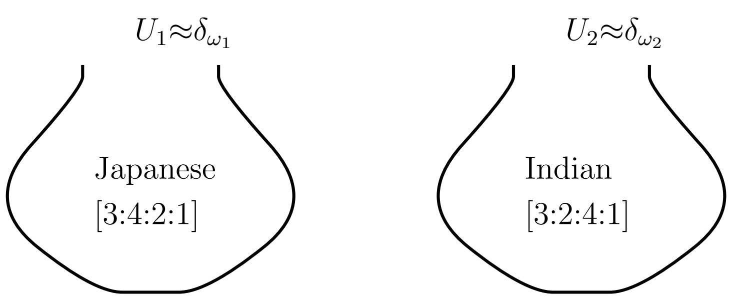


Figure 2.9: Population(=system)≈urn

Define the blood type observable $\mathsf{O}_{\mathrm{BT}} = (\{O, A, B, AB\}, 2^{\{O,A,B,AB\}}, F_{\mathrm{BT}})$ in $L^\infty(\Omega, \nu)$ such that

$$[F_{\mathrm{BT}}(\{O\})](\omega_1) = 0.3, \qquad\qquad [F_{\mathrm{BT}}(\{A\})](\omega_1) = 0.4$$

[4] Note that "population" = "system" (*cf.* Table 2.1).

$$[F_{\mathrm{BT}}(\{B\})](\omega_1)=0.2, \qquad [F_{\mathrm{BT}}(\{AB\})](\omega_1)=0.1 \tag{2.71}$$

and,

$$\begin{aligned}
&[F_{\mathrm{BT}}(\{O\})](\omega_2)=0.3, && [F_{\mathrm{BT}}(\{A\})](\omega_2)=0.2\\
&[F_{\mathrm{BT}}(\{B\})](\omega_2)=0.4, && [F_{\mathrm{BT}}(\{AB\})](\omega_2)=0.1
\end{aligned} \tag{2.72}$$

Thus we get the measurement $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O}_{\mathrm{BT}}, S_{[\delta_{\omega_2}]})$. Hence, the above (a) is translated to the following statement (in terms of quantum language):

(b) The probability that a measured value $\begin{bmatrix} O \\ A \\ B \\ AB \end{bmatrix}$ is obtained by the measurement $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O}_{\mathrm{BT}}, S_{[\delta_{\omega_2}]})$ is given by

$$\begin{bmatrix}
{}_{C_0(\Omega)^*}\Big(\delta_{\omega_2}, F_{\mathrm{BT}}(\{O\})\Big)_{L^\infty(\Omega,\nu)} = [F_{\mathrm{BT}}(\{O\})](\omega_2) = 0.3\\
{}_{C_0(\Omega)^*}\Big(\delta_{\omega_2}, F_{\mathrm{BT}}(\{A\})\Big)_{L^\infty(\Omega,\nu)} = [F_{\mathrm{BT}}(\{A\})](\omega_2) = 0.2\\
{}_{C_0(\Omega)^*}\Big(\delta_{\omega_2}, F_{\mathrm{BT}}(\{B\})\Big)_{L^\infty(\Omega,\nu)} = [F_{\mathrm{BT}}(\{B\})](\omega_2) = 0.4\\
{}_{C_0(\Omega)^*}\Big(\delta_{\omega_2}, F_{\mathrm{BT}}(\{AB\})\Big)_{L^\infty(\Omega,\nu)} = [F_{\mathrm{BT}}(\{AB\})](\omega_2) = 0.1
\end{bmatrix}$$

♠**Note 2.7.** Readers may feel that Example 2.34–Example 2.35 are too easy. However, as mentioned in (a) of Sec. 2.8.1, what we can do is

- $\begin{cases} \text{to be faithful to Axioms} \\ \\ \text{to trust in Man's linguistic competence} \end{cases}$

If some find the other language that is more powerful than quantum language, it will be praised as the greatest discovery in the history of science. That is because this discovery is regarded as beyond the discovery of quantum mechanics.

2.9 Simple quantum examples (Stern=Gerlach experiment)

2.9.1 Stern=Gerlach experiment

Example 2.36. [Quantum measurement(Schtern–Gerlach experiment (1922))]

Assume that we examine the beam (of silver particles(or simply, electrons) after passing through the magnetic field. Then, as seen in the following figure, we see that all particles are deflected either equally upwards or equally downwards in a 50:50 ratio. See Figure 2.10.

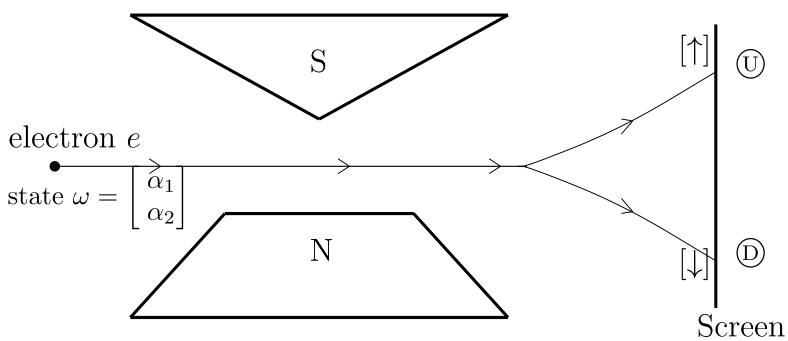


Figure 2.10: Stern–Gerlach experiment (1922)

Consider the two dimensional Hilbert space $H = \mathbb{C}^2$, And therefore, we get the non-commutative basic algebra $B(H)$, that is, the algebra composed of all 2×2 matrices. Thus, we have the quantum basic structure:

$$[\mathcal{C}(H) \subseteq B(H) \subseteq B(H)] = [B(\mathbb{C}^2) \subseteq B(\mathbb{C}^2) \subseteq B(\mathbb{C}^2)]$$

since the dimension of H is finite.

The spin state of an electron P is represented by $\rho(= |\omega\rangle\langle\omega|)$, where $\omega \in \mathbb{C}^2$ such that $\|\omega\| = 1$. Put $\omega = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$ (where, $||\omega||^2 = |\alpha_1|^2 + |\alpha_2|^2 = 1$).

Define $\mathsf{O}_z \equiv (Z, 2^Z, F_z)$, the spin observable concerning the z-axis, such that, $Z = \{\uparrow, \downarrow\}$ and

$$F_z(\{\uparrow\}) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad F_z(\{\downarrow\}) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \tag{2.73}$$

$$F_z(\emptyset) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad F_z(\{\uparrow, \downarrow\}) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Here, Born's quantum measurement theory (the probabilistic interpretation of quantum mechanics) says that

($\sharp$) When a quantum measurement$\mathsf{M}_{B(\mathbb{C}^2)}(\mathsf{O}, S_{[\rho]})$ is taken, the probability that

$$\text{a measured value } \begin{bmatrix} \uparrow \\ \downarrow \end{bmatrix} \text{ is obtained is given by } \begin{bmatrix} \langle \omega, F^z(\{\uparrow\})\omega \rangle = |\alpha_1|^2 \\ \\ \langle \omega, F^z(\{\downarrow\})\omega \rangle = |\alpha_2|^2 \end{bmatrix}$$

That is, putting ω ($= \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$), we says that

When the electron with a spin state state ρ progresses in a magnetic field, the probability that the Geiger counter $\begin{bmatrix} Ⓤ \\ Ⓓ \end{bmatrix}$ sounds

$$\text{is give by } \begin{bmatrix} \begin{bmatrix} \overline{\alpha}_1 & \overline{\alpha}_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = |\alpha_1|^2 \\ \\ \begin{bmatrix} \overline{\alpha}_1 & \overline{\alpha}_2 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = |\alpha_2|^2 \end{bmatrix}$$

Also, we can define $\mathbf{O}^x \equiv (X, 2^X, F^x)$, the spin observable concerning the x-axis, such that, $X = \{\uparrow_x, \downarrow_x\}$ and

$$F^x(\{\uparrow_x\}) = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}, \quad F^x(\{\downarrow_x\}) = \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix}. \tag{2.74}$$

And furthermore, we can define $\mathbf{O}^y \equiv (Y, 2^Y, F^y)$, the spin observable concerning the y-axis, such that, $Y = \{\uparrow_y, \downarrow_y\}$ and

$$F^y(\{\uparrow_y\}) = \begin{bmatrix} 1/2 & i/2 \\ -i/2 & 1/2 \end{bmatrix}, \quad F^y(\{\downarrow_y\}) = \begin{bmatrix} 1/2 & -i/2 \\ i/2 & 1/2 \end{bmatrix}, \tag{2.75}$$

where $i = \sqrt{-1}$.

Here, putting

$$\hat{S}_x = F_x(\{\uparrow\}) - F_x(\{\downarrow\}), \quad \hat{S}_y = F_y(\{\uparrow\}) - F_y(\{\downarrow\}), \quad \hat{S}_z = F_z(\{\uparrow\}) - F_z(\{\downarrow\})$$

we have the following commutation relation:

$$\hat{S}_y\hat{S}_z - \hat{S}_z\hat{S}_y = 2i\hat{S}_x, \quad \hat{S}_z\hat{S}_x - \hat{S}_x\hat{S}_z = 2i\hat{S}_y, \quad \hat{S}_x\hat{S}_y - \hat{S}_y\hat{S}_x = 2i\hat{S}_z \tag{2.76}$$

2.10 de Broglie paradox in $B(\mathbb{C}^2)$

Axiom 1(measurement) includes the paradox (that is, so called de Broglie paradox "there is something faster than light"). In what follows, we shall explain de Broglie paradox in $B(\mathbb{C}^2)$, though the original idea is mentioned in $B(L^2(\mathbb{R}))$ (*cf.* §11.3, and refs.[12, 68]). Also, it should be noted that the argument below is essentially the same as the Stern=Gerlach experiment.

Example 2.37. [de Broglie paradox in $B(\mathbb{C}^2)$] Let H be a two dimensional Hilbert space, i.e., $H = \mathbb{C}^2$. Consider the quantum basic structure:

$$[B(\mathbb{C}^2) \subseteq B(\mathbb{C}^2) \subseteq B(\mathbb{C}^2)]$$

Now consider the situation in the following Figure 2.11.

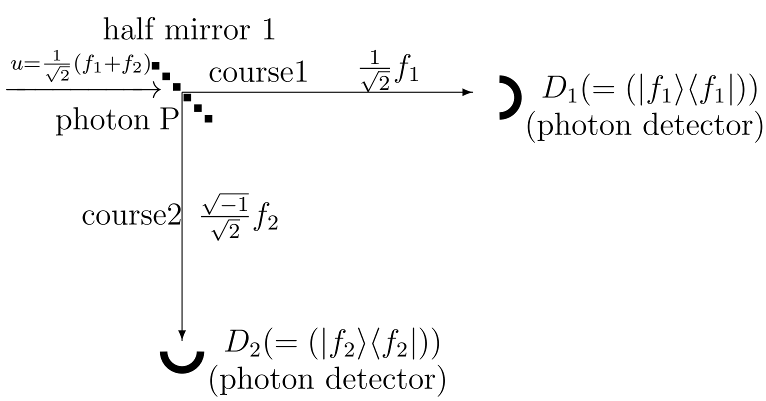


Figure 2.11: $[D_2 + D_1]$ = observable O

Let us explain this figure in what follows. Let $f_1, f_2 \in H$ such that

$$f_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \in \mathbb{C}^2, \qquad f_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \in \mathbb{C}^2$$

Put

$$u = \frac{f_1 + f_2}{\sqrt{2}}$$

Thus, we have the state $\rho = |u\rangle\langle u|$ $(\in \mathfrak{S}^p(B(\mathbb{C}^2)))$.

Let $U(\in B(\mathbb{C}^2))$ be an unitary operator such that

$$U = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\pi/2} \end{bmatrix}$$

and let $\Phi : B(\mathbb{C}^2) \to B(\mathbb{C}^2)$ be the homomorphism such that

$$\Phi(F) = U^* F U \qquad (\forall F \in B(\mathbb{C}^2))$$

Consider the observable $\mathsf{O}_f = (\{1,2\}, 2^{\{1,2\}}, F)$ in $B(\mathbb{C}^2)$ such that

$$F(\{1\}) = |f_1\rangle\langle f_1|, \quad F(\{2\}) = |f_2\rangle\langle f_2|$$

and thus, define the observable $\Phi\mathsf{O}_f = (\{1,2\}, 2^{\{1,2\}}, \Phi F)$ by

$$\Phi F(\Xi) = U^* F(\Xi) U \qquad (\forall \Xi \subseteq \{1,2\})$$

Let us explain Figure 2.11. The photon P with the state $u = \frac{1}{\sqrt{2}}(f_1 + f_2)$ (precisely, $|u\rangle\langle u|$) rushed into the half-mirror 1

(A_1) the f_1 part in u passes through the half-mirror 1, and goes along the course 1 to the photon detector D_1.

(A_2) the f_2 part in u rebounds on the half-mirror 1 (and strictly saying, the f_2 changes to $\sqrt{-1}f_2$, we are not concerned with it), and goes along the course 2 to the photon detector D_2.

Thus, we have the measurement:

$$\mathsf{M}_{B(\mathbb{C}^2)}(\Phi\mathsf{O}_f, S_{[\rho]}) \tag{2.77}$$

And thus, we see:

(B) The probability that a $\begin{bmatrix} \text{measured value } 1 \\ \text{measured value } 2 \end{bmatrix}$ is obtained by the measurement $\mathsf{M}_{B(\mathbb{C}^2)}(\Phi\mathsf{O}_f, S_{[\rho]})$ is given by

$$\begin{bmatrix} \mathrm{Tr}(\rho \cdot \Phi F(\{1\})) \\ \mathrm{Tr}(\rho \cdot \Phi F(\{2\})) \end{bmatrix} = \begin{bmatrix} \langle u, \Phi F(\{1\}) u\rangle \\ \langle u, \Phi F(\{2\}) u\rangle \end{bmatrix} = \begin{bmatrix} \langle Uu, F(\{1\}) Uu\rangle \\ \langle Uu, F(\{2\}) Uu\rangle \end{bmatrix} = \begin{bmatrix} |\langle u, f_1\rangle|^2 \\ |\langle u, f_2\rangle|^2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$$

This is easy, but it is deep in the following sense.

(C) Assume that

Detector D_1 and Detector D_2 are very far.

And assume that the photon P is discovered at the detector D_1. Then, we are troubled if the photon P is also discovered at the detector D_2. Thus, in order to avoid this difficulty, the photon P (discovered at the detector D_1) has to eliminate the wave function $\frac{\sqrt{-1}}{\sqrt{2}} f_2$ in an instant. In this sense, the (B) implies that

there may be something faster than light

KSTS/RR-16/001
January 8, 2016

This is the de Broglie paradox (*cf.* [12, 68]). From the view point of quantum language, we give up to solve the paradox, that is, we declare that

Stop to be bothered!

(Also, see [61]).

♠**Note 2.8.** The de Broglie paradox (i.e., there may be something faster than light) always appears in quantum mechanics. For example, the readers should confirm that it appears in Example 2.36 (Schtern-Gerlach experiment). I think that

- the de Broglie paradox is the only paradox in quantum mechanics

Chapter 3

The linguistic interpretation

Measurement theory (= quantum language) is formulated as follows.

- $\underset{\text{(=quantum language)}}{\boxed{\text{measurement theory}}} := \underbrace{\overset{\text{[Axiom 1]}}{\underset{(cf.\ \S 2.7)}{\boxed{\text{Measurement}}}} + \overset{\text{[Axiom 2]}}{\underset{(cf.\ \S 10.3)}{\boxed{\text{Causality}}}}}_{\text{a kind of spell(a priori judgment)}} + \underbrace{\overset{\text{[quantum linguistic interpretation]}}{\underset{(cf.\ \S 3.1)}{\boxed{\text{Linguistic interpretation}}}}}_{\text{manual to use spells}}$

Measurement theory says that

- Describe every phenomenon modeled on Axioms 1 and 2 (by a hint of the linguistic interpretation)!

Since we dealt with simple examples in the previous chapter, we did not need the linguistic interpretation. In this chapter, we study several more difficult problems with the linguistic interpretation. Also, the linguistic interpretation may be called "the linguistic Copenhagen interpretation" since we believe that it is the true colors of so called Copenhagen interpretation (*cf.* Section 1.1.1).

3.1 The linguistic interpretation

3.1.1 The review of Axiom 1 (measurement: §2.7)

In the previous chapter, we introduced Axiom 1 (measurement) as follows.

(A): Axiom 1(measurement) pure type

(*cf.* It was able to read under the preparation to §2.7))

With any system S, a basic structure $[\mathcal{A} \subseteq \overline{\mathcal{A}}]_{B(H)}$ can be associated in which measurement theory of that system can be formulated. In $[\mathcal{A} \subseteq \overline{\mathcal{A}}]_{B(H)}$, consider a **$W^*$-measurement** $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}{=}(X, \mathcal{F}, F), S_{[\rho]})$ $\Big($ or, **C^*-measurement** $\mathsf{M}_{\mathcal{A}}(\mathsf{O}{=}(X, \mathcal{F}, F), S_{[\rho]})$ $\Big)$. That is, consider

- a W^*-measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}, S_{[\rho]})$ $\Big($ or, C^*-measurement $\mathsf{M}_{\mathcal{A}}(\mathsf{O}{=}(X, \mathcal{F}, F), S_{[\rho]})$ $\Big)$ of an **observable** $\mathsf{O}{=}(X, \mathcal{F}, F)$ for a **state** $\rho (\in \mathfrak{S}^p(\mathcal{A}^*)$: state space)

Then, the probability that a measured value x $(\in X)$ obtained by the W^*-measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}, S_{[\rho]})$ $\Big($ or, C^*-measurement $\mathsf{M}_{\mathcal{A}}(\mathsf{O}{=}(X, \mathcal{F}, F), S_{[\rho]})$ $\Big)$ belongs to Ξ $(\in \mathcal{F})$ is given by

$$\rho(F(\Xi)) (\equiv {}_{\mathcal{A}^*}(\rho, F(\Xi))_{\overline{\mathcal{A}}})$$

(if $F(\Xi)$ is essentially continuous at ρ, or see Definition 2.14).

Here, note that

(B$_1$) **the above axiom is a kind of spell (i.e., incantation, magic words, metaphysical statement), and thus, it is impossible to verify them experimentally.**

In this sense, the above axiom corresponds to "a priori synthetic judgment" in Kant's philosophy (*cf.* [53]). And thus, we say:

(B$_2$) After we learn the spell (= Axiom 1) by rote, we have to exercise and lesson the spell (= Axiom 1). Since quantum language is a language, it may be unable to use well at first.

It will make progress gradually, while applying a trial-and-error method.

However,

(C$_1$) if we would like to make speed of acquisition of a quantum language as quick as possible, we may want the good manual to use the axioms.

Here, we think that

**(C$_2$) the linguistic interpretation
= the manual to use the spells (Axiom 1 and 2)**

3.1.2 Descartes figure (in the linguistic interpretation)

In what follows, let us explain the linguistic interpretation.

The concept of "measurement" can be, for the first time, understood in dualism. Let us explain it. The image of "measurement" is as shown in Figure 3.1.

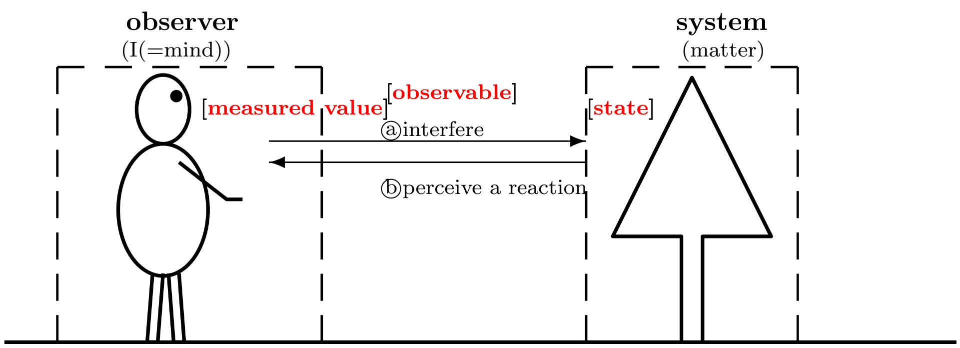


Figure 3.1[**Descartes Figure**]:The image of "measurement(=ⓐ+ⓑ)" in dualism

In the above,

(D_1) ⓐ: it suffices to understand that "interfere" is, for example, "apply light".
ⓑ: perceive the reaction.

That is, "measurement" is characterized as the interaction between "observer" and "measuring object". However,

(D_2) In measurement theory, "interaction" must not be emphasized.

Therefore, in order to avoid confusion, it might better to omit the interaction "ⓐ and ⓑ" in Figure 3.1.

After all, we think that:

(D_3) It is clear that there is no measured value without observer (i.e., brain). Thus, we consider that measurement theory is composed of three key-words:

$$\underset{\text{(observer,brain, mind)}}{\boxed{\text{measured value}}}, \quad \underset{\text{(thermometer, eye, ear, body, polar star (\textit{cf.} Note 3.1 later))}}{\boxed{\text{observable (= measuring instrument)}}}, \quad \underset{\text{(matter)}}{\boxed{\text{state}}}, \tag{3.1}$$

and thus, it might be called "trialism" (and not "dualism"). But, according to the custom, it is called **"dualism"** in this note.

3.1.3 The linguistic interpretation [(E_1)-(E_7)]

The linguistic interpretation is "the manual to use Axiom 1 and 2". Thus, there are various explanations for the linguistic interpretations. However, it is usual to consider that the linguistic interpretation is characterized as the following (E). And the most important is

Only one measurement is permitted

(E):The linguistic interpretation (=quantum language interpretation)

With **Descartes figure 3.1 (and (E_1)-(E_7))** in mind,
describe every phenomenon in terms of Axioms 1 and 2

(E_1) Consider the dualism composed of "observer" and "system(=measuring object)". And therefore, "observer" and "system" must be absolutely separated. If it says for a metaphor, we say "Audience should not be up to the stage".

(E_2) Of course, "matter(=measuring object)" has the space-time. On the other hand, the observer does not have the space-time. Thus, the question: "When and where is a measured value obtained?" is out of measurement theory, Thus, there is no tense in measurement theory. This implies that there is no tense in science.

(E_3) In measurement theory, "interaction" must not be emphasized.

(E_4) **Only one measurement is permitted**. Thus, the state after measurement (or, wave function collapse, the influence of measurement) is meaningless. (*cf.* Projection Postulate 11.6)

(E_5) There is no probability without measurement.

(E_6) State never moves,

and so on.

Also, since our assertion is

quantum language is the final goal of dualistic idealism (="Descartes=Kant philosophy")

(*cf.* ⑧ in Figure 1.1), we have to assert that

(E_7) **Many of maxims of the philosophers (particularly, the dualistic idealism) can be regarded as a part of the linguistic interpretation.**

Some may think that the (E_7) is unbelievable. However,

(F) Since the purpose of philosophies and that of quantum language are the same, that is, the non-realistic world view, it is natural to consider that maims of philosophers $\approx$ the linguistic interpretation

Recall the following figure:

Figure 3.1. [=Figure 1.1:The location of quantum language in the history of world-description]

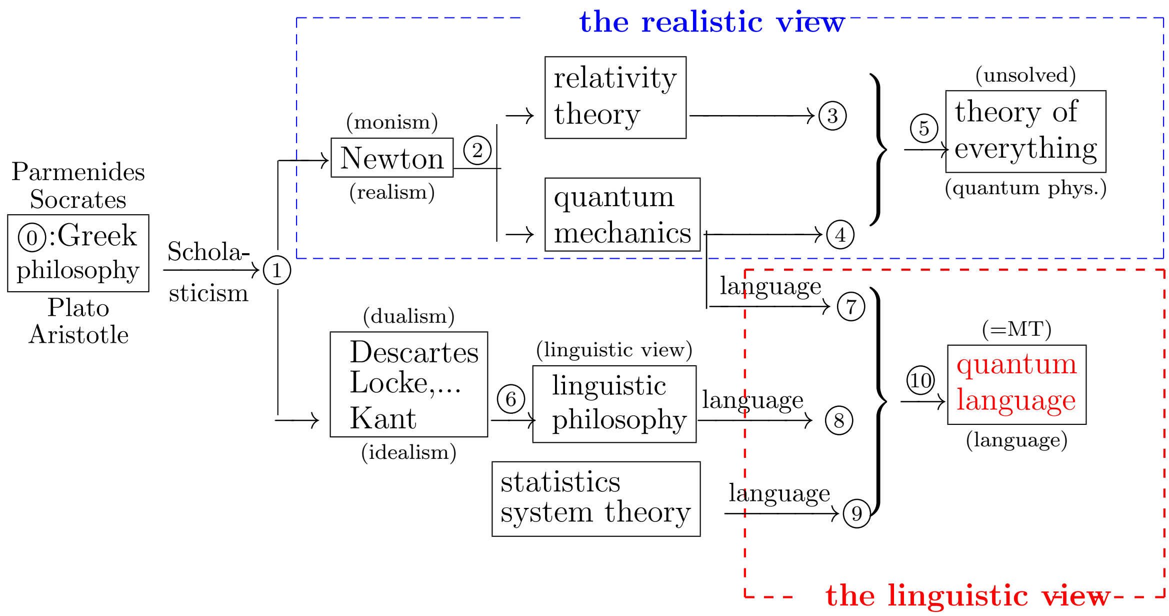


Figure 1.1: The history of the world-view

In the above, we regard

$$[⓪ \longrightarrow ① \longrightarrow ⑥ \longrightarrow ⑧ \longrightarrow ⑩] \tag{3.2}$$

as a genealogy of the dualistic idealism. Talking cynically, we say that

- Philosophers continued investigating "linguistic interpretation" (="how to use Axioms 1 and 2") without Axioms 1 and 2.

For example, "Only one measurement is permitted" and "State never moves" may be related to Parmenides' words;

$$\begin{cases} \textbf{There are no "plurality", but only "one".} \\ \textbf{And therefore, there is no movement.} \end{cases} \tag{3.3}$$

Table 3.1: Trialism (i.e., dualism) in world-views (*cf.* Table 2.1)

Quantum language	measured value	observable	state (system)
Plato	/	idea (*cf.* Note 3.1)	/
Aristotle	/	/	edios (hyle)
Thomas Aquinas	universale post rem	universale ante rem	/ (universale in re)
Descartes	I, mind, brain	body (*cf.* Note 3.1)	/ (matter)
Locke	/	secondary quality	primary quality (/)
Newton	/	/	state (point mass)
statistics	sample space	/	parameter (population)
quantum mechanics	measured value	observable	state (particle)

Thus, we want to assert that Parmenides (born around BC. 515) is the oldest discoverer of the linguistic interpretation. Also, we propose the following table:

♠**Note 3.1.** In the above table, Newtonian mechanics may be the most understandable. We regard "Plato idea" as "absolute standard". And, we want to understand that Newton is similar to Aristotle, since their assertions belong to the realistic world view(*cf.* Figure 1.1). Also, recall the formula (3.1), that is, "observable"="measuring instrument"="body". Thus, as the examples of "observable", we think:

eyes, ears, glasses, telescope, compass, etc.

If "compass" is accepted, "the polar star" should be also accepted as the example of the observable. In the same sense, "the jet stream to an airplane" is a kind of observable (*cf.* Section 8.1 (pp.129-135) in [38]). Also, if it is certain that Descartes is the first discoverer of "I", I have to retract my understanding of Scholasticism in Table 3.1. Although I have no confidence about Scholasticism, the discover of three words ("post rem", "ante rem", "in re") should be remarkable.

3.2 Tensor operator algebra

3.2.1 Tensor product of Hilbert space

The linguistic interpretation (§3.1) says

"Only one measurement is permitted"

which implies "only one measuring object" or "only one state". Thus, if there are several states, these should be regarded as "only one state". In order to do it, we have to prepare "tensor operator algebra". That is,

(A) "several states" $\xrightarrow[\text{by tensor operator algebra}]{\text{combine several into one}}$ **"one state"**

In what follows, we shall introduce the tensor operator algebra.

Let H, K be Hilbert spaces. We shall define the tensor Hilbert space $H \otimes K$ as follows. Let $\{e_m \mid m \in \mathbb{N} \equiv \{1, 2, \ldots\}\}$ be the CONS (i.e, complete orthonormal system) in H. And, let $\{f_n \mid n \in \mathbb{N} \equiv \{1, 2, \ldots\}\}$ be the CONS in K. For each $(m, n) \in \mathbb{N}^2$, consider the symbol "$e_m \otimes f_n$". Here, consider the following "space":

$$H \otimes K = \Big\{ g = \sum_{(m,n) \in \mathbb{N}^2} \alpha_{m,n} e_m \otimes f_n \ \Big| \ ||g||_{H \otimes K} \equiv [\sum_{(m,n) \in \mathbb{N}^2} |\alpha_{m,m}|^2]^{1/2} < \infty \Big\} \tag{3.4}$$

Also, the inner product $\langle \cdot, \cdot \rangle_{H \otimes K}$ is represented by

$$\begin{aligned} \langle e_{m_1} \otimes f_{n_1}, e_{m_2} \otimes f_{n_2} \rangle_{H \otimes K} &\equiv \langle e_{m_1}, e_{m_2} \rangle_H \cdot \langle f_{n_1}, f_{n_2} \rangle_K \\ &= \begin{cases} 1 & (m_1, n_1) = (m_2, n_2) \\ 0 & (m_1, n_1) \neq (m_2, n_2) \end{cases} \end{aligned} \tag{3.5}$$

Thus, summing up, we say

(B) the tensor Hilbert space $H \otimes K$ is defined by the Hilbert space with the CONS $\{e_m \otimes f_n \mid (m, n) \in \mathbb{N}^2\}$.

For example, for any $e = \sum_{m=1}^{\infty} \alpha_m e_m \in H$ and any $f = \sum_{n=1}^{\infty} \beta_n f_m \in H$, the tensor $e \otimes f$ is defined by

$$e \otimes f = \sum_{(m,n) \in \mathbb{N}^2} \alpha_m \beta_n (e_m \otimes f_n)$$

Also, the tensor norm $||\widehat{u}||_{H \otimes K}$ $(\widehat{u} \in H \otimes K)$ is defined by

$$||\widehat{u}||_{H \otimes K} = |\langle \widehat{u}, \widehat{u} \rangle_{H \otimes K}|^{1/2}$$

Example 3.2. [Simple example:tensor Hilbert space $\mathbb{C}^2 \otimes \mathbb{C}^3$] Consider the 2-dimensional Hilbert space $H = \mathbb{C}^2$ and the 3-dimensional Hilbert space $K = \mathbb{C}^3$. Now we shall define the tensor Hilbert space $H \otimes K = \mathbb{C}^2 \otimes \mathbb{C}^3$ as follows.

Consider the CONS $\{e_1, e_2\}$ in H such as

$$e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

And, consider the CONS $\{f_1.f_2, f_3\}$ in K such as

$$f_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad f_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad f_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Therefore, the tensor Hilbert space $H \otimes K = \mathbb{C}^2 \otimes \mathbb{C}^3$ has the CONS such as

$$e_1 \otimes f_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad e_1 \otimes f_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, e_1 \otimes f_3 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$
$$e_2 \otimes f_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad e_2 \otimes f_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, e_2 \otimes f_3 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Thus, we see that

$$H \otimes K = \mathbb{C}^2 \otimes \mathbb{C}^3 = \mathbb{C}^6$$

That is because the CONS $\{e_i \otimes f_j \mid i = 1, 2, 3, \quad j = 1, 2\}$ in $H \otimes K$ can be regarded as $\{g_k \mid k = 1, 2, ..., 6\}$ such that

$$g_1 = e_1 \otimes f_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad g_2 = e_1 \otimes f_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, g_3 = e_1 \otimes f_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$
$$g_4 = e_2 \otimes f_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad g_5 = e_2 \otimes f_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, g_6 = e_2 \otimes f_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

This Example 3.2 can be easily generalized as follows.

Theorem 3.3. [Finite tensor Hilbert space]

$$\mathbb{C}^{m_1} \otimes \mathbb{C}^{m_2} \otimes \cdots \otimes \otimes \mathbb{C}^{m_n} = \mathbb{C}^{\sum_{k=1}^{n} m_k} \tag{3.6}$$

Theorem 3.4. [Concrete tensor Hilbert space]

$$L^2(\Omega_1, \nu_1) \otimes L^2(\Omega_2, \nu_2) = L^2(\Omega_1 \times \Omega_2, \nu_1 \otimes \nu_2) \tag{3.7}$$

where, $\nu_1 \otimes \nu_2$ is the product measure.

Definition 3.5. [Infinite tensor Hilbert space] Let $H_1, H_2, ..., H_k, ...$ be Hilbert spaces. Then, the infinite tensor Hilbert space $\bigotimes_{k=1}^{\infty} H_k$ can be defined as follows. For each $k(\in \mathbb{N})$, consider the CONS $\{e_k^j\}_{j=1}^{\infty}$ in a Hilbert space H_k. For any map $b : \mathbb{N} \to \mathbb{N}$, define the symbol $\bigotimes_{k=1}^{\infty} e_k^{b(k)}$ such that

$$\bigotimes_{k=1}^{\infty} e_k^{b(k)} = e_1^{b(1)} \otimes e_2^{b(2)} \otimes e_3^{b(3)} \otimes \cdots$$

Then, we have:

$$\Big\{ \bigotimes_{k=1}^{\infty} e_k^{b(k)} \;\Big|\; b : \mathbb{N} \to \mathbb{N} \text{ is a map} \Big\} \tag{3.8}$$

Hence we can define the infinite Hilbert space $\bigotimes_{k=1}^{\infty} H_k$ such that it has the CONS (3.8).

3.2.2 Tensor basic structure

For each continuous linear operators $F \in B(H), G \in B(K)$, the tensor operator $F \otimes G \in B(H \otimes K)$ is defined by

$$(F \otimes G)(e \otimes f) = Fe \otimes Gf \quad (\forall e \in H, f \in K)$$

Definition 3.6. [Tensor C^*-algebra and Tensor W^*-algebra] Consider basic structures

$$[\mathcal{A}_1 \subseteq \overline{\mathcal{A}_1} \subseteq B(H_1)] \text{ and } [\mathcal{A}_2 \subseteq \overline{\mathcal{A}_2} \subseteq B(H_2)]$$

[I]: The tensor C^*-algebra $\mathcal{A}_1 \otimes \mathcal{A}_2$ is defined by the smallest C^*-algebra $\widehat{\mathcal{A}}$ such that

$$\{F \otimes G \ (\in B(H_1 \otimes H_2)) \mid F \in \mathcal{A}_1, \ G \in \mathcal{A}_2\} \subseteq \widehat{\mathcal{A}} \subseteq B(H_1 \otimes H_2)$$

[II]: The tensor W^*-algebra $\overline{\mathcal{A}}_1 \otimes \overline{\mathcal{A}}_2$ is defined by the smallest W^*-algebra $\widetilde{\mathcal{A}}$ such that

$$\{F \otimes G \ (\in B(H_1 \otimes H_2)) \mid F \in \overline{\mathcal{A}}_1, \ G \in \overline{\mathcal{A}}_2\} \subseteq \widetilde{\mathcal{A}} \subseteq B(H_1 \otimes H_2)$$

Here, note that $\overline{\mathcal{A}}_1 \otimes \overline{\mathcal{A}}_2 = \overline{\mathcal{A}_1 \otimes \mathcal{A}_2}$.

Theorem 3.7. [Tensor basic structure] [I]: Consider basic structures

$$[\mathcal{A}_1 \subseteq \overline{\mathcal{A}_1} \subseteq B(H_1)] \text{ and } [\mathcal{A}_2 \subseteq \overline{\mathcal{A}_2} \subseteq B(H_2)]$$

Then, we have the tensor basic structure:

$$[\mathcal{A}_1 \otimes \mathcal{A}_2 \subseteq \overline{\mathcal{A}}_1 \otimes \overline{\mathcal{A}}_2 \subseteq B(H_1 \otimes H_2)]$$

[II]: Consider quantum basic structures $[\mathcal{C}(H_1) \subseteq B(H_1) \subseteq B(H_1)]$ and $[\mathcal{C}(H_2) \subseteq B(H_2) \subseteq B(H_2)]$. Then, we have tensor quantum basic structure:

$$\begin{aligned}&[\mathcal{C}(H_1) \subseteq B(H_1) \subseteq B(H_1)] \otimes [\mathcal{C}(H_2) \subseteq B(H_2) \subseteq B(H_2)]\\ =&[\mathcal{C}(H_1 \otimes H_2) \subseteq B(H_1 \otimes H_2) \subseteq B(H_1 \otimes H_2)]\end{aligned}$$

[III]: Consider classical basic structures $[C_0(\Omega_1) \subseteq L^\infty(\Omega_1, \nu_1) \subseteq B(L^2(\Omega_1, \nu_1))]$ and $[C_0(\Omega_2) \subseteq L^\infty(\Omega_2, \nu_2) \subseteq B(L^2(\Omega_2 \nu_2))]$. Then, we have tensor classical basic structure:

$$[C_0(\Omega_1) \subseteq L^\infty(\Omega_1 \subseteq \nu_1) \subseteq B(L^2(\Omega_1, \nu_1))] \otimes [C_0(\Omega_2) \subseteq L^\infty(\Omega_2 \subseteq \nu_2) \subseteq B(L^2(\Omega_2, \nu_2))]$$

$$=[C_0(\Omega_1 \times \Omega_2) \subseteq L^\infty(\Omega_1 \times \Omega_2, \nu_1 \otimes \nu_2) \subseteq B(L^2(\Omega_1 \times \Omega_2, \nu_1 \otimes \nu_2))]$$

Theorem 3.8. The $\bigotimes_{k=1}^\infty B(H_k)$ $(\subseteq B(\bigotimes_{k=1}^\infty H_k))$ is defined by the smallest C^*-algebra that contains

$$F_1 \otimes F_2 \otimes \cdots \otimes F_n \otimes I \otimes I \otimes \cdots \Big(\in B(\bigotimes_{k=1}^\infty H_k)\Big)$$
$$(\forall F_k \in B(H_k),\ k = 1, 2, ..., n, n = 1, 2, ...)$$

Then, it holds that

$$\bigotimes_{k=1}^\infty B(H_k) = B(\bigotimes_{k=1}^\infty H_k) \tag{3.9}$$

Theorem 3.9. The followings hold:

$$\begin{aligned}
\text{(i)}: &\quad \rho_k \in \mathcal{A}_k^* \Longrightarrow \bigotimes_{k=1}^n \rho_k \in (\bigotimes_{k=1}^n \mathcal{A}_k)^*\\
\text{(ii)}: &\quad \rho_k \in \mathfrak{S}^m(\mathcal{A}_k^*) \Longrightarrow \bigotimes_{k=1}^n \rho_k \in \mathfrak{S}^m((\bigotimes_{k=1}^n \mathcal{A}_k)^*)\\
\text{(iii)}: &\quad \rho_k \in \mathfrak{S}^p(\mathcal{A}_k^*) \Longrightarrow \bigotimes_{k=1}^n \rho_k \in \mathfrak{S}^p((\bigotimes_{k=1}^n \mathcal{A}_k)^*)
\end{aligned}$$

♠**Note 3.2.** The theory of operator algebra is a deep mathematical theory. However, in this note, we do not use more than the above preparation.

3.3 The linguistic interpretation — Only one measurement is permitted

In this section, we examine the linguistic interpretation (§3.1), i.e., "Only one measurement is permitted". "Only one measurement" implies that "only one observable" and "only one state". That is, we see:

$$
\text{[only one measurement]} \Longrightarrow \begin{cases} \text{only one observable (=measuring instrument)} \\ \\ \text{only one state} \end{cases} \tag{3.10}
$$

♠**Note 3.3.** Although there may be several opinions, I believe that the standard Copenhagen interpretation also says "only one measurement is permitted". Thus, some think that this spirit is inherited to quantum language. However, our assertion is reverse, namely, the Copenhagen interpretation is due to the linguistics interpretation. That is, we assert that

$$
\text{not "}\boxed{\text{Copenhagen interpretation}} \Longrightarrow \boxed{\text{Linguistic interpretation}}\text{"}
$$

$$
\text{but "}\boxed{\text{Linguistic interpretation}} \Longrightarrow \boxed{\text{Copenhagen interpretation}}\text{"}
$$

3.3.1 "Observable is only one" and simultaneous measurement

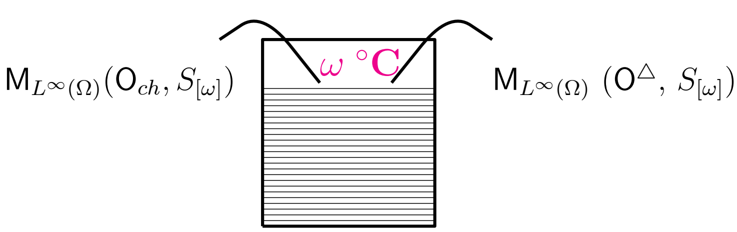
Recall the measurement Example 2.31 (Cold or hot?) and Example 2.32 (Approximate temperature), and consider the following situation:

(a) There is a cup in which water is filled. Assume that the temperature is ω °C ($0 \leqq \omega \leqq 100$). Consider two questions:

$$
\begin{cases} \text{"Is this water cold or hot?"} \\ \\ \text{"How many degrees(°C) is roughly the water?"} \end{cases}
$$

This implies that we take two measurements such that

$$
\begin{cases} (\sharp_1)\text{: } \mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{ch}{=}(\{c,h\}, 2^{\{c,h\}}, F_{ch}), S_{[\omega]}) \text{ in Example2.31} \\ \\ (\sharp_2) : \mathsf{M}_{L^\infty(\Omega)} \; (\mathsf{O}^{\triangle} = (\mathbb{N}_{10}^{100}, 2^{\mathbb{N}_{10}^{100}}, G^{\triangle}), S_{[\omega]}) \text{ in Example2.32} \end{cases}
$$



However, as mentioned in the linguistic interpretation,

"only one measurement" $\Longrightarrow$ "only one observable"

Thus, we have the following problem.

Problem 3.10. Represent two measurements $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{ch}{=}(\{c,h\},\ 2^{\{c,h\}}, F_{ch}), S_{[\omega]})$ and $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}^{\triangle}{=}(\mathbb{N}_{10}^{100},\ 2^{\mathbb{N}_{10}^{100}}, G^{\triangle}),\ S_{[\omega]})$ by only one measurement.

This will be answered in what follows.

Definition 3.11. [Product measurable space] For each $k=1,2,\ldots,n$, consider a measurable $(X_k,\ \mathcal{F}_k)$. The product space $\times_{k=1}^n X_k$ of X_k $(k=1,2,\ldots,n)$ is defined by

$$\mathop{\times}_{k=1}^{n} X_k = \{(x_1,x_2,\ldots,x_n) \mid x_k \in X_k \ \ (k=1,2,\ldots,n)\}$$

Similarly, define the product $\times_{k=1}^n \Xi_k$ of $\Xi_k(\in \mathcal{F}_k)$ $(k=1,2,\ldots,n)$ by

$$\mathop{\times}_{k=1}^{n} \Xi_k = \{(x_1,x_2,\ldots,x_n) \mid x_k \in \Xi_k \ \ (k=1,2,\ldots,n)\}$$

Further, the σ-field $\boxtimes_{k=1}^n \mathcal{F}_k$ on the product space $\times_{k=1}^n X_k$ is defined by

($\sharp$) $\boxtimes_{k=1}^n \mathcal{F}_k$ is the smallest field including $\{\times_{k=1}^n \Xi_k \mid \Xi_k \in \mathcal{F}_k \ \ (k=1,2,\ldots,n)\}$

$(\times_{k=1}^n X_k,\ \boxtimes_{k=1}^n \mathcal{F}_k)$ is called the *product measurable space*. Also, in the case that $(X,\mathcal{F}) = (X_k,\mathcal{F}_k)$ $(k=1,2,\ldots,n)$, the product space $\times_{k=1}^n X_k$ is denoted by X^n, and the product measurable space $(\times_{k=1}^n X_k,\ \boxtimes_{k=1}^n \mathcal{F}_k)$ is denoted by $(X^n,\mathcal{F}^n)$.

Definition 3.12. [Simultaneous observable , simultaneous measurement] Consider the basic structure $[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]$. Let $\rho \in \mathfrak{S}^p(\mathcal{A}^*)$. For each $k=1,2,\ldots,n$, consider a measurement $\mathsf{M}_{\overline{\mathcal{A}}}\ (\mathsf{O}_k = (X_k,\mathcal{F}_k,F_k), S_{[\rho]})$ in $\overline{\mathcal{A}}$. Let $(\times_{k=1}^n X_k,\ \boxtimes_{k=1}^n \mathcal{F}_k)$ be the product measurable space. An observable $\widehat{\mathsf{O}} = (\times_{k\in K} X_k,\ \boxtimes_{k=1}^n \mathcal{F}_k,\ \widehat{F})$ in $\overline{\mathcal{A}}$ is called the simultaneous observable of $\{\mathsf{O}_k \ : \ k=1,2,...,n\}$, if it satisfies the following condition:

$$\widehat{F}(\Xi_1 \times \Xi_2 \times \cdots \times \Xi_n) = F_1(\Xi_1)\cdot F_2(\Xi_2)\cdots F_n(\Xi_n) \qquad (3.11)$$
$$(\ \forall \Xi_k \in \mathcal{F}_k \ (k=1,2,\ldots,n))$$

$\widehat{\mathsf{O}}$ is also denoted by $\times_{k=1}^n \mathsf{O}_k$, $\widehat{F} = \times_{k=1}^n F_k$. Also, the measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\times_{k=1}^n \mathsf{O}_k, S_{[\rho]})$ is called the simultaneous measurement. Here, it should be noted that

- the existence of the simultaneous observable $\times_{k=1}^n \mathsf{O}_k$ is not always guaranteed.

though it always exists in the case that $\overline{\mathcal{A}}$ is commutative (this is, $\overline{\mathcal{A}} = L^\infty(\Omega)$).

In what follows, we shall explain the meaning of "simultaneous observable".

Let us explain the simultaneous measurement. We want to take two measurements $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}_1, S_{[\rho]})$ and measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}_2, S_{[\rho]})$. That is, it suffices to image the following:

$$
\text{(b)}\quad
\begin{tikzcd}
 & \boxed{\text{observable}}_{\mathsf{O}_1=(X_1,\mathcal{F}_1,F_1)} \arrow[r, "\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}_1,S_{[\rho]})"'] & \boxed{\text{measured value}}_{x_1(\in X_1)} \\
\boxed{\text{state}}_{\rho(\in\mathfrak{S}^p(\mathcal{A}^*))} \arrow[ur] \arrow[dr] & & \\
 & \boxed{\text{observable}}_{\mathsf{O}_2=(X_2,\mathcal{F}_2,F_2)} \arrow[r, "\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}_2,S_{[\rho]})"'] & \boxed{\text{measured value}}_{x_2(\in X_2)}
\end{tikzcd}
$$

However, according to the linguistic interpretation (§3.1), two measurements $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}_1, S_{[\rho]})$ and $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}_2, S_{[\rho]})$ can not be taken. That is,

The (b) is impossible

Therefore, combining two observables O_1 and O_2, we construct the simultaneous observable $\mathsf{O}_1 \times \mathsf{O}_2$, and take the simultaneous measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}_1 \times \mathsf{O}_2, S_{[\rho]})$ in what follows.

$$
\text{(c)}\quad
\boxed{\text{state}}_{\rho(\in\mathfrak{S}^p(\mathcal{A}^*))} \longrightarrow \boxed{\text{simultaneous observable}}_{\mathsf{O}_1\times\mathsf{O}_2} \xrightarrow[\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}_1\times\mathsf{O}_2,S_{[\rho]})]{} \boxed{\text{measured value}}_{(x_1,x_2)(\in X_1\times X_2)}
$$

The (c) is possible if $\mathsf{O}_1 \times \mathsf{O}_2$ exists

Answer 3.13. [The answer to Problem3.10] Consider the state space Ω such that $\Omega = [0, 100]$, the closed interval. And consider two observables, that is, [C-H]-observable $\mathsf{O}_{ch} = (X{=}\{c, h\}, 2^X, F_{ch})$ (in Example2.31) and triangle observable $\mathsf{O}^{\triangle} = (Y(=\mathbb{N}_{10}^{100}), 2^Y, G^{\triangle})$ (in Example2.32). Thus, we get the simultaneous observable $\mathsf{O}_{ch}{\times}\mathsf{O}^{\triangle} = (\{c, h\}{\times}\mathbb{N}_{10}^{100}, 2^{\{c,h\}\times\mathbb{N}_{10}^{100}}, F_{ch}\times G^{\triangle})$, and we can take the simultaneous measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{ch} \times \mathsf{O}^{\triangle}, S_{[\omega]})$. For example, putting $\omega = 55$, we see

(d) when the simultaneous measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{ch} \times \mathsf{O}^{\triangle}, S_{[55]})$ is taken, the probability

$$
\text{that the measured value } \begin{bmatrix} (\text{c, about } 50\ ^\circ\text{C}) \\ (\text{c, about } 60\ ^\circ\text{C}) \\ (\text{h, about } 50\ ^\circ\text{C}) \\ (\text{h, about } 60\ ^\circ\text{C}) \end{bmatrix} \text{ is obtained is given by } \begin{bmatrix} 0.125 \\ 0.125 \\ 0.375 \\ 0.375 \end{bmatrix} \tag{3.12}
$$

That is because

$$
[(F_{ch} \times G^{\triangle})(\{(\text{c, about } 50\ ^\circ\text{C})\})](55)
$$

$$=[F_{ch}(\{\text{c}\})](55) \cdot [G^{\triangle}(\{\text{about } 50\,°\text{C}\})](55) = 0.25 \cdot 0.5 = 0.125$$

and similarly,

$$[(F_{ch} \times G^{\triangle})(\{(\text{c}, \text{about } 60\,°\text{C})\})](55) = 0.25 \cdot 0.5 = 0.125$$
$$[(F_{ch} \times G^{\triangle})(\{(\text{h}, \text{about } 50\,°\text{C})\})](55) = 0.75 \cdot 0.5 = 0.375$$
$$[(F_{ch} \times G^{\triangle})(\{(\text{h}, \text{about } 60\,°\text{C})\})](55) = 0.75 \cdot 0.5 = 0.375$$

♠**Note 3.4.** The above argument is not always possible. In quantum mechanics, a simultaneous observable $\mathsf{O}_1 \times \mathsf{O}_2$ does not always exist (See the following Example 3.14 and Heisenberg's uncertainty principle in Sec.4.4).

Example 3.14. [The non-existence of the simultaneous spin observables] Assume that the electron P has the (spin) state $\rho = |u\rangle\langle u| \in \mathfrak{S}^p(B(\mathbb{C}^2))$, where

$$u = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \quad (\text{where, } |u| = (|\alpha_1|^2 + |\alpha_2|^2)^{1/2} = 1)$$

Let $\mathsf{O}_z = (X(=\{\uparrow, \downarrow\}), 2^X, F^z)$ be the spin observable concerning the z-axis such that

$$F^z(\{\uparrow\}) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad F^z(\{\downarrow\}) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Thus, we have the measurement $\mathsf{M}_{B(\mathbb{C}^2)}(\mathsf{O}_z = (X, 2^X, F^z), S_{[\rho]})$.

Let $\mathsf{O}_x = (X, 2^X, F^x)$ be the spin observable concerning the x-axis such that

$$F^x(\{\uparrow\}) = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}, \quad F^x(\{\downarrow\}) = \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix}$$

Thus, we have the measurement $\mathsf{M}_{B(\mathbb{C}^2)}(\mathsf{O}_x = (X, 2^X, F^x), S_{[\rho]})$

Then we have the following problem:

(a) Two measurements $\mathsf{M}_{B(\mathbb{C}^2)}(\mathsf{O}_z = (X, 2^X, F^z), S_{[\rho]})$ and $\mathsf{M}_{B(\mathbb{C}^2)}(\mathsf{O}_x = (X, 2^X, F^x), S_{[\rho]})$ are taken simultaneously?

This is impossible. That is because the two observable O_z and O_x do not commute. For example, we see

$$F^z(\{\uparrow\})F^x(\{\uparrow\}) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 \\ 0 & 0 \end{bmatrix}$$

$$F^x(\{\uparrow\})F^z(\{\uparrow\}) = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1/2 & 0 \\ 1/2 & 0 \end{bmatrix}$$

And thus,

$$F^x(\{\uparrow\})F^z(\{\uparrow\}) \neq F^z(\{\uparrow\})F^x(\{\uparrow\})$$

///

The following theorem is clear. For completeness, we add the proof to it.

Theorem 3.15. [Exact measurement and system quantity] Consider the classical basic structure:

$$[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))]$$

Let $\mathsf{O}_0^{(\text{exa})} = (X, \mathcal{F}, F^{(\text{exa})})$ (i.e., $(X, \mathcal{F}, F^{(\text{exa})}) = (\Omega, \mathcal{B}_\Omega, \chi)$) be the exact observable in $L^\infty(\Omega, \nu)$. Let $\mathsf{O}_1 = (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, G)$ be the observable that is induced by a quantity $\widetilde{g} : \Omega \to \mathbb{R}$ as in Example 2.26(system quantity). Consider the simultaneous observable $\mathsf{O}_0^{(\text{exa})} \times \mathsf{O}_1$. Let (x, y) $(\in X \times \mathbb{R})$ be a measured value obtained by the simultaneous measurement $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O}_0^{(\text{exa})} \times \mathsf{O}_1, S_{[\delta_\omega]})$. Then, we can surely believe that $x = \omega$, and $y = \widetilde{g}(\omega)$.

Proof. Let $D_0 (\in \mathcal{B}_\Omega)$ be arbitrary open set such that $\omega (\in D_0 \subseteq \Omega {=} X)$. Also, let $D_1 (\in \mathcal{B}_{\mathbb{R}})$ be arbitrary open set such that $\widetilde{g}(\omega) \in D_1$. The probability that a measured value (x, y) obtained by the measurement $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O}_0^{(\text{exa})} \times \mathsf{O}_1, S_{[\delta_\omega]})$ belongs to $D_0 \times D_1$ is given by $\chi_{D_0}(\omega) \cdot \chi_{\widetilde{g}^{-1}(D_1)}(\omega) = 1$. Since D_0 and D_1 are arbitrary, we can surely believe that $x = \omega$ and $y = \widetilde{g}(\omega)$. □

3.3.2 "State does not move" and quasi-product observable

We consider that

"only one measurement" $\Longrightarrow$ "state does not move"

That is because

(a) In order to see the state movement, we have to take measurement at least more than twice. However, the "plural measurement" is prohibited. Thus, we conclude "state does not move"

Review 3.16. [= Example 2.34:urn problem] There are two urns U_1 and U_2. The urn U_1 [resp. U_2] contains 8 white and 2 black balls [resp. 4 white and 6 black balls] (*cf.* Figure 3.2).

Table 3.2: urn problem

Urn＼ w·b	white ball	black ball
Urn U_1	8	2
Urn U_2	4	6

Here, consider the following statement (a):

(a) When one ball is picked up from the urn U_2, the probability that the ball is white is 0.4.

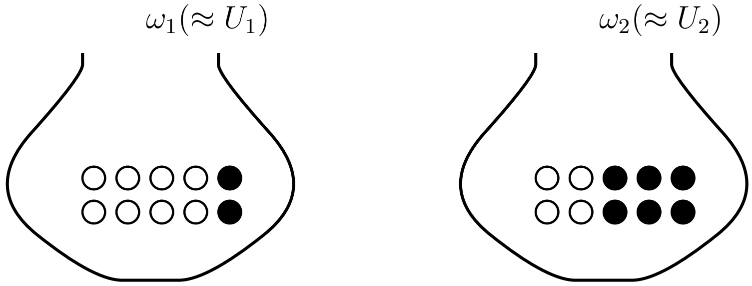


Figure 3.2: Urn problem

In measurement theory, the statement (a) is formulated as follows: Assuming

$$
\begin{array}{lll}
U_1 & \cdots & \text{“the urn with the state } \omega_1\text{”} \\
U_2 & \cdots & \text{“the urn with the state } \omega_2\text{”}
\end{array}
$$

define the state space Ω by $\Omega = \{\omega_1, \omega_2\}$ with discrete metric and counting measure ν. That is, we assume the identification;

$$U_1 \approx \omega_1, \quad U_2 \approx \omega_2,$$

Thus, consider the classical basic structure:

$$[C_0(\Omega) \subseteq L^\infty(\Omega,\nu) \subseteq B(L^2(\Omega,\nu))]$$

Put “w” = “white”, “b” = “black”, and put $X = \{w, b\}$. And define the observable $\mathsf{O}_{wb}(\equiv (X \equiv \{w,b\}, 2^{\{w,b\}}, F_{wb}))$ in $L^\infty(\Omega)$ by

$$
\begin{array}{ll}
[F_{wb}(\{w\})](\omega_1) = 0.8, & [F_{wb}(\{b\})](\omega_1) = 0.2, \\
[F_{wb}(\{w\})](\omega_2) = 0.4, & [F_{wb}(\{b\})](\omega_2) = 0.6.
\end{array}
\tag{3.13}
$$

Thus, we get the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{wb}, S_{[\delta_{\omega_2}]})$. Here, Axiom 1 (§2.7) says that

(b) the probability that a measured value w is obtained by $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{wb}, S_{[\delta_{\omega_2}]})$ is given by

$$F_{wb}(\{b\})(\omega_2) = 0.4$$

Thus, the above statement (b) can be rewritten in the terms of quantum language as follows.

(c) the probability that a measured value $\begin{bmatrix} w \\ b \end{bmatrix}$ is obtained by the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{wb}, S_{[\omega_2]})$ is given by

$$
\begin{bmatrix}
\int_\Omega [F_{wb}(\{w\})](\omega)\delta_{\omega_2}(d\omega) = [F_{wb}(\{w\})](\omega_2) = 0.4 \\
\int_\Omega [F_{wb}(\{b\})](\omega)\delta_{\omega_2}(d\omega) = [F_{wb}(\{b\})](\omega_2) = 0.6
\end{bmatrix}
$$

Problem 3.17. (a) [Sampling with replacement]: Pick out one ball from the urn U_2, and recognize the color (“white” or “black”) of the ball. And **the ball is returned to the**

urn. And again, Pick out one ball from the urn U_2, and recognize the color of the ball. Therefore, we have four possibilities such that.

$$(w,w) \quad (w,b) \quad (b,w) \quad (b,b)$$

It is a common sense that

$$\text{the probability that} \begin{bmatrix} (w,w) \\ (w,b) \\ (b,w) \\ (b,b) \end{bmatrix} \text{ is given by } \begin{bmatrix} 0.16 \\ 0.24 \\ 0.24 \\ 0.36 \end{bmatrix}$$

Now, we have the following problem:

(a) How do we describe the above fact in term of quantum language?

Answer Is suffices to consider the simultaneous measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{wb}^2, S_{[\delta_{\omega_2}]})$ $(= \mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{wb} \times \mathsf{O}_{wb}, S_{[\delta_{\omega_2}]})$), where $\mathsf{O}_{wb}^2 = (\{w,b\} \times \{w,b\}, 2^{\{w,b\}\times\{w,b\}}, F_{wb}^2 (= F_{wb} \times F_{wb}))$. The, we calculate as follows.

$$F_{wb}^2(\{(w,w)\})(\omega_1) = 0.64, \qquad F_{wb}^2(\{(w,b)\})(\omega_1) = 0.16$$
$$F_{wb}^2(\{(b,w)\})(\omega_1) = 0.16, \qquad F_{wb}^2(\{(b,b)\})(\omega_1) = 0.4$$

and

$$F_{wb}^2(\{(w,w)\})(\omega_2) = 0.16, \qquad F_{wb}^2(\{(w,b)\})(\omega_2) = 0.24$$
$$F_{wb}^2(\{(b,w)\})(\omega_2) = 0.24, \qquad F_{wb}^2(\{(b,b)\})(\omega_2) = 0.36$$

Thus, we conclude that

(b) the probability that a measured value $\begin{bmatrix} (w,w) \\ (w,b) \\ (b,w) \\ (b,b) \end{bmatrix}$ is obtained by $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{wb} \times \mathsf{O}_{wb}, S_{[\delta_{\omega_2}]})$

$$\text{is given by} \begin{bmatrix} [F_{wb}(\{w\})](\omega_2) \cdot [F_{wb}(\{w\})](\omega_2) = 0.16 \\ [F_{wb}(\{w\})](\omega_2) \cdot [F_{wb}(\{b\})](\omega_2) = 0.24 \\ [F_{wb}(\{b\})](\omega_2) \cdot [F_{wb}(\{w\})](\omega_2) = 0.24 \\ [F_{wb}(\{b\})](\omega_2) \cdot [F_{wb}(\{b\})](\omega_2) = 0.36 \end{bmatrix}$$

Problem 3.18. (a) [Sampling without replacement]: Pick out one ball from the urn U_2, and recognize the color ("white" or "black") of the ball. And **the ball is not returned to the urn.** And again, Pick out one ball from the urn U_2, and recognize the color of the ball. Therefore, we have four possibilities such that.

$$(w,w) \quad (w,b) \quad (b,w) \quad (b,b)$$

It is a common sense that

$$\text{the probability that} \begin{bmatrix} (w,w) \\ (w,b) \\ (b,w) \\ (b,b) \end{bmatrix} \text{ is given by } \begin{bmatrix} 12/90 \\ 24/90 \\ 24/90 \\ 30/90 \end{bmatrix}$$

Now, we have the following problem:

(a) How do we describe the above fact in term of quantum language?

Now, recall the simultaneous observable (Definition3.12) as follows. Let $\mathsf{O}_k = (X_k, \mathcal{F}_k, F_k)$ $(k = 1, 2, \ldots, n\ )$ be observables in $\overline{\mathcal{A}}$. The simultaneous observable $\widehat{\mathsf{O}} = (\times_{k=1}^n X_k,\ \boxtimes_{k=1}^n \mathcal{F}_k,\ \widehat{F})$ is defined by

$$\widehat{F}(\Xi_1 \times \Xi_2 \times \cdots \times \Xi_n) = F_1(\Xi_1) F_2(\Xi_2) \cdots F_n(\Xi_n) \qquad (\forall \Xi_k \in \mathcal{F}_k, \forall k = 1, 2, \ldots, n)$$

The following definition ("quasi-product observable") is a kind of simultaneous observable:

Definition 3.19. [quasi-product observable] Let $\mathsf{O}_k = (X_k, \mathcal{F}_k, F_k)$ $(k = 1, 2, \ldots, n\ )$ be observables in a W^*-algebra $\overline{\mathcal{A}}$. Assume that an observable $\mathsf{O}_{12\ldots n} = (\times_{k=1}^n X_k,\ \boxtimes_{k=1}^n \mathcal{F}_k,\ F_{12\ldots n})$ satisfies

$$F_{12\ldots n}(X_1 \times \cdots \times X_{k-1} \times \Xi_k \times X_{k+1} \times \cdots \times X_n) = F_k(\Xi_k) \qquad (\forall \Xi_k \in \mathcal{F}_k, \forall k = 1, 2, \ldots, n) \tag{3.14}$$

The observable $\mathsf{O}_{12\ldots n} = (\times_{k=1}^n X_k,\ \boxtimes_{k=1}^n \mathcal{F}_k,\ F_{12\ldots n})$ is called a **quasi-product observable** of $\{\mathsf{O}_k \mid k = 1, 2, \ldots, n\}$, and denoted by

$$\mathop{\times}_{k=1,2,\ldots,n}^{\mathrm{qp}} \mathsf{O}_k = (\mathop{\times}_{k=1}^{n} X_k,\ \boxtimes_{k=1}^n \mathcal{F}_k,\ \mathop{\times}_{k=1,2,\ldots,n}^{\mathrm{qp}} F_k)$$

Of course, a simultaneous observable is a kind of quasi-product observable. Therefore, quasi-product observable is not uniquely determined. Also, in quantum systems, the existence of the quasi-product observable is not always guaranteed.

Answer 3.20. [The answer to Problem 3.17] Define the quasi-product observable $\mathsf{O}_{wb} \overset{\mathrm{qp}}{\times} \mathsf{O}_{wb} = (\{w,b\} \times \{w,b\}, 2^{\{w,b\} \times \{w,b\}}, F_{12}(= F_{wb} \overset{\mathrm{qp}}{\times} F_{wb}))$ of $\mathsf{O}_{wb} = (\{w,b\}, 2^{\{w,b\}}, F)$ in $L^\infty(\Omega)$ such that

$$F_{12}(\{(w,w)\})(\omega_1) = \frac{8 \times 7}{90}, \qquad F_{12}(\{(w,b)\})(\omega_1) = \frac{8 \times 2}{90}$$
$$F_{12}(\{(b,w)\})(\omega_1) = \frac{2 \times 8}{90}, \qquad F_{12}(\{(b,b)\})(\omega_1) = \frac{2 \times 1}{90}$$
$$F_{12}(\{(w,w)\})(\omega_2) = \frac{4 \times 3}{90}, \qquad F_{12}(\{(w,b)\})(\omega_2) = \frac{4 \times 6}{90}$$

$$F_{12}(\{(b,w)\})(\omega_2)=\frac{6\times 4}{90}, \qquad F_{12}(\{(b,b)\})(\omega_2)=\frac{6\times 5}{90}$$

Thus, we have the (quasi-product) measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{12}, S_{[\omega]})$

Therefore, in terms of quantum language, we describe as follows.

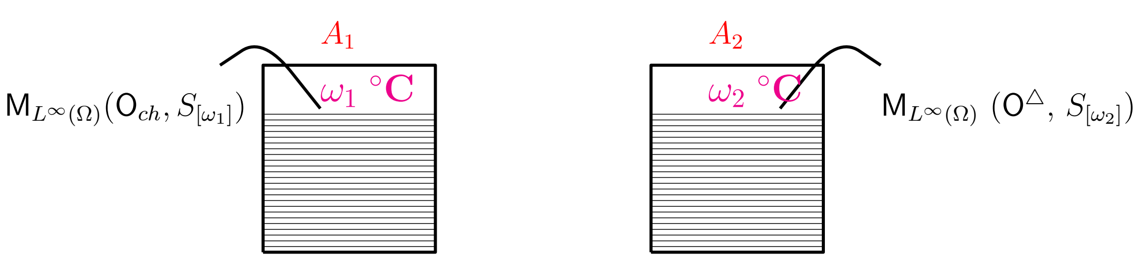
(b) the probability that a measured value $\begin{bmatrix} (w,w) \\ (w,b) \\ (b,w) \\ (b,b) \end{bmatrix}$ is obtained dy $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{wb} \overset{\text{qp}}{\times} \mathsf{O}_{wb}, S_{[\delta_{\omega_2}]})$ is given by $\begin{bmatrix} [F_{12}(\{(w,w)\})](\omega_2)=\frac{4\times 3}{90} \\ [F_{12}(\{(w,b)\})](\omega_2)=\frac{4\times 6}{90} \\ [F_{12}(\{(b,w)\})](\omega_2)=\frac{4\times 6}{90} \\ [F_{12}(\{(b,b)\})](\omega_2)=\frac{6\times 5}{90} \end{bmatrix}$

3.3.3 Only one state and parallel measurement

For example, consider the following situation:

(a) There are two cups A_1 and A_2 in which water is filled. Assume that the temperature of the water in the cup A_k $(k=1,2)$ is ω_k °C $(0 \leqq \omega_k \leqq 100)$. Consider two questions "Is the water in the cup A_1 cold or hot?" and "How many degrees(°C) is roughly the water in the cup A_2?". This implies that we take two measurements such that

$$\begin{cases} (\sharp_1)\text{: } \mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{ch}{=}(\{c,h\}, 2^{\{c,h\}}, F_{ch}), S_{[\omega_1]}) \text{ in Example2.31} \\ (\sharp_2) : \mathsf{M}_{L^\infty(\Omega)}\ (\mathsf{O}^{\triangle} = (\mathbb{N}_{10}^{100}, 2^{\mathbb{N}_{10}^{100}}, G^{\triangle}), S_{[\omega_2]}) \text{ in Example2.32} \end{cases}$$



However, as mentioned in the above,

"only one state" must be demanded.

Thus, we have the following problem.

Problem 3.21. Represent two measurements $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{ch}{=}(\{c,h\},\ 2^{\{c,h\}}, F_{ch}), S_{[\omega_1]})$ and $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}^{\triangle} = (\mathbb{N}_{10}^{100},\ 2^{\mathbb{N}_{10}^{100}}, G^{\triangle}),\ S_{[\omega_2]})$ by only one measurement.

This will be answered in what follows.

Definition 3.22. [Parallel observable] For each $k = 1, 2, \ldots, n$, consider a basic structure $[\mathcal{A}_k \subseteq \overline{\mathcal{A}}_k \subseteq B(H_k)]$, and an observable $\mathsf{O}_k = (X_k, \mathcal{F}_k, F_k)$ in $\overline{\mathcal{A}}_k$. Define the observable $\widetilde{\mathsf{O}} = (\times_{k=1}^n X_k,\ \boxtimes_{k=1}^n \mathcal{F}_k, \widetilde{F})$ in $\bigotimes_{k=1}^n \overline{\mathcal{A}}_k$ such that

$$\widetilde{F}(\Xi_1 \times \Xi_2 \times \cdots \times \Xi_n) = F_1(\Xi_1) \otimes F_2(\Xi_2) \otimes \cdots \otimes F_n(\Xi_n) \qquad \forall \Xi_k \in \mathcal{F}_k \ (k = 1, 2, \ldots, n) \tag{3.15}$$

Then, the observable $\widetilde{\mathsf{O}} = (\times_{k=1}^n X_k,\ \boxtimes_{k=1}^n \mathcal{F}_k, \widetilde{F})$ is called the parallel observable in $\bigotimes_{k=1}^n \overline{\mathcal{A}}_k$, and denoted by $\widetilde{F} = \bigotimes_{k=1}^n F_k$, $\widetilde{\mathsf{O}} = \bigotimes_{k=1}^n \mathsf{O}_k$. the measurement of the parallel observable $\widetilde{\mathsf{O}} = \bigotimes_{k=1}^n \mathsf{O}_k$, that is, the measurement $\mathsf{M}_{\bigotimes_{k=1}^n \overline{\mathcal{A}}_k}(\widetilde{\mathsf{O}},\ S_{[\bigotimes_{k=1}^n \rho_k]})$ is called a **parallel measurement**, and denoted by $\mathsf{M}_{\bigotimes_{k=1}^n \overline{\mathcal{A}}_k}(\bigotimes_{k=1}^n \mathsf{O}_k,\ S_{[\bigotimes_{k=1}^n \rho_k]})$ or $\bigotimes_{k=1}^n \mathsf{M}_{\overline{\mathcal{A}}_k}(\mathsf{O}_k,\ S_{[\rho_k]})$.

The meaning of the parallel measurement is as follows.

Our present purpose is

- to take both measurements $\mathsf{M}_{\overline{\mathcal{A}}_1}(\mathsf{O}_1, S_{[\rho_1]})$ and $\mathsf{M}_{\overline{\mathcal{A}}_2}(\mathsf{O}_2, S_{[\rho_2]})$

Then. image the following:

$$\text{(b)} \quad \left\{ \begin{array}{lll} \underset{\rho_1(\in \mathfrak{S}^p(\mathcal{A}_1^*))}{\boxed{\text{state}}} \xrightarrow{\hspace{2cm}} & \underset{\mathsf{O}_1}{\boxed{\text{observable}}} \xrightarrow[\mathsf{M}_{\overline{\mathcal{A}}_1}(\mathsf{O}_1, S_{[\rho_1]})]{} & \underset{x_1(\in X_1)}{\boxed{\text{measured value}}} \\ \\ \underset{\rho_2(\in \mathfrak{S}^p(\mathcal{A}_2^*))}{\boxed{\text{state}}} \xrightarrow{\hspace{2cm}} & \underset{\mathsf{O}_2}{\boxed{\text{observable}}} \xrightarrow[\mathsf{M}_{\overline{\mathcal{A}}_2}(\mathsf{O}_2, S_{[\rho_2]})]{} & \underset{x_2(\in X_2)}{\boxed{\text{measured value}}} \end{array} \right.$$

However, according to the linguistic interpretation (§3.1), two measurements can not be taken. Hence,

The (b) is impossible

Thus, two states ρ_1 and ρ_1 are regarded as one state $\rho_1 \otimes \rho_2$, and further, combining two observables O_1 and O_2, we construct the parallel observable $\mathsf{O}_1 \otimes \mathsf{O}_2$, and take the parallel measurement $\mathsf{M}_{\overline{\mathcal{A}}_1 \otimes \overline{\mathcal{A}}_2}(\mathsf{O}_1 \otimes \mathsf{O}_2, S_{[\rho_1 \otimes \rho_2]})$ in what follows.

$$\text{(c)} \quad \underset{\rho_1 \otimes \rho_2(\in \mathfrak{S}^p(\mathcal{A}_1^*) \otimes \mathfrak{S}^p(\mathcal{A}_2^*))}{\boxed{\text{state}}} \rightarrow \underset{\mathsf{O}_1 \otimes \mathsf{O}_2}{\boxed{\text{parallel observable}}} \xrightarrow[\mathsf{M}_{\overline{\mathcal{A}}_1 \otimes \overline{\mathcal{A}}_2}(\mathsf{O}_1 \otimes \mathsf{O}_2, S_{[\rho_1 \otimes \rho_2]})]{} \underset{(x_1, x_2)(\in X_1 \times X_2)}{\boxed{\text{measured value}}}$$

The (c) is always possible

Example 3.23. [The answer to Problem 3.21] Put $\Omega_1 = \Omega_2 = [0, 100]$, and define the state space $\Omega_1 \times \Omega_2$. And consider two observables, that is, the [C-H]-observable $\mathsf{O}_{ch} = (X{=}\{c, h\}, 2^X, F_{ch})$ in $C(\Omega_1)$ (in Example2.31) and triangle-observable $\mathsf{O}^{\triangle} = (Y(=\mathbb{N}_{10}^{100}), 2^Y, G^{\triangle})$ in $L^\infty(\Omega_2)$ (in Example2.32). Thus, we get the parallel observable $\mathsf{O}_{ch} \otimes \mathsf{O}^{\triangle} = (\{c, h\} \times \mathbb{N}_{10}^{100}, 2^{\{c,h\}\times\mathbb{N}_{10}^{100}}, F_{ch} \otimes G^{\triangle})$ in $L^\infty(\Omega_1 \times \Omega_2)$, take the parallel measurement $\mathsf{M}_{L^\infty(\Omega_1\times\Omega_2)}(\mathsf{O}_{ch} \otimes \mathsf{O}^{\triangle}, S_{[(\omega_1,\omega_2)]})$. Here, note that

$$\delta_{\omega_1} \otimes \delta_{\omega_2} = \delta_{(\omega_1,\omega_2)} \approx (\omega_1, \omega_2).$$

For example, putting $(\omega_1, \omega_2) = (25, 55)$, we see the following.

(d) When the parallel measurement $\mathsf{M}_{L^\infty(\Omega_1\times\Omega_2)}(\mathsf{O}_{ch} \otimes \mathsf{O}^{\triangle}, S_{[(25,55)]})$ is taken, the probability

$$\text{that the measured value } \begin{bmatrix} (\text{c, about } 50\ {}^\circ\text{C}) \\ (\text{c, about } 60\ {}^\circ\text{C}) \\ (\text{h, about } 50\ {}^\circ\text{C}) \\ (\text{h, about } 60\ {}^\circ\text{C}) \end{bmatrix} \text{ is obtained is given by } \begin{bmatrix} 0.375 \\ 0.375 \\ 0.125 \\ 0.125 \end{bmatrix}$$

That is because

$$\begin{aligned} &[(F_{ch} \otimes G^{\triangle})(\{(\text{c, about } 50\ {}^\circ\text{C})\})](25, 55) \\ =&[F_{ch}(\{\text{c}\})](25) \cdot [G^{\triangle}(\{\text{about } 50\ {}^\circ\text{C}\})](55) = 0.75 \cdot 0.5 = 0.375 \end{aligned}$$

Thus, similarly,

$$[(F_{ch} \otimes G^{\triangle})(\{(\text{c, about } 60\ {}^\circ\text{C})\})](25, 55) = 0.75 \cdot 0.5 = 0.375$$
$$[(F_{ch} \otimes G^{\triangle})(\{(\text{h, about } 50\ {}^\circ\text{C})\})](25, 55) = 0.25 \cdot 0.5 = 0.125$$
$$[(F_{ch} \otimes G^{\triangle})(\{(\text{h, about } 60\ {}^\circ\text{C})\})](25, 55) = 0.25 \cdot 0.5 = 0.125$$

Remark 3.24. Also, for example, putting $(\omega_1, \omega_2) = (55, 55)$, we see:

(e) the probability that a measured value $\begin{bmatrix} (\text{c, about } 50\ {}^\circ\text{C}) \\ (\text{c, about } 60\ {}^\circ\text{C}) \\ (\text{h, about } 50\ {}^\circ\text{C}) \\ (\text{h, about } 60\ {}^\circ\text{C}) \end{bmatrix}$ is obtained by parallel measurement $\mathsf{M}_{L^\infty(\Omega_1\times\Omega_2)}(\mathsf{O}_{ch} \otimes \mathsf{O}^{\triangle}, S_{[(55,55)]})$ is given by $\begin{bmatrix} 0.125 \\ 0.125 \\ 0.375 \\ 0.375 \end{bmatrix}$

That is because, we similarly, see

$$
\begin{cases}
[F_{ch}(\{\mathrm{c}\})](55) \cdot [G^{\triangle}(\{\text{about } 50\,^\circ\mathrm{C}\})](55) = 0.25 \cdot 0.5 = 0.125 \\
[F_{ch}(\{\mathrm{c}\})](55) \cdot [G^{\triangle}(\{\text{about } 60\,^\circ\mathrm{C}\})](55) = 0.25 \cdot 0.5 = 0.125 \\
[F_{ch}(\{\mathrm{h}\})](55) \cdot [G^{\triangle}(\{\text{about } 50\,^\circ\mathrm{C}\})](55) = 0.75 \cdot 0.5 = 0.375 \\
[F_{ch}(\{\mathrm{h}\})](55) \cdot [G^{\triangle}(\{\text{about } 60\,^\circ\mathrm{C}\})](55) = 0.75 \cdot 0.5 = 0.375
\end{cases}
\tag{3.16}
$$

Note that this is the same as Answer 3.13 (cf. Note 3.5 later).

The following theorem is clear. But, the assertion is significant.

Theorem 3.25. [Ergodic property] For each $k = 1, 2, \cdots, n$, consider a measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_k(:= (X_k, \mathcal{F}_k, F_k)), S_{[\delta_\omega]})$ with the sample probability space $(X_k, \mathcal{F}_k, P_k^\omega)$. Then, the sample probability spaces of the simultaneous measurement $\mathsf{M}_{L^\infty(\Omega)}(\times_{k=1}^n \mathsf{O}_k, S_{[\delta_\omega]})$ and the parallel measurement $\mathsf{M}_{L^\infty(\Omega^n)}\ (\bigotimes_{k=1}^n \mathsf{O}_k,\ S_{[\otimes_{k=1}^n \delta_\omega]})$ are the same, that is, these are the same as the product probability space

$$
(\mathop{\times}_{k=1}^{n} X_k,\ \boxtimes_{k=1}^{n} \mathcal{F}_k, \bigotimes_{k=1}^{n} P_k^\omega)
\tag{3.17}
$$

Proof. It is clear, and thus we omit the proof. (Also, see Note 3.5 later.) □

Example 3.26. [The parallel measurement is always meaningful in both classical and quantum systems] The electron P_1 has the (spin) state $\rho_1 = |u_1\rangle\langle u_1| \in \mathfrak{S}^p(B(\mathbb{C}^2))$ such that

$$
u_1 = \begin{bmatrix} \alpha_1 \\ \beta_1 \end{bmatrix} \quad (\text{where, } \|u_1\| = (|\alpha_1|^2 + |\beta_1|^2)^{1/2} = 1)
$$

Let $\mathsf{O}_z = (X(=\{\uparrow, \downarrow\}), 2^X, F^z)$ be the spin observable concerning the z-axis such that

$$
F^z(\{\uparrow\}) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad F^z(\{\downarrow\}) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}
$$

Thus, we have the measurement $\mathsf{M}_{B(\mathbb{C}^2)}(\mathsf{O}_z = (X, 2^X, F^z), S_{[\rho_1]})$.

The electron P_2 has the (spin) state $\rho_2 = |u_2\rangle\langle u_2| \in \mathfrak{S}^p(B(\mathbb{C}^2))$ such that

$$
u = \begin{bmatrix} \alpha_2 \\ \beta_2 \end{bmatrix} \quad (\text{where, } \|u_2\| = (|\alpha_2|^2 + |\beta_2|^2)^{1/2} = 1)
$$

Let $\mathsf{O}_x = (X, 2^X, F^x)$ be the spin observable concerning the x-axis such that

$$
F^x(\{\uparrow\}) = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}, \quad F^x(\{\downarrow\}) = \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix}
$$

Thus, we have the measurement $\mathsf{M}_{B(\mathbb{C}^2)}(\mathsf{O}_x = (X, 2^X, F^x), S_{[\rho_2]})$

Then we have the following problem:

(a) Two measurements $\mathsf{M}_{B(\mathbb{C}^2)}(\mathsf{O}_z = (X, 2^X, F^z), S_{[\rho_1]})$ and $\mathsf{M}_{B(\mathbb{C}^2)}(\mathsf{O}_x = (X, 2^X, F^x), S_{[\rho_2]})$ are taken simultaneously?

This is possible. It can be realized by the parallel measurement

$$\mathsf{M}_{B(\mathbb{C}^2)\otimes B(\mathbb{C}^2)}(\mathsf{O}_z \otimes \mathsf{O}_z = (X \times X, 2^{X\times X}, F^z \otimes F^x), S_{[\rho\otimes\rho]})$$

That is,

(b) The probability that a measured value $\begin{bmatrix} (\uparrow,\uparrow) \\ (\uparrow,\downarrow) \\ (\downarrow,\uparrow) \\ (\downarrow,\downarrow) \end{bmatrix}$ is obtained by the parallel measurement $\mathsf{M}_{B(\mathbb{C}^2)\otimes B(\mathbb{C}^2)}(\mathsf{O}_z \otimes \mathsf{O}_z, S_{[\rho\otimes\rho]})$ is given by

$$\begin{bmatrix} \langle u, F^z(\{\uparrow\})u\rangle\langle u, F^x(\{\uparrow\})u\rangle = p_1 p_2 \\ \langle u, F^z(\{\uparrow\})u\rangle\langle u, F^x(\{\downarrow\})u\rangle = p_1(1-p_2) \\ \langle u, F^z(\{\downarrow\})u\rangle\langle u, F^x(\{\uparrow\})u\rangle = (1-p_1)p_2 \\ \langle u, F^z(\{\downarrow\})u\rangle\langle u, F^x(\{\downarrow\})u\rangle = (1-p_1)(1-p_2) \end{bmatrix}$$

where $p_1 = |\alpha_1|^2, \quad p_2 = \frac{1}{2}(|\alpha_1|^2 + \widehat{\alpha}_1\alpha_2 + \alpha_1\widehat{\alpha}_2 + |\alpha_2|^2)$

♠**Note 3.5.** Theorem 3.25 is rather deep in the following sense. For example, "To toss a coin 10 times" is a simultaneous measurement. On the other hand, "To toss 10 coins once" is characterized as a parallel measurement. The two have the same sample space. That is,

$$\text{"spatial average"} = \text{"time average"}$$

which is called the **ergodic property**. This means that the two are not distinguished by the sample space and not the measurements (i.e., a simultaneous measurement and a parallel measurement). However, this is peculiar to classical pure measurements. It does not hold in classical mixed measurements and quantum measurement.

Chapter 4

Linguistic interpretation of quantum systems

Measurement theory (= quantum language) is formulated as follows.

$$\bullet \quad \underset{\text{(=quantum language)}}{\boxed{\text{measurement theory}}} := \underbrace{\overset{\text{[Axiom 1]}}{\underset{(cf.\ \S 2.7)}{\boxed{\text{Measurement}}}} + \overset{\text{[Axiom 2]}}{\underset{(cf.\ \S 10.3)}{\boxed{\text{Causality}}}}}_{\text{a kind of spell(a priori judgment)}} + \underbrace{\overset{\text{[quantum linguistic interpretation]}}{\underset{(cf.\ \S 3.1)}{\boxed{\text{Linguistic interpretation}}}}}_{\text{manual to use spells}}$$

Measurement theory says that

- Describe every phenomenon modeled on Axioms 1 and 2 (by a hint of the linguistic interpretation)!

In this chapter, we devote ourselves to the linguistic interpretation (§3.1) for general (or, quantum) systems.

4.1 Kolmogorov's extension theorem and the linguistic interpretation

Kolmogorov's probability theory (*cf.* [54]) starts from the following spell:

($\sharp$) Let $(X, \mathcal{F}, P)$ be a probability space. Then, the probability that a event Ξ $(\in \mathcal{F})$ happens is given by $P(\Xi)$

And, through trial and error, Kolmogorov found his extension theorem, which says that

($\sharp$) **"Only one probability space is permitted"**

which surely corresponds to

($\sharp$) **"Only one measurement is permitted" in the linguistic interpretation (§3.1)**

Therefore, we want to say that

($\sharp$) Parmenides (born around BC. 515) and Kolmogorov (1903-1987) said about the same thing

(*cf.* Parmenides' words (3.3)).

Let $\widehat{\Lambda}$ be a set (called an index set). For each $\lambda \in \widehat{\Lambda}$, consider a set X_λ. For any subsets $\Lambda_1 \subseteq \Lambda_2 (\subseteq \widehat{\Lambda})$, $\pi_{\Lambda_1,\Lambda_2}$ is the natural map such that:

$$\pi_{\Lambda_1,\Lambda_2} : \mathop{\times}_{\lambda \in \Lambda_2} X_\lambda \longrightarrow \mathop{\times}_{\lambda \in \Lambda_1} X_\lambda. \tag{4.1}$$

Especially, put $\pi_\Lambda = \pi_{\Lambda,\widehat{\Lambda}}$. Consider the basic structure

$$[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]$$

For each $\lambda \in \widehat{\Lambda}$, consider an observable $(X_\lambda, \mathcal{F}_\lambda, F_\lambda)$ in $\overline{\mathcal{A}}$. Note that the quasi-product observable $\overline{\mathbf{O}} \equiv (\times_{\lambda\in\widehat{\Lambda}} X_\lambda, \ \times_{\lambda\in\widehat{\Lambda}} \mathcal{F}_\lambda, \ F_{\widehat{\Lambda}})$ of $\{ \ (X_\lambda, \mathcal{F}_\lambda, \ F_\lambda) \ \mid \ \lambda \in \widehat{\Lambda} \ \}$ is characterized as the observable such that:

$$F_{\widehat{\Lambda}}(\pi^{-1}_{\{\lambda\}}(\Xi_\lambda)) = F_\lambda(\Xi_\lambda) \qquad (\forall \Xi_\lambda \in \mathcal{F}_\lambda, \forall \lambda \in \widehat{\Lambda}), \tag{4.2}$$

though the existence and the uniqueness of a quasi-product observable are not guaranteed in general. The following theorem says something about the existence and uniqueness of the quasi-product observable.

Let $\widetilde{\Lambda}$ be a set. For each $\lambda \in \widetilde{\Lambda}$, consider a set X_λ. For any subset $\Lambda_1 \subseteq \Lambda_2 (\subseteq \widetilde{\Lambda})$, define the natural map $\pi_{\Lambda_1,\Lambda_2} : \times_{\lambda\in\Lambda_2} X_\lambda \longrightarrow \times_{\lambda\in\Lambda_1} X_\lambda$ by

$$\mathop{\times}_{\lambda \in \Lambda_2} X_\lambda \ni (x_\lambda)_{\lambda\in\Lambda_2} \mapsto (x_\lambda)_{\lambda\in\Lambda_1} \in \mathop{\times}_{\lambda \in \Lambda_1} X_\lambda \tag{4.3}$$

The following theorem guarantees the existence and uniqueness of the observable. It should be noted that this is due to the the linguistic interpretation (§3.1), i.e., "only one measurement is permitted".

Theorem 4.1. [Kolmogorov extension theorem in measurement theory (cf. [27, 29])] Consider the basic structure

$$[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]$$

For each $\lambda \in \widehat{\Lambda}$, consider a Borel measurable space $(X_\lambda, \mathcal{F}_\lambda)$, where X_λ is a separable complete metric space. Define the set $\mathcal{P}_0(\widehat{\Lambda})$ such as $\mathcal{P}_0(\widehat{\Lambda}) \equiv \{\Lambda \subseteq \widehat{\Lambda} \mid \Lambda \text{ is finite } \}$. Assume that the family of the observables $\big\{ \ \overline{\mathbf{O}}_\Lambda \equiv (\ \times_{\lambda\in\Lambda} X_\lambda, \ \times_{\lambda\in\Lambda} \mathcal{F}_\lambda, \ F_\Lambda \ ) \ \mid \ \Lambda \in \mathcal{P}_0(\widehat{\Lambda}) \ \big\}$ in $\overline{\mathcal{A}}$ satisfies the following "**consistency condition**":

- for any $\Lambda_1, \Lambda_2 \in \mathcal{P}_0(\widehat{\Lambda})$ such that $\Lambda_1 \subseteq \Lambda_2$,

$$F_{\Lambda_2}\big(\pi^{-1}_{\Lambda_1,\Lambda_2}(\Xi_{\Lambda_1})\big) = F_{\Lambda_1}\big(\Xi_{\Lambda_1}\big) \quad (\forall \Xi_{\Lambda_1} \in \mathop{\times}_{\lambda\in\Lambda_1} \mathcal{F}_\lambda). \tag{4.4}$$

Then, there uniquely exists the observable $\widetilde{\mathbf{O}}_{\widehat{\Lambda}} \equiv \big(\times_{\lambda\in\widehat{\Lambda}} X_\lambda,\ \times_{\lambda\in\widehat{\Lambda}} \mathcal{F}_\lambda,\ \widetilde{F}_{\widehat{\Lambda}}\big)$ in $\overline{\mathcal{A}}$ such that:

$$\widetilde{F}_{\widehat{\Lambda}}\big(\pi^{-1}_{\Lambda}(\Xi_\Lambda)\big) = F_\Lambda\big(\Xi_\Lambda\big) \quad (\forall \Xi_\Lambda \in \mathop{\times}_{\lambda\in\Lambda} \mathcal{F}_\lambda,\ \forall \Lambda \in \mathcal{P}_0(\widehat{\Lambda})).$$

Proof. For the proof, see refs.[27, 29].

□

Corollary 4.2. [Infinite simultaneous observable] Consider the basic structure

$$[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]$$

Let $\widetilde{\Lambda}$ be a set. For each $\lambda \in \widetilde{\Lambda}$, assume that X_λ is a separable complete metric space, $\mathcal{F}_\lambda$ is its Borel field. For each $\lambda \in \widetilde{\Lambda}$, consider an observable $\mathsf{O}_\lambda = (X_\lambda, \mathcal{F}_\lambda, F_\lambda)$ in $\overline{\mathcal{A}}$ such that it satisfies the commutativity condition, that is,

$$F_{k_1}(\Xi_{k_1})F_{k_2}(\Xi_{k_2}) = F_{k_2}(\Xi_{k_2})F_{k_1}(\Xi_{k_1}) \quad (\forall \Xi_{k_1} \in \mathcal{F}_{k_1},\ \forall \Xi_{k_2} \in \mathcal{F}_{k_2},\ k_1 \neq k_2) \tag{4.5}$$

Then, a simultaneous observable $\widehat{\mathsf{O}} = (\times_{\lambda\in\widetilde{\Lambda}} X_\lambda,\ \boxtimes_{\lambda\in\widetilde{\Lambda}}\mathcal{F}_\lambda,\ \widehat{F}= \times_{\lambda\in\widetilde{\Lambda}} F_\lambda)$ uniquely exists. That is, for any finite set $\Lambda_0 (\subseteq \widetilde{\Lambda})$, it holds that

$$\widehat{F}\big((\mathop{\times}_{\lambda\in\Lambda_0} \Xi_\lambda) \times (\mathop{\times}_{\lambda\in\widetilde{\Lambda}\setminus\Lambda_0} X_\lambda)\big) = \mathop{\times}_{\lambda\in\Lambda_0} F_\lambda(\Xi_\lambda) \qquad (\forall \Xi_\lambda \in \mathcal{F}_\lambda, \forall \lambda \in \Lambda_0)$$

Proof. The proof is a direct consequence of Theorem 4.1. Thus, it is omitted. □

Remark 4.3. Now we can answer the following question:

(B) Why is Kolmogorov's extension theory fundamental in probability theory ?

That is, I can assert the following chain:

$$\underset{\text{(Linguistic interpretation)}}{\boxed{\text{Only one measurement is permitted}}} \longrightarrow \underset{\text{(Kolmogorov's extension theorem 4.1 in quantum language)}}{\boxed{\text{The existence of measurement}}} \longrightarrow \underset{\text{(Kolmogorov's extension theorem)}}{\boxed{\text{The existence of sample space}}}$$

///

4.2 The law of large numbers in quantum language

4.2.1 The sample space of infinite parallel measurement $\bigotimes_{k=1}^{\infty} \mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}=(X,\mathcal{F},F), S_{[\rho]})$

Consider the basic structure

$$[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]$$
$$\Big(\text{that is, } [\mathcal{C}(H) \subseteq B(H) \subseteq B(H)], \text{ or } [C_0(\Omega) \subseteq L^\infty(\Omega,\nu) \subseteq B(L^2(\Omega,\nu))]\Big)$$

and measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}=(X,\mathcal{F},F), S_{[\rho]})$, which has the sample probability space $(X,\mathcal{F},P_\rho)$

Note that the existence of the infinite parallel observable $\widetilde{\mathsf{O}}$ $(=\bigotimes_{k=1}^{\infty}\mathsf{O}) = (X^{\mathbb{N}},\ \boxtimes_{k=1}^{\infty}\mathcal{F},$ $\widetilde{F}(=\bigotimes_{k=1}^{\infty}F))$ in an infinite tensor W^*-algebra $\bigotimes_{k=1}^{\infty}\overline{\mathcal{A}}$ is assured by Kolmogorov's extension theorem (Corollary4.2).

For completeness, let us calculate the sample probability space of the parallel measurement $\mathsf{M}_{\bigotimes_{k=1}^{\infty}\overline{\mathcal{A}}}(\widetilde{\mathsf{O}}, S_{[\bigotimes_{k=1}^{\infty}\rho]})$ in both cases (i.e., quantum case and classical case):

Preparation 4.4. [I]: quantum system: The quantum infinite tensor basic structure is defined by

$$[\mathcal{C}(\otimes_{k=1}^{\infty}H) \subseteq B(\otimes_{k=1}^{\infty}H) \subseteq B(\otimes_{k=1}^{\infty}H)]$$

Therefore, infinite tensor state space is characterized by

$$\mathfrak{S}^p(\mathcal{T}r(\otimes_{k=1}^{\infty}H)) \subset \mathfrak{S}^m(\mathcal{T}r(\otimes_{k=1}^{\infty}H)) = \overline{\mathfrak{S}}^m(\mathcal{T}r(\otimes_{k=1}^{\infty}H)) \tag{4.6}$$

Since Definition 2.17 says that $\mathcal{F}=\mathcal{F}_\rho$ $(\forall \rho \in \mathfrak{S}^p(\mathcal{T}r(H)))$, the sample probability space $(X^{\mathbb{N}},$ $\boxtimes_{k=1}^{\infty}\mathcal{F},\ P_{\bigotimes_{k=1}^{\infty}\rho})$ of the infinite parallel measurement $\mathsf{M}_{\bigotimes_{k=1}^{\infty}B(H)}(\otimes_{k=1}^{\infty}\mathsf{O} = (X^{\mathbb{N}},\ \boxtimes_{k=1}^{\infty}\mathcal{F},$ $\otimes k=1^{\infty}F),\ S_{[\bigotimes_{k=1}^{\infty}\rho]})$ is characterized by

$$P_{\bigotimes_{k=1}^{\infty}\rho}(\Xi_1\times\Xi_2\times\cdots\times\Xi_n\times(\mathop{\times}_{k=n+1}^{\infty}X)) = \mathop{\times}_{k=1}^{n}{}_{\mathcal{T}r(H)}\Big(\rho, F(\Xi_k)\Big)_{B(H)} \tag{4.7}$$
$$(\ \forall \Xi_k \in \mathcal{F}=\mathcal{F}_\rho, (\ k=1,2,\ldots,n), n=1,2,3\cdots)$$

which is equal to the infinite product probability measure $\bigotimes_{k=1}^{n}P_\rho$.

[II]: classical system: Without loss of generality, we assume that the state space Ω is compact, and $\nu(\Omega)=1$ (*cf.* Note 2.1). Then, the classical infinite tensor basic structure is defined by

$$[C_0(\times_{k=1}^{\infty}\Omega) \subseteq L^\infty(\times_{k=1}^{\infty}\Omega, \otimes_{k=1}^{\infty}\nu) \subseteq B(L^2(\times_{k=1}^{\infty}\Omega, \otimes_{k=1}^{\infty}\nu))] \tag{4.8}$$

Therefore, the infinite tensor state space is characterized by

$$\mathfrak{S}^p(C_0(\times_{k=1}^{\infty}\Omega)^*)\Big(\approx \mathop{\times}_{k=1}^{\infty}\Omega\Big) \tag{4.9}$$

Put $\rho = \delta_\omega$. the sample probability space $(X^{\mathbb{N}},\ \boxtimes_{k=1}^{\infty}\mathcal{F},\ P_{\bigotimes_{k=1}^{\infty}\rho})$ of the infinite parallel measurement $\mathsf{M}_{L^\infty(\times_{k=1}^{\infty}\Omega, \otimes_{k=1}^{\infty}\nu)}(\otimes_{k=1}^{\infty}\mathsf{O} = (X^{\mathbb{N}},\ \boxtimes_{k=1}^{\infty}\mathcal{F}, \otimes k = 1^{\infty}F), S_{[\bigotimes_{k=1}^{\infty}\rho]})$ is characterized by

$$P_{\bigotimes_{k=1}^{\infty}\rho}(\Xi_1 \times \Xi_2 \times \cdots \times \Xi_n \times (\bigtimes_{k=n+1}^{\infty} X)) = \bigtimes_{k=1}^{n}[F(\Xi_k)](\omega) \tag{4.10}$$

$$(\ \forall \Xi_k \in \mathcal{F} = \mathcal{F}_\rho, (\ k = 1, 2, \ldots, n), n = 1, 2, 3 \cdots)$$

which is equal to the infinite product probability measure $\bigotimes_{k=1}^{n} P_\rho$.
[III]: Conclusion: Therefore, we can conclude

(♯) **in both cases, the sample probability space $(X^{\mathbb{N}},\ \boxtimes_{k=1}^{\infty}\mathcal{F},\ P_{\bigotimes_{k=1}^{\infty}\rho})$ is defined by the infinite product probability space $(X^{\mathbb{N}},\ \boxtimes_{k=1}^{\infty}\mathcal{F},\ \bigotimes_{k=1}^{\infty} P_\rho)$**

Summing up, we have the following theorem (the law of large numbers).

Theorem 4.5. [The law of large numbers] Consider the measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O} = (X, \mathcal{F}, F), S_{[\rho]})$ with the sample probability space $(X, \mathcal{F}, P_\rho)$. Then, by Kolmogorov's extension theorem (Corollary4.2), we have the infinite parallel measurement:

$$\mathsf{M}_{\bigotimes_{k=1}^{\infty}\overline{\mathcal{A}}}(\otimes_{k=1}^{\infty}\mathsf{O} = (X^{\mathbb{N}},\ \boxtimes_{k=1}^{\infty}\mathcal{F}, \otimes_{k=1}^{\infty}F), S_{[\bigotimes_{k=1}^{\infty}\rho]})$$

The sample probability space $(X^{\mathbb{N}},\ \boxtimes_{k=1}^{\infty}\mathcal{F},\ P_{\bigotimes_{k=1}^{\infty}\rho})$ is characterized by the infinite probability space $(X^{\mathbb{N}},\ \boxtimes_{k=1}^{\infty}\mathcal{F},\ \bigotimes_{k=1}^{\infty} P_\rho)$. Further, we see

(A) for any $f \in L^1(X, P_\rho)$, put

$$D_f = \Big\{(x_1, x_2, \ldots) \in X^{\mathbb{N}} \mid \lim_{n\to\infty} \frac{f(x_1) + f(x_2) + \cdots + f(x_n)}{n} = E(f)\Big\} \tag{4.11}$$

$$(\text{ where, } \ E(f) = \int_X f(x) P_\rho(dx)\ )$$

Then, it holds that

$$P_{\bigotimes_{k=1}^{\infty}\rho}(D_f) = 1 \tag{4.12}$$

That is, we see, almost surely,

$$\underset{\text{(population mean)}}{\boxed{\int_X f(x) P_\rho(dx)}} = \underset{\text{(sample mean)}}{\boxed{\lim_{n\to\infty} \frac{f(x_1)+f(x_2)+\cdots+f(x_n)}{n}}} \tag{4.13}$$

Remark 4.6. [Frequency probability] In the above, consider the case that

$$f(x) = \chi_\Xi(x) = \begin{cases} 1 & (x \in \Xi) \\ 0 & (x \notin \Xi) \end{cases} \qquad (\Xi \in \mathcal{F})$$

Then, put

$$D_{\chi_\Xi} = \Big\{ (x_1, x_2, \ldots) \in X^{\mathbb{N}} \mid \lim_{n\to\infty} \frac{\sharp[\{k \mid x_k \in \Xi, 1 \le k \le n\}}{n} = P_\rho(\Xi) \Big\} \tag{4.14}$$

(where, $\sharp[A]$ is the number of the elements of the set A)

Then, it holds that

$$P_{\bigotimes_{k=1}^\infty \rho}(D_{\chi_\Xi}) = 1 \tag{4.15}$$

Therefore, the law of large numbers (Theorem 4.5) says that

($\sharp_1$) the probability in Axiom 1 (§2.7) can be regarded as "frequency probability"

Thus, we have the following opinion:

$$(\sharp_2) \begin{cases} \text{G. Galileo} & \cdots \text{the originator of the realistic world view} \\ \\ \text{J. Bernoulli} & \cdots \text{the originator of the linguistic world view} \end{cases}$$

4.2.2 Mean, variance, unbiased variance

Consider the measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O} = (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, F), S_{[\rho]})$. Let $(\mathbb{R}, \mathcal{B}_{\mathbb{R}}, P_\rho)$ be its sample probability space. That is, consider the case that a measured value space $X = \mathbb{R}$.

Here, define:

$$\textbf{population mean}(\mu_{\mathsf{O}}^\rho) : E[\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O} = (\mathbb{R}, \mathcal{B}_{\mathbb{R}} F), S_{[\rho]})] = \int_{\mathbb{R}} x P_\rho(dx) (= \mu) \tag{4.16}$$

$$\textbf{population variance}((\sigma_{\mathsf{O}}^\rho)^2) : \quad V[\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O} = (\mathbb{R}, \mathcal{B}_{\mathbb{R}} F), S_{[\rho]})] = \int_{\mathbb{R}} (x-\mu)^2 P_\rho(dx) \tag{4.17}$$

Assume that a measured value $(x_1, x_2, x_3, ..., x_n) (\in \mathbb{R}^n)$ is obtained by the parallel measurement $\otimes_{k=1}^n \mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}, S_{[\rho]})$. Put

$$\textbf{sample distribution}(\nu_n) : \nu_n = \frac{\delta_{x_1} + \delta_{x_2} + \cdots + \delta_{x_n}}{n} \in \mathcal{M}_{+1}(X)$$

$$\begin{aligned} \textbf{sample mean}(\overline{\mu}_n) : \overline{E}[\otimes_{k=1}^n \mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}, S_{[\rho]})] &= \frac{x_1 + x_2 + \cdots + x_n}{n} (= \overline{\mu}) \\ &= \int_{\mathbb{R}} x \nu_n(dx) \end{aligned}$$

$$\begin{aligned} \textbf{sample variance}(s_n^2) : \overline{V}[\otimes_{k=1}^n \mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}, S_{[\rho]})] &= \frac{(x_1 - \overline{\mu})^2 + (x_2 - \overline{\mu})^2 + \cdot + (x_2 - \overline{\mu})^2}{n} \\ &= \int_{\mathbb{R}} (x - \overline{\mu})^2 \nu_n(dx) \end{aligned}$$

$$\begin{aligned} \textbf{unbiased variance}(u_n^2) : \overline{U}[\otimes_{k=1}^n \mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}, S_{[\rho]})] &= \frac{(x_1 - \overline{\mu})^2 + (x_2 - \overline{\mu})^2 + \cdot + (x_2 - \overline{\mu})^2}{n-1} \\ &= \frac{n}{n-1} \int_{\mathbb{R}} (x - \overline{\mu})^2 \nu_n(dx) \end{aligned}$$

Under the above preparation, we have:

Theorem 4.7. [Population mean, population variance, sample mean, sample variance] Assume that a measured value $(x_1, x_2, x_3, \cdots)(\in \mathbb{R}^{\mathbb{N}})$ is obtained by the infinite parallel measurement $\bigotimes_{k=1}^{\infty} \mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O} = (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, F), S_{[\rho]})$. Then, the law of large numbers (Theorem4.5) says that

$$
\begin{aligned}
(4.16) = \text{population mean}(\mu_{\mathsf{O}}^{\rho}) &= \lim_{n\to\infty} \frac{x_1 + x_2 + \cdots + x_n}{n} =: \overline{\mu} = \text{sample mean} \\
(4.17) = \text{population variance}(\sigma_{\mathsf{O}}^{\rho}) &= \lim_{n\to\infty} \frac{(x_1 - \mu_{\mathsf{O}}^{\rho})^2 + (x_2 - \mu_{\mathsf{O}}^{\rho})^2 + \cdots + (x_n - \mu_{\mathsf{O}}^{\rho})^2}{n} \\
&= \lim_{n\to\infty} \frac{(x_1 - \overline{\mu})^2 + (x_2 - \overline{\mu})^2 + \cdots + (x_n - \overline{\mu})^2}{n} =: \text{sample variance}
\end{aligned}
$$

Example 4.8. [Spectrum decomposition] Consider the quantum basic structure

$$[\mathcal{C}(H) \subseteq B(H) \subseteq B(H)]$$

Let A be a self-adjoint operator on H, which has the spectrum decomposition (i.e., projective observable) $\mathsf{O}_A = (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, F_A)$ such that

$$A = \int_{\mathbb{R}} \lambda F_A(d\lambda)$$

That is, under the identification:

$$\text{self-adjoint operator: } A \underset{\text{identification}}{\longleftrightarrow} \text{spectrum decomposition:} \mathsf{O}_A = (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, F_A)$$

the self-adjoint operator A is regarded as the projective observable $\mathsf{O}_A = (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, F_A)$. Fix the state $\rho_u = |u\rangle\langle u| \in \mathfrak{S}^p(\mathcal{T}r(H))$. Consider the measurement $\mathsf{M}_{B(H)}(\mathsf{O}_A, S_{[|u\rangle\langle u|]})$. Then, we see

$$\text{population mean}(\mu_{\mathsf{O}_A}^{\rho_u}) : E[\mathsf{M}_{B(H)}(\mathsf{O}_A, S_{[|u\rangle\langle u|]})] = \int_{\mathbb{R}} \lambda \langle u, F_A(d\lambda)u\rangle = \langle u, Au\rangle \tag{4.18}$$

$$
\begin{aligned}
\text{population variance}((\sigma_{\mathsf{O}_A}^{\rho_u})^2) : \quad V[\mathsf{M}_{B(H)}(\mathsf{O}_A, S_{[|u\rangle\langle u|]})] &= \int_{\mathbb{R}} (\lambda - \langle u, Au\rangle)^2 \langle u, F_A(d\lambda)u\rangle \\
&= \|(A - \langle u, Au\rangle)u\|^2
\end{aligned} \tag{4.19}
$$

4.2.3 Robertson's uncertainty principle

Now we can introduce Robertson's uncertainty principle as follows.

Theorem 4.9. [Robertson's uncertainty principle (parallel measurement) (*cf.* [65])] Consider the quantum basic structure $[\mathcal{C}(H) \subseteq B(H) \subseteq B(H)]$. Let A_1 and A_2 be unbounded self-adjoint operators on a Hilbert space H, which respectively has the spectrum decomposition:

$$\mathsf{O}_{A_1} = (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, F_{A_1}) \quad \text{to} \quad \mathsf{O}_{A_1} = (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, F_{A_1})$$

Thus, we have two measurements $\mathsf{M}_{B(H)}(\mathsf{O}_{A_1}, S_{[\rho_u]})$ and $\mathsf{M}_{B(H)}(\mathsf{O}_{A_2}, S_{[\rho_u]})$, where $\rho_u = |u\rangle\langle u| \in \mathfrak{S}^p(\mathcal{C}(H)^*)$. To take two measurements means to take the **parallel measurement**: $\mathsf{M}_{B(\mathbb{C}^n)}(\mathsf{O}_{A_1}, S_{[\rho_u]}) \otimes \mathsf{M}_{B(\mathbb{C}^n)}(\mathsf{O}_{A_2}, S_{[\rho_u]})$, namely,

$$\mathsf{M}_{B(H)\otimes B(H)}(\mathsf{O}_{A_1} \otimes \mathsf{O}_{A_2}, S_{[\rho_u \otimes \rho_u]})$$

Then, the following inequality (i.e., Robertson's uncertainty principle) holds that

$$\sigma_{A_1}^{\rho_u} \cdot \sigma_{A_2}^{\rho_u} \geqq \frac{1}{2}|\langle u, (A_1 A_2 - A_2 A_1)u\rangle| \qquad (\forall |u\rangle\langle u| = \rho_u, \ \ \|u\|_H = 1)$$

where $\sigma_{A_1}^{\rho_u}$ and $\sigma_{A_2}^{\rho_u}$ are shown in (4.19), namely,

$$\begin{cases} \sigma_{A_1}^{\rho_u} = [\langle A_1 u, A_1 u\rangle - |\langle u, A_1 u\rangle|^2]^{1/2} = \|(A_1 - \langle u, A_1 u\rangle)u\| \\ \sigma_{A_2}^{\rho_u} = [\langle A_2 u, A_2 u\rangle - |\langle u, A_2 u\rangle|^2]^{1/2} = \|(A_2 - \langle u, A_2 u\rangle)u\| \end{cases}$$

Therefore, putting $[A_1, A_2] \equiv A_1 A_2 - A_2 A_1$, we rewrite Robertson's uncertainty principle as follows:

$$\|A_1 u\| \cdot \|A_2 u\| \geq \|(A_1 - \langle u, A_1 u\rangle)u\| \cdot \|(A_2 - \langle u, A_2 u\rangle)u\| \geq |\langle u, [A_1, A_2]u\rangle|/2 \tag{4.20}$$

For example, when $A_1(= Q)$ [resp. $A_2(= P)$] is the position observable [resp. momentum observable] (i.e., $QP - PQ = \hbar\sqrt{-1}$), it holds that

$$\sigma_Q^{\rho_u} \cdot \sigma_P^{\rho_u} \geqq \frac{1}{2}\hbar$$

Proof. Robertson's uncertainty principle (4.20) is essentially the same as Schwarz inequality, that is,

$$\begin{aligned} & |\langle u, [A_1, A_2]u\rangle| = |\langle u, (A_1 A_2 - A_2 A_1)u\rangle| \\ =& \Big|\Big\langle u, \Big((A_1 - \langle u, A_1 u\rangle)(A_2 - \langle u, A_2 u\rangle) - (A_2 - \langle u, A_2 u\rangle)(A_1 - \langle u, A_1 u\rangle)\Big)u\Big\rangle\Big| \\ \leq& 2\|(A_1 - \langle u, A_1 u\rangle)u\| \cdot \|(A_2 - \langle u, A_2 u\rangle)u\| \end{aligned}$$

$\square$

4.3 Heisenberg's uncertainty principle

4.3.1 Why is Heisenberg's uncertainty principle famous?

Heisenberg's uncertainty principle is as follows.

Proposition 4.10. [Heisenberg's uncertainty principle (*cf.* [18]:1927)]

(i) The position x of a particle P can be measured exactly. Also similarly, the momentum p of a particle P can be measured exactly. However, the position x and momentum p of a particle P can not be measured simultaneously and exactly, namely, the both errors Δ_x and Δ_p can not be equal to 0. That is, the position x and momentum p of a particle P can be measured simultaneously and approximately,

(ii) And, Δ_x and Δ_p satisfy Heisenberg's uncertainty principle as follows.

$$\Delta_x \cdot \Delta_p \doteqdot \hbar (= \text{Plank constant}/2\pi \doteqdot 1.5547 \times 10^{-34} Js). \tag{4.21}$$

This was discovered by Heisenberg's thought experiment due to γ-ray microscope. It is

(A) **one of the most famous statements in the 20-th century.**

But, we think that it is doubtful in the following sense.

♠**Note 4.1.** I think, strictly speaking, that Heisenberg's uncertainty principle(Proposition 4.8) is meaningless. That is because, for example,

(♯) The approximate measurement and "error" in Proposition 4.8 are not defined.

This will be improved in Theorem 4.15 in the framework of quantum mechanics. That is, Heisenberg's thought experiment is an excellent idea before the discovery of quantum mechanics. Some may ask that

> If it be so, why is Heisenberg's uncertainty principle (Proposition 4.8) famous?

I think that

> Heisenberg's uncertainty principle (Proposition 4.8) was used as the slogan for advertisement of quantum mechanics in order to emphasize the difference between classical mechanics and quantum mechanics.

And, this slogan was completely successful. This kind of slogan is not rare in the history of science. For example, recall "cogito proposition (due to Descartes)", that is,

I think, therefore I am.

which is also meaningless (*cf.* §8.3). However, it is certain that the cogito proposition built the foundation of modern science.

♠**Note 4.2.** Heisenberg's uncertainty principle(Proposition 4.8) may include contradiction (*cf.* ref. [22]), if we think as follows

($\sharp$) it is "natural" to consider that

$$\Delta_x = |x - \widetilde{x}|, \quad \Delta_p = |p - \widetilde{p}|,$$

where

$$\begin{cases} \text{Position:} & [x : \text{exact measured value (=true value)}, \widetilde{x} : \text{measured value}] \\ \text{Momentum:} & [p : \text{exact measured value (=true value)}, \widetilde{p} : \text{measured value}] \end{cases}$$

However, this is in contradiction with Heisenberg's uncertainty principle (4.21). That is because (4.21) says that the exact measured value (x, p) can not be measured.

4.3.2 The mathematical formulation of Heisenberg's uncertainty principle

In this section, we shall propose the mathematical formulation of Heisenberg's uncertainty principle 4.10.

Consider the quantum basic structure:

$$[\mathcal{C}(H) \subseteq B(H) \subseteq B(H)]$$

Let A_i $(i = 1, 2)$ be arbitrary self-adjoint operator on H. For example, it may satisfy that

$$[A_1, A_2](:= A_1A_2 - A_2A_1) = \hbar\sqrt{-1}I$$

Let $\mathsf{O}_{A_i} = (\mathbb{R}, \mathcal{B}, F_{A_i})$ be the spectral representation of A_i, i.e., $A_i = \int_{\mathbb{R}} \lambda F_{A_i}(d\lambda)$, which is regarded as the projective observable in $B(H)$. Let $\rho_0 = |u\rangle\langle u|$ be a state, where $u \in H$ and $\|u\| = 1$. Thus, we have two measurements:

$$(\mathrm{B}_1) \ \ \mathsf{M}_{B(H)}(\mathsf{O}_{A_1} {:=} (\mathbb{R}, \mathcal{B}, F_{A_1}), \ S_{[\rho_u]}) \qquad \xrightarrow[\text{expectation}]{\text{by (4.18)}} \langle u, A_1 u\rangle$$

$$(\mathrm{B}_2) \ \ \mathsf{M}_{B(H)}(\mathsf{O}_{A_2} {:=} (\mathbb{R}, \mathcal{B}, F_{A_2}), \ S_{[\rho_u]}) \qquad \xrightarrow[\text{expectation}]{\text{by (4.18)}} \langle u, A_2 u\rangle$$

$$(\forall \rho_u = |u\rangle\langle u| \in \mathfrak{S}^p(\mathcal{C}(H)^*))$$

However, since it is not always assumed that $A_1A_2 - A_2A_1 = 0$, we can not expect the existence of the simultaneous observable $\mathsf{O}_{A_1} \times \mathsf{O}_{A_2}$, namely,

- **in general, two observables O_{A_1} and O_{A_2} can not be simultaneously measured**

That is,

(B$_3$) the measurement $\mathsf{M}_{B(H)}(\mathsf{O}_{A_1} \times \mathsf{O}_{A_2}, S_{[\rho_u]})$ is impossible, Thus, we have the question:

Then, what should be done?

In what follows, we shall answer this.

Let K be another Hilbert space, and let s be in K such that $\|s\| = 1$. Thus, we also have two observables $\mathsf{O}_{A_1 \otimes I} :=(\mathbb{R}, \mathcal{B}, F_{A_1} \otimes I)$ and $\mathsf{O}_{A_2 \otimes I} :=(\mathbb{R}, \mathcal{B}, F_{A_2} \otimes I)$ in the tensor algebra $B(H \otimes K)$.

Put

$$\text{the tensor state } \widehat{\rho}_{us} = |u \otimes s\rangle\langle u \otimes s|$$

And we have the following two measurements:

(C$_1$) $\mathsf{M}_{B(H\otimes K)}(\mathsf{O}_{A_1\otimes I}, S_{[\widehat{\rho}_{us}]}) \xrightarrow[\text{expectation}]{\text{by (4.18)}} \langle u \otimes s, (A_1 \otimes I)(u \otimes s)\rangle = \langle u, A_1 u\rangle$

(C$_2$) $\mathsf{M}_{B(H\otimes K)}(\mathsf{O}_{A_2\otimes I}, S_{[\widehat{\rho}_{us}]}) \xrightarrow[\text{expectation}]{\text{by (4.18)}} \langle u \otimes s, (A_2 \otimes I)(u \otimes s)\rangle = \langle u, A_2 u\rangle$

It is a matter of course that

$$(\text{C}_1)=(\text{B}_1) \qquad (\text{C}_2)=(\text{B}_2)$$

and

(C$_3$) $\mathsf{M}_{B(H\otimes K)}(\mathsf{O}_{A_1\otimes I} \times \mathsf{O}_{A_2\otimes I}, S_{[\widehat{\rho}_{us}]})$ is impossible.

Thus, overcoming this difficulty, we prepare the following idea:

Preparation 4.11. Let $\widehat{A}_i$ $(i = 1, 2)$ be arbitrary self-adjoint operator on the tensor Hilbert space $H \otimes K$, where it is assumed that

$$[\widehat{A}_1, \widehat{A}_2](:= \widehat{A}_1\widehat{A}_2 - \widehat{A}_2\widehat{A}_1) = 0 \quad \text{(i.e., the commutativity)} \tag{4.22}$$

Let $\mathsf{O}_{\widehat{A}_i} = (\mathbb{R}, \mathcal{B}, F_{\widehat{A}_i})$ be the spectral representation of $\widehat{A}_i$, i.e.$\widehat{A}_i = \int_{\mathbb{R}} \lambda F_{\widehat{A}_i}(d\lambda)$, which is regarded as the projective observable in $B(H \otimes K)$. Thus, we have two measurements as follows:

(D$_1$) $\mathsf{M}_{B(H\otimes K)}(\mathsf{O}_{\widehat{A}_1}, S_{[\widehat{\rho}_{us}]}) \xrightarrow[\text{expectation}]{\text{by (4.18)}} \langle u\otimes s, \widehat{A}_1(u\otimes s)\rangle$

(D$_2$) $\mathsf{M}_{B(H\otimes K)}(\mathsf{O}_{\widehat{A}_2}, S_{[\widehat{\rho}_{us}]}) \xrightarrow[\text{expectation}]{\text{by (4.18)}} \langle u\otimes s, \widehat{A}_2(u\otimes s)\rangle$

Note, by the commutative condition (4.22), that the two can be measured by the simultaneous measurement $\mathsf{M}_{B(H\otimes K)}(\mathsf{O}_{\widehat{A}_1}\times \mathsf{O}_{\widehat{A}_2}, S_{[\widehat{\rho}_{us}]})$, where $\mathsf{O}_{\widehat{A}_1}\times \mathsf{O}_{\widehat{A}_2} = (\mathbb{R}^2, \mathcal{B}^2, F_{\widehat{A}_1}\times F_{\widehat{A}_2})$.

Again note that any relation between $A_i\otimes I$ and $\widehat{A}_i$ is not assumed. However,

- we want to regard this simultaneous measurement as the substitute of the above two (C$_1$) and (C$_2$). That is, we want to regard

(D$_1$) and (D$_2$) as the substitute of (C$_1$) and (C$_2$)

For this, we have to prepare Hypothesis 4.9 below.

Putting

$$\widehat{N}_i := \widehat{A}_i - A_i\otimes I \quad (\text{and thus, } \widehat{A}_i = \widehat{N}_i + A_i\otimes I) \tag{4.23}$$

we define the $\Delta^{\widehat{\rho}_{us}}_{\widehat{N}_i}$ and $\overline{\Delta}^{\widehat{\rho}_{us}}_{\widehat{N}_i}$ such that

$$\begin{aligned}\Delta^{u\otimes s}_{\widehat{N}_i} =&\|\widehat{N}_i(u\otimes s)\| = \|(\widehat{A}_i - A_i\otimes I)(u\otimes s)\| \\ \overline{\Delta}^{u\otimes s}_{\widehat{N}_i} =&\|(\widehat{N}_i - \langle u\otimes s, \widehat{N}_i(u\otimes s)\rangle)(u\otimes s)\| \\ =&\|((\widehat{A}_i - A_i\otimes I) - \langle u\otimes s, (\widehat{A}_i - A_i\otimes I)(u\otimes s)\rangle)(u\otimes s)\|\end{aligned} \tag{4.24}$$

where the following inequality:

$$\Delta^{\widehat{\rho}_{us}}_{\widehat{N}_i} \geq \overline{\Delta}^{\widehat{\rho}_{us}}_{\widehat{N}_i} \tag{4.25}$$

is common sense.

By the commutative condition (4.22), (4.23) implies that

$$[\widehat{N}_1, \widehat{N}_2] + [\widehat{N}_1, A_2\otimes I] + [A_1\otimes I, \widehat{N}_2] = -[A_1\otimes I, A_2\otimes I] \tag{4.26}$$

Here, we should note that the first term (or, precisely, $|\langle u\otimes s, [\text{the first term}](u\otimes s)\rangle|$) of (4.26) can be, by the Robertson uncertainty relation (*cf.* Theorem4.9), estimated as follows:

$$2\overline{\Delta}^{\widehat{\rho}_{us}}_{\widehat{N}_1}\cdot \overline{\Delta}^{\widehat{\rho}_{us}}_{\widehat{N}_2} \geq |\langle u\otimes s, [\widehat{N}_1, \widehat{N}_2](u\otimes s)\rangle| \tag{4.27}$$

4.3.2.1 Average value coincidence conditions; approximately simultaneous measurement

However, it should be noted that

In the above, any relation between $A_i \otimes I$ and $\widehat{A}_i$ is not assumed.

Thus, we think that the following hypothesis is natural.

Hypothesis 4.12. [Average value coincidence conditions]. We assume that

$$\langle u \otimes s, \widehat{N}_i(u \otimes s) \rangle = 0 \qquad (\forall u \in H, i = 1, 2) \tag{4.28}$$

or equivalently,

$$\langle u \otimes s, \widehat{A}_i(u \otimes s) \rangle = \langle u, A_i u \rangle \qquad (\forall u \in H, i = 1, 2) \tag{4.29}$$

That is,

$$\begin{aligned}
&\text{the average measured value of } \mathsf{M}_{B(H\otimes K)}(\mathsf{O}_{\widehat{A}_i}, S_{[\widehat{\rho}_{us}]}) \\
=&\langle u \otimes s, \widehat{A}_i(u \otimes s) \rangle \\
=&\langle u, A_i u \rangle \\
=&\text{the average measured value of } \mathsf{M}_{B(H)}(\mathsf{O}_{A_i}, S_{[\rho_u]}) \\
&\quad (\forall u \in H, ||u||_H = 1, i = 1, 2)
\end{aligned}$$

Hence, we have the following definition.

Definition 4.13. [Approximately simultaneous measurement] Let A_1 and A_2 be (unbounded) self-adjoint operators on a Hilbert space H. The quartet $(K, s, \widehat{A}_1, \widehat{A}_2)$ is called an approximately simultaneous observable of A_1 and A_2, if it satisfied that

(E$_1$) K is a Hilbert space. $s \in K$, $\|s\|_K = 1$, $\widehat{A}_1$ and $\widehat{A}_2$ are commutative self-adjoint operators on a tensor Hilbert space $H \otimes K$ that satisfy the average value coincidence condition (4.28), that is,

$$\langle u \otimes s, \widehat{A}_i(u \otimes s) \rangle = \langle u, A_i u \rangle \qquad (\forall u \in H, i = 1, 2) \tag{4.30}$$

Also, the measurement $\mathsf{M}_{B(H\otimes K)}(\mathsf{O}_{\widehat{A}_1} \times \mathsf{O}_{\widehat{A}_2}, S_{[\widehat{\rho}_{us}]})$ is called the approximately simultaneous measurement of $\mathsf{M}_{B(H)}(\mathsf{O}_{A_1}, S_{[\rho_u]})$ and $\mathsf{M}_{B(H)}(\mathsf{O}_{A_2}, S_{[\rho_u]})$.
Thus, under the average coincidence condition, we regard

(D$_1$) and (D$_2$) as the substitute of (C$_1$) and (C$_2$)

And

(E$_2$) $\Delta^{\widehat{\rho}_{us}}_{\widehat{N}_1}$ $(= \|(\widehat{A}_1 - A_1 \otimes I)(u \otimes s)\|)$ and $\Delta^{\widehat{\rho}_{us}}_{\widehat{N}_2}$ $(= \|(\widehat{A}_2 - A_2 \otimes I)(u \otimes s)\|)$ are called errors of the approximate simultaneous measurement measurement $\mathsf{M}_{B(H \otimes K)}(\mathsf{O}_{\widehat{A}_1} \times \mathsf{O}_{\widehat{A}_2}, S_{[\widehat{\rho}_{us}]})$

Lemma 4.14. Let A_1 and A_2 be (unbounded) self-adjoint operators on a Hilbert space H. And let $(K, s, \widehat{A}_1, \widehat{A}_2)$ be an approximately simultaneous observable of A_1 and A_2. Then, it holds that

$$\Delta^{\widehat{\rho}_{us}}_{\widehat{N}_i} = \overline{\Delta}^{\widehat{\rho}_{us}}_{\widehat{N}_i} \tag{4.31}$$

$$\langle u \otimes s, [\widehat{N}_1, A_2 \otimes I](u \otimes s)\rangle = 0 \qquad (\forall u \in H) \tag{4.32}$$

$$\langle u \otimes s, [A_1 \otimes I, \widehat{N}_2](u \otimes s)\rangle = 0 \quad (\forall u \in H) \tag{4.33}$$

The proof is easy, thus, we omit it.

Under the above preparations, we can easily get "Heisenberg's uncertainty principle" as follows.

$$\Delta^{\widehat{\rho}_{us}}_{\widehat{N}_1} \cdot \Delta^{\widehat{\rho}_{us}}_{\widehat{N}_2} (= \overline{\Delta}^{\widehat{\rho}_{us}}_{\widehat{N}_1} \cdot \overline{\Delta}^{\widehat{\rho}_{us}}_{\widehat{N}_2}) \geq \frac{1}{2}|\langle u, [A_1, A_2]u\rangle| \quad (\forall u \in H \text{ such that } ||u|| = 1) \tag{4.34}$$

Summing up, we have the following theorem:

Theorem 4.15. [The mathematical formulation of Heisenberg's uncertainty principle] Let A_1 and A_2 be (unbounded) self-adjoint operators on a Hilbert space H. Then. we have the followings:

(i) There exists an approximately simultaneous observable$(K, s, \widehat{A}_1, \widehat{A}_2)$ of A_1 and A_2, that is, $s \in K$, $\|s\|_K = 1$, $\widehat{A}_1$ and $\widehat{A}_2$ are commutative self-adjoint operators on a tensor Hilbert space $H \otimes K$ that satisfy the average value coincidence condition (4.28). Therefore, the approximately simultaneous measurement $\mathsf{M}_{B(H \otimes K)}(\mathsf{O}_{\widehat{A}_1} \times \mathsf{O}_{\widehat{A}_2}, S_{[\widehat{\rho}_{us}]})$ exists.

(ii) And further, we have the following inequality (i.e., Heisenberg's uncertainty principle).

$$\begin{aligned}\Delta^{\widehat{\rho}_{us}}_{\widehat{N}_1} \cdot \Delta^{\widehat{\rho}_{us}}_{\widehat{N}_2} (= \overline{\Delta}^{\widehat{\rho}_{us}}_{\widehat{N}_1} \cdot \overline{\Delta}^{\widehat{\rho}_{us}}_{\widehat{N}_2}) &= \|(\widehat{A}_1 - A_1 \otimes I)(u \otimes s)\| \cdot \|(\widehat{A}_2 - A_2 \otimes I)(u \otimes s)\| \\ &\geq \frac{1}{2}|\langle u, [A_1, A_2]u\rangle| \quad (\forall u \in H \text{ such that } ||u|| = 1)\end{aligned} \tag{4.35}$$

(iii) In addition, if $A_1 A_2 - A_2 A_1 = \hbar\sqrt{-1}$, we see that

$$\Delta^{\widehat{\rho}_{us}}_{\widehat{N}_1} \cdot \Delta^{\widehat{\rho}_{us}}_{\widehat{N}_2} \geq \hbar/2 \quad (\forall u \in H \text{ such that } ||u|| = 1) \tag{4.36}$$

Proof. For the proof of (i) and (ii), see

- Ref. [22]: S. Ishikawa, Rep. Math. Phys. Vol.29(3), 1991, pp.257–273,

As shown in the above (4.34), the proof (ii) is easy (*cf.* [29, 62]), but the proof (i) is not easy (*cf.* [7, 29]).

4.3.3 Without the average value coincidence condition

Now we have the complete form of Heisenberg's uncertainty relation as Theorem 4.15, To be compared with Theorem 4.15, we should note that the conventional Heisenberg's uncertainty relation (= Proposition 4.10) is ambiguous. Wrong conclusions are sometimes derived from the ambiguous statement (= Proposition 4.10). For example, in some books of physics, it is concluded that EPR-experiment (Einstein, Podolosky and Rosen [13], or, see the following section) conflicts with Heisenberg's uncertainty relation. That is,

[I] Heisenberg's uncertainty relation says that the position and the momentum of a particle can not be measured simultaneously and exactly.

On the other hand,

[II] EPR-experiment says that the position and the momentum of a certain "particle"can be measured simultaneously and exactly (Also, see Note 4.4.)

Thus someone may conclude that the above [I] and [II] includes a paradox, and therefore, EPR-experiment is in contradiction with Heisenberg's uncertainty relation. Of course, this is a misunderstanding. This "paradox"was solved in [22, 29]. Now we shall explain the solution of the paradox.

[Concerning the above [I]] Put $H = L^2(\mathbb{R}_q)$. Consider two-particles system in $H \otimes H = L^2(\mathbb{R}^2_{(q_1,q_2)})$. In the EPR problem, we, for example, consider the state u_e ($\in H \otimes H = L^2(\mathbb{R}^2_{(q_1,q_2)})$) $\Big($or precisely, $|u_e\rangle\langle u_e|\Big)$ such that:

$$u_e(q_1, q_2) = \sqrt{\frac{1}{2\pi\epsilon\sigma}} e^{-\frac{1}{8\sigma^2}(q_1-q_2-a)^2-\frac{1}{8\epsilon^2}(q_1+q_2-b)^2} \cdot e^{i\phi(q_1,q_2)} \tag{4.37}$$

where ϵ is assumed to be a sufficiently small positive number and $\phi(q_1, q_2)$ is a real-valued function. Let $A_1 \colon L^2(\mathbb{R}^2_{(q_1,q_2)}) \to L^2(\mathbb{R}^2_{(q_1,q_2)})$ and $A_2 \colon L^2(\mathbb{R}^2_{(q_1,q_2)}) \to L^2(\mathbb{R}^2_{(q_1,q_2)})$ be (unbounded) self-adjoint operators such that

$$A_1 = q_1, \qquad A_2 = \frac{\hbar\partial}{i\partial q_1}. \tag{4.38}$$

Then, Theorem 4.15 says that there exists an approximately simultaneous observable$(K, s, \widehat{A}_1, \widehat{A}_2)$ of A_1 and A_2. And thus, the following Heisenberg's uncertainty relation (= Theorem 4.15) holds,

$$\|\widehat{A}_1 u_e - A_1 u_e\| \cdot \|\widehat{A}_2 u_e - A_2 u_e\| \geq \hbar/2 \tag{4.39}$$

[Concerning the above [II]] However, it should be noted that, in the above situation we assume that the state u_e is known before the measurement. In such a case, we may take another measurement as follows: Put $K = \mathbb{C}$, $s = 1$. Thus, $(H \otimes H) \otimes K = H \otimes H$, $u \otimes s = u \otimes 1 = u$. Define the self-adjoint operators $\widehat{A}_1 : L^2(\mathbb{R}^2_{(q_1,q_2)}) \to L^2(\mathbb{R}^2_{(q_1,q_2)})$ and $\widehat{A}_2 : L^2(\mathbb{R}^2_{(q_1,q_2)}) \to L^2(\mathbb{R}^2_{(q_1,q_2)})$ such that

$$\widehat{A}_1 = b - q_2, \qquad \widehat{A}_2 = A_2 = \frac{\hbar\partial}{i\partial q_1} \tag{4.40}$$

Note that these operators commute. Therefore,

($\sharp$) we can take an exact simultaneous measurement of $\widehat{A}_1$ and $\widehat{A}_2$ (for the state u_e).

And moreover, we can easily calculate as follows:

$$\begin{aligned}
&\|\widehat{A}_1 u_e - A_1 u_e\| \\
=&\Big[\iint_{\mathbb{R}^2} \Big| ((b-q_2)-q_1) \sqrt{\frac{1}{2\pi\epsilon\sigma}} e^{-\frac{1}{8\sigma^2}(q_1-q_2-a)^2 - \frac{1}{8\epsilon^2}(q_1+q_2-b)^2} \cdot e^{i\phi(q_1,q_2)} \Big|^2 dq_1 dq_2 \Big]^{1/2} \\
=&\Big[\iint_{\mathbb{R}^2} \Big| ((b-q_2)-q_1) \sqrt{\frac{1}{2\pi\epsilon\sigma}} e^{-\frac{1}{8\sigma^2}(q_1-q_2-a)^2 - \frac{1}{8\epsilon^2}(q_1+q_2-b)^2} \Big|^2 dq_1 dq_2 \Big]^{1/2} \\
=&\sqrt{2}\epsilon,
\end{aligned} \tag{4.41}$$

and

$$\|\widehat{A}_2 u_e - A_2 u_e\| = 0. \tag{4.42}$$

Thus we see

$$\|\widehat{A}_1 u_e - A_1 u_e\| \cdot \|\widehat{A}_2 u_e - A_2 u_e\| = 0. \tag{4.43}$$

However it should be again noted that, the measurement ($\sharp$) is made from the knowledge of the state u_e.

[[I] and [II] are consistent] The above conclusion (4.43) does not contradict Heisenberg's uncertainty relation (4.39), since the measurement ($\sharp$) is not an approximate simultaneous measurement of A_1 and A_2. In other words, the $(K, s, \widehat{A}_1, \widehat{A}_2)$ is not an approximately simultaneous observable of A_1 and A_2. Therefore, we can conclude that

(F) Heisenberg's uncertainty principle is violated without the average value coincidence condition

(*cf.* Remark 3 in ref.[22], or p.316 in [29]).

♠**Note 4.3.** Some may consider that the formulas (4.41) and (4.42) imply that the statement [II] is true. However, it is not true. This is answered in Remark 8.15.

Also, we add the following remark.

Remark 4.16. Calculating the second term (precisely , $\langle u\otimes s,$"the second term"$(u\otimes s)\rangle$) and the third term (precisely , $\langle u \otimes s,$"the third term"$(u \otimes s)\rangle$) in (4.26), we get, by Robertson's uncertainty principle (4.20),

$$
\begin{aligned}
&2\overline{\Delta}^{\widehat{\rho}_{us}}_{\widehat{N}_1} \cdot \sigma(A_2;u) \geq |\langle u\otimes s, [\widehat{N}_1, A_2\otimes I](u\otimes s)\rangle| && (4.44)\\
&2\overline{\Delta}^{\widehat{\rho}_{us}}_{\widehat{N}_2} \cdot \sigma(A_1;u) \geq |\langle u\otimes s, [A_{\otimes} I, \widehat{N}_2](u\otimes s)\rangle| && (4.45)\\
&\qquad (\forall u \in H \text{ such that } ||u||=1)
\end{aligned}
$$

and, from (4.26), (4.27), (4.44),(4.45), we can get the following inequality

$$
\begin{aligned}
&\Delta^{\widehat{\rho}_{us}}_{\widehat{N}_1} \cdot \Delta^{\widehat{\rho}_{us}}_{\widehat{N}_2} + \Delta^{\widehat{\rho}_{us}}_{\widehat{N}_2} \cdot \sigma(A_1;u) + \Delta^{\widehat{\rho}_{us}}_{\widehat{N}_1} \cdot \sigma(A_2;u)\\
\geq&\overline{\Delta}^{\widehat{\rho}_{us}}_{\widehat{N}_1} \cdot \overline{\Delta}^{\widehat{\rho}_{us}}_{\widehat{N}_2} + \overline{\Delta}^{\widehat{\rho}_{us}}_{\widehat{N}_2} \cdot \sigma(A_1;u) + \overline{\Delta}^{\widehat{\rho}_{us}}_{\widehat{N}_1} \cdot \sigma(A_2;u)\\
\geq&\frac{1}{2}|\langle u, [A_1, A_2]u\rangle| \quad (\forall u \in H \text{ such that } ||u||=1) && (4.46)
\end{aligned}
$$

Since we do not assume the average value coincidence condition, it is a matter of course that this (4.46) is more rough than Heisenberg's uncertainty principle (4.35)

If a certain interpretation is adopted such that $\Delta^{\widehat{\rho}_{us}}_{\widehat{N}_1}$ and $\Delta^{\widehat{\rho}_{us}}_{\widehat{N}_2}$ mean "error:$\epsilon(A_1,u)$" and "disturbance:$\eta(A_2,u)$", respectively, then the inequality (4.46), i.e.,

$$
\epsilon(A_1,u)\eta(A_2,u) + \epsilon(A_1,u)\sigma(A_2,u) + \sigma(A_1,u)\eta(A_2,u) \geq \frac{1}{2}|\langle u, [A_1, A_2]u\rangle|
$$

is called Ozawa's inequality (*cf.* [63]). He asserted that this inequality is a faithful description of Heisenberg's thought experiment (due to γ-ray microscope).

4.4 EPR-paradox (1935) and faster-than-light

4.4.1 EPR-paradox

Next, let us explain EPR-paradox (Einstein–Poolside–Rosen: [13, 68]). Consider Two electrons P_1 and P_2 and their spins. The tensor Hilbert space $H = \mathbb{C}^2 \otimes \mathbb{C}^2$ is defined in what follows. That is,

$$e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

(i.e., the complete orthonormal system $\{e_1, e_2\}$ in the $\mathbb{C}^2$),

$$\mathbb{C}^2 \otimes \mathbb{C}^2 = \{ \sum_{i,j=1,2} \alpha_{ij} e_i \otimes e_j \mid \alpha_{ij} \in \mathbb{C}, i, j = 1, 2\}$$

Put $u = \sum_{i,j=1,2} \alpha_{ij} e_i \otimes e_j$ and $v = \sum_{i,j=1,2} \beta_{ij} e_i \otimes e_j$. And the inner product $\langle u, v \rangle_{\mathbb{C}^2 \otimes \mathbb{C}^2}$ is defined by

$$\langle u, v \rangle_{\mathbb{C}^2 \otimes \mathbb{C}^2} = \sum_{i,j=1,2} \overline{\alpha}_{i,j} \cdot \beta_{i,j}$$

Therefore, we have the tensor Hilbert space $H = \mathbb{C}^2 \otimes \mathbb{C}^2$ with the complete orthonormal system $\{e_1 \otimes e_1, e_1 \otimes e_2, e_2 \otimes e_1, e_2 \otimes e_2\}$.

For each $F \in B(\mathbb{C}^2)$ and $G \in B(\mathbb{C}^2)$, define the $F \otimes G \in B(\mathbb{C}^2 \otimes \mathbb{C}^2)$ (i.e., linear operator $F \otimes G : \mathbb{C}^2 \otimes \mathbb{C}^2 \to \mathbb{C}^2 \otimes \mathbb{C}^2$) such that

$$(F \otimes G)(u \otimes v) = Fu \otimes Gv$$

Let us define the entangled state $\rho = |s\rangle\langle s|$ of two particles P_1 and P_2 such that

$$s = \frac{1}{\sqrt{2}}(e_1 \otimes e_2 - e_2 \otimes e_1)$$

Here, we see that $\langle s, s \rangle_{\mathbb{C}^2 \otimes \mathbb{C}^2} = \frac{1}{2}\langle e_1 \otimes e_2 - e_2 \otimes e_1, e_1 \otimes e_2 - e_2 \otimes e_1 \rangle_{\mathbb{C}^2 \otimes \mathbb{C}^2} = \frac{1}{2}(1 + 1) = 1$, and thus, ρ is a state. Also, assume that

two particles P_1 and P_2 are far.

Let $\mathsf{O} = (X, 2^X, F^z)$ in $B(\mathbb{C}^2)$ (where $X = \{\uparrow, \downarrow\}$) be the spin observable concerning the z-axis such that

$$F^z(\{\uparrow\}) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad F^z(\{\downarrow\}) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

The parallel observable $\mathsf{O} \otimes \mathsf{O} = (X^2, 2^X \times 2^X, F^z \otimes F^z)$ in $B(\mathbb{C}^2 \otimes \mathbb{C}^2)$ is defined by

$$
\begin{aligned}
(F^z \otimes F^z)(\{(\uparrow,\uparrow)\}) &= F^z(\{\uparrow\}) \otimes F^z(\{\uparrow\}) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \\
(F^z \otimes F^z)(\{(\downarrow,\uparrow)\}) &= F^z(\{\downarrow\}) \otimes F^z(\{\uparrow\}) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \\
(F^z \otimes F^z)(\{(\uparrow,\downarrow)\}) &= F^z(\{\uparrow\}) \otimes F^z(\{\downarrow\}) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \otimes \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \\
(F^z \otimes F^z)(\{(\downarrow,\downarrow)\}) &= F^z(\{\downarrow\}) \otimes F^z(\{\downarrow\}) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}
\end{aligned}
$$

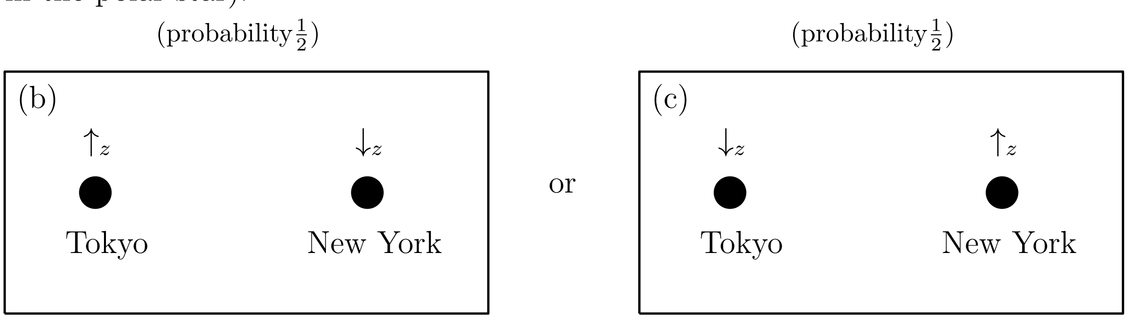
Thus, we get the measurement $\mathsf{M}_{B(\mathbb{C}^2 \otimes \mathbb{C}^2)}(\mathsf{O} \otimes \mathsf{O}, S_{[\rho]})$ The, Born's quantum measurement theory says that

> When the parallel measurementmeasurement $\mathsf{M}_{B(\mathbb{C}^2 \otimes \mathbb{C}^2)}(\mathsf{O} \otimes \mathsf{O}, S_{[s]})$ is taken,
>
> the probability that the measured value $\begin{bmatrix} (\uparrow,\uparrow) \\ (\downarrow,\uparrow) \\ (\uparrow,\downarrow) \\ (\downarrow,\downarrow) \end{bmatrix}$ is obtained
>
> is given by $\begin{bmatrix} \langle s, (F^z \otimes F^z)(\{(\uparrow,\uparrow)\})s \rangle_{\mathbb{C}^2 \otimes \mathbb{C}^2} = 0 \\ \langle s, (F^z \otimes F^z)(\{(\downarrow,\uparrow)\})s \rangle_{\mathbb{C}^2 \otimes \mathbb{C}^2} = 0.5 \\ \langle s, (F^z \otimes F^z)(\{(\uparrow,\downarrow)\})s \rangle_{\mathbb{C}^2 \otimes \mathbb{C}^2} = 0.5 \\ \langle s, (F^z \otimes F^z)(\{(\downarrow,\downarrow)\})s \rangle_{\mathbb{C}^2 \otimes \mathbb{C}^2} = 0 \end{bmatrix}$

That is because, $F^z(\{\uparrow\})e_1 = e_1$, $F^z(\{\downarrow\})e_2 = e_2$, $F^z(\{\uparrow\})e_2 = F^z(\{\downarrow\})e_1 = 0$ For example,

$$
\begin{aligned}
&\langle s, (F^z \otimes F^z)(\{(\uparrow,\downarrow)\})s \rangle_{\mathbb{C}^2 \otimes \mathbb{C}^2} \\
=&\frac{1}{2} \langle (e_1 \otimes e_2 - e_2 \otimes e_1), (F^z(\{\uparrow\}) \otimes F^z(\{\downarrow\})(e_1 \otimes e_2 - e_2 \otimes e_1) \rangle_{\mathbb{C}^2 \otimes \mathbb{C}^2} \\
=&\frac{1}{2} \langle (e_1 \otimes e_2 - e_2 \otimes e_1), e_1 \otimes e_2 \rangle_{\mathbb{C}^2 \otimes \mathbb{C}^2} = \frac{1}{2}
\end{aligned}
$$

Here, it should be noted that we can assume that the x_1 and the x_2 (in $(x_1, x_2) \in \{ (\uparrow_z, \uparrow_z), (\uparrow_z, \downarrow_z), (\downarrow_z, \uparrow_z), (\downarrow_z, \downarrow_z)\}$) are respectively obtained in Tokyo and in New York (or, in the earth and in the polar star).



This fact is, figuratively speaking, explained as follows:

- Immediately after the particle in Tokyo is measured and the measured value $\uparrow_z$ [resp. $\downarrow_z$] is observed, the particle in Tokyo informs the particle in New York "Your measured value has to be $\downarrow_z$ [resp. $\uparrow_z$]."

Therefore, the above fact implies that quantum mechanics says that *there is something faster than light.* This is essentially the same as *the de Broglie paradox* (*cf.* [68]). That is,

- if we admit quantum mechanics, we must also admit the fact that there is something faster than light (i.e., so called "non-locality").

♠**Note 4.4.** EPR-paradox is closely related to the fact that quantum syllogism does not hold in general. This will be discussed in Chapter 8. The Bohr-Einstein debates were a series of public disputes about quantum mechanics between Albert Einstein and Niels Bohr. Although there may be several opinions, I regard this debates as

$$\underset{\text{(realistic view)}}{\boxed{\text{Einstein}}} \underset{\text{v.s.}}{\longleftrightarrow} \underset{\text{(linguistic view)}}{\boxed{\text{Bohr}}}$$

For the further argument, see Section 10.7 (Leibniz-Clarke debates).

4.5 Bell's inequality(1966)

4.5.1 Bell's inequality is violated in classical and quantum systems

J. Bell's inequality is important in the relation of "the hidden variable". J. Bell showed that, if Bell's inequality is violated, then the hidden variable does not exist. However, it shoulde be noted that even if Bell's inequality is violated, it does not imply that quantum mechanics is wrong. In this section I would like to mention some of the things about Bell's inequality, though I am not concerned with "the hidden variable".

Firstly, let us mention Bell's inequality in mathematics.[1]

Theorem 4.17. [Bell's inequality] Let $(Y, \mathcal{G}, \mu)$ be a probability space. Consider measurable functions $f_k : Y \to \{-1, 1\}$, $(k = 1, 2, 3, 4)$, and define the correlations: $C_{13} = \int_Y f_1(y) \cdot f_3(y)\mu(dy)$, $C_{14} = \int_Y f_1(y) \cdot f_4(y)\mu(dy)$, $C_{23} = \int_Y f_2(y) \cdot f_3(y)\mu(dy)$, $C_{24} = \int_Y f_2(y) \cdot f_4(y)\mu(dy)$. Then, we have *Bell's inequality* such that

$$|C_{13} - C_{14}| + |C_{23} + C_{24}| \leqq 2. \tag{4.47}$$

Proof. It is easy as follows.

$$\begin{aligned} & |C_{13} - C_{14}| + |C_{23} + C_{24}| \\ \le & \int_Y f_1(y) \cdot |f_3(y) - f_4(y)|\mu(dy) + \int_Y f_2(y) \cdot |f_3(y) + f_4(y)|\mu(dy) = 2. \end{aligned}$$

□

The purpose of this section is

- to discuss the classical version of Aspects' experiment (which proves that "Bell's inequality" is violated).

Here, let us prepare three steps (I∼III) as follows.

[Step I]: Consider the basic structure:

$$[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]$$

Define the measured value space $X^2 = \{-1, 1\}^2$ such that $X^2 = \{-1, 1\}^2 = \{(1, 1), (1, -1), (-1, 1), (-1, -1)\}$.

---

[1] This section is extracted from the following paper:

Ref. [30]: S. Ishikawa, "A New Interpretation of Quantum Mechanics," Journal of Quantum Information Science, Vol. 1 No. 2, 2011, pp. 35-42. doi: 10.4236/jqis.2011.12005

Consider two complex numbers $a = \alpha_1 + \alpha_2\sqrt{-1}$ and $b = \beta_1 + \beta_2\sqrt{-1}$ such that $|a| \equiv \sqrt{|\alpha_1|^2 + |\alpha_2|^2} = 1$ and $|b| \equiv \sqrt{|\beta_1|^2 + |\beta_2|^2} = 1$. Define the probability space $(X^2, \mathcal{P}(X^2), \nu_{ab})$ such that

$$\begin{aligned}
&\nu_{ab}(\{(1,1)\}) = \nu_{ab}(\{(-1,-1)\}) = (1 - \alpha_1\beta_1 - \alpha_2\beta_2)/4 \\
&\nu_{ab}(\{(1,-1)\}) = \nu_{ab}(\{(-1,1)\}) = (1 + \alpha_1\beta_1 + \alpha_2\beta_2)/4.
\end{aligned} \tag{4.48}$$

The correlation function $P(a, b)$ is calculated as

$$P(a,b) \equiv \sum_{(x_1,x_2)\in X\times X} x_1 \cdot x_2 \nu_{ab}(\{(x_1,x_2)\}) = -\alpha_1\beta_1 - \alpha_2\beta_2 \tag{4.49}$$

Our present problem is as follows.

(D):Problem

Find the measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}_{ab} := (X^2,\ \mathcal{P}(X^2),\ F_{ab}),\ S_{[\rho_0]})$ that satisfies

$$\nu_{ab}(\Xi) = \rho_0(F_{ab}(\Xi)) \qquad (\forall \Xi \in \mathcal{P}(X^2))$$

This will be answered in the following step [II].

[Step: II]: Consider the problem in the two cases. That is,

$$\begin{cases} \text{(i):quantum case: } [\mathcal{A} = B(\mathbb{C}^2 \otimes \mathbb{C}^2)] \\ \text{(ii):classical case: } [\mathcal{A} = C_0(\Omega \times \Omega)] \end{cases}$$

(i): quantum case $[\mathcal{A} = B(\mathbb{C}^2) \otimes B(\mathbb{C}^2) = B(\mathbb{C}^2 \otimes \mathbb{C}^2)\ ]$

Put

$$e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (\in \mathbb{C}^2).$$

For each $c \in \{a, b\}$, define the observable $\mathsf{O}_c \equiv \big(X, \mathcal{P}(X), G_c\big)$ in $B(\mathbb{C}^2)$ such that

$$G_c(\{1\}) = \frac{1}{2}\begin{bmatrix} 1 & \bar{c} \\ c & 1 \end{bmatrix}, \quad G_c(\{-1\}) = \frac{1}{2}\begin{bmatrix} 1 & -\bar{c} \\ -c & 1 \end{bmatrix}.$$

Consider the two particles quantum system in $B(\mathbb{C}^2 \otimes \mathbb{C}^2)$.

Consider two states $\rho_s = |\psi_s\rangle\langle\psi_s|$ and $\rho_0 = |\psi_0\rangle\langle\psi_0|$ $\big(\in \mathfrak{S}^p(B(\mathbb{C}^2 \otimes \mathbb{C}^2)^*)\big)$. Here, put $\psi_s = (e_1 \otimes e_2 - e_2 \otimes e_1)/\sqrt{2}$ and $\psi_0 = e_1 \otimes e_1$.

Consider the unitary operator U $(\in B(\mathbb{C}^2 \otimes \mathbb{C}^2)$ such that $U\psi_0 = \psi_s$.

Consider an observable $\mathsf{O}_{ab} = (X^2, \mathcal{P}(X^2), F_{ab} := U^*(G_a \otimes G_b)U)$ in $B(\mathbb{C}^2 \otimes \mathbb{C}^2)$, and get the measurement $\mathsf{M}_{B(\mathbb{C}^2\otimes\mathbb{C}^2)}(\mathsf{O}_{ab}, S_{[\rho_0]})$.

This clearly satisfies (D). That is because we see that, for each $(x_1, x_2) \in X^2$,

$$\rho_0(F_{ab}(\{(x_1, x_2)\})) = \langle \psi_0, F_{ab}(\{(x_1, x_2)\})\psi_0 \rangle$$
$$= \langle \psi_s, (G_a(\{x_1\}) \otimes G_b(\{x_2\}))\psi_s \rangle = \nu_{ab}(\{(x_1, x_2)\}).$$

(ii): classical case: $[\mathcal{A} = C_0(\Omega) \otimes C_0(\Omega) = C_0(\Omega \times \Omega)]$

Put $\omega_0 (= (\omega_0', \omega_0'')) \in \Omega \times \Omega$, and $\rho_0 = \delta_{\omega_0}$ $(\in \mathfrak{S}^p(C_0(\Omega \times \Omega)^*)\ )$.

Define the observable $\mathsf{O}_{ab} := (X^2, \mathcal{P}(X^2), F_{ab})$ in $L^\infty(\Omega \times \Omega)$ such that

$$[F_{ab}(\{(x_1, x_2)\})](\omega_0) = \nu_{ab}(\{(x_1, x_2)\})$$

Therefore, we get the measurement $\mathsf{M}_{L^\infty(\Omega\times\Omega)}(\mathsf{O}_{ab}, S_{[\delta_{\omega_0}]})$, which clearly satisfies (D).

[Step III]: For each $k = 1, 2$, consider two complex numbers $a^k (= \alpha_1^k + \alpha_2^k \sqrt{-1})$ and $b^k (= \beta_1^k + \beta_2^k \sqrt{-1})$ such that $|a^k| = |b^k| = 1$.

Consider the tensor parallel measurement $\otimes_{i,j=1,2} \mathsf{M}_{\mathcal{A}}(\mathsf{O}_{a^i b^j} := (X^2, \mathcal{P}(X^2), F_{a^i b^j}), S_{[\rho_0]})$ in the tensor W^*-algebra $\otimes_{i,j=1,2} \overline{\mathcal{A}}$. Assume the measured value $x (\in X^8)$. That is,

$$x = \big((x_1^{11}, x_2^{11}), (x_1^{12}, x_2^{12}), (x_1^{21}, x_2^{21}), (x_1^{22}, x_2^{22})\big)$$
$$\in \mathop{\times}_{i,j=1,2} X^2$$

Here, we see, by (4.49), that, for any $i, j = 1, 2$,

$$P(a^i, b^j) = \sum_{(x_1^{ij}, x_2^{ij}) \in X \times X} x_1^{ij} \cdot x_2^{ij} \rho_0(F_{a^i b^j}(\{(x_1^{ij}, x_2^{ij})\}))$$
$$= -\alpha_1^i \beta_1^j - \alpha_2^i \beta_2^j$$

Putting

$$a^1 = \sqrt{-1},\ b^1 = \frac{1 + \sqrt{-1}}{\sqrt{2}},\ a^2 = 1,\ b^2 = \frac{1 - \sqrt{-1}}{\sqrt{2}},$$

we get the following equality:

$$|P(a^1, b^1) - P(a^1, b^2)| \ + \ |P(a^2, b^1) + P(a^2, b^2)| = 2\sqrt{2} \tag{4.50}$$

Thus, in both cases (i.e., quantum case $[\mathcal{A} = B(\mathbb{C}^2 \otimes \mathbb{C}^2)]$ and classical case $[\mathcal{A} = C_0(\Omega \times \Omega)]$), the formula (4.50) holds. This fact is often said that

Bell's inequality is violated

though we omit the explanation of the relation between the equality (4.50) and Bell's inequality (4.47).

Remark 4.18. [Shut up and calculate]. The above argument may suggest that there is something faster than light. However, when faster-than-light appears, our standing point is

Stop being bothered

This is not only our opinion but also most physicists'. In fact, in Mermin's book [61], he said

(a) "Most physicists, I think it is fair to say, are not bothered."

(b) If I were forced to sum up in one sentence what the Copenhagen interpretation says to me, it would be **"Shut up and calculate"**

If it is so, we want to assert that the linguistic interpretation (§3.1) is the true colors of "the Copenhagen interpretation". That is because I also consider that

(c) If I were forced to sum up in one sentence what the linguistic interpretation says to me, it would be *"Shut up and calculate."*

KSTS/RR-16/001
January 8, 2016

Chapter 5

Fisher statistics (I)

Measurement theory (= quantum language) is formulated as follows.

$$\bullet \quad \underset{\text{(=quantum language)}}{\boxed{\text{measurement theory}}} := \underbrace{\overset{\text{[Axiom 1]}}{\underset{(cf.\ \S2.7)}{\boxed{\text{Measurement}}}} + \overset{\text{[Axiom 2]}}{\underset{(cf.\ \S10.3)}{\boxed{\text{Causality}}}}}_{\text{a kind of spell(a priori judgment)}} + \underbrace{\overset{\text{[quantum linguistic interpretation]}}{\underset{(cf.\ \S3.1)}{\boxed{\text{Linguistic interpretation}}}}}_{\text{manual to use spells}}$$

Measurement theory says that

- Describe every phenomenon modeled on Axioms 1 and 2 (by a hint of the linguistic interpretation)!

In this chapter, we study Fisher statistics in terms of Axiom 1 (measurement: §2.7). We shall emphasize

the reverse relation between measurement and inference

(such as "the two sides of a coin").

The readers can read this chapter without the knowledge of statistics.

5.1 Statistics is, after all, urn problems

5.1.1 Population(=system)↔state

Example 5.1. The density functions of the whole Japanese male's height and the whole American male's height is respectively defined by f_J and f_A. That is,

$$\int_\alpha^\beta f_J(x)dx = \frac{\text{A Japanese male's population whose height is from } \alpha\text{(cm) to } \beta\text{(cm)}}{\text{A Japanese male's overall population}}$$

$$\int_\alpha^\beta f_A(x)dx = \frac{\text{An American male's population whose height is from } \alpha\text{(cm) to } \beta\text{(cm)}}{\text{An American male's overall population}}$$

Let the density functions f_J and f_A be regarded as the probability density functions f_J and f_A such as

(A) From $\left[\begin{array}{l} \text{the set of all Japanese males} \\ \text{the set of all American males} \end{array}\right]$, choose a person (at random). Then, the probability that his height is from α(cm) to β(cm) is given by

$$\left[\begin{array}{l} [F_h([\alpha,\beta))](\omega_J) = \int_\alpha^\beta f_J(x)dx \\ [F_h([\alpha,\beta))](\omega_A) = \int_\alpha^\beta f_A(x)dx \end{array}\right]$$

Now, let us represent the statements (A_1) and (A_2) in terms of quantum language: Define the state space Ω by $\Omega = \{\omega_J, \omega_A\}$ with the discrete metric d_D and the counting measure ν such that

$$\nu(\{\omega_J\}) = 1, \ \ \nu(\{\omega_A\}) = 1$$

$\Big($It does not matter, even if $\nu(\{\omega_J\}) = a, \ \ \nu(\{\omega_A\}) = b \ \ (a, b > 0)\Big)$.

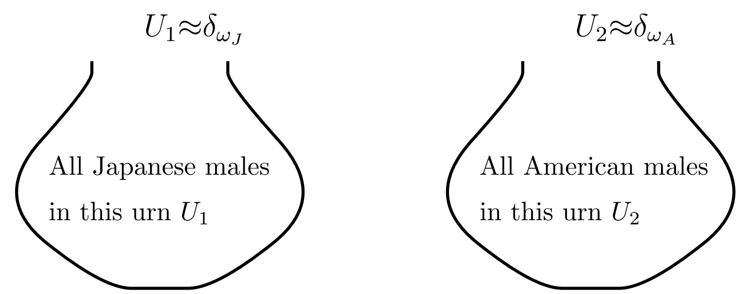


Figure 5.1: Population≈urn(↔state)

Thus, we have the classical basic structure:

$$\text{Classical basic structure}[C_0(\Omega) \subseteq L^\infty(\Omega,\nu) \subseteq B(L^2(\Omega,\nu))]$$

The pure state space is defined by

$$\mathfrak{S}^p(C_0(\Omega)^*) = \{\delta_{\omega_J}, \delta_{\omega_A}\} \approx \{\omega_J, \omega_A\} = \Omega$$

Here, we consider that

$$\begin{array}{lcl} \delta_{\omega_J} & \cdots & \text{“the state of the set } U_1 \text{ of all Japanese males”,} \\ \delta_{\omega_A} & \cdots & \text{“the state of the set } U_2 \text{ of all American males”,} \end{array}$$

and thus, we have the following identification (that is, Figure 5.1):

$$U_1 \approx \delta_{\omega_J}, \qquad U_2 \approx \delta_{\omega_A}$$

The observable $\mathsf{O}_h = (\mathbb{R}, \mathcal{B}, F_h)$ in $L^\infty(\Omega,\nu)$ is already defined by (A). Thus, we have the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_h, S_{[\delta_\omega]})$ $(\omega \in \Omega = \{\omega_J, \omega_A\})$. The statement(A) is represented in terms of quantum language by

(B) The probability that a measured value obtained by the measurement $\left[\begin{array}{l} \mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_h, S_{[\omega_J]}) \\ \mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_h, S_{[\omega_A]}) \end{array}\right]$ belongs to an interval $[\alpha, \beta)$ is given by

$$\left[\begin{array}{l} {}_{C_0(\Omega)^*}\Big(\delta_{\omega_J}, F_h([\alpha,\beta))\Big)_{L^\infty(\omega,\nu)} = [F_h([\alpha,\beta))](\omega_J) \\ {}_{C_0(\Omega)^*}\Big(\delta_{\omega_A}, F_h([\alpha,\beta))\Big)_{L^\infty(\omega,\nu)} = [F_h([\alpha,\beta))](\omega_A) \end{array}\right]$$

Therefore, we get:

$$\underset{\text{(ordinary language)}}{\boxed{\text{statement (A)}}} \xrightarrow[\text{translation}]{} \underset{\text{(quantum language)}}{\boxed{\text{statement (B)}}}$$

5.1.2 Normal observable and student t-distribution

Consider the classical basic structure:

$$[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))]$$

where $\Omega = \mathbb{R}$ (=the real line) with the Lebesgue measure ν. Let $\sigma > 0$ be a standard deviation, which is assumed to be fixed. Define the measured value space X by $\mathbb{R}$ (i.e., $X = \mathbb{R}$). Define the **normal observable** $\mathsf{O}_{G_\sigma} = (X(=\mathbb{R}), \mathcal{B}_{\mathbb{R}}, G_\sigma)$ in $L^\infty(\Omega, \nu)$ such that

$$[G_\sigma(\Xi)](\omega) = \frac{1}{\sqrt{2\pi}\sigma} \int_\Xi \exp\left[-\frac{1}{2\sigma^2}(x-\omega)^2 \right] dx \tag{5.1}$$
$$(\forall \Xi \in \mathcal{B}_X(=\mathcal{B}_{\mathbb{R}}),\ \forall \omega \in \Omega(=\mathbb{R}))$$

where $\mathcal{B}_{\mathbb{R}}$ is the Borel field. For example,

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\sigma}^{\sigma} e^{-\frac{x^2}{2\sigma^2}} dx = 0.683..., \qquad \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-2\sigma}^{2\sigma} e^{-\frac{x^2}{2\sigma^2}} dx = 0.954...,$$
$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-1.96\sigma}^{1.96\sigma} e^{-\frac{x^2}{2\sigma^2}} dx \doteqdot 0.95$$

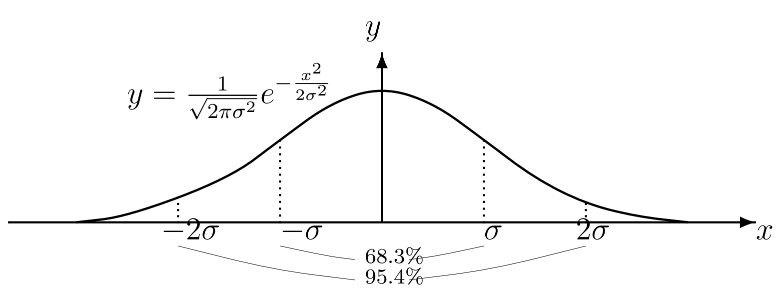


Figure 5.2: Error function

Next, consider the parallel observable $\bigotimes_{k=1}^n \mathsf{O}_{G_\sigma} = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}, \bigotimes_{k=1}^n G_\sigma)$ in $L^\infty(\Omega^n, \nu^{\otimes n})$ and restrict it on

$$K = \{(\omega, \omega, \ldots, \omega) \in \Omega^n \mid \omega \in \Omega\} (\subseteq \Omega^n)$$

This is essentially the same as the simultaneous observable $\mathsf{O}^n = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}, \times_{k=1}^n G_\sigma)$ in $L^\infty(\Omega)$. That is,

$$\begin{aligned} &[(\mathop{\times}_{k=1}^n G_\sigma)(\Xi_1 \times \Xi_2 \times \cdots \times \Xi_n)](\omega) = \mathop{\times}_{k=1}^n [G_\sigma(\Xi_k)](\omega) \\ =& \mathop{\times}_{k=1}^n \frac{1}{\sqrt{2\pi}\sigma} \int_{\Xi_k} \exp\Big[-\frac{1}{2\sigma^2}(x_k - \omega)^2 \Big] dx_k \\ & (\forall \Xi_k \in \mathcal{B}_X (= \mathcal{B}_{\mathbb{R}}), \ \forall \omega \in \Omega (= \mathbb{R})) \end{aligned} \tag{5.2}$$

Then, for each $(x_1, x_2, \cdots, x_n) \in X^n (= \mathbb{R}^n)$, define

$$\begin{aligned} \overline{x}_n &= \frac{x_1 + x_2 + \cdots + x_n}{n} \\ U_n^2 &= \frac{(x_1 - \overline{x}_n)^2 + (x_2 - \overline{x}_n)^2 + \cdots + (x_n - \overline{x}_n)^2}{n-1} \end{aligned}$$

and define the map $\psi : \mathbb{R}^n \to \mathbb{R}$ such that

$$\psi(x_1, x_2, \ldots, x_n) = \frac{\overline{x}_n - \omega}{U_n/\sqrt{n}}$$

Then, we have the observable $\mathsf{O}_{T_n^\sigma} = (X(=\mathbb{R}), \mathcal{B}_{\mathbb{R}}, T_n^\sigma)$ in $L^\infty(\mathbb{R})$ such that

$$[T_n^\sigma(\Xi)](\omega) = \Big[G_\sigma \Big(\{ (x_1, x_2, ..., x_n) \in \mathbb{R}^n \ \mid \ \frac{\overline{x}_n - \omega}{U_n/\sqrt{n}} \in \Xi \} \Big) \Big](\omega) \quad (\forall \Xi \in \mathcal{F}) \tag{5.3}$$

The observable $\mathsf{O}_{T_n^\sigma} = (X(=\mathbb{R}), \mathcal{B}_{\mathbb{R}}, T_n^\sigma)$ in $L^\infty(\mathbb{R})$ is called the student t observable .

Here, putting

$$f_n^\sigma(x) = \frac{\{\Gamma(n/2)}{\sqrt{(n-1)\pi}\Gamma((n-1)/2)} (1 + \frac{x^2}{n-1})^{-n/2} \qquad (\Gamma \text{ is Gamma function}) \tag{5.4}$$

we see that

$$[T_n^\sigma(\Xi)](\omega) = \int_\Xi f_n^\sigma(x) dx \quad (\forall \Xi \in \mathcal{F}) \tag{5.5}$$

which is independent of ω and σ. Also note that

$$\begin{aligned} \lim_{n\to\infty} f_n^\sigma(x) =& \lim_{n\to\infty} \frac{\Gamma(n/2)}{\sqrt{(n-1)\pi}\Gamma((n-1)/2)} (1 + \frac{x^2}{n-1})^{-n/2} \\ =& \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \end{aligned}$$

thus, if $n \geq 30$, it can be regarded as the normal distribution $N(0,1)$(that is, mean 0, the standard deviation 1).

5.2 The reverse relation between Fisher (=inference) and Born (=measurement)

In this section, we consider the reverse relation between Fisher (=inference) and Born (=measurement)

5.2.1 Inference problem (Statistical inference)

Before we mention Fisher's maximum likelihood method, we exercise the following problem:

Problem 5.2. [Urn problem(=Example2.34), A simplest example of Fisher's maximum likelihood method]
There are two urns U_1 and U_2. The urn U_1 [resp. U_2] contains 8 white and 2 black balls [resp. 4 white and 6 black balls].

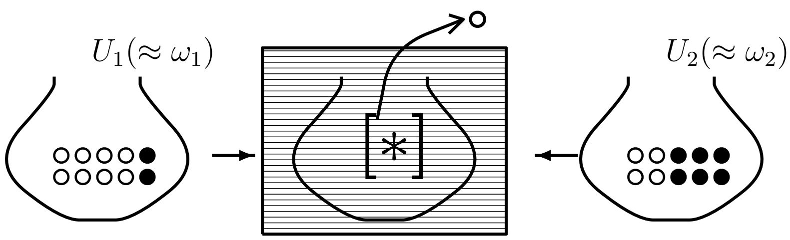


Figure 5.3: Pure measurement (Fisher's maximum likelihood method)

Here consider the following procedures (i) and (ii).

(i) One of the two (i.e., U_1 or U_2) is chosen and is settled behind a curtain. Note, for completeness, that you do not know whether it is U_1 or U_2.

(ii) Pick up a ball out of the unknown urn behind the curtain. And you find that the ball is white.

Here, we have the following problem:

(iii) **Infer the urn behind the curtain, U_1 or U_2?**

The answer is easy, that is, the urn behind the curtain is U_1. That is because the urn U_1 has more white balls than U_2. The above problem is too easy, but it includes the essence of Fisher maximum likelihood method.

5.2.2 Fisher's maximum likelihood method in measurement theory

We begin with the following notation:

Notation 5.3. $[\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}, S_{[*]})]$: Consider the measurement $\mathsf{M}_{\overline{\mathcal{A}}}$ $(\mathsf{O}{=}(X, \mathcal{F}, F),\ S_{[\rho]})$ formulated in the basic structure $[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]$. Here, note that

(A$_1$) In most cases that the measurement $\mathsf{M}_{\overline{\mathcal{A}}}$ $(\mathsf{O}{=}(X, \mathcal{F}, F),\ S_{[\rho]})$ is taken, it is usual to think that the state ρ $(\in \mathfrak{S}^p(\mathcal{A}^*))$ is unknown.

That is because

(A$_2$) the measurement $\mathsf{M}_{\mathcal{A}}(\mathsf{O}, S_{[\rho]})$ may be taken in order to know the state ρ.

Therefore, when we want to stress that

we do not know the state ρ

The measurement $\mathsf{M}_{\overline{\mathcal{A}}}$ $(\mathsf{O}{=}(X, \mathcal{F}, F),\ S_{[\rho]})$ is often denoted by

(A$_3$) $\mathsf{M}_{\overline{\mathcal{A}}}$ $(\mathsf{O}{=}(X, \mathcal{F}, F),\ S_{[*]})$

Further, consider the subset $K (\subseteq \mathfrak{S}^p(\mathcal{A}^*))$. When we know that the state ρ belongs to K, $\mathsf{M}_{\overline{\mathcal{A}}}$ $(\mathsf{O}{=}(X, \mathcal{F}, F),\ S_{[*]})$ is denoted by $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}, S_{[*]}(\!(K)\!))$. Therefore, it suffices to consider that

$$\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}, S_{[*]}) = \mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}, S_{[*]}(\!(\mathfrak{S}^p(\mathcal{A}^*))\!))$$

Using this notation $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}, S_{[*]})$, we characterize our problem (i.e., inference) as follows.

Problem 5.4. [Inference problem]

(a) Assume that a measured value obtained by $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}{=}(X, \mathcal{F}, F), S_{[*]}(\!(K)\!))$ belongs to $\Xi (\in \mathcal{F})$. Then, infer the unknown state $[*]$ $(\in \Omega)$

or,

(b) Assume that a measured value (x, y) obtained by $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}{=}(X \times Y, \mathcal{F} \boxtimes \mathcal{G}, H), S_{[*]}(\!(K)\!))$ belongs to $\Xi \times Y$ $(\Xi \in \mathcal{F})$. Then, infer the probability that $y \in \Gamma$.

Before we answer the problem, we emphasize the reverse relation between "inference" and "measurement".

The measurement is "the view from the front", that is,

(B$_1$) $$(\text{observable}[\mathsf{O}], \text{state}[\omega(\in \Omega)]) \xrightarrow[\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[\omega]})]{\text{measurement}} \text{measured value}[x(\in X)]$$

On the other hand, the inference is "the view from the back", that is,

(B$_2$) $$(\text{observable}[\mathsf{O}], \text{measured value}[x \in \Xi(\in \mathcal{F})]) \xrightarrow[\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[*]})]{\text{inference}} \text{state } [\omega(\in \Omega)]$$

In this sense, we say that

the inference problem is the reverse problem of measurement

Therefore, it suffices to image Fig. 5.4.

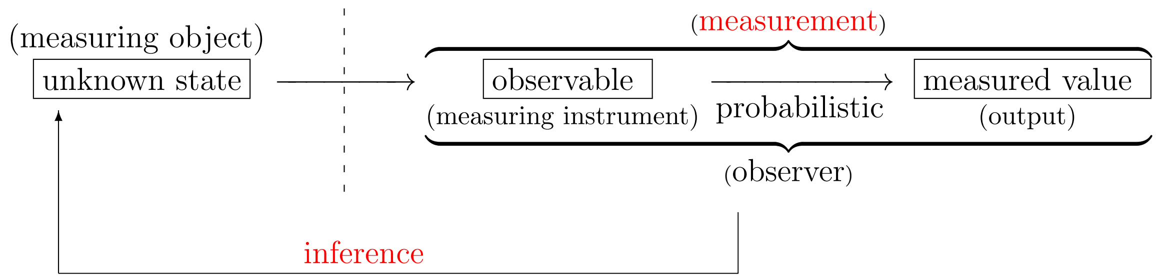


Figure 5.4: The image of inference

In order to answer the above problem 5.4, we shall describe Fisher maximum likelihood method in terms of measurement theory.

Theorem 5.5. [(Answer to Problem 5.4(b)): Fisher's maximum likelihood method(the general case)] Consider the basic structure

$$[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]$$

Assume that a measured value(x, y) obtained by a measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}{=}(X{\times}Y, \mathcal{F} \boxtimes \mathcal{G}, H), S_{[*]}((K)))$ belongs to $\Xi \times Y$ $(\Xi \in \mathcal{F})$. Then, there is reason to infer that the probability $P(\Gamma)$ that $y \in \Gamma$ is equal to

$$P(\Gamma) = \frac{\rho_0(H(\Xi \times \Gamma))}{\rho_0(H(\Xi \times Y))} \quad (\forall \Gamma \in \mathcal{G})$$

where, $\rho_0 \in K$ is determined by.

$$\rho_0(H(\Xi \times Y)) = \max_{\rho \in K} \rho(H(\Xi \times Y)) \tag{5.6}$$

Proof. Assume that $\rho_1, \rho_2 \in K$ and $\rho_1(H(\Xi \times Y)) < \rho_2(H(\Xi \times Y))$. By Axiom 1 (measurement: §2.7)

(i) the probability that a measured value(x, y) obtained by a measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}, S_{[\rho_1]})$ belongs to $\Xi \times Y$ is equal to $\rho_1(H(\Xi \times Y))$

(ii) the probability that a measured value(x, y) obtained by a measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}, S_{[\rho_2]})$ belongs to $\Xi \times Y$ is equal to $\rho_2(H(\Xi \times Y))$

Since we assume that $\rho_1(H(\Xi \times Y)) < \rho_2(H(\Xi \times Y))$, we can conclude that "(i) is more rare than (ii)". Thus, there is a reason to infer that $[*] = \omega_2$. Therefore, the ρ_0 in (5.6) is reasonable. Since the probability that a measured value(x, y) obtained by $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}, S_{[\rho_0]})$ belongs to $\Xi \times \Gamma$ is given by $\rho_0(H(\Xi \times \Gamma))$, we complete the proof of Theorem 5.5. □

Theorem 5.6. [(Answer to 5.4(a)): Fisher's maximum likelihood method **in classical case**]
(i): Consider a measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O} = (X, \mathcal{F}, F),\ S_{[*]}((K)))$. Assume that we know that a measured value obtained by a measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[*]}((K)))$ belongs to Ξ $(\in \mathcal{F})$. Then, there is a reason to infer that the unknown state state $[*]$ is ω_0 $(\in \Omega)$ such that

$$[F(\Xi)](\omega_0) = \max_{\omega \in \Omega} [F(\Xi)](\omega)$$

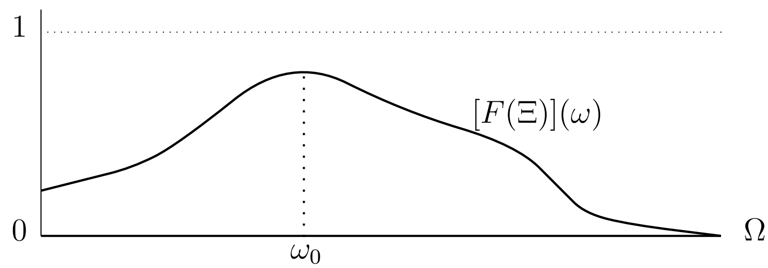


Figure 5.5: Fisher maximum likelihood method

(ii): Assume that a measured value x_0 $(\in X)$ is obtained by a measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O} = (X, \mathcal{F}, F),\ S_{[*]}((K)))$. Define the likelihood function $f(x, \omega)$ by

$$f(x, \omega) = \inf_{\omega_1 \in K} \Big[\lim_{\Xi \ni x, [F(\Xi)](\omega_1) \neq 0, \Xi \to \{x\}} \frac{[F(\Xi)](\omega)}{[F(\Xi)](\omega_1)} \Big] \tag{5.7}$$

Then, there is a reason to infer that $[*] = \omega_0 (\in K)$ such that $f(x_0, \omega_0) = 1$.

Proof. Consider Theorem 5.5 in the case that

$$[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)] = [C_0(\Omega) \subseteq L^\infty(\Omega) \subseteq B(L^2(\Omega))]$$

Thus, in the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O} {=} (X \times Y, \mathcal{F} \boxtimes \mathcal{G}, H), S_{[*]}((K)))$, consider the case that

$$\text{Fixed } \mathsf{O}_1 {=} (X, \mathcal{F}, F), \quad \text{any } \mathsf{O}_2 {=} (Y, \mathcal{G}, G),$$
$$\mathsf{O} {=} \mathsf{O}_1 \times \mathsf{O}_2 = (X \times Y, \mathcal{F} \boxtimes \mathcal{G}, F \times G), \quad \rho_0 = \delta_{\omega_0}$$

Then, we see

$$P(\Gamma) = \frac{[H(\Xi)](\omega_0) \times [G(\Gamma)](\omega_0)}{[H(\Xi)](\omega_0) \times [G(Y)](\omega_0)} = [G(\Gamma)](\omega_0) \quad (\forall \Gamma \in \mathcal{G}) \tag{5.8}$$

And, from the arbitrariness of O_2, there is a reason to infer that

$$[*] = \delta_{\omega_0} (\underset{\text{identification}}{\approx} \omega_0)$$

□

♠**Note 5.1.** The linguistic interpretation says that the state after measurement is non-sense. In this sense, the readers may consider that

($\sharp_1$) Theorem 5.6 is also non-sense

However, we say that

($\sharp_2$) in the sense of (5.8), Theorem 5.6 should be accepted.

or

($\sharp_3$) as far as classical system, it suffices to believe in Theorem 5.6

Answer 5.7. [The answer to Problem 5.2 by Fisher's maximum likelihood method]
You do not know which the urn behind the curtain is, U_1 or U_2.
Assume that you pick up a white ball from the urn.
The urn is U_1 or U_2? Which do you think?

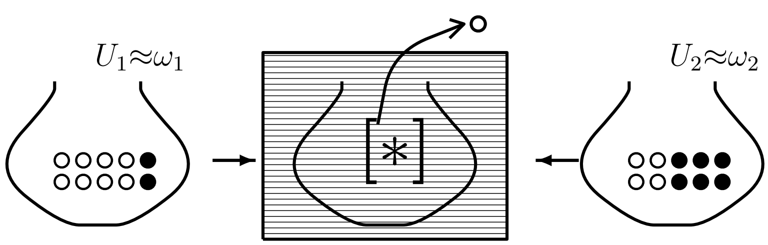


Figure 5.6: Pure measurement (Fisher's maximum likelihood method)

Answer: Consider the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}{=}\ (\{w,\ b\},\ 2^{\{w,b\}},\ F), S_{[*]})$, where the observable $\mathsf{O}_{wb} = (\{w, b\}, 2^{\{w,b\}}, F_{wb})$ in $L^\infty(\Omega)$ is defined by

$$
\begin{aligned}
&[F_{wb}(\{w\})](\omega_1) = 0.8, && [F_{wb}(\{b\})](\omega_1) = 0.2 \\
&[F_{wb}(\{w\})](\omega_2) = 0.4, && [F_{wb}(\{b\})](\omega_2) = 0.6
\end{aligned}
\tag{5.9}
$$

Here, we see:

$$
\begin{aligned}
&\max\{[F_{wb}(\{w\})](\omega_1), [F_{wb}(\{w\})](\omega_2)\} \\
=&\max\{0.8, 0.4\} = 0.8 = F_{wb}(\{w\})](\omega_1)
\end{aligned}
$$

Then, Fisher's maximum likelihood method (Theorem 5.6) says that

$$[*] = \omega_1$$

Therefore, there is a reason to infer that the urn behind the curtain is U_1. □

♠**Note 5.2.** As seen in Figure 5.4 , inference (Fisher maximum likelihood method) is the reverse of measurement (i.e., Axiom 1 due to Born). Here note that

(a) Born's discovery "the probabilistic interpretation of quantum mechanics" in [6] (1926)

(b) Fisher's great book *"Statistical Methods for Research Workers"* (1925)

Thus, it is surprising that Fisher and Born investigated the same thing in the different fields in the same age.

5.3 Examples of Fisher's maximum likelihood method

All examples mentioned in this section are easy for the readers who studied the elementary of statistics. However, it should be noted that these are consequence of Axiom 1 (measurement: §2.7).

Example 5.8. [Urn problem] Each urn U_1, U_2, U_3 contains many white balls and black ball such as:

Table 5.1: urn problem

w·b╲ Urn	Urn U_1	Urn U_2	Urn U_3
white ball	80%	40%	10%
black ball	20%	60%	90%

Here,

(i) one of three urns is chosen, but you do not know it. Pick up one ball from the unknown urn. And you find that its ball is white. Then, how do you infer the unknown urn, i.e., U_1, U_2 or U_3?

Further,

(ii) And further, you pick up another ball from the unknown urn (in (i)). And you find that its ball is black. That is, after all, you have one white ball and one black ball. Then, how do you infer the unknown urn, i.e., U_1, U_2 or U_3?

In what follows, we shall answer the above problems (i) and (ii) in terms of measurement theory.

Consider the classical basic structure:

$$[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))]$$

Put

$$\delta_{\omega_j}(\approx \omega_j) \longleftrightarrow \text{[the state such that urn } U_j \text{ is chosen]} \quad (j = 1, 2, 3)$$

Thus, we have the state space Ω ($=\{\omega_1, \omega_2, \omega_3\}$) with the counting measure ν. Further, define the observable $\mathsf{O} = (\{w, b\}, 2^{\{w,b\}}, F)$ in $C(\Omega)$ such that

$$F(\{w\})(\omega_1) = 0.8, \qquad F(\{w\})(\omega_2) = 0.4, \qquad F(\{w\})(\omega_3) = 0.1$$
$$F(\{b\})(\omega_1) = 0.2, \qquad F(\{b\})(\omega_2) = 0.6, \qquad F(\{b\})(\omega_3) = 0.9$$

Answer to (i): Consider the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[*]})$, by which a measured value "w" is obtained. Therefore, we see

$$[F(\{w\})](\omega_1) = 0.8 = \max_{\omega \in \Omega}[F(\{w\})](\omega) = \max\{0.8,\ 0.4,\ 0.1\}$$

Hence, by Fisher's maximum likelihood method (Theorem5.6) we see that

$$[*] = \omega_1$$

Thus, we can infer that the unknown urn is U_1.

Answer to (ii): Next, consider the simultaneous measurement $\mathsf{M}_{L^\infty(\Omega)}(\times_{k=1}^2 \mathsf{O} = (X^2, 2^{X^2}, \widehat{F} = \times_{k=1}^2 F), S_{[*]})$, by which a measured value (w, b) is obtained. Here, we see

$$[\widehat{F}(\{(w,b)\})](\omega) = [F(\{w\})](\omega) \cdot [F(\{b\})](\omega)$$

thus,

$$[\widehat{F}(\{(w,b)\})](\omega_1) = 0.16, \quad [\widehat{F}(\{(w,b)\})](\omega_2) = 0.24, \quad [\widehat{F}(\{(w,b)\})](\omega_3) = 0.09$$

Hence, by Fisher's maximum likelihood method (Theorem5.6), we see that

$$[*] = \omega_2$$

Thus, we can infer that the unknown urn is U_2. □

Example 5.9. **[Normal observable(i): $\Omega = \mathbb{R}$]** As mentioned before, we again discuss the normal observable in what follows. Consider the classical basic structure:

$$[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))] \qquad (\text{where, } \ \Omega = \mathbb{R})$$

Fix $\sigma > 0$, and consider the normal observable $\mathsf{O}_{G_\sigma} = (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, G_\sigma)$ in $L^\infty(\mathbb{R})$ (where $\Omega = \mathbb{R}$) such that

$$[G_\sigma(\Xi)](\mu) = \frac{1}{\sqrt{2\pi}\sigma} \int_{\Xi} \exp[-\frac{1}{2\sigma^2}(x-\mu)^2]dx$$
$$(\forall \Xi \in \mathcal{B}_{\mathbb{R}}, \quad \forall \mu \in \Omega = \mathbb{R})$$

Thus, the simultaneous observable $\times_{k=1}^3 \mathsf{O}_{G_\sigma}$ (in short, $\mathsf{O}^3_{G_\sigma}$) $= (\mathbb{R}^3, \mathcal{B}_{\mathbb{R}^3}, G^3_\sigma)$ in $L^\infty(\mathbb{R})$ is defined by

$$
\begin{aligned}
&[G_\sigma^3(\Xi_1 \times \Xi_2 \times \Xi_3)](\mu) = [G_\sigma(\Xi_1)](\mu) \cdot [G_\sigma(\Xi_2)](\mu) \cdot [G_\sigma(\Xi_3)](\mu) \\
=&\frac{1}{(\sqrt{2\pi}\sigma)^3} \iiint_{\Xi_1 \times \Xi_2 \times \Xi_3} \exp[-\frac{(x_1-\mu)^2+(x_2-\mu)^2+(x_3-\mu)^2}{2\sigma^2}] \\
&\qquad \times dx_1 dx_2 dx_3 \\
&\qquad\qquad (\forall \Xi_k \in \mathcal{B}_{\mathbb{R}}, k=1,2,3, \quad \forall \mu \in \Omega = \mathbb{R})
\end{aligned}
$$

Thus, we get the measurement $\mathsf{M}_{L^\infty(\mathbb{R})}(\mathsf{O}_{G_\sigma}^3, S_{[*]})$

Now we consider the following problem:

(a) Assume that a measured value (x_1^0, x_2^0, x_3^0) $(\in \mathbb{R}^3)$ is obtained by the measurement $\mathsf{M}_{L^\infty(\mathbb{R})}(\mathsf{O}_{G_\sigma}^3, S_{[*]})$. Then, infer the unknown state $[*](\in \mathbb{R})$.

Answer(a) Put

$$
\Xi_i = [x_i^0 - \frac{1}{N}, x_i^0 + \frac{1}{N}] \qquad (i=1,2,3)
$$

Assume that N is sufficiently large. Fisher's maximum likelihood method (Theorem5.6) says that the unknown state$[\,*\,] = \mu_0$ is found in what follows.

$$
[G_\sigma^3(\Xi_1 \times \Xi_2 \times \Xi_3)](\mu_0) = \max_{\mu \in \mathbb{R}} [G_\sigma^3(\Xi_1 \times \Xi_2 \times \Xi_3)](\mu)
$$

Since N is sufficiently large, we see

$$
\begin{aligned}
&\frac{1}{(\sqrt{2\pi}\sigma)^3} \exp[-\frac{(x_1^0-\mu_0)^2+(x_2^0-\mu_0)^2+(x_3^0-\mu_0)^2}{2\sigma^2}] \\
=&\max_{\mu \in \mathbb{R}} \Big[\frac{1}{(\sqrt{2\pi}\sigma)^3} \exp[-\frac{(x_1^0-\mu)^2+(x_2^0-\mu)^2+(x_3^0-\mu)^2}{2\sigma^2}]\Big]
\end{aligned}
$$

That is,

$$
(x_1^0-\mu_0)^2+(x_2^0-\mu_0)^2+(x_3^0-\mu_0)^2 = \min_{\mu \in \mathbb{R}} \big\{(x_1^0-\mu)^2+(x_2^0-\mu)^2+(x_3^0-\mu)^2\big\}
$$

Therefore, solving $\frac{d}{d\mu}\{\cdots\}=0$, we conclude that

$$
\mu_0 = \frac{x_1^0+x_2^0+x_3^0}{3}
$$

□

[Normal observable(ii)] Next consider the classical basic structure:

$$
[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))] \qquad (\text{where, } \ \Omega = \mathbb{R} \times \mathbb{R}_+)
$$

and consider the case:

- we know that the length of the pencil μ is satisfied that 10cm μ L cm $\leq$30.

And we assume that

($\sharp$) the length of the pencil μ and the roughness σ of the ruler are unknown.

That is, assume that the state space $\Omega = [10,30] \times \mathbb{R}_+$ $\big(=\{\mu \in \mathbb{R} \mid 10 \leqq \mu \leqq 30\} \times \{\sigma \in \mathbb{R} \mid \sigma > 0\}\big)$

Define the observable $\mathsf{O} = (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, G)$ in $L^\infty([10,30] \times \mathbb{R}_+)$ such that

$$[G(\Xi)](\mu,\sigma) = [G_\sigma(\Xi)](\mu) \quad (\forall \Xi \in \mathcal{B}_{\mathbb{R}}, \ \forall (\mu,\sigma) \in \Omega = [10,30] \times \mathbb{R}_+)$$

Therefore, the simultaneous observable $\mathsf{O}^3 = (\mathbb{R}^3, \mathcal{B}_{\mathbb{R}^3}, G^3)$ in $C([10,30] \times \mathbb{R}_+)$ is defined by

$$\begin{aligned}
&[G^3(\Xi_1 \times \Xi_2 \times \Xi_3)](\mu,\sigma) = [G(\Xi_1)](\mu,\sigma) \cdot [G(\Xi_2)](\mu,\sigma) \cdot [G(\Xi_3)](\mu,\sigma) \\
=&\frac{1}{(\sqrt{2\pi}\sigma)^3} \int_{\Xi_1 \times \Xi_2 \times \Xi_3} \exp[-\frac{(x_1-\mu)^2+(x_2-\mu)^2+(x_3-\mu)^2}{2\sigma^2}] dx_1 dx_2 dx_3 \\
&\qquad (\forall \Xi_k \in \mathcal{B}_{\mathbb{R}}, k=1,2,3, \quad \forall (\mu,\sigma) \in \Omega = [10,30] \times \mathbb{R}_+)
\end{aligned}$$

Thus, we get the simultaneous measurement $\mathsf{M}_{L^\infty([10,30]\times\mathbb{R}_+)}(\mathsf{O}^3, S_{[*]})$. Here, we have the following problem:

(b) When a measured value (x_1^0, x_2^0, x_3^0) $(\in \mathbb{R}^3)$ is obtained by the measurement $\mathsf{M}_{L^\infty([10,30]\times\mathbb{R}_+)}$ $(\mathsf{O}^3,\ S_{[*]})$, infer the unknown state $[*](=(\mu_0,\sigma_0) \in [10,30] \times \mathbb{R}_+)$, i.e., the length μ_0 of the pencil and the roughness σ_0 of the ruler.

Answer (b) By the same way of (a), Fisher's maximum likelihood method (Theorem5.6) says that the unknownstate $[\,*\,] = (\mu_0,\sigma_0)$ such that

$$\begin{aligned}
&\frac{1}{(\sqrt{2\pi}\sigma_0)^3} \exp[-\frac{(x_1^0-\mu_0)^2+(x_2^0-\mu_0)^2+(x_3^0-\mu_0)^2}{2\sigma_0^2}] \\
=&\max_{(\mu,\sigma)\in[10,30]\times\mathbb{R}_+} \Big\{ \frac{1}{(\sqrt{2\pi}\sigma)^3} \exp[-\frac{(x_1^0-\mu)^2+(x_2^0-\mu)^2+(x_3^0-\mu)^2}{2\sigma^2}] \Big\}
\end{aligned} \tag{5.10}$$

Thus, solving $\frac{\partial}{\partial\mu}\{\cdots\}=0$, $\frac{\partial}{\partial\sigma}\{\cdots\}=0$ we see

$$\mu_0 = \begin{cases} 10 & (\text{when } (x_1^0+x_2^0+x_3^0)/3 < 10\ ) \\ (x_1^0+x_2^0+x_3^0)/3 & (\text{when } 10 \leqq (x_1^0+x_2^0+x_3^0)/3 \leqq 30\ ) \\ 30 & (\text{when } 30 < (x_1^0+x_2^0+x_3^0)/3\ ) \end{cases} \tag{5.11}$$

$$\sigma_0 = \sqrt{\{(x_1^0-\widetilde{\mu})^2+(x_2^0-\widetilde{\mu})^2+(x_3^0-\widetilde{\mu})^2\}/3}$$

where

$$\widetilde{\mu} = (x_1^0 + x_2^0 + x_3^0)/3$$

□

Example 5.10. [Fisher's maximum likelihood method for the simultaneous normal measurement]. Consider the simultaneous normal observable $\mathsf{O}_G^n = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n, G^n)$ in $L^\infty(\mathbb{R} \times \mathbb{R}_+)$ (such as defined in formula (5.2)). This is essentially the same as the simultaneous observable $\mathsf{O}^n = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}, \times_{k=1}^n G_\sigma)$ in $L^\infty(\mathbb{R} \times \mathbb{R}_+)$. That is,

$$[(\mathop{\times}_{k=1}^n G_\sigma)(\Xi_1 \times \Xi_2 \times \cdots \times \Xi_n)](\omega) = \mathop{\times}_{k=1}^n [G_\sigma(\Xi_k)](\omega)$$
$$= \mathop{\times}_{k=1}^n \frac{1}{\sqrt{2\pi}\sigma} \int_{\Xi_k} \exp\Big[- \frac{1}{2\sigma^2}(x_k - \mu)^2 \Big] dx_k$$
$$(\forall \Xi_k \in \mathcal{B}_X (= \mathcal{B}_{\mathbb{R}}),\ \forall \omega = (\mu, \sigma) \in \Omega (= \mathbb{R} \times \mathbb{R}_+))$$

Assume that a measured value $x = (x_1, x_2, \ldots, x_n) (\in \mathbb{R}^n)$ is obtained by the measurement $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}(\mathsf{O}^n = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n, G_\sigma^n), S_{[*]})$. The likelihood function $L_x(\mu, \sigma) (= L(x, (\mu, \sigma))$ is equal to

$$L_x(\mu, \sigma) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp\Big[- \frac{\sum_{k=1}^n (x_k - \mu)^2}{2\sigma^2} \Big]$$

or, in the sense of (5.7),

$$L_x(\mu, \sigma) = \frac{\frac{1}{(\sqrt{2\pi}\sigma)^n} \exp\big[- \frac{\sum_{k=1}^n (x_k - \mu)^2}{2\sigma^2} \big]}{\frac{1}{(\sqrt{2\pi}\overline{\sigma}(x))^n} \exp\big[- \frac{\sum_{k=1}^n (x_k - \overline{\mu}(x))^2}{2\overline{\sigma}(x)^2} \big]} \tag{5.12}$$
$$(\forall x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n, \quad \forall \omega = (\mu, \sigma) \in \Omega = \mathbb{R} \times \mathbb{R}_+).$$

Therefore, we get the following likelihood equation:

$$\frac{\partial L_x(\mu, \sigma)}{\partial \mu} = 0, \quad \frac{\partial L_x(\mu, \sigma)}{\partial \sigma} = 0 \tag{5.13}$$

which is easily solved. That is, Fisher's maximum likelihood method (Theorem5.6) says that the unknown state $[*] = (\mu, \sigma)$ $(\in \mathbb{R} \times \mathbb{R}_+)$ is inferred as follows.

$$\mu = \overline{\mu}(x) = \frac{x_1 + x_2 + \ldots + x_n}{n}, \tag{5.14}$$

$$\sigma = \overline{\sigma}(x) = \sqrt{\frac{\sum_{k=1}^n (x_k - \overline{\mu}(x))^2}{n}} \tag{5.15}$$

5.4 Moment method: useful but artificial

Let us explain the moment method (*cf.* [29]), which as well as Fisher's maximum likelihood method are frequently used.

Consider the measurement $\mathsf{M}_{\mathcal{A}}\big(\mathsf{O} \equiv (X, \mathcal{F}, F), S_{[\rho]}\big)$, and its parallel measurement $\otimes_{k=1}^{n} \mathsf{M}_{\mathcal{A}}\big(\mathsf{O} \equiv (X, \mathcal{F}, F), S_{[\rho]}\big)$ $(= \mathsf{M}_{\otimes \mathcal{A}}\big(\bigotimes_{k=1}^{n} \mathsf{O} := (X^n, \mathcal{F}^n, \bigotimes_{k=1}^{n} F), S_{[\otimes_{k=1}^{n} \rho]}\big)$. Assume that the measured value $(x_1, x_2, ..., x_n) (\in X^n)$ is obtained by the parallel measurement. Assume that n is sufficiently large. By the law of large numbers (Theorem 4.5), we can assure that

$$\mathcal{M}_{+1}(X) \ni \nu_n\Big(\equiv \frac{\delta_{x_1} + \delta_{x_2} + \cdots + \delta_{x_n}}{n}\Big) \doteqdot \rho(F(\cdot)) \in \mathcal{M}_{+1}(X) \tag{5.16}$$

Thus,

(A) in order to infer the unknown state $\rho(\in \mathfrak{S}^p(\mathcal{A}^*))$, it suffices to solve the equation (5.16)

For example, we have several methods to solve the equation (5.16) as follows.

(B_1) Solve the following equation:

$$\|\nu_n(\cdot) - \rho(F(\cdot))\|_{\mathcal{M}(X)} = \min\{\|\nu_n(\cdot) - \rho_1(F(\cdot))\|_{\mathcal{M}(X)} \mid \rho_1(\in \mathfrak{S}^p(\mathcal{A}^*))\} \tag{5.17}$$

(B_2) For some $f_1, f_2, \cdots, f_n \in C(X)$ (= the set of all continuous functions on X), it suffices to find $\rho(\in \mathfrak{S}^p(\mathcal{A}^*))$ such that $\Delta(\rho) = \min_{\rho_1(\in \mathfrak{S}^p(\mathcal{A}^*))} \Delta(\rho_1)$, where

$$\begin{aligned}\Delta(\rho) =& \sum_{k=1}^{n} \Big| \int_X f_k(\xi)\nu_n(d\xi) - \int_X f_k(\xi)\rho(F(d\xi))\Big| \\ =& \sum_{k=1}^{n} \Big| \frac{f_k(x_1) + f_k(x_2) + \cdots + f_k(x_n)}{n} - \int_X f_k(\xi)\rho(F(d\xi))\Big|\end{aligned}$$

(B_3) In the cases of the classical measurement $\mathsf{M}_{L^\infty(\Omega)}\big(\mathsf{O} \equiv (X, \mathcal{F}, F), S_{[\rho]}\big)$ (putting $\rho = \delta_\omega$), it suffices to solve

$$0 = \sum_{k=1}^{n} \Big| \frac{f_k(x_1) + f_k(x_2) + \cdots + f_k(x_n)}{n} - \int_X f_k(\xi)[F(d\xi)](\omega)\Big| \tag{5.18}$$

or, it suffices to solve

$$\begin{cases} \frac{f_1(x_1)+f_1(x_2)+\cdots+f_1(x_n)}{n} - \int_X f_1(\xi)[F(d\xi)](\omega) = 0 \\ \frac{f_2(x_1)+f_2(x_2)+\cdots+f_2(x_n)}{n} - \int_X f_2(\xi)[F(d\xi)](\omega) = 0 \\ \qquad \cdots\cdots \\ \qquad \cdots\cdots \\ \frac{f_m(x_1)+f_m(x_2)+\cdots+f_m(x_n)}{n} - \int_X f_m(\xi)[F(d\xi)](\omega) = 0 \end{cases}$$

(B$_4$) Particularly, in the case that $X = \{\xi_1, \xi_2, \cdots, \xi_m\}$ is finite, define $f_1, f_2, \cdots, f_m \in C(X)$ by

$$f_k(\xi) = \chi_{\{\xi_k\}}(\xi) = \begin{cases} 1 & (\xi = \xi_k) \\ 0 & (\xi \neq \xi_k) \end{cases}$$

and, it suffices to find the $\rho (= \delta_\omega)$ such that

$$\sum_{k=1}^{n} \Big| \frac{\chi_{\{\xi_k\}}(x_1) + \chi_{\{\xi_k\}}(x_2) + \cdots + \chi_{\{\xi_k\}}(x_n)}{n} - \int_X \chi_{\{\xi_k\}}(\xi)\rho(F(d\xi)) \Big|$$
$$= \sum_{k=1}^{n} \Big| \frac{\sharp[\{x_m \; : \; \xi_k = x_m\}]}{n} - [F(\{\xi_k\}](\omega)) \Big| = 0$$

The above methods are all **the moment method**. Note that

(C$_1$) It is desirable that n is sufficiently large, but the moment method may be valid even when $n = 1$.

(C$_2$) The choice of f_k is artificial (on the other hand, Fisher' maximum likelihood method is natural).

Problem 5.11. [=Problem5.2: Urn problem: by the moment method]
You do not know which the urn behind the curtain is, U_1 or U_2.
Assume that you pick up a white ball from the urn.
The urn is U_1 or U_2? Which do you think?

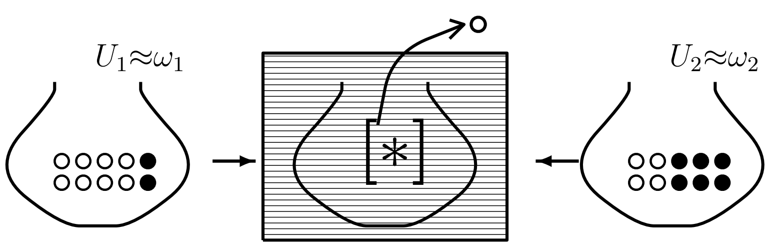


Figure 5.7: Inference(by moment method)

Answer: Consider the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O} = (\{w, b\}, 2^{\{w,b\}}, F), S_{[*]})$. Here, recall that the observable $\mathsf{O}_{wb} = (\{w, b\}, 2^{\{w,b\}}, F_{wb})$ in $L^\infty(\Omega)$ is defined by

$$[F_{wb}(\{w\})](\omega_1) = 0.8, \qquad [F_{wb}(\{b\})](\omega_1) = 0.2$$

$$[F_{wb}(\{w\})](\omega_2) = 0.4, \qquad [F_{wb}(\{b\})](\omega_2) = 0.6$$

Since a measured value "w" is obtained, the approximate sample space $(\{w, b\}, 2^{\{w,b\}}, \nu_1)$ is obtained as

$$\nu_1(\{w\}) = 1, \quad \nu_1(\{b\}) = 0$$

[when the unknown state $[*]$ is ω_1]

$$(5.17) = |1 - 0.8| + |0 - 0.2|$$

[when the unknown state $[*]$ is ω_2]

$$(5.17) = |1 - 0.4| + |0 - 0.6|$$

Thus, by the moment method, we can infer that $[*] = \omega_1$, that is, the urn behind the curtain is U_1.

[II] The above may be too easy. Thus, we add the following problem.

Problem 5.12. [Sampling with replacement]: As mentioned in the above, assume that "white ball" is picked. and the ball is returned to the urn. And further, we pick "black ball", and it is returned to the urn. Repeat this, after all, assume that we get

$$\text{“}w\text{”}, \text{“}b\text{”}, \text{“}b\text{”}, \text{“}w\text{”}, \text{“}b\text{”}, \text{“}w\text{”}, \text{“}b\text{”},$$

Then, we have the following problem:

(a) Which the urn behind the curtain is U_1 or U_2?

Answer: Consider the simultaneous measurement $\mathsf{M}_{L^\infty(\Omega)}(\times_{k=1}^{7}\mathsf{O}= (\{w, b\}^7, 2^{\{w,b\}^7}, \times_{k=1}^{7} F), S_{[*]})$. And assume that the measured value is (w, b, b, w, b, w, b). Then,

[when $[*]$ is ω_1]

$$(5.17) = |3/7 - 0.8| + |4/7 - 0.2| = 52/70$$

[when $[*]$ is ω_2]

$$(5.17) = |3/7 - 0.4| + |4/7 - 0.6| = 10/70$$

Thus, by the moment method, we can infer that $[*] = \omega_2$, that is, the urn behind the curtain is U_2.

□

Example 5.13. [The most important example of moment method] Putting $\Omega = \mathbb{R} \times \mathbb{R}_+ = \{\omega = (\mu, \sigma) \mid \mu \in \mathbb{R}, \sigma > 0\}$ with Lebesgue measure ν, Consider the classical basic structure

$$[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))]$$

Assume that the observable $\mathsf{O}_G = (X(=\mathbb{R}), \mathcal{B}_{\mathbb{R}}, G)$ in $L^\infty(\Omega, \nu)$ satisfies that

$$\int_{\mathbb{R}} \xi[G(d\xi)](\mu, \sigma) = \mu, \quad \int_{\mathbb{R}} (\xi - \mu)^2 [G(d\xi)](\mu, \sigma) = \sigma^2$$
$$(\forall \omega = (\mu, \sigma) \in \Omega(= \mathbb{R} \times \mathbb{R}_+))$$

Here, assume that a measured value $(x_1, x_2, x_3)(\in \mathbb{R}^3)$ is obtained by the simultaneous measurement $\times_{k=1}^3 \mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_G, S_{[*]})$. That is, we have the 3-sample distribution ν_3 such that

$$\nu_3 = \frac{\delta_{x_1} + \delta_{x_2} + \delta_{x_3}}{3} \in \mathcal{M}_{+1}(\mathbb{R})$$

Put $f_1(\xi) = \xi, f_2(\xi) = \xi^2$. Then, by the moment method (5.18), we see:

$$\begin{aligned}
0 =& \sum_{k=1}^{2} \Big| \int_{\mathbb{R}} \xi^k \nu_3(d\xi) - \int_{\mathbb{R}} \xi^k [G(d\xi)](\omega) \Big| \\
=& \sum_{k=1}^{2} \Big| \frac{(x_1)^k + (x_2)^k + (x_n)^k}{3} - \int_{\mathbb{R}} \xi^k [G(d\xi)](\mu, \sigma) \Big| \\
=& \Big| \frac{x_1 + x_2 + x_3}{3} - \mu \Big| + \Big| \frac{(x_1)^2 + (x_2)^2 + (x_3)^2}{3} - (\sigma^2 + \mu^2) \Big|
\end{aligned}$$

Thus, we get:

$$\begin{aligned}
\mu =& \frac{x_1 + x_2 + x_n}{3} \\
\sigma^2 =& \frac{(x_1)^2 + (x_2)^2 + (x_3)^2}{3} - \mu^2 \\
=& \frac{(x_1 - \frac{x_1+x_2+x_n}{3})^2 + (x_2 - \frac{x_1+x_2+x_n}{3})^2 + (x_3 - \frac{x_1+x_2+x_n}{3})^2}{3}
\end{aligned}$$

which is the same as the (5.11) concerning the normal measurement.

♠**Note 5.3.** Consider the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}{=}(X, 2^X,\ F),\ S_{[*]})$, where $X = \{x_1, x_2, ..., x_n\}$ is finite. Then, we see that

"Fisher's maximum likelihood method"="moment method"

.

[**Answer**] Assume that a measured value$x_m (\in X)$ is obtained by the measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}{=}(X, 2^X,\ F),\ S_{[*]})$

[Fisher's maximum likelihood method]:

(a) Find $\omega_0 (\in \Omega)$ such that

$$[F(\{x_m\})](\omega_0) = \max_{\omega \in \Omega} [F(\{x_m\})](\omega)$$

[Moment method]:

(b) Since we get the approximate sample probability space $(X, 2^X, \delta_{x_m})$, we see

$$\begin{aligned}
&|0 - [F(\{x_1\})](\omega)| + \cdots + |0 - [F(\{x_{m-1}\})](\omega)| + |1 - [F(\{x_m\})](\omega)| \\
&\qquad + |0 - [F(\{x_{m+1}\})](\omega)| + \cdots + |0 - [F(\{x_n\})](\omega)| \\
=&[F(\{x_1\})](\omega) + \cdots + [F(\{x_{m-1}\})](\omega) + [F(\{x_m\})](\omega) \\
&\qquad + [F(\{x_{m+1}\})](\omega) + \cdots + [F(\{x_n\})](\omega) \\
=&1 - 2[F(\{x_m\})](\omega)
\end{aligned}$$

Thus, it suffice to find $\omega_0 (\in \Omega)$ such that

$$1 - 2[F(\{x_m\})](\omega_0) = \min_{\omega} (1 - 2[F(\{x_m\})](\omega))$$

Thus, Fisher's maximum likelihood method and the moment method are the same in this case.

5.5 Monty Hall problem—High school student puzzle—

Monty Hall problem is as follows[1].

Problem 5.14. [Monty Hall problem]

You are on a game show and you are given the choice of three doors. Behind one door is a car, and behind the other two are goats. You choose, say, door 1, and the host, who knows where the car is, opens another door, behind which is a goat. For example, the host says that

($\flat$) the door 3 has a goat.

And further, he now gives you the choice of sticking with door 1 or switching to door 2? **What should you do?**

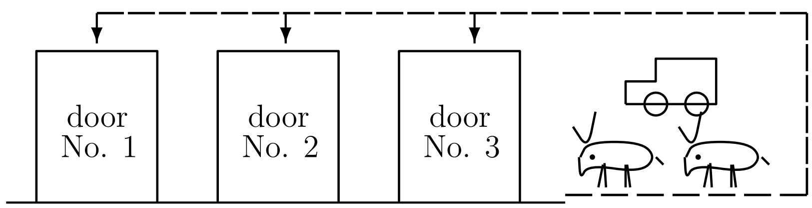


Figure 5.8: Monty Hall problem

Answer: Put $\Omega = \{\omega_1, \omega_2, \omega_3\}$ with the discrete topology d_D and the counting measure ν. Thus consider the classical basic structure:

$$[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))]$$

Assume that each state $\delta_{\omega_m} (\in \mathfrak{S}^p(C(\Omega)^*))$ means

$$\delta_{\omega_m} \Leftrightarrow \text{ the state that the car is behind the door } m \quad (m = 1, 2, 3)$$

Define the observable $\mathsf{O}_1 \equiv (\{1, 2, 3\}, 2^{\{1,2,3\}}, F_1)$ in $L^\infty(\Omega)$ such that

$$[F_1(\{1\})](\omega_1) = 0.0, \qquad [F_1(\{2\})](\omega_1) = 0.5, \qquad [F_1(\{3\})](\omega_1) = 0.5,$$
$$[F_1(\{1\})](\omega_2) = 0.0, \qquad [F_1(\{2\})](\omega_2) = 0.0, \qquad [F_1(\{3\})](\omega_2) = 1.0,$$

[1]This section is extracted from the followings:

(a) Ref. [29]: S. Ishikawa, "Mathematical Foundations of Measurement Theory," Keio University Press Inc. 2006.

(b) Ref. [33]: S. Ishikawa, "Monty Hall Problem and the Principle of Equal Probability in Measurement Theory," Applied Mathematics, Vol. 3 No. 7, 2012, pp. 788-794. doi: 10.4236/am.2012.37117.

$$[F_1(\{1\})](\omega_3) = 0.0, \qquad [F_1(\{2\})](\omega_3) = 1.0, \qquad [F_1(\{3\})](\omega_3) = 0.0, \tag{5.19}$$

where it is also possible to assume that $F_1(\{2\})(\omega_1) = \alpha$, $F_1(\{3\})(\omega_1) = 1-\alpha$ $(0 < \alpha < 1)$. The fact that you say "the door 1" clearly means that you take a measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_1, S_{[*]})$. Here, we assume that

a) "a measured value 1 is obtained by the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_1, S_{[*]})$"
$\Leftrightarrow$ The host says "Door 1 has a goat"

b) "measured value 2 is obtained by the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_1, S_{[*]})$ "
$\Leftrightarrow$ The host says "Door 2 has a goat"

c) "measured value 3 is obtained by the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_1, S_{[*]})$ "
$\Leftrightarrow$ The host says "Door 3 has a goat"

Recall that, in Problem 5.14, the host said "Door 3 has a goat." This implies that you get the measured value "3" by the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_1, S_{[*]})$. Therefore, Theorem 5.6 (Fisher's maximum likelihood method) says that *you should pick door number 2.* That is because we see that

$$\begin{aligned}\max\{[F_1(\{3\})](\omega_1), [F_1(\{3\})](\omega_2), [F_1(\{3\})](\omega_3)\} &= \max\{0.5, \;\; 1.0, \;\; 0.0\} \\ &= 1.0 = [F_1(\{3\})](\omega_2)\end{aligned}$$

and thus, there is a reason to infer that $wquaualweigh[*] = \delta_{\omega_2}$. Thus, you should switch to door 2. This is the first answer to Problem 5.14 (Monty-Hall problem). $\square$

♠**Note 5.4.** Examining the above example, the readers should understand that the problem "What is measurement?" is an unreasonable demand. Thus,

we abandon the realistic approach, and accept the metaphysical approach.

Also, for a Bayesian approach to Monty Hall problem, see Chapter 9 and Chapter 19.

Remark 5.15. [The answer by the moment method] In the above, a measured value "3" is obtained by the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}{=}(\{1,2,3\}, 2^{\{1,2,3\}}, F),\ S_{[*]})$. Thus, the approximate sample space $(\{1,2,3\}, 2^{\{1,2,3\}}, \nu_1)$ is obtained such that $\nu_1(\{1\}) = 0$, $\nu_1(\{2\}) = 0$, $\nu_1(\{3\}) = 1$. Therefore, [when the unknown [$*$] is ω_1]

$$(5.17) = |0-0| + |0-0.5| + |1-0.5| = 1,$$

[when the unknown [$*$] is ω_2]

$$(5.17) = |0-0| + |0-0| + |1-1| = 0$$

[when the unknown [$*$] is ω_3]

$$(5.17) = |0-0| + |0-1| + |1-0| = 2.$$

Thus, we can infer that $[*] = \omega_2$. That is, you should change to the Door 2.

$\square$

5.6 The two envelope problem —High school student puzzle—

This section is extracted from the following:

Ref. [46]: S. Ishikawa; The two envelopes paradox in non-Bayesian and Bayesian statistics (arXiv:1408.4916v4 [stat.OT] 2014)

Also, for a Bayesian approach to the two envelope problem, see Chapter 9.

5.6.1 Problem(the two envelope problem)

The following problem is the famous "two envelope problem(*cf.* [59])".

Problem 5.16. [The two envelope problem]
The host presents you with a choice between two envelopes (i.e., Envelope A and Envelope B). You know one envelope contains twice as much money as the other, but you do not know which contains more. That is, Envelope A [resp. Envelope B] contains V_1 dollars [resp. V_2 dollars]. You know that

(a) $\frac{V_1}{V_2} = 1/2$ or, $\frac{V_1}{V_2} = 2$

Define the exchanging map $\overline{x} : \{V_1, V_2\} \to \{V_1, V_2\}$ by

$$\overline{x} = \begin{cases} V_2, & (\text{ if } x = V_1), \\ V_1 & (\text{ if } x = V_2) \end{cases}$$

You choose randomly (by a fair coin toss) one envelope, and you get x_1 dollars (i.e., if you choose Envelope A [resp. Envelope B], you get V_1 dollars [resp. V_2 dollars]). And the host gets $\overline{x}_1$ dollars. Thus, you can infer that $\overline{x}_1 = 2x_1$ or $\overline{x}_1 = x_1/2$. Now the host says "You are offered the options of keeping your x_1 or switching to my $\overline{x}_1$". **What should you do?**

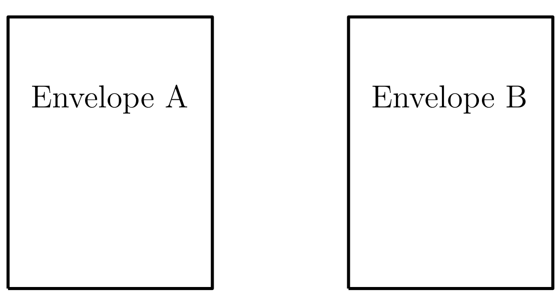


Figure 5.9: Two envelope problem

[(P1):Why is it paradoxical?]. You get $\alpha = x_1$. Then, you reason that, with probability 1/2, $\overline{x}_1$ is equal to either $\alpha/2$ or 2α dollars. Thus the expected value (denoted $E_{\text{other}}(\alpha)$ at this moment) of the other envelope is

$$E_{\text{other}}(\alpha) = (1/2)(\alpha/2) + (1/2)(2\alpha) = 1.25\alpha \tag{5.20}$$

This is greater than the α in your current envelope A. Therefore, you should switch to B. But this seems clearly wrong, as your information about A and B is symmetrical. This is the famous two-envelope paradox (i.e., "The Other Person's Envelope is Always Greener").

5.6.2 Answer: the two envelope problem 5.16

Consider the classical basic structure

$$[C_0(\Omega) \subseteq L^\infty(\Omega,\nu) \subseteq B(L^2(\Omega,\nu))]$$

where the locally compact space Ω is arbitrary, that is, it may be $\overline{\mathbb{R}}_+ = \{\omega \mid \omega \geq 0\}$ or the one point set $\{\omega_0\}$ or $\Omega = \{2^n \mid n = 0, \pm 1, \pm 2, \ldots\}$. Put $X = \overline{\mathbb{R}}_+ = \{x \mid x \geq 0\}$. Consider two continuous (or generally, measurable) functions $V_1 : \Omega \to \overline{\mathbb{R}}_+$ and $V_2 : \Omega \to \overline{\mathbb{R}}_+$. such that

$$V_2(\omega) = 2V_1(\omega) \text{ or, } \ 2V_2(\omega) = V_1(\omega) \quad (\forall \omega \in \Omega)$$

For each $k = 1, 2$, define the observable $\mathsf{O}_k = (X(=\overline{\mathbb{R}}_+), \mathcal{F}(=\mathcal{B}_{\overline{\mathbb{R}}_+} : \text{the Borel field}), F_k)$ in $L^\infty(\Omega,\nu)$ such that

$$[F_k(\Xi)](\omega) = \begin{cases} 1 & (\text{ if } V_k(\omega) \in \Xi) \\ 0 & (\text{ if } V_k(\omega) \notin \Xi) \end{cases}$$
$$(\forall \omega \in \Omega, \forall \Xi \in \mathcal{F} = \mathcal{B}_{\overline{\mathbb{R}}_+} \text{ i.e., the Bore field in } X(=\overline{\mathbb{R}}_+) \ )$$

Further, define the observable $\mathsf{O} = (X, \mathcal{F}, F)$ in $L^\infty(\Omega,\nu)$ such that

$$F(\Xi) = \frac{1}{2}\Big(F_1(\Xi) + F_2(\Xi)\Big) \quad (\forall \Xi \in \mathcal{F}) \tag{5.21}$$

That is,

$$[F(\Xi)](\omega) = \begin{cases} 1 & (\text{ if } V_1(\omega) \in \Xi, \ \ V_2(\omega) \in \Xi) \\ 1/2 & (\text{ if } V_1(\omega) \in \Xi, \ \ V_2(\omega) \notin \Xi) \\ 1/2 & (\text{ if } V_1(\omega) \notin \Xi, \ \ V_2(\omega) \in \Xi) \\ 0 & (\text{ if } V_1(\omega) \notin \Xi, \ \ V_2(\omega) \notin \Xi) \end{cases}$$
$$(\forall \omega \in \Omega, \forall \Xi \in \mathcal{F} = \mathcal{B}_X \text{ i.e., } \Xi \text{ is a Borel set in } X(=\overline{\mathbb{R}}_+) \ )$$

Fix a state $\omega(\in \Omega)$, which is assumed to be unknown. Consider the measurement $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O} = (X, \mathcal{F}, F), S_{[\omega]})$. Axiom 1 (§2.7) says that

(A$_1$) the probability that a measured value $\left\{ \begin{array}{c} V_1(\omega) \\ V_2(\omega) \end{array} \right\}$ is obtained by the measurement $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O}$ $= (X, \mathcal{F}, F), S_{[\omega]})$ is given by $\left\{ \begin{array}{c} 1/2 \\ 1/2 \end{array} \right\}$

If you switch to $\left\{ \begin{array}{c} V_2(\omega) \\ V_1(\omega) \end{array} \right\}$, your gain is $\left\{ \begin{array}{c} V_2(\omega) - V_1(\omega) = \omega \\ V_1(\omega) - V_2(\omega) = -\omega \end{array} \right\}$. Therefore, the expectation of switching is

$$(V_2(\omega) - V_1(\omega))/2 + (V_1(\omega) - V_2(\omega))/2 = 0$$

That is, it is wrong *"The Other Person's envelope is Always Greener"*.

Remark 5.17. The condition (a) in Problem 5.16 is not needed. This condition plays a role to confuse the essence of the problem.

5.6.3 Another answer: the two envelope problem 5.16

For the preparation of the following section (§ 5.6.4), consider the state space Ω such that

$$\Omega = \overline{\mathbb{R}}_+$$

with Lebesgue measure ν. Thus, we start from the classical basic structure

$$[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))]$$

Also, putting $\widehat{\Omega} = \{(\omega, 2\omega) \mid \omega \in \overline{\mathbb{R}}_+\}$, we consider the identification:

$$\Omega \ni \omega \underset{\text{(identification)}}{\longleftrightarrow} (\omega, 2\omega) \in \widehat{\Omega} \tag{5.22}$$

Further, define $V_1 : \Omega(\equiv \overline{\mathbb{R}}_+) \to X(\equiv \overline{\mathbb{R}}_+)$ and $V_2 : \Omega(\equiv \overline{\mathbb{R}}_+) \to X(\equiv \overline{\mathbb{R}}_+)$ such that

$$V_1(\omega) = \omega, \quad V_2(\omega) = 2\omega \qquad (\forall \omega \in \Omega)$$

And define the observable $\mathsf{O} = (X(= \overline{\mathbb{R}}_+), \mathcal{F}(= \mathcal{B}_{\overline{\mathbb{R}}_+} :$ the Borel field$), F)$ in $L^\infty(\Omega, \nu)$ such that

$$[F(\Xi)](\omega) = \begin{cases} 1 & (\text{ if } \omega \in \Xi, \ \ 2\omega \in \Xi) \\ 1/2 & (\text{ if } \omega \in \Xi, \ \ 2\omega \notin \Xi) \\ 1/2 & (\text{ if } \omega \notin \Xi, \ \ 2\omega \in \Xi) \\ 0 & (\text{ if } \omega \notin \Xi, \ \ 2\omega \notin \Xi) \end{cases} \qquad (\forall \omega \in \Omega, \forall \Xi \in \mathcal{F})$$

Fix a state $\omega (\in \Omega)$, which is assumed to be unknown. Consider the measurement $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O} = (X, \mathcal{F}, F), S_{[\omega]})$. Axiom 1 (measurement: §2.7) says that

(A_2) the probability that a measured value $\left\{ \begin{array}{l} x = V_1(\omega) = \omega \\ x = V_2(\omega) = 2\omega \end{array} \right\}$ is obtained by $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O} = (X, \mathcal{F}, F), S_{[\omega]})$ is given by $\left\{ \begin{array}{l} 1/2 \\ 1/2 \end{array} \right\}$

If you switch to $\left\{ \begin{array}{l} V_2(\omega) \\ V_1(\omega) \end{array} \right\}$, your gain is $\left\{ \begin{array}{l} V_2(\omega) - V_1(\omega) \\ V_1(\omega) - V_2(\omega) \end{array} \right\}$. Therefore, the expectation of switching is

$$(V_2(\omega) - V_1(\omega))/2 + (V_1(\omega) - V_2(\omega))/2 = 0$$

That is, it is wrong *"The Other Person's envelope is Always Greener"*.

Remark 5.18. The readers should note that Fisher's maximum likelihood method is not used in the two answers (in §5.6.2 and §5.6.3). If we try to apply Fisher's maximum likelihood method to Problem 5.16 (Two envelope problem), we get into a dead end. This is shown below.

5.6.4 Where do we mistake in (P1) of Problem 5.16?

Now we can answer to the question:

Where do we mistake in (P1) of Problem 5.16?

Let us explain it in what follows.
Assume that

(a) a measured value α is obtained by the measurement $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O} = (X, \mathcal{F}, F), S_{[*]})$

Then, we get the likelihood function $f(\alpha, \omega)$ such that

$$f(\alpha, \omega) \equiv \inf_{\omega_1 \in \Omega} \Big[\lim_{\Xi \to \{x\}, [F(\Xi)](\omega_1) \neq 0} \frac{[F(\Xi)](\omega)}{[F(\Xi)](\omega_1)} \Big] = \left\{ \begin{array}{ll} 1 & (\omega = \alpha/2 \text{ or } \alpha) \\ 0 & (\text{ elsewhere }) \end{array} \right.$$

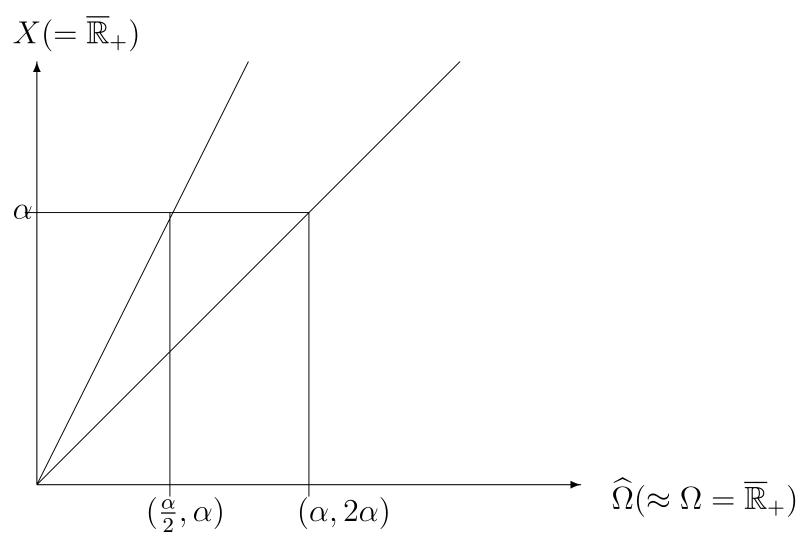


Figure 5.10: Two envelope problem

Therefore, Fisher's maximum likelihood method says that

(B$_1$) unknown state $[*]$ is equal to $\alpha/2$ or α

$\Big($ If $[*] = \alpha/2$ [resp. $[*] = \alpha$], then the switching gain is $(\alpha/2 - \alpha)$ [resp. $(2\alpha - \alpha)$] $\Big)$.

However, Fisher's maximum likelihood method does not say

$$(\mathrm{B}_2) \begin{cases} \text{“the probability that } [*] = \alpha/2\text{”}=1/2 \\ \text{“the probability that } [*] = \alpha\text{”}=1/2 \\ \text{“the probability that } [*] \text{ is otherwise”}=0 \end{cases}$$

Therefore, we can not calculate (such as (5.20)):

$$(\alpha/2 - \alpha) \times \frac{1}{2} + (2\alpha - \alpha) \times \frac{1}{2} = 1.25\alpha$$

(C$_1$) Thus, the sentence “with probability 1/2” in [(P1):Why is it paradoxical?] is wrong.

Hence, we can conclude that

(C$_2$) **If “state space” is specified, there will be no method of a mistake.**

since the state space is not declared in [(P1):Why is it paradoxical?].

After all, we see

(D) *If “state space” is specified, there will be no room to make a mistake.*

since the state space is not declared in [(P1):Why is it paradoxical ?].

Remark 5.19. The condition (b) in Problem 5.16 is indispensable. Without this condition, we can not difine the observable $\mathsf{O} = (X, \mathcal{F}, F)$ by the formula (5.23), and thus we can not solve Problem 5.16. However, it is usual to assume the principle of equal weight (i.e., *no information is interpreted as a fair coin toss*), or more precisely,

($\sharp$) *the principle that, in the absence of any reason to expect one event rather than another, all the possible events should be assigned the same probability*

Under this hypothesis, the condition (b) may be often omitted. Also, we will again discuss the principle of equal weight in Chapters 9 and 18.

♠**Note 5.5.** The readers may think that

($\sharp_1$) the answer of Problem 5.16 is a direct consequence of the fact that the information about A and B is symmetrical (as mentioned in [(P1): Why is it paradoxical?] in Problem 5.16). That is, it suffices to point out the symmetry.

This answer ($\sharp_1$) may not be wrong. But we think that the ($\sharp_1$) is not sufficient. That is because

($\sharp_2$) in the above answer ($\sharp_1$), the problem "What kind of theory (or, language, world view) is used?" is not clear. On the other hand, the answer presented in Section 5.6.2 is based on quantum language.

This is quite important. For example, someone may paradoxically assert that it is impossible to decide "Geocentric model vs. Heliocentrism", since motion is relative. However, we can say, at least, that

($\sharp_3$) Heliocentrism is more handy (than Geocentric model) under Newtonian mechanics.

That is, I think that

($\sharp_4$) Geocentric model may not be wrong under Aristotle's world view.

Therefore, I think that the true meaning of the Copernican revolution is

$$\boxed{\text{Aristotle's world view}} \xrightarrow[\text{(the Copernican revolution)}]{} \boxed{\text{Newtonian mechanical world view}} \tag{5.23}$$

and not

$$\boxed{\text{Geocentric model}} \xrightarrow[\text{(the Copernican revolution)}]{} \boxed{\text{Heliocentrism}} \tag{5.24}$$

Thus, this (5.24) is merely one of the symbolic events in the Copernican revolution (5.23). The readers should recall my only one assertion in this note, i.e., Figure 1.1 (The history of the world views).

KSTS/RR-16/001
January 8, 2016

Chapter 6

The confidence interval and statistical hypothesis testing

The standard university course of statistics is as follows:

$$\underset{\substack{\text{(maximum likelihood method)}\\ \text{(moment method)}}}{\overset{①}{\boxed{\text{Inference}}}} \longrightarrow \overset{②}{\boxed{\text{confidence interval}}} \longrightarrow \overset{③}{\boxed{\text{statistical hypothesis testing}}}$$

$$\longrightarrow \overset{④}{\boxed{\text{ANOVA (Analysis of Variance)}}}$$

In the previous chapter, we are concerned with ① (inference) in quantum language. In this chapter, we devote ourselves to ② and ③ (confidence interval and statistical hypothesis testing).

This chapter is extracted from

Ref. [40]: S. Ishikawa; A quantum linguistic characterization of the reverse relation between confidence interval and hypothesis testing (arXiv:1401.2709 [math.ST] 2014)

6.1 Review: classical quantum language(Axiom 1)

Firstly, we review classical measurement theory as follows.

(A): Axiom 1(measurement) classical pure type

(*cf.* This can be read under the preparation to §2.7))

With any classical system S, a basic structure $[C_0(\Omega) \subseteq L^\infty(\Omega,\nu) \subseteq B(L^2(\Omega,\nu))]$ can be associated in which measurement theory of that classical system can be formulated. In $[C_0(\Omega) \subseteq L^\infty(\Omega,\nu) \subseteq B(L^2(\Omega,\nu))]$, consider a **$W^*$-measurement** $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O}{=}(X,\mathcal{F},F), S_{[\delta_\omega]})$ $\Big($ or, **C^*-measurement** $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}{=}(X,\mathcal{F},F), S_{[\delta_\omega]})$ $\Big)$. That is, consider

- a W^*-measurement $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O}, S_{[\delta_\omega]})$ $\Big($ or, C^*-measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}{=}(X,\mathcal{F},F), S_{[\delta_\omega]})$ $\Big)$ of an **observable** $\mathsf{O}{=}(X,\mathcal{F},F)$ for a **state** $\delta_\omega(\in \mathcal{M}^p(\Omega)$: state space)

Then, the probability that a measured value x $(\in X)$ obtained by the W^*-measurement $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O}, S_{[\delta_\omega]})$ $\Big($ or, C^*-measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}{=}(X,\mathcal{F},F), S_{[\delta_\omega]})$ $\Big)$ belongs to Ξ $(\in \mathcal{F})$ is given by

$$\delta_\omega(F(\Xi))(\equiv [F(\Xi)](\omega) = {}_{\mathcal{M}(\Omega)}(\delta_\omega, F(\Xi))_{L^\infty(\Omega.\nu)})$$

(if $F(\Xi)$ is essentially continuous at δ_ω, or see Definition 2.14).

In this chapter, we devote ourselves to the simultaneous normal measurement as follows.

Example 6.1. [Normal observable]. Let $\mathbb{R}$ be the real axis. Define the state space $\Omega = \mathbb{R}\times\mathbb{R}_+$, where $\mathbb{R}_+ = \{\sigma \in \mathbb{R} | \sigma > 0\}$ with the Lebesgue measure ν. Consider the classical basic structure:

$$[C_0(\Omega) \subseteq L^\infty(\Omega,\nu) \subseteq B(L^2(\Omega,\nu))]$$

The normal observable $\mathsf{O}_G = (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, G)$ in $L^\infty(\Omega(\equiv \mathbb{R}\times\mathbb{R}_+))$ is defined by

$$[G(\Xi)](\omega) = \frac{1}{\sqrt{2\pi}\sigma}\int_\Xi \exp[-\frac{(x-\mu)^2}{2\sigma^2}]dx \tag{6.1}$$
$$(\forall \Xi \in \mathcal{B}_{\mathbb{R}}(= \text{the Borel field in } \mathbb{R})), \quad \forall\omega = (\mu,\sigma) \in \Omega = \mathbb{R}\times\mathbb{R}_+).$$

Example 6.2. [Simultaneous normal observable]. Let n be a natural number. Let $\mathsf{O}_G = (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, G)$ be the normal observable in $L^\infty(\mathbb{R}\times\mathbb{R}_+)$. Define the n-th simultaneous normal observable $\mathsf{O}_G^n = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n, G^n)$ in $L^\infty(\mathbb{R}\times\mathbb{R}_+)$ such that

$$\begin{aligned}&[G^n(\times_{k=1}^n \Xi_k)](\omega) = \times_{k=1}^n [G(\Xi_k)](\omega)\\ =&\frac{1}{(\sqrt{2\pi}\sigma)^n}\underset{\times_{k=1}^n \Xi_k}{\int\cdots\int} \exp[-\frac{\sum_{k=1}^n (x_k-\mu)^2}{2\sigma^2}]dx_1dx_2\cdots dx_n\end{aligned} \tag{6.2}$$

$$(\forall \Xi_k \in \mathcal{B}_{\mathbb{R}}(k=1,2,\ldots,n), \quad \forall \omega=(\mu,\sigma) \in \Omega = \mathbb{R}\times\mathbb{R}_+).$$

Thus, we have the simultaneous normal measurement $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}(\mathsf{O}_G^n = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n, G^n), S_{[(\mu,\sigma)]})$.

Consider the maps $\overline{\mu}:\mathbb{R}^n \to \mathbb{R}$, $\overline{SS}:\mathbb{R}^n \to \mathbb{R}$ and $\overline{\sigma}:\mathbb{R}^n \to \mathbb{R}$ such that

$$\overline{\mu}(x)=\overline{\mu}(x_1,x_2,\ldots,x_n)=\frac{x_1+x_2+\cdots+x_n}{n} \quad (\forall x=(x_1,x_2,\ldots,x_n)\in\mathbb{R}^n) \tag{6.3}$$

$$\overline{SS}(x)=\overline{SS}(x_1,x_2,\ldots,x_n)=\sum_{k=1}^n (x_k-\overline{\mu}(x))^2 \quad (\forall x=(x_1,x_2,\ldots,x_n)\in\mathbb{R}^n) \tag{6.4}$$

$$\overline{\sigma}(x)=\overline{\sigma}(x_1,x_2,\ldots,x_n)=\sqrt{\frac{\sum_{k=1}^n (x_k-\overline{\mu}(x))^2}{n}} \quad (\forall x=(x_1,x_2,\ldots,x_n)\in\mathbb{R}^n) \tag{6.5}$$

Therefore, we get and calculate (by the formulas of Gauss integrals (in § 7.4)) two image observables $\overline{\mu}(\mathsf{O}_G^n)=(\mathbb{R},\mathcal{B}_{\mathbb{R}}, G^n\circ\overline{\mu}^{-1})$ and $\overline{SS}(\mathsf{O}_G^n)=(\mathbb{R}_+,\mathcal{B}_{\mathbb{R}_+}, G^n\circ\overline{SS}^{-1})$ in $L^\infty(\mathbb{R}\times\mathbb{R}_+)$ as follows.

$$\begin{aligned}
&[(G^n\circ\overline{\mu}^{-1})(\Xi_1)](\omega)\\
=&\frac{1}{(\sqrt{2\pi}\sigma)^n}\underset{\{x\in\mathbb{R}^n\,:\,\overline{\mu}(x)\in\Xi_1\}}{\int\cdots\int}\exp[-\frac{\sum_{k=1}^n (x_k-\mu)^2}{2\sigma^2}]dx_1dx_2\cdots dx_n\\
=&\frac{\sqrt{n}}{\sqrt{2\pi}\sigma}\int_{\Xi_1}\exp[-\frac{n(x-\mu)^2}{2\sigma^2}]dx\\
&(\forall \Xi_1\in\mathcal{B}_{\mathbb{R}}, \qquad \forall\omega=(\mu,\sigma)\in\Omega\equiv\mathbb{R}\times\mathbb{R}_+).
\end{aligned} \tag{6.6}$$

and,

$$\begin{aligned}
&[(G^n\circ\overline{SS}^{-1})(\Xi_2)](\omega)\\
=&\frac{1}{(\sqrt{2\pi}\sigma)^n}\underset{\{x\in\mathbb{R}^n\,:\,\overline{SS}(x)\in\Xi_2\}}{\int\cdots\int}\exp[-\frac{\sum_{k=1}^n (x_k-\mu)^2}{2\sigma^2}]dx_1dx_2\cdots dx_n\\
=&\int_{\Xi_2/\sigma^2} p_{n-1}^{\chi^2}(x)dx\\
&(\ \forall \Xi_2\in\mathcal{B}_{\mathbb{R}_+}, \quad \forall\omega=(\mu,\sigma)\in\Omega\equiv\mathbb{R}\times\mathbb{R}_+).
\end{aligned} \tag{6.7}$$

where $p_{n-1}^{\chi^2}(x)$ is the probability density function of χ^2-distribution with $(n-1)$ degree of freedom. That is,

$$p_{n-1}^{\chi^2}(x)=\frac{x^{(n-1)/2-1}e^{-x/2}}{2^{(n-1)/2}\Gamma((n-1)/2)} \quad (x>0) \tag{6.8}$$

where, Γ is the Gamma function.

6.2 The reverse relation between confidence interval method and statistical hypothesis testing

In what follows, we shall mention the reverse relation (such as "the two sides of a coin") between confidence interval method and statistical hypothesis testing.

We devote ourselves to the classical systems, i.e., the classical basic structure:

$$[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))]$$

6.2.1 The confidence interval method

Consider an observable $\mathsf{O} = (X, \mathcal{F}, F)$ in $L^\infty(\Omega)$. Let Θ be a locally compact space (called the second state space), which has the semi-metric d^x_Θ ($\forall x \in X$) such that,

($\sharp$) for each $x \in X$, the map $d^x_\Theta : \Theta^2 \to [0, \infty)$ satisfies (i):$d^x_\Theta(\theta, \theta) = 0$, (ii):$d^x_\Theta(\theta_1, \theta_2) = d^x_\Theta(\theta_2, \theta_1)$, (ii):$d^x_\Theta(\theta_1, \theta_3) \le d^x_\Theta(\theta_1, \theta_2) + d^x_\Theta(\theta_2, \theta_3)$.

Further, consider two maps $E : X \to \Theta$ and $\pi : \Omega \to \Theta$. Here, $E : X \to \Theta$ and $\pi : \Omega \to \Theta$ is respectively called an estimator and a system quantity.

Theorem 6.3. [Confidence interval method]. Let a positive number α be $0 < \alpha \ll 1$, for example, $\alpha = 0.05$. For any state $\omega(\in \Omega)$, define the positive number $\delta^{1-\alpha}_\omega$ (> 0) such that:

$$\delta^{1-\alpha}_\omega = \inf\{\delta > 0 : [F(\{x \in X \ : \ d^x_\Theta(E(x), \pi(\omega)) < \delta\})](\omega) \ge 1 - \alpha\} \tag{6.9}$$

Then we say that:

(A) the probability, that the measured value x obtained by the measurement $\mathsf{M}_{L^\infty(\Omega)}\big(\mathsf{O} := (X, \mathcal{F}, F),\ S_{[\omega_0]}\big)$ satisfies the following condition (6.10), is more than or equal to $1 - \alpha$ (e.g., $1 - \alpha = 0.95$).

$$d^x_\Theta(E(x), \pi(\omega_0)) \le \delta^{1-\alpha}_{\omega_0} \tag{6.10}$$

And further, put

$$D^{1-\alpha,\Theta}_x = \{\pi(\omega)(\in \Theta) : d^x_\Theta(E(x), \pi(\omega)) \le \delta^{1-\alpha}_\omega\}. \tag{6.11}$$

which is called *the* $(1 - \alpha)$*-confidence interval.* Here, we see the following equivalence:

$$(6.10) \iff D^{1-\alpha,\Theta}_x \ni \pi(\omega_0). \tag{6.12}$$

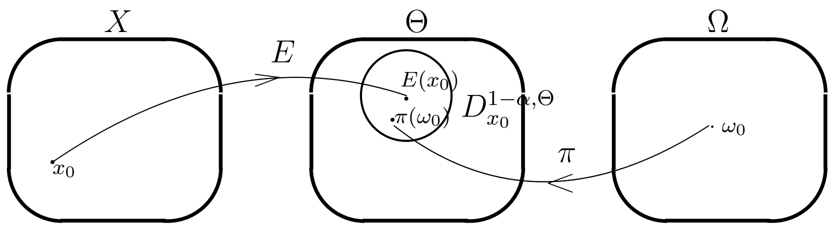


Figure 6.1 Confidence interval $D_{x_0}^{1-\alpha,\Theta}$

Remark 6.4. [(B_1):The meaning of confidence interval]. Consider the parallel measurement $\bigotimes_{j=1}^J \mathsf{M}_{L^\infty(\Omega)}(\mathsf{O} := (X, \mathcal{F}, F), S_{[\omega_0]})$, and assume that a measured value $x = (x_1, x_2, \ldots, x_J)(\in X^J)$ is obtained by the parallel measurement. Recall the formula (6.12). Then, it surely holds that

$$\lim_{J \to \infty} \frac{\mathrm{Num}[\{j \mid D_{x_j}^{1-\alpha,\Theta} \ni \pi(\omega_0)]}{J} \geq 1 - \alpha (= 0.95) \tag{6.13}$$

where Num$[A]$ is the number of the elements of the set A. Hence Theorem 6.3 can be tested by numerical analysis (with random number). Similarly, Theorem 6.5 (mentioned later) can be tested.

[(B_2)] Also, note that

$$\begin{aligned}(6.9) = \delta_\omega^{1-\alpha} &= \inf\{\delta > 0 : [F(\{x \in X \ : \ d_\Theta^x(E(x), \pi(\omega)) < \delta\})](\omega) \geq 1 - \alpha\} \\ &= \inf\{\eta > 0 : [F(\{x \in X \ : \ d_\Theta^x(E(x), \pi(\omega)) \geq \eta\})](\omega) \leq \alpha\}\end{aligned} \tag{6.14}$$

6.2.2 Statistical hypothesis testing

Next, we shall explain the statistical hypothesis testing, which is characterized as the reverse of the confident interval method.

Theorem 6.5. [Statistical hypothesis testing]. Let α be a real number such that $0 < \alpha \ll 1$, for example, $\alpha = 0.05$. For any state $\omega(\in \Omega)$, define the positive number η_ω^α (> 0) such that:

$$\eta_\omega^\alpha = \inf\{\eta > 0 : [F(\{x \in X \ : \ d_\Theta^x(E(x), \pi(\omega)) \geq \eta\})](\omega) \leq \alpha\} \tag{6.15}$$
$$\text{(by the (6.14), note that } \delta_\omega^{1-\alpha} = \eta_\omega^\alpha)$$

Then we say that:

(C) the probability, that the measured value x obtained by the measurement $\mathsf{M}_{L^\infty(\Omega)}\big(\mathsf{O} := (X, \mathcal{F}, F),\ S_{[\omega_0]}\big)$ satisfies the following condition (6.16), is less than or equal to α (e.g., $\alpha = 0.05$).

$$d^x_\Theta(E(x), \pi(\omega_0)) \geq \eta^\alpha_{\omega_0}. \tag{6.16}$$

Further, consider a subset H_N of Θ, which is called a "null hypothesis". Put

$$\widehat{R}^{\alpha,\Theta}_{H_N} = \bigcap_{\omega \in \Omega \text{ such that } \pi(\omega) \in H_N} \{E(x)(\in \Theta) : d^x_\Theta(E(x), \pi(\omega)) \geq \eta^\alpha_\omega\}. \tag{6.17}$$

which is called *the* (α)*-rejection region of the null hypothesis* H_N. Then we say that:

(D) the probability, that the measured value x obtained by the measurement $\mathsf{M}_{L^\infty(\Omega)}\big(\mathsf{O} := (X, \mathcal{F}, F),\ S_{[\omega_0]}\big)$ (where $\pi(\omega_0) \in H_N$) satisfies the following condition (6.18), is less than or equal to α (e.g., $\alpha = 0.05$).

$$\widehat{R}^\alpha_{H_N} \ni E(x). \tag{6.18}$$

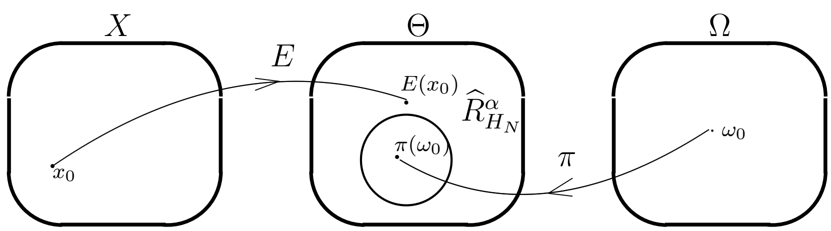


Figure 6.2: Rejection region $\widehat{R}^\alpha_{H_N}$ (when $H_N = \{\pi(\omega_0)\}$

Corollary 6.6. [The reverse relation between Confidence interval and statistical hypothesis testing]. Let $0 < \alpha \ll 1$. Consider an observable $\mathsf{O} = (X, \mathcal{F}, F)$ in $L^\infty(\Omega)$, and the second state space Θ (i.e., locally compact space with a semi-metric $d^x_\Theta (x \in X)$). And consider the estimator $E : X \to \Theta$ and the system quantity $\pi : \Omega \to \Theta$. Define $\delta^{1-\alpha}_\omega$ by (6.9), and define η^α_ω by (6.15) (and thus, $\delta^{1-\alpha}_\omega = \eta^\alpha_\omega$).

(E) [Confidence interval method]. for each $x \in X$, define $(1-\alpha)$-confidence interval by

$$D^{1-\alpha,\Theta}_x = \{\pi(\omega)(\in \Theta) : d^x_\Theta(E(x), \pi(\omega)) < \delta^{1-\alpha}_\omega\} \tag{6.19}$$

Also,

$$D^{1-\alpha,\Omega}_x = \{\omega(\in \Omega) : d^x_\Theta(E(x), \pi(\omega)) < \delta^{1-\alpha}_\omega\} \tag{6.20}$$

KSTS/RR-16/001
January 8, 2016

Here, assume that a measured value $x(\in X)$ is obtained by the measurement $\mathsf{M}_{L^\infty(\Omega)}\big(\mathsf{O} := (X, \mathcal{F}, F),\ S_{[\omega_0]}\big)$. Then, we see that

(E$_1$) the probability that

$$D_x^{1-\alpha,\Theta} \ni \pi(\omega_0) \qquad \text{or, in the same sense} \qquad D_x^{1-\alpha,\Omega} \ni \omega_0$$

is more than $1-\alpha$.

(F) [statistical hypothesis testing]. Consider the null hypothesis $H_N(\subseteq \Theta)$. Assume that the state $\omega_0(\in \Omega)$ satisfies:

$$\pi(\omega_0) \in H_N(\subseteq \Theta)$$

Here, put,

$$\widehat{R}_{H_N}^{\alpha;\Theta} = \bigcap_{\omega \in \Omega \text{ such that } \pi(\omega) \in H_N} \{E(x)(\in \Theta) : d_\Theta^x(E(x), \pi(\omega)) \geq \eta_\omega^\alpha\}. \tag{6.21}$$

or,

$$\widehat{R}_{H_N}^{\alpha;X} = E^{-1}(\widehat{R}_{H_N}^{\alpha;\Theta}) = \bigcap_{\omega \in \Omega \text{ such that } \pi(\omega) \in H_N} \{x(\in X) : d_\Theta^x(E(x), \pi(\omega)) \geq \eta_\omega^\alpha\}. \tag{6.22}$$

which is called the (α)-rejection region of the null hypothesis H_N.

Assume that a measured value $x(\in X)$ is obtained by the measurement $\mathsf{M}_{L^\infty(\Omega)}\big(\mathsf{O} := (X, \mathcal{F}, F),\ S_{[\omega_0]}\big)$. Then, we see that

(F$_1$) the probability that

$$\text{“}E(x) \in \widehat{R}_{H_N}^{\alpha;\Theta}\text{”} \qquad \text{or, in the same sense,} \qquad \text{“}x \in \widehat{R}_{H_N}^{\alpha;X}\text{”} \tag{6.23}$$

is less than α.

6.3 Confidence interval and statistical hypothesis testing for population mean

Consider the classical basic structure:

$$[C_0(\Omega) \subseteq L^\infty(\Omega,\nu) \subseteq B(L^2(\Omega,\nu))]$$

Fix a positive number α such that $0 < \alpha \ll 1$, for example, $\alpha = 0.05$.

6.3.1 Preparation (simultaneous normal measurement)

Example 6.7. Consider the simultaneous normal measurement $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}(\mathsf{O}_G^n = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n, G^n), S_{[(\mu,\sigma)]})$ in $L^\infty(\mathbb{R}\times\mathbb{R}_+)$. Here, the simultaneous normal observable $\mathsf{O}_G^n = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n, G^n)$ is defined by

$$\begin{aligned}
&[G^n(\times_{k=1}^n \Xi_k)](\omega) = \times_{k=1}^n [G(\Xi_k)](\omega) \\
=&\frac{1}{(\sqrt{2\pi}\sigma)^n} \underset{\times_{k=1}^n \Xi_k}{\int \cdots \int} \exp[-\frac{\sum_{k=1}^n (x_k-\mu)^2}{2\sigma^2}] dx_1 dx_2 \cdots dx_n \\
&(\forall \Xi_k \in \mathcal{B}_{\mathbb{R}} (k=1,2,\ldots,n), \quad \forall \omega = (\mu,\sigma) \in \Omega = \mathbb{R}\times\mathbb{R}_+).
\end{aligned} \tag{6.24}$$

Therefore, the state space Ω and the measured value space X are defined by

$$\Omega = \mathbb{R}\times\mathbb{R}_+$$
$$X = \mathbb{R}^n$$

Also, the second state space Θ is defined by

$$\Theta = \mathbb{R}$$

The estimator $E : \mathbb{R}^n \to \Theta(\equiv \mathbb{R})$ and the system quantity$\pi : \Omega \to \Theta$ are respectively defined by

$$E(x) = E(x_1, x_2, \ldots, x_n) = \overline{\mu}(x) = \frac{x_1 + x_2 + \cdots + x_n}{n}$$
$$\Omega = \mathbb{R}\times\mathbb{R}_+ \ni \omega = (\mu,\sigma) \mapsto \pi(\omega) = \mu \in \Theta = \mathbb{R}$$

Also, the semi-metric $d_\Theta^{(1)}$ in Θ is defined by

$$d_\Theta^{(1)}(\theta_1,\theta_2) = |\theta_1 - \theta_2| \qquad (\forall \theta_1, \theta_2 \in \Theta = \mathbb{R})$$

6.3.2 Confidence interval

Our present problem is as follows.

Problem 6.8. [Confidence interval]. Consider the simultaneous normal measurement $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}$ $(\mathsf{O}_G^n = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n, G^n),\ S_{[(\mu,\sigma)]})$. Assume that a measured value$x \in X = \mathbb{R}^n$ is obtained by the measurement. Let $0 < \alpha \ll 1$.
Then, find the $D_x^{1-\alpha;\Theta}(\subseteq \Theta)$ (which may depend on σ) such that

- the probability that $\mu \in D_x^{1-\alpha;\Theta}$ is more than $1-\alpha$.

Here, the more $D_x^{1-\alpha;\Theta}(\subseteq \Theta)$ is small, the more it is desirable.

Consider the following semi-distance $d_\Omega^{(1)}$ in the state space $\mathbb{R}\times\mathbb{R}_+$:

$$d_\Omega^{(1)}((\mu_1,\sigma_1),(\mu_2,\sigma_2)) = |\mu_1 - \mu_2| \tag{6.25}$$

For any $\omega = (\mu,\sigma)(\in \Omega = \mathbb{R}\times\mathbb{R}_+)$, define the positive number $\delta_\omega^{1-\alpha}$ (>0) such that:

$$\delta_\omega^{1-\alpha} = \inf\{\eta > 0 : [F(E^{-1}(\mathrm{Ball}_{d_\Omega^{(1)}}(\omega;\eta))](\omega) \geq 1-\alpha\}$$

where $\mathrm{Ball}_{d_\Omega^{(1)}}(\omega;\eta) = \{\omega_1(\in\Omega) : d_\Omega^{(1)}(\omega,\omega_1) \leq \eta\} = [\mu-\eta,\mu+\eta]\times\mathbb{R}_+$

Hence we see that

$$\begin{aligned}
&E^{-1}(\mathrm{Ball}_{d_\Omega^{(1)}}(\omega;\eta)) = E^{-1}([\mu-\eta,\mu+\eta]\times\mathbb{R}_+)\\
=&\{(x_1,\ldots,x_n)\in\mathbb{R}^n \ :\ \mu-\eta \leq \frac{x_1+\ldots+x_n}{n} \leq \mu+\eta\}
\end{aligned} \tag{6.26}$$

Thus,

$$\begin{aligned}
&[G^n(E^{-1}(\mathrm{Ball}_{d_\Omega^{(1)}}(\omega;\eta))](\omega)\\
=&\frac{1}{(\sqrt{2\pi}\sigma)^n}\underset{\mu-\eta\leq\frac{x_1+\ldots+x_n}{n}\leq\mu+\eta}{\int\cdots\int}\exp[-\frac{\sum_{k=1}^n(x_k-\mu)^2}{2\sigma^2}]dx_1dx_2\cdots dx_n\\
=&\frac{1}{(\sqrt{2\pi}\sigma)^n}\underset{-\eta\leq\frac{x_1+\ldots+x_n}{n}\leq\eta}{\int\cdots\int}\exp[-\frac{\sum_{k=1}^n(x_k)^2}{2\sigma^2}]dx_1dx_2\cdots dx_n\\
=&\frac{\sqrt{n}}{\sqrt{2\pi}\sigma}\int_{-\eta}^{\eta}\exp[-\frac{nx^2}{2\sigma^2}]dx = \frac{1}{\sqrt{2\pi}}\int_{-\sqrt{n}\eta/\sigma}^{\sqrt{n}\eta/\sigma}\exp[-\frac{x^2}{2}]dx
\end{aligned} \tag{6.27}$$

Solving the following equation:

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-z(\alpha/2)}\exp[-\frac{x^2}{2}]dx = \frac{1}{\sqrt{2\pi}}\int_{z(\alpha/2)}^{\infty}\exp[-\frac{x^2}{2}]dx = \frac{\alpha}{2} \tag{6.28}$$

we define that

$$\delta_\omega^{1-\alpha} = \frac{\sigma}{\sqrt{n}} z(\frac{\alpha}{2}) \tag{6.29}$$

Then, for any x ($\in \mathbb{R}^n$), we get $D_x^{1-\alpha,\Omega}$ (the $(1-\alpha)$-confidence interval of x) as follows:

$$\begin{aligned} D_x^{1-\alpha,\Omega} &= \{\omega(\in \Omega) : d_\Omega(E(x), \omega) \le \delta_\omega^{1-\alpha}\} \\ &= \{(\mu, \sigma) \in \mathbb{R} \times \mathbb{R}_+ \ : \ |\mu - \overline{\mu}(x)| = |\mu - \frac{x_1 + \ldots + x_n}{n}| \le \frac{\sigma}{\sqrt{n}} z(\frac{\alpha}{2})\} \end{aligned} \tag{6.30}$$

Also,

$$\begin{aligned} D_x^{1-\alpha,\Theta} &= \{\pi(\omega)(\in \Theta) : d_\Omega(E(x), \omega) \le \delta_\omega^{1-\alpha}\} \\ &= \{\mu \in \mathbb{R} \ : \ |\mu - \overline{\mu}(x)| = |\mu - \frac{x_1 + \ldots + x_n}{n}| \le \frac{\sigma}{\sqrt{n}} z(\frac{\alpha}{2})\} \end{aligned}$$

which depends on σ.

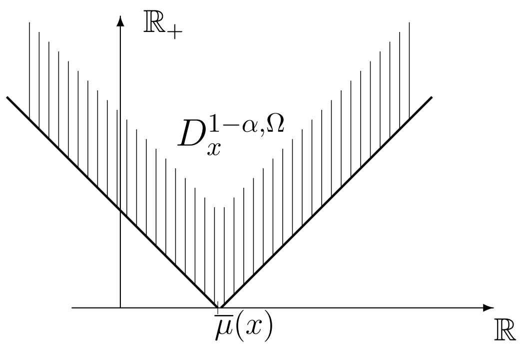


Figure 6.3: Confidence interval $D_x^{1-\alpha,\Omega}$ for the semi-distance $d_\Omega^{(1)}$

6.3.3 Statistical hypothesis testing[null hypothesis$H_N = \{\mu_0\}(\subseteq \Theta = \mathbb{R})$]

Problem 6.9. [Statistical hypothesis testing]. Consider the simultaneous normal measurement $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}$ $(\mathsf{O}_G^n = (\mathbb{R}^n, \mathcal{B}_\mathbb{R}^n, G^n),\ S_{[(\mu,\sigma)]})$. Assume the null hypothesis H_N such that

$$H_N = \{\mu_0\}(\subseteq \Theta = \mathbb{R}))$$

Let $0 < \alpha \ll 1$.
Then, find the rejection region $\widehat{R}_{H_N}^{\alpha;\Theta}(\subseteq \Theta)$ (which may depend on σ) such that

- the probability that a measured value $x(\in \mathbb{R}^n)$ obtained by $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}$ $(\mathsf{O}_G^n = (\mathbb{R}^n, \mathcal{B}_\mathbb{R}^n, G^n),\ S_{[(\mu_0,\sigma)]})$ satisfies that

$$E(x) \in \widehat{R}_{H_N}^{\alpha;\Theta}$$

is less than α.

Here, the more the rejection region $\widehat{R}_{H_N}^{\alpha;\Theta}$ is large, the more it is desirable.

Define the null hypothesis H_N such that

$$H_N = \{\mu_0\}(\subseteq \Theta(=\mathbb{R}))$$

For any $\omega = (\mu, \sigma)(\in \Omega = \mathbb{R} \times \mathbb{R}_+)$, define the positive number η^α_ω (>0) such that:

$$\eta^\alpha_\omega = \inf\{\eta > 0 : [F(E^{-1}(\mathrm{Ball}^C_{d^{(1)}_\Theta}(\pi(\omega);\eta))](\omega) \le \alpha\}$$

where $\mathrm{Ball}^C_{d^{(1)}_\Theta}(\pi(\omega);\eta) = \{\theta(\in\Theta) : d^{(1)}_\Theta(\mu,\theta) \ge \eta\} = \Big((-\infty, \mu-\eta] \cup [\mu+\eta, \infty)\Big)$

Hence we see that

$$\begin{aligned}
&E^{-1}(\mathrm{Ball}^C_{d^{(1)}_\Theta}(\pi(\omega);\eta)) = E^{-1}\Big((-\infty, \mu-\eta] \cup [\mu+\eta, \infty)\Big)\\
=&\{(x_1,\ldots,x_n) \in \mathbb{R}^n \ : \ \frac{x_1+\ldots+x_n}{n} \le \mu-\eta \text{ or } \mu+\eta \le \frac{x_1+\ldots+x_n}{n}\}\\
=&\{(x_1,\ldots,x_n) \in \mathbb{R}^n \ : \ |\frac{(x_1-\mu)+\ldots+(x_n-\mu)}{n}| \ge \eta\}
\end{aligned} \tag{6.31}$$

Thus,

$$\begin{aligned}
&[G^n(E^{-1}(\mathrm{Ball}^C_{d^{(1)}_\Theta}(\pi(\omega);\eta))](\omega)\\
=&\frac{1}{(\sqrt{2\pi}\sigma)^n} \underset{|\frac{(x_1-\mu)+\ldots+(x_n-\mu)}{n}|\ge\eta}{\int\cdots\int} \exp[-\frac{\sum_{k=1}^n (x_k-\mu)^2}{2\sigma^2}]dx_1 dx_2 \cdots dx_n\\
=&\frac{1}{(\sqrt{2\pi}\sigma)^n} \underset{|\frac{x_1+\ldots+x_n}{n}|\ge\eta}{\int\cdots\int} \exp[-\frac{\sum_{k=1}^n (x_k)^2}{2\sigma^2}]dx_1 dx_2 \cdots dx_n\\
=&\frac{\sqrt{n}}{\sqrt{2\pi}\sigma}\int_{x\ge\eta} \exp[-\frac{nx^2}{2\sigma^2}]dx = \frac{1}{\sqrt{2\pi}}\int_{x\ge\sqrt{n}\eta/\sigma} \exp[-\frac{x^2}{2}]dx
\end{aligned} \tag{6.32}$$

Solving the following equation:

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-z(\alpha/2)} \exp[-\frac{x^2}{2}]dx = \frac{1}{\sqrt{2\pi}}\int_{z(\alpha/2)}^{\infty} \exp[-\frac{x^2}{2}]dx = \frac{\alpha}{2} \tag{6.33}$$

we define that

$$\eta^\alpha_\omega = \frac{\sigma}{\sqrt{n}} z(\frac{\alpha}{2}) \tag{6.34}$$

Therefore, we get $\widehat{R}^{\alpha}_{H_N}$ (the (α)-rejection region of $H_N(=\{\mu_0\} \subseteq \Theta(=\mathbb{R}))$) as follows:

$$
\begin{aligned}
\widehat{R}^{\alpha,\Theta}_{\{\mu_0\}} &= \bigcap_{\pi(\omega)=\mu\in\{\mu_0\}} \{E(x)(\in \Theta=\mathbb{R}) : d^{(1)}_{\Theta}(E(x),\pi(\omega)) \geq \eta^{\alpha}_{\omega}\} \\
&= \{E(x)(=\frac{x_1+\ldots+x_n}{n}) \in \mathbb{R} \ : \ \overline{\mu}(x) - \mu_0 = \frac{x_1+\ldots+x_n}{n} - \mu_0 \geq \frac{\sigma}{\sqrt{n}} z(\frac{\alpha}{2})\} \qquad (6.35)
\end{aligned}
$$

Remark 6.10. Note that the $\widehat{R}^{\alpha,\Theta}_{\{\mu_0\}}$ (the (α)-rejection region of $\{\mu_0\}$) depends on σ. Thus, putting

$$
\widehat{R}^{\alpha}_{\{\mu_0\}\times\mathbb{R}_+} = \{(\overline{\mu}(x),\sigma) \in \mathbb{R}\times\mathbb{R}_+ \ : \ |\mu_0 - \overline{\mu}(x)| = |\mu_0 - \frac{x_1+\ldots+x_n}{n}| \geq \frac{\sigma}{\sqrt{n}} z(\frac{\alpha}{2})\} \qquad (6.36)
$$

we see that $\widehat{R}^{\alpha}_{\{\mu_0\}\times\mathbb{R}_+}$="the slash part in Figure 6.4".

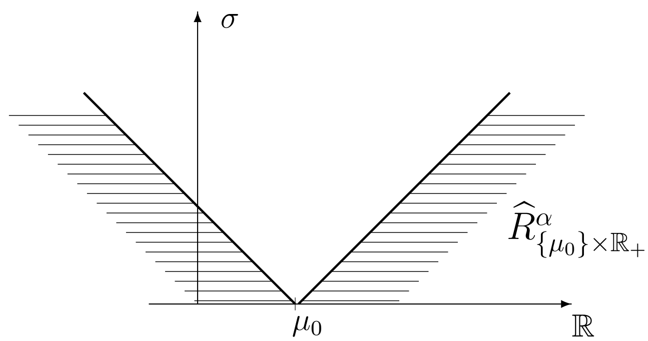


Figure 6.4: Rejection region $\widehat{R}^{\alpha}_{\{\mu_0\}}$ (which depends on σ)

6.3.4 Statistical hypothesis testing[null hypothesis$H_N = (-\infty,\mu_0](\subseteq \Theta(=\mathbb{R}))$]

Our present problem was as follows

Problem 6.11. [Statistical hypothesis testing]. Consider the simultaneous normal measurement $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}$ $(\mathsf{O}^n_G = (\mathbb{R}^n, \mathcal{B}^n_{\mathbb{R}}, G^n),\ S_{[(\mu,\sigma)]})$. Assume the null hypothesis H_N such that

$$
H_N = (-\infty,\mu_0](\subseteq \Theta=\mathbb{R}))
$$

Let $0<\alpha \ll 1$.
Then, find the rejection region $\widehat{R}^{\alpha;\Theta}_{H_N}(\subseteq \Theta)$ (which may depend on σ) such that

- the probability that a measured value $x(\in \mathbb{R}^n)$ obtained by $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}$ $(\mathsf{O}^n_G =$

$(\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n, G^n)$, $S_{[(\mu_0,\sigma)]})$ satisfies that

$$E(x) \in \widehat{R}_{H_N}^{\alpha;\Theta}$$

is less than α.

Here, the more the rejection region $\widehat{R}_{H_N}^{\alpha;\Theta}$ is large, the more it is desirable.

[Rejection region of $H_N = (-\infty, \mu_0] \subseteq \Theta(= \mathbb{R})$]. Consider the simultaneous measurement $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}$ $(\mathsf{O}_N^n = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n, G^n)$, $S_{[(\mu,\sigma)]})$ in $L^\infty(\mathbb{R} \times \mathbb{R}_+)$. Thus, we consider that $\Omega = \mathbb{R} \times \mathbb{R}$, $X = \mathbb{R}^n$. Assume that the real σ in a state $\omega = (\mu, \sigma) \in \Omega$ is fixed and known. Put

$$\Theta = \mathbb{R}$$

The formula (6.3) urges us to define the estimator $E : \mathbb{R}^n \to \Theta(\equiv \mathbb{R})$ such that

$$E(x) == \overline{\mu}(x) = \frac{x_1 + x_2 + \cdots + x_n}{n} \tag{6.37}$$

And consider the quantity $\pi : \Omega \to \Theta$ such that

$$\Omega = \mathbb{R} \times \mathbb{R}_+ \ni \omega = (\mu, \sigma) \mapsto \pi(\omega) = \mu \in \Theta = \mathbb{R}$$

Consider the following semi-distance $d_\Theta^{(2)}$ in $\Theta(= \mathbb{R})$:

$$d_\Theta^{(2)}((\theta_1, \theta_2) = \begin{cases} |\theta_1 - \theta_2| & \theta_0 \le \theta_1, \theta_2 \\ |\theta_2 - \theta_0| & \theta_1 \le \theta_0 \le \theta_2 \\ |\theta_1 - \theta_0| & \theta_2 \le \theta_0 \le \theta_1 \\ 0 & \theta_1, \theta_2 \le \theta_0 \end{cases} \tag{6.38}$$

Define the null hypothesis H_N such that

$$H_N = (-\infty, \mu_0](\subseteq \Theta(= \mathbb{R}))$$

For any $\omega = (\mu, \sigma)(\in \Omega = \mathbb{R} \times \mathbb{R}_+)$, define the positive number η_ω^α (> 0) such that:

$$\eta_\omega^\alpha = \inf\{\eta > 0 : [F(E^{-1}(\mathrm{Ball}_{d_\Theta^{(2)}}^C(\pi(\omega); \eta))](\omega) \le \alpha\}$$

where $\mathrm{Ball}_{d_\Theta^{(2)}}^C(\pi(\omega); \eta) = \{\theta(\in \Theta) : d_\Theta^{(2)}(\mu, \theta) \ge \eta\} = \Big((-\infty, \mu - \eta] \cup [\mu + \eta, \infty)\Big)$

Hence we see that

$$\begin{aligned} & E^{-1}(\mathrm{Ball}_{d_\Theta^{(2)}}^C(\pi(\omega); \eta)) = E^{-1}\Big([\mu + \eta, \infty)\Big) \\ =& \{(x_1, \ldots, x_n) \in \mathbb{R}^n \ : \ \mu + \eta \le \frac{x_1 + \ldots + x_n}{n}\} \end{aligned}$$

$$=\{(x_1,\ldots,x_n)\in\mathbb{R}^n \ : \ \frac{(x_1-\mu)+\ldots+(x_n-\mu)}{n}\geq\eta\} \tag{6.39}$$

Thus,

$$\begin{aligned}
&[G^n(E^{-1}(\mathrm{Ball}^C_{d^{(2)}_\Theta}(\pi(\omega);\eta))](\omega) \\
=&\frac{1}{(\sqrt{2\pi}\sigma)^n}\underset{\frac{(x_1-\mu)+\ldots+(x_n-\mu)}{n}\geq\eta}{\int\cdots\int}\exp[-\frac{\sum_{k=1}^n(x_k-\mu)^2}{2\sigma^2}]dx_1dx_2\cdots dx_n \\
=&\frac{1}{(\sqrt{2\pi}\sigma)^n}\underset{\frac{x_1+\ldots+x_n}{n}\geq\eta}{\int\cdots\int}\exp[-\frac{\sum_{k=1}^n(x_k)^2}{2\sigma^2}]dx_1dx_2\cdots dx_n \\
=&\frac{\sqrt{n}}{\sqrt{2\pi}\sigma}\int_{|x|\geq\eta}\exp[-\frac{nx^2}{2\sigma^2}]dx=\frac{1}{\sqrt{2\pi}}\int_{|x|\geq\sqrt{n}\eta/\sigma}\exp[-\frac{x^2}{2}]dx
\end{aligned} \tag{6.40}$$

Solving the following equation:

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-z(\alpha/2)}\exp[-\frac{x^2}{2}]dx=\frac{1}{\sqrt{2\pi}}\int_{z(\alpha/2)}^{\infty}\exp[-\frac{x^2}{2}]dx=\alpha \tag{6.41}$$

we define that

$$\eta^\alpha_\omega=\frac{\sigma}{\sqrt{n}}z(\alpha) \tag{6.42}$$

Then, we get $\widehat{R}^{\alpha,\Theta}_{H_N}$ (the (α)-rejection region of $H_N(=(-\infty,\mu_0]\subseteq\Theta(=\mathbb{R}))$) as follows:

$$\begin{aligned}
\widehat{R}^{\alpha,\Theta}_{(-\infty,\mu_0]}=&\bigcap_{\pi(\omega)=\mu\in(-\infty,\mu_0]}\{E(x)(\in\Theta=\mathbb{R}) : d^{(2)}_\Theta(E(x),\pi(\omega))\geq\eta^\alpha_\omega\} \\
=&\{E(x)(=\frac{x_1+\ldots+x_n}{n})\in\mathbb{R} \ : \ \frac{x_1+\ldots+x_n}{n}-\mu_0\geq\frac{\sigma}{\sqrt{n}}z(\alpha)\}
\end{aligned} \tag{6.43}$$

Thus, in a similar way of Remark 6.10, we see that $\widehat{R}^{\alpha}_{(-\infty,\mu_0]\times\mathbb{R}_+}$="the slash part in Figure 6.5", where

$$\widehat{R}^{\alpha}_{(-\infty,\mu_0]\times\mathbb{R}_+}=\{(E(x)(=\frac{x_1+\ldots+x_n}{n}),\sigma)\in\mathbb{R}\times\mathbb{R}_+ \ : \ \frac{x_1+\ldots+x_n}{n}-\mu_0\geq\frac{\sigma}{\sqrt{n}}z(\alpha)\} \tag{6.44}$$

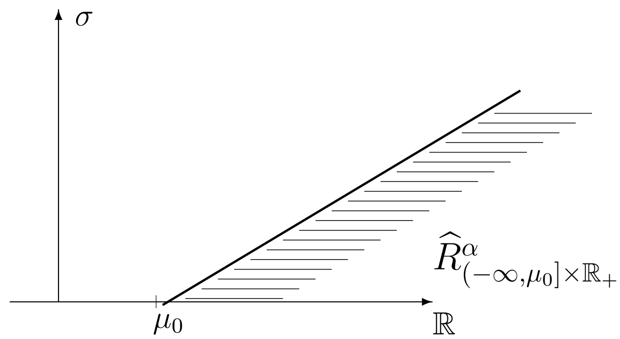


Figure 6.5: Rejection region $\widehat{R}^{\alpha,\Theta}_{(-\infty,\mu_0]}$ (which depends on σ)

6.4 Confidence interval and statistical hypothesis testing for population variance

6.4.1 Preparation (simultaneous normal measurement)

Consider the simultaneous normal measurement $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}$ $(\mathsf{O}_G^n = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n, G^n),\ S_{[(\mu,\sigma)]})$ in $L^\infty(\mathbb{R}\times\mathbb{R}_+)$. Here, recall that the simultaneous normal observable $\mathsf{O}_G^n = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n, G^n)$ is defined by

$$
\begin{aligned}
&[G^n(\times_{k=1}^n \Xi_k)](\omega) = \times_{k=1}^n [G(\Xi_k)](\omega) \\
=&\frac{1}{(\sqrt{2\pi}\sigma)^n} \underset{\times_{k=1}^n \Xi_k}{\int\cdots\int} \exp[-\frac{\sum_{k=1}^n (x_k-\mu)^2}{2\sigma^2}] dx_1 dx_2 \cdots dx_n \qquad (6.45)\\
&(\forall \Xi_k \in \mathcal{B}_{\mathbb{R}} (k=1,2,\ldots,n), \quad \forall \omega=(\mu,\sigma) \in \Omega = \mathbb{R}\times\mathbb{R}_+).
\end{aligned}
$$

where, note that

$$
\Omega = \mathbb{R}\times\mathbb{R}_+
$$
$$
X = \mathbb{R}^n
$$

The second state space Θ is

$$
\Theta = \mathbb{R}_+
$$

Putting

$$
\overline{\mu}(x) = \frac{x_1 + x_2 + \cdots + x_n}{n}
$$

we define the estimator $E : \mathbb{R}^n \to \Theta (\equiv \mathbb{R}_+)$ by

$$
E(x) = E(x_1, x_2, \ldots, x_n) = \sqrt{\frac{(x_1 - \overline{\mu}(x))^2 + (x_2 - \overline{\mu}(x))^2 + \cdots + (x_n - \overline{\mu}(x))^2}{n}}
$$

and the system quantity $\pi : \Omega \to \Theta$ by

$$
\Omega = \mathbb{R}\times\mathbb{R}_+ \ni \omega = (\mu,\sigma) \mapsto \pi(\omega) = \sigma \in \Theta = \mathbb{R}_+
$$

6.4.2 Confidence interval

Our present problem is as follows.

Problem 6.12. [Confidence interval for population variance]. Consider the simultaneous normal measurement $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}$ $(\mathsf{O}_G^n = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n, G^n), S_{[(\mu,\sigma)]})$. Assume that a measured value$x \in X = \mathbb{R}^n$ is obtained by the measurement. Let $0 < \alpha \ll 1$.
Then, find the $D_x^{1-\alpha;\Theta}(\subseteq \Theta)$ (which may depend on μ) such that

- the probability that $\sigma \in D_x^{1-\alpha;\Theta}$ is more than $1-\alpha$

Here, the more $D_x^{1-\alpha;\Theta}(\subseteq \Theta)$ is small, the more it is desirable.

Consider the following semi-distance $d_\Theta^{(1)}$ in $\Theta(= \mathbb{R}_+)$:

$$d_\Theta^{(1)}(\theta_1, \theta_2) = |\int_{\sigma_1}^{\sigma_2} \frac{1}{\sigma} d\sigma| = |\log \sigma_1 - \log \sigma_2| \tag{6.46}$$

For any $\omega = (\mu, \sigma)(\in \Omega = \mathbb{R} \times \mathbb{R}_+)$, define the positive number $\delta_\omega^{1-\alpha}$ (> 0) such that:

$$\begin{aligned}\delta_\omega^{1-\alpha} &= \inf\{\eta > 0 : [F(E^{-1}(\mathrm{Ball}_{d_\Theta^{(1)}}(\omega; \eta))](\omega) \geq 1 - \alpha\} \\ &= \inf\{\eta > 0 : [F(E^{-1}(\mathrm{Ball}^C_{d_\Theta^{(1)}}(\omega; \eta))](\omega) \leq \alpha\}\end{aligned} \tag{6.47}$$

where

$$\mathrm{Ball}^C_{d_\Theta^{(1)}}(\omega;\eta) = \mathrm{Ball}^C_{d_\Theta^{(1)}}((\mu;\sigma),\eta) = \mathbb{R} \times \{\sigma' \ : \ |\log(\sigma'/\sigma)| \geq \eta\} = \mathbb{R} \times \Big((0, \sigma e^{-\eta}] \cup [\sigma e^{\eta}, \infty)\Big) \tag{6.48}$$

Then,

$$\begin{aligned}&E^{-1}(\mathrm{Ball}^C_{d_\Theta^{(1)}}(\omega;\eta)) = E^{-1}\Big(\mathbb{R} \times \big((0, \sigma e^{-\eta}] \cup [\sigma e^{\eta}, \infty)\big)\Big) \\ =&\{(x_1, \ldots, x_n) \in \mathbb{R}^n \ : \ \Big(\frac{\sum_{k=1}^n (x_k - \overline{\mu}(x))^2}{n}\Big)^{1/2} \leq \sigma e^{-\eta} \text{ or } \sigma e^{\eta} \leq \Big(\frac{\sum_{k=1}^n (x_k - \overline{\mu}(x))^2}{n}\Big)^{1/2}\}\end{aligned} \tag{6.49}$$

Hence we see, by the Gauss integral (6.7), that

$$\begin{aligned}&[G^n(E^{-1}(\mathrm{Ball}^C_{d_\Theta^{(1)}}(\omega;\eta))](\omega) \\ =&\frac{1}{(\sqrt{2\pi}\sigma)^n} \underset{E^{-1}\Big(\mathbb{R}\times\big((0,\sigma e^{-\eta}]\cup[\sigma e^{\eta},\infty)\big)\Big)}{\int \cdots \int} \exp[-\frac{\sum_{k=1}^n (x_k - \mu)^2}{2\sigma^2}] dx_1 dx_2 \cdots dx_n \\ =&\int_0^{ne^{-2\eta}} p^{\chi^2}_{n-1}(x) dx + \int_{ne^{2\eta}}^{\infty} p^{\chi^2}_{n-1}(x) dx = 1 - \int_{ne^{-2\eta}}^{ne^{2\eta}} p^{\chi^2}_{n-1}(x) dx\end{aligned} \tag{6.50}$$

Using the chi-squared distribution $p^{\chi^2}_{n-1}(x)$ (with $n-1$ degrees of freedom) in (6.8), define the $\delta^{1-\alpha}_\omega$ such that

$$1-\alpha = \int_{ne^{-2\delta^{1-\alpha}_\omega}}^{ne^{2\delta^{1-\alpha}_\omega}} p^{\chi^2}_{n-1}(x)dx \tag{6.51}$$

where it should be noted that the $\delta^{1-\alpha}_\omega$ depends on only α and n. Thus, put

$$\delta^{1-\alpha}_\omega = \delta^{1-\alpha}_n \tag{6.52}$$

Hence we get, for any x ($\in X$), the $D_x^{1-\alpha,\Omega}$ (the $(1-\alpha)$-confidence interval of x) as follows:

$$\begin{aligned} D_x^{1-\alpha,\Omega} &= \{\omega(\in \Omega) : d^{(1)}_\Theta(E(x),\pi(\omega)) \le \delta^{1-\alpha}_n\} \\ &= \{(\mu,\sigma)\in \mathbb{R}\times\mathbb{R}_+ \ : \ \sigma e^{-\delta^{1-\alpha}_n} \le \Big(\frac{\sum_{k=1}^n (x_k-\overline{\mu}(x))^2}{n}\Big)^{1/2} \le \sigma e^{\delta^{1-\alpha}_n}\} \end{aligned} \tag{6.53}$$

Recalling (6.4), i.e., $\overline{\sigma}(x) = \Big(\frac{\sum_{k=1}^n (x_k-\overline{\mu}(x))^2}{n}\Big)^{1/2} = (\frac{\overline{SS}(x)}{n})^{1/2}$, we conclude that

$$\begin{aligned} D_x^{1-\alpha,\Omega} &= \{(\mu,\sigma)\in \mathbb{R}\times\mathbb{R}_+ \ : \ \overline{\sigma}(x)e^{-\delta^{1-\alpha}_n} \le \sigma \le \overline{\sigma}(x)e^{\delta^{1-\alpha}_n}\} \\ &= \{(\mu,\sigma)\in \mathbb{R}\times\mathbb{R}_+ \ : \ \frac{e^{-2\delta^{1-\alpha}_n}}{n}\overline{SS}(x) \le \sigma^2 \le \frac{e^{2\delta^{1-\alpha}_n}}{n}\overline{SS}(x)\} \end{aligned} \tag{6.54}$$

And

$$\begin{aligned} D_x^{1-\alpha,\Theta} &= \{\sigma\in\mathbb{R}_+ \ : \ \overline{\sigma}(x)e^{-\delta^{1-\alpha}_n} \le \sigma \le \overline{\sigma}(x)e^{\delta^{1-\alpha}_n}\} \\ &= \{(\mu,\sigma)\in \mathbb{R}\times\mathbb{R}_+ \ : \ \frac{e^{-2\delta^{1-\alpha}_n}}{n}\overline{SS}(x) \le \sigma^2 \le \frac{e^{2\delta^{1-\alpha}_n}}{n}\overline{SS}(x)\} \end{aligned}$$

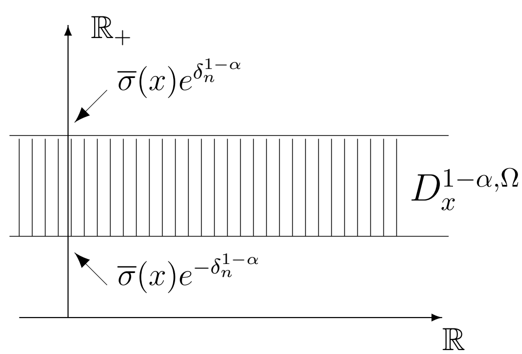


Figure 6.6: Confidence interval $D_x^{1-\alpha,\Omega}$ for the semi-distance $d^{(1)}_\Theta$

6.4.3 Statistical hypothesis testing[null hypothesis$H_N = \{\sigma_0\} \subseteq \Theta = \mathbb{R}_+$]

Our present problem is as follows.

Problem 6.13. [Statistical hypothesis testing]. Consider the simultaneous normal measurement $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}$ $(\mathsf{O}_G^n = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n, G^n), S_{[(\mu,\sigma)]})$. Assume the null hypothesis H_N such that

$$H_N = \{\sigma_0\}(\subseteq \Theta = \mathbb{R}))$$

Let $0 < \alpha \ll 1$.
Then, find the rejection region $\widehat{R}_{H_N}^{\alpha;\Theta}(\subseteq \Theta)$ (which may depend on μ) such that

- the probability that a measured value$x(\in \mathbb{R}^n)$ obtained by $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}$ $(\mathsf{O}_G^n = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n, G^n), S_{[(\mu_0,\sigma)]})$ satisfies that

$$E(x) \in \widehat{R}_{H_N}^{\alpha;\Theta}$$

  is less that α.

Here, the more the rejection region $\widehat{R}_{H_N}^{\alpha;\Theta}$ is large, the more it is desirable.

For any $\omega = (\mu, \sigma)(\in \Omega = \mathbb{R} \times \mathbb{R}_+)$, define the positive number η_ω^α (> 0) such that:

$$\eta_\omega^\alpha = \inf\{\eta > 0 : [F(E^{-1}(\mathrm{Ball}_{d_\Theta^{(1)}}^C(\omega;\eta))](\omega) \le \alpha\}$$

Recall that

$$\eta_\omega^\alpha = \delta_\omega^{1-\alpha} = \delta_n^{1-\alpha} (= \eta_n^\alpha)$$

Hence we get the $\widehat{R}_{H_N}^{\alpha,\Theta}$ (the (α)-rejection region of $H_N = \{\sigma_0\} \subseteq \Theta = \mathbb{R}_+$) as follows:

$$\begin{aligned}\widehat{R}_{H_N}^{\alpha,\Theta} = \widehat{R}_{\{\sigma_0\}}^{\alpha,\Theta} &= \bigcap_{\pi(\omega)=\sigma\in\{\sigma_0\}} \{E(x)(\in \Theta) : d_\Theta^{(1)}(E(x), \pi(\omega)) \ge \eta_\omega^\alpha\} \\ &= \{E(x)(\in \Theta = \mathbb{R}_+) : d_\Theta^{(1)}(E(x), \sigma_0) \ge \eta_n^\alpha\} \\ &= \{\overline{\sigma}(x)(\in \Theta = \mathbb{R}_+) \ : \ \overline{\sigma}(x) \le \sigma_0 e^{-\eta_n^\alpha} \text{ or } \sigma_0 e^{\eta_n^\alpha} \le \overline{\sigma}(x)\} \end{aligned} \tag{6.55}$$

where $\overline{\sigma}(x) = \left(\frac{\sum_{k=1}^n (x_k - \overline{\mu}(x))^2}{n}\right)^{1/2}$.

Thus, in a similar way of Remark 6.10, we see that $\widehat{R}_{\mathbb{R}\times\{\sigma_0\}}^{\alpha}$=“the slash part in Figure 6.7”, where

$$\widehat{R}_{\mathbb{R}\times\{\sigma_0\}}^{\alpha} = \{(\mu, \overline{\sigma}(x)) \in \mathbb{R} \times \mathbb{R}_+ \ : \ \overline{\sigma}(x) \le \sigma_0 e^{-\eta_n^\alpha} \text{ or } \sigma_0 e^{\eta_n^\alpha} \le \overline{\sigma}(x)\} \tag{6.56}$$

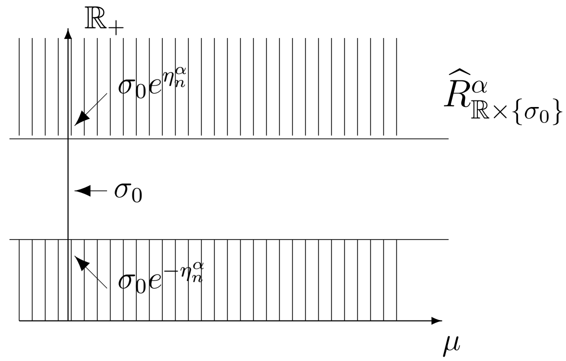


Figure 6.7: Rejection region $\widehat{R}^\alpha_{\mathbb{R}\times\{\sigma_0\}}$

6.4.4 Statistical hypothesis testing[null hypothesis$H_N = (0,\sigma_0] \subseteq \Theta = \mathbb{R}_+$]

Our present problem is as follows.

Problem 6.14. [Statistical hypothesis testing]. Consider the simultaneous normal measurement $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}$ $(\mathsf{O}_G^n = (\mathbb{R}^n, \mathcal{B}_\mathbb{R}^n, G^n),\ S_{[(\mu,\sigma)]})$. Assume the null hypothesis H_N such that

$$H_N = (0,\sigma_0](\subseteq \Theta = \mathbb{R}))$$

Let $0 < \alpha \ll 1$.
Then, find the rejection region $\widehat{R}^{\alpha;\Theta}_{H_N}(\subseteq \Theta)$ (which may depend on μ) such that

- the probability that a measured value$x(\in \mathbb{R}^n)$ obtained by $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}$ $(\mathsf{O}_G^n = (\mathbb{R}^n, \mathcal{B}_\mathbb{R}^n, G^n),\ S_{[(\mu_0,\sigma)]})$ satisfies that
$$E(x) \in \widehat{R}^{\alpha;\Theta}_{H_N}$$
is less that α.

Here, the more the rejection region $\widehat{R}^{\alpha;\Theta}_{H_N}$ is large, the more it is desirable.

Consider the following semi-distance $d_\Theta^{(2)}$ in $\Theta(=\mathbb{R}_+)$:

$$d_\Theta^{(2)}(\sigma_1,\sigma_2) = \begin{cases} |\int_{\sigma_1}^{\sigma_2} \frac{1}{\sigma} d\sigma| = |\log\sigma_1 - \log\sigma_2| & (\sigma_0 \le \sigma_1,\sigma_2) \\ |\int_{\sigma_0}^{\sigma_2} \frac{1}{\sigma} d\sigma| = |\log\sigma_0 - \log\sigma_2| & (\sigma_1 \le \sigma_0 \le \sigma_2) \\ |\int_{\sigma_0}^{\sigma_1} \frac{1}{\sigma} d\sigma| = |\log\sigma_0 - \log\sigma_1| & (\sigma_2 \le \sigma_0 \le \sigma_1) \\ 0 & (\sigma_1,\sigma_2 \le \sigma_0) \end{cases} \tag{6.57}$$

For any $\omega = (\mu,\sigma)(\in \Omega = \mathbb{R}\times\mathbb{R}_+)$, define the positive number η^α_ω (>0) such that:

$$\eta^\alpha_\omega = \inf\{\eta > 0 : [F(E^{-1}(\mathrm{Ball}^C_{d_\Theta^{(2)}}(\omega;\eta))](\omega) \le \alpha\} \tag{6.58}$$

where

$$\mathrm{Ball}^C_{d^{(2)}_\Theta}(\omega;\eta)=\mathrm{Ball}^C_{d^{(2)}_\Theta}((\mu;\sigma),\eta)=\mathbb{R}\times[\sigma e^{\eta},\infty) \tag{6.59}$$

Then,

$$\begin{aligned}&E^{-1}(\mathrm{Ball}^C_{d^{(2)}_\Theta}(\omega;\eta))=E^{-1}\Big([\sigma e^{\eta},\infty)\Big)\\ =&\{(x_1,\ldots,x_n)\in\mathbb{R}^n\ :\ \sigma e^{\eta}\le\overline{\sigma}(x)=\Big(\frac{\sum_{k=1}^n(x_k-\overline{\mu}(x))^2}{n}\Big)^{1/2}\}\end{aligned} \tag{6.60}$$

Hence we see, by the Gauss integral (6.7), that

$$\begin{aligned}&[G^n(E^{-1}(\mathrm{Ball}^C_{d^{(2)}_\Theta}(\omega;\eta))](\omega)\\ =&\frac{1}{(\sqrt{2\pi}\sigma)^n}\int\cdots\int_{\sigma_0 e^{\eta}\le\overline{\sigma}(x)}\exp[-\frac{\sum_{k=1}^n(x_k-\mu)^2}{2\sigma^2}]dx_1dx_2\cdots dx_n\\ =&\int_{\frac{ne^{2\eta}\sigma^2}{\sigma^2}}^{\infty}p^{\chi^2}_{n-1}(x)dx\\ \le&\int_{ne^{2\eta}}^{\infty}p^{\chi^2}_{n-1}(x)dx\end{aligned} \tag{6.61}$$

Solving the following equation, define the $(\eta^\alpha_n)'(>0)$ such that

$$\alpha=\int_{ne^{2(\eta^\alpha_n)'}}^{\infty}p^{\chi^2}_{n-1}(x)dx \tag{6.62}$$

Hence we get the $\widehat{R}^{\alpha,\Theta}_{H_N}$ (the (α)-rejection region of $H_N=(0,\sigma_0]$) as follows:

$$\begin{aligned}\widehat{R}^{\alpha,\Theta}_{H_N}=\widehat{R}^{\alpha,\Theta}_{(0,\sigma_0]}=&\bigcap_{\pi(\omega)\in(0,\sigma_0]}\{E(x)(\in\Theta=\mathbb{R}_+):d^{(2)}_\Theta(E(x),\pi(\omega))\ge\eta^\alpha_\omega\}\\ =&\bigcap_{\pi(\omega)\in(0,\sigma_0]}\{E(x)(\in\Theta):d^{(2)}_\Theta(E(x),\pi(\omega))\ge(\eta^\alpha_n)'\}\\ =&\{\sigma(=\overline{\sigma}(x))\in\mathbb{R}_+\ :\ \sigma_0e^{(\eta^\alpha_n)'}\le\overline{\sigma}(x)\}\end{aligned} \tag{6.63}$$

where $\overline{\sigma}(x)=\Big(\frac{\sum_{k=1}^n(x_k-\overline{\mu}(x))^2}{n}\Big)^{1/2}$.

Thus, in a similar way of Remark 6.10, we see that $\widehat{R}^{\alpha}_{\mathbb{R}\times(0,\sigma_0]}$="the slash part in Figure 6.8", where

$$\widehat{R}^{\alpha}_{\mathbb{R}\times(0,\sigma_0]}=\{(\mu,\overline{\sigma}(x))\in\mathbb{R}\times\mathbb{R}_+\ :\ \sigma_0e^{(\eta^\alpha_n)'}\le\overline{\sigma}(x)\} \tag{6.64}$$

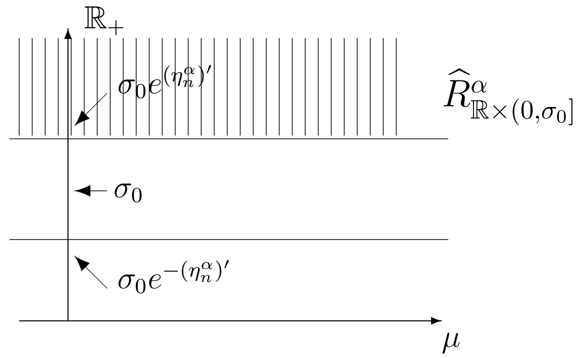


Figure 6.8: Rejection region $\widehat{R}^\alpha_{\mathbb{R}\times(0,\sigma_0]}$

6.5 Confidence interval and statistical hypothesis testing for the difference of population means

6.5.1 Preparation (simultaneous normal measurement)

Consider the parallel measurement $\mathsf{M}_{L^\infty((\mathbb{R}\times\mathbb{R}_+)\times(\mathbb{R}\times\mathbb{R}_+))}$ $(\mathsf{O}_G^n \otimes \mathsf{O}_G^m = (\mathbb{R}^n \times \mathbb{R}^m, \mathcal{B}_{\mathbb{R}}^n \boxtimes \mathcal{B}_{\mathbb{R}}^m, G^n \otimes G^m), S_{[(\mu_1,\sigma_1,\mu_2,\sigma_2)]})$ (in $L^\infty((\mathbb{R}\times\mathbb{R}_+)\times(\mathbb{R}\times\mathbb{R}_+)))$ of two normal measurements.

Assume that σ_1 and σ_2 are fixed and known. Thus, this parallel measurement is represented by $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R})}$ $(\mathsf{O}_{G_{\sigma_1}}^n \otimes \mathsf{O}_{G_{\sigma_1}}^m = (\mathbb{R}^n \times \mathbb{R}^m, \mathcal{B}_{\mathbb{R}}^n \boxtimes \mathcal{B}_{\mathbb{R}}^m, {G_{\sigma_1}}^n \otimes {G_{\sigma_2}}^m), S_{[(\mu_1,\mu_2)]})$ in $L^\infty(\mathbb{R}\times\mathbb{R})$. Here, recall the normal observable (6.1), i.e.,

$$[G_\sigma(\Xi)](\mu) = \frac{1}{\sqrt{2\pi}\sigma}\int_\Xi \exp[-\frac{(x-\mu)^2}{2\sigma^2}]dx \quad (\forall \Xi \in \mathcal{B}_{\mathbb{R}}(=\text{Borel field in } \mathbb{R})), \quad \forall \mu \in \mathbb{R}). \tag{6.65}$$

Therefore, we have the state space $\Omega = \mathbb{R}^2 = \{\omega = (\mu_1, \mu_2) \ : \ \mu_1, \mu_2 \in \mathbb{R}\}$. Put $\Theta = \mathbb{R}$ with the distance $d_\Theta^{(1)}(\theta_1, \theta_2) = |\theta_1 - \theta_2|$ and consider the quantity $\pi : \mathbb{R}^2 \to \mathbb{R}$ by

$$\pi(\mu_1, \mu_2) = \mu_1 - \mu_2 \tag{6.66}$$

The estimator $E : \widehat{X}(= X \times Y = \mathbb{R}^n \times \mathbb{R}^m) \to \Theta(= \mathbb{R})$ is defined by

$$E(x_1, \ldots, x_n, y_1, \ldots, y_m) = \frac{\sum_{k=1}^n x_k}{n} - \frac{\sum_{k=1}^m y_k}{m} \tag{6.67}$$

For any $\omega = (\mu_1, \mu_2)(\in \Omega = \mathbb{R}\times\mathbb{R})$, define the positive number $\eta_\omega^\alpha(= \delta_\omega^{1-\alpha})$ (> 0) such that:

$$\eta_\omega^\alpha(= \delta_\omega^{1-\alpha}) = \inf\{\eta > 0 : [F(E^{-1}(\mathrm{Ball}_{d_\Theta^{(1)}}^C(\pi(\omega); \eta))](\omega) \geq \alpha\}$$

where $\mathrm{Ball}_{d_\Theta^{(1)}}^C(\pi(\omega); \eta) = (-\infty, \mu_1 - \mu_2 - \eta] \cup [\mu_1 - \mu_2 + \eta, \infty)$. Define the null hypothesis H_N ($\subseteq \Theta = \mathbb{R}$) such that

$$H_N = \{\theta_0\}$$

Now let us calculate the η_ω^α as follows:

$$\begin{aligned}
&E^{-1}(\mathrm{Ball}_{d_\Theta^{(1)}}^C(\pi(\omega); \eta)) = E^{-1}((-\infty, \mu_1 - \mu_2 - \eta] \cup [\mu_1 - \mu_2 + \eta, \infty)) \\
=&\{(x_1, \ldots, x_n, y_1, \ldots, y_m) \in \mathbb{R}^n \times \mathbb{R}^m \ : \ |\frac{\sum_{k=1}^n x_k}{n} - \frac{\sum_{k=1}^m y_k}{m} - (\mu_1 - \mu_2)| \geq \eta\} \\
=&\{(x_1, \ldots, x_n, y_1, \ldots, y_m) \in \mathbb{R}^n \times \mathbb{R}^m \ : \ |\frac{\sum_{k=1}^n (x_k - \mu_1)}{n} - \frac{\sum_{k=1}^m (y_k - \mu_2)}{m}| \geq \eta\}
\end{aligned} \tag{6.68}$$

Thus,

$$
\begin{aligned}
&[(N_{\sigma_1}{}^n \otimes N_{\sigma_2}{}^m)(E^{-1}(\mathrm{Ball}^C_{d^{(1)}_\Theta}(\pi(\omega);\eta))](\omega) \\
=&\frac{1}{(\sqrt{2\pi}\sigma_1)^n(\sqrt{2\pi}\sigma_2)^m} \\
&\times \underset{|\frac{\sum_{k=1}^n (x_k-\mu_1)}{n} - \frac{\sum_{k=1}^m (y_k-\mu_2)}{m}|\geq \eta}{\int\cdots\int} \exp[-\frac{\sum_{k=1}^n (x_k-\mu_1)^2}{2\sigma_1^2} - \frac{\sum_{k=1}^m (y_k-\mu_2)^2}{2\sigma_2^2}]dx_1 dx_2\cdots dx_n dy_1 dy_2 \cdots dy_m \\
=&\frac{1}{(\sqrt{2\pi}\sigma_1)^n(\sqrt{2\pi}\sigma_2)^m} \underset{|\frac{\sum_{k=1}^n x_k}{n} - \frac{\sum_{k=1}^m y_k}{m}|\geq \eta}{\int\cdots\int} \exp[-\frac{\sum_{k=1}^n {x_k}^2}{2\sigma_1^2} - \frac{\sum_{k=1}^m {y_k}^2}{2\sigma_2^2}]dx_1 dx_2\cdots dx_n dy_1 dy_2 \cdots dy_m \\
=&1 - \frac{1}{\sqrt{2\pi}(\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m})^{1/2}} \int_{-\eta}^{\eta} \exp[-\frac{x^2}{2(\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m})}]dx
\end{aligned}
\tag{6.69}
$$

Using the $z(\alpha/2)$ in (6.33), we get that

$$
\eta^\alpha_\omega = \delta^{1-\alpha}_\omega = (\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m})^{1/2} z(\frac{\alpha}{2}) \tag{6.70}
$$

6.5.2 Confidence interval

Our present problem is as follows

Problem 6.15. [Confidence interval for the difference of population means]. Let σ_1 and σ_2 be positive numbers which are assumed to be fixed. Consider the parallel measurement $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R})}$ $(\mathsf{O}^n_{G_{\sigma_1}} \otimes \mathsf{O}^m_{G_{\sigma_1}} = (\mathbb{R}^n \times \mathbb{R}^m, \mathcal{B}^n_{\mathbb{R}} \boxtimes \mathcal{B}^m_{\mathbb{R}}, G_{\sigma_1}{}^n \otimes G_{\sigma_2}{}^m), S_{[(\mu_1,\mu_2)]})$. Assume that a measured value $\widehat{x} = (x,y) = (x_1,\ldots,x_n,y_1,\ldots,y_m)$ ($\in \mathbb{R}^n \times \mathbb{R}^m$) is obtained by the measurement. Let $0 < \alpha \ll 1$.
Then, find the confidence interval $D^{1-\alpha;\Theta}_{(x,y)} (\subseteq \Theta)$ (which may depend on σ_1 and σ_2) such that

- the probability that $\mu_1 - \mu_2 \in D^{1-\alpha;\Theta}_{(x,y)}$ is more than $1-\alpha$.

Here, the more the confidence interval $D^{1-\alpha;\Theta}_{(x,y)}$ is small, the more it is desirable.

Therefore, for any $\widehat{x} = (x,y) = (x_1,\ldots,x_n,y_1,\ldots,y_m)$ ($\in \mathbb{R}^n \times \mathbb{R}^m$), we get $D^{1-\alpha}_{\widehat{x}}$ (the $(1-\alpha)$-confidence interval of $\widehat{x}$) as follows:

$$
\begin{aligned}
D^{1-\alpha,\Omega}_{\widehat{x}} &= \{\omega(\in \Omega) : d_\Theta(E(\widehat{x}), \pi(\omega)) \leq \delta^{1-\alpha}_\omega\} \\
&= \{(\mu_1,\mu_2) \in \mathbb{R}\times\mathbb{R} \ : \ |\frac{\sum_{k=1}^n x_k}{n} - \frac{\sum_{k=1}^m y_k}{m} - (\mu_1-\mu_2)| \leq (\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m})^{1/2} z(\frac{\alpha}{2})\}
\end{aligned}
\tag{6.71}
$$

6.5.3 Statistical hypothesis testing[rejection region: null hypothesis$H_N = \{\mu_0\} \subseteq \Theta = \mathbb{R}$]

Our present problem is as follows

Problem 6.16. [Statistical hypothesis testing for the difference of population means]. Consider the parallel measurement $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R})}$ $(\mathsf{O}^n_{G_{\sigma_1}} \otimes \mathsf{O}^m_{G_{\sigma_1}} = (\mathbb{R}^n \times \mathbb{R}^m\ , \mathcal{B}^n_{\mathbb{R}} \boxtimes \mathcal{B}^m_{\mathbb{R}}, {G_{\sigma_1}}^n \otimes {G_{\sigma_2}}^m)$, $S_{[(\mu_1,\mu_2)]})$. Assume that

$$\pi(\mu_1, \mu_2) = \mu_1 - \mu_2 = \theta_0 \in \Theta = \mathbb{R}$$

that is, assume the null hypothesisH_N such that

$$H_N = \{\theta_0\}(\subseteq \Theta = \mathbb{R}))$$

Let $0 < \alpha \ll 1$.
Then, find the rejection region $\widehat{R}^{\alpha;\Theta}_{H_N}(\subseteq \Theta)$ (which may depend on μ) such that

- the probability that a measured value$(x, y)(\in \mathbb{R}^n \times \mathbb{R}^m)$ obtained by $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R})}$ $(\mathsf{O}^n_{G_{\sigma_1}} \otimes \mathsf{O}^m_{G_{\sigma_1}} = (\mathbb{R}^n \times \mathbb{R}^m\ , \mathcal{B}^n_{\mathbb{R}} \boxtimes \mathcal{B}^m_{\mathbb{R}}, {G_{\sigma_1}}^n \otimes {G_{\sigma_2}}^m)$, $S_{[(\mu_1,\mu_2)]})$ satisfies

$$E(x, y) = \frac{x_1 + x_2 + \cdots + x_n}{n} - \frac{y_1 + y_2 + \cdots + y_m}{m} \in \widehat{R}^{\alpha;\Theta}_{H_N}$$

  is less than α.

Here, the more the rejection region $\widehat{R}^{\alpha;\Theta}_{H_N}$ is large, the more it is desirable.

By the formula (6.70), we see that the rejection region$\widehat{R}^{\alpha}_{\widehat{x}}$ ((α)-rejection region of $H_N = \{\theta_0\}(\subseteq \Theta)$) is defined by

$$\widehat{R}^{\alpha,\Theta}_{H_N} = \bigcap_{\omega=(\mu_1,\mu_2)\in\Omega(=\mathbb{R}^2) \text{ such that } \pi(\omega)=\mu_1-\mu_2\in H_N(=\{\theta_0\})} \{E(\widehat{x})(\in \Theta) : d^{(1)}_{\Theta}(E(\widehat{x}), \pi(\omega)) \geq \eta^\alpha_\omega\}$$
$$= \{\overline{\mu}(x) - \overline{\mu}(y) \in \Theta(= \mathbb{R})\ :\ |\overline{\mu}(x) - \overline{\mu}(y) - \theta_0| \geq (\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m})^{1/2} z(\frac{\alpha}{2})\} \tag{6.72}$$

or,

$$\widehat{R}^{\alpha,X}_{H_N} = \bigcap_{\omega=(\mu_1,\mu_2)\in\Omega(=\mathbb{R}^2) \text{ such that } \pi(\omega)=\mu_1-\mu_2\in H_N(=\{\theta_0\})} \{\widehat{x}(\in \mathbb{R}^n \times \mathbb{R}^m) : d^{(1)}_{\Theta}(E(\widehat{x}), \pi(\omega)) \geq \eta^\alpha_\omega\}$$
$$= \{\widehat{x}(\in \mathbb{R}^n \times \mathbb{R}^m)\ :\ |\overline{\mu}(x) - \overline{\mu}(y) - \theta_0| \geq (\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m})^{1/2} z(\frac{\alpha}{2})\} \tag{6.73}$$

Here,

$$\overline{\mu}(x) = \frac{\sum_{k=1}^n x_k}{n}, \quad \overline{\mu}(y) = \frac{\sum_{k=1}^m y_k}{m}$$

6.5.4 Statistical hypothesis testing[rejection region: null hypothesis$H_N = (-\infty, \theta_0] \subseteq \Theta = \mathbb{R}$]

Our present problem is as follows

Problem 6.17. [Statistical hypothesis testing for the difference of population means]. Consider the parallel measurement $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R})}$ $(\mathsf{O}^n_{G_{\sigma_1}} \otimes \mathsf{O}^m_{G_{\sigma_1}} = (\mathbb{R}^n \times \mathbb{R}^m, \mathcal{B}^n_{\mathbb{R}} \boxtimes \mathcal{B}^m_{\mathbb{R}}, {G_{\sigma_1}}^n \otimes {G_{\sigma_2}}^m)$, $S_{[(\mu_1,\mu_2)]})$. Assume that

$$\pi(\mu_1, \mu_2) = \mu_1 - \mu_2 = (-\infty, \theta_0] \subseteq \Theta = \mathbb{R}$$

that is, assume the null hypothesisH_N such that

$$H_N = (-\infty, \theta_0] (\subseteq \Theta = \mathbb{R}))$$

Let $0 < \alpha \ll 1$.
Then, find the rejection region $\widehat{R}^{\alpha;\Theta}_{H_N} (\subseteq \Theta)$ (which may depend on μ) such that

- the probability that a measured value$(x, y) (\in \mathbb{R}^n \times \mathbb{R}^m)$ obtained by $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R})}$ $(\mathsf{O}^n_{G_{\sigma_1}} \otimes \mathsf{O}^m_{G_{\sigma_1}} = (\mathbb{R}^n \times \mathbb{R}^m, \mathcal{B}^n_{\mathbb{R}} \boxtimes \mathcal{B}^m_{\mathbb{R}}, {G_{\sigma_1}}^n \otimes {G_{\sigma_2}}^m)$, $S_{[(\mu_1,\mu_2)]})$ satisfies

$$E(x, y) = \frac{x_1 + x_2 + \cdots + x_n}{n} - \frac{y_1 + y_2 + \cdots + y_m}{m} \in \widehat{R}^{\alpha;\Theta}_{H_N}$$

  is less than α.

Here, the more the rejection region $\widehat{R}^{\alpha;\Theta}_{H_N}$ is large, the more it is desirable.

Since the null hypothesis H_N is assumed as follows:

$$H_N = (-\infty, \theta_0],$$

it suffices to define the semi-distance $d^{(1)}_\Theta$ in $\Theta (= \mathbb{R})$ such that

$$d^{(1)}_\Theta(\theta_1, \theta_2) = \begin{cases} |\theta_1 - \theta_2| & (\forall \theta_1, \theta_2 \in \Theta = \mathbb{R} \text{ such that } \theta_0 \le \theta_1, \theta_2) \\ \max\{\theta_1, \theta_2\} - \theta_0 & (\forall \theta_1, \theta_2 \in \Theta = \mathbb{R} \text{ such that } \min\{\theta_1, \theta_2\} \le \theta_0 \le \max\{\theta_1, \theta_2\}) \\ 0 & (\forall \theta_1, \theta_2 \in \Theta = \mathbb{R} \text{ such that } \theta_1, \theta_2 \le \theta_0) \end{cases} \tag{6.74}$$

Then, we can easily see that

$$\begin{aligned} \widehat{R}^{\alpha,\Theta}_{H_N} &= \bigcap_{\omega = (\mu_1,\mu_2) \in \Omega (=\mathbb{R}^2) \text{ such that } \pi(\omega) = \mu_1 - \mu_2 \in H_N (=(-\infty, \theta_0])} \{E(\widehat{x}) (\in \Theta) : d^{(1)}_\Theta(E(\widehat{x}), \pi(\omega)) \ge \eta^\alpha_\omega\} \\ &= \{\overline{\mu}(x) - \overline{\mu}(y) \in \mathbb{R} \ : \ \overline{\mu}(x) - \overline{\mu}(y) - \theta_0 \ge (\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m})^{1/2} z(\alpha)\} \end{aligned} \tag{6.75}$$

6.6 Student t-distribution of population mean

6.6.1 Preparation

Example 6.18. [Student t-distribution]. Consider the simultaneous measurement $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}(\mathsf{O}_G^n = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n, G^n),\ S_{[(\mu,\sigma)]})$ in $L^\infty(\mathbb{R}\times\mathbb{R}_+)$. Thus, we consider that $\Omega = \mathbb{R}\times\mathbb{R}_+$, $X = \mathbb{R}^n$. Put $\Theta = \mathbb{R}$ with the semi-distance $d_\Theta^x (\forall x \in X)$ such that

$$d_\Theta^x(\theta_1, \theta_2) = \frac{|\theta_1 - \theta_2|}{\overline{\sigma}'(x)/\sqrt{n}} \qquad (\forall x \in X = \mathbb{R}^n, \forall \theta_1, \theta_2 \in \Theta = \mathbb{R}) \tag{6.76}$$

where $\overline{\sigma}'(x) = \sqrt{\frac{n}{n-1}}\overline{\sigma}(x)$. The quantity $\pi : \Omega(= \mathbb{R}\times\mathbb{R}_+) \to \Theta(= \mathbb{R})$ is defined by

$$\Omega(= \mathbb{R}\times\mathbb{R}_+) \ni \omega = (\mu, \sigma) \mapsto \pi(\mu, \sigma) = \mu \in \Theta(= \mathbb{R}) \tag{6.77}$$

Also, define the estimator $E : X(= \mathbb{R}^n) \to \Theta(= \mathbb{R})$ such that

$$E(x) = E(x_1, x_2, \ldots, x_n) = \overline{\mu}(x) = \frac{x_1 + x_2 + \cdots + x_n}{n} \tag{6.78}$$

Define the null hypothesis H_N $(\subseteq \Theta = \mathbb{R}))$ such that

$$H_N = \{\mu_0\} \tag{6.79}$$

Thus, for any $\omega = (\mu_0, \sigma)(\in \Omega = \mathbb{R}\times\mathbb{R}_+)$, we see that

$$\begin{aligned}
&[G^n(\{x \in X(= \mathbb{R}^n) \ :\ d_\Theta^x(E(x), \pi(\omega)) \ge \eta\})](\omega)\\
=&[G^n(\{x \in X \ :\ \frac{|\overline{\mu}(x) - \mu_0|}{\overline{\sigma}'(x)/\sqrt{n}} \ge \eta\})](\omega)\\
=&\frac{1}{(\sqrt{2\pi}\sigma)^n} \underset{\eta \le \frac{|\overline{\mu}(x)-\mu_0|}{\overline{\sigma}'(x)/\sqrt{n}}}{\int\cdots\int} \exp[-\frac{\sum_{k=1}^n (x_k - \mu_0)^2}{2\sigma^2}] dx_1 dx_2 \cdots dx_n\\
=&\frac{1}{(\sqrt{2\pi})^n} \underset{\eta \le \frac{|\overline{\mu}(x)|}{\overline{\sigma}'(x)/\sqrt{n}}}{\int\cdots\int} \exp[-\frac{\sum_{k=1}^n (x_k)^2}{2}] dx_1 dx_2 \cdots dx_n\\
=&1 - \int_{-\eta}^{\eta} p_{n-1}^t(x) dx
\end{aligned} \tag{6.80}$$

where p_{n-1}^t is the t-distribution with $n-1$ degrees of freedom. Solving the equation $1 - \alpha = \int_{-\eta_\omega^\alpha}^{\eta_\omega^\alpha} p_{n-1}^t(x) dx$, we get

$$\delta_\omega^{1-\alpha} = \eta_\omega^\alpha = t(\alpha/2)$$

6.6.2 Confidence interval

Our present problem is as follows

Problem 6.19. [Confidence interval]. Consider the simultaneous normal measurement $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}$ $(\mathsf{O}_G^n = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n, G^n),\ S_{[(\mu,\sigma)]})$. Assume that a measured value$x \in X = \mathbb{R}^n$ is obtained by the measurement. Let $0 < \alpha \ll 1$.
Then, find the confidence interval $D_x^{1-\alpha;\Theta}(\subseteq \Theta)$ (which does not depend on σ) such that

- the probability that $\mu \in D_x^{1-\alpha;\Theta}$ is more than $1-\alpha$

Here, the more the confidence interval $D_x^{1-\alpha;\Theta}$ is small, the more it is desirable.

Therefore, for any x ($\in X$), we get $D_x^{1-\alpha,\Theta}$(*the* $(1-\alpha)$*-confidence interval of* x) as follows:

$$
\begin{aligned}
D_x^{1-\alpha} &= \{\pi(\omega)(\in \Theta) : \omega \in \Omega,\ \ d_\Theta^x(E(x), \pi(\omega)) \le \delta_\omega^{1-\alpha}\} \\
&= \{\mu \in \Theta(= \mathbb{R})\ :\ \overline{\mu}(x) - \frac{\overline{\sigma}'(x)}{\sqrt{n}} t(\alpha/2) \le \mu \le \overline{\mu}(x) + \frac{\overline{\sigma}'(x)}{\sqrt{n}} t(\alpha/2)\}
\end{aligned}
\tag{6.81}
$$

$$
\begin{aligned}
D_x^{1-\alpha,\Omega} &= \{\omega = (\mu,\sigma)(\in \Omega) : \omega \in \Omega,\ \ d_\Theta^x(E(x), \pi(\omega)) \le \delta_\omega^{1-\alpha}\} \\
&= \{\omega = (\mu,\sigma)(\in \Omega)\ :\ \overline{\mu}(x) - \frac{\overline{\sigma}'(x)}{\sqrt{n}} t(\alpha/2) \le \mu \le \overline{\mu}(x) + \frac{\overline{\sigma}'(x)}{\sqrt{n}} t(\alpha/2)\}
\end{aligned}
\tag{6.82}
$$

6.6.3 Statistical hypothesis testing[null hypothesis$H_N = \{\mu_0\}(\subseteq \Theta = \mathbb{R})$]

Our present problem was as follows

Problem 6.20. [Statistical hypothesis testing]. Consider the simultaneous normal measurement $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}$ $(\mathsf{O}_G^n = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n, G^n),\ S_{[(\mu,\sigma)]})$. Assume that

$$\mu = \mu_0$$

That is, assume the null hypothesis H_N such that

$$H_N = \{\mu_0\}(\subseteq \Theta = \mathbb{R}))$$

Let $0 < \alpha \ll 1$.
Then, find the rejection region $\widehat{R}_{H_N}^{\alpha;\Theta}(\subseteq \Theta)$ (which does not depend on σ) such that

- the probability that a measured value$x(\in \mathbb{R}^n)$ obtained by $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}$ $(\mathsf{O}_G^n = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n, G^n),\ S_{[(\mu_0,\sigma)]})$ satisfies

$$E(x) \in \widehat{R}_{H_N}^{\alpha;\Theta}$$

is less than α.

Here, the more the rejection region $\widehat{R}_{H_N}^{\alpha;\Theta}$ is large, the more it is desirable.

The rejection region$\widehat{R}_{H_N}^{\alpha,\Theta}$(($\alpha$)-rejection region of null hypothesis $H_N(=\{\mu_0\})$) is calculated as follows:

$$\widehat{R}_{H_N}^{\alpha,\Theta} = \bigcap_{\omega=(\mu,\sigma)\in\Omega(=\mathbb{R}\times\mathbb{R}_+) \text{ such that } \pi(\omega)=\mu\in H_N(=\{\mu_0\})} \{E(x)(\in\Theta): \ d_\Theta^x(E(x),\pi(\omega)) \geq \eta_\omega^\alpha\}$$
$$= \{\overline{\mu}(x)\in\Theta(=\mathbb{R}) \ : \ \frac{|\overline{\mu}(x)-\mu_0|}{\overline{\sigma}'(x)/\sqrt{n}} \geq t(\alpha/2)\}$$
$$= \{\overline{\mu}(x)\in\Theta(=\mathbb{R}) \ : \ \mu_0 \leq \overline{\mu}(x) - \frac{\overline{\sigma}'(x)}{\sqrt{n}}t(\alpha/2) \text{ or } \overline{\mu}(x) + \frac{\overline{\sigma}'(x)}{\sqrt{n}}t(\alpha/2) \leq \mu_0\} \tag{6.83}$$

Also,

$$\widehat{R}_{H_N}^{\alpha,X} = \bigcap_{\omega=(\mu,\sigma)\in\Omega(=\mathbb{R}\times\mathbb{R}_+) \text{ such that } \pi(\omega)=\mu\in H_N(=\{\mu_0\})} \{x\in X: \ d_\Theta^x(E(x),\pi(\omega)) \geq \eta_\omega^\alpha\}$$
$$= \{x\in X=\mathbb{R}^n \ : \ \frac{|\overline{\mu}(x)-\mu_0|}{\overline{\sigma}'(x)/\sqrt{n}} \geq t(\alpha/2)\}$$
$$= \{x\in X=\mathbb{R}^n \ : \ \mu_0 \leq \overline{\mu}(x) - \frac{\overline{\sigma}'(x)}{\sqrt{n}}t(\alpha/2) \text{ or } \overline{\mu}(x) + \frac{\overline{\sigma}'(x)}{\sqrt{n}}t(\alpha/2) \leq \mu_0\} \tag{6.84}$$

6.6.4 Statistical hypothesis testing[null hypothesis $H_N = (-\infty,\mu_0](\subseteq \Theta=\mathbb{R}\ )$]

Our present problem was as follows

Problem 6.21. [Statistical hypothesis testing]. Consider the simultaneous normal measurement $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}$ $(\mathsf{O}_G^n=(\mathbb{R}^n,\mathcal{B}_\mathbb{R}^n,G^n),\ S_{[(\mu,\sigma)]})$. Assume that

$$\mu\in(-\infty,\mu_0]$$

That is, assume the null hypothesis H_N such that

$$H_N=(-\infty,\mu_0](\subseteq\Theta=\mathbb{R}))$$

Let $0<\alpha\ll 1$.
Then, find the rejection region $\widehat{R}_{H_N}^{\alpha;\Theta}(\subseteq\Theta)$ (which does not depend on σ) such that

- the probability that a measured value$x(\in\ \mathbb{R}^n)$ obtained by $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}$ $(\mathsf{O}_G^n = (\mathbb{R}^n,\mathcal{B}_\mathbb{R}^n,G^n),\ S_{[(\mu_0,\sigma)]})$ satisfies

$$E(x)\in\widehat{R}_{H_N}^{\alpha;\Theta}$$

is less than α.

Here, the more the rejection region $\widehat{R}_{H_N}^{\alpha;\Theta}$ is large, the more it is desirable.

Since the null hypothesis H_N is assumed as follows:

$$H_N = (-\infty, \mu_0],$$

it suffices to define the semi-distance d_Θ^x in $\Theta(=\mathbb{R})$ such that

$$d_\Theta^x(\theta_1,\theta_2) = \begin{cases} \frac{|\theta_1-\theta_2|}{\overline{\sigma}'(x)/\sqrt{n}} & (\forall \theta_1,\theta_2 \in \Theta = \mathbb{R} \text{ such that } \mu_0 \le \theta_1,\theta_2) \\ \frac{\max\{\theta_1,\theta_2\}-\mu_0}{\overline{\sigma}'(x)/\sqrt{n}} & (\forall \theta_1,\theta_2 \in \Theta = \mathbb{R} \text{ such that } \min\{\theta_1,\theta_2\} \le \mu_0 \le \max\{\theta_1,\theta_2\}) \\ 0 & (\forall \theta_1,\theta_2 \in \Theta = \mathbb{R} \text{ such that } \theta_1,\theta_2 \le \mu_0) \end{cases} \tag{6.85}$$

for any $x \in X = \mathbb{R}^n$.

Then, (α)-rejection region$\widehat{R}_{H_N}^{\alpha,\Theta}$ is calculated as follows.

$$\begin{aligned} \widehat{R}_{H_N}^{\alpha,\Theta} = & \bigcap_{\omega=(\mu,\sigma)\in\Omega(=\mathbb{R}\times\mathbb{R}_+) \text{ such that } \pi(\omega)=\mu\in H_N(=(-\infty,\mu_0])} \{E(x)(\in\Theta) : \ d_\Theta^x(E(x),\pi(\omega)) \ge \eta_\omega^\alpha\} \\ = & \{\overline{\mu}(x) \in \Theta(=\mathbb{R}) \ : \ \mu_0 \le \overline{\mu}(x) - \frac{\overline{\sigma}'(x)}{\sqrt{n}} t(\alpha)\} \end{aligned} \tag{6.86}$$

Also,

$$\begin{aligned} \widehat{R}_{H_N}^{\alpha,X} = & \bigcap_{\omega=(\mu,\sigma)\in\Omega(=\mathbb{R}\times\mathbb{R}_+) \text{ such that } \pi(\omega)=\mu\in H_N(=(-\infty,\mu_0])} \{x(\in X=\mathbb{R}^n) \ : \ d_\Theta^x(E(x),\pi(\omega)) \ge \eta_\omega^\alpha\} \\ = & \{x(\in X = \mathbb{R}^n) \ : \ \mu_0 \le \overline{\mu}(x) - \frac{\overline{\sigma}'(x)}{\sqrt{n}} t(\alpha)\} \end{aligned} \tag{6.87}$$

Remark 6.22. There are many ideas of statistical hypothesis testing. The most natural idea is the likelihood-ratio, which is discussed in

(a) Ref. [29]: S. Ishikawa, "Mathematical Foundations of Measurement Theory," Keio University Press Inc. 2006.

(b) Ref. [32]: S. Ishikawa, "A Measurement Theoretical Foundation of Statistics," Applied Mathematics, Vol. 3, No. 3, 2012, pp. 283-292. doi: 10.4236/am.2012.33044

Also, we think that the arguments concerning "null hypothesis vs. alternative hypothesis" and "one-sided test and two-sided test" are practical and not theoretical.

Chapter 7

ANOVA(= Analysis of Variance)

The standard university course of statistics is as follows:

$$\underset{\text{(likelihood method, moment method)}}{\overset{①}{\boxed{\text{Inference}}}} \longrightarrow \overset{②}{\boxed{\text{confidence interval}}} \longrightarrow \overset{③}{\boxed{\text{statistical hypothesis testing}}}$$

$$\longrightarrow \overset{④}{\boxed{\text{ANOVA}}}$$

In the previous chapters, we studied ①,② and ③. In this chapter, we devote ourselves to ④(ANOVA). This chapter is extracted from the following.

Ref. [41]: S. Ishikawa, ANOVA (analysis of variance) in the quantum linguistic formulation of statistics (arXiv:1402.0606 [math.ST] 2014)

7.1 Zero way ANOVA (Student t-distribution)

In the previous chapter, we introduced the statistical hypothesis testing for student t-distribution, which is characterized as "zero" way ANOVA (analysis of variance). In this section, we review "zero" way ANOVA (analysis of variance).

Consider the classical basic structure

$$[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))]$$

where

$$\Omega = \mathbb{R} \times \mathbb{R}_+ = \{(\mu, \sigma) \mid \mu \text{ is real}, \sigma \text{ is positive real}\}$$

Consider the simultaneous normal measurement $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}$ $(\mathsf{O}_G^n = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n, G^n),\ S_{[(\mu,\sigma)]})$ (in $L^\infty(\mathbb{R} \times \mathbb{R}_+)$). For completeness, recall that

$$
\begin{aligned}
&[G^n(\times_{k=1}^n \Xi_k)](\omega) = \times_{k=1}^n [G(\Xi_k)](\omega) \\
=&\frac{1}{(\sqrt{2\pi}\sigma)^n} \int \cdots \int_{\times_{k=1}^n \Xi_k} \exp[-\frac{\sum_{k=1}^n (x_k-\mu)^2}{2\sigma^2}] dx_1 dx_2 \cdots dx_n \qquad (7.1)\\
&(\forall \Xi_k \in \mathcal{B}_{\mathbb{R}}(k=1,2,\ldots,n), \quad \forall \omega = (\mu,\sigma) \in \Omega = \mathbb{R} \times \mathbb{R}_+).
\end{aligned}
$$

And recall the state space $\Omega = \mathbb{R} \times \mathbb{R}_+$, the measured value space $X = \mathbb{R}^n$, the second state space(=parameter space) $\Theta = \mathbb{R}$. Also, recall the estimator $E : X(= \mathbb{R}^n) \to \Theta(= \mathbb{R})$ defined by

$$
E(x) = E(x_1, x_2, \ldots, x_n) = \overline{\mu}(x) = \frac{x_1 + x_2 + \cdots + x_n}{n} \tag{7.2}
$$

and the system quantity $\pi : \Omega(= \mathbb{R} \times \mathbb{R}_+) \to \Theta(= \mathbb{R})$ defined by

$$
\Omega(= \mathbb{R} \times \mathbb{R}_+) \ni \omega = (\mu, \sigma) \mapsto \pi(\mu, \sigma) = \mu \in \Theta(= \mathbb{R}) \tag{7.3}
$$

The essence of "studentized" is to define the semi-metric $d_\Theta^x (\forall x \in X)$ in the second state space $\Theta(= \mathbb{R})$such that

$$
d_\Theta^x(\theta^{(1)}, \theta^{(2)}) = \frac{|\theta^{(1)} - \theta^{(2)}|}{\sqrt{n}\overline{\sigma}(x)} = \frac{|\theta^{(1)} - \theta^{(2)}|}{\sqrt{\overline{SS}(x)}} \qquad (\forall x \in X = \mathbb{R}^n, \forall \theta^{(1)}, \theta^{(2)} \in \Theta = \mathbb{R}) \tag{7.4}
$$

where

$$
\overline{SS}(x) = \overline{SS}(x_1, x_2, \ldots, x_n) = \sum_{k=1}^n (x_k - \overline{\mu}(x))^2 \quad (\forall x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n)
$$

Thus, as mentioned in the previous chapter, our problem is characterized as follows.

Problem 7.1. [The zero-way ANOVA]. Consider the simultaneous normal measurement $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}$ $(\mathsf{O}_G^n = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n, G^n),\ S_{[(\mu,\sigma)]})$ Here, assume that

$$
\mu = \mu_0
$$

That is, the null hypothesis H_N is defined by $H_N = \{\mu_0\}\ (\subseteq \Theta = \mathbb{R})$. Consider $0 < \alpha \ll 1$. **Then, find** the largest $\widehat{R}_{H_N}^{\alpha;\Theta} (\subseteq \Theta)$ (independent of σ) such that

(A$_1$) the probability that a measured value $x (\in \mathbb{R}^n)$ (obtained by $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}(\mathsf{O}_G^n = (X(\equiv \mathbb{R}^n), \mathcal{B}_{\mathbb{R}}^n, G^n), S_{[(\mu_0,\sigma)]}))$ satisfies

$$
E(x) \in \widehat{R}_{H_N}^{\alpha;\Theta} \tag{7.5}
$$

is less than α.

We see, for any $\omega = (\mu_0, \sigma)(\in \Omega = \mathbb{R} \times \mathbb{R}_+)$,

$$
\begin{aligned}
&[G^n(\{x \in X \ : \ d^x_\Theta(E(x), \pi(\omega)) \geq \eta\})](\omega) \\
=&[G^n(\{x \in X \ : \ \frac{|\overline{\mu}(x) - \mu_0|}{\sqrt{\overline{SS}(x)}} \geq \eta\})](\omega) \\
=&\frac{1}{(\sqrt{2\pi}\sigma)^n} \underset{\eta\sqrt{n-1} \leq \frac{|\overline{\mu}(x)-\mu_0|}{\sqrt{\overline{SS}(x)}/\sqrt{n-1}}}{\int \cdots \int} \exp[-\frac{\sum_{k=1}^n (x_k - \mu_0)^2}{2\sigma^2}] dx_1 dx_2 \cdots dx_n \\
=&\frac{1}{(\sqrt{2\pi})^n} \underset{\eta^2 n(n-1) \leq \frac{n(\overline{\mu}(x))^2}{\overline{SS}(x)/(n-1)}}{\int \cdots \int} \exp[-\frac{\sum_{k=1}^n (x_k)^2}{2}] dx_1 dx_2 \cdots dx_n
\end{aligned}
\tag{7.6}
$$

(A_2) by the formula of Gauss integrals (Formula 7.8(A)(§7.4)), we see

$$
= \int_{\eta^2 n(n-1)}^{\infty} p^F_{(1,n-1)}(t) dt = \alpha \ \ (\text{ e.g., } \alpha = 0.05) \tag{7.7}
$$

where $p^F_{(1,n-1)}$ is the probability density function of F-distribution with $(1, n-1)$ degree of freedom.

Note that the probability density function $p^F_{(n_1,n_2)}(t)$ of F-distribution with (n_1, n_2) degree of freedom is defined by

$$
p^F_{(n_1,n_2)}(t) = \frac{1}{B(n_1/2, n_2/2)} \Big(\frac{n_1}{n_2}\Big)^{n_1/2} \frac{t^{(n_1-2)/2}}{(1 + n_1 t/n_2)^{(n_1+n_2)/2}} \qquad (t \geq 0) \tag{7.8}
$$

where $B(\cdot, \cdot)$ is the Beta function.

The α-point: $F^{n_2}_{n_1,\alpha}$ (> 0) is defined by

$$
\int_{F^{n_2}_{n_1,\alpha}}^{\infty} p^F_{(n_1,n_2)}(t) dt = \alpha \qquad (0 < \alpha \ll 1. \text{ e.g., } \alpha = 0.05) \tag{7.9}
$$

Thus, it suffices to solve the following equation:

$$
\eta^2 n(n-1) = F^1_{n-1,\alpha} \tag{7.10}
$$

Therefore,

$$
(\eta^\alpha_\omega)^2 = \frac{F^1_{n-1,\alpha}}{n(n-1)} \tag{7.11}
$$

Then, the rejection region$\widehat{R}^{\alpha;\Theta}_{H_N}$((or $\widehat{R}^{\alpha;X}_{H_N}$) is calculated as

$$
\widehat{R}^{\alpha;\Theta}_{H_N} = \bigcap_{\omega=(\mu,\sigma)\in\Omega(=\mathbb{R}\times\mathbb{R}_+) \text{ such that } \pi(\omega)=\mu\in H_N(=\{\mu_0\})} \{E(x)(\in \Theta) : \ d^x_\Theta(E(x), \pi(\omega)) \geq \eta^\alpha_\omega\}
$$

$$= \{\overline{\mu}(x) \in \Theta(=\mathbb{R}) \ : \ \frac{|\overline{\mu}(x)-\mu_0|}{\sqrt{\overline{SS}(x)}} \geq \eta^\alpha_\omega\} = \{\overline{\mu}(x) \in \Theta(=\mathbb{R}) \ : \ \frac{|\overline{\mu}(x)-\mu_0|}{\overline{\sigma}(x)} \geq \eta^\alpha_\omega \sqrt{n}\}$$

$$= \Big\{\overline{\mu}(x) \in \Theta(=\mathbb{R}) \ : \ \frac{|\overline{\mu}(x)-\mu_0|}{\overline{\sigma}(x)} \geq \sqrt{\frac{F^1_{n-1,\alpha}}{n-1}} \ \Big\}$$

$$= \Big\{\overline{\mu}(x) \in \Theta(=\mathbb{R}) \ : \ \mu_0 \leq \overline{\mu}(x) - \overline{\sigma}(x)\sqrt{\frac{F^1_{n-1,\alpha}}{n-1}} \text{ or } \overline{\mu}(x) + \overline{\sigma}(x)\sqrt{\frac{F^1_{n-1,\alpha}}{n-1}} \leq \mu_0 \Big\} \quad (7.12)$$

and,

$$\widehat{R}^{\alpha;X}_{H_N} = E^{-1}(\widehat{R}^{\alpha;\Theta}_{H_N})$$

$$= \Big\{x \in X(=\mathbb{R}^n) \ : \ \mu_0 \leq \overline{\mu}(x) - \overline{\sigma}(x)\sqrt{\frac{F^1_{n-1,\alpha}}{n-1}} \text{ or } \overline{\mu}(x) + \overline{\sigma}(x)\sqrt{\frac{F^1_{n-1,\alpha}}{n-1}} \leq \mu_0 \Big\} \quad (7.13)$$

♠**Note 7.1.** (i): It should be noted that the mathematical part is only the (A_2).

(ii): Also, note that

($\sharp$) F-distribution with $(1, n-1)$ degree of freedom
= the student t-distribution with $(n-1)$ degree of freedom

Thus, we conclude that

$$(7.12) = (6.83) \qquad (7.13) = (6.84)$$

7.2 The one way ANOVA

For each $i = 1, 2, \cdots, a$, a natural number n_i is determined. And put, $n = \sum_{i=1}^{a} n_i$. Consider the parallel simultaneous normal observable $\mathsf{O}_G^n = (X(\equiv \mathbb{R}^n), \mathcal{B}_{\mathbb{R}}^n, G^n)$ (in $L^\infty(\Omega(\equiv (\mathbb{R}^a \times \mathbb{R}_+))$) such that

$$[G^n(\widehat{\Xi})](\omega) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \underbrace{\int \cdots \int}_{\widehat{\Xi}} \exp[-\frac{\sum_{i=1}^{a}\sum_{k=1}^{n_i}(x_{ik}-\mu_i)^2}{2\sigma^2}] \mathop{\times}_{i=1}^{a} \mathop{\times}_{k=1}^{n_i} dx_{ik} \tag{7.14}$$

$$(\forall \omega = (\mu_1, \mu_2, \ldots, \mu_a, \sigma) \in \Omega = \mathbb{R}^a \times \mathbb{R}_+, \widehat{\Xi} \in \mathcal{B}_{\mathbb{R}}^n)$$

That is, consider

$$\mathsf{M}_{L^\infty(\mathbb{R}^a \times \mathbb{R}_+)}(\mathsf{O}_G^n = (X(\equiv \mathbb{R}^n), \mathcal{B}_{\mathbb{R}}^n, G^n), S_{[(\mu=(\mu_1,\mu_2,\cdots,\mu_a),\sigma)]})$$

Put a_i as follows.

$$\alpha_i = \mu_i - \frac{\sum_{i=1}^{a} \mu_i}{a} \qquad (\forall i = 1, 2, \ldots, a) \tag{7.15}$$

and put,

$$\Theta = \mathbb{R}^a$$

Thus,, the system quantity $\pi : \Omega \to \Theta$ is defined as follows.

$$\Omega = \mathbb{R}^a \times \mathbb{R}_+ \ni \omega = (\mu_1, \mu_2, \ldots, \mu_a, \sigma) \mapsto \pi(\omega) = (\alpha_1, \alpha_2, \ldots, \alpha_a) \in \Theta = \mathbb{R}^a \tag{7.16}$$

Define the null hypothesis $H_N (\subseteq \Theta = \mathbb{R}^a)$ as follows.

$$\begin{aligned} H_N &= \{(\alpha_1, \alpha_2, \ldots, \alpha_a) \in \Theta = \mathbb{R}^a \ : \ \alpha_1 = \alpha_2 = \ldots = \alpha_a = \alpha\} \\ &= \{(\overbrace{0, 0, \ldots, 0}^{a})\} \end{aligned} \tag{7.17}$$

Here, note the following equivalence:

$$\text{“}\mu_1 = \mu_2 = \ldots = \mu_a\text{”} \Leftrightarrow \text{“}\alpha_1 = \alpha_2 = \ldots = \alpha_a = 0\text{”} \Leftrightarrow \text{“(7.17)”}$$

Hence, our problem is as follows.

Problem 7.2. [The one-way ANOVA]. Put $n = \sum_{i=1}^{a} n_i$. Consider the parallel simultaneous normal measurement $\mathsf{M}_{L^\infty(\mathbb{R}^a \times \mathbb{R}_+)}(\mathsf{O}_G^n = (X(\equiv \mathbb{R}^n), \mathcal{B}_{\mathbb{R}}^n, G^n), S_{[(\mu=(\mu_1,\mu_2,\cdots,\mu_a),\sigma)]})$ Here, assume that

$$\mu_1 = \mu_2 = \cdots = \mu_a$$

that is,

$$\pi(\mu_1, \mu_2, \cdots, \mu_a) = (0, 0, \cdots, 0)$$

Namely, assume that the null hypothesis is $H_N = \{(0, 0, \cdots, 0)\}$ $(\subseteq \Theta = \mathbb{R}))$. Consider $0 < \alpha \ll 1$.
Then, find the largest $\widehat{R}_{H_N}^{\alpha;\Theta}(\subseteq \Theta)$ (independent of σ) such that

(A$_1$) the probability that a measured value $x(\in \mathbb{R}^n)$ (obtained by $\mathsf{M}_{L^\infty(\mathbb{R}^a \times \mathbb{R}_+)}(\mathsf{O}_G^n = (X(\equiv \mathbb{R}^n), \mathcal{B}_{\mathbb{R}}^n, G^n), S_{[(\mu=(\mu_1,\mu_2,\cdots,\mu_a),\sigma)]}))$ satisfies

$$E(x) \in \widehat{R}_{H_N}^{\alpha;\Theta}$$

is less than α.

Consider the weighted Euclidean norm $\|\theta^{(1)} - \theta^{(2)}\|_\Theta$ in $\Theta = \mathbb{R}^a$ as follows.

$$\|\theta^{(1)} - \theta^{(2)}\|_\Theta = \sqrt{\sum_{i=1}^{a} n_i \Big(\theta_i^{(1)} - \theta_i^{(2)}\Big)^2}$$
$$(\forall \theta^{(\ell)} = (\theta_1^{(\ell)}, \theta_2^{(\ell)}, \ldots, \theta_a^{(\ell)}) \in \mathbb{R}^a, \ \ell = 1, 2)$$

Also, put

$$X = \mathbb{R}^n \ni x = ((x_{ik})_{k=1,2,\ldots,n_i})_{i=1,2,\ldots,a}$$
$$x_{i\bullet} = \frac{\sum_{k=1}^{n_i} x_{ik}}{n_i}, \qquad x_{\bullet\bullet} = \frac{\sum_{i=1}^{a} \sum_{k=1}^{n_i} x_{ik}}{n_i}, \tag{7.18}$$

Theorem 5.6 (Fisher's maximum likelihood method) urges us to calculate $\overline{\sigma}(x)(= \sqrt{\frac{\overline{SS}(x)}{n}})$ as follows.
For $x \in X = \mathbb{R}^n$,

$$\begin{aligned}
&\overline{SS}(x) = \overline{SS}(((x_{ik})_{\ k=1,2,\ldots,n_i})_{i=1,2,\ldots,a}\ ) \\
=&\sum_{i=1}^{a} \sum_{k=1}^{n_i} (x_{ik} - x_{i\bullet})^2 \\
=&\sum_{i=1}^{a} \sum_{k=1}^{n_i} (x_{ik} - \frac{\sum_{k=1}^{n_i} x_{ik}}{n_i})^2 \\
=&\sum_{i=1}^{a} \sum_{k=1}^{n_i} ((x_{ik} - \mu_i) - \frac{\sum_{k=1}^{n_i} (x_{ik} - \mu_i)}{n_i})^2
\end{aligned}$$

$$=\overline{SS}(((x_{ik}-\mu_i)_{\;k=1,2,\ldots,n_i})_{i=1,2,\ldots,a}\;) \tag{7.19}$$

For each $x \in X = \mathbb{R}^n$, define the semi-norm d_Θ^x in Θ such that

$$d_\Theta^x(\theta^{(1)},\theta^{(2)}) = \frac{\|\theta^{(1)}-\theta^{(2)}\|_\Theta}{\sqrt{\overline{SS}(x)}} \qquad (\forall \theta^{(1)},\theta^{(2)} \in \Theta)). \tag{7.20}$$

Further, define the estimator $E : X(=\mathbb{R}^n) \to \Theta(=\mathbb{R}^a)$ as follows.

$$\begin{aligned}
E(x) =&E((x_{ik})_{i=1,2,\ldots,a,k=1,2,\ldots,n})\\
=&\Big(\frac{\sum_{k=1}^{n_i} x_{1k}}{n} - \frac{\sum_{i=1}^{a}\sum_{k=1}^{n_i} x_{ik}}{n}, \frac{\sum_{k=1}^{n_i} x_{2k}}{n} - \frac{\sum_{i=1}^{a}\sum_{k=1}^{n_i} x_{ik}}{n}, \ldots, \frac{\sum_{k=1}^{n_i} x_{ak}}{n} - \frac{\sum_{i=1}^{a}\sum_{k=1}^{n_i} x_{ik}}{n}\Big)\\
=&\Big(\frac{\sum_{k=1}^{n_i} x_{ik}}{n} - \frac{\sum_{i=1}^{a}\sum_{k=1}^{n_i} x_{ik}}{n}\Big)_{i=1,2,\ldots,a} = (x_{i\bullet} - x_{\bullet\bullet})_{i=1,2,\ldots,a}
\end{aligned} \tag{7.21}$$

Thus, we get

$$\begin{aligned}
&\|E(x)-\pi(\omega)\|_\Theta^2\\
=&||\Big(\frac{\sum_{k=1}^{n_i} x_{ik}}{n} - \frac{\sum_{i=1}^{a}\sum_{k=1}^{n_i} x_{ik}}{n}\Big)_{i=1,2,\ldots,a} - (\alpha_i)_{i=1,2,\ldots,a}||_\Theta^2\\
=&||\Big(\frac{\sum_{k=1}^{n_i} x_{ik}}{n} - \frac{\sum_{i=1}^{a}\sum_{k=1}^{n_i} x_{ik}}{n} - (\mu_i - \frac{\sum_{i=1}^{a}\mu_i}{a})\Big)_{i=1,2,\ldots,a}||_\Theta^2
\end{aligned}$$

remarking the null hypothesis H_N (i.e., $\mu_i - \frac{\sum_{k=1}^{a}\mu_i}{a} = \alpha_i = 0 (i=1,2,\ldots,a)$),

$$=||\Big(\frac{\sum_{k=1}^{n_i} x_{ik}}{n} - \frac{\sum_{i=1}^{a}\sum_{k=1}^{n_i} x_{ik}}{n}\Big)_{i=1,2,\ldots,a}||_\Theta^2 = \sum_{i=1}^{a} n_i(x_{i\bullet} - x_{\bullet\bullet})^2 \tag{7.22}$$

Therefore, for any $\omega = ((\mu_{ik})_{i=12,\ldots,a,\;k=1,2,\ldots,n},\sigma)(\in \Omega = \mathbb{R}^n \times \mathbb{R}_+)$, define the positive real η^α_ω (> 0) such that

$$\eta^\alpha_\omega = \inf\{\eta > 0 : [G^n(E^{-1}(\mathrm{Ball}^C_{d^x_\Theta}(\pi(\omega);\eta))](\omega) \ge \alpha\} \tag{7.23}$$

where

$$\mathrm{Ball}^C_{d^x_\Theta}(\pi(\omega);\eta) = \{\theta \in \Theta \;:\; d^x_\Theta(\pi(\omega),\theta) > \eta\} \tag{7.24}$$

Recalling the null hypothesis H_N (i.e., $\mu_i - \frac{\sum_{k=1}^{a}\mu_i}{a} = \alpha_i = 0 (i=1,2,\ldots,a)$) , calculate η^α_ω as follows.

$$\begin{aligned}
&E^{-1}(\mathrm{Ball}^C_{d^x_\Theta}(\pi(\omega);\eta)) = \{x \in X = \mathbb{R}^n \;:\; d^x_\Theta(E(x),\pi(\omega)) > \eta\}\\
=&\{x \in X = \mathbb{R}^n \;:\; \frac{\|E(x)-\pi(\omega)\|_\Theta^2}{\overline{SS}(x)} = \frac{\sum_{i=1}^{a} n_i(x_{i\bullet} - x_{\bullet\bullet})^2}{\sum_{i=1}^{a}\sum_{k=1}^{n_i}(x_{ik} - x_{i\bullet})^2} > \eta^2\}
\end{aligned} \tag{7.25}$$

For any $\omega = (\mu_1, \mu_2, \ldots, \mu_a, \sigma) \in \Omega = \mathbb{R}^a \times \mathbb{R}_+$ such that $\pi(\omega)(= (\alpha_1, \alpha_2, \ldots, \alpha_a)) \in H_N(= \{0, 0, \ldots, 0)\})$, we see

$$
\begin{aligned}
&[G^n(E^{-1}(\mathrm{Ball}^C_{d^x_\Theta}(\pi(\omega);\eta)))(\omega)\\
=&\frac{1}{(\sqrt{2\pi}\sigma)^n} \underset{\frac{\sum_{i=1}^a n_i(x_{i\bullet}-x_{\bullet\bullet})^2}{\sum_{i=1}^a\sum_{k=1}^{n_i}(x_{ik}-x_{i\bullet})^2}>\eta^2}{\int\cdots\int} \exp[-\frac{\sum_{i=1}^a\sum_{k=1}^{n_i}(x_{ik}-\mu_i)^2}{2\sigma^2}] \mathop{\times}_{i=1}^{a}\mathop{\times}_{k=1}^{n_i} dx_{ik}\\
=&\frac{1}{(\sqrt{2\pi})^n} \underset{\frac{(\sum_{i=1}^a n_i(x_{i\bullet}-x_{\bullet\bullet})^2/(a-1)}{(\sum_{i=1}^a\sum_{k=1}^{n_i}(x_{ik}-x_{i\bullet})^2)/(n-a)}>\frac{\eta^2(n-a)}{(a-1)}}{\int\cdots\int} \exp[-\frac{\sum_{i=1}^a\sum_{k=1}^{n_i}(x_{ik})^2}{2}] \mathop{\times}_{i=1}^{a}\mathop{\times}_{k=1}^{n_i} dx_{ik}
\end{aligned}
$$

(A$_2$) By the formula of Gauss integrals (Formula 7.8(B)(§7.4)), we see

$$
= \int_{\frac{\eta^2(n-a)}{(a-1)}}^{\infty} p^F_{(a-1,n-a)}(t)dt = \alpha \ \ (\text{ e.g., } \alpha{=}0.05) \tag{7.26}
$$

where, $p^F_{(a-1,n-a)}$ is a probability density function of the F-distribution with $p^F_{(a-1,n-a)}$ degree of freedom.

Therefore, it suffices to solve the following equation

$$
\frac{\eta^2(n-a)}{(a-1)} = F^{a-1}_{n-a,\alpha}(= \text{“}\alpha\text{-point”}) \tag{7.27}
$$

This is solved,

$$
(\eta^\alpha_\omega)^2 = F^{a-1}_{n-a,\alpha}(a-1)/(n-a) \tag{7.28}
$$

Then, we get $\widehat{R}^{\alpha;\Theta}_{\widehat{x}}$ (or, $\widehat{R}^{\alpha;X}_{\widehat{x}}$; the (α)-rejection region of $H_N = \{(0.0.\ldots,0)\}(\subseteq \Theta = \mathbb{R}^a)$) as follows:

$$
\begin{aligned}
\widehat{R}^{\alpha;\Theta}_{H_N} =& \bigcap_{\omega=((\mu_i)_{i=1}^a,\sigma)\in\Omega(=\mathbb{R}^a\times\mathbb{R}_+) \text{ such that } \pi(\omega)=(\mu)_{i=1}^a\in H_N=\{(0,0,\ldots,0)\}} \{E(x)(\in\Theta) : d^x_\Theta(E(x),\pi(\omega)) \geq \eta^\alpha_\omega\}\\
=& \{E(x)(\in\Theta) : \frac{(\sum_{i=1}^a n_i(x_{i\bullet}-x_{\bullet\bullet})^2)/(a-1)}{(\sum_{i=1}^a\sum_{k=1}^{a_i}(x_{ik}-x_{i\bullet})^2))/(n-a)} \geq F^{a-1}_{n-a,\alpha}\}
\end{aligned} \tag{7.29}
$$

Thus,

$$
\widehat{R}^{\alpha;X}_{\widehat{x}} = E^{-1}(\widehat{R}^{\alpha;\Theta}_{H_N}) = \{x \in X \ : \ \frac{(\sum_{i=1}^a n_i(x_{i\bullet}-x_{\bullet\bullet})^2)/(a-1)}{(\sum_{i=1}^a\sum_{k=1}^{n_i}(x_{ik}-x_{i\bullet})^2)/(n-a)} \geq F^{a-1}_{n-a,\alpha}\} \tag{7.30}
$$

♠**Note 7.2.** It should be noted that the mathematical part is only the (A$_2$).

7.3 The two way ANOVA

7.3.1 Preparation

As one of generalizations of the simultaneous normal observable (7.14), we consider a kind of observable $\mathsf{O}_G^{abn} = (X(\equiv \mathbb{R}^{abn}), \mathcal{B}_{\mathbb{R}}^{abn}, G^{abn})$ in $L^\infty(\Omega(\equiv (\mathbb{R}^{ab} \times \mathbb{R}_+))$.

$$
\begin{aligned}
&[G^{abn}(\widehat{\Xi})](\omega) \\
=&\frac{1}{(\sqrt{2\pi}\sigma)^{abn}} \int \cdots \int_{\widehat{\Xi}} \exp[-\frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (x_{ijk} - \mu_{ij})^2}{2\sigma^2}] \mathop{\times}_{k=1}^{n} \mathop{\times}_{j=1}^{b} \mathop{\times}_{i=1}^{a} dx_{ijk} \\
&\qquad (\forall \omega = ((\mu_{ij})_{i=1,2,\ldots,a, j=1,2,\ldots,b}, \sigma) \in \Omega = \mathbb{R}^{ab} \times \mathbb{R}_+, \widehat{\Xi} \in \mathcal{B}_{\mathbb{R}}^{abn})
\end{aligned} \tag{7.31}
$$

Therefore, consider the parallel simultaneous normal measurement:

$$
\mathsf{M}_{L^\infty(\mathbb{R}^{ab} \times \mathbb{R}_+)}(\mathsf{O}_G^{abn} = (X(\equiv \mathbb{R}^{abn}), \mathcal{B}_{\mathbb{R}}^{abn}, G^{abn}), S_{[(\mu=(\mu_{ij} \mid i=1,2,\cdots,a, j=1,2,\cdots,b),\sigma)]})
$$

Here,

$$
\begin{aligned}
\mu_{ij} = \overline{\mu}(= \mu_{\bullet\bullet} = &\frac{\sum_{i=1}^a \sum_{j=1}^b \mu_{ij}}{ab}) \\
&+ \alpha_i (= \mu_{i\bullet} - \mu_{\bullet\bullet} = \frac{\sum_{j=1}^b \mu_{ij}}{b} - \frac{\sum_{i=1}^a \sum_{j=1}^b \mu_{ij}}{ab}) \\
&+ \beta_j (= \mu_{\bullet j} - \mu_{\bullet\bullet} = \frac{\sum_{i=1}^a \mu_{ij}}{a} - \frac{\sum_{i=1}^a \sum_{j=1}^b \mu_{ij}}{ab}) \\
&+ (\alpha\beta)_{ij} (= \mu_{ij} - \mu_{i\bullet} - \mu_{\bullet j} + \mu_{\bullet\bullet})
\end{aligned} \tag{7.32}
$$

And put,

$$
\begin{aligned}
&X = \mathbb{R}^{abn} \ni x = (x_{ijk})_{i=1,2,\ldots,a,\ j=1,2,\ldots,b,\ k=1,2,\ldots,n} \\
&x_{ij\bullet} = \frac{\sum_{k=1}^n x_{ijk}}{n}, \quad x_{i\bullet\bullet} = \frac{\sum_{j=1}^b \sum_{k=1}^n x_{ijk}}{bn}, \quad x_{\bullet j\bullet} = \frac{\sum_{i=1}^a \sum_{k=1}^n x_{ijk}}{an}, \\
&x_{\bullet\bullet\bullet} = \frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n x_{ijk}}{abn}
\end{aligned} \tag{7.33}
$$

7.3.2 The null hypothesis: $\mu_{1\bullet} = \mu_{2\bullet} = \cdots = \mu_{a\bullet} = \mu_{\bullet\bullet}$

Now put,

$$
\Theta = \mathbb{R}^a \tag{7.34}
$$

define the system quantity $\pi_1 : \Omega(=\mathbb{R}^{ab} \times \mathbb{R}_+) \to \Theta(=\mathbb{R}^a)$ by

$$\Omega = \mathbb{R}^{ab} \times \mathbb{R}_+ \ni \omega = ((\mu_{ij})_{i=1,2,\ldots,a,j=1,2,\ldots,b}, \sigma) \mapsto \pi_1(\omega) = (\alpha_i)_{i=1}^a (= (\mu_{i\bullet} - \mu_{\bullet\bullet})_{i=1}^a) \in \Theta = \mathbb{R}^a \tag{7.35}$$

Define the null hypothesis $H_N (\subseteq \Theta = \mathbb{R}^a)$ such that

$$H_N = \{(\alpha_1, \alpha_2, \ldots, \alpha_a) \in \Theta = \mathbb{R}^a \ : \ \alpha_1 = \alpha_2 = \ldots = \alpha_a = \alpha\} \tag{7.36}$$

$$= \{(\overbrace{0, 0, \ldots, 0}^{a})\} \tag{7.37}$$

Here, "(7.36)⇔(7.37)" is derived from

$$a\alpha = \sum_{i=1}^a \alpha_i = \sum_{i=1}^a (\mu_{i\bullet} - \mu_{\bullet\bullet}) = \frac{\sum_{i=1}^a \sum_{j=1}^b \mu_{ij}}{b} - \sum_{i=1}^a \frac{\sum_{i=1}^a \sum_{j=1}^b \mu_{ij}}{ab} = 0 \tag{7.38}$$

Also, define the estimator $E : X(=\mathbb{R}^{abn}) \to \Theta(=\mathbb{R}^a)$ by

$$E(x) = \Big(\frac{\sum_{j=1}^b \sum_{k=1}^n x_{ijk}}{bn} - \frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n x_{ijk}}{abn} \Big)_{i=1,2,\ldots,a} = \Big(x_{i\bullet\bullet} - x_{\bullet\bullet\bullet} \Big)_{i=1,2,\ldots,a} \tag{7.39}$$

Now we have the following problem:

Problem 7.3. [The two-way ANOVA]. Consider the parallel simultaneous normal measurement:

$$\mathsf{M}_{L^\infty(\mathbb{R}^{ab} \times \mathbb{R}_+)} (\mathsf{O}_G^{abn} = (X(\equiv \mathbb{R}^{abn}), \mathcal{B}_{\mathbb{R}}^{abn}, G^{abn}), S_{[(\mu=(\mu_{ij} \mid i=1,2,\cdots,a,j=1,2,\cdots,b),\sigma)]})$$

where we assume that

$$\mu_{1\bullet} = \mu_{2\bullet} = \cdots = \mu_{a\bullet} = \mu_{\bullet\bullet}$$

that is,

$$\pi_1(\mu_1, \mu_2, \cdots, \mu_a) = (0, 0, \cdots, 0)$$

namely, consider the null hypothesis $H_N = \{(0, 0, \cdots, 0)\} \ (\subseteq \Theta = \mathbb{R}^a)$. Let $0 < \alpha \ll 1$.
Then, find the largest $\widehat{R}_{H_N}^{\alpha;\Theta} (\subseteq \Theta)$(independent of σ) such that

(A$_1$) the probability that a measured value $x(\in \mathbb{R}^{abn})$ obtained by $\mathsf{M}_{L^\infty(\mathbb{R}^{ab} \times \mathbb{R}_+)} (\mathsf{O}_G^{abn} = (X(\equiv \mathbb{R}^{abn}), \mathcal{B}_{\mathbb{R}}^{abn}, G^{abn}), S_{[(\mu=(\mu_{ij} \mid i=1,2,\cdots,a,j=1,2,\cdots,b),\sigma)]})$ satisfies that

$$E(x) \in \widehat{R}_{H_N}^{\alpha;\Theta}$$

is less than α.

Further,

$$\|\theta^{(1)}-\theta^{(2)}\|_\Theta=\sqrt{\sum_{i=1}^a\Big(\theta_i^{(1)}-\theta_i^{(2)}\Big)^2}$$
$$(\forall\theta^{(\ell)}=(\theta_1^{(i)},\theta_2^{(\ell)},\ldots,\theta_a^{(\ell)})\in\mathbb{R}^a,\ \ell=1,2)$$

Motivated by Theorem 5.6 (Fisher's maximum likelihood method), define and calculate $\overline{\sigma}(x)\Big(=\sqrt{\overline{SS}(x)/(abn)}\Big)$ as follows.

$$\begin{aligned}
&\overline{SS}(x)=\overline{SS}((x_{ijk})_{i=1,2,\ldots,a,\ j=1,2,\ldots,b,k=1,2,\ldots,n})\\
:=&\sum_{i=1}^a\sum_{j=1}^b\sum_{k=1}^n(x_{ijk}-x_{ij\bullet})^2=\sum_{i=1}^a\sum_{j=1}^b\sum_{k=1}^n(x_{ijk}-\frac{\sum_{k=1}^n x_{ijk}}{n})^2\\
=&\sum_{i=1}^a\sum_{j=1}^b\sum_{k=1}^n((x_{ijk}-\mu_{ij})-\frac{\sum_{k=1}^n(x_{ijk}-\mu_{ij})}{n})^2\\
=&\overline{SS}(((x_{ijk}-\mu_{ij})_{i=1,2,\ldots,a,\ j=1,2,\ldots,b})_{k=1,2,\cdots,n})
\end{aligned}\tag{7.40}$$

Define the semi-distance d_Θ^x (in $\Theta=\mathbb{R}^a$) such that

$$d_\Theta^x(\theta^{(1)},\theta^{(2)})=\frac{\|\theta^{(1)}-\theta^{(2)}\|_\Theta}{\sqrt{\overline{SS}(x)}}\qquad(\forall\theta^{(1)},\theta^{(2)}\in\Theta=\mathbb{R}^a,\forall x\in X=\mathbb{R}^{abn})\tag{7.41}$$

Define the estimator $E:X(=\mathbb{R}^{abn})\to\Theta(=\mathbb{R}^a)$ such that

$$E(x)=\Big(\frac{\sum_{j=1}^b\sum_{k=1}^n x_{ijk}}{bn}-\frac{\sum_{i=1}^a\sum_{j=1}^b\sum_{k=1}^n x_{ijk}}{abn}\Big)_{i=1,2,\ldots,a}=\Big(x_{i\bullet\bullet}-x_{\bullet\bullet\bullet}\Big)_{i=1,2,\ldots,a}$$

Therefore,

$$\begin{aligned}
&\|E(x)-\pi(\omega)\|_\Theta^2\\
=&||\Big(\frac{\sum_{j=1}^b\sum_{k=1}^n x_{ijk}}{bn}-\frac{\sum_{i=1}^a\sum_{j=1}^b\sum_{k=1}^n x_{ijk}}{abn}\Big)_{i=1,2,\ldots,a}-\Big(\alpha_i\Big)_{i=1,2,\ldots,a}||_\Theta^2\\
=&||\Big(\frac{\sum_{j=1}^b\sum_{k=1}^n x_{ijk}}{bn}-\frac{\sum_{i=1}^a\sum_{j=1}^b\sum_{k=1}^n x_{ijk}}{abn}\Big)_{i=1,2,\ldots,a}-\Big(\frac{\sum_{j=1}^b\mu_{ij}}{b}-\frac{\sum_{i=1}^a\sum_{j=1}^b\mu_{ij}}{ab}\Big)_{i=1,2,\ldots,a}||_\Theta^2\\
=&||\Big(\frac{\sum_{k=1}^n\sum_{j=1}^b(x_{ijk}-\mu_{ij})}{bn}-\frac{\sum_{i=1}^a\sum_{j=1}^b\sum_{k=1}^n(x_{ijk}-\mu_{ij})}{abn}\Big)_{i=1,2,\ldots,a}||_\Theta^2
\end{aligned}$$

and thus, if the null hypothesis H_N is assumed (i.e., $\mu_{i\cdot}-\mu_{\cdot\cdot}=\alpha_i=0$ $(\forall i=1,2,\ldots,a)$)

$$=||\Big(\frac{\sum_{k=1}^n\sum_{j=1}^b x_{ijk}}{bn}-\frac{\sum_{i=1}^a\sum_{j=1}^b\sum_{k=1}^n x_{ijk}}{abn}\Big)_{i=1,2,\ldots,a}||_\Theta^2=\sum_{i=1}^a(x_{ij\bullet}-x_{\bullet\bullet\bullet})^2\tag{7.42}$$

Thus, for any $\omega = (\mu_1, \mu_2)(\in \Omega = \mathbb{R} \times \mathbb{R})$, define the positive number η^α_ω (> 0) such that:

$$\eta^\alpha_\omega = \inf\{\eta > 0 : [G(E^{-1}(\mathrm{Ball}^C_{d^x_\Theta}(\pi(\omega); \eta))](\omega) \geq \alpha\} \tag{7.43}$$

Assume the null hypothesis H_N. Now let us calculate the η^α_ω as follows:

$$\begin{aligned} & E^{-1}(\mathrm{Ball}^C_{d^x_\Theta}(\pi(\omega); \eta)) = \{x \in X = \mathbb{R}^{abn} \ : \ d^x_\Theta(E(x), \pi(\omega)) > \eta\} \\ =& \{x \in X = \mathbb{R}^{abn} \ : \ \frac{abn \sum_{i=1}^a \sum_{j=1}^b (x_{ij\bullet} - x_{\bullet\bullet\bullet})^2}{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (x_{ijk} - x_{ij\bullet})^2} > \eta\} \end{aligned} \tag{7.44}$$

That is, for any $\omega = ((\mu_{ij})_{i=1,2,\ldots,a,\ j=1,2,\ldots,b}, \ ,\sigma) \in \Omega$ such that $\pi(\omega)(= (\alpha_1, \alpha_2, \ldots, \alpha_a)) \in H_N(= \{0, 0, \ldots, 0)\})$,

$$\begin{aligned} & [G^{abn}(E^{-1}(\mathrm{Ball}^C_{d^x_\Theta}(\pi(\omega); \eta)))(\omega) \\ =& \frac{1}{(\sqrt{2\pi}\sigma)^{abn}} \int \cdots \int_{E^{-1}(\mathrm{Ball}^C_{d^x_\Theta}(\pi(\omega);\eta))} \exp[- \frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (x_{ijk} - \mu_{ij})^2}{2\sigma^2}] \mathop{\times}_{k=1}^{n} \mathop{\times}_{j=1}^{b} \mathop{\times}_{i=1}^{a} dx_{ijk} \\ =& \frac{1}{(\sqrt{2\pi}\sigma)^{abn}} \int \cdots \int_{\frac{abn \sum_{i=1}^a \sum_{j=1}^b (x_{ij\bullet} - x_{\bullet\bullet\bullet})^2}{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (x_{ijk} - x_{ij\bullet})^2} > \eta^2} \exp[- \frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (x_{ijk} - \mu_{ij})^2}{2\sigma^2}] \mathop{\times}_{k=1}^{n} \mathop{\times}_{j=1}^{b} \mathop{\times}_{i=1}^{a} dx_{ijk} \\ =& \frac{1}{(\sqrt{2\pi})^{abn}} \int \cdots \int_{\frac{\frac{\sum_{i=1}^a \sum_{j=1}^b (x_{ij\bullet} - x_{\bullet\bullet\bullet})^2)}{(a-1)}}{\frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (x_{ijk} - x_{ij\bullet})^2}{ab(n-1)}} > \frac{\eta^2(ab(n-1))}{abn(a-1)}} \exp[- \frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (x_{ijk})^2}{2}] \mathop{\times}_{k=1}^{n} \mathop{\times}_{j=1}^{b} \mathop{\times}_{i=1}^{a} dx_{ijk} \end{aligned} \tag{7.45}$$

(A$_2$) using the formula of Gauss integrals derived in Kolmogorov's probability theory, we finally get as follows.

$$= \int_{\frac{\eta^2(n-1)}{n(a-1)}}^\infty p^F_{(a-1, ab(n-1))}(t) dt = \alpha \quad (\text{e.g., } \alpha = 0.05) \tag{7.46}$$

where $p^F_{(a-1,ab(n-1))}$ is the F-distribution with $(a-1, ab(n-1))$ degrees of freedom. Thus, it suffices to calculate the α-point $F^{a-1}_{ab(n-1),\alpha}$ Thus, we see

$$(\eta^\alpha_\omega)^2 = F^{a-1}_{ab(n-1),\alpha} \cdot n(a-1)/(n-1) \tag{7.47}$$

Therefore, we get $\widehat{R}_{\widehat{x}}^{\alpha;\Theta}$ (or, $\widehat{R}_{\widehat{x}}^{\alpha;X}$; the (α)-rejection region of $H_N = \{(0.0.\ldots,0)\}(\subseteq \Theta = \mathbb{R}^a)$) as follows:

$$\widehat{R}_{H_N}^{\alpha;\Theta} = \bigcap_{\omega=((\mu_i)_{i=1}^a,\sigma)\in\Omega(=\mathbb{R}^a\times\mathbb{R}_+) \text{ such that } \pi(\omega)=(\alpha_i)_{i=1}^a\in H_N=\{(0,0,\ldots,0)\}} \{E(x)(\in\Theta) : d_\Theta^x(E(x),\pi(\omega)) \geq \eta_\omega^\alpha\}$$

$$= \{E(x)(\in\Theta) : \frac{(\sum_{i=1}^a \sum_{j=1}^b (x_{ij\cdot} - x_{\cdot\cdot\cdot})^2)/(a-1)}{(\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (x_{ijk} - x_{ij\cdot})^2)/(ab(n-1))} \geq F_{ab(n-1),\alpha}^{a-1}\} \tag{7.48}$$

Thus,

$$\widehat{R}_{H_N}^{\alpha;X} = E^{-1}(\widehat{R}_{H_N}^{\alpha;\Theta}) = \{x(\in X) : \frac{(\sum_{i=1}^a \sum_{j=1}^b (x_{ij\cdot} - x_{\cdot\cdot\cdot})^2)/(a-1)}{(\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (x_{ijk} - x_{ij\cdot})^2)/(ab(n-1))} \geq F_{ab(n-1),\alpha}^{a-1}\} \tag{7.49}$$

□

♠**Note 7.3.** It should be noted that the mathematical part is only the (A_2).

7.3.3 Null hypothesis: $\mu_{\bullet 1} = \mu_{\bullet 2} = \cdots = \mu_{\bullet b} = \mu_{\bullet\bullet}$

Our present problem is as follows

Problem 7.4. [The two-way ANOVA]. Consider the parallel simultaneous normal measurement:

$$\mathsf{M}_{L^\infty(\mathbb{R}^{ab}\times\mathbb{R}_+)}(\mathsf{O}_G^{abn} = (X(\equiv \mathbb{R}^{abn}), \mathcal{B}_{\mathbb{R}}^{abn}, G^{abn}), S_{[(\mu=(\mu_{ij}\mid i=1,2,\cdots,a, j=1,2,\cdots,b),\sigma)]})$$

where the null hypothesis

$$\mu_{\bullet 1} = \mu_{\bullet 2} = \cdots = \mu_{\bullet b} = \mu_{\bullet\bullet}$$

is assumed. Let $0 < \alpha \ll 1$.
Then, find the largest $\widehat{R}_{H_N}^{\alpha;\Theta}(\subseteq \Theta)$(independent of σ) such that

(B)$'$ the probability that a measured value $x(\in \mathbb{R}^{abn})$ obtained by $\mathsf{M}_{L^\infty(\mathbb{R}^{ab}\times\mathbb{R}_+)}(\mathsf{O}_G^{abn} = (X(\equiv \mathbb{R}^{abn}), \mathcal{B}_{\mathbb{R}}^{abn}, G^{abn}), S_{[(\mu=(\mu_{ij}\mid i=1,2,\cdots,a, j=1,2,\cdots,b),\sigma)]})$ satisfies that

$$E(x) \in \widehat{R}_{H_N}^{\alpha;\Theta}$$

is less than α.

Since a and b have the same role, by the similar way of §7.3.2, we can easily solve Problem 7.4.

7.3.4 Null hypothesis: $(\alpha\beta)_{ij} = 0$ $(\forall i = 1, 2, \ldots, a,\ j = 1, 2, \ldots, b\ )$

Now, put

$$\Theta = \mathbb{R}^{ab} \tag{7.50}$$

And, define the system quantity$\pi : \Omega \to \Theta$ by

$$\Omega = \mathbb{R}^{ab} \times \mathbb{R}_+ \ni \omega = ((\mu_{ij})_{i=1,2,\ldots,a,\ j=1,2,\ldots,b}, \sigma) \mapsto \pi(\omega) = ((\alpha\beta)_{ij})_{i=1,2,\ldots,a,\ j=1,2,\ldots,b} \in \Theta = \mathbb{R}^{ab} \tag{7.51}$$

Here, recall:

$$(\alpha\beta)_{ij} = \mu_{ij} - \mu_{i\bullet} - \mu_{\bullet j} + \mu_{\bullet\bullet} \tag{7.52}$$

Also, the estimator $E : X(= \mathbb{R}^{abn}) \to \Theta(= \mathbb{R}^{ab})$ is defined by

$$\begin{aligned} & E((x_{ijk})_{i=1,\ldots,a,\ j=1,2,\ldots b,\ k=1,2,\ldots,n}) \\ =& \Big(\frac{\sum_{k=1}^n x_{ijk}}{n} - \frac{\sum_{j=1}^b \sum_{k=1}^n x_{ijk}}{bn} - \frac{\sum_{j=1}^b \sum_{k=1}^n x_{ijk}}{an} + \frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n x_{ijk}}{abn}\Big)_{i=1,2,\ldots,a\ j=1,2,\ldots b,} \\ =& \Big(x_{ij\bullet} - x_{i\bullet\bullet} - x_{\bullet j\bullet} + x_{\bullet\bullet\bullet}\Big)_{i=1,2,\ldots,a\ j=1,2,\ldots b,} \end{aligned} \tag{7.53}$$

Our present problem is as follows

Problem 7.5. [The two way ANOVA]. Consider the parallel simultaneous normal measurement:

$$\mathsf{M}_{L^\infty(\mathbb{R}^{ab}\times\mathbb{R}_+)}(\mathsf{O}_G^{abn} = (X(\equiv \mathbb{R}^{abn}), \mathcal{B}_{\mathbb{R}}^{abn}, G^{abn}), S_{[(\mu=(\mu_{ij}\ \mid\ i=1,2,\cdots,a, j=1,2,\cdots,b),\sigma)]})$$

The null hypothesis $H_N(\subseteq \Theta = \mathbb{R}^{ab})$ is defined by

$$H_N = \{((\alpha\beta)_{ij})_{i=1,2,\ldots,a,\ j=1,2,\ldots,b} \in \Theta = \mathbb{R}^{ab}\ :\ (\alpha\beta)_{ij} = 0, (\forall i = 1, 2, \ldots, a,\ \ j = 1, 2, \ldots, b)\} \tag{7.54}$$

That is,

$$(\alpha\beta)_{ij} = \mu_{ij} - \mu_{i\bullet} - \mu_{\bullet j} + \mu_{\bullet\bullet} = 0 \quad (i = 1, 2, \cdots, a, \quad j = 1, 2, \cdots, b) \tag{7.55}$$

Let $0 < \alpha \ll 1$.
Then, find the largest $\widehat{R}_{H_N}^{\alpha;\Theta}(\subseteq \Theta)$(independent of σ) such that

(C$_1$) the probability that a measured value $x(\in \mathbb{R}^{abn})$ obtained by $\mathsf{M}_{L^\infty(\mathbb{R}^{ab}\times\mathbb{R}_+)}(\mathsf{O}_G^{abn} = (X(\equiv \mathbb{R}^{abn}), \mathcal{B}_{\mathbb{R}}^{abn}, G^{abn}), S_{[(\mu=(\mu_{ij} \mid i=1,2,\cdots,a,j=1,2,\cdots,b),\sigma)]})$ satisfies that

$$E(x) \in \widehat{R}_{H_N}^{\alpha;\Theta}$$

is less than α.

Now,

$$\|\theta^{(1)} - \theta^{(2)}\|_\Theta = \sqrt{\sum_{i=1}^a \sum_{j=1}^b \Big(\theta_{ij}^{(\ell)} - \theta_{ij}^{(\ell)}\Big)^2} \tag{7.56}$$
$$(\forall \theta^{(\ell)} = (\theta_{ij}^{(\ell)})_{i=1,2,\ldots,a,\ j=1,2,\ldots,b} \in \mathbb{R}^{ab},\ \ell=1,2)$$

and, define the semi-distance d_Θ^x in Θ by

$$d_\Theta^x(\theta^{(1)}, \theta^{(2)}) = \frac{\|\theta^{(1)} - \theta^{(2)}\|_\Theta}{\sqrt{\overline{SS}(x)}} \qquad (\forall \theta^{(1)}, \theta^{(2)} \in \Theta, \forall x \in X) \tag{7.57}$$

$$
\begin{aligned}
&E((x_{ijk} - \mu_{ij})_{i=1,\ldots,a,\ j=1,2,\ldots b,\ k=1,2,\ldots,n})\\
=&\Big(\frac{\sum_{k=1}^n (x_{ijk} - \mu_{ij})}{n} - \frac{\sum_{j=1}^b \sum_{k=1}^n (x_{ijk} - \mu_{ij})}{bn}\\
&\quad - \frac{\sum_{j=1}^b \sum_{k=1}^n (x_{ijk} - \mu_{ij})}{an} + \frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (x_{ijk} - \mu_{ij})}{abn}\Big)_{i=1,2,\ldots,a\ j=1,2,\ldots b,}\\
=&\Big((x_{ij\bullet} - \mu_{ij}) - (x_{i\bullet\bullet} - \mu_{i\bullet}) - (x_{\bullet j\bullet} - \mu_{\bullet j}) + (x_{\bullet\bullet\bullet} - \mu_{\bullet\bullet})\Big)_{i=1,2,\ldots,a\ j=1,2,\ldots b,}\\
=&\Big(x_{ij\bullet} - x_{i\bullet\bullet} - x_{\bullet j\bullet} + x_{\bullet\bullet\bullet}\Big)_{i=1,2,\ldots,a\ j=1,2,\ldots b} \qquad \text{(Remark:null hypothesis } (\alpha\beta)_{ij} = 0)
\end{aligned} \tag{7.58}
$$

Therefore,

$$E((x_{ijk})_{i=1,\ldots,a,\ j=1,2,\ldots b,\ k=1,2,\ldots,n}) = E((x_{ijk} - \mu_{ij})_{i=1,\ldots,a,\ j=1,2,\ldots b,\ k=1,2,\ldots,n}) \tag{7.59}$$

Thus, for each $i = 1, \ldots, a,\ j = 1, 2, \ldots b$,

$$
\begin{aligned}
&E_{ij}(x_{ijk} - \mu_{ij})\\
=&\frac{\sum_{k=1}^n (x_{ijk} - \mu_{ij})}{n} - \frac{\sum_{j=1}^b \sum_{k=1}^n (x_{ijk} - \mu_{ij})}{bn} - \frac{\sum_{j=1}^b \sum_{k=1}^n (x_{ijk} - \mu_{ij})}{an}\\
&\quad + \frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (x_{ijk} - \mu_{ij})}{abn}\\
=&E_{ij}(x) - (\alpha\beta)_{ij}
\end{aligned}
$$

$$=x_{ij\bullet} - x_{i\bullet\bullet} - x_{\bullet j\bullet} + x_{\bullet\bullet\bullet} - (\alpha\beta)_{ij} \tag{7.60}$$

And, we see:

$$\begin{aligned}&\|E(x) - \pi(\omega)\|^2_\Theta \\ =&\|\Big(E_{ij}(x) - (\alpha\beta)_{ij}\Big)_{i=1,2,\ldots,a\ j=1,2,\ldots b}\|^2_\Theta\end{aligned} \tag{7.61}$$

Recalling that the null hypothesis H_N (i.e., $(\alpha\beta)_{ij} = 0$ $(\forall i = 1,2,\ldots,a,\ j = 1,2,\ldots,b)$), we see

$$=\sum_{i=1}^a\sum_{j=1}^b(x_{ij\bullet} - x_{i\bullet\bullet} - x_{\bullet j\bullet} + x_{\bullet\bullet\bullet})^2 \tag{7.62}$$

Thus, for each $\omega = (\mu,\sigma)(\in \Omega = \mathbb{R}^{ab} \times \mathbb{R})$, define the positive real η^α_ω (>0) such that

$$\eta^\alpha_\omega = \inf\{\eta > 0 : [G(E^{-1}(\mathrm{Ball}^C_{d^x_\Theta}(\pi(\omega);\eta))](\omega) \geq \alpha\} \tag{7.63}$$

Recalling the null hypothesisH_N (i.e., $(\alpha\beta)_{ij} = 0$ $(\forall i = 1,2,\ldots,a,\ j = 1,2,\ldots,b)$), calculate the η^α_ωas follows.

$$\begin{aligned}&E^{-1}(\mathrm{Ball}^C_{d^x_\Theta}(\pi(\omega);\eta)) = \{x \in X = \mathbb{R}^{abn}\ :\ d^x_\Theta(E(x),\pi(\omega)) > \eta\} \\ =&\{x \in X = \mathbb{R}^{abn}\ :\ \frac{abn\sum_{i=1}^a\sum_{j=1}^b(x_{ij\bullet} - x_{i\bullet\bullet} - x_{\bullet j\bullet} + x_{\bullet\bullet\bullet})^2}{\sum_{i=1}^a\sum_{j=1}^b\sum_{k=1}^n(x_{ijk} - x_{ij\bullet})^2} > \eta^2\}\end{aligned} \tag{7.64}$$

Thus, for any $\omega = ((\mu_{ij})_{i=1,2,\ldots,a,\ j=1,2,\ldots,b},\ ,\sigma) \in \Omega = \mathbb{R}^{ab} \times \mathbb{R}_+$ such that $\pi(\omega) \in H_N(\subseteq \mathbb{R}^{ab})$ (i.e., $(\alpha\beta)_{ij} = 0$ $(\forall i = 1,2,\ldots,a,\ j = 1,2,\ldots,b)$), we see:

$$\begin{aligned}&[G^{abn}(E^{-1}(\mathrm{Ball}^C_{d^x_\Theta}(\pi(\omega);\eta)))(\omega) \\ =&\frac{1}{(\sqrt{2\pi}\sigma)^{abn}}\underset{E^{-1}(\mathrm{Ball}^C_{d^x_\Theta}(\pi(\omega);\eta))}{\int\cdots\int}\exp[-\frac{\sum_{i=1}^a\sum_{j=1}^b\sum_{k=1}^n(x_{ijk}-\mu_{ij})^2}{2\sigma^2}]\mathop{\times}_{k=1}^n\mathop{\times}_{j=1}^b\mathop{\times}_{i=1}^a dx_{ijk} \\ =&\frac{1}{(\sqrt{2\pi}\sigma)^{abn}}\underset{\{x\in X\,:\,d^x_\Theta(E(x),\pi(\omega)\geq\eta\}}{\int\cdots\int}\exp[-\frac{\sum_{i=1}^a\sum_{j=1}^b\sum_{k=1}^n(x_{ijk}-\mu_{ij})^2}{2\sigma^2}]\mathop{\times}_{k=1}^n\mathop{\times}_{j=1}^b\mathop{\times}_{i=1}^a dx_{ijk} \\ =&\frac{1}{(\sqrt{2\pi})^{abn}}\underset{\frac{\sum_{i=1}^a\sum_{j=1}^b(x_{ij\bullet}-x_{i\bullet\bullet}-x_{\bullet j\bullet}+x_{\bullet\bullet\bullet})^2}{\sum_{i=1}^a\sum_{j=1}^b\sum_{k=1}^n(x_{ijk}-x_{ij\bullet})^2}>\frac{\eta^2}{abn}}{\int\cdots\int}\exp[-\frac{\sum_{i=1}^a\sum_{j=1}^b\sum_{k=1}^n(x_{ijk})^2}{2}]\mathop{\times}_{k=1}^n\mathop{\times}_{j=1}^b\mathop{\times}_{i=1}^a dx_{ijk}\end{aligned}$$

$$=\frac{1}{(\sqrt{2\pi})^{abn}} \underset{\frac{\frac{\sum_{i=1}^a \sum_{j=1}^b (x_{ij\bullet} - x_{i\bullet\bullet} - x_{\bullet j\bullet} + x_{\bullet\bullet\bullet})^2}{(a-1)(b-1)}}{\frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (x_{ijk} - x_{ij\bullet})^2}{ab(n-1)}} > \frac{\eta^2(ab(n-1))}{abn(a-1)(b-1)}}{\int \cdots \int} \exp[- \frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (x_{ijk})^2}{2}] \mathop{\times}_{k=1}^{n} \mathop{\times}_{j=1}^{b} \mathop{\times}_{i=1}^{a} dx_{ijk} \tag{7.65}$$

(C$_2$) Then, by the formula of Gauss integrals 7.8(D) (§7.4) , we see

$$= \int_{\frac{\eta^2(n-1)}{n(a-1)(b-1)}}^\infty p^F_{((a-1)(b-1),ab(n-1))}(t)dt = \alpha(\text{ e.g., } \alpha = 0.05) \tag{7.66}$$

where $p^F_{((a-1)(b-1),ab(n-1))}$ is a probability density function of the F-distribution with $((a-1)(b-1), ab(n-1))$ degrees of freedom.

Hence, it suffices to the following equation:

$$\frac{\eta^2(n-1)}{n(a-1)(b-1)} = F^{(a-1)(b-1)}_{ab(n-1),\alpha} (= \text{“}\alpha\text{-point”}) \tag{7.67}$$

thus, we see,

$$(\eta^\alpha_\omega)^2 = F^{(a-1)(b-1)}_{ab(n-1),\alpha} n(a-1)(b-1)/(n-1) \tag{7.68}$$

Therefore, we get the (α)-rejection region $\widehat{R}^{\alpha;\Theta}_{\widehat{x}}$ (or, $\widehat{R}^{\alpha;X}_{\widehat{x}}$; $H_N = \{((\alpha\beta)_{ij})_{i=1,2,\cdots,a, j=1,2,\cdots,b} : (\alpha\beta)_{ij} = 0 \ (i = 1, 2, \cdots, a, j = 1, 2, \cdots, b)\} (\subseteq \Theta = \mathbb{R}^{ab})$):

$$\begin{aligned} \widehat{R}^{\alpha;\Theta}_{H_N} = & \bigcap_{\omega = ((\mu_{ij})_{i=1}^a{}_{j=1}^b, \sigma) \in \Omega (= \mathbb{R}^a \times \mathbb{R}_+) \text{ such that } \pi(\omega) = (\alpha\beta)_{ij} \in H_N} \{E(x) (\in \Theta) : d^x_\Theta (E(x), \pi(\omega)) \geq \eta^\alpha_\omega\} \\ = & \{E(x)(\in \Theta) : \frac{(\sum_{i=1}^a \sum_{j=1}^b (x_{ij\bullet} - x_{\bullet\bullet\bullet})^2)/((a-1)(b-1))}{(\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (x_{ijk} - x_{ij\bullet})^2)/(ab(n-1))} \geq F^{(a-1)(b-1)}_{ab(n-1),\alpha}\} \end{aligned} \tag{7.69}$$

Also,

$$\widehat{R}^{\alpha;X}_{H_N} = E^{-1}(\widehat{R}^{\alpha;\Theta}_{H_N}) = \{x (\in X) : \frac{(\sum_{i=1}^a \sum_{j=1}^b (x_{ij\bullet} - x_{\bullet\bullet\bullet})^2)/((a-1)(b-1))}{(\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (x_{ijk} - x_{ij\bullet})^2)/(ab(n-1))} \geq F^{(a-1)(b-1)}_{ab(n-1),\alpha}\} \tag{7.70}$$

♠**Note 7.4.** It should be noted that the mathematical part is only the (C$_2$).

7.4 Supplement(the formulas of Gauss integrals)

7.4.1 Normal distribution, chi-squared distribution, Student t-distribution, F-distribution

Definition 7.6. [Fdistribution]. Let $t \geq 0$, and n_1 and n_2 be natural numbers. The probability density function $p^F_{(n_1,n_2)}(t)$ of F-distribution with the degree of freedom(n_1, n_2) is defined by

$$p^F_{(n_1,n_2)}(t) = \frac{1}{B(n_1/2, n_2/2)} \Big(\frac{n_1}{n_2}\Big)^{n_1/2} \frac{t^{(n_1-2)/2}}{(1+n_1 t/n_2)^{(n_1+n_2)/2}} \qquad (t \geq 0) \tag{7.71}$$

where, $B(\cdot,\cdot)$ is the Beta function, that is, for $x, y > 0$,

$$B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1} dt$$

Note that

$$\begin{aligned} &F\text{-distribution with degree of freedom}(1, n-1) \\ =\ &\text{Student } t\text{-distribution with the degree of freedom}(n-1) \end{aligned}$$

Define two maps $\overline{\mu} : \mathbb{R}^n \to \mathbb{R}$ and $\overline{SS} : \mathbb{R}^n \to \mathbb{R}$ as follows.

$$\begin{aligned} \overline{\mu}(x) = \overline{\mu}(x_1, x_2, \cdots, x_n) &= \frac{\sum_{k=1}^n x_k}{n} \\ \overline{SS}(x) = \overline{SS}(x_1, x_2, \cdots, x_n) &= \sum_{k=1}^n (x_k - \overline{\mu}(x))^2 \\ (\forall x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n) \end{aligned}$$

Formula 7.7. [Gauss integral(normal distribution and chi-squared distribution)]. This was already mentioned in (6.6) and (6.7).

Formula 7.8. [Gauss integral(F-distribution)]. For $c \geq 0$,

$$\text{(A):}\quad \frac{1}{(\sqrt{2\pi})^n} \underset{c \leq \frac{n(\overline{\mu}(x))^2}{\overline{SS}(x)/(n-1)}}{\int\cdots\int} \exp[-\frac{\sum_{k=1}^n (x_k)^2}{2}] dx_1 dx_2 \cdots dx_n = \int_c^\infty p^F_{(1,n-1)}(t) dt \tag{7.72}$$

(B): For $n = \sum_{i=1}^a n_i$,

$$\frac{1}{(\sqrt{2\pi})^n} \underset{\frac{(\sum_{i=1}^a n_i (x_{i\bullet} - x_{\bullet\bullet})^2/(a-1)}{(\sum_{i=1}^a \sum_{k=1}^{n_i} (x_{ik} - x_{i\bullet})^2)/(n-a)} > c}{\int\cdots\int} \exp[-\frac{\sum_{i=1}^a \sum_{k=1}^{n_i} (x_{ik})^2}{2}] \mathop{\times}_{i=1}^{a} \mathop{\times}_{k=1}^{n_i} dx_{ik}$$

$$= \int_c^\infty p^F_{(a-1,n-a)}(t)dt \tag{7.73}$$

$$\text{(C)}: \quad \frac{1}{(\sqrt{2\pi})^{abn}} \underset{\frac{\frac{\sum_{i=1}^a \sum_{j=1}^b (x_{ij\bullet} - x_{\bullet\bullet\bullet})^2}{(a-1)}}{\frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (x_{ijk} - x_{ij\bullet})^2}{ab(n-1)}} > c}{\int \cdots \int} \exp[-\frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (x_{ijk})^2}{2}] \mathop{\times}_{k=1}^n \mathop{\times}_{j=1}^b \mathop{\times}_{i=1}^a dx_{ijk}$$

$$= \int_c^\infty p^F_{(a-1,ab(n-1))}(t)dt \tag{7.74}$$

Or, equivalently,

$$\text{(D)}: \quad \frac{1}{(\sqrt{2\pi})^{abn}} \underset{\frac{\frac{\sum_{i=1}^a \sum_{j=1}^b (x_{ij\bullet} - x_{i\bullet\bullet} - x_{\bullet j\bullet} + x_{\bullet\bullet\bullet})^2}{(a-1)(b-1)}}{\frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (x_{ijk} - x_{ij\bullet})^2}{ab(n-1)}} > c}{\int \cdots \int} \exp[-\frac{\sum_{i=1}^a \sum_{j=1}^b \sum_{k=1}^n (x_{ijk})^2}{2}] \mathop{\times}_{k=1}^n \mathop{\times}_{j=1}^b \mathop{\times}_{i=1}^a dx_{ijk}$$

$$= \int_c^\infty p^F_{((a-1)(b-1),ab(n-1))}(t)dt \tag{7.75}$$

Chapter 8

Practical logic–Do you believe in syllogism?–

The term "practical logic" means the logic in measurement theory. It is certain that pure logic (=mathematical logic) is merely a kind of rule in mathematics (or meta-mathematics). If it is so, the mathematical logic is not guaranteed to be applicable to our world. For instance, mathematical syllogism ("$A \Rightarrow B$" and "$B \Rightarrow C$" imply "$A \Rightarrow C$") does not assure the following famous statement:

($\sharp_1$) *Since Socrates is a man and all men are mortal, it follows that Socrates is mortal.*

That is, we think that

($\sharp_2$) the above ($\sharp_1$) is not clarified yet.

In this chapter, we prove the ($\sharp_1$) in classical systems. Also, we point out that syllogism does not hold in quantum systems[1]

8.1 Marginal observable and quasi-product observable

Definition 8.1. [(=Definition 3.19):quasi-product product observable] Let $\mathsf{O}_k = (X_k, \mathcal{F}_k, F_k)$ ($k = 1, 2, \ldots, n$) be observables in a W^*-algebra $\overline{\mathcal{A}}$. Assume that an observable $\mathsf{O}_{12\ldots n} = (\times_{k=1}^n X_k, \ \boxtimes_{k=1}^n \mathcal{F}_k, \ F_{12\ldots n})$ satisfies

$$F_{12\ldots n}(X_1 \times \cdots \times X_{k-1} \times \Xi_k \times X_{k+1} \times \cdots \times X_n) = F_k(\Xi_k). \tag{8.1}$$

[1] This chapter is mostly extracted from the following:

($\sharp$) Ref. [25]: S. Ishikawa, "Fuzzy Inferences by Algebraic Method," Fuzzy Sets and Systems, Vol. 87, No. 2, 1997, pp. 181-200. doi:10.1016/S0165-0114(96)00035-8

$$(\forall \Xi_k \in \mathcal{F}_k, \forall k = 1, 2, \ldots, n)$$

The observable $\mathsf{O}_{12\ldots n} = (\times_{k=1}^n X_k, \ \boxtimes_{k=1}^n \mathcal{F}_k, \ F_{12\ldots n})$ is called a *quasi-product observable* of $\{\mathsf{O}_k \mid k = 1, 2, \ldots, n\}$, and denoted by

$$\underset{k=1,2,\ldots,n}{\overset{\text{qp}}{\times}} \mathsf{O}_k = (\underset{k=1}{\overset{n}{\times}} X_k, \ \boxtimes_{k=1}^n \mathcal{F}_k, \ \underset{k=1,2,\ldots,n}{\overset{\text{qp}}{\times}} F_k).$$

Of course, a simultaneous observable is a kind of quasi-product observable. Therefore, quasi-product observable is not uniquely determined. Also, in quantum systems, the existence of the quasi-product observable is not always guaranteed.

Definition 8.2. [Image observable, marginal observable] Consider the basic structure $[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]$. And consider the observable $\mathsf{O} = (X, \mathcal{F}, F)$ in $\overline{\mathcal{A}}$. Let $(Y, \mathcal{G})$ be a measurable space, and let $f : X \to Y$ be a measurable map. Then, we can define the image observable $f(\mathsf{O}) = (X, \mathcal{F}, F \circ f^{-1})$ in $\overline{\mathcal{A}}$, where $F \circ f^{-1}$ is defined by

$$(F \circ f^{-1})(\Gamma) = F(f^{-1}(\Gamma)) \qquad (\forall \Gamma \in \mathcal{G}).$$

[Marginal observable] Consider the basic structure $[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]$. And consider the observable $\mathsf{O}_{12\ldots n} = (\times_{k=1}^n X_k, \ \boxtimes_{k=1}^n \mathcal{F}_k, \ F_{12\ldots n})$ in $\overline{\mathcal{A}}$. For any natural number j such that $1 \leqq j \leqq n$, define $F_{12\ldots n}^{(j)}$ such that

$$F_{12\ldots n}^{(j)}(\Xi_j) = F_{12\ldots n}(X_1 \times \cdots \times X_{j-1} \times \Xi_j \times X_{j+1} \times \cdots \times X_n) \quad (\forall \Xi_j \in \mathcal{F}_j).$$

Then we have the observable $\mathsf{O}_{12\ldots n}^{(j)} = (X_j, \mathcal{F}_j, F_{12\ldots n}^{(j)})$ in $\overline{\mathcal{A}}$. The $\mathsf{O}_{12\ldots n}^{(j)}$ is called a marginal observable of $\mathsf{O}_{12\ldots n}$ (or, precisely, (j)-marginal observable). Consider a map $P_j : \times_{k=1}^n X_k \to X_j$ such that

$$\underset{k=1}{\overset{n}{\times}} \ni (x_1, x_2, ..., x_j, ..., x_n) \mapsto x_j \in X_j.$$

Then, the marginal observable $\mathsf{O}_{12\ldots n}^{(j)}$ is characterized as the image observable $P_j(\mathsf{O}_{12\ldots n})$. The above can be easily generalized as follows. For example, define $\mathsf{O}_{12\ldots n}^{(12)} = (X_1 \times X_2, \mathcal{F}_1 \boxtimes \mathcal{F}_2, F_{12\ldots n}^{(12)})$ such that

$$F_{12\ldots n}^{(12)}(\Xi_1 \times \Xi_2) = F_{12\ldots n}^{(12)}(\Xi_1 \times \Xi_2 \times X_3 \times \cdots \times X_n) \quad (\forall \Xi_1 \in \mathcal{F}_1, \forall \Xi_2 \in \mathcal{F}_2).$$

Then, we have the (12)-marginal observable $\mathsf{O}_{12\ldots n}^{(12)} = (X_1 \times X_2, \mathcal{F}_1 \boxtimes \mathcal{F}_2, F_{12\ldots n}^{(12)})$. Of course, we also see that $F_{12\ldots n} = F_{12\ldots n}^{(12\ldots n)}$.

The following theorem is often used:

Theorem 8.3. Consider the basic structure

$$[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]$$

Let $\overline{\mathcal{A}}$ be a C^*-algebra. Let $\mathbf{O}_1 \equiv (X_1, \mathcal{F}_1, F_1)$ and $\mathbf{O}_2 \equiv (X_2, \mathcal{F}_2, F_2)$ be W^*-observables in $\overline{\mathcal{A}}$ such that at least one of them is a projective observable. $\Big($So, without loss of generality, we assume that $\mathbf{O}_2$ is projective, i.e., $F_2 = (F_2)^2$ $\Big)$. Then, the following statements are equivalent:

(i) There exists a quasi-product observable $\mathbf{O}_{12} \equiv (X_1 \times X_2, \mathcal{F}_1 \times \mathcal{F}_2, F_1 \overset{\text{qp}}{\times} F_2)$ with marginal observables $\mathbf{O}_1$ and $\mathbf{O}_2$.

(ii) $\mathbf{O}_1$ and $\mathbf{O}_2$ commute, that is, $F_1(\Xi_1)F_2(\Xi_2) = F_2(\Xi_2)F_1(\Xi_1)$ $(\forall \Xi_1 \in \mathcal{F}_1, \forall \Xi_2 \in \mathcal{F}_2)$.

Furthermore, if the above statements (i) and (ii) hold, the uniqueness of the quasi-product observable $\mathbf{O}_{12}$ of $\mathbf{O}_1$ and $\mathbf{O}_2$ is guaranteed.

Proof. See refs. [11, 25, 29].

Consider the measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}_{12}{=}(X_1 \times X_2, \mathcal{F}_1 \boxtimes \mathcal{F}_2, F_{12}), S_{[\rho]})$ with the sample probability space $(X_1 \times X_2, \mathcal{F}_1 \boxtimes \mathcal{F}_2, {}_{\mathcal{A}^*}\big(\rho, F_{12}(\cdot)\big)_{\overline{\mathcal{A}}})$.

Put

$$\mathrm{Rep}_\rho^{\Xi_1 \times \Xi_2}[\mathsf{O}_{12}] = \begin{bmatrix} {}_{\mathcal{A}^*}\big(\rho, F_{12}(\Xi_1 \times \Xi_2)\big)_{\overline{\mathcal{A}}} & {}_{\mathcal{A}^*}\big(\rho, F_{12}(\Xi_1 \times \Xi_2^c)\big)_{\overline{\mathcal{A}}} \\ {}_{\mathcal{A}^*}\big(\rho, F_{12}(\Xi_1^c \times \Xi_2)\big)_{\overline{\mathcal{A}}} & {}_{\mathcal{A}^*}\big(\rho, F_{12}(\Xi_1^c \times \Xi_2^c)\big)_{\overline{\mathcal{A}}} \end{bmatrix} \quad (\forall \Xi_1 \in \mathcal{F}_1, \forall \Xi_2 \in \mathcal{F}_2)$$

where, Ξ^c is the complement of Ξ $\{x \in X \mid x \notin \Xi\}$. Also, note that

$$\begin{aligned}
&{}_{\mathcal{A}^*}\big(\rho, F_{12}(\Xi_1 \times \Xi_2)\big)_{\overline{\mathcal{A}}} + {}_{\mathcal{A}^*}\big(\rho, F_{12}(\Xi_1 \times \Xi_2^c)\big)_{\overline{\mathcal{A}}} = {}_{\mathcal{A}^*}\big(\rho, F_{12}^{(1)}](\Xi_1)\big)_{\overline{\mathcal{A}}} \\
&{}_{\mathcal{A}^*}\big(\rho, F_{12}(\Xi_1^c \times \Xi_2^c)\big)_{\overline{\mathcal{A}}} + {}_{\mathcal{A}^*}\big(\rho, F_{12}(\Xi_1^c \times \Xi_2)\big)_{\overline{\mathcal{A}}} = {}_{\mathcal{A}^*}\big(\rho, F_{12}^{(1)}(\Xi_1^c)\big)_{\overline{\mathcal{A}}} \\
&{}_{\mathcal{A}^*}\big(\rho, F_{12}(\Xi_1^c \times \Xi_2^c)\big)_{\overline{\mathcal{A}}} + {}_{\mathcal{A}^*}\big(\rho, F_{12}(\Xi_1 \times \Xi_2^c)\big)_{\overline{\mathcal{A}}} = {}_{\mathcal{A}^*}\big(\rho, F_{12}^{(2)}(\Xi_2^c)\big)_{\overline{\mathcal{A}}} \\
&{}_{\mathcal{A}^*}\big(\rho, F_{12}(\Xi_1 \times \Xi_2^c)\big)_{\overline{\mathcal{A}}} + {}_{\mathcal{A}^*}\big(\rho, F_{12}(\Xi_1^c \times \Xi_2^c)\big)_{\overline{\mathcal{A}}} = {}_{\mathcal{A}^*}\big(\rho, F_{12}^{(2)}(\Xi_2^c)\big)_{\overline{\mathcal{A}}}
\end{aligned}$$

□

We have the following lemma.

Lemma 8.4. [The condition of quasi-product observables] Consider the general basic structure

$$[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)].$$

Let $\mathsf{O}_1 = (X_1, \mathcal{F}_1, F_1)$ and $\mathsf{O}_2 = (X_2, \mathcal{F}_2, F_2)$ be observables in $C(\Omega)$. Let $\mathsf{O}_{12} = (X_1 \times X_2, \mathcal{F}_1 \times \mathcal{F}_2, F_{12} {=} F_1 \overset{\mathrm{qp}}{\times} F_2)$ be a quasi-product observable of O_1 and O_2. That is, it holds that

$$F_1 = F_{12}^{(1)}, \qquad F_2 = F_{12}^{(2)}$$

Then, putting $\alpha_\rho^{\Xi_1 \times \Xi_2} = {}_{\mathcal{A}^*}\big(\rho, F_{12}(\Xi_1 \times \Xi_2)\big)_{\overline{\mathcal{A}}} = \rho(F_{12}(\Xi_1 \times \Xi_2))$, we see

$$\mathrm{Rep}_\rho^{\Xi_1 \times \Xi_2}[\mathsf{O}_{12}] = \begin{bmatrix} {}_{\mathcal{A}^*}\big(\rho, F_{12}(\Xi_1 \times \Xi_2)\big)_{\overline{\mathcal{A}}} & {}_{\mathcal{A}^*}\big(\rho, F_{12}(\Xi_1 \times \Xi_2^c)\big)_{\overline{\mathcal{A}}} \\ {}_{\mathcal{A}^*}\big(\rho, F_{12}(\Xi_1^c \times \Xi_2)\big)_{\overline{\mathcal{A}}} & {}_{\mathcal{A}^*}\big(\rho, F_{12}(\Xi_1^c \times \Xi_2^c)\big)_{\overline{\mathcal{A}}} \end{bmatrix}$$
$$= \begin{bmatrix} \alpha_\rho^{\Xi_1 \times \Xi_2} & \rho(F_1(\Xi_1)) - \alpha_\rho^{\Xi_1 \times \Xi_2} \\ \rho(F_2(\Xi_2)) - \alpha_\rho^{\Xi_1 \times \Xi_2} & 1 + \alpha_\rho^{\Xi_1 \times \Xi_2} - \rho(F_1(\Xi_1)) - \rho(F_2(\Xi_2)) \end{bmatrix} \tag{8.2}$$

and

$$\max\{0, \rho(F_1(\Xi_1)) + \rho(F_2(\Xi_2)) - 1\} \leqq \alpha_\rho^{\Xi_1 \times \Xi_2} \leqq \min\{\rho(F_1(\Xi_1)),\ \rho(F_2(\Xi_2))\}$$
$$(\forall \Xi_1 \in \mathcal{F}_1, \forall \Xi_2 \in \mathcal{F}_2, \forall \rho \in \mathfrak{S}^p(\mathcal{A}^*)) \tag{8.3}$$

Reversely, for any $\alpha_\rho^{\Xi_1 \times \Xi_2}$ satisfying (8.3), the observable O_{12} defined by (8.2) is a quasi-product observable of O_1 and O_2. Also, it holds that

$$\rho(F(\Xi_1 \times \Xi_2^c)) = 0 \iff \alpha_\rho^{\Xi_1 \times \Xi_2} = \rho(F_1(\Xi_1))$$
$$\Longrightarrow \rho(F_1(\Xi_1)) \leqq \rho(F_2(\Xi_2)) \tag{8.4}$$

Proof. Though this lemma is easy, we add a brief proof for completeness. $0 \leqq \rho(F((\Xi_1' \times \Xi_2'))) \leqq 1$, $(\forall \Xi_1' \in \mathcal{F}_1, \Xi_2' \in \mathcal{F}_2)$ we see, by (8.2) that

$$0 \leqq \alpha_\rho^{\Xi_1 \times \Xi_2} \leqq 1$$
$$0 \leqq 1 + \alpha_\rho^{\Xi_1 \times \Xi_2} - \rho(F_1(\Xi_1)) - \rho(F_2(\Xi_2)) \leqq 1$$
$$0 \leqq \rho(F_2(\Xi_2)) - \alpha_\rho^{\Xi_1 \times \Xi_2} \leqq 1$$
$$0 \leqq \rho(F_1(\Xi_1)) - \alpha_\rho^{\Xi_1 \times \Xi_2} \leqq 1$$

which clearly implies (8.3). Conversely. if α satisfies (8.3),then we easily see (8.2),Also, (8.4) is obvious. This completes the proof. $\square$

Let $\mathsf{O}_{12} = (X_1 \times X_2, \mathcal{F}_1 \boxtimes \mathcal{F}_2, F_{12} {=} F_1 \overset{\mathrm{qp}}{\times} F_2)$ be a quasi-product observable of $\mathsf{O}_1 = (X_1, \mathcal{F}_1, F_1)$ and $\mathsf{O}_2 = (X_2, \mathcal{F}_2, F_2)$ in $\overline{\mathcal{A}}$. Consider the measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}_{12} {=} (X_1 \times X_2, \mathcal{F}_1 \boxtimes \mathcal{F}_2, F_{12} {=} F_1 \overset{\mathrm{qp}}{\times} F_2),$

$S_{[\rho]})$). And assume that a measured value(x_1, x_2) $(\in X_1 \times X_2)$ is obtained. And assume that we know that $x_1 \in \Xi_1$. Then, the probability (i.e., the conditional probability) that $x_2 \in \Xi_2$ is given by

$$P = \frac{\rho(F_{12}(\Xi_1 \times \Xi_2))}{\rho(F_1(\Xi_1))} = \frac{\rho(F_{12}(\Xi_1 \times \Xi_2))}{\rho(F_{12}(\Xi_1 \times \Xi_2)) + \rho(F_{12}(\Xi_1 \times \Xi_2^c))}$$

And further, it is, by (8.3), estimated as follows.

$$\frac{\max\{0, \rho(F_1(\Xi_1)) + \rho(F_2(\Xi_2)) - 1\}}{\rho(F_{12}(\Xi_1 \times \Xi_2)) + \rho(F_{12}(\Xi_1 \times \Xi_2^c))} \leqq P \leqq \frac{\min\{\rho(F_1(\Xi_1)),\ \rho(F_2(\Xi_2))\}}{\rho(F_{12}(\Xi_1 \times \Xi_2)) + \rho(F_{12}(\Xi_1 \times \Xi_2^c))}$$

Example 8.5. [Example of tomatoes] Let $\Omega = \{\omega_1, \omega_2, ..., \omega_N\}$ be a set of tomatoes, which is regarded as a compact Hausdorff space with the discrete topology. Consider the classical basic structure

$$[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))]$$

Consider yes-no observables $\mathsf{O}_{\rm RD} \equiv (X_{\rm RD}, 2^{X_{\rm RD}}, F_{\rm RD})$ and $\mathsf{O}_{\rm SW} \equiv (X_{\rm SW}, 2^{X_{\rm SW}}, F_{\rm SW})$ in $C(\Omega)$ such that:

$$X_{\rm RD} = \{y_{\rm RD}, n_{\rm RD}\} \text{ and } X_{\rm SW} = \{y_{\rm SW}, n_{\rm SW}\},$$

where we consider that "$y_{\rm RD}$" and "$n_{\rm RD}$" respectively mean "RED" and "NOT RED". Similarly, "$y_{\rm SW}$" and "$n_{\rm SW}$" respectively mean "SWEET" and "NOT SWEET".
For example, the ω_1 is red and not sweet, the ω_2 is red and sweet, etc. as follows.

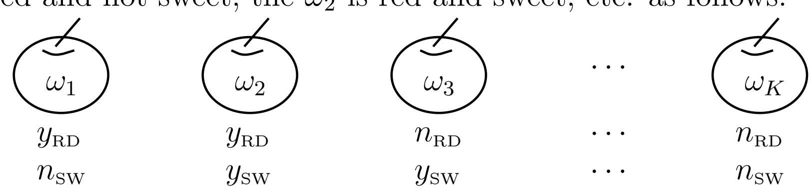


Figure 8.1: Tomatoes (Red or Sweet?)

Next, consider the quasi-product observable as follows.

$$\mathsf{O}_{12} = (X_{\rm RD} \times X_{\rm SW}, 2^{X_{\rm RD} \times X_{\rm SW}}, F{=}F_{\rm RD} \overset{\rm qp}{\times} F_{\rm SW})$$

That is,

$$\mathrm{Rep}_{\omega_k}^{\{(y_{\rm RD}, y_{\rm SW})\}}[\mathsf{O}_{12}] = \begin{bmatrix} [F(\{(y_{\rm RD}, y_{\rm SW})\})](\omega_k) & [F(\{(y_{\rm RD}, n_{\rm SW})\})](\omega_k) \\ [F(\{(n_{\rm RD}, y_{\rm SW})\})](\omega_k) & [F(\{(n_{\rm RD}, n_{\rm SW})\})](\omega_k) \end{bmatrix}$$

$$
= \begin{bmatrix} \alpha_{\{(y_{\rm RD}, y_{\rm SW})\}} & [F_{\rm RD}(\{y_{\rm RD}\})] - \alpha_{\{(y_{\rm RD}, y_{\rm SW})\}} \\ [F_{\rm SW}(\{y_{\rm SW}\})] - \alpha_{\{(y_{\rm RD}, y_{\rm SW})\}} & 1 + \alpha_{\{(y_{\rm RD}, y_{\rm SW})\}} - [F_{\rm RD}(\{y_{\rm RD}\})] - [F_{\rm SW}(\{y_{\rm SW}\})] \end{bmatrix}
$$

where $\alpha_{\{(y_{\rm RD}, y_{\rm SW})\}}(\omega_k)$ satisfies the (8.3). When we know that a tomato ω_k is red, the probability P that the tomato ω_k is sweet is given by

$$
P = \frac{[F(\{(y_{\rm RD}, y_{\rm SW})\})](\omega_k)}{[F(\{(y_{\rm RD}, y_{\rm SW})\})](\omega_k) + [F(\{(y_{\rm RD}, n_{\rm SW})\})](\omega_k)} = \frac{[F(\{(y_{\rm RD}, y_{\rm SW})\})](\omega_k)}{[F_{\rm RD}(\{y_{\rm RD}\})](\omega_k)}
$$

Since $[F(\{(y_{\rm RD}, y_{\rm SW})\})](\omega_k) = \alpha_{\{(y_{\rm RD}, y_{\rm SW})\}}(\omega_k)$, the conditional probability P is estimated by

$$
\frac{\max\{0, [F_1(\{y_{\rm RD}\})](\omega_k) + [F_2(\{y_{\rm SW}\})](\omega_k) - 1\}}{[F_{\rm RD}(\{y_{\rm RD}\})](\omega_k)} \leqq P \leqq \frac{\min[F_1(\{y_{\rm SW}\})](\omega_k),\ [F_2(\{y_{\rm SW}\})](\omega_k)\}}{[F_{\rm RD}(\{y_{\rm RD}\})](\omega_k)}
$$

8.2 Implication—the definition of "⇒"

8.2.1 Implication and contraposition

In Example 8.5, consider the case that $[F(\{(y_{\rm RD}, n_{\rm SW})\})](\omega) = 0$. In this case, we see

$$\frac{[F(\{(y_{\rm RD}, y_{\rm SW})\})](\omega)}{[F(\{(y_{\rm RD}, y_{\rm SW})\})](\omega) + [F(\{(y_{\rm RD}, n_{\rm SW})\})](\omega)} = 1$$

Therefore, when we know that a tomato ω is red, the probability, that the tomato ω is sweet, is equal to 1. That is,

$$\text{"}[F(\{(y_{\rm RD}, n_{\rm SW})\})](\omega) = 0\text{"} \quad \Longleftrightarrow \quad \Big[\text{"Red"} \Longrightarrow \text{"Sweet"}\Big]$$

Motivated by the above argument, we have the following definition.

Definition 8.6. [Implication] Consider the general basic structure

$$[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]$$

Let $\mathsf{O}_{12} = (X_1 \times X_2, \mathcal{F}_1 \boxtimes \mathcal{F}_2, F_{12}{=}F_1 \overset{\rm qp}{\times} F_2)$ be a quasi-observable in $\overline{\mathcal{A}}$ Let $\rho \in \mathfrak{S}^p(\mathcal{A}^*)$, $\Xi_1 \in \mathcal{F}_1$, $\Xi_2 \in \mathcal{F}_2$. Then, if it holds that

$$\rho(F_{12}(\Xi_1 \times (\Xi_2^c))) = 0$$

this is denoted by

$$[\mathsf{O}_{12}^{(1)}; \Xi_1] \underset{\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}_{12}, S_{[\rho]})}{\Longrightarrow} [\mathsf{O}_{12}^{(2)}; \Xi_2] \tag{8.5}$$

Of course, this (8.5) should be read as follows.

(A) Assume that a measured value $(x_1, x_2)(\in X_1 \times X_2)$ is obtained by a measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{12}, S_{[\omega]})$. When we know that $x_1 \in \Xi_1$, then we can assure that $x_2 \in \Xi_2$.

The above argument is generalized as follows. Let $\mathsf{O}_{12\dots n} = (\times_{k=1}^n X_k, \boxtimes_{k=1}^n \mathcal{F}_k, F_{12\dots n} = \overset{\rm qp}{\underset{k=1,2,\dots,n}{\times}} F_k)$ be a quasi-product observable in $\overline{\mathcal{A}}$. Let $\Xi_1 \in \mathcal{F}_i$ and $\Xi_2 \in \mathcal{F}_j$. Then, the condition

$${}_{\mathcal{A}^*}\big(\rho, F_{12\dots n}^{(ij)}(\Xi_i \times (\Xi_j^c))\big)_{\overline{\mathcal{A}}} = 0$$

(where, $\Xi^c = X \setminus \Xi$) is denoted by

$$[\mathsf{O}_{12\dots n}^{(i)}; \Xi_i] \underset{\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}_{12\dots n}, S_{[\rho]})}{\Longrightarrow} [\mathsf{O}_{12\dots n}^{(j)}; \Xi_j] \tag{8.6}$$

Theorem 8.7. [Contraposition] Let $\mathsf{O}_{12} = (X_1 \times X_2, \mathcal{F}_1 \times \mathcal{F}_2, F_{12}{=}F_1 \overset{\mathrm{qp}}{\times} F_2)$ be a quasi-product observable in $\overline{\mathcal{A}}$. Let $\rho \in \mathfrak{S}^p(\mathcal{A}^*)$. Let $\Xi_1 \in \mathcal{F}_1$ and $\Xi_2 \in \mathcal{F}_2$. If it holds that

$$[\mathsf{O}_{12}^{(1)}; \Xi_1] \underset{\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}_{12}, S_{[\rho]})}{\Longrightarrow} [\mathsf{O}_{12}^{(2)}; \Xi_2] \tag{8.7}$$

then we see:

$$[\mathsf{O}_{12}^{(1)}; \Xi_1^c] \underset{\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}_{12}, S_{[\rho]})}{\Longleftarrow} [\mathsf{O}_{12}^{(2)}; \Xi_2^c]$$

Proof. The proof is easy, but we add it. Assume the condition (8.7). That is,

$$_{\mathcal{A}^*}\big(\rho, F_{12}(\Xi_1 \times (X_2 \setminus \Xi_2))\big)_{\overline{\mathcal{A}}} = 0$$

Since $\Xi_1 \times {\Xi_2}^c = (\Xi_1^c)^c \times \Xi_2^c$ we see

$$_{\mathcal{A}^*}\big(\rho, F_{12}((\Xi_1^c)^c \times \Xi_2^c)\big)_{\overline{\mathcal{A}}} = 0$$

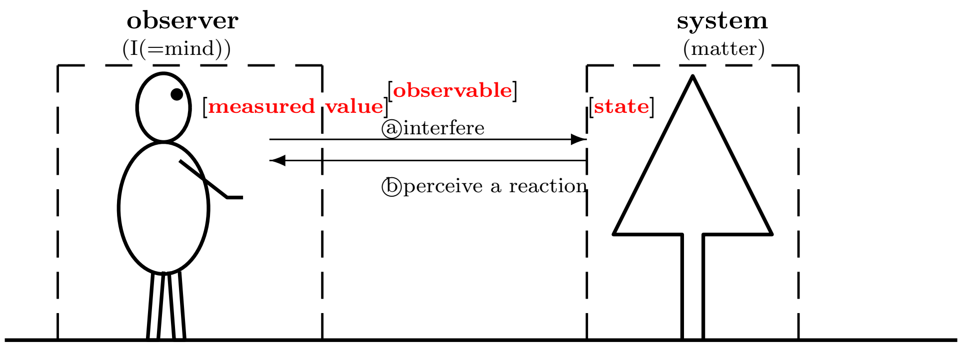
Therefore, we get

$$[\mathsf{O}_{12}^{(1)}; \Xi_1^c] \underset{\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}_{12}, S_{[\rho]})}{\Longleftarrow} [\mathsf{O}_{12}^{(2)}; \Xi_2^c]$$

□

8.3 Cogito— I think, therefore I am—

Recall the following figure.



[Descartes Figure 8.2 (=Figure 3.1)]:The image of "measurement(=ⓐ+ⓑ)" in dualism

The following example may be rather unnatural, but this is indispensable for the well-understanding of dualism.

Example 8.8. [Brain death(*cf.* ref. p.89 in [38])] Consider the classical basic structure

$$[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))]$$

Let ω_n $(\in \Omega = \{\omega_1, \omega_2, \ldots, \omega_N\})$ be the state of Peter. Let $\mathsf{O}_{12} = (X_1 \times X_2,\ 2^{X_1 \times X_2},\ F_{12}{=}F_1 \overset{\text{qp}}{\times} F_2)$ be the brain death observable in $L^\infty(\Omega)$ such that $X_1 = \{T, \overline{T}\}$ $X_2 = \{L, \overline{L}\}$, where $T =$ "think", $\overline{T} =$ "not think", $L =$ "live", $\overline{L} =$ "not live". For each ω_n $(n = 1, 2, \ldots, N)$, O_{12} satisfies the condition in Table 8.2.

[Table 8.2]: Brain death observable $\mathsf{O}_{12} = (X_1 \times X_2,\ 2^{X_1 \times X_2},\ F_{12})$

$F_1 \diagdown F_2$	$[F_2(\{\mathrm{L}\})](\omega_n)$	$[F_2(\{\overline{\mathrm{L}}\})](\omega_n)$
$[F_1(\{\mathrm{T}\})](\omega_n)$	$(1+(-1)^n)/2$ $(=[F_{12}(\{\mathrm{T}\}\times\{\mathrm{L}\})](\omega_n))$	0 $(=[F_{12}(\{\mathrm{T}\}\times\{\overline{\mathrm{L}}\})](\omega_n))$
$[F_1(\{\overline{\mathrm{T}}\})](\omega_n)$	0 $(=[F_{12}(\{\overline{\mathrm{T}}\}\times\{\mathrm{L}\})](\omega_n))$	$(1-(-1)^n)/2$ $(=[F_{12}(\{\overline{\mathrm{T}}\}\times\{\overline{\mathrm{L}}\})](\omega_n))$

Since $[F_{12}(\{\mathrm{T}\} \times \{\overline{\mathrm{L}}\})](\omega_n) = 0$, the following formula holds:

$$[\mathsf{O}_{12}^{(1)}; \{\mathrm{T}\}] \underset{\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{12}, S_{[\omega_n]})}{\Longrightarrow} [\mathsf{O}_{12}^{(2)}; \{\mathrm{L}\}]$$

Of course, this implies that

(A$_1$) Peter thinks, therefore, Peter lives.

This is the same as the statement concerning brain death. Note that in the above example, we see that

$$\text{observer} \longleftrightarrow \text{doctor}, \qquad \text{system} \longleftrightarrow \text{Peter},$$

The above (A_1) should not be confused with the following famous Descartes' saying (= cogito proposition):

(A_2) *"I think, therefore I am".*

in which the following identification may be assumed:

$$\text{observer} \longleftrightarrow \text{I}, \qquad \text{system} \longleftrightarrow \text{I}$$

And thus, the above is not a statement in dualism (=measurement theory). In order to propose Figure 8.2 (i.e., dualism) (that is, in order to establish the concept "I" in science), he started from the ambiguous statement "I think, therefore I am". Summing up, we want to say the following irony:

(B) Descartes proposed the dualism (i.e., Figure 8.2) by the cogito proposition (A_2) which is not understandable in dualism.

♠**Note 8.1.** It is not true to consider that every phenomena can be describe in terns of quantum language. Although readers may think that the following can be described in measurement theory, but we believe that it is impossible. For example, the followings can not be written by quantum language:

① : tense—past, present, future — ② : Heidegger's saying "In-der-Welt-sein"
③ : the measurement of a measurement, ④ : Bergson's subjective time
⑤ : observer's space-time,
⑥ : Only the present exists (due to Augustinus(354-430))

If we want to understand the above words, we have to propose the other scientific languages (except quantum language). We have to recall Wittgenstein's sayings

The limits of my language mean the limits of my world

8.4 Combined observable —Only one measurement is permitted —

8.4.1 Combined observable — only one observable

The linguistic interpretation says that

"Only one measurement is permitted"

$\Rightarrow$ "only one observable" $\Rightarrow$ "the necessity of the combined observable"

Thus, we prepare the following theorem.

Theorem 8.9. [The existence theorem of classical combined observable(*cf.*refs.[25, 29])] Consider the classical basic structure

$$[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))]$$

And consider observables O_{12}=$(X_1 \times X_2, \mathcal{F}_1 \boxtimes \mathcal{F}_2, F_{12})$ and O_{23}= $(X_2 \times X_3,\ \mathcal{F}_2 \boxtimes \mathcal{F}_3, F_{23})$ in $L^\infty(\Omega, \nu)$. Here, for simplicity, assume that X_i=$\{x_i^1, x_i^2, \ldots, x_i^{n_i}\}$ $(i = 1, 2, 3)$ is finite, Also, assume that $\mathcal{F}_i = 2^{X_i}$. Further assume that

$$\mathsf{O}_{12}^{(2)} = \mathsf{O}_{23}^{(2)} \quad (\text{That is,} \quad F_{12}(X_1 \times \Xi_2) = F_{23}(\Xi_2 \times X_3) \quad (\forall \Xi_2 \in 2^{X_2}))$$

Then, we have the observable O_{123}=$(X_1 \times X_2 \times X_3, \mathcal{F}_1 \times \mathcal{F}_2 \times \mathcal{F}_3, F_{123})$ in $L^\infty(\Omega)$ such that

$$\mathsf{O}_{123}^{(12)} = \mathsf{O}_{12}, \quad \mathsf{O}_{123}^{(23)} = \mathsf{O}_{23}$$

That is,

$$F_{123}^{(12)}(\Xi_1 \times \Xi_2 \times X_3) = F_{12}(\Xi_1 \times \Xi_2), \ \ F_{123}^{(23)}(X_1 \times \Xi_2 \times \Xi_3) = F_{23}(\Xi_2 \times \Xi_3) \tag{8.8}$$
$$(\forall \Xi_1 \in \mathcal{F}_1, \forall \Xi_2 \in \mathcal{F}_2, \forall \Xi_3 \in \mathcal{F}_3))$$

The O_{123} is called the combined observable of O_{12} and O_{23}.

Proof. $\mathsf{O}_{123} = (X_1 \times X_2 \times X_3,\ \mathcal{F}_1 \times \mathcal{F}_2 \times \mathcal{F}_3,\ F_{123})$ is, for example, defined by

$$[F_{123}(\{(x_1, x_2, x_3)\})](\omega) = \begin{cases} \dfrac{[F_{12}(\{(x_1, x_2)\})](\omega) \cdot [F_{23}(\{(x_2, x_3)\})](\omega)}{[F_{12}(X_1 \times \{x_2\})](\omega)} \\ \qquad\qquad\qquad\qquad ([F_{12}(X_1 \times \{x_2\})](\omega) \ \neq 0 \text{ and }) \\ 0 \\ \qquad\qquad\qquad\qquad ([F_{12}(X_1 \times \{x_2\})](\omega) = 0 \text{ and }) \end{cases}$$

$$(\forall \omega \in \Omega, \forall (x_1, x_2, x_3) \in X_1 \times X_2 \times X_3)$$

This clearly satisfies (8.8). □

Counter example 8.10. [Counter example in quantum systems] Theorem 8.9 does not hold in the quantum basic structure

$$[\mathcal{C}(H) \subseteq B(H) \subseteq B(H)]$$

For example, put $H = \mathbb{C}^n$, and consider the three Hermitian $(n \times n)$-matrices T_1, T_2, T_3 in $B(H)$ such that

$$T_1 T_2 = T_2 T_1, \quad T_2 T_3 = T_3 T_2, \quad T_1 T_3 \neq T_3 T_1 \tag{8.9}$$

For each $k = 1, 2, 3$, define the spectrum decomposition $\mathsf{O}_k = (X_k, \mathcal{F}_k, F_k)$ in H (which is regarded as a projective observable) such that

$$T_k = \int_{X_k} x_k F_k(dx_k) \tag{8.10}$$

where $X_k = \mathbb{R}, \mathcal{F}_k = \mathcal{B}_{\mathbb{R}}$.

From the commutativity, we have the simultaneous observables

$$\mathsf{O}_{12} {=} \mathsf{O}_1 \times \mathsf{O}_2 = (X_1 \times X_2, \mathcal{F}_1 \boxtimes \mathcal{F}_2, F_{12} = F_1 \times F_2)$$

and

$$\mathsf{O}_{23} {=} \mathsf{O}_2 \times \mathsf{O}_3 = (X_2 \times X_3, \mathcal{F}_2 \boxtimes \mathcal{F}_3, F_{23} = F_2 \times F_3)$$

It is clear that

$$\mathsf{O}_{12}^{(2)} = \mathsf{O}_{23}^{(2)} \quad (\text{that is, } \; F_{12}(X_1 \times \Xi_2) = F_2(\Xi_2) = F_{23}(\Xi_2 \times X_3) \quad (\forall \Xi_2 \in \mathcal{F}_2))$$

However, it should be noted that there does not exist the observable $\mathsf{O}_{123}{=}(X_1 \times X_2 \times X_3, \mathcal{F}_1 \boxtimes \mathcal{F}_2 \boxtimes \mathcal{F}_3, F_{123})$ in $B(H)$ such that

$$\mathsf{O}_{123}^{(12)} = \mathsf{O}_{12}, \quad \mathsf{O}_{123}^{(23)} = \mathsf{O}_{23}$$

That is because, if O_{123} exists, Theorem 8.3 says that O_1 and O_3 commute, and it is in contradiction with the (8.9). Therefore, the **combined observable** O_{123} of O_{12} and O_{23} does not exist.

8.4.2 Combined observable and Bell's inequality

Now we consider the following problem:

Problem 8.11. [combined observable and Bell's inequality (*cf.* [38])] Consider the basic structure

$$[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]$$

Put $X_1 = X_2 = X_3 = X_4 = \{-1, 1\}$. Let O_{13}=$(X_1 \times X_3, 2^{X_1} \times 2^{X_3}, F_{13})$, O_{14}=$(X_1 \times X_4, 2^{X_1} \times 2^{X_4}, F_{14})$, O_{23}= $(X_2 \times X_3,\ 2^{X_2} \times 2^{X_3}, F_{23})$ and O_{24}= $(X_2 \times X_3,\ 2^{X_2} \times 2^{X_4}, F_{24})$ be observables in $L^\infty(\Omega)$ such that

$$\mathsf{O}_{13}^{(1)} = \mathsf{O}_{14}^{(1)}, \quad \mathsf{O}_{23}^{(2)} = \mathsf{O}_{24}^{(2)}, \quad \mathsf{O}_{13}^{(3)} = \mathsf{O}_{23}^{(3)}, \quad \mathsf{O}_{14}^{(4)} = \mathsf{O}_{24}^{(4)}$$

Define the probability measure ν_{ab} on $\{-1,1\}^2$ by the formula (4.48). Assume that there exists a state $\rho_0 \in \mathfrak{S}^p(\mathcal{A}^*)$ such that

$$
\begin{aligned}
{}_{\mathcal{A}^*}\big(\rho_0, F_{13}(\{(x_1,x_3)\})\big)_{\overline{\mathcal{A}}} &= \nu_{a^1b^1}(\{(x_1,x_3)\}, \\
{}_{\mathcal{A}^*}\big(\rho_0, F_{14}(\{(x_1,x_4)\})\big)_{\overline{\mathcal{A}}} &= \nu_{a^1b^2}(\{(x_1,x_4)\} \\
{}_{\mathcal{A}^*}\big(\rho_0, F_{23}(\{(x_2,x_3)\})\big)_{\overline{\mathcal{A}}} &= \nu_{a^2b^1}(\{(x_2,x_3)\}, \\
{}_{\mathcal{A}^*}\big(\rho_0, F_{24}(\{(x_2,x_4)\})\big)_{\overline{\mathcal{A}}} &= \nu_{a^2b^2}(\{(x_2,x_4)\}
\end{aligned}
$$

Now we have the following problem:

(a) Does the observable O_{1234}=$(\times_{k=1}^4 X_k, \times_{k=1}^4 \mathcal{F}_k, F_{1234})$ in $\overline{\mathcal{A}}$ satisfying the following (♯) exist?

(♯) $\mathsf{O}_{1234}^{(13)} = \mathsf{O}_{13}, \quad \mathsf{O}_{1234}^{(14)} = \mathsf{O}_{14}, \quad \mathsf{O}_{1234}^{(23)} = \mathsf{O}_{23}, \quad \mathsf{O}_{1234}^{(24)} = \mathsf{O}_{24}$

In what follows, we show that the above observable O_{1234} does not exist.

Assume that the observable O_{1234}=$(\times_{k=1}^4 X_k,\ \times_{k=1}^4 \mathcal{F}_k,\ F_{1234})$ exists. Then, it suffices to show the contradiction. Define $C_{13}(\rho_0)$, $C_{14}(\rho_0)$, $C_{23}(\rho_0)$ and $C_{24}(\rho_0)$ such that

$$
\begin{aligned}
C_{13}(\rho_0) &= \int_{\times_{k=1}^4 X_k} x_1 \cdot x_3 \ {}_{\mathcal{A}^*}\big(\rho_0, F_{1234}(\mathop{\times}_{k=1}^{4} dx_k)\big)_{\overline{\mathcal{A}}} \quad \Big(= \int_{X_1 \times X_3} x_1 \cdot x_3 \ \nu_{a^1b^1}(dx_1 dx_3)\Big) \\
C_{14}(\rho_0) &= \int_{\times_{k=1}^4 X_k} x_1 \cdot x_4 \ {}_{\mathcal{A}^*}\big(\rho_0, F_{1234}(\mathop{\times}_{k=1}^{4} dx_k)\big)_{\overline{\mathcal{A}}} \quad \Big(= \int_{X_1 \times X_4} x_1 \cdot x_4 \ \nu_{a^1b^2}(dx_1 dx_4)\Big) \\
C_{23}(\rho_0) &= \int_{\times_{k=1}^4 X_k} x_2 \cdot x_3 \ {}_{\mathcal{A}^*}\big(\rho_0, F_{1234}(\mathop{\times}_{k=1}^{4} dx_k)\big)_{\overline{\mathcal{A}}} \quad \Big(= \int_{X_2 \times X_3} x_2 \cdot x_3 \ \nu_{a^2b^1}(dx_2 dx_3)\Big) \\
C_{24}(\rho_0) &= \int_{\times_{k=1}^4 X_k} x_2 \cdot x_4 \ {}_{\mathcal{A}^*}\big(\rho_0, F_{1234}(\mathop{\times}_{k=1}^{4} dx_k)\big)_{\overline{\mathcal{A}}} \quad \Big(= \int_{X_2 \times X_4} x_2 \cdot x_4 \ \nu_{a^2b^2}(dx_2 dx_4)\Big)
\end{aligned}
$$

Then, we can easily get the following Bell's inequality: (cf. Bell's inequality 4.17).

$$
\begin{aligned}
&|C_{13}(\rho_0) - C_{14}(\rho_0)| + |C_{23}(\rho_0) + C_{24}(\rho_0)| \\
\leqq &\int_{\times_{k=1}^4 X_k} |x_1| \cdot |x_3 - x_4| \ \ + |x_2| \cdot |x_3 + x_4| \big[F_{1234}(\mathop{\times}_{k=1}^{4} dx_k)\big](\rho_0) \\
\leqq &\, 2 \qquad (\text{since } x_k \in \{-1, 1\})
\end{aligned}
\tag{8.11}
$$

However, the formula (4.50) says that this (8.11) must be $2\sqrt{2}$. Thus, by contradiction, we says that O_{1234} satisfying (a) does not exist. Thus we can not take a measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}_{1234}, S_{[\rho_0]})$.

However, it should be noted that

(b) instead of $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}_{1234}, S_{[\rho_0]})$. we can take a parallel measurement $\mathsf{M}_{\otimes_{k=1}^4 \overline{\mathcal{A}}}(\mathsf{O}_{13} \otimes \mathsf{O}_{14} \otimes \mathsf{O}_{23} \otimes \mathsf{O}_{24},\ S_{[\otimes_{k=1}^4 \rho_0]})$. In this case, we easily see that $(8.11) = 2\sqrt{2}$ as the formula (4.50).

That is,

(c) in the case of a parallel measurement, Bell's inequality is broken in both quantum and classical systems.

♠**Note 8.2.** In the above argument, Bell's inequality is used in the framework of measurement theory. This is of course true. Also I can guess that

($\sharp$) Problem 8.11 is related to the theory of "hidden variables"?

8.5 Syllogism and EPR-paradox — Does Socrates die?

8.5.1 Syllogism and its variations

Next, we shall discuss practical syllogism (i.e., measurement theoretical theorem concerning implication (Definition8.6)). Before the discussion, we note that

(♯) Since Theorem8.9 (The existence of the combined observable) does not hold in quantum system, (*cf.* Counter Example8.10), syllogism does not hold.

On the other hand, in classical system, we can expect that syllogism holds. This will be proved in the following theorem.

Theorem 8.12. [Practical syllogism in classical systems] Consider the classical basic structure

$$[C_0(\Omega) \subseteq L^\infty(\Omega,\nu) \subseteq B(L^2(\Omega,\nu))]$$

Let $\mathsf{O}_{123} = (X_1 \times X_2 \times X_3,\ \mathcal{F}_1 \times \mathcal{F}_2 \times \mathcal{F}_3,\ F_{123}{=}\overset{\mathrm{qp}}{\times}_{k=1,2,3}F_k)$ be an observable in $L^\infty(\Omega)$ Fix $\omega \in \Omega$, $\Xi_1 \in \mathcal{F}_1$, $\Xi_2 \in \mathcal{F}_2$, $\Xi_3 \in \mathcal{F}_3$ Then, we see the following (i) – (iii).

(i).(practical syllogism)

$$[\mathsf{O}_{123}^{(1)};\Xi_1] \underset{\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{123},S_{[\omega]})}{\Longrightarrow} [\mathsf{O}_{123}^{(2)};\Xi_2], \quad [\mathsf{O}_{123}^{(2)};\Xi_2] \underset{\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{123},S_{[\omega]})}{\Longrightarrow} [\mathsf{O}_{123}^{(3)};\Xi_3]$$

implies

$$\mathrm{Rep}_\omega^{\Xi_1\times\Xi_3}[\mathsf{O}_{123}^{(13)}] = \begin{bmatrix} [F_{123}^{(13)}(\Xi_1 \times \Xi_3)](\omega) & [F_{123}^{(13)}(\Xi_1 \times \Xi_3^c)](\omega) \\ [F_{123}^{(13)}(\Xi_1^c \times \Xi_3)](\omega) & [F_{123}^{(13)}(\Xi_1^c \times \Xi_3^c)](\omega) \end{bmatrix}$$

$$= \begin{bmatrix} [F_{123}^{(1)}(\Xi_1)](\omega) & 0 \\ [F_{123}^{(3)}(\Xi_3)](\omega) - [F_{123}^{(1)}(\Xi_1)](\omega) & 1 - [F_{123}^{(3)}(\Xi_3)](\omega) \end{bmatrix}$$

That is, it holds:

$$[\mathsf{O}_{123}^{(1)};\Xi_1] \underset{\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{123},S_{[\omega]})}{\Longrightarrow} [\mathsf{O}_{123}^{(3)};\Xi_3] \tag{8.12}$$

(ii).

$$[\mathsf{O}_{123}^{(1)};\Xi_1] \underset{\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{123},S_{[\omega]})}{\Longleftarrow} [\mathsf{O}_{123}^{(2)};\Xi_2], \quad [\mathsf{O}_{123}^{(2)};\Xi_2] \underset{\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{123},S_{[\omega]})}{\Longrightarrow} [\mathsf{O}_{123}^{(3)};\Xi_3]$$

implies

$$\mathrm{Rep}_\omega^{\Xi_1\times\Xi_3}[\mathsf{O}_{123}^{(13)}] = \begin{bmatrix} [F_{123}^{(13)}(\Xi_1 \times \Xi_3)](\omega) & [F_{123}^{(13)}(\Xi_1 \times \Xi_3^c)](\omega) \\ [F_{123}^{(13)}(\Xi_1^c \times \Xi_3)](\omega) & [F_{123}^{(13)}(\Xi_1^c \times \Xi_3^c)](\omega) \end{bmatrix}$$

$$= \begin{bmatrix} \alpha_{\Xi_1 \times \Xi_3} & [F_{123}^{(1)}(\Xi_1)](\omega) - \alpha_{\Xi_1 \times \Xi_3} \\ [F_{123}^{(3)}(\Xi_3)](\omega) - \alpha_{\Xi_1 \times \Xi_3} & 1 - \alpha_{\Xi_1 \times \Xi_3} - [F_{123}^{(1)}(\Xi_1)] - [F_{123}^{(3)}(\Xi_3)] \end{bmatrix}$$

where

$$\max\{[F_{123}^{(2)}(\Xi_2)](\omega), [F_{123}^{(1)}(\Xi_1)](\omega) + [F_{123}^{(3)}(\Xi_3)](\omega) - 1\} \\ \leqq \alpha_{\Xi_1 \times \Xi_3}(\omega) \leqq \min\{[F_{123}^{(1)}(\Xi_1)](\omega), [F_{123}^{(3)}(\Xi_3)](\omega)\} \tag{8.13}$$

(iii).

$$[\mathsf{O}_{123}^{(1)}; \Xi_1] \underset{\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{123}, S_{[\omega]})}{\Longrightarrow} [\mathsf{O}_{123}^{(2)}; \Xi_2], \quad [\mathsf{O}_{123}^{(2)}; \Xi_2] \underset{\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{123}, S_{[\omega]})}{\Longleftarrow} [\mathsf{O}_{123}^{(3)}; \Xi_3]$$

implies

$$\mathrm{Rep}_\omega^{\Xi_1 \times \Xi_3}[\mathsf{O}_{123}^{(13)}] = \begin{bmatrix} [F_{123}^{(13)}(\Xi_1 \times \Xi_3)](\omega) & [F_{123}^{(13)}(\Xi_1 \times \Xi_3^c)](\omega) \\ [F_{123}^{(13)}(\Xi_1^c \times \Xi_3)](\omega) & [F_{123}^{(13)}(\Xi_1^c \times \Xi_3^c)](\omega) \end{bmatrix}$$

$$= \begin{bmatrix} \alpha_{\Xi_1 \times \Xi_3}(\omega) & [F_{123}^{(1)}(\Xi_1)](\omega) - \alpha_{\Xi_1 \times \Xi_3}(\omega) \\ [F_{123}^{(3)}(\Xi_3)](\omega) - \alpha_{\Xi_1 \times \Xi_3}(\omega) & 1 - \alpha_{\Xi_1 \times \Xi_3}(\omega) - [F_{123}^{(1)}(\Xi_1)](\omega) - [F_{123}^{(3)}(\Xi_3)](\omega) \end{bmatrix}$$

where

$$\max\{0, [F_{123}^{(1)}(\Xi_1)](\omega) + [F_{123}^{(3)}(\Xi_3)](\omega) - [F_{123}^{(2)}(\Xi_2)](\omega)\} \\ \leqq \alpha_{\Xi_1 \times \Xi_3}(\omega) \leqq \min\{[F_{123}^{(1)}(\Xi_1)](\omega), [F_{123}^{(3)}(\Xi_3)](\omega)\}$$

Proof. (i): By the condition, we see

$$0 = [F_{123}^{(12)}(\Xi_1 \times \Xi_2^c)](\omega) = [F_{123}(\Xi_1 \times \Xi_2^c \times \Xi_3)](\omega) + [F_{123}(\Xi_1 \times \Xi_2^c \times \Xi_3^c)](\omega)$$
$$0 = [F_{123}^{(23)}(\Xi_2 \times \Xi_3^c)](\omega) = [F_{123}(\Xi_1 \times \Xi_2 \times \Xi_3^c)](\omega) + [F_{123}(\Xi_1^c \times \Xi_2 \times \Xi_3^c)](\omega)$$

Therefore,

$$0 = [F_{123}(\Xi_1 \times \Xi_2^c \times \Xi_3)](\omega) = [F_{123}(\Xi_1 \times \Xi_2^c \times \Xi_3^c)](\omega)$$
$$0 = [F_{123}(\Xi_1 \times \Xi_2 \times \Xi_3^c)](\omega) = [F_{123}(\Xi_1^c \times \Xi_2 \times \Xi_3^c)](\omega)$$

Hence,

$$[F_{123}^{(13)}(\Xi_1 \times \Xi_3^c)](\omega) = [F_{123}(\Xi_1 \times \Xi_2 \times \Xi_3^c)](\omega) + [F_{123}^{(13)}(\Xi_1 \times \Xi_2^c \times \Xi_3^c)](\omega) = 0$$

Thus, we get, (8.12).
For the proof of (ii) and (iii), see refs. [25, 29]. □

Example 8.13. [Continued from Example 8.5] Let $\mathsf{O}_1 = \mathsf{O}_{\rm SW} = (X_{\rm SW},\, 2^{X_{\rm SW}},\, F_{\rm SW})$ and $\mathsf{O}_3 = \mathsf{O}_{\rm RD} = (X_{\rm RD},\, 2^{X_{\rm RD}},\, F_{\rm RD})$ be as in Example 8.5. Putting $X_{\rm RP} = \{y_{\rm RP}, n_{\rm RP}\}$, consider the new observable $\mathsf{O}_2 = \mathsf{O}_{\rm RP} = (X_{\rm RP}, 2^{X_{\rm RP}}, F_{\rm RP})$. Here, "$y_{\rm RP}$" and "$n_{\rm RP}$" respectively means "ripe" and "not ripe". Put

$$
\begin{aligned}
{\rm Rep}[\mathsf{O}_1] &= \big[[F_{\rm SW}(\{y_{\rm SW}\})](\omega_k), [F_{\rm SW}(\{n_{\rm SW}\})](\omega_k)\big]\\
{\rm Rep}[\mathsf{O}_2] &= \big[[F_{\rm RP}(\{y_{\rm RP}\})](\omega_k), [F_{\rm RP}(\{n_{\rm RP}\})](\omega_k)\big]\\
{\rm Rep}[\mathsf{O}_3] &= \big[[F_{\rm RD}(\{y_{\rm RD}\})](\omega_k), [F_{\rm RD}(\{n_{\rm RD}\})](\omega_k)\big]
\end{aligned}
$$

Consider the following quasi-product observable:

$$
\begin{aligned}
\mathsf{O}_{12} &= (X_{\rm SW} \times X_{\rm RP}, 2^{X_{\rm SW}\times X_{\rm RP}}, F_{12}{=}F_{\rm SW}\overset{\rm qp}{\times} F_{\rm RP})\\
\mathsf{O}_{23} &= (X_{\rm RP} \times X_{\rm RD}, 2^{X_{\rm RP}\times X_{\rm RD}}, F_{23}{=}F_{\rm RP}\overset{\rm qp}{\times} F_{\rm RD})
\end{aligned}
$$

Let $\omega_k \in \Omega$. And assume that

$$
\begin{aligned}
&[\mathsf{O}_{123}^{(1)}; \{y_{\rm SW}\}] \underset{\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{123}, S_{[\omega_k]})}{\Longrightarrow} [\mathsf{O}_{123}^{(2)}; \{y_{\rm RP}\}],\\
&[\mathsf{O}_{123}^{(2)}; \{y_{\rm RP}\}] \underset{\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{123}, S_{[\omega_k]})}{\Longrightarrow} [\mathsf{O}_{123}^{(3)}; \{y_{\rm RD}\}]
\end{aligned}
\tag{8.14}
$$

Then, by Theorem 8.12(i), we get:

$$
\begin{aligned}
{\rm Rep}[\mathsf{O}_{13}] &= \begin{bmatrix} [F_{13}(\{y_{\rm SW}\}\times\{y_{\rm RD}\})](\omega_k) & [F_{13}(\{y_{\rm SW}\}\times\{n_{\rm RD}\})](\omega_k) \\ [F_{13}(\{n_{\rm SW}\}\times\{y_{\rm RD}\})](\omega_k) & [F_{13}(\{n_{\rm SW}\}\times\{n_{\rm RD}\})](\omega_k) \end{bmatrix}\\
&= \begin{bmatrix} [F_{\rm SW}(\{y_{\rm SW}\})](\omega_k) & 0 \\ [F_{\rm RD}(\{y_{\rm RD}\})](\omega_k) - [F_{\rm SW}(\{y_{\rm SW}\})](\omega_k) & 1-[F_{\rm RD}(\{y_{\rm RD}\})](\omega_k) \end{bmatrix}
\end{aligned}
$$

Therefore, when we know that the tomato ω_k is sweet by measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{123}, S_{[\omega_k]})$, the probability that ω_k is red is given by

$$
\frac{[F_{13}(\{y_{\rm SW}\}\times\{y_{\rm RD}\})](\omega_k)}{[F_{13}(\{y_{\rm SW}\}\times\{y_{\rm RD}\})](\omega_k) + [F_{13}(\{y_{\rm SW}\}\times\{n_{\rm RD}\})](\omega_k)} = \frac{[F_{\rm RD}(\{y_{\rm RD}\})](\omega_k)}{[F_{\rm RD}(\{y_{\rm RD}\})](\omega_k)} = 1 \tag{8.15}
$$

Of course, (8.14) means

"Sweet" $\Longrightarrow$ "Ripe" "Ripe" $\Longrightarrow$ "Red"

Therefore, by (8.12), we get the following conclusion.

"Sweet" $\Longrightarrow$ "Red"

However, it is not useful in the market. What we want to know is such as

"Red" $\Longrightarrow$ "Sweet"

This will be discussed in the following example.

Example 8.14. [Continued from Example 8.5] Instead of (8.14), assume that

$$\mathsf{O}_1^{\{y_1\}} \underset{\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{12}, S_{[\delta_{\omega_n}]})}{\Longleftarrow} \mathsf{O}_2^{\{y_2\}}, \qquad \mathsf{O}_2^{\{y_2\}} \underset{\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{23}, S_{[\delta_{\omega_n}]})}{\Longrightarrow} \mathsf{O}_3^{\{y_3\}}. \tag{8.16}$$

When we observe that the tomato ω_n is "RED", we can infer, by the fuzzy inference $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{13}, S_{[\delta_{\omega_n}]})$, the probability that the tomato ω_n is "SWEET" is given by

$$Q = \frac{[F_{13}(\{y_{\mathrm{SW}}\} \times \{y_{\mathrm{RD}}\})](\omega_n)}{[F_{13}(\{y_{\mathrm{SW}}\} \times \{y_{\mathrm{RD}}\})](\omega_n) + [F_{13}(\{n_{\mathrm{SW}}\} \times \{y_{\mathrm{RD}}\})](\omega_n)}$$

which is, by (8.3), estimated as follows:

$$\max\left\{\frac{[F_{\mathrm{RP}}(\{y_{\mathrm{RP}}\})](\omega_n)}{[F_{\mathrm{RD}}(\{y_{\mathrm{RD}}\})](\omega_n)}, \frac{[F_{\mathrm{SW}}(\{y_{\mathrm{SW}}\})] + [F_{\mathrm{RD}}(\{y_{\mathrm{RD}}\})] - 1}{[F_{\mathrm{RD}}(\{y_{\mathrm{RD}}\})](\omega_n)}\right\} \le Q \le \min\{\frac{[F_{\mathrm{SW}}(\{y_{\mathrm{SW}}\})](\omega_n)}{[F_{\mathrm{RD}}(\{y_{\mathrm{RD}}\})](\omega_n)}, 1\}. \tag{8.17}$$

Note that (8.16) implies (and is implied by)

$$\text{“RIPE”} \Longrightarrow \text{“SWEET”} \qquad \text{and} \qquad \text{“RIPE”} \Longrightarrow \text{“RED”} .$$

And note that the conclusion (8.17) is somewhat like

$$\text{“RED”} \Longrightarrow \text{“SWEET”} .$$

Therefore, this conclusion is peculiar to "fuzziness".

///

Remark 8.15. [Syllogism does not hold in quantum system (*cf.* ref. [35])]

Concerning EPR's paper[13], we shall add some remark as follows. Let A and B be particles with the same masses m. Consider the situation described in the following figure:

Figure 8.3: The case that "the velocity of A"$= -$"the velocity of B."

The position q_A (at time t_0) of the particle A can be exactly measured, and moreover, the velocity of v_B (at time t_0) of the particle B can be exactly measured. Thus, we may conclude that

(A) the position and momentum (at time t_0) of the particle A are respectively and exactly equal to q_A and $-mv_B$?

(As mentioned in Section 4.4.3, this is not in contradiction with Heisenberg' uncertainty principle).

However, we have the following question:

Is the conclusion (A) true?

Now we shall describe the above arguments in quantum system:

A quantum two particles system S is formulated in a tensor Hilbert space $H = H_1 \otimes H_1 = L^2(\mathbb{R}_{q_1}) \otimes L^2(\mathbb{R}_{q_2}) = L^2(\mathbb{R}^2_{(q_1,q_2)})$. The state u_0 ($\in H = H_1 \otimes H_1 = L^2(\mathbb{R}^2_{(q_1,q_2)})$) $\Big($ or precisely, $\rho_0 = |u_0\rangle\langle u_0|\Big)$ of the system S is assumed to be

$$u_0(q_1, q_2) = \sqrt{\frac{1}{2\pi\epsilon\sigma}} e^{-\frac{1}{8\sigma^2}(q_1-q_2-2a)^2 - \frac{1}{8\epsilon^2}(q_1+q_2)^2} \tag{8.18}$$

where a positive number ϵ is sufficiently small. For each $k = 1, 2$, define the self-adjoint operators $Q_k\colon L^2(\mathbb{R}^2_{(q_1,q_2)}) \to L^2(\mathbb{R}^2_{(q_1,q_2)})$ and $P_k\colon L^2(\mathbb{R}^2_{(q_1,q_2)}) \to L^2(\mathbb{R}^2_{(q_1,q_2)})$ by

$$\begin{aligned} Q_1 &= q_1, \qquad P_1 = \frac{\hbar\partial}{i\partial q_1} \\ Q_2 &= q_2, \qquad P_2 = \frac{\hbar\partial}{i\partial q_2} \end{aligned} \tag{8.19}$$

($\sharp_1$) Let $\mathsf{O}_1 = (\mathbb{R}^3, \mathcal{B}_{\mathbb{R}^3}, F_1)$ be the observable representation of the self-adjoint operator $(Q_1 \otimes P_2) \times (I \otimes P_2)$. And consider the measurement $\mathsf{M}_{B(H)}(\mathsf{O}_1 = (\mathbb{R}^3, \mathcal{B}_{\mathbb{R}^3}, F_1), S_{[|u_0\rangle\langle u_0|]})$. Assume that the measured value $(x_1, p_2, p_2) (\in \mathbb{R}^3)$. That is,

$$\underset{\text{(the position of } A_1\text{, the momentum of } A_2)}{(x_1, p_2)} \underset{\mathsf{M}_{B(H)}(\mathsf{O}_1, S_{[\rho_0]})}{\Longrightarrow} \underset{\text{the momentum of } A_2}{p_2}$$

($\sharp_2$) Let $\mathsf{O}_2 = (\mathbb{R}^2, \mathcal{B}_{\mathbb{R}^2}, F_2)$ be the observable representation of $(I \otimes P_2) \times (P_1 \otimes I)$. And consider the measurement $\mathsf{M}_{B(H)}(\mathsf{O}_2 = (\mathbb{R}^2, \mathcal{B}_{\mathbb{R}^2}, F_2), S_{[|u_0\rangle\langle u_0|]})$. Assume that the measured value $(p_2, -p_2) (\in \mathbb{R}^3)$. That is,

$$\underset{\text{the momentum of } A_2}{p_2} \underset{\mathsf{M}_{B(H)}(\mathsf{O}_2, S_{[\rho_0]})}{\Longrightarrow} \underset{\text{the momentum of } A_1}{-p_2}$$

($\sharp_3$) Therefore, by ($\sharp_1$) and ($\sharp_2$), **"syllogism"** may say that

$$\underset{\text{the momentum of } A_1}{-p_2} \qquad \Big(\text{that is, the momentum of } A_1 \text{ is equal to } -p_2\Big)$$

Hence, some assert that

(B) **The (A) is true**

But, the above argument (particularly, "syllogism") is not true, thus,

The (A) is not true

That is because

($\sharp_4$) $(Q_1 \otimes P_2) \times (I \otimes P_2)$ and $(I \otimes P_2) \times (P_1 \otimes I)$ (Therefore, O_1 and O_2) do not commute, and thus, the simultaneous observable does not exist.
Thus, we can not test the ($\sharp_3$) experimentally.

Remark 8.16. After all, we think that EPR-paradox says the following two:

(C_1) syllogism does not necessarily hold in quantum systems,

(C_2) there is something faster than light.

We think that (C_1) is not serious. Thus, we do not need to investigate how to understand the fact (C_1). On the other hand, (C_2) is serious. Although we have to make efforts to understand the "fact (C_2)", this is the problem in physics (i.e., in ⑤ in Figure 1.1). Recall that the spirit of quantum language (i.e., in ⑩ in Figure 1.1) is

"Stop being bothered."

Chapter 9

Mixed measurement theory (⊃Bayesian statistics)

Quantum language (= measurement theory) is classified as follows.

$$
(\sharp) \;\; \underset{\text{(=quantum language)}}{\text{measurement theory}} \begin{cases} \underset{(\sharp_1)}{\text{pure type}} \begin{cases} \text{classical system : Fisher statistics} \\ \text{quantum system : usual quantum mechanics} \end{cases} \\ \\ \underset{(\sharp_2)}{\text{mixed type}} \begin{cases} \text{classical system : including Bayesian statistics, Kalman filter} \\ \text{quantum system : quantum decoherence} \end{cases} \end{cases}
$$

In this chapter, we study mixed measurement theory, which includes Bayesian statistics.

9.1 Mixed measurement theory(⊃Bayesian statistics)

9.1.1 Axiom$^{(m)}$ 1 (mixed measurement)

In the previous chapters, we studied Axiom 1 (pure measurement: §2.7), that is,

$$
\underset{\text{(=quantum language)}}{\boxed{\text{pure measurement theory}}} := \underbrace{\overset{\text{[(pure)Axiom 1]}}{\underset{(cf.\ \S 2.7)}{\boxed{\text{pure measurement}}}} + \overset{\text{[Axiom 2]}}{\underset{(cf.\ \S 10.3)}{\boxed{\text{Causality}}}}}_{\text{a kind of spells (a priori judgment)}} + \underbrace{\overset{\text{[quantum linguistic interpretation]}}{\underset{(cf.\ \S 3.1)}{\boxed{\text{Linguistic interpretation}}}}}_{\text{manual to use spells}} \tag{9.1}
$$

In this chapter, we shall study "Axiom$^{(m)}$ 1 (mixed measurement)" in mixed measurement theory, that is,

$$
\underset{\text{(=quantum language)}}{\boxed{\text{mixed measurement theory}}} := \underbrace{\overset{\text{[(mixed)Axiom}^{(m)}\text{ 1]}}{\underset{(cf.\ \S 9.1\ )}{\boxed{\text{mixed measurement}}}} + \overset{\text{[Axiom 2]}}{\underset{(cf.\ \S 10.3)}{\boxed{\text{Causality}}}}}_{\text{a kind of spells (a priori judgment)}} + \underbrace{\overset{\text{[quantum linguistic interpretation]}}{\underset{(cf.\ \S 3.1)}{\boxed{\text{Linguistic interpretation}}}}}_{\text{manual to use spells}} \tag{9.2}
$$

In the previous chapters, we mainly discussed pure measurements listed in Review 9.1, especially W^*-measurement (A$_1$).

Review 9.1. [=Preparation 2.30].

(A$_1$) W^*-measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}{=}\ (X, \mathcal{F}, F),\ S_{[\rho]})$, where $\mathsf{O}{=}\ (X, \mathcal{F}, F)$ is a W^*-observable in $\overline{\mathcal{A}}$, and $\rho(\in \mathfrak{S}^p(\mathcal{A}^*))$ is a pure state. Here, "W^*-measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O},\ S_{[\rho]})$" is also denoted by

"measurement$^{W^*}$ $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}.\ S_{[\rho]})$" , or "measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}.\ S_{[\rho]})$" ,

(A$_2$) C^*-measurement $\mathsf{M}_{\mathcal{A}}(\mathsf{O}{=}\ (X, \mathcal{F}, F),\ S_{[\rho]})$, where $\mathsf{O}{=}\ (X, \mathcal{F}, F)$ is a C^*-observable in $\mathcal{A}$, and $\rho(\in \mathfrak{S}^p(\mathcal{A}^*))$ is a pure state. Here, "C^*-measurement $\mathsf{M}_{\mathcal{A}}(\mathsf{O},\ S_{[\rho]})$" is also denoted by

"measurement$^{C^*}$ $\mathsf{M}_{\mathcal{A}}(\mathsf{O}.\ S_{[\rho]})$" , or "measurement $\mathsf{M}_{\mathcal{A}}(\mathsf{O}.\ S_{[\rho]})$" .

In this chapter, we introduce four "mixed measurements" as follows.

Preparation 9.2.

(B$_1$) W^*-mixed measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}{=}\ (X, \mathcal{F}, F),\ \overline{S}_{[*]}(w_0))$, where $\mathsf{O}{=}\ (X, \mathcal{F}, F)$ is a W^*-observable in $\overline{\mathcal{A}}$, and $w_0(\in \overline{\mathfrak{S}}^m(\overline{\mathcal{A}}_*))$ is a W^*-mixed state. Here, "W^*-mixed measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O},\ \overline{S}_{[*]}(w_0))$" is also denoted by

"W^*-mixed measurement$^{W^*}$ $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}.\ \overline{S}_{[*]}(w_0))$", or

"mixed measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}.\ \overline{S}_{[*]}(w_0))$"

(B$_2$) C^*-mixed measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}{=}\ (X, \mathcal{F}, F),\ S_{[*]}(\rho_0))$, where $\mathsf{O}{=}\ (X, \mathcal{F}, F)$ is a W^*-observable in $\overline{\mathcal{A}}$, and $\rho_0(\in \mathfrak{S}^m(\mathcal{A}^*))$ is a C^*-mixed state. Here, "C^*-mixed measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O},\ S_{[*]}(\rho_0))$" is also denoted by

"C^*-mixed measurement$^{W^*}$ $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}.\ S_{[*]}(\rho_0))$", or

"mixed measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}.\ S_{[*]}(\rho_0))$"

Although we mainly devote ourselves to the above two, we add the followings.

(B$_3$) W^*-mixed measurement $\mathsf{M}_{\mathcal{A}}(\mathsf{O}{=}\ (X, \mathcal{F}, F),\ \overline{S}_{[*]}(w_0))$, where $\mathsf{O}{=}\ (X, \mathcal{F}, F)$ is a C^*-observable in $\mathcal{A}$, and $w_0(\in \overline{\mathfrak{S}}^m(\overline{\mathcal{A}}_*))$ is a W^*-mixed state. Here, "W^*-mixed measurement $\mathsf{M}_{\mathcal{A}}(\mathsf{O},\ \overline{S}_{[*]}(w_0))$" is also denoted by

"W^*-mixed measurement$^{C^*}$ $\mathsf{M}_{\mathcal{A}}(\mathsf{O}.\ \overline{S}_{[*]}(w_0))$", or

"mixed measurement $\mathsf{M}_{\mathcal{A}}(\mathsf{O}.\ \overline{S}_{[*]}(w_0))$"

(B$_4$) C^*-mixed measurement $\mathsf{M}_{\mathcal{A}}\big(\mathsf{O}{=}\ (X, \mathcal{F}, F),\ S_{[*]}(\rho_0)\big)$, where $\mathsf{O}{=}\ (X, \mathcal{F}, F)$ is a C^*-observable in $\mathcal{A}$, and $\rho_0(\in \mathfrak{S}^m(\mathcal{A}^*))$ is a C^*-mixed state. Here, "C^*-mixed measurement $\mathsf{M}_{\mathcal{A}}\big(\mathsf{O},\ S_{[*]}(\rho_0)\big)$" is also denoted by

"C^*-mixed measurement$^{C^*}$ $\mathsf{M}_{\mathcal{A}}\big(\mathsf{O}.\ \overline{S}_{[*]}(\rho_0)\big)$", or

"mixed measurement $\mathsf{M}_{\mathcal{A}}\big(\mathsf{O}.S_{[*]}(\rho_0)\big)$"

We now give Axiom$^{(m)}$ 1 for mixed measurements. We will discuss (C$_1$) mainly, and (C$_2$) when necessary.

(C):Axiom$^{(m)}$ 1 (mixed measurement)

Let $\mathsf{O}{=}\ (X, \mathcal{F}, F)$ be an observable in $\overline{\mathcal{A}}$
(C$_1$): Let $w_0 \in \overline{\mathfrak{S}}^m(\overline{\mathcal{A}}_*)$. The probability that a measured value obtained by W^*-mixed measurement $\mathsf{M}_{\overline{\mathcal{A}}}\big(\mathsf{O}{=}\ (X, \mathcal{F}, F),\ S_{[*]}(w_0)\big)$ belongs to Ξ $(\in \mathcal{F})$ is given by

$$ {}_{\overline{\mathcal{A}}_*}(w_0, F(\Xi))_{\overline{\mathcal{A}}} \quad \Big(\equiv w_0(F(\Xi))\Big) $$

(C$_2$): Let $\rho_0 \in \mathfrak{S}^m(\mathcal{A}^*)$. The probability that a measured value obtained by C^*-mixed measurement $\mathsf{M}_{\overline{\mathcal{A}}}\big(\mathsf{O}{=}\ (X, \mathcal{F}, F),\ S_{[*]}(\rho_0)\big)$ belongs to Ξ $(\in \mathcal{F})$ is given by

$$ {}_{\mathcal{A}^*}(\rho_0, F(\Xi))_{\overline{\mathcal{A}}} \quad \Big(\equiv \rho(F(\Xi))\Big) $$

As we learned Axiom 1 by rote in pure measurement theory,

we have to learn Axiom$^{(m)}$ 1 by rote, and exercise a lot of examples

The practices will be done in this chapter.

Remark 9.3. In the above Axiom$^{(m)}$ 1, (C$_1$) and (C$_2$) are not so different.

($\sharp_1$) In the quantum case, (C$_1$)=(C$_2$) clearly holds, since $\mathfrak{S}^m(\mathcal{T}r(H)) = \overline{\mathfrak{S}}^m(\mathcal{T}r(H))$ in (2.17).

($\sharp_2$) In the classical case, we see

$$ L^1_{+1}(\Omega.\nu) \ni w_0 \xrightarrow{\rho_0(D)=\int_D w_0(\omega)\nu(d\omega)} \rho_0 \in \mathcal{M}_{+1}(\Omega) $$

Therefore, in this case, we consider that

$$ \mathsf{M}_{L^\infty(\Omega.\nu)}\big(\mathsf{O}{=}(X, \mathcal{F}, F), S_{[*]}(w_0)\big) = \mathsf{M}_{L^\infty(\Omega.\nu)}\big(\mathsf{O}{=}(X, \mathcal{F}, F), S_{[*]}(\rho_0)\big) $$

Hence, (C_1) and (C_2) are not so different. In oder to avoid the confusion, we use the following notation:

$$
\begin{cases}
W^*\text{-state } w_0 \; (\in \overline{\mathfrak{S}}^m(\overline{\mathcal{A}}_*) \text{ is written by Roman alphabet (e.g., } w_0, w, v, ...) \\
C^*\text{-state } \rho_0 \; (\in \mathfrak{S}^m(\mathcal{A}^*) \text{ is written by Greek alphabet (e.g., } \rho_0, \rho, ...)
\end{cases}
$$

///

9.2 Simple examples in mixed measurement theory

Recall the following wise sayings:

experience is the best teacher, or custom makes all things

Thus, we exercise the following problem.

Review 9.4. [Answer 5.7 to Problem 5.2 by Fisher's maximum likelihood method]
You do not know the urn behind the curtain. Assume that you pick up a white ball from the urn. Which urn do you think is more likely, U_1 or U_2 ?

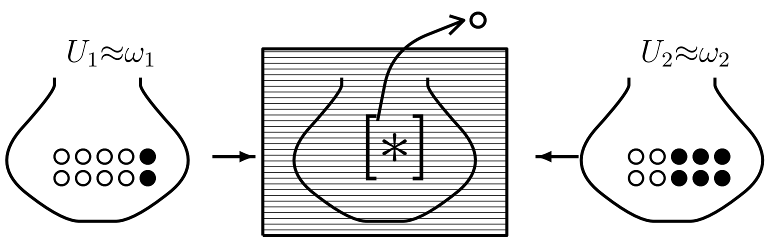


Figure 9.1 (= Figure 5.6:): Pure measurement (Fisher's maximum likelihood method)

Answer Consider the state space $\Omega = \{\omega_1, \omega_2\}$ with the discrete topology and the measure ν such that

$$\nu(\{\omega_1\}) = 1, \qquad \nu(\{\omega_2\}) = 1 \tag{9.3}$$

In the classical basic structure $[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))]$, consider the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}{=}\ (\{W, B\}, 2^{\{W,B\}}, F_{WB}), S_{[*]})$, where the observable $\mathsf{O}_{WB} = (\{W, B\}, 2^{\{W,B\}}, F_{WB})$ in $L^\infty(\Omega)$ is defined by

$$\begin{aligned} &[F_{WB}(\{W\})](\omega_1) = 0.8, && [F_{WB}(\{B\})](\omega_1) = 0.2 \\ &[F_{WB}(\{W\})](\omega_2) = 0.4, && [F_{WB}(\{B\})](\omega_2) = 0.6. \end{aligned} \tag{9.4}$$

Here, we see:

$$\begin{aligned} &\max\{[F_{WB}(\{W\})](\omega_1), [F_{WB}(\{W\})](\omega_2)\} \\ =&\max\{0.8, 0.4\} = 0.8 = F_{WB}(\{W\})](\omega_1). \end{aligned}$$

Then, Fisher's maximum likelihood method (Theorem 5.6) says that

$$[*] = \omega_1.$$

Therefore, there is a reason to infer that the urn behind the curtain is U_1. □

Thus, we exercise the following problem.

Problem 9.5. [mixed measurement $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O}=(X,\mathcal{F},F), S_{[*]}(w))$]

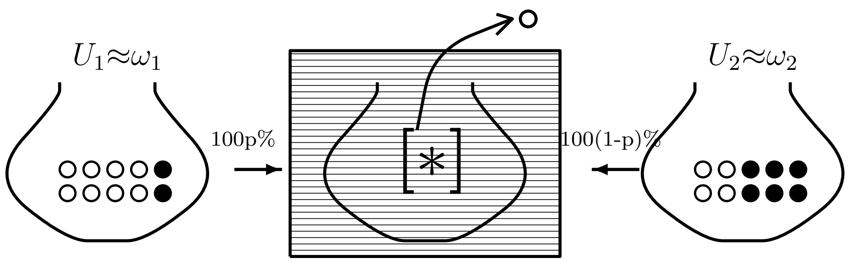


Figure 9.2: Mixed measurement (Urn problem)

($\sharp_1$) Assume an unfair coin-tossing ($T_{p,1-p}$) such that ($0 \leqq p \leqq 1$): That is,

$$\begin{cases} \text{the possibility that "head" appears is } 100p\% \\ \text{the possibility that "tail" appears is } 100(1-p)\% \end{cases}$$

If "head" [resp. "tail"] appears, put an urn $U_1(\approx\omega_1)$ [resp. $U_2(\approx\omega_2)$] behind the curtain. Assume that you do not know which urn is behind the curtain, U_1 or U_2). The unknown urn is denoted by $[*](\in\{\omega_1,\omega_2\})$.

This situation is represented by $w \in L^1_{+1}(\Omega,\nu)$ (with the counting measure ν), that is,

$$w(\omega) = \begin{cases} p & (\text{ if } \omega=\omega_1) \\ 1-p & (\text{ if } \omega=\omega_2) \end{cases}$$

($\sharp_2$) Consider the "measurement" such that a ball is picked out from the unknown urn. This "measurement" is denoted by $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O}, \overline{S}_{[*]}(w))$, and called a mixed measurement.

Then, we have the following problems:

(a) Calculate the probability that a white ball is picked from the unknown urn behind the curtain !

And further,

(b) when a white ball is picked, calculate the probability that the unknown urn behind the curtain is U_1 !

We would like to remark

- the term "subjective probability" is not used in the above problem.

Answer: Assume that the state space$\Omega = \{\omega_1, \omega_2\}$ is defined by the discrete metric with the following measure ν:

$$\nu(\{\omega_1\}) = 1, \qquad \nu(\{\omega_2\}) = 1. \tag{9.5}$$

Thus, we start from the classical basic structure:

$$[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))], \tag{9.6}$$

in which we consider the mixed measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}{=}\ (\{W,\ B\},\ 2^{\{W,B\}},\ F), S_{[*]}(w))$. Here, the observable $\mathsf{O}_{WB} = (\{W, B\}, 2^{\{W,B\}}, F_{WB})$ in $L^\infty(\Omega)$ is defined by

$$\begin{aligned} &[F_{WB}(\{W\})](\omega_1) = 0.8, && [F_{WB}(\{B\})](\omega_1) = 0.2 \\ &[F_{WB}(\{W\})](\omega_2) = 0.4, && [F_{WB}(\{B\})](\omega_2) = 0.6. \end{aligned} \tag{9.7}$$

Also, the mixed state $w_0 \in L^1_{+1}(\Omega, \nu)$ is defined by

$$w_0(\omega_1) = p, \qquad w_0(\omega_2) = 1 - p. \tag{9.8}$$

Then, by Axiom$^{(m)}$ 1, we see

(a): the probability that a measured value x $(\in \{W, B\})$ is obtained by $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}{=}\ (\{W,\ B\},\ 2^{\{W,B\}},\ F), S_{[*]}(w))$ is given by

$$\begin{aligned} P(\{x\}) &= {}_{L^1(\Omega)}\big(w_0, F(\{x\})\big)_{L^\infty(\Omega)} = \int_\Omega [F(\{x\})](\omega) \cdot w_0(\omega) \nu(d\omega) \\ &= p[F(\{x\})](\omega_1) + (1-p)[F(\{x\})](\omega_2) \\ &= \begin{cases} 0.8p + 0.4(1-p) & (\text{when } \ x = W) \\ 0.2p + 0.6(1-p) & (\text{when } \ x = B) \end{cases} \end{aligned} \tag{9.9}$$

The question (b) will be answered in Answer 9.13. □

♠**Note 9.1.** The following question is natural. That is,

($\sharp_1$) In the above (i), why is "the possibility that $[\,*\,] = \omega_1$ is $100p\%$ $\cdots$" replaced by "the probability that $[\,*\,] = \omega_1$ is $100p\%$ $\cdots$" ?

However, the linguistic interpretation says that

($\sharp_2$) there is no probability without measurements.

This is the reason why the term "probability" is not used in (i). However, from the practical point of view, we are not sensitive to the difference between "probability" and "possibility".

Example 9.6. [Mixed spin measurement $\mathsf{M}_{B(\mathbb{C}^2)}(\mathsf{O} = (X = \{\uparrow, \downarrow\}, 2^X, F^z), S_{[*]}(w))$] Consider the quantum basic structure:

$$[B(\mathbb{C}^2) \subseteq B(\mathbb{C}^2) \subseteq B(\mathbb{C}^2)]$$

And consider a particle P_1 with spin state $\rho_1 = |a\rangle\langle a| \in \mathfrak{S}^p(B(\mathbb{C}^2))$, where

$$a = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \in \mathbb{C}^2 \quad (\|a\| = (|\alpha_1|^2 + |\alpha_2|^2)^{1/2} = 1)$$

And consider another particle P_2 with spin state $\rho_2 = |b\rangle\langle b| \in \mathfrak{S}^p(B(\mathbb{C}^2))$, where

$$b = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \in \mathbb{C}^2 \quad (\|b\| = (|\beta_1|^2 + |\beta_2|^2)^{1/2} = 1)$$

Here, assume that

- the "probability" that the "particle" P is $\left\{ \begin{array}{l} \text{a particle } P_1 \\ \text{a particle } P_2 \end{array} \right\}$ is given by $\left\{ \begin{array}{l} p \\ 1-p \end{array} \right\}$

That is,

$$\underset{\text{(Particle } P_1)}{\boxed{\text{state } \rho_1}} \xrightarrow[\text{“probability” } p]{} \underset{\text{(Particle } P)}{\boxed{\text{unknown state } [*]}} \xleftarrow[\text{“probability” } 1-p]{} \underset{\text{(Particle } P_2)}{\boxed{\text{state } \rho_2}}$$

Here, the unknown state $[*]$ of Particle P is represented by the mixed state$w$ $(\in \mathfrak{S}^m(\mathcal{T}r(\mathbb{C}^2)))$ such that

$$w = p\rho_1 + (1-p)\rho_2 = p|a\rangle\langle a| + (1-p)|b\rangle\langle b|$$

Therefore, we have the mixed measurement $\mathsf{M}_{B(\mathbb{C}^2)}(\mathsf{O}_z = (X, 2^X, F^z), S_{[*]}(w))$ of the z-axis spin observable $\mathsf{O}_z = (X, \mathcal{F}, F^z)$, where

$$F^z(\{\uparrow\}) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad F^z(\{\downarrow\}) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

And we say that

(a) the probability that a measured value $\left\{ \begin{array}{c} \uparrow \\ \downarrow \end{array} \right\}$ is obtained by the mixed measurement $\mathsf{M}_{B(\mathbb{C}^2)}(\mathsf{O}_z = (X, 2^X, F^z), S_{[*]}(w))$ is given by

$$\left\{ \begin{array}{l} {}_{\mathcal{T}r(\mathbb{C}^2)}\Big(w, F^z(\{\uparrow\})\Big)_{B(\mathbb{C}^2)} = p|\alpha_1|^2 + (1-p)|\beta_1|^2 \\ \\ {}_{\mathcal{T}r(\mathbb{C}^2)}\Big(w, F^z(\{\downarrow\})\Big)_{B(\mathbb{C}^2)} = p|\alpha_2|^2 + (1-p)|\beta_2|^2 \end{array} \right\}$$

Remark 9.7. As seen in the above, we say that

(a) Pure measurement theory is fundamental. Adding the concept of "mixed state", we can construct mixed measurement theory as follows.

$$\underset{\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[*]}(w))}{\boxed{\text{mixed measurement theory}}} := \underset{\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[*]})}{\boxed{\text{pure measurement theory}}} + \underset{w}{\boxed{\text{mixed state}}}$$

Therefore,

There is no mixed measurement without pure measurement

That is, in quantum language, there is no confrontation between "frequency probability" and "subjective probability". The reason that a coin-tossing is used in Problem 9.5 is to emphasize that the naming of "subjective probability" is improper.

9.3 St. Petersburg two envelope problem

This section is extracted from the following:

Ref. [46]: S. Ishikawa; The two envelopes paradox in non-Bayesian and Bayesian statistics
(arXiv:1408.4916v4 [stat.OT] 2014)

Now, we shall review the St. Petersburg two envelope problem (*cf.* [9][1]).

Problem 9.8. [The St. Petersburg two envelope problem] The host presents you with a choice between two envelopes (i.e., Envelope A and Envelope B). You are told that each of them contains an amount determined by the following procedure, performed separately for each envelope:

(♯) a coin was flipped until it came up heads, and if it came up heads on the k-th trial, 2^k is put into the envelope. This procedure is performed separately for each envelope.

You choose randomly (by a fair coin toss) one envelope. For example, assume that the envelope is Envelope A. And therefore, the host get Envelope B. You find 2^m dollars in the envelope A. Now you are offered the options of keeping A (=your envelope) or switching to B (= host's envelope). **What should you do?**

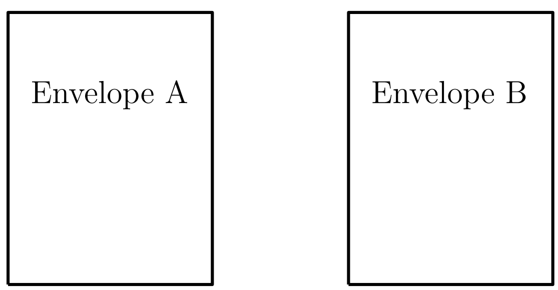


Figure 9.2: Two envelope problem

[(P2):Why is it paradoxical?].
You reason that, before opening the envelopes A and B, the expected values $E(x)$ and $E(y)$ in A and B is infinite respectively. That is because

$$1 \times \frac{1}{2} + 2 \times \frac{1}{2^2} + 2^2 \times \frac{1}{2^3} + \cdots = \infty$$

For any 2^m, if you knew that A contained $x = 2^m$ dollars, then the expected value $E(y)$ in B would still be infinite. Therefore, you should switch to B. But this seems clearly wrong, as your information about A and B is symmetrical. This is the famous St. Petersburg two-envelope paradox (i.e., "The Other Person's Envelope is Always Greener").

[1] D.J. Chalmers, "The St. Petersburg Two-Envelope Paradox," Analysis, Vol.62, 155-157, (2002)

9.3.1 (P2): St. Petersburg two envelope problem: classical mixed measurement

Define the state space Ω such that $\Omega = \{\omega = 2^k \mid k = 1, 2, \cdots\}$, with the discrete metric and the counting measure ν. And define the exact observable $\mathsf{O} = (X, \mathcal{F}, F)$ in $L^\infty(\Omega, \nu)$ such that

$$
\begin{aligned}
& X = \Omega, \quad \mathcal{F} = 2^X \equiv \{\Xi \mid \Xi \subseteq X\} \\
& [F(\Xi)](\omega) = \chi_\Xi(\omega) \equiv \begin{cases} 1 & (\omega \in \Xi) \\ 0 & (\omega \notin \Xi) \end{cases} \qquad (\forall \Xi \in \mathcal{F}, \forall \omega \in \Omega)
\end{aligned}
$$

Define the mixed state w ($\in L^1_{+1}(\Omega, \nu)$, i.e., the probability density function on Ω) such that

$$
w_0(\omega) = 2^{-k} \quad (\forall \omega = 2^k \in \Omega).
$$

Consider the mixed measurement $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O} = (X, \mathcal{F}, F), \overline{S}_{[*]}(w_0))$. Axiom$^{(\mathrm{m})}$ 1(C_1) (§9.1) says that

(A) the probability that a measured value 2^k is obtained by $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O} = (X, \mathcal{F}, F), \overline{S}_{[*]}(w_0))$ is given by 2^{-k}.

Therefore, the expectation of the measured value is calculated as follows.

$$
E = \sum_{k=1}^{\infty} 2^k \cdot 2^{-k} = \infty
$$

Note that you knew that A contained $x = 2^m$ dollars (and thus, $E = \infty > 2^m$). There is a reason to consider that the switching to B is an advantage.

Remark 9.9. After you get a measured value 2^m from the envelope A, you can guess (also see Bayes theorem later) that the probability density function w_0 changes to the new w_1 such that $w_1(2^m) = 1, w_1(2^k) = 0 (k \neq m)$. Thus, now your information about $A : w_1$ and $B : w_0$ is not symmetrical. Hence, in this case, it is true: *"The Other Person's envelope is Always Greener"*.

♠**Note 9.2.** There are various criterions except the expectaion. For example, consider the criterion such that

($\sharp$) "the probability that the switching is disadvantageous" $< \frac{1}{2}$

Under this criterion, it is reasonable to judge that

$$
\begin{cases} m = 1 & \Longrightarrow \text{switching to } B \\ m = 2, 3, \ldots & \Longrightarrow \text{keeping } A \end{cases}
$$

9.4 Bayesian statistics is to use Bayes theorem

Although there may be several opinions for the question "What is Bayesian statistics?", we think that

Bayesian statistics is to use Bayes theorem

Thus,

let us start from Bayes theorem.

The following is clear.

Theorem 9.10. [The conditional probability]. Consider the mixed measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}{=}\ (X \times Y, \mathcal{F} \boxtimes \mathcal{G}, H),\ S_{[*]}(w))$, which is formulated in the basic structure

$$[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]$$

Assume that a measured value (x, y) $(\in X \times Y)$ is obtained by the mixed measurement$\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}{=}\ (X \times Y, \mathcal{F} \boxtimes \mathcal{G}, H),\ S_{[*]}(w))$ belongs to $\Xi \times Y$ $(\in \mathcal{F})$. Then, the probability that $y \in \Gamma$ is given by

$$\frac{{}_{\overline{\mathcal{A}}_*}(w, H(\Xi \times \Gamma))_{\overline{\mathcal{A}}}}{{}_{\overline{\mathcal{A}}_*}(w, H(\Xi \times Y))_{\overline{\mathcal{A}}}} \quad (\forall \Gamma \in \mathcal{G})$$

Proof. This is due to the property (or, common sense) of conditional probability. $\square$

In the classical case, this is rewritten as follows.

Theorem 9.11. [Bayes' Theorem (in classical mixed measurement)]. Consider the simultaneous measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}{=}\ (X \times Y, \mathcal{F} \boxtimes \mathcal{G}, F \times G),\ S_{[*]}(w_0))$ formulated in the classical basic structure $[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))]$. Here the observable $\mathsf{O}_{12}{=}(X \times Y, \mathcal{F} \boxtimes \mathcal{G}, F \times G)$ is defined by the simultaneous observable of the two observables $\mathsf{O}_1{=}(X, \mathcal{F}, F)$ and $\mathsf{O}_2{=}(Y, \mathcal{G}, G)$. That is,

$$(F \times G)(\Xi \times \Gamma) = F(\Xi) \cdot G(\Gamma) \qquad (\forall \Xi \in \mathcal{F}, \forall \Gamma \in \mathcal{G}). \tag{9.10}$$

Assume that

(a) a measured value (x, y) $(\in X \times Y)$ obtained by the mixed measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{12}{=}$

$(X \times Y, \mathcal{F} \boxtimes \mathcal{G}, F \times G)$, $S_{[*]}(w_0)\big)$ belongs to $\Xi \times Y$ (where, $\Xi \in \mathcal{F}$).

Then, the probability such that "$y \in \Gamma$" is given by

$$\frac{{}_{L^1(\Omega)}(w_0, H(\Xi \times \Gamma))_{L^\infty(\Omega)}}{{}_{L^1(\Omega)}(w_0, H(\Xi \times Y))_{L^\infty(\Omega)}} \Big(= \frac{\int_\Omega [F(\Xi)](\omega) \cdot [G(\Gamma)](\omega) \cdot w_0(\omega)\nu(d\omega)}{\int_\Omega [F(\Xi)](\omega) \cdot w_0(\omega)\nu(d\omega)} \Big). \tag{9.11}$$

Here, putting

(b) $w_{\rm new}(\omega) = \frac{[F(\Xi)](\omega) \cdot w_0(\omega)}{\int_\Omega [F(\Xi)](\omega) \cdot w_0(\omega)\nu(d\omega)}$ $(\forall \omega \in \Omega)$.

we see:

$$(9.23) = \int_\Omega [G(\Gamma)](\omega) w_{\rm new}(\omega)\nu(d\omega) \quad (\forall \Gamma \in \mathcal{G}). \tag{9.12}$$

Remark 9.12. [How to understand Bayes' Theorem] Bayes' theorem 9.11 is usually read as follows.

(b′) If a measured value x $(\in X)$ obtained by the mixed measurement $\mathsf{M}_{L^\infty(\Omega)}\big(\mathsf{O}_1 = (X, \mathcal{F}, F)$, $S_{[*]}(w_0)\big)$ belongs to Ξ $(\in \mathcal{F})$, then, the following state collapse happens:

$$\underset{\text{pre-state}}{\boxed{w_0}} \xrightarrow[x \in \Xi]{} \underset{\text{post-state}}{\boxed{w_{\rm new}}}$$

The above (d) superficially contradicts the linguistic interpretation, which says

A state never moves.

In this sense, the above (b) or (b′) (i.e., Bayes' theorem) is convenient and makeshift.

Answer 9.13. [Bayes' Theorem (=Problem 9.5 and the answer to (c_2))]

Assume that the state space $\Omega = \{\omega_1, \omega_2\}$ is defined by the discrete metric with the following measure ν:

$$\nu(\{\omega_1\}) = 1, \qquad \nu(\{\omega_2\}) = 1. \tag{9.13}$$

Thus, we start from the classical basic structure:

$$[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))], \tag{9.14}$$

in which we consider the mixed measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O} = (\{W, B\}, 2^{\{W,B\}}, F), S_{[*]}(w))$. Here, the observable $\mathsf{O}_{WB} = (\{W, B\}, 2^{\{W,B\}}, F_{WB})$ in $L^\infty(\Omega)$ is defined by

$$[F_{WB}(\{W\})](\omega_1) = 0.8, \qquad\qquad [F_{WB}(\{B\})](\omega_1) = 0.2,$$

$$[F_{WB}(\{W\})](\omega_2) = 0.4, \qquad [F_{WB}(\{B\})](\omega_2) = 0.6. \tag{9.15}$$

Also, the mixed state $w_0 \in L^1_{+1}(\Omega, \nu)$ is defined by

$$w_0(\omega_1) = p, \qquad w_0(\omega_2) = 1 - p. \tag{9.16}$$

Then, by Axiom$^{(m)}$ 1, we see

(a): the probability that a measured value x ($\in \{W, B\}$) is obtained by $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}= (\{W, B\}, 2^{\{W,B\}}, F), S_{[*]}(w))$ is given by

$$\begin{aligned}P(\{x\}) &= {}_{L^1(\Omega)}\big(w_0, F(\{x\})\big)_{L^\infty(\Omega)} = \int_\Omega [F(\{x\})](\omega) \cdot w_0(\omega)\nu(d\omega) \\ &= p[F(\{x\})](\omega_1) + (1-p)[F(\{x\})](\omega_2) \\ &= \begin{cases} 0.8p + 0.4(1-p) & (\text{when } \ x = W) \\ 0.2p + 0.6(1-p) & (\text{when } \ x = B) \end{cases}\end{aligned} \tag{9.17}$$

[W^*-algebraic answer to Problem 9.5(c_2) in Sec. 9.1.2]
Since "white ball" is obtained by a mixed measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[*]}(w_0))$, a new mixed state $w_{\text{new}}(\in L^1_{+1}(\Omega))$ is given by

$$w_{\text{new}}(\omega) = \frac{[F(\{W\})](\omega)w_0(\omega)}{\int_\Omega [F(\{W\})](\omega)w_0(\omega)\nu(d\omega)} = \begin{cases} \dfrac{0.8p}{0.8p + 0.2(1-p)} & (\text{when } \omega = \omega_1) \\ \\ \dfrac{0.2(1-p)}{0.8p + 0.2(1-p)} & (\text{when } \omega = \omega_2) \end{cases}$$

[C^*-algebraic answer to Problem 9.5 (c_2) in Sec. 9.1.2]
Since "white ball" is obtained by a mixed measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[*]}(\rho_0))$, a new mixed state $\rho_{\text{new}}(\in \mathcal{M}_{+1}(\Omega))$ is given by

$$\rho_{\text{new}} = \frac{F(\{W\})\rho_0}{\int_\Omega [F(\{W\})](\omega)\rho_0(d\omega)} = \frac{0.8p}{0.8p + 0.2(1-p)}\delta_{\omega_1} + \frac{0.2(1-p)}{0.8p + 0.2(1-p)}\delta_{\omega_2}.$$

9.5 Two envelope problem (Bayes' method)

This section is extracted from the following:

ref. [46]: S. Ishikawa; The two envelopes paradox in non-Bayesian and Bayesian statistics (arXiv:1408.4916v4 [stat.OT] 2014)

Problem 9.14. [(=**Problem5.16**): **the two envelope problem**]
The host presents you with a choice between two envelopes (i.e., Envelope A and Envelope B). You know one envelope contains twice as much money as the other, but you do not know which contains more. That is, Envelope A [resp. Envelope B] contains V_1 dollars [resp. V_2 dollars]. You know that

(a) $\frac{V_1}{V_2} = 1/2$ or, $\frac{V_1}{V_2} = 2$

Define the exchanging map $\overline{x} : \{V_1, V_2\} \to \{V_1, V_2\}$ by

$$\overline{x} = \begin{cases} V_2, & (\text{ if } x = V_1), \\ V_1 & (\text{ if } x = V_2) \end{cases}$$

You choose randomly (by a fair coin toss) one envelope, and you get x_1 dollars (i.e., if you choose Envelope A [resp. Envelope B], you get V_1 dollars [resp. V_2 dollars]). And the host gets $\overline{x}_1$ dollars. Thus, you can infer that $\overline{x}_1 = 2x_1$ or $\overline{x}_1 = x_1/2$. Now the host says "You are offered the options of keeping your x_1 or switching to my $\overline{x}_1$". **What should you do?**

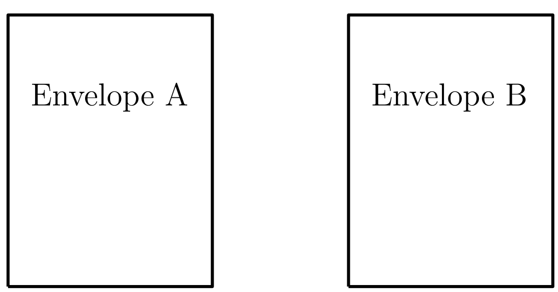


Figure 9.4: Two envelope problem

[(P1):Why is it paradoxical?]. You get $\alpha = x_1$. Then, you reason that, with probability 1/2, $\overline{x}_1$ is equal to either $\alpha/2$ or 2α dollars. Thus the expected value (denoted $E_{\text{other}}(\alpha)$ at this moment) of the other envelope is

$$E_{\text{other}}(\alpha) = (1/2)(\alpha/2) + (1/2)(2\alpha) = 1.25\alpha \tag{9.18}$$

This is greater than the α in your current envelope A. Therefore, you should switch to B. But this seems clearly wrong, as your information about A and B is symmetrical. This is the famous two-envelope paradox (i.e., "The Other Person's Envelope is Always Greener").

9.5.1 (P1): Bayesian approach to the two envelope problem

Consider the state space Ω such that

$$\Omega = \overline{\mathbb{R}}_+ (= \{\omega \in \mathbb{R} \mid \omega \geq 0\})$$

with Lebesgue measure ν. Thus, we start from the classical basic structure

$$[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))]$$

Also, putting $\widehat{\Omega} = \{(\omega, 2\omega) \mid \omega \in \overline{\mathbb{R}}_+\}$, we consider the identification:

$$\Omega \ni \omega \underset{\text{(identification)}}{\longleftrightarrow} (\omega, 2\omega) \in \widehat{\Omega} \tag{9.19}$$

Further, define $V_1 : \Omega (\equiv \overline{\mathbb{R}}_+) \to X(\equiv \overline{\mathbb{R}}_+)$ and $V_2 : \Omega (\equiv \overline{\mathbb{R}}_+) \to X(\equiv \overline{\mathbb{R}}_+)$ such that

$$V_1(\omega) = \omega, \quad V_2(\omega) = 2\omega \qquad (\forall \omega \in \Omega)$$

And define the observable $\mathsf{O} = (X(= \overline{\mathbb{R}}_+), \mathcal{F}(= \mathcal{B}_{\overline{\mathbb{R}}_+} : \text{ the Borel field}), F)$ in $L^\infty(\Omega, \nu)$ such that

$$[F(\Xi)](\omega) = \begin{cases} 1 & (\text{ if } \omega \in \Xi, \ \ 2\omega \in \Xi) \\ 1/2 & (\text{ if } \omega \in \Xi, \ \ 2\omega \notin \Xi) \\ 1/2 & (\text{ if } \omega \notin \Xi, \ \ 2\omega \in \Xi) \\ 0 & (\text{ if } \omega \notin \Xi, \ \ 2\omega \notin \Xi) \end{cases} \qquad (\forall \omega \in \Omega, \forall \Xi \in \mathcal{F})$$

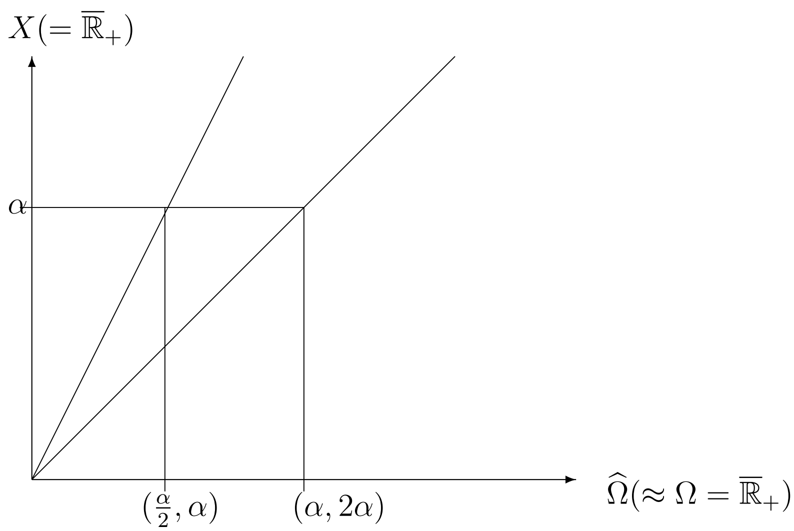


Figure 9.5: Two envelope problem

Recalling the identification : $\widehat{\Omega} \ni (\omega, 2\omega) \longleftrightarrow \omega \in \Omega = \overline{\mathbb{R}}_+$, assume that

$$\rho_0(D) = \int_D w_0(\omega) d\omega \quad (\forall D \in \mathcal{B}_\Omega = \mathcal{B}_{\overline{\mathbb{R}}_+})$$

where the probability density function $w_0 : \Omega(\approx \overline{\mathbb{R}}_+) \to \overline{\mathbb{R}}_+$ is assumed to be continuous positive function. That is, the mixed state $\rho_0 (\in \mathcal{M}_{+1}(\Omega(= \overline{\mathbb{R}}_+)))$ has the probability density function w_0.

Axiom$^{(m)}$ 1(§9.1) says that

(A$_1$) The probability $P(\Xi)$ ($\Xi \in \mathcal{B}_X = \mathcal{B}_{\overline{\mathbb{R}}_+}$) that a measured value obtained by the mixed measurement $\mathsf{M}_{L^\infty(\Omega,d\omega)}(\mathsf{O}=(X,\mathcal{F},F), S_{[*]}(\rho_0))$ belongs to $\Xi(\in \mathcal{B}_X = \mathcal{B}_{\overline{\mathbb{R}}_+})$ is given by

$$
\begin{aligned}
P(\Xi) &= \int_\Omega [F(\Xi)](\omega)\rho_0(d\omega) = \int_\Omega [F(\Xi)](\omega) w_0(\omega) d\omega \\
&= \int_\Xi \frac{w_0(x/2)}{4} + \frac{w_0(x)}{2} \; dx \quad (\forall \Xi \in \mathcal{B}_{\overline{\mathbb{R}}_+})
\end{aligned} \tag{9.20}
$$

Therefore, the expectation is given by

$$
\int_{\overline{\mathbb{R}}_+} xP(dx) = \frac{1}{2}\int_0^\infty x \cdot \Big(w_0(x/2)/2 + w_0(x) \Big) dx = \frac{3}{2}\int_{\overline{\mathbb{R}}_+} x w_0(x) dx
$$

Further, Theorem 9.11 (Bayes' theorem) says that

(A$_2$) When a measured value α is obtained by the mixed measurement $\mathsf{M}_{L^\infty(\Omega,d\omega)}(\mathsf{O}=(X,\mathcal{F},F),$ $S_{[*]}(\rho_0))$, then the post-state $\rho_{\rm post}(\in \mathcal{M}_{+1}(\Omega))$ is given by

$$
\rho_{\rm post}^\alpha = \frac{\frac{w_0(\alpha/2)}{2}}{\frac{h(\alpha/2)}{2} + w_0(\alpha)} \delta_{(\frac{\alpha}{2},\alpha)} + \frac{w_0(\alpha)}{\frac{w_0(\alpha/2)}{2} + w_0(\alpha)} \delta_{(\alpha,2\alpha)} \tag{9.21}
$$

Hence,

(A$_3$) if $[*] = \left\{ \begin{array}{l} \delta_{(\frac{\alpha}{2},\alpha)} \\ \delta_{(\alpha,2\alpha)} \end{array} \right\}$, then you change $\left\{ \begin{array}{l} \alpha \longrightarrow \frac{\alpha}{2} \\ \alpha \longrightarrow 2\alpha \end{array} \right\}$, and thus you get the switching gain $\left\{ \begin{array}{l} \frac{\alpha}{2} - \alpha (= -\frac{\alpha}{2}) \\ 2\alpha - \alpha (= \alpha) \end{array} \right\}$.

Therefore, the expectation of the switching gain is calculated as follows:

$$
\begin{aligned}
&\int_{\overline{\mathbb{R}}_+} \Big((-\frac{\alpha}{2}) \frac{\frac{w_0(\alpha/2)}{2}}{\frac{w_0(\alpha/2)}{2} + w_0(\alpha)} + \alpha \frac{w_0(\alpha)}{\frac{w_0(\alpha/2)}{2} + w_0(\alpha)} \Big) P(d\alpha) \\
= &\int_{\overline{\mathbb{R}}_+} (-\frac{\alpha}{2}) \frac{w_0(\alpha/2)}{4} + \alpha \cdot \frac{w_0(\alpha)}{2} \; d\alpha = 0
\end{aligned} \tag{9.22}
$$

Therefore, we see that the swapping is even, i.e., no advantage and no disadvantage.

9.6 Monty Hall problem (The Bayesian approach)

9.6.1 The review of Problem5.14 (Monty Hall problem in pure measurement)

Problem 9.15. [**Monty Hall problem (The answer to Fisher's maximum likelihood method)**]

You are on a game show and you are given the choice of three doors. Behind one door is a car, and behind the other two are goats. You choose, say, door 1, and the host, who knows where the car is, opens another door, behind which is a goat. For example, the host says that

($\flat$) the door 3 has a goat.

And further, He now gives you the choice of sticking with door 1 or switching to door 2? **What should you do?**

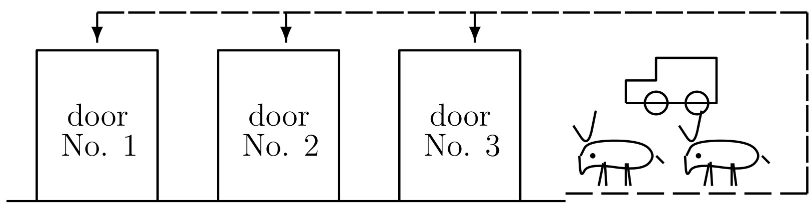


Figure 9.6: Monty Hall problem

Answer: Put $\Omega = \{\omega_1, \omega_2, \omega_3\}$ with the discrete topology d_D and the counting measure ν. Thus consider the classical basic structure:

$$[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))]$$

Assume that each state $\delta_{\omega_m} (\in \mathfrak{S}^p(C_0(\Omega)^*))$ means

$$\delta_{\omega_m} \Leftrightarrow \text{ the state that the car is behind the door 1} \quad (m = 1, 2, 3)$$

Define the observable $\mathsf{O}_1 \equiv (\{1, 2, 3\}, 2^{\{1,2,3\}}, F_1)$ in $L^\infty(\Omega)$ such that

$$[F_1(\{1\})](\omega_1) = 0.0, \qquad [F_1(\{2\})](\omega_1) = 0.5, \qquad [F_1(\{3\})](\omega_1) = 0.5,$$
$$[F_1(\{1\})](\omega_2) = 0.0, \qquad [F_1(\{2\})](\omega_2) = 0.0, \qquad [F_1(\{3\})](\omega_2) = 1.0,$$

$$[F_1(\{1\})](\omega_3) = 0.0, \qquad [F_1(\{2\})](\omega_3) = 1.0, \qquad [F_1(\{3\})](\omega_3) = 0.0, \tag{9.23}$$

where it is also possible to assume that $F_1(\{2\})(\omega_1) = \alpha$, $F_1(\{3\})(\omega_1) = 1 - \alpha$ $(0 < \alpha < 1)$. The fact that you say "the door 1" means that we have a measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_1, S_{[*]})$. Here, we assume that

a) "a measured value 1 is obtained by the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_1, S_{[*]})$"
$\Leftrightarrow$ The host says "Door 1 has a goat"

b) "measured value 2 is obtained by the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_1, S_{[*]})$ "
$\Leftrightarrow$ The host says "Door 2 has a goat"

c) "measured value 3 is obtained by the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_1, S_{[*]})$ "
$\Leftrightarrow$ The host says "Door 3 has a goat"

Since the host said "Door 3 has a goat," this implies that you get the measured value "3" by the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_1, S_{[*]})$. Therefore, Theorem 5.6 (Fisher's maximum likelihood method) says that *you should pick door number 2.* That is because we see that

$$\begin{aligned}\max\{[F_1(\{3\})](\omega_1), [F_1(\{3\})](\omega_2), [F_1(\{3\})](\omega_3)\} &= \max\{0.5, \ \ 1.0, \ \ 0.0\}\\ &= 1.0 = [F_1(\{3\})](\omega_2)\end{aligned}$$

and thus, there is a reason to infer that $[*] = \delta_{\omega_2}$. Thus, you should switch to door 2. This is the first answer to Monty-Hall problem. □

9.6.2 Monty Hall problem in mixed measurement

Next, let us study Monty Hall problem in mixed measurement theory (particularly, Bayesian statistics).

Problem 9.16. [Monty Hall problem(The answer by Bayes' method)]

Suppose you are on a game show, and you are given the choice of three doors (i.e., "number 1," "number 2," "number 3"). Behind one door is a car, behind the others, goats. You pick a door, say number 1. Then, the host, who set a car behind a certain door, says

($\sharp_1$) the car was set behind the door decided by the cast of the distorted dice. That is, the host set the car behind the k-th door (i.e., "number k") with probability p_k (or, weight such that $p_1 + p_2 + p_3 = 1$, $0 \le p_1, p_2, p_3 \le 1$).

And further, the host says, for example,

($\flat$) the door 3 has a goat.

He says to you, "Do you want to pick door number 2?" Is it to your advantage to switch your choice of doors?

Answer: In the same way as we did in Problem 9.15 (Monty Hall problem:the answer by Fisher's maximum likelihood method), consider the state space $\Omega = \{\omega_1, \omega_2, \omega_3\}$ with the discrete metric d_D and the observable O_1. Under the hypothesis ($\sharp_1$), define the mixed state ν_0 $(\in \mathcal{M}_{+1}(\Omega))$ such that

$$\nu_0 = p_1 \delta_{\omega_1} + p_2 \delta_{\omega_2} + p_3 \delta_{\omega_3}$$

namely,

$$\nu_0(\{\omega_1\}) = p_1, \quad \nu_0(\{\omega_2\}) = p_2, \quad \nu_0(\{\omega_3\}) = p_3$$

Thus we have a mixed measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_1, S_{[*]}(\nu_0))$. Note that

a) "measured value 1 is obtained by the mixed measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_1, S_{[*]}(\nu_0))$"
$\Leftrightarrow$ the host says "Door 1 has a goat"

b) "measured value 2 is obtained by the mixed measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_1, S_{[*]}(\nu_0))$"
$\Leftrightarrow$ the host says "Door 2 has a goat"

c) "measured value 3 is obtained by the mixed measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_1, S_{[*]}(\nu_0))$"
$\Leftrightarrow$ the host says "Door 3 has a goat"

Here, assume that, by the mixed measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_1, S_{[*]}(\nu_0))$, you obtain a measured value 3, which corresponds to the fact that the host said "Door 3 has a goat." Then, Theorem 9.11 (Bayes' theorem) says that the posterior state ν_{post} $(\in \mathcal{M}_{+1}(\Omega))$ is given by

$$\nu_{\text{post}} = \frac{F_1(\{3\}) \times \nu_0}{\big\langle \nu_0, F_1(\{3\}) \big\rangle}.$$

That is,

$$\nu_{\text{post}}(\{\omega_1\}) = \frac{\frac{p_1}{2}}{\frac{p_1}{2} + p_2}, \quad \nu_{\text{post}}(\{\omega_2\}) = \frac{p_2}{\frac{p_1}{2} + p_2}, \quad \nu_{\text{post}}(\{\omega_3\}) = 0.$$

Particularly, we see that

($\sharp_2$) if $p_1 = p_2 = p_3 = 1/3$, then it holds that $\nu_{\text{post}}(\{\omega_1\}) = 1/3$, $\nu_{\text{post}}(\{\omega_2\}) = 2/3$, $\nu_{\text{post}}(\{\omega_3\}) = 0$, *and thus, you should pick Door 2.*

□

♠**Note 9.3.** It is not natural to assume the rule ($\sharp_1$) in Problem 9.16. That is because the host may intentionally set the car behind a certain door. Thus we think that Problem 9.16 is temporary. For our formal assertion, see Problem 9.17 latter.

9.7 Monty Hall problem (The principle of equal weight)

9.7.1 The principle of equal weight— The most famous unsolved problem

Let us reconsider Monty Hall problem (Problem 9.14, Problem9.15) in what follows. We think that the following is one of the most reasonable answers (also, see Problem 19.5).

Problem 9.17. [**Monty Hall problem (The principle of equal weight)**]

Suppose you are on a game show, and you are given the choice of three doors (i.e., "number 1," "number 2," "number 3"). Behind one door is a car, behind the others, goats.

($\sharp_2$) *You choose a door by the cast of the fair dice, i.e., with probability* 1/3.

According to the rule ($\sharp_2$), you pick a door, say number 1, and the host, who knows where the car is, opens another door, behind which is a goat. For example, the host says that

($\flat$) *the door 3 has a goat.*

He says to you, "Do you want to pick door number 2?" Is it to your advantage to switch your choice of doors?

Answer: By the same way of Problem9.15 and Problem9.16 (Monty Hall problem), define the state space $\Omega = \{\omega_1, \omega_2, \omega_3\}$ and the observable $\mathsf{O} = (X, \mathcal{F}, F)$. And the observable $\mathsf{O} = (X, \mathcal{F}, F)$ is defined by the formula (9.23). The map $\phi : \Omega \to \Omega$ is defined by

$$\phi(\omega_1) = \omega_2, \quad \phi(\omega_2) = \omega_3, \quad \phi(\omega_3) = \omega_1$$

we get a causal operator $\Phi : L^\infty(\Omega) \to L^\infty(\Omega)$ by $[\Phi(f)](\omega) = f(\phi(\omega))$ $(\forall f \in L^\infty(\Omega),\ \forall \omega \in \Omega)$. Assume that a car is behind the door k $(k = 1, 2, 3)$. Then, we say that

(a) By the dice-throwing, you get $\begin{bmatrix} 1,2 \\ 3,4 \\ 5,6 \end{bmatrix}$, then, take a measurement $\begin{bmatrix} \mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[\omega_k]}) \\ \mathsf{M}_{L^\infty(\Omega)}(\Phi\mathsf{O}, S_{[\omega_k]}) \\ \mathsf{M}_{L^\infty(\Omega)}(\Phi^2\mathsf{O}, S_{[\omega_k]}) \end{bmatrix}$

We, by the argument in Chapter 11 (*cf.* the formula (11.7))[2], see the following identifications:

$$\mathsf{M}_{L^\infty(\Omega)}(\Phi\mathsf{O}, S_{[\omega_k]}) = \mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[\phi(\omega_k)]}),\ \mathsf{M}_{L^\infty(\Omega)}(\Phi^2\mathsf{O},\ S_{[\omega_k]}) = \mathsf{M}_{L^\infty(\Omega)}(\mathsf{O},\ S_{[\phi^2(\omega_k)]}).$$

Thus, the above (a) is equal to

---

[2]Thus, from the pure theoretical point of view, this problem should be discussed after Chapter 11

(b) By the dice-throwing, you get $\begin{bmatrix} 1,2 \\ 3,4 \\ 5,6 \end{bmatrix}$ then, take a measurement $\begin{bmatrix} \mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[\omega_k]}) \\ \mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[\phi(\omega_k)]}) \\ \mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[\phi^2(\omega_k)]}) \end{bmatrix}$

Here, note that $\frac{1}{3}(\delta_{\omega_k} + \delta_{\phi(\omega_k)} + \delta_{\phi^2(\omega_k)}) = \frac{1}{3}(\delta_{\omega_1} + \delta_{\omega_2} + \delta_{\omega_3})$ $(\forall k = 1,2,3)$. Thus, this (b) is identified with the mixed measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[*]}(\nu_e))$, where

$$\nu_e = \frac{1}{3}(\delta_{\omega_1} + \delta_{\omega_2} + \delta_{\omega_3})$$

Therefore, Problem 9.17 is the same as Problem 9.16. Hence, you should choose the door 2. □

♠**Note 9.4.** The above argument is easy. That is, since you have no information, we choose the door by a fair dice throwing. In this sense, the principle of equal weight — unless we have sufficient reason to regard one possible case as more probable than another, we treat them as equally probable — is clear in measurement theory. However, it should be noted that the above argument is based on **dualism**.

From the above argument, we have the following theorem.

Theorem 9.18. [The principle of equal weight] Consider a finite state space Ω, that is, $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$. Let $\mathsf{O} = (X, \mathcal{F}, F)$ be an observable in $L^\infty(\Omega, \nu)$, where ν is the counting measure. Consider a measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[*]})$. If the observer has no information for the state $[*]$, there is a reason to that this measurement is identified with the mixed measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[*]}(w_e))$ $\Big($ or, $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[*]}(\nu_e))$ $\Big)$, where

$$w_e(\omega_k) = 1/n \ \ (\forall k = 1,2,...,n) \qquad \text{or} \qquad \nu_e = \frac{1}{n}\sum_{k=1}^{n} \delta_{\omega_k}$$

Proof. The proof is a easy consequence of the above Monty Hall problem (or, see [29, 32]). □

♠**Note 9.5.** Concerning the principle of equal weight, we deal the following three kinds:

($\sharp_1$) the principle of equal weight in Remark 5.19

($\sharp_2$) the principle of equal weight in Theorem 9.18

($\sharp_3$) the principle of equal weight in Proclaim 19.4

9.8 Averaging information (Entropy)

As one of applications (of Bayes theorem), we now study the "entropy (cf. [69])" of the measurement. This section is due to the following refs.

(♯) Ref. [26]: S. Ishikawa, *A Quantum Mechanical Approach to Fuzzy Theory,* Fuzzy Sets and Systems, Vol. 90, No. 3, 277-306, 1997, doi: 10.1016/S0165-0114(96)00114-5

(♯) Ref. [29]: S. Ishikawa, "Mathematical Foundations of Measurement Theory," Keio University Press Inc. 2006.

Let us begin with the following definition.

Definition 9.19. [Entropy (*cf.* [26, 29])] Assume

$$\text{Classical basic structure } [C_0(\Omega) \subseteq L^\infty(\Omega,\nu) \subseteq B(L^2(\Omega,\nu))]$$

Consider a mixed measurement $\mathsf{M}_{L^\infty(\Omega,\nu)}$ $(\mathsf{O}=(X,2^X,F),\ S_{[*]}(w_0))$ with a countable measured value space $X=\{x_1,x_2,\ldots\}$. The probability $P(\{x_n\})$ that a measured value x_n is obtained by the mixed measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O},S_{[*]}(w_0))$ is given by

$$P(\{x_n\})=\int_\Omega [F(\{x_n\})](\omega)w_0(\omega)\nu(d\omega) \tag{9.24}$$

Further, when a measured value x_n is obtained, the information $I(\{x_n\})$ is, from Bayes' theorem 9.11, is calculated as follows.

$$I(\{x_n\})=\int_\Omega \frac{[F(\{x_n\})](\omega)}{\int_\Omega [F(\{x_n\})](\omega)w_0(\omega)\nu(d\omega)}\log\frac{[F(\{x_n\})](\omega)}{\int_\Omega [F(\{x_n\})](\omega)w_0(\omega)\nu(d\omega)}w_0(\omega)\nu(d\omega)$$

Therefore, the averaging information $H\big(\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O},S_{[*]}(w_0))\big)$ of the mixed measurement $\mathsf{M}_{L^\infty(\Omega)}$ $(\mathsf{O},\ S_{[*]}(w_0))$ is naturally defined by

$$H\big(\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O},S_{[*]}(w_0))\big)=\sum_{n=1}^\infty P(\{x_n\})\cdot I(\{x_n\}) \tag{9.25}$$

Also, the following is clear:

$$\begin{aligned}H\big(\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O},S_{[*]}(w_0))\big)=&\sum_{n=1}^\infty\int_\Omega [F(\{x_n\})](\omega)\log[F(\{x_n\})](\omega)w_0(\omega)\nu(d\omega)\\ &-\sum_{n=1}^\infty P(\{x_n\})\log P(\{x_n\})\end{aligned} \tag{9.26}$$

Example 9.20. [The offender is man or female? fast or slow?] Assume that

(a) There are 100 suspected persons such as $\{s_1, s_2, \ldots, s_{100}\}$, in which there is one criminal.

Define the state space $\Omega = \{\omega_1, \omega_2, \ldots, \omega_{100}\}$ such that

$$\text{state}\omega_n \cdots \text{the state such that suspect } s_n \text{ is a criminal} \qquad (n = 1, 2, ..., 100)$$

Assume the counting measure ν such that $\nu(\{\omega_k\}) = 1 (\forall k = 1, 2, \cdots, 100)$ Define a male-observable $\mathsf{O}_{\rm m} = (X = \{y_{\rm m}, n_{\rm m}\}, 2^X, M)$ in $L^\infty(\Omega)$ by

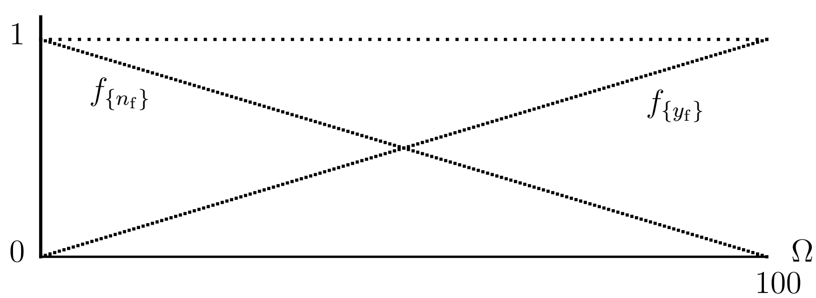
$$[M(\{y_{\rm m}\})](\omega_n) = m_{y_{\rm m}}(\omega_n) = \begin{cases} 0 & (n \text{ is odd}) \\ 1 & (n \text{ is even}) \end{cases}$$
$$[M(\{n_{\rm m}\})](\omega_n) = m_{n_{\rm m}}(\omega_n) = 1 - [M(\{y_{\rm m}\})](\omega_n)$$

For example,

> Taking a measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{\rm m}, S_{[\omega_{17}]})$ — the sex of the criminal s_{17} —, we get the measured value $n_{\rm m}$(=female).

Also, define the fast-observable $\mathsf{O}_{\rm f} = (Y = \{y_{\rm f}, n_{\rm f}\}, 2^Y, F)$ in $L^\infty(\Omega)$ by

$$[F(\{y_{\rm f}\})](\omega_n) = f_{y_{\rm f}}(\omega_n) = \frac{n-1}{99},$$
$$[F(\{n_{\rm f}\})](\omega_n) = f_{n_{\rm f}}(\omega_n) = 1 - [F(\{y_{\rm f}\})](\omega_n)$$



According to the principle of equal weight (=Theorem 9.18), there is a reason to consider that a mixed state w_0 $(\in L^1_{+1}(\Omega))$ is equal to the state w_e such that $w_0(\omega_n) = w_e(\omega_n) = 1/100$ $(\forall n)$. Thus, consider two mixed measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{\rm m}, S_{[*]}(w_e))$ and $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{\rm f}, S_{[*]}(w_e))$. Then, we see:

$$\begin{aligned} H\big(\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{\rm m}, S_{[*]}(w_e))\big) = & \int_\Omega m_{y_{\rm m}}(\omega) w_e(\omega) \nu(d\omega) \cdot \log \int_\Omega m_{y_{\rm m}}(\omega) w_e(\omega) \nu(d\omega) \\ & - \int_\Omega m_{\{n_{\rm m}\}}(\omega) w_e(\omega) \nu(d\omega) \cdot \log \int_\Omega m_{n_{\rm m}}(\omega) w_e(\omega) \nu(d\omega) \end{aligned}$$

$$= \quad -\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2} = \log_2 2 = 1 \ \ (\text{bit})^3.$$

Also,

$$\begin{aligned} H\big(\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{\mathrm{f}}, S_{[*]}(w_e))\big) \ &= \ \int_\Omega f_{y_{\mathrm{f}}}(\omega)\log f_{y_{\mathrm{f}}}(\omega) w_e(\omega)\nu(d\omega) \\ &+ \int_\Omega f_{n_{\mathrm{f}}}(\omega)\log f_{n_{\mathrm{f}}}(\omega) w_e(\omega)\nu(d\omega) - \int_\Omega f_{y_{\mathrm{f}}}(\omega) w_e(\omega)\nu(d\omega)\cdot\log\int_\Omega f_{y_{\mathrm{f}}}(\omega) w_e(\omega)\nu(d\omega) \\ &- \int_\Omega f_{n_{\mathrm{f}}}(\omega) w_e(d\omega)\cdot\log\int_\Omega f_{n_{\mathrm{f}}}(\omega) w_e(\omega)\nu(d\omega) \\ &\fallingdotseq 2\int_0^1 \lambda\log_2\lambda d\lambda + 1 = -\frac{1}{2\log_e 2} + 1 = 0.278\cdots(\text{bit}) \end{aligned}$$

Therefore, as eyewitness information, “male or female” has more valuable than “fast or slow”.

9.9 Fisher statistics:Monty Hall problem [three prisoners problem]

This section is extracted from the following:

Ref. [45]: S. Ishikawa; The Final Solutions of Monty Hall Problem and Three Prisoners Problem (arXiv:1408.0963v1 [stat.OT] 2014)

It is usually said that

Monty Hall problem and three prisoners problem are so-called isomorphism problem

But, we think that the meaning of "isomorphism problem" is not clarified, or, it is not able to be clarified without measurement (or, the dualism).

Therefore, in order to understand "isomorphism", we simultaneously discuss the two

- $\begin{cases} \text{Monty Hall problem} \\ \text{three prisoners problem} \end{cases}$

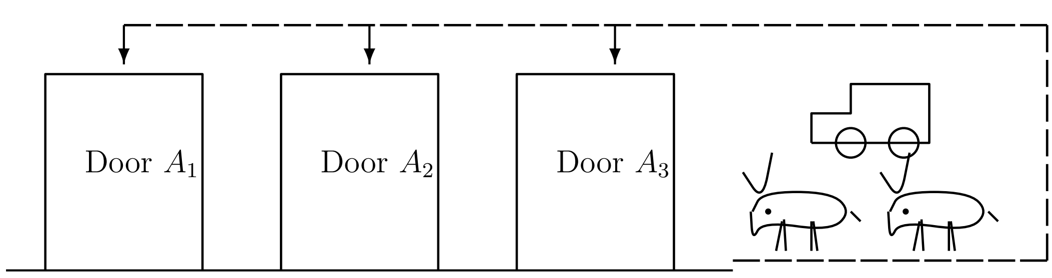
9.9.1 Fisher statistics: Monty Hall problem [resp. three prisoners problem]

Problem 9.21. (=Problem9.15: [Monty Hall problem]).

Suppose you are on a game show, and you are given the choice of three doors (i.e., "Door A_1," "Door A_2," "Door A_3"). Behind one door is a car, behind the others, goats. You do not know what's behind the doors

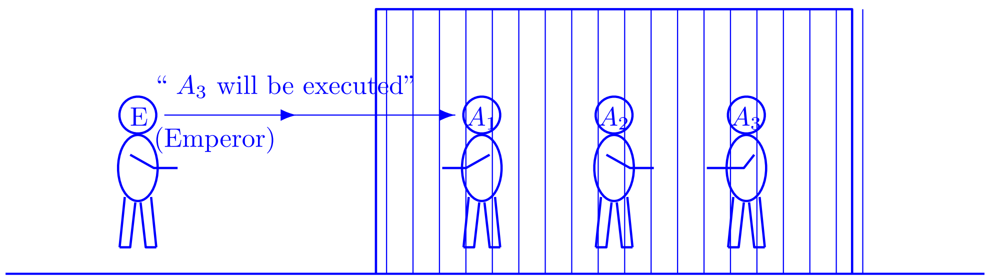
However, you pick a door, say "Door A_1", and the host, who knows what's behind the doors, opens another door, say "Door A_3," which has a goat.

He says to you, "Do you want to pick Door A_2?" Is it to your advantage to switch your choice of doors?



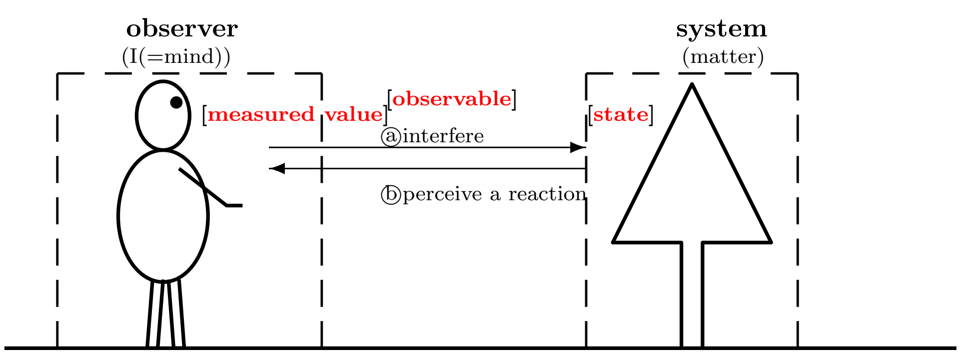
Problem 9.22. [three prisoners problem].

Three prisoners, A_1, A_2, and A_3 were in jail. They knew that one of them was to be set free and the other two were to be executed. They did not know who was the one to be spared, but the emperor did know. A_1 said to the emperor, "I already know that at least one the other two prisoners will be executed, so if you tell me the name of one who will be executed, you won't have given me any information about my own execution". After some thinking, the emperor said, "A_3 will be executed." Thereupon A_1 felt happier because his chance had increased from $\frac{1}{3(=\mathrm{Num}\{A_1,A_2,A_3\}])}$ to $\frac{1}{2(=\mathrm{Num}\{A_1,A_2\}])}$. This prisoner A_1's happiness may or may not be reasonable?



9.9.2 The answer in Fisher statistics: Monty Hall problem [resp. three prisoners problem]

Let rewrite the spirit of dualism (Descartes figure) as follows.



Descartes Figure 9.7: The image of "measurement(=ⓐ+ⓑ)" in dualism

In the dualism, we have the confrontation

$$\text{“observer} \longleftrightarrow \text{system”}$$

as follows.

Table 9.1: Correspondence: observer · system

Problems＼ dualism	Mind(=I=Observer)	Matter(=System)
Monty Hall problem	you	Three doors
Three prisoners problem	Prisoner A_1	Emperor's mind

In what follows, we present the first answer to $\left[\begin{array}{l} \text{Problem 9.21 (Monty-Hall problem)} \\ \text{Problem 9.22 (Three prisoners problem)} \end{array} \right]$ in classical pure measurement theory. The two will be simultaneously solved as follows. The spirit of dualism (in Figure 9.7) urges us to declare that

$$\text{(A)} \left[\begin{array}{l} \text{“observer} \approx \text{you” and “system} \approx \text{three doors” in Problem 9.21} \\ \text{“observer} \approx \text{prisoner } A_1\text{” and “system} \approx \text{emperor's mind” in Problem 9.22} \end{array} \right]$$

Put $\Omega = \{\omega_1, \omega_2, \omega_3\}$ with the discrete topology. Assume that each state $\delta_{\omega_m} (\in \mathfrak{S}^p(C(\Omega)^*))$ means

$$\left[\begin{array}{ll} \delta_{\omega_m} \Leftrightarrow & \text{the state that the car is behind the door } A_m \\ \delta_{\omega_m} \Leftrightarrow & \text{the state that the prisoner } A_m \text{ is will be executed} \end{array} \right] \quad (m = 1, 2, 3) \tag{9.27}$$

Define the observable $\mathsf{O}_1 \equiv (\{1,2,3\}, 2^{\{1,2,3\}}, F_1)$ in $L^\infty(\Omega)$ such that

$$\begin{aligned}
&[F_1(\{1\})](\omega_1) = 0.0, \quad &&[F_1(\{2\})](\omega_1) = 0.5, \quad &&[F_1(\{3\})](\omega_1) = 0.5, \\
&[F_1(\{1\})](\omega_2) = 0.0, \quad &&[F_1(\{2\})](\omega_2) = 0.0, \quad &&[F_1(\{3\})](\omega_2) = 1.0, \\
&[F_1(\{1\})](\omega_3) = 0.0, \quad &&[F_1(\{2\})](\omega_3) = 1.0, \quad &&[F_1(\{3\})](\omega_3) = 0.0,
\end{aligned} \tag{9.28}$$

where it is also possible to assume that $F_1(\{2\})(\omega_1) = \alpha$, $F_1(\{3\})(\omega_1) = 1 - \alpha$ $(0 < \alpha < 1)$. Thus we have a measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_1, S_{[*]})$, which should be regarded as the measurement theoretical representation of the measurement that $\left[\begin{array}{l} \textit{you say “Door } A_1\textit{”} \\ \textit{“Prisoner } A_1\textit{” asks to the emperor} \end{array} \right]$. Here, we assume that

a) “measured value 1 is obtained by the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_1, S_{[*]})$”
$\Leftrightarrow \left[\begin{array}{l} \text{the host says “Door } A_1 \text{ has a goat”} \\ \text{the emperor says “Prisoner } A_1 \text{ will be executed”} \end{array} \right]$

b) “measured value 2 is obtained by the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_1, S_{[*]})$ ”
$\Leftrightarrow \left[\begin{array}{l} \text{the host says “Door } A_2 \text{ has a goat”} \\ \text{the emperor says “Prisoner } A_2 \text{ will be executed”} \end{array} \right]$

c) "measured value 3 is obtained by the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_1, S_{[*]})$ "
$\Leftrightarrow \left[\begin{array}{l}\text{the host says "Door } A_3 \text{ has a goat"} \\ \text{the emperor says "Prisoner } A_3 \text{ will be executed"}\end{array}\right]$

Recall that $\left[\begin{array}{c}\text{the host said "Door 3 has a goat"} \\ \text{the emperor said "Prisoner } A_3 \text{ will be executed"}\end{array}\right]$.

This implies that $\left[\begin{array}{l}\text{you} \\ \text{Prisoner } A_1\end{array}\right]$ get the measured value "3" by the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_1, S_{[*]})$. Note that

$$\begin{aligned}[F_1(\{3\})](\omega_2) &= 1.0 = \max\{0.5, \ \ 1.0, \ \ 0.0\} \\ &= \max\{[F_1(\{3\})](\omega_1), [F_1(\{3\})](\omega_2), [F_1(\{3\})](\omega_3)\},\end{aligned} \tag{9.29}$$

Therefore, Theorem 5.6 (Fisher's maximum likelihood method) says that

(B$_1$) In Problem 9.21 (Monty-Hall problem), there is a reason to infer that $[*] = \delta_{\omega_2}$. Thus, you should switch to Door A_2.

(B$_2$) In Problem 9.22 (Three prisoners problem), there is a reason to infer that $[*] = \delta_{\omega_2}$. However, there is no reasonable answer for the question: whether Prisoner A_1's happiness increases. That is, Problem 9.22 is not within Fisher's maximum likelihood method.

9.10 Bayesian statistics: Monty Hall problem [three prisoners problem]

This section is extracted from the following:

Ref. [45]: S. Ishikawa; The Final Solutions of Monty Hall Problem and Three Prisoners Problem (arXiv:1408.0963v1 [stat.OT] 2014)

9.10.1 Bayesian statistics: Monty Hall problem [resp. three prisoners problem]

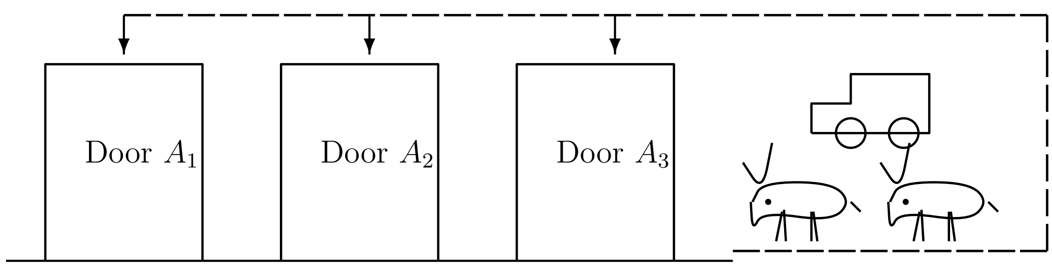
Problem 9.23. [(=Problem9.16)Monty Hall problem (the case that the host throws the dice)].

Suppose you are on a game show, and you are given the choice of three doors (i.e., "Door A_1," "Door A_2," "Door A_3"). Behind one door is a car, behind the others, goats. You do not know what's behind the doors.

However, you pick a door, say "Door A_1", and the host, who knows what's behind the doors, opens another door, say "Door A_3," which has a goat. And he adds that

($\sharp_1$) *the car was set behind the door decided by the cast of the (distorted) dice. That is, the host set the car behind Door A_m with probability p_m (where $p_1 + p_2 + p_3 = 1$, $0 \le p_1, p_2, p_3 \le 1$).*

He says to you, "Do you want to pick Door A_2?" Is it to your advantage to switch your choice of doors?

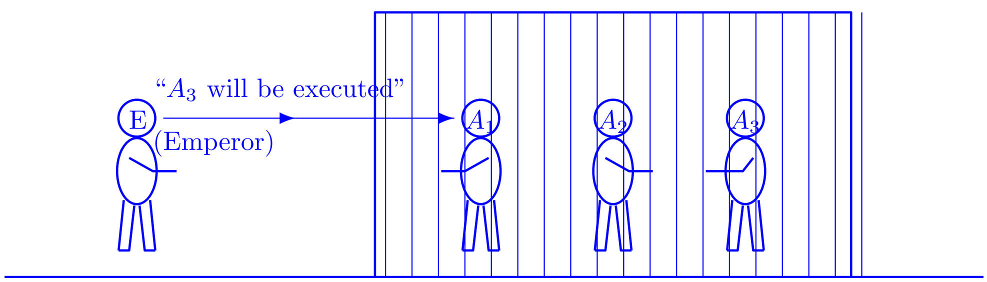


Problem 9.24. [three prisoners problem].

Three prisoners, A_1, A_2, and A_3 were in jail. They knew that one of them was to be set free and the other two were to be executed. They did not know who was the one to be spared, but they know that

($\sharp_2$) *the one to be spared was decided by the cast of the (distorted) dice. That is, Prisoner A_m is to be spared with probability p_m (where $p_1+p_2+p_3=1$, $0 \le p_1, p_2, p_3 \le 1$).*

but the emperor did know the one to be spared. A_1 said to the emperor, "I already know that at least one the other two prisoners will be executed, so if you tell me the name of one who will be executed, you won't have given me any information about my own execution". After some thinking, the emperor said, "A_3 will be executed." Thereupon A_1 felt happier because his chance had increased from $\frac{1}{3(=\mathrm{Num}[\{A_1,A_2,A_3\}])}$ to $\frac{1}{2(=\mathrm{Num}[\{A_1,A_2\}])}$. This prisoner A_1's happiness may or may not be reasonable?



9.10.2 The answer in Bayesian statistics: Monty Hall problem [resp. three prisoners problem]

In the dualism, we have the confrontation

$$\text{"observer} \longleftrightarrow \text{system"}$$

as follows.

Table 9.2: Correspondence: observer · system

Problems＼ dualism	Mind(=I=Observer)	Matter(=System)
Monty Hall problem	you	Three doors
Three prisoners problem	Prisoner A	Emperor's mind

In what follows we study these problems. Let Ω and O_1 be as in Section 9.8. Under the hypothesis $\left\{ \begin{array}{c} (\sharp_1) \\ (\sharp_2) \end{array} \right\}$, define the mixed state ν_0 ($\in \mathcal{M}_{+1}^m(\Omega)$) such that:

$$\nu_0(\{\omega_1\}) = p_1, \quad \nu_0(\{\omega_2\}) = p_2, \quad \nu_0(\{\omega_3\}) = p_3 \tag{9.30}$$

Thus we have a mixed measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_1, S_{[*]}(\nu_0))$. Note that

a) "measured value 1 is obtained by the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_1, S_{[*]})$"
$\Leftrightarrow \left[\begin{array}{l}\text{the host says "Door } A_1 \text{ has a goat"} \\ \text{the emperor says "Prisoner } A_1 \text{ will be executed"}\end{array}\right]$

b) "measured value 2 is obtained by the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_1, S_{[*]})$ "
$\Leftrightarrow \left[\begin{array}{l}\text{the host says "Door } A_2 \text{ has a goat"} \\ \text{the emperor says "Prisoner } A_2 \text{ will be executed"}\end{array}\right]$

c) "measured value 3 is obtained by the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_1, S_{[*]})$ "
$\Leftrightarrow \left[\begin{array}{l}\text{the host says "Door } A_3 \text{ has a goat"} \\ \text{the emperor says "Prisoner } A_3 \text{ will be executed"}\end{array}\right]$

Here, assume that, by the statistical measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_1, S_{[*]}(\nu_0))$, you obtain a measured value 3, which corresponds to the fact that $\left[\begin{array}{l}\text{the host said "Door } A_3 \text{ has a goat"} \\ \text{the emperor said "Prisoner } A_3 \text{ is to be executed"}\end{array}\right]$
Then, Bayes' theorem 9.11 says that the posterior state ν_{post} $(\in \mathcal{M}_{+1}^m(\Omega))$ is given by

$$\nu_{\text{post}} = \frac{F_1(\{3\}) \times \nu_0}{\langle \nu_0, F_1(\{3\}) \rangle}. \tag{9.31}$$

That is,

$$\nu_{\text{post}}(\{\omega_1\}) = \frac{\frac{p_1}{2}}{\frac{p_1}{2} + p_2}, \quad \nu_{\text{post}}(\{\omega_2\}) = \frac{p_2}{\frac{p_1}{2} + p_2}, \quad \nu_{\text{post}}(\{\omega_3\}) = 0. \tag{9.32}$$

Then,

(I1) In Problem 9.23,

$$\begin{cases} \text{if } \nu_{\text{post}}(\{\omega_1\}) < \nu_{\text{post}}(\{\omega_2\}) \text{ (i.e., } p_1 < 2p_2\text{), you should pick Door } A_2 \\ \text{if } \nu_{\text{post}}(\{\omega_1\}) = \nu_{\text{post}}(\{\omega_2\}) \text{ (i.e., } p_1 = 2p_2\text{), you may pick Doors } A_1 \text{ or } A_2 \\ \text{if } \nu_{\text{post}}(\{\omega_1\}) > \nu_{\text{post}}(\{\omega_2\}) \text{ (i.e., } p_1 > 2p_2\text{), you should not pick Door } A_2 \end{cases}$$

(I2) In Problem 9.24,

$$\begin{cases} \text{if } \nu_0(\{\omega_1\}) < \nu_{\text{post}}(\{\omega_1\}) \text{ (i.e., } p_1 < 1 - 2p_2\text{), the prisoner } A_1\text{'s happiness increases} \\ \text{if } \nu_0(\{\omega_1\}) = \nu_{\text{post}}(\{\omega_1\}) \text{ (i.e., } p_1 = 1 - 2p_2\text{), the prisoner } A_1\text{'s happiness is invariant} \\ \text{if } \nu_0(\{\omega_1\}) > \nu_{\text{post}}(\{\omega_1\}) \text{ (i.e., } p_1 > 1 - 2p_2\text{), the prisoner } A_1\text{'s happiness decreases} \end{cases}$$

9.11 Equal probability: Monty Hall problem [three prisoners problem]

This section is extracted from the following:

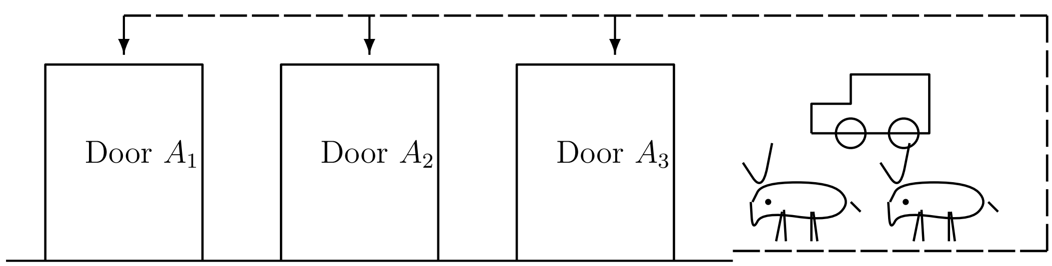
ref. [45]: S. Ishikawa; The Final Solutions of Monty Hall Problem and Three Prisoners Problem (arXiv:1408.0963v1 [stat.OT] 2014)

Problem 9.25. [(=Problem9.16)Monty Hall problem (the case that you throws the dice)].

Suppose you are on a game show, and you are given the choice of three doors (i.e., "Door A_1," "Door A_2," "Door A_3"). Behind one door is a car, behind the others, goats. You do not know what's behind the doors. Thus,

($\sharp_1$) *you select Door A_1 by the cast of the fair dice. That is, you say "Door A_1" with probability 1/3.*

The host, who knows what's behind the doors, opens another door, say "Door A_3," which has a goat. He says to you, "Do you want to pick Door A_2?" Is it to your advantage to switch your choice of doors?

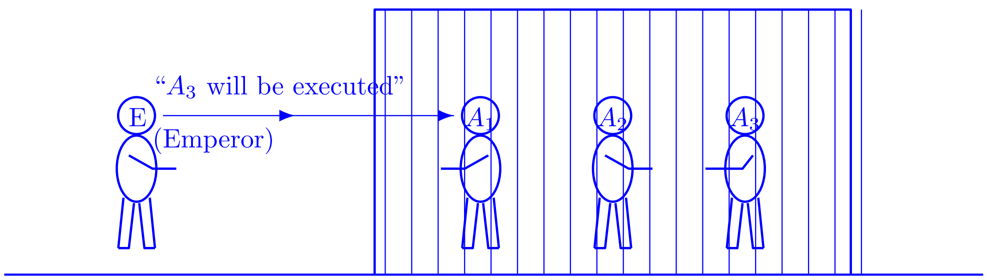


Problem 9.26. [three prisoners problem(the case that the prisoner throws the dice)].

Three prisoners, A_1, A_2, and A_3 were in jail. They knew that one of them was to be set free and the other two were to be executed. They did not know who was the one to be spared, but the emperor did know. Since three prisoners wanted to ask the emperor,

($\sharp_2$) *the questioner was decided by the fair die throw. And Prisoner A_1 was selected with probability* 1/3

Then, A_1 said to the emperor, "I already know that at least one the other two prisoners will be executed, so if you tell me the name of one who will be executed, you won't have given me any information about my own execution". After some thinking, the emperor said, "A_3 will be executed." Thereupon A_1 felt happier because his chance had increased from $\frac{1}{3(=\mathrm{Num}[\{A_1,A_2,A_3\}])}$ to $\frac{1}{2(=\mathrm{Num}[\{A_1,A_2\}])}$. This prisoner A_1's happiness may or may not be reasonable?



Answer By Theorem 9.18(The principle of equal weight), the above Problems 9.25 and 9.26 is respectively the same as Problems 9.23 and 9.24 in the case that $p_1 = p_2 = p_3 = 1/3$. Then, the formulas (9.30) and (9.32) say that

(A_1) In Problem9.25, since $\nu_{\mathrm{post}}(\{\omega_1\}) = 1/3 < 2/3 = \nu_{\mathrm{post}}(\{\omega_2\})$, you should pick Door A_2.

(A_2) In Problem9.26, since $\nu_0(\{\omega_1\}) = 1/3 = \nu_{\mathrm{post}}(\{\omega_1\})$, the prisoner A_1's happiness is invariant.

Therefore,

(B_1) Problem9.25 [Monty Hall problem (the case that you throw a fair dice)]

$$\nu_{\mathrm{post}}(\{\omega_1\}) < \nu_{\mathrm{post}}(\{\omega_2\}) \text{ (i.e., } p_1 = 1/3 < 2/3 = 2p_2),$$

thus, you should choose the door A_2

(B_2) Problem9.26 [three prisoners problem (the case that the emperor throws a fair dice)],

$$\nu_0(\{\omega_1\}) = \nu_{\mathrm{post}}(\{\omega_1\}) \text{ (i.e., } p_1 = 1/3 = 1 - 2p_2),$$

Thus, the happiness of the prisoner A_1 is invariant

♠**Note 9.6.** These problems (i.e., Monty Hall problem and the three prisoners problem) continued attracting the philosopher's interest. This is not due to that these are easy to make a mistake for high school students, but

these problems include the essence of "dualism".

9.12 Bertrand's paradox("randomness" depends on how you look at)

Theorem9.18(the principle of equal weight) implies that

- the "randomness" may be related to the invariant probability measure.

However, this is due to the finiteness of the state space. In the case of infinite state space,

"randomness" depends on how you look at

This is explained in this section.

9.12.1 Bertrand's paradox("randomness" depends on how you look at)

Let us explain Bertrand's paradox as follows.

Consider classical basic structure:

$$[C_0(\Omega) \subseteq L^\infty(\Omega, m) \subseteq B(L^2(\Omega, m))]$$

We can define the exact observable $\mathsf{O}_E = (\Omega, \mathcal{B}_\Omega, F_E)$ in $L^\infty(\Omega, m)$ such that

$$[F_E(\Xi)](\omega) = \chi_\Xi(\omega) = \begin{cases} 1 & (\omega \in \Xi) \\ 0 & (\omega \notin \Xi) \end{cases}$$
$$(\forall \omega \in \Omega, \quad \Xi \in \mathcal{B}_\Omega)$$

Here, we have the following problem:

(A) Can the measurement $\mathsf{M}_{L^\infty(\Omega,m)}(\mathsf{O}_E, S_{[*]}(\rho))$ that represents "at random" be determined uniquely?

This question is of course denied by so-called Bertrand paradox. Here, let us review the argument about the Bertrand paradox (*cf.* [21, 29, 43]). Consider the following problem:

Problem 9.27. (Bertrand paradox) Given a circle with the radius 1. Suppose a chord of the circle is chosen **at random**. What is the probability that the chord is shorter than $\sqrt{3}$?

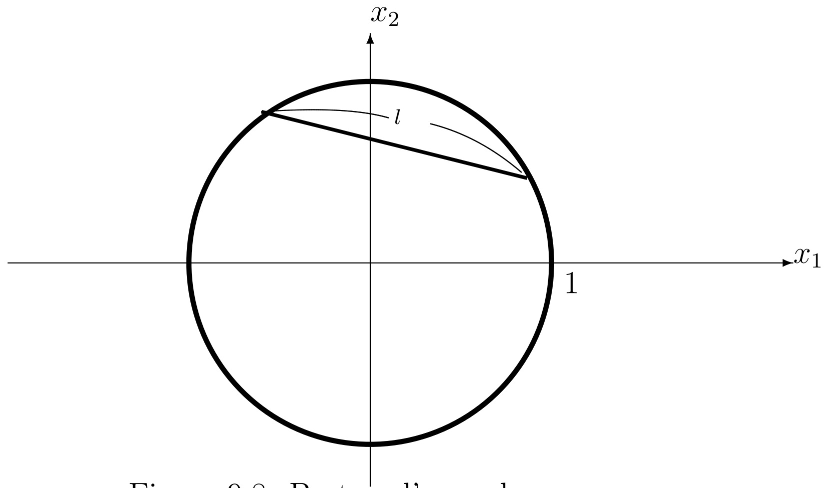


Figure 9.8: Bertrand' paradox

Define the rotation map $T_{\rm rot}^{\theta} : \mathbb{R}^2 \to \mathbb{R}^2$ $(0 \le \theta < 2\pi)$ and the reverse map $T_{\rm rev} : \mathbb{R}^2 \to \mathbb{R}^2$ such that

$$T_{\rm rot}^{\theta} x = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad T_{\rm rev} x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Problem 9.28. (Bertrand paradox and its answer) Given a circle with the radius 1.

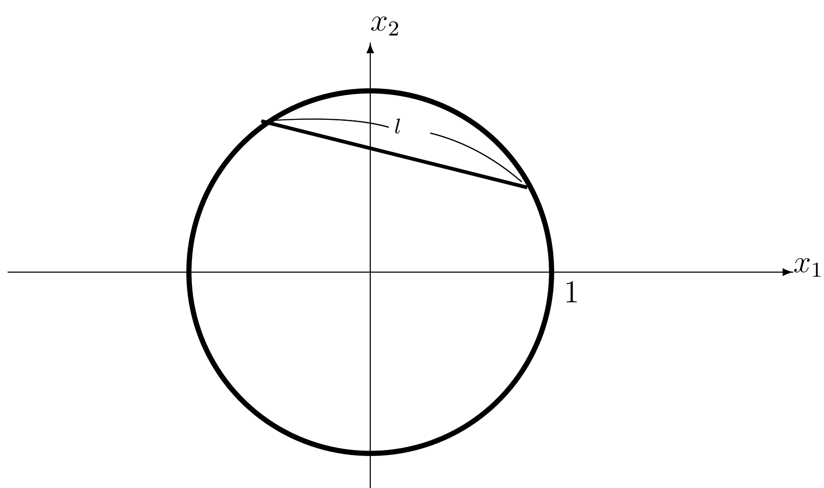


Figure 9.9: Bertrand' paradox

Put $\Omega = \{ l \mid l \text{ is a chord}\}$, that is, **the set of all chords**.

(B) Can we uniquely define an invariant probability measure on Ω?

Here, "invariant" means "invariant concerning the rotation map $T_{\rm rot}^{\theta}$ and reverse map $T_{\rm rev}$".

In what follows, we show that the above invariant measure exists but it is not determined uniquely.

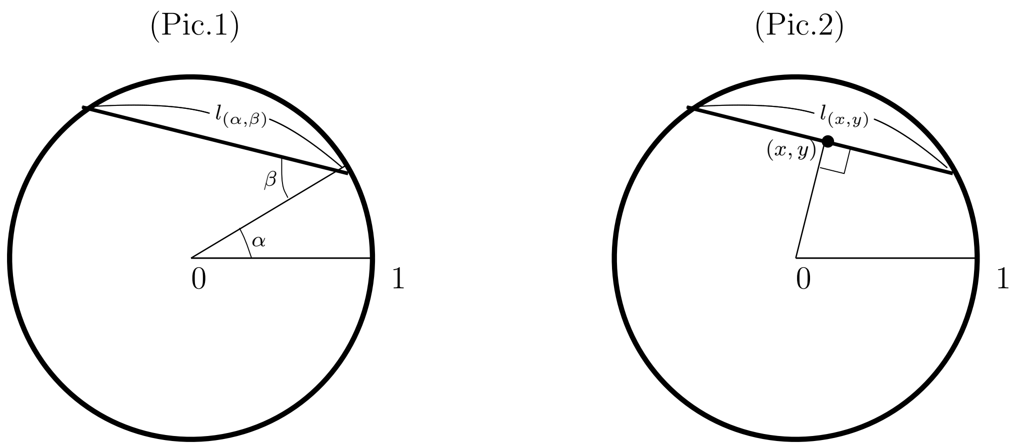


Figure 9.10: Two cases in Bertrand' paradox

[The first answer (Pic.1(in Figure 9.10))]. In Pic.1, we see that the chord l is represented by a point (α, β) in the rectangle $\Omega_1 \equiv \{(\alpha,\beta) \mid 0 < \alpha \le 2\pi, \ 0 < \beta \le \pi/2(\text{radian})\}$. That is, we have the following identification:

$$\Omega(= \text{the set of all chords}) \ni l_{(\alpha,\beta)} \underset{\text{identification}}{\longleftrightarrow} (\alpha,\beta) \in \Omega_1(\subset \mathbb{R}^2).$$

Note that we have the natural probability measure nu_1 on Ω_1 such that $\nu_1(A) = \frac{\text{Meas}[A]}{\text{Meas}[\Omega_1]} = \frac{\text{Meas}[A]}{\pi^2}$ $(\forall A \in \mathcal{B}_{\Omega_1})$, where " Meas" = " Lebesgue measure". Transferring the probability measure ν_1 on Ω_1 to Ω, we get ρ_1 on Ω. That is,

$$\mathcal{M}_{+1}(\Omega) \ni \rho_1 \underset{\text{identification}}{\longleftrightarrow} \nu_1 \in \mathcal{M}_{+1}(\Omega_1)$$

($\sharp$) It is clear that the measure ρ_1 is invariant concerning the rotation map T_{rot}^{θ} and reverse map T_{rev}.

Therefore, we have a natural measurement $\mathsf{M}_{L^\infty(\Omega,m)}(\mathsf{O}_E \equiv (\Omega, \mathcal{B}_\Omega, F_E), S_{[*]}(\rho_1))$. Consider the identification:

$$\Omega \supseteq \Xi_{\sqrt{3}} \underset{\text{identification}}{\longleftrightarrow} \{(\alpha,\beta) \in \Omega_1 \ : \ \text{"the length of } l_{(\alpha,\beta)}\text{"} < \sqrt{3}\} \subseteq \Omega_1$$

Then, Axiom$^{(\text{m})}$ 1 says that the probability that a measured value belongs to $\Xi_{\sqrt{3}}$ is given by

$$\begin{aligned}
&\int_\Omega [F_E(\Xi_{\sqrt{3}})](\omega) \ \rho_1(d\omega) = \int_{\Xi_{\sqrt{3}}} 1 \ \rho_1(d\omega) \\
=&m_1(\{l_{(\alpha,\beta)} \approx (\alpha,\beta) \in \Omega_1 \mid \ \text{"the length of } l_{(\alpha,\beta)}\text{"} \le \sqrt{3}\}) \\
=&\frac{\text{Meas}[\{(\alpha,\beta) \mid 0 \le \alpha \le 2\pi, \ \pi/6 \le \beta \le \pi/2\}]}{\text{Meas}[\{(\alpha,\beta) \mid 0 \le \alpha \le 2\pi, \ 0 \le \beta \le \pi/2\}]}
\end{aligned}$$

$$=\frac{2\pi \times (\pi/3)}{\pi^2} = \frac{2}{3}.$$

[The second answer (Pic.2(in Figure 9.10))]. In Pic.2, we see that the chord l is represented by a point (x, y) in the circle $\Omega_2 \equiv \{(x, y) \mid x^2 + y^2 < 1\}$.

That is, we have the following identification:

$$\Omega(= \text{the set of all chords}) \ni l_{(x,y)} \underset{\text{identification}}{\longleftrightarrow} (x, y) \in \Omega_2(\subset \mathbb{R}^2).$$

We have the natural probability measure ν_2 on Ω_2 such that $\nu_2(A) = \frac{\text{Meas}[A]}{\text{Meas}[\Omega_2]} = \frac{\text{Meas}[A]}{\pi}$ $(\forall A \in \mathcal{B}_{\Omega_2})$. Transferring the probability measure ν_2 on Ω_2 to Ω, we get ρ_2 on Ω. That is,

$$\mathcal{M}_{+1}(\Omega) \ni \rho_2 \underset{\text{identification}}{\longleftrightarrow} \nu_2 \in \mathcal{M}_{+1}(\Omega_2)$$

(♯) It is clear that the measure ρ_2 is invariant concerning the rotation map T^θ_{rot} and reverse map T_{rev}.

Therefore, we have a natural measurement $\mathsf{M}_{L^\infty(\Omega,m)}(\mathsf{O}_E \equiv (\Omega, \mathcal{B}_\Omega, F_E), S_{[*]}(\rho_2))$.

Consider the identification:

$$\Omega \supseteq \Xi_{\sqrt{3}} \underset{\text{identification}}{\longleftrightarrow} \{(x, y) \in \Omega_2 \ : \ \text{“the length of } l_{(\alpha,\beta)}\text{”} < \sqrt{3}\} \subseteq \Omega_1$$

Then, Axiom$^{(\text{m})}$ 1 says that the probability that a measured value belongs to $\Xi_{\sqrt{3}}$ is given by

$$\begin{aligned}
&\int_\Omega [F_E(\Xi_{\sqrt{3}})](\omega)\ \rho_2(d\omega) = \int_{\Xi_{\sqrt{3}}} 1 \ \ \rho_2(d\omega) \\
=&\nu_2(\{l_{(x,y)} \approx (x, y) \in \Omega_2 \mid \ \text{“the length of } l_{(x,y)}\text{”} \le \sqrt{3}\}) \\
=&\frac{\text{Meas}[\{(x, y) \mid 1/4 \le x^2 + y^2 \le 1\}]}{\pi} = \frac{3}{4}.
\end{aligned}$$

Conclusion 9.29. Thus, even if there is a custom to regard a natural probability measure (i.e., an invariant measure concerning natural maps) as “random”, the first answer and the second answer say that

(♯) **the uniqueness in (B) of Problem 9.28 is denied.**

Chapter 10

Axiom 2—causality

Measurement theory has the following classification:

$$
\text{(A)} \quad \underset{\text{(=quantum language)}}{\text{measurement theory}} \begin{cases} \underset{(\mathrm{A}_1)}{\text{pure type}} \begin{cases} \text{classical system : Fisher statistics} \\ \text{quantum system : usual quantum mechanics} \end{cases} \\ \\ \underset{(\mathrm{A}_2)}{\text{mixed type}} \begin{cases} \text{classical system : including Bayesian statistics, Kalman filter} \\ \text{quantum system : quantum decoherence} \end{cases} \end{cases}
$$

This is formulated as follows.

$$
\text{(B)} \begin{cases} (\mathrm{B}_1) : \underset{\text{(=quantum language)}}{\boxed{\text{pure measurement theory}}} \\ \quad := \underbrace{\overset{\text{[(pure)Axiom 1]}}{\underset{(cf.\ \S 2.7)}{\boxed{\text{pure measurement}}}} + \overset{\text{[Axiom 2]}}{\underset{(cf.\ \S 10.3)}{\boxed{\text{Causality}}}}}_{\text{a kind of spell(a priori judgment)}} + \underbrace{\overset{\text{[quantum linguistic interpretation]}}{\underset{(cf.\ \S 3.1)}{\boxed{\text{Linguistic interpretation}}}}}_{\text{the manual to use spells}} \\ \\ (\mathrm{B}_2) : \underset{\text{(=quantum language)}}{\boxed{\text{mixed measurement theory}}} \\ \quad := \underbrace{\overset{\text{[(mixed)Axiom}^{(m)}\text{ 1]}}{\underset{(cf.\ \S 9.1)}{\boxed{\text{mixed measurement}}}} + \overset{\text{[Axiom 2]}}{\underset{(cf.\ \S 10.3)}{\boxed{\text{Causality}}}}}_{\text{a kind of spell(a priori judgment)}} + \underbrace{\overset{\text{[quantum linguistic interpretation]}}{\underset{(cf.\ \S 3.1)}{\boxed{\text{Linguistic interpretation}}}}}_{\text{the manual to use spells}} \end{cases}
$$

In this chapter, we devote ourselves to the last theme $\overset{\text{[Axiom 2]}}{\underset{(cf.\ \S 10.3)}{\boxed{\text{Causality}}}}$, which is common to both (B_1) and (B_2).

10.1 The most important unsolved problem—what is causality?

The importance of "measurement" and "causality" should be reconfirmed in the following famous maxims:

(C_1) There is no science without measurement.

(C_2) Science is the knowledge about causal relationship.

They should be also regarded as one of the linguistic interpretation in a wider sense.

10.1.1 Modern science started from the discovery of "causality."

When a certain thing happens, the cause always exists. This is called **causality**. You should just remember the proverb of

"smoke is not located on the place which does not have fire."

It is not so simple although you may think that it is natural. For example, if you consider

> This morning I feel good. Is it because that I slept sound yesterday? or is it because I go to favorite golf from now on?

you may be able to understand the difficulty of how to use the word "causality". In daily conversation, it is used in many cases, mixing up "a cause (past)", "a reason (connotation)", and "the purpose and a motive (future)."

It may be supposed that the pioneers of research of movement and change are

$$\begin{cases} \text{Heraclitus(BC.540 -BC.480): "Everything changes."} \\ \\ \text{Parmenides (born around BC. 515): "Movement does not exist."} \\ \text{(Zeno's teacher)} \end{cases}$$

though their assertions are not clear. However, these two pioneers (i.e., Heraclitus and Parmenides) noticed first that "movement and change" were the primary importance keywords in science(= "world description") , i.e., it is

> **[The beginning of World description]**
>
> $=$**[The discovery of movement and change]** $= \begin{cases} \text{Heraclitus(BC.540 -BC.480)} \\ \\ \text{Parmenides(born around BC. 515)} \end{cases}$

However, Aristotle(BC384–BC322) further investigated about the essence of movement and change, and he thought that

all the movements had the "purpose."

For example, supposing a stone falls, that is because the stone has the purpose that the stone tries to go downward. Supposing smoke rises, that is because smoke has the purpose that smoke rises upwards. Under the influence of Aristotle, "**Purpose**" continued remaining as a mainstream idea of "Movement" for a long time of 1500 years or more.

Although "the further investigation" of Aristotle was what should be praised, it was not able to be said that "the purpose was to the point." In order to free ourselves from Purpose and for human beings to discover that the essence of movement and change is "causal relationship", we had to wait for the appearance of Galileo, Bacon, Descartes, Newton, etc.

Revolution to "Causality" from "Purpose"

is the greatest history-of-science top paradigm shift. It is not an overstatement even if we call it "**birth of modern science**".

$$
\underset{\text{(Heraclitus, Parmenides, Zeno)}}{\overset{\text{the birth of world description}}{\boxed{\text{Movement}}}}
\xrightarrow[\text{Aristotle :(About 1500 years)}]{\text{“purpose”}}
\underset{\text{(Galileo, Bacon, Descartes, Newton)}}{\overset{\text{the birth of modern science}}{\boxed{\text{Causality}}}}
$$

♠**Note 10.1.** I cannot emphasize too much the importance of the discovery of the term: "causality". That is,

(♯) Science is the discipline about phenomena can be represented by the term "causality". (i.e., "No smoke without fire")

Thus, I consider that the discovery of "causality" is equal to that of science.

10.1.2 Four answers to "what is causality?"

As mentioned above, about "what is an essence of movement and change?", it was once settled with the word "causality." However, not all were solved now. We do not yet understand "causality" fully. In fact,

Problem 10.1. Problem:

"What is causality?"

is the most important outstanding problems in modern science.

Answer this problem!

There may be some readers who are surprised with saying like this, although it is the outstanding problems in the present. Below, I arrange the history of the answer to this problem.

(a) **[Realistic causality]:** Newton advocated the realistic describing method of Newtonian mechanics as a final settlement of accounts of ideas, such as Galileo, Bacon, and Descartes, and he thought as follows. :

> "Causality" actually exists in the world. Newtonian equation described faithfully this "causality". That is, Newtonian equation is the equation of a causal chain.

This realistic causality may be a very natural idea, and you may think that you cannot think in addition to this. In fact, probably, we may say that the current of the realistic causal relationship which continues like

"Newtonian mechanics$\longrightarrow$ Electricity and magnetism$\longrightarrow$ Theory of relativity$\longrightarrow$ $\cdots$"

is a scientific flower.

However, there are also other ideas, i.e., three "non-realistic causalities" as follows.

(b) **[Cognitive causality]:** David Hume, Immanuel Kant, etc. who are philosophers thought as follows. :

> We can not say that "Causality" actually exists in the world, or that it does not exist in the world. And when we think that "something" in the world is "causality", we should just believe that the it has "causality".

Most readers may regard this as "a kind of rhetoric", however, several readers may be convinced in "Now that you say that, it may be so." Surely, since you are looking through the prejudice "causality", you may look such. This is Kant's famous "Copernican revolution", that is,

"recognition constitutes the world."

which is considered that the recognition circuit of causality is installed in the brain, and when it is stimulated by "something" and reacts, "there is causal relationship." Probably, many readers doubt about the substantial influence which this (b) had on the science after it. However, in this book, I adopted the friendly story to the utmost to Kant.

(c) **[Mathematical causality(Dynamical system theory)]:** Since dynamical system theory has developed as the mathematical technique in engineering, they have not investigated "What is causality?" thoroughly. However,

In dynamical system theory, we start from the state equation (i.e., simultaneous ordinary differential equation of the first order) such that

$$\begin{cases} \frac{d\omega_1}{dt}(t) = v_1(\omega_1(t), \omega_2(t), \ldots, \omega_n(t), t) \\ \frac{d\omega_2}{dt}(t) = v_2(\omega_1(t), \omega_2(t), \ldots, \omega_n(t), t) \\ \cdots\cdots \\ \frac{d\omega_n}{dt}(t) = v_n(\omega_1(t), \omega_2(t), \ldots, \omega_n(t), t) \end{cases} \tag{10.1}$$

and, we think that

($\sharp$) the phenomenon described by the state equation has "causality."

This is the spirit of dynamical system theory (= statistics). Although this is proposed under the confusion of mathematics and world description, it is quite useful. In this sense, I think that (c) should be evaluated more.

(d) **[Linguistic causal relationship (MeasurementTheory)]:** The causal relationship of measurement theory is decided by the Axiom 2 (causality; §10.3) of this chapter. If I say in detail,:

Although measurement theory consists of the two Axioms 1 and 2, it is the Axiom 2 that is concerned with causal relationship. When describing a certain phenomenon in quantum language (i.e., a language called measurement theory) and using Axiom 2 (causality; §10.3) , we think that the phenomenon has causality.

Summary 10.2. The above is summarized as follows.

(a) World is first
(b) Recognition is first
(c) Mathematics(buried into ordinary language) is first
(d) Language (= quantum language) is first

Now, in measurement theory, we assert the next as said repeatedly:

Quantum language is a basic language which describes various sciences.

Supposing this is recognized, we can assert the next. Namely,

KSTS/RR-16/001
January 8, 2016

In science, causality is just as mentioned in the above (d).

This is my answer to "What is causality ?". I explain this in detail in the following.

♠**Note 10.2.** Consider the following problems:

($\sharp_1$) What is time (space, causality, probability, etc.) ?

There are two ways to answer.

$$(\sharp_2)\quad \text{The answer of "What is XX ?"} \begin{cases} \text{(a): To show the definition of XX} \\ \\ \text{(b): To show how to use the term "XX"} \end{cases}$$

In this note, the answer to the question ($\sharp_1$) is presented from the linguistic point of view (b).

10.2 Causality—Mathematical preparation

10.2.1 The Heisenberg picture and the Schrödinger picture

First, let us review the general basic structure (*cf.* §2.1.3) as follows.

(A): General basic structure and State spaces

$$
\begin{tikzcd}
\underset{C^*\text{-pure state}}{\mathfrak{S}^p(\mathcal{A}^*)} \arrow[r, phantom, "\subset"] & \underset{C^*\text{-mixed state}}{\mathfrak{S}^m(\mathcal{A}^*)} \arrow[r, phantom, "\subset"] & \mathcal{A}^* & & \\
& & \boxed{\mathcal{A}} \arrow[u, "\text{dual}"'] \arrow[r, "\subseteq", "\text{subalgebra}\cdot\text{weak-closure}"'] & \boxed{\overline{\mathcal{A}}} \arrow[d, "\text{pre-dual}"] \arrow[r, "\subseteq", "\text{subalgebra}"'] & \boxed{B(H)} \\
& & \underset{W^*\text{-mixed state}}{\overline{\mathfrak{S}}^m(\overline{\mathcal{A}}_*)} \arrow[r, phantom, "\subset"] & \overline{\mathcal{A}}_* &
\end{tikzcd}
\tag{10.2}
$$

Remark 10.3. $[\overline{\mathcal{A}}_* \subseteq \mathcal{A}^*]$: Consider the basic structure $[\mathcal{A} \subseteq \overline{\mathcal{A}}]_{B(H)}$. For each $\rho \in \overline{\mathcal{A}}_*$, $F \in \mathcal{A}(\subseteq \overline{\mathcal{A}} \subseteq B(H))$, we see that

$$\Big| {}_{\overline{\mathcal{A}}_*}\Big(\rho, F\Big)_{\overline{\mathcal{A}}} \Big| \leq C\|F\|_{B(H)} = C\|F\|_{\mathcal{A}} \tag{10.3}$$

Thus, we can consider that $\rho \in \mathcal{A}^*$. That is, in the sense of (10.3), we consider that

$$\overline{\mathcal{A}}_* \subseteq \mathcal{A}^*$$

When $\rho(\in \overline{\mathcal{A}}_*)$ is regarded as the element of $\mathcal{A}^*$, it is sometimes denoted by $\widehat{\rho}$. Therefore,

$${}_{\overline{\mathcal{A}}_*}\Big(\rho, F\Big)_{\overline{\mathcal{A}}} = {}_{\mathcal{A}^*}\Big(\widehat{\rho}, F\Big)_{\mathcal{A}} \qquad (\forall F \in \mathcal{A}(\subseteq \overline{\mathcal{A}})) \tag{10.4}$$

Definition 10.4. [Causal operator (= Markov causal operator)] Consider two basic structures:

$$[\mathcal{A}_1 \subseteq \overline{\mathcal{A}}_1 \subseteq B(H_1)] \text{ and } [\mathcal{A}_2 \subseteq \overline{\mathcal{A}}_2 \subseteq B(H_2)]$$

A continuous linear operator $\Phi_{1,2} : \overline{\mathcal{A}}_2 \to \overline{\mathcal{A}}_1$ is called a **causal operator**(or, Markov causal operator , the Heisenberg picture of "causality"), if it satisfies the following (i)—(iv):

(i) $F_2 \in \overline{\mathcal{A}}_2 \;\; F_2 \geqq 0 \Longrightarrow \Phi_{12}F_2 \geqq 0$

(ii) $\Phi_{12} I_{\overline{\mathcal{A}}_2} = I_{\overline{\mathcal{A}}_1}$ (where, $I_{\overline{\mathcal{A}}_1}(\in \overline{\mathcal{A}}_1)$ is the identity)

(iii) there exists the continuous linear operator $(\Phi_{1,2})_* : (\overline{\mathcal{A}}_1)_* \to (\overline{\mathcal{A}}_2)_*$ such that

(a) $${}_{(\overline{\mathcal{A}}_1)_*}\Big(\rho_1, \Phi_{1,2}F_2\Big)_{\overline{\mathcal{A}}_1} = {}_{(\overline{\mathcal{A}}_2)_*}\Big((\Phi_{1,2})_*\rho_1, F_2\Big)_{\overline{\mathcal{A}}_2} \qquad (\forall \rho_1 \in (\overline{\mathcal{A}}_1)_*, \forall F_2 \in \overline{\mathcal{A}}_2) \tag{10.5}$$

(b) $(\Phi_{1,2})_*(\overline{\mathfrak{S}}^m((\overline{\mathcal{A}}_1)_*)) \subseteq \overline{\mathfrak{S}}^m((\overline{\mathcal{A}}_2)_*)$ (10.6)

This $(\Phi_{1,2})_*$ is called the **pre-dual causal operator** of $\Phi_{1,2}$.

(iv) there exists the continuous linear operator $\Phi_{1,2}^* : \mathcal{A}_1^* \to \mathcal{A}_2^*$ such that

(a) $${}_{(\overline{\mathcal{A}}_1)_*}\Big(\rho_1, \Phi_{1,2}F_2\Big)_{\overline{\mathcal{A}}_1} = {}_{\mathcal{A}_2^*}\Big(\Phi_{1,2}^*\widehat{\rho}_1, F_2\Big)_{\mathcal{A}_2} \qquad (\forall \rho_1 = \widehat{\rho}_1 \in (\overline{\mathcal{A}}_1)_*(\subseteq \mathcal{A}_1^*), \forall F_2 \in \mathcal{A}_2) \quad (10.7)$$

(b) $(\Phi_{1,2})^*(\mathfrak{S}^p(\mathcal{A}_1^*)) \subseteq \mathfrak{S}^m(\mathcal{A}_2^*)$ (10.8)

This $\Phi_{1,2}^*$ is called the **dual operator** of $\Phi_{1,2}$.

In addition, the causal operator $\Phi_{1,2}$ is called a **deterministic causal operator** , if it satisfies that

$$(\Phi_{1,2})^*(\mathfrak{S}^p(\mathcal{A}_1^*)) \subseteq \mathfrak{S}^p(\mathcal{A}_2^*) \qquad (10.9)$$

♠**Note 10.3.** [Causal operator in Classical systems] Consider the two basic structures:

$$[C_0(\Omega_1) \subseteq L^\infty(\Omega_1, \nu_1)]_{B(H_1)} \text{ and } [C_0(\Omega_2) \subseteq L^\infty(\Omega_2, \nu_2)]_{B(H_2)}$$

A continuous linear operator $\Phi_{1,2} : L^\infty(\Omega_2) \to L^\infty(\Omega_1)$ called a **causal operator**, if it satisfies the following (i)—(iii):

(i) $f_2 \in L^\infty(\Omega_2), \quad f_2 \geqq 0 \Longrightarrow \Phi_{12}f_2 \geqq 0$

(ii) $\Phi_{12}1_2 = 1_1$ where, $1_k(\omega_k) = 1$ $(\forall \omega_k \in \Omega_k, k = 1, 2)$

(iii) There exists a continuous linear operator $(\Phi_{1,2})_* : L^1(\Omega_1) \to L^1(\Omega_2)$ (and $(\Phi_{1,2})_* : L^1_{+1}(\Omega_1) \to L^1_{+1}(\Omega_2)$) such that

$$\int_{\Omega_1} [\Phi_{1,2}f_2](\omega_1) \ \ \rho_1(\omega_1)\nu_1(d\omega_1) = \int_{\Omega_2} f_2(\omega_2) \ \ [(\Phi_{1,2})_*\rho_1](\omega_2)\nu_2(d\omega_2)$$
$$(\forall \rho_1 \in L^1(\Omega_1), \forall f_2 \in L^\infty(\Omega_2))$$

This $(\Phi_{1,2})_*$ is called a **pre-dual causal operator** of $\Phi_{1,2}$.

(iv) There exists a continuous linear operator $\Phi_{1,2}^* : \mathcal{M}(\Omega_1) \to \mathcal{M}(\Omega_2)$ (and $\Phi_{1,2}^* : \mathcal{M}_{+1}(\Omega_1) \to \mathcal{M}_{+1}(\Omega_2)$) such that

$${}_{L^1(\Omega_1)}\Big(\rho_1, \Phi_{1,2}F_2\Big)_{L^\infty(\Omega_1)} = {}_{\mathcal{M}(\Omega_2)}\Big(\Phi_{1,2}^*\widehat{\rho}_1, F_2\Big)_{C_0(\Omega_2)} \qquad (\forall \rho_1 = \widehat{\rho}_1 \in \mathcal{M}(\Omega_1), \forall F_2 \in C_0(\Omega_2))$$

where, $\widehat{\rho}_1(D) = \int_D \rho_1(\omega_1)\nu_1(d\omega_1)$ $(\forall D \in \mathcal{B}_{\Omega_1})$. This $(\Phi_{1,2})^*$ is called a **dual causal operator** of $\Phi_{1,2}$.

In addition, a causal operator $\Phi_{1,2}$ is called a **deterministic causal operator**, if there exists a continuous map $\phi_{1,2} : \Omega_1 \to \Omega_2$ such that

$$[\Phi_{1,2}f_2](\omega_1) = f_2(\phi_{1,2}(\omega_1)) \quad (\forall f_2 \in C(\Omega_2), \forall \omega_1 \in \Omega_1) \qquad (10.10)$$

This $\phi_{1,2} : \Omega_1 \to \Omega_2$ is called a **deterministic causal map**. Here, it is clear that

$$\Omega_1 \approx \mathfrak{S}^p(C_0(\Omega_1)^*) \ni \delta_{\omega_1} \xrightarrow[\Phi_{12}^*]{} \delta_{\phi_{12}(\omega_1)} \in \mathfrak{S}^p(C_0(\Omega_2)^*) \approx \Omega_2$$

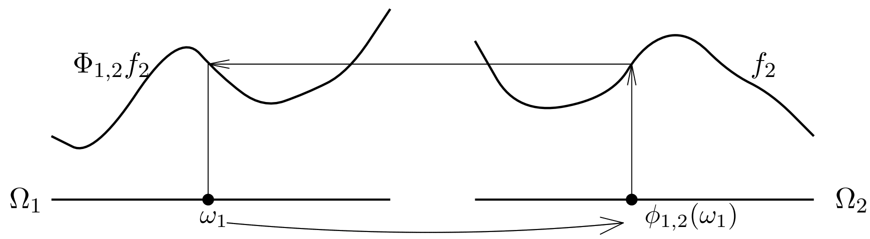


Figure 10.1: Deterministic causal map $\phi_{1,2}$ and deterministic causal operator $\Phi_{1,2}$

Theorem 10.5. [Continuous map and deterministic causal map] Let $(\Omega_1, \mathcal{B}_{\Omega_1}, \nu_1)$ and $(\Omega_2, \mathcal{B}_{\Omega_2}, \nu_2)$ be measure spaces. Assume that a continuous map $\phi_{1,2} : \Omega_1 \to \Omega_2$ satisfies:

$$D_2 \in \mathcal{B}_{\Omega_2}, \ \ \nu_2(D_2) = 0 \quad \Longrightarrow \quad \nu_1(\phi_{1,2}^{-1}(D_2)) = 0.$$

Then, the continuous map $\phi_{1,2} : \Omega_1 \to \Omega_2$ is deterministic, that is, the operator $\Phi_{1,2} : L^\infty(\Omega_2, \nu_2) \to L^\infty(\Omega_1, \nu_1)$ defined by (10.10) is a deterministic causal operator.

Proof. For each $\overline{\rho}_1 \in L^1(\Omega_1, \nu_1)$, define a measure μ_2 on $(\Omega_2, \mathcal{B}_{\Omega_2})$ such that

$$\mu_2(D_2) = \int_{\phi_{1,2}^{-1}(D_2)} \overline{\rho}_1(\omega_1) \, \nu_1(d\omega_1) \qquad (\forall D_2 \in \mathcal{B}_{\Omega_2})$$

Then, it suffices to consider the Radon-Nikodym derivative (*cf.* [74]) $[\Phi_{1,2}]_*(\overline{\rho}_1) = d\mu_2/d\nu_2$. That is because

$$D_2 \in \mathcal{B}_{\Omega_2}, \ \ \nu_2(D_2) = 0 \quad \Longrightarrow \quad \nu_1(\phi_{1,2}^{-1}(D_2)) = 0 \quad \Longrightarrow \quad \mu_2(D_2) = 0 \tag{10.11}$$

Thus, by the Radon-Nikodym theorem, we get a continuous linear operator $[\Phi_{1,2}]_* : L^1(\Omega_1, \nu_1) \to L^1(\Omega_2, \nu_2)$. □

Theorem 10.6. Let $\Phi_{1,2} : L^\infty(\Omega_2) \to L^\infty(\Omega_1)$ be a deterministic causal operator. Then, it holds that

$$\Phi_{1,2}(f_2 \cdot g_2) = \Phi_{1,2}(f_2) \cdot \Phi_{1,2}(g_2) \qquad (\forall f_2, \forall g_2 \in L^\infty(\Omega_2))$$

Proof. Let f_2, g_2 be in $L^\infty(\Omega_2)$. Let $\phi_{1,2} : \Omega_1 \to \Omega_2$ be the deterministic causal map of the deterministic causal operator $\Phi_{1,2}$. Then, we see

$$\begin{aligned} & [\Phi_{1,2}(f_2 \cdot g_2)](\omega_1) = (f_2 \cdot g_2)(\phi_{1,2}(\omega_1)) = f_2(\phi_{1,2}(\omega_1)) \cdot g_2(\phi_{1,2}(\omega_1)) \\ =& [\Phi_{1,2}(f_2)](\omega_1) \cdot [\Phi_{1,2}(g_2)](\omega_1) = [\Phi_{1,2}(f_2) \cdot \Phi_{1,2}(g_2)](\omega_1) \qquad (\forall \omega_1 \in \Omega_1) \end{aligned}$$

This completes the theorem. □

10.2.2 Simple example—Finite causal operator is represented by matrix

Example 10.7. [Deterministic causal operator, deterministic dual causal operator, deterministic causal map] Define the two states space Ω_1 and Ω_2 such that $\Omega_1 = \Omega_2 = \mathbb{R}$ with the Lebesgue measure ν. Thus we have the classical basic structures:

$$[C_0(\Omega_k) \subseteq L^\infty(\Omega_k, \nu) \subseteq B(L^2(\Omega_k, \nu))] \quad (k = 1, 2)$$

Define the deterministic causal map $\phi_{1,2} : \Omega_1 \to \Omega_2$ such that

$$\omega_2 = \phi_{1,2}(\omega_1) = 3(\omega_1)^2 + 2 \qquad (\forall \omega_1 \in \Omega_1 = \mathbb{R})$$

Then, by (10.10), we get the deterministic dual causal operator $\Phi^*_{1,2} : \mathcal{M}(\Omega_1) \to \mathcal{M}(\Omega_2)$ such that

$$\Phi^*_{1,2}\delta_{\omega_1} = \delta_{3(\omega_1)^2+2} \qquad (\forall \omega_1 \in \Omega_1)$$

where $\delta_{(\cdot)}$ is the point measure. Also, the deterministic causal operator$\Phi_{1,2} : L^\infty(\Omega_2) \to L^\infty(\Omega_1)$ is defined by

$$[\Phi_{1,2}(f_2)](\omega_1) = f_2(3(\omega_1)^2 + 2) \qquad (\forall f_2 \in C_0(\Omega_2), \forall \omega_1 \in \Omega_1)$$

Example 10.8. [Dual causal operator, causal operator] Recall Remark 2.13, that is, if Ω $(= \{1, 2, ..., n\})$ is finite set (with the discrete metric d_D and the counting measure ν,), we can consider that

$$C_0(\Omega) = L^\infty(\Omega, \nu) = \mathbb{C}^n, \qquad \mathcal{M}(\Omega) = L^1(\Omega, \nu) = \mathbb{C}^n, \qquad \mathcal{M}_{+1}(\Omega) = L^1_{+1}(\Omega, \nu)$$

For example, put $\Omega_1 = \{\omega_1^1, \omega_1^2, \omega_1^3\}$ and $\Omega_2 = \{\omega_2^1, \omega_2^2\}$. And define $\rho_1 (\in \mathcal{M}_{+1}(\Omega_1))$ such that

$$\rho_1 = a_1\delta_{\omega_1^1} + a_2\delta_{\omega_1^2} + a_3\delta_{\omega_1^3} \quad (0 \leqq a_1, a_2, a_3 \leqq 1, a_1 + a_2 + a_3 = 1)$$

Then, the dual causal operator $\Phi^*_{1,2} : \mathcal{M}_{+1}(\Omega_1) \to \mathcal{M}_{+1}(\Omega_2)$ is represented by

$$\begin{aligned}\Phi^*_{1,2}(\rho_1) =&(c_{11}a_1 + c_{12}a_2 + c_{13}a_3)\delta_{\omega_2^1} + (c_{21}a_1 + c_{22}a_2 + c_{23}a_3)\delta_{\omega_2^2} \\ &(0 \leqq c_{ij} \leqq 1, \sum_{i=1}^{2} c_{ij} = 1)\end{aligned}$$

and, consider the identification:$\mathcal{M}(\Omega_1) \approx \mathbb{C}^3$, $\mathcal{M}(\Omega_2) \approx \mathbb{C}^2$, That is,

$$\mathcal{M}(\Omega_1) \ni \alpha_1\delta_{\omega_1^1} + \alpha_2\delta_{\omega_1^2} + \alpha_3\delta_{\omega_1^3} \underset{\text{(identification)}}{\longleftrightarrow} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} \in \mathbb{C}^3$$

$$\mathcal{M}(\Omega_2) \ni \beta_1 \delta_{\omega_2^1} + \beta_2 \delta_{\omega_2^2} \underset{\text{(identification)}}{\longleftrightarrow} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} \in \mathbb{C}^2$$

Then, putting

$$\Phi_{1,2}^*(\rho_1) = \beta_1 \delta_{\omega_2^1} + \beta_2 \delta_{\omega_2^1} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix},$$

$$\rho_1 = \alpha_1 \delta_{\omega_1^1} + \alpha_2 \delta_{\omega_1^2} + \alpha_3 \delta_{\omega_1^3} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}$$

write, by matrix representation, as follows.

$$\Phi_{1,2}^*(\rho_1) = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}$$

Next, from this dual causal operator $\Phi_{1,2}^* : \mathcal{M}(\Omega_1) \to \mathcal{M}(\Omega_2)$, we shall construct a causal operator $\Phi_{1,2} : C_0(\Omega_2) \to C_0(\Omega_1)$. Consider the identification:$C_0(\Omega_1) \approx \mathbb{C}^3$, $C_0(\Omega_2) \approx \mathbb{C}^2$, that is,

$$C_0(\Omega_1) \ni f_1 \underset{\text{(identification)}}{\longleftrightarrow} \begin{bmatrix} f_1(\omega_1^1) \\ f_1(\omega_1^2) \\ f_1(\omega_1^3) \end{bmatrix} \in \mathbb{C}^3, \qquad C_0(\Omega_2) \ni f_2 \underset{\text{(identification)}}{\longleftrightarrow} \begin{bmatrix} f_2(\omega_2^1) \\ f_2(\omega_2^2) \end{bmatrix} \in \mathbb{C}^2$$

Let $f_2 \in C_0(\Omega_2)$, $f_1 = \Phi_{1,2} f_2$. Then, we see

$$\begin{bmatrix} f_1(\omega_1^1) \\ f_1(\omega_1^2) \\ f_1(\omega_1^3) \end{bmatrix} = f_1 = \Phi_{1,2}(f_2) = \begin{bmatrix} c_{11} & c_{21} \\ c_{12} & c_{22} \\ c_{13} & c_{23} \end{bmatrix} \begin{bmatrix} f_2(\omega_2^1) \\ f_2(\omega_2^2) \end{bmatrix}$$

Therefore, the relation between the dual causal operator$\Phi_{1,2}^*$ and causal operator$\Phi_{1,2}$ is represented as the the transposed matrix.

Example 10.9. [Deterministic dual causal operator, deterministic causal map, deterministic causal operator] Consider the case that dual causal operator $\Phi_{1,2}^* : \mathcal{M}(\Omega_1)(\approx\mathbb{C}^3) \to \mathcal{M}(\Omega_2)(\approx\mathbb{C}^2)$ ha s the matrix representation such that

$$\Phi_{1,2}^*(\rho_1) = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

In this case, it is the deterministic dual causal operator. This deterministic causal operator $\Phi_{1,2}$: $C_0(\Omega_2) \to C_0(\Omega_1)$ is represented by

$$\begin{bmatrix} f_1(\omega_1^1) \\ f_1(\omega_1^2) \\ f_1(\omega_1^3) \end{bmatrix} = f_1 = \Phi_{1,2}(f_2) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} f_2(\omega_2^1) \\ f_2(\omega_2^2) \end{bmatrix}$$

with the deterministic causal map $\phi_{1,2} : \Omega_1 \to \Omega_2$ such that

$$\phi_{1,2}(\omega_1^1) = \omega_2^2, \quad \phi_{1,2}(\omega_1^2) = \omega_2^1, \quad \phi_{1,2}(\omega_1^3) = \omega_2^1$$

10.2.3 Sequential causal operator — A chain of causalities

Let $(T, \leq)$ be a finite tree[1], i.e., a tree like semi-ordered finite set such that "$t_1 \leq t_3$ and $t_2 \leq t_3$" implies "$t_1 \leq t_2$ or $t_2 \leq t_1$". Assume that there exists an element $t_0 \in T$, called the *root* of T, such that $t_0 \leq t$ $(\forall t \in T)$ holds.

Put $T^2_{\leq} = \{(t_1, t_2) \in T^2 : t_1 \leq t_2\}$. An element $t_0 \in T$ is called a *root* if $t_0 \leq t$ $(\forall t \in T)$ holds. Since we usually consider the subtree T_{t_0} $(\subseteq T)$ with the root t_0, we assume that the tree has a root. In this chapter, assume, for simplicity, that T is finite (though it is sometimes infinite in applications).

For simplicity, assume that T is finite, or a finite subtree of a whole tree. Let T $(= \{0, 1, ..., N\})$ be a tree with the root 0. Define the *parent map* $\pi : T \setminus \{0\} \to T$ such that $\pi(t) = \max\{s \in T : s < t\}$. It is clear that the tree $(T \equiv \{0, 1, ..., N\}, \leq)$ can be identified with the pair $(T \equiv \{0, 1, ..., N\}, \pi : T \setminus \{0\} \to T)$. Also, note that, for any $t \in T \setminus \{0\}$, there uniquely exists a natural number $h(t)$ (called the *height* of t) such that $\pi^{h(t)}(t) = 0$. Here, $\pi^2(t) = \pi(\pi(t))$, $\pi^3(t) = \pi(\pi^2(t))$, etc. Also, put $\{0, 1, ..., N\}^2_{\leq} = \{(m, n) \mid 0 \leq m \leq n \leq N\}$. In Fig. 10.2, see the root t_0, the parent map: $\pi(t_3) = \pi(t_4) = t_2$, $\pi(t_2) = \pi(t_5) = t_1$, $\pi(t_1) = \pi(t_6) = \pi(t_7) = t_0$

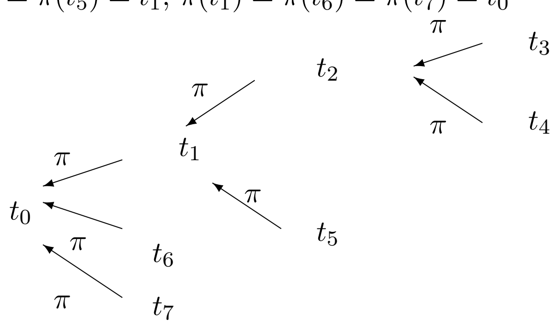


Figure 10.2: Tree: $(T = \{t_0, t_1, ..., t_7\}, \pi : T \setminus \{t_0\} \to T)$

Definition 10.10. [Sequential causal operator; Heisenberg picture of causality] The family $\{\Phi_{t_1,t_2} : \overline{\mathcal{A}}_{t_2} \to \overline{\mathcal{A}}_{t_1}\}_{(t_1,t_2) \in T^2_{\leqq}}$ $\Big($ or, $\{ \overline{\mathcal{A}}_{t_2} \overset{\Phi_{t_1,t_2}}{\to} \overline{\mathcal{A}}_{t_1}\}_{(t_1,t_2) \in T^2_{\leqq}}$ $\Big)$ is called a **sequential causal operator**, if it satisfies that

(i) For each t $(\in T)$, a basic structure $[\mathcal{A}_t \subseteq \overline{\mathcal{A}}_t \subseteq B(H_t)]$ is determined.

(ii) For each $(t_1, t_2) \in T^2_{\leqq}$, a causal operator $\Phi_{t_1,t_2} : \overline{\mathcal{A}}_{t_2} \to \overline{\mathcal{A}}_{t_1}$ is defined such as $\Phi_{t_1,t_2}\Phi_{t_2,t_3} = \Phi_{t_1,t_3}$ $(\forall (t_1, t_2), \forall (t_2, t_3) \in T^2_{\leqq})$. Here, $\Phi_{t,t} : \overline{\mathcal{A}}_t \to \overline{\mathcal{A}}_t$ is the identity operator.

[1] In Chapter 14, we discuss the infinite case

KSTS/RR-16/001
January 8, 2016

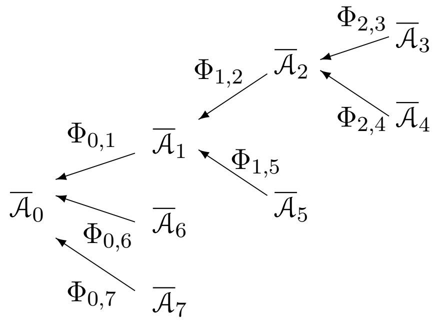
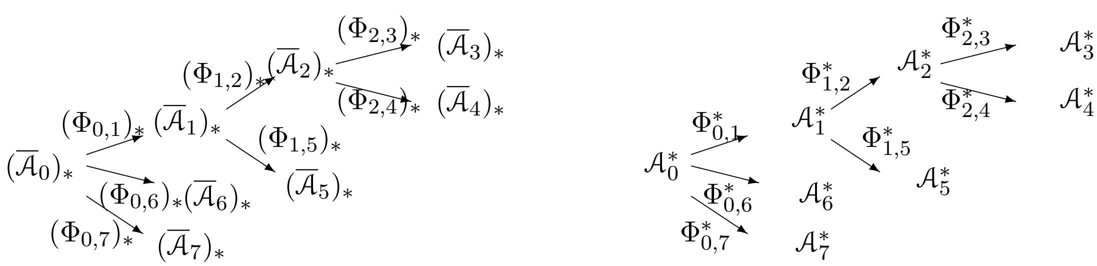


Figure 10.3: Heisenberg picture(sequential causal operator)

Definition 10.11. (i): [pre-dual sequential causal operator : Schrödinger picture of causality] The sequence $\{(\Phi_{t_1,t_2})_* : (\overline{\mathcal{A}}_{t_1})_* \to (\overline{\mathcal{A}}_{t_1})_*\}_{(t_1,t_2)\in T^2_{\leqq}}$ is called a **pre-dual sequential causal operator** of $\{\Phi_{t_1,t_2} : \overline{\mathcal{A}}_{t_2} \to \overline{\mathcal{A}}t_1\}_{(t_1,t_2)\in T^2_{\leqq}}$

(ii): [Dual sequential causal operator : Schrödinger picture of causality] A sequence $\{\Phi^*_{t_1,t_2} : \mathcal{A}^*_{t_1} \to \mathcal{A}^*_{t_1}\}_{(t_1,t_2)\in T^2_{\leqq}}$ is called a **dual sequential causal operator** of $\{\Phi_{t_1,t_2} : \overline{\mathcal{A}}_{t_2} \to \overline{\mathcal{A}}_{t_1}\}_{(t_1,t_2)\in T^2_{\leqq}}$.



(i):pre-dual sequential causal operator (ii):dual sequential causal operator

Figure 10.4: Schrödinger picture (dual sequential causal operator)

Remark 10.12. [The Heisenberg picture is formal; the Schrödinger picture is makeshift] The Schrödinger picture is intuitive and handy. Consider the Schrödinger picture$\{\Phi^*_{t_1,t_2} : \mathcal{A}^*_{t_1} \to \mathcal{A}^*_{t_1}\}_{(t_1,t_2)\in T^2_{\leqq}}$. For C^*-mixed state $\rho_{t_1} (\in \mathfrak{S}^m(\mathcal{A}^*_{t_1})$ (i.e., a state at time t_1),

- C^*-mixed state $\rho_{t_2} (\in \mathfrak{S}^m(\mathcal{A}^*_{t_2}))$ (at time $t_2(\geq t_1)$) is defined by

$$\rho_{t_2} = \Phi^*_{t_1,t_2}\rho_{t_1}$$

However, the linguistic interpretation says "state does not move", and thus, we consider that

- $\begin{cases} \text{the Heisenberg picture is formal} \\ \text{the Schrödinger picture is makeshift} \end{cases}$

10.3 Axiom 2 —Smoke is not located on the place which does not have fire

10.3.1 Axiom 2 (A chain of causal relations)

Now we can propose Axiom 2 (i.e., causality), which is the measurement theoretical representation of the maxim (Smoke is not located on the place which does not have fire):

(C): Axiom 2 (A chain of causalities)

(Under the preparation to this section, we can read this)

For each $t(\in T=$"tree")), consider the basic structure:

$$[\mathcal{A}_t \subseteq \overline{\mathcal{A}}_t \subseteq B(H_t)]$$

Then, the **chain of causalities** is represented by a **sequential causal operator** $\{\Phi_{t_1,t_2} : \overline{\mathcal{A}}_{t_2} \to \overline{\mathcal{A}}_{t_1}\}_{(t_1,t_2)\in T^2_{\leqq}}$.

♠**Note 10.4.** Axiom 2 (causality) as well as Axiom 1 (measurement) are a kind of spells. There are several spells concerning "motion". For example,

($\sharp_1$) [Aristotle]: final cause

($\sharp_2$) [Darwin]: evolution theory (survival of the fittest)

($\sharp_3$) [Hegel]: dialectic (Thesis, antithesis, synthesis)

($\sharp_4$) law of entropy increase

($\sharp_1$)–($\sharp_3$) are non-quantitative, but ($\sharp_4$) is quantitative. Everybody agrees that these (($\sharp_1$)–($\sharp_4$)) move the world.

10.3.2 Sequential causal operator—State equation, etc.

In what follows, we shall exercise the chain of causality in terms of quantum language.

Example 10.13. [State equation] Let $T = \mathbb{R}$ be a tree which represents the time axis. (Don't mind the infinity of T. *Cf.* Chapter 14.) For each $t(\in T)$, consider the state space $\Omega_t = \mathbb{R}^n$ (n-dimensional real space). And consider simultaneous ordinary differential equation of the first order

$$\begin{cases} \frac{d\omega_1}{dt}(t) = v_1(\omega_1(t), \omega_2(t), \ldots, \omega_n(t), t) \\ \frac{d\omega_2}{dt}(t) = v_2(\omega_1(t), \omega_2(t), \ldots, \omega_n(t), t) \\ \cdots\cdots \\ \frac{d\omega_n}{dt}(t) = v_n(\omega_1(t), \omega_2(t), \ldots, \omega_n(t), t) \end{cases} \tag{10.12}$$

which is called a **state equation** . Let $\phi_{t_1,t_2} : \Omega_{t_1} \to \Omega_{t_2}$, $(t_1 \leqq t_2)$ be a deterministic causal map induced by the state equation (10.12). It is clear that $\phi_{t_2,t_3}(\phi_{t_1,t_2}(\omega_{t_1})) = \phi_{t_1,t_3}(\omega_{t_1})$ $(\omega_{t_1} \in \Omega_{t_1}, t_1 \leqq t_2 \leqq t_3)$. Therefore, we have the deterministic sequential causal operator $\{\Phi_{t_1,t_2} : L^\infty(\Omega_{t_2}) \to L^\infty(\Omega_{t_1})\}_{(t_1,t_2) \in T^2_{\leqq}}$.

Example 10.14. [Difference equation of the second order] Consider the discrete time $T = \{0, 1, 2, \ldots\}$ with the parent map $\pi : T \setminus \{0\} \to T$ such that $\pi(t) = t - 1$ $(\forall t = 1, 2, ...)$. For each $t (\in T)$, consider a state space Ω_t such that $\Omega_t = \mathbb{R}$ (with the Lebesgue measure). For example, consider the following difference equation, that is, $\phi : \Omega_t \times \Omega_{t+1} \to \Omega_{t+2}$ satisfies as follows.

$$\omega_{t+2} = \phi(\omega_t, \omega_{t+1}) = \omega_t + \omega_{t+1} + 2 \qquad (\forall t \in T)$$

Here, note that the state ω_{t+2} depends on both ω_{t+1} and ω_t (i.e., multiple markov property). This must be modified as follows. For each $t (\in T)$ consider a new state space $\widetilde{\Omega}_t = \Omega_t \times \Omega_{t+1} = \mathbb{R} \times \mathbb{R}$. And define the deterministic causal map $\widetilde{\phi}_{t,t+1} : \widetilde{\Omega}_t \to \widetilde{\Omega}_{t+1}$ as follows.

$$(\omega_{t+1}, \omega_{t+2}) = \widetilde{\phi}_{t,t+1}(\omega_t, \omega_{t+1}) = (\omega_{t+1}, \omega_t + \omega_{t+1} + 2)$$
$$(\forall (\omega_t, \omega_{t+1}) \in \widetilde{\Omega}_t, \forall t \in T)$$

Therefore, by Theorem 10.5, the deterministic causal operator $\widetilde{\Phi}_{t,t+1} : L^\infty(\widetilde{\Omega}_{t+1}) \to L^\infty(\widetilde{\Omega}_t)$ is defined by

$$[\widetilde{\Phi}_{t,t+1} \tilde{f}_t](\omega_t, \omega_{t+1}) = \tilde{f}_t(\omega_{t+1}, \omega_t + \omega_{t+1} + 2)$$
$$(\forall (\omega_t, \omega_{t+1}) \in \widetilde{\Omega}_t, \forall \tilde{f}_t \in L^\infty(\widetilde{\Omega}_{t+1}), \forall t \in T \setminus \{0\}))$$

Thus, we get the deterministic sequential causal operator $\{\widetilde{\Phi}_{t,t+1} : L^\infty(\widetilde{\Omega}_{t+1}) \to L^\infty(\widetilde{\Omega}_t)\}_{t \in T \setminus \{0\}}$.

♠**Note 10.5.** In order to analyze multiple markov process and time-lag process, such ideas in Example 10.14 are needed.

10.4 Kinetic equation (in classical mechanics and quantum mechanics)

10.4.1 Hamiltonian (Time-invariant system)

In this section, we consider the simplest kinetic equation in classical system and quantum system.

Consider the state space Ω such that $\Omega = \mathbb{R}^2$, that is,

$$\mathbb{R}^2 = \mathbb{R}_q \times \mathbb{R}_p = \{(q, p) = (\text{position} , \text{momentum}) \mid q, p \in \mathbb{R}\} \tag{10.13}$$

Hamiltonian $\mathcal{H}(q, p)$ is defined by the total energy, for example, as the typical case (m: particle mass), we consider that

$$\begin{aligned} &[\text{Hamiltonian } (= \mathcal{H}(q, p))] \\ =&[\text{kinetic energy}(= \frac{p^2}{2m})] + [\text{potential energy}(= V(q))] \end{aligned} \tag{10.14}$$

10.4.2 Newtonian equation(=Hamilton's canonical equation)

Concerning Hamiltonian $\mathcal{H}(q, p)$, **Hamilton's canonical equation** is defined by

$$\text{Hamilton's canonical equation} = \begin{cases} \frac{dp}{dt} = -\frac{\mathcal{H}(q,p)}{\partial q} \\ \\ \frac{dq}{dt} = \frac{\mathcal{H}(q,p)}{\partial p} \end{cases} \tag{10.15}$$

And thus, in the case of (10.14), we get

$$\text{Hamilton's canonical equation} = \begin{cases} \frac{dp}{dt} = -\frac{\mathcal{H}(q,p)}{\partial q} = -\frac{\partial V(q,p)}{\partial q} \\ \\ \frac{dq}{dt} = \frac{\partial \mathcal{H}(q,p)}{\partial p} = \frac{p}{m} \end{cases} \tag{10.16}$$

which is the same as Newtonian equation. That is,

$$m\frac{d^2 q}{dt^2} = [\text{Mass}] \times [\text{Acceleration}] = -\frac{\partial V(q, p)}{\partial q} (= \text{Force})$$

Now, let us describe the above (10.16) in terms of quantum language. For each $t \in T = \mathbb{R}$, define the state space Ω_t by

$$\Omega_t = \Omega = \mathbb{R}^2 = \mathbb{R}_q \times \mathbb{R}_p = \{(q, p) = (\text{position} , \text{momentum}) \mid q, p \in \mathbb{R}\} \tag{10.17}$$

and assume Lebesgue measure ν.

Then, we have the classical basic structure:

$$[C_0(\Omega_t) \subseteq L^\infty(\Omega_t) \subseteq B(L^2(\Omega_t))] \qquad (\forall t \in T = \mathbb{R})$$

The solution of the canonical equation (10.16) is defined by

$$\Omega_{t_1} \ni \omega_{t_1} \mapsto \phi_{t_1,t_2}(\omega_{t_1}) = \omega_{t_2} \in \Omega_{t_2} \tag{10.18}$$

Since (10.18) determines the deterministic causal map, we have the deterministic sequential causal operator $\{\Phi_{t_1,t_2} : L^\infty(\Omega_{t_2}) \to L^\infty(\Omega_{t_1})\}_{(t_1,t_2)\in T^2_{\leq}}$ such that

$$[\Phi_{t_1,t_2}(f_{t_2})](\omega_{t_1}) = f_{t_2}(\phi_{t_1,t_2}(\omega_{t_1})) \qquad (\forall f_{t_2} \in L^\infty(\Omega_2), \forall \omega_{t_1} \in \Omega_{t_1}, t_1 \leq t_2) \tag{10.19}$$

10.4.3 Schrödinger equation (quantizing Hamiltonian)

The quantization is the following procedure:

$$\textbf{quantization}^2 \begin{cases} \text{total energy} E & \xrightarrow[\text{quantumization}]{} \frac{\hbar\sqrt{-1}\partial}{\partial t} \\ \text{momentum } p & \xrightarrow[\text{quantumization}]{} \frac{\hbar \partial}{\sqrt{-1}\partial q} \\ \text{position } q & \xrightarrow[\text{quantumization}]{} q \end{cases} \tag{10.20}$$

Substituting the quantumization (10.20) to the classical Hamiltonian:

$$E = \mathcal{H}(q,p) = \frac{p^2}{2m} + V(q)$$

we get

$$\hbar\sqrt{-1}\frac{\partial}{\partial t} = \mathcal{H}(q, \frac{\hbar}{\sqrt{-1}}\frac{\partial}{\partial q}) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial q^2} + V(q) \tag{10.21}$$

And therefore, we get the **Schrödinger equation**:

$$\hbar\sqrt{-1}\frac{\partial u(t,q)}{\partial t} = \mathcal{H}(q, \frac{\hbar}{\sqrt{-1}}\frac{\partial}{\partial q})u(t,q) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial q^2}u(t,q) + V(q)u(t,q) \tag{10.22}$$

Putting $u(t,\cdot) = u_t \in L^2(\mathbb{R})$ $(\forall t \in T = \mathcal{R})$ we denote the Schrödinger equation (10.22) by

$$u_t = \frac{1}{\hbar\sqrt{-1}}\mathcal{H}u_t$$

[2] Learning the (10.20) by rote, we can derive Schrödinger equation (10.22). However, the meaning of "quantumization" is not clear.

Solving this formally, we see

$$u_t = e^{\frac{\mathcal{H}}{\hbar\sqrt{-1}}t} u_0 \quad \text{(Thus, the state representation is } |u_t\rangle\langle u_t| = |e^{\frac{\mathcal{H}}{\hbar\sqrt{-1}}t} u_0\rangle\langle e^{\frac{\mathcal{H}}{\hbar\sqrt{-1}}t} u_0| \text{)} \tag{10.23}$$

where, $u_0 \in L^2(\mathbb{R})$ is an initial condition.

Now, put Hilbert space $H_t = L^2(\mathbb{R})$ $(\forall t \in T = \mathbb{R})$, and consider the quantum basic structure:

$$[\mathcal{C}(L^2(\mathbb{R})) \subseteq B(L^2(\mathbb{R})) \subseteq B(L^2(\mathbb{R}))]$$

The dual sequential causal operator $\{\Phi^*_{t_1,t_2} : \mathcal{T}r(H_{t_1}) \to \mathcal{T}r(H_{t_2})\}_{(t_1,t_2)\in T^2_{\le}}$ is defined by

$$\Phi^*_{t_1,t_2}(\rho) = e^{\frac{\mathcal{H}}{\hbar\sqrt{-1}}(t_2-t_1)} \rho e^{\frac{-\mathcal{H}}{\hbar\sqrt{-1}}(t_2-t_1)} \quad (\forall \rho \in \mathcal{T}r(H_{t_1}) = (B(H_{t_1}))_* = \mathcal{C}(H_{t_1})^*) \tag{10.24}$$

And therefore, the sequential causal operator $\{\Phi_{t_1,t_2} : B(H_{t_2}) \to B(H_{t_1})\}_{(t_1,t_2)\in T^2_{\le}}$ is defined by

$$\Phi_{t_1,t_2}(A) = e^{\frac{-\mathcal{H}}{\hbar\sqrt{-1}}(t_2-t_1)} A e^{\frac{\mathcal{H}}{\hbar\sqrt{-1}}(t_2-t_1)} \quad (\forall A \in B(H_{t_2})) \tag{10.25}$$

Also, since

$$\Phi^*_{t_1,t_2}(\mathfrak{S}^p(\mathcal{C}(H_{t_1})^*) \subseteq \mathfrak{S}^p(\mathcal{C}(H_{t_2})^*),$$

the sequential causal operator $\{\Phi_{t_1,t_2} : B(H_{t_2}) \to B(H_{t_1})\}_{(t_1,t_2)\in T^2_{\le}}$ is deterministic. Since we deal with the time-invariant system, putting $t = t_2 - t_1$, we see that (10.25) is equal to

$$A_t = \Phi_t(A_0) = e^{\frac{-\mathcal{H}}{\hbar\sqrt{-1}}t} A_0 e^{\frac{\mathcal{H}}{\hbar\sqrt{-1}}t} \tag{10.26}$$

And thus, we get the differential equation:

$$\begin{aligned} \frac{dA_t}{dt} &= \frac{-\mathcal{H}}{\hbar\sqrt{-1}} e^{\frac{-\mathcal{H}}{\hbar\sqrt{-1}}t} A_0 e^{\frac{\mathcal{H}}{\hbar\sqrt{-1}}t} + \frac{-\mathcal{H}}{\hbar\sqrt{-1}} e^{\frac{-\mathcal{H}}{\hbar\sqrt{-1}}t} A_0 e^{\frac{\mathcal{H}}{\hbar\sqrt{-1}}t} \frac{\mathcal{H}}{\hbar\sqrt{-1}} \\ &= \frac{-\mathcal{H}}{\hbar\sqrt{-1}} A_t + A_t \frac{\mathcal{H}}{\hbar\sqrt{-1}} = \frac{1}{\hbar\sqrt{-1}} \Big(A_t\mathcal{H} - \mathcal{H}A_t \Big) \end{aligned} \tag{10.27}$$

which is just *Heisenberg's kinetic equation.* In quantum language, we say that

- Heisenberg's kinetic equation is formal, and Schrödinger equation is makeshift,

though the two are usually said to be equivalent.

10.5 Exercise:Solve Schrödinger equation by variable separation method

Consider a particle with the mass m in the box (i.e., the closed interval $[0,2]$) in the one dimensional space $\mathbb{R}$. The motion of this particle (i.e., the wave function of the particle) is represented by the following Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\psi(q,t) = -\frac{\hbar^2\partial^2}{2m\partial q^2}\psi(q,t) + V_0(q)\psi(q,t) \qquad (\text{ in } H = L^2(\mathbb{R}))$$

where

$$V_0(q) = \begin{cases} 0 & (0 \le q \le 2) \\ \infty & (\text{ otherwise }) \end{cases}$$

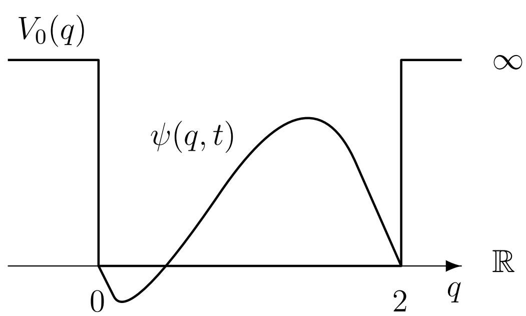


Figure 10.5: Particle in a box

Put

$$\phi(q,t) = T(t)X(q) \qquad (0 \le q \le 2).$$

And consider the following equation:

$$i\hbar\frac{\partial}{\partial t}\phi(q,t) = -\frac{\hbar^2\partial^2}{2m\partial q^2}\phi(q,t).$$

Then, we see

$$\frac{iT'(t)}{T(t)} = -\frac{X''(q)}{2mX(q)} = K(= \text{ constant }).$$

Then,

$$\phi(q,t) = T(t)X(q) = C_3\exp(iKt)\Big(C_1\exp(i\sqrt{2mK/\hbar}\, q) + C_2\exp(-i\sqrt{2mK/\hbar}\, q).\Big)$$

Since $X(0) = X(2) = 0$ (perfectly elastic collision), putting $K = \frac{n^2\pi^2\hbar}{8m}$, we see

$$\phi(q,t) = T(t)X(q) = C_3 \exp(\frac{in^2\pi^2\hbar t}{8m}) \sin(n\pi q/2) \qquad (n = 1, 2, ...).$$

Assume the initial condition:

$$\psi(q,0) = c_1 \sin(\pi q/2) + c_2 \sin(2\pi q/2) + c_3 \sin(3\pi q/2) + \cdots .$$

where $\int_{\mathbb{R}} |\psi(q,0)|^2 dq = 1$. Then we see

$$\begin{aligned}&\psi(q,t)\\ =&c_1 \exp(\frac{i\pi^2\hbar t}{8m}) \sin(\pi q/2) + c_2 \exp(\frac{i4\pi^2\hbar t}{8m}) \sin(2\pi q/2) + c_3 \exp(\frac{i9\pi^2\hbar t}{8m}) \sin(3\pi q/2) + \cdots .\end{aligned}$$

And thus, we have the time evolution of the state by

$$\rho_t = |\psi(\cdot,t)\rangle\langle\psi(\cdot,t)| \qquad (\in \mathfrak{S}^p(Tr(H)) \subseteq B(H)) \qquad\qquad (\forall t \geq 0)$$

10.6 Random walk and quantum decoherence

10.6.1 Diffusion process

Example 10.15. [Random walk] Let the state space Ω be $\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$ with the counting measure ν. Define the dual causal operator $\Phi^* : \mathcal{M}_{+1}(\mathbb{Z}) \to \mathcal{M}_{+1}(\mathbb{Z})$ such that

$$\Phi^*(\delta_i) = \frac{\delta_{i-1} + \delta_{i+1}}{2} \qquad (i \in \mathbb{Z})$$

where $\delta_{(\cdot)}(\in \mathcal{M}_{+1}(\mathbb{Z}))$ is a point measure. Therefore, the causal operator $\Phi : L^\infty(\mathbb{Z}) \to L^\infty(\mathbb{Z})$ is defined by

$$[\Phi(F)](i) = \frac{F(i-1) + F(i+1)}{2} \qquad (\forall F \in L^\infty(\mathbb{Z}), \forall i \in \mathbb{Z})$$

and the pre-dual causal operator $\Phi_* : L^1(\mathbb{Z}) \to L^1(\mathbb{Z})$ is defined by

$$[\Phi_*(f)](i) = \frac{f(i-1) + F(i+1)}{2} \qquad (\forall f \in L^1(\mathbb{Z}), \forall i \in \mathbb{Z})$$

Now, consider the discrete time $T = \{0, 1, 2, \ldots, N\}$, where the parent map $\pi : T \setminus \{0\} \to T$ is defined by $\pi(t) = t - 1$ $(t = 1, 2, \ldots)$. For each $t(\in T)$, a state space Ω_t is define by $\Omega_t = \mathbb{Z}$. Then, we have the sequential causal operator $\{\Phi_{\pi(t),t}(= \Phi) : L^\infty(\Omega_t) \to L^\infty(\Omega_{\pi(t)})\}_{t \in T \setminus \{0\}}$

.

10.6.2 Quantum decoherence: non-deterministic causal operator

Consider the quantum basic structure:

$$[\mathcal{C}(H) \subseteq B(H) \subseteq B(H)]$$

Let $\mathbb{P} = \{P_n\}_{n=1}^\infty$ be the spectrum decomposition in $B(H)$, that is,

$$P_n \text{ is a projection (i.e., } P_n = (P_n)^2 \text{), and, } \sum_{n=1}^\infty P_n = I$$

Define the operator $(\Psi_{\mathbb{P}})_* : \mathcal{T}r(H) \to \mathcal{T}r(H)$ such that

$$(\Psi_{\mathbb{P}})_*(|u\rangle\langle u|) = \sum_{n=1}^\infty |P_n u\rangle\langle P_n u| \quad (\forall u \in H)$$

Clearly we see

$$\langle v, (\Psi_{\mathbb{P}})_*(|u\rangle\langle u|)v\rangle = \langle v, (\sum_{n=1}^\infty |P_n u\rangle\langle P_n u|)v\rangle = \sum_{n=1}^\infty |\langle v, |P_n u\rangle|^2 \geq 0 \qquad (\forall u, v \in H)$$

and,

$$
\begin{aligned}
&\mathrm{Tr}((\Psi_{\mathbb{P}})_*(|u\rangle\langle u|)) \\
=&\mathrm{Tr}(\sum_{n=1}^{\infty}|P_n u\rangle\langle P_n u|) = \sum_{n=1}^{\infty}\sum_{k=1}^{\infty}|\langle e_k, P_n u\rangle|^2 = \sum_{n=1}^{\infty}\|P_n u\|^2 = \|u\|^2 \qquad (\forall u \in H)
\end{aligned}
$$

where $\{e_k\}_{k=1}^{\infty}$ is CONS in H.
And so,

$$
(\Psi_{\mathbb{P}})_*(\mathfrak{T}r_{+1}^p(H)) \subseteq \mathfrak{T}r_{+1}(H)
$$

Therefore, $\Psi_{\mathbb{P}}(= ((\Psi_{\mathbb{P}})_*)^*) : B(H) \to B(H)$ is a causal operator, but it is not deterministic. In this note, a non-deterministic (sequential) causal operator is called a quantum decoherence.

Remark 10.16. [Quantum decoherence] For the relation between quantum decoherence and quantum Zeno effect, see § 11.4. Also, for the relation between quantum decoherence and Schrödinger's cat, see § 11.5.

In tis note, we assume that the don-deterministic causal operator belongs to the mixed measurement theory. Thus, we consider that quantum language (= measurement theory) is classified as follows.

$$
\text{(A)}\ \underset{\text{(=quantum language)}}{\text{measurement theory}}
\begin{cases}
\underset{(\mathrm{A}_1)}{\text{pure type}} \begin{cases} \text{classical system : Fisher statistics} \\ \text{quantum system : usual quantum mechanics} \end{cases} \\
\\
\underset{(\mathrm{A}_2)}{\text{mixed type}} \begin{cases} \text{classical system : including Bayesian statistics, Kalman filter} \\ \text{quantum system : quantum decoherence} \end{cases}
\end{cases}
$$

10.7 Leibniz=Clarke Correspondence: What is space-time?

The problems ("What is space?" and "What is time?") are the most important in modern science as well as the traditional philosophies. In this section, we give my answer to this problem.

10.7.1 "What is space?" and "What is time?")

10.7.1.1 Space in quantum language
(How to describe "space" in quantum language)

In what follows, let us explain "space" in measurement theory (= quantum language).

For example, consider the simplest case, that is,

(A) "space"$=\mathbb{R}_q$(one dimensional space)

Since classical system and quantum system must be considered, we see

$$
\text{(B)} \quad \begin{cases} (\mathrm{B}_1)\text{: a classical particle in the one dimensional space } \mathbb{R}_q \\ \\ (\mathrm{B}_2)\text{: a quantum particle in the one dimensional space } \mathbb{R}_q \end{cases}
$$

In the classical case, we start from the following state:

$$(q,p) = (\text{"position"}, \text{"momentum"}) \in \mathbb{R}_q \times \mathbb{R}_p$$

Thus, we have the classical basic structure:

$$(\mathrm{C}_1) \qquad [C_0(\mathbb{R}_q \times \mathbb{R}_p) \subseteq L^\infty(\mathbb{R}_q \times \mathbb{R}_p) \subseteq B(L^2(\mathbb{R}_q \times \mathbb{R}_p)]$$

Also, concerning quantum system, we have the quantum basic structure:

$$(\mathrm{C}_2) \qquad [\mathcal{C}(L^2(\mathbb{R}_q) \subseteq B(L^2(\mathbb{R}_q) \subseteq B(L^2(\mathbb{R}_q)]$$

Summing up, we have the basic structure

$$
\text{(C)} \ \ [\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)] \begin{cases} (\mathrm{C}_1)\text{: classical } [C_0(\mathbb{R}_q \times \mathbb{R}_p) \subseteq L^\infty(\mathbb{R}_q \times \mathbb{R}_p) \subseteq B(L^2(\mathbb{R}_q \times \mathbb{R}_p)] \\ \\ (\mathrm{C}_2)\text{: quantum } [\mathcal{C}(L^2(\mathbb{R}_q) \subseteq B(L^2(\mathbb{R}_q) \subseteq B(L^2(\mathbb{R}_q)] \end{cases}
$$

Since we always start from a basic structure in quantum language, we consider that

$$
\begin{aligned}
&\text{How to describe "space" in quantum language} \\
\Leftrightarrow\ &\text{How to describe [(A):space] by [(C):basic structure]}
\end{aligned}
\tag{10.28}
$$

This is done in the following steps.

Assertion 10.17. How to describe "space" in quantum language

(D$_1$) Begin with the basic structure:

$$[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]$$

(D$_2$) Next, consider a certain commutative C^*-algebra $\mathcal{A}_0 (= C_0(\Omega))$ such that

$$\mathcal{A}_0 \subseteq \overline{\mathcal{A}}$$

(D$_3$) Lastly, the spectrum Ω $(\approx \mathfrak{S}^p(\mathcal{A}_*))$ is used to represent "space".

For example,

(E$_1$) in the classical case (C$_1$):

$$[C_0(\mathbb{R}_q \times \mathbb{R}_p) \subseteq L^\infty(\mathbb{R}_q \times \mathbb{R}_p) \subseteq B(L^2(\mathbb{R}_q \times \mathbb{R}_p))]$$

we have the commutative $C_0(\mathbb{R}_q)$ such that

$$C_0(\mathbb{R}_q) \subseteq L^\infty(\mathbb{R}_q \times \mathbb{R}_p)$$

And thus, we get the space $\mathbb{R}_q$ as mentioned in (A)

(E$_2$) in the quantum case (C$_2$):

$$[\mathcal{C}(L^2(\mathbb{R}_q) \subseteq B(L^2(\mathbb{R}_q)) \subseteq B(L^2(\mathbb{R}_q))]$$

we have the commutative $C_0(\mathbb{R}_q)$ such that

$$C_0(\mathbb{R}_q) \subseteq B(L^2(\mathbb{R}_q))$$

And thus, we get the space $\mathbb{R}_q$ as mentioned in (A)

10.7.1.2 Time in quantum language (How to describe "time" in quantum language)

In what follows, let us explain "time" in measurement theory (= quantum language).

This is easily done in the following steps.

Assertion 10.18. How to describe "time" in quantum language

KSTS/RR-16/001
January 8, 2016

(F$_1$) Let T be a tree. (Don't mind the finiteness or infinity of T. *Cf.* Chapter 14.) For each $t \in T$, consider the basic structure:

$$[\mathcal{A}_t \subseteq \overline{\mathcal{A}}_t \subseteq B(H_t)]$$

(F$_2$) Next, consider a certain linear subtree $T'(\subseteq T)$, which can be used to represent "time".

10.7.2 Leibniz-Clarke Correspondence

The above argument urges us to recall Leibniz-Clarke Correspondence (1715–1716: *cf.* [1]), which is important to know both Leibniz's and Clarke's (=Newton's) ideas concerning space and time.

(G) [The realistic space-time]
Newton's absolutism says that the space-time should be regarded as a receptacle of a "thing." Therefore, even if "thing" does not exits, the space-time exists.

On the other hand,

(H) [The metaphysical space-time]
Leibniz's relationalism says that

- (H$_1$) Space is a kind of state of "thing".
- (H$_2$) Time is an order of occurring in succession which changes one after another.

Therefore, I regard this correspondence as

$$\underset{\text{(realistic view)}}{\boxed{\text{Newton } (\approx \text{Clarke})}} \quad \underset{\text{v.s.}}{\longleftrightarrow} \quad \underset{\text{(linguistic view)}}{\boxed{\text{Leibniz}}}$$

which should be compared to

$$\underset{\text{(realistic view)}}{\boxed{\text{Einstein}}} \quad \underset{\text{v.s.}}{\longleftrightarrow} \quad \underset{\text{(linguistic view)}}{\boxed{\text{Bohr}}}$$

(also, recall Note 4.4).

♠**Note 10.6.** Many scientists may think that

Newton's assertion is understandable, in fact, his idea was inherited by Einstein. On the other, Leibniz's assertion is incomprehensible and literary. Thus, his idea is not related to science.

However, recall the classification of the world-description (Figure 1.1):

$$\begin{cases} ① : \underset{\text{(realistic world view)}}{\text{Newton, Clarke}} \cdots \underset{\text{“What is space-time?”}}{\overset{\text{(space-time in physics)}}{\boxed{\text{realistic space-time}}}} \quad \text{(successors: Einstein, etc.)} \\ ② : \underset{\text{(linguistic world view)}}{\text{Leibniz}} \cdots \underset{\text{“How should space-time be represented?”}}{\overset{\text{(space-time in measurement theory)}}{\boxed{\text{linguistic space-time}}}} \quad \text{(i.e., spectrum, tree)} \end{cases}$$

in which Newton and Leibniz respectively devotes himself to ① and ②. Although Leibniz's assertion is not clear, we believe that

- Leibniz found the importance of "linguistic space and time" in science,

Also, it should be noted that

(♯) Newton proposed the scientific language called Newtonian mechanics,
on the other hand,
Leibniz could not propose a scientific language

♠Note 10.7. I want to believe that "realistic" vs. "linguistic" is always hidden behind the great disputes in the history of the world view. That is,

$$\boxed{\text{realistic world view}} \underset{\text{v.s.}}{\longleftrightarrow} \underset{\text{(idealistic)}}{\boxed{\text{linguistic world view}}}$$

For example,

Table 10.1 : The realistic world view vs the linguistic world view

Dispute ＼ R vs. L	the realistic world view	the linguistic world view
Greek philosophy	Aristotle	Plato
Problem of universals	Realismus(Anselmus)	Nominalisme(William of Ockham)
Space·times	Clarke(Newton)	Liebniz
Quantum mechanics	Einstein (*cf.* [13])	Bohr (*cf.* [5])

I want to believe that "realistic" vs. "linguistic" is always hidden behind the greatest disputes in the history of the world view.

♠Note 10.8. The space-time in measuring object is well discussed in the above. However, we have to say something about "observer's time". We conclude that observer's time is meaningless in measurement theory as mentioned the linguistic interpretation in Chap. 1. That is, the following question is nonsense in measurement theory:

($\sharp_1$) When and where does an observer take a measurement

($\sharp_2$) Therefore, there is no tense (present, past, future) in sciences.

Thus, some may recall

McTaggart's paradox: "Time does not exist"

(*cf.* ref.[58]). Although McTaggart s logic is not clear, we believe that his assertion is the same as "Subjective time (e.g., Augustinus' times, Bergson's times, etc.) does not exist in science". If it be so,

($\sharp_3$) **McTaggart's assertion as well as Leibniz' assertion are one of the linguistic interpretation.**

After all, we conclude that

($\sharp_4$) **the cause of philosophers' failure is not to propose a language**.

Talking cynically, we say that

($\sharp_5$) Philosophers continued investigating "linguistic interpretation" (="how to use Axioms 1 and 2") without language (i.e., Axiom 1(measurement:§2.7) and Axiom 2(causality:§10.3)).

Chapter 11

Simple measurement and causality

Until the previous chapter, we studied all of quantum language, that is,

$$
(\sharp)\begin{cases}
(\sharp_1): \underset{\text{(=quantum language)}}{\boxed{\text{pure measurement theory}}} \\
\qquad := \underbrace{\overset{\text{[(pure)Axiom 1]}}{\underset{(\textit{cf.}\ \S 2.7)}{\boxed{\text{pure measurement}}}} + \overset{\text{[Axiom 2]}}{\underset{(\textit{cf.}\ \S 10.3)}{\boxed{\text{Causality}}}}}_{\text{a kind of spell(a priori judgment)}} + \underbrace{\overset{\text{[quantum linguistic interpretation]}}{\underset{(\textit{cf.}\ \S 3.1)}{\boxed{\text{Linguistic interpretation}}}}}_{\text{the manual to use spells}} \\
\\
(\sharp_2): \underset{\text{(=quantum language)}}{\boxed{\text{mixed measurement theory}}} \\
\qquad := \underbrace{\overset{\text{[(mixed)Axiom}^{(m)}\text{ 1]}}{\underset{(\textit{cf.}\ \S 9.1)}{\boxed{\text{mixed measurement}}}} + \overset{\text{[Axiom 2]}}{\underset{(\textit{cf.}\ \S 10.3)}{\boxed{\text{Causality}}}}}_{\text{a kind of spell(a priori judgment)}} + \underbrace{\overset{\text{[quantum linguistic interpretation]}}{\underset{(\textit{cf.}\ \S 3.1)}{\boxed{\text{Linguistic interpretation}}}}}_{\text{the manual to use spells}}
\end{cases}
$$

However, what is important is

- **to exercise the relationship of measurement and causality**

Since measurement theory is a language, we have to note the following wise sayings:

- experience is the best teacher, or custom makes all things

11.1 The Heisenberg picture and the Schrödinger picture

11.1.1 State does not move— the Heisenberg picture —

We consider that

"only one measurement" $\Longrightarrow$"state does not move"

That is because

(a) In order to see the state movement, we have to take measurement at least more than twice. However, the "plural measurement" is prohibited. Thus, we conclude "state does not move"

We want to believe that this is associated with Parmenides' words:

There is no movement

which is related to the Heisenberg picture. This will be explained in what follows.

Theorem 11.1. [Causal operator and observable] Consider the basic structure:

$$[\mathcal{A}_k \subseteq \overline{\mathcal{A}}_k \subseteq B(H_k)] \qquad (k=1,2)$$

Let $\Phi_{1,2} : \overline{\mathcal{A}}_2 \to \overline{\mathcal{A}}_1$ be a causal operator, and let $\mathsf{O}_2 = (X, \mathcal{F}, F_2)$ be an observable in $\overline{\mathcal{A}}_2$. Then, $\Phi_{1,2}\mathsf{O}_2 = (X, \mathcal{F}, \Phi_{1,2}F_2)$ is an observable in $\overline{\mathcal{A}}_2$.

Proof. Let $\Xi \ (\in \mathcal{F})$. And consider the countable decomposition $\{\Xi_1, \Xi_2, \ldots, \Xi_n, \ldots\}$ of Ξ $\Big($ i.e., $\Xi = \bigcup_{n=1}^{\infty} \Xi_n$, $\Xi_n \in \mathcal{F}, (n=1,2,\ldots)$, $\Xi_m \cap \Xi_n = \emptyset \quad (m \neq n)$ $\Big)$. Then we see, for any $\rho_1 (\in (\mathcal{A}_1)_*)$,

$$
\begin{aligned}
&{}_{(\overline{\mathcal{A}}_1)_*}\Big(\rho_1, \Phi_{1,2}F_2(\bigcup_{n=1}^{\infty}\Xi_n)\Big)_{\overline{\mathcal{A}}_1} = {}_{(\overline{\mathcal{A}}_1)_*}\Big((\Phi_{1,2})_*\rho_1, F_2(\bigcup_{n=1}^{\infty}\Xi_n)\Big)_{\overline{\mathcal{A}}_2} \\
=&\sum_{n=1}^{\infty} {}_{(\overline{\mathcal{A}}_1)_*}\Big((\Phi_{1,2})_*\rho_1, F_2(\Xi_n)\Big)_{\overline{\mathcal{A}}_2} = \sum_{n=1}^{\infty} {}_{(\overline{\mathcal{A}}_1)_*}\Big(\rho_1, \Phi_{1,2}F_2(\Xi_n)\Big)_{\overline{\mathcal{A}}_2}
\end{aligned}
$$

Thus,$\Phi_{1,2}\mathsf{O}_2 = (X, \mathcal{F}, \Phi_{1,2}F_2)$ is an observable in $\overline{\mathcal{A}}_1$. □

Let us begin from the simplest case. Consider a tree $T = \{0, 1\}$. For each $t \in T$, consider the basic structure:

$$[\mathcal{A}_t \subseteq \overline{\mathcal{A}}_t \subseteq B(H_t)] \qquad (t=0,1)$$

And consider the causal operator $\Phi_{0,1} : \overline{\mathcal{A}}_1 \to \overline{\mathcal{A}}_0$. That is,

$$\overline{\mathcal{A}}_0 \xleftarrow{\Phi_{0,1}} \overline{\mathcal{A}}_1 \tag{11.1}$$

Therefore, we have the pre-dual operator $(\Phi_{0,1})_*$ and the dual operator $\Phi_{0,1}^*$:

$$(\overline{\mathcal{A}}_0)_* \xrightarrow[(\Phi_{0,1})_*]{} (\overline{\mathcal{A}}_1)_* \qquad\qquad \mathcal{A}_0^* \xrightarrow[\Phi_{0,1}^*]{} \mathcal{A}_1^* \tag{11.2}$$

If $\Phi_{0,1} : \overline{\mathcal{A}}_1 \to \overline{\mathcal{A}}_0$ is deterministic, we see that

$$\mathcal{A}_0^* \supset \mathfrak{S}^p(\mathcal{A}_0^*) \ni \rho \underset{\Phi_{0,1}^*}{\longrightarrow} \Phi_{0,1}^* \rho \in \mathfrak{S}^p(\mathcal{A}_1^*) \subset \mathcal{A}_1^* \tag{11.3}$$

Under the above preparation, we shall explain the Heisenberg picture and the Schrödinger picture in what follows.

Assume that

(A_1) Consider a deterministic causal operator $\Phi_{0,1} : \overline{\mathcal{A}}_1 \to \overline{\mathcal{A}}_0$.

(A_2) a state $\rho_0 \in \mathfrak{S}^p(\mathcal{A}_0^*)$: pure state

(A_3) Let $\mathsf{O}_1 = (X_1, \mathcal{F}_1, F_1)$ be an observable in $\overline{\mathcal{A}}_1$.

Explanation 11.2. **[the Heisenberg picture].**
The Heisenberg picture is just the following (a):

(a1) **To identify an observable O_1 in $\overline{\mathcal{A}}_1$ with an $\Phi_{0,1}\mathsf{O}_1$ in $\overline{\mathcal{A}}_0$** . That is,

$$\underset{(\text{ in } \overline{\mathcal{A}}_0)}{\Phi_{0,1}\overline{\mathsf{O}}_1} \quad \xleftarrow[\text{identification}]{\Phi_{0,1}} \quad \underset{(\text{ in } \overline{\mathcal{A}}_1)}{\mathsf{O}_1}$$

Therefore,

(a2) a measurement of an observable O_1 (at time $t = 1$) for a pure state ρ_0 (at time $t = 0$) $\in \mathfrak{S}^p(\mathcal{A}_0^*)$ is represented by

$$\mathsf{M}_{\overline{\mathcal{A}}_0}(\Phi_{0,1}\mathsf{O}_1, S_{[\rho_0]})$$

Thus, Axiom 1 (measurement: §2.7) says that

(a3) the probability that a measured value belongs to $\Xi (\in \mathcal{F})$ is given by

$$_{\mathcal{A}_0^*}\Big(\rho_0, \Phi_{0,1}(F_1(\Xi)\Big)_{\overline{\mathcal{A}}_0} \tag{11.4}$$

Explanation 11.3. **[the Schrödinger picture].** The Schrödinger picture is just the following (b):

(b1) **To identify a pure state $\Phi_{0,1}^*\rho_0 (\in \mathfrak{S}^p(\mathcal{A}_1^*))$ with $\rho_0 (\in \mathfrak{S}^p(\mathcal{A}_0^*))$,** That is,

$$\mathcal{A}_0^* \supset \mathfrak{S}^p(\mathcal{A}_0^*) \ni \rho_0 \xrightarrow[\text{identification}]{\Phi_{0,1}^*} \Phi_{0,1}^* \rho_0 \in \mathfrak{S}^p(\mathcal{A}_1^*) \subset \mathcal{A}_1^*$$

Therefore, Axiom 1 (measurement: §2.7) says that (b2) a measurement of an observable O_1 (at time $t=1$) for a pure state ρ_0 (at time $t=0$) $\in \mathfrak{S}^p(\mathcal{A}_1^*)$ is represented by

$$\mathsf{M}_{\overline{\mathcal{A}}_1}(\mathsf{O}_1, S_{[\Phi_{0,1}^*\rho_0]})$$

Thus,

(a3) the probability that a measured value belongs to $\Xi (\in \mathcal{F})$ is given by

$$ {}_{\mathcal{A}_1^*}\Big(\Phi_{0,1}^*\rho_0, F_1(\Xi)\Big)_{\overline{\mathcal{A}}_1} \tag{11.5}$$

which is equal to

$$ {}_{\mathcal{A}_0^*}\Big(\rho_0, \Phi_{0,1}(F_1(\Xi))\Big)_{\overline{\mathcal{A}}_0} \tag{11.6}$$

In the above sense (i.e., (11.5) and (11.6)), we conclude that, under the condition (A_1),

the Heisenberg picture and the Schrödinger picture are equivalent

That is,

$$\underset{\text{(Heisenberg picture)}}{\boxed{\mathsf{M}_{\overline{\mathcal{A}}_0}(\Phi_{0,1}\mathsf{O}_1, S_{[\rho_0]})}} \quad \underset{\text{(identification)}}{\longleftrightarrow} \quad \underset{\text{(Schrödenger picture)}}{\boxed{\mathsf{M}_{\overline{\mathcal{A}}_1}(\mathsf{O}_1, S_{[\Phi_{0,1}^*\rho_0]})}} \tag{11.7}$$

Remark 11.4. In the above, the conditions (A_1) is indispensable, that is,

(A_1) Consider a deterministic causal operator $\Phi_{0,1} : \overline{\mathcal{A}}_1 \to \overline{\mathcal{A}}_0$.

Without the deterministic conditions (A_1), the Schrödinger picture can not be formulated completely. That is because $\Phi_{0,1}^*\rho_0$ is not necessarily a pure state. In this sense, we consider that

- $\begin{cases} \text{the Heisenberg picture is formal} \\ \text{the Schrödinger picture is makeshift} \end{cases}$

11.2 The wave function collapse (i.e., the projection postulate)

The lingistic interpretation says that the post measurement state is meaningless. However, considering a tricky measurement, we can realize the wave function collapse. In this section, we shall explain this idea in the following paper:

- [49] S. Ishikawa, *Linguistic interpretation of quantum mechanics; Projection Postulate,* Journal of quantum information science, Vol. 5, No.4 , 150-155, 2015,
DOI: 10.4236/jqis.2015.54017
(`http://www.scirp.org/Journal/PaperInformation.aspx?PaperID=62464`)

11.2.1 Problem: The von Neumann-Lüders projection postulate

Let $[\mathcal{C}(H), B(H)]_{B(H)}$ be a quantum basic structure. Let Λ be a countable set.

Consider the projection valued observable $\mathsf{O}_P = (\Lambda, 2^\lambda, P)$ in $B(H)$. Put

$$P_\lambda = P(\{\lambda\}) \qquad (\forall \lambda \in \Lambda) \tag{11.8}$$

Axiom 1 (measurement; §2.7) says:

(A_1) The probability that a measured value λ_0 $(\in \Lambda)$ is obtained by the measurement $\mathsf{M}_{B(H)}(\mathsf{O}_P$ $:=(\Lambda, 2^\lambda, P)$, $S_{[\rho]})$ is given by

$$\mathrm{Tr}_H(\rho P_{\lambda_0})(= \langle u, P_{\lambda_0} u\rangle = \|P_{\lambda_0} u\|^2), \quad (\text{ where } \rho = |u\rangle\langle u|) \tag{11.9}$$

Also, the von Neumann-Lüders projection postulate (in the Copenhagen interpretation, *cf.* [70, 57]) says:

(A_2) When a measured value λ_0 $(\in \Lambda)$ is obtained by the measurement $\mathsf{M}_{B(H)}(\mathsf{O}_P := (\Lambda, 2^\lambda, P)$, $S_{[\rho]})$, the post-measurement state $\rho_{\rm post}$ is given by

$$\rho_{\rm post} = \frac{P_{\lambda_0}|u\rangle\langle u|P_{\lambda_0}}{\|P_{\lambda_0}u\|^2} \tag{11.10}$$

And therefore, when a next measurement $\mathsf{M}_{B(H)}(\mathsf{O}_F := (X, \mathcal{F}, F)$, $S_{[\rho_{\rm post}]})$ is taken (where O_F is arbitrary observable in $B(H)$), the probability that a measured value belongs to $\Xi (\in \mathcal{F})$ is given by

$$\mathrm{Tr}_H(\rho_{\rm post} F(\Xi))\Big(= \langle \frac{P_{\lambda_0}u}{\|P_{\lambda_0}u\|}, F(\Xi)\frac{P_{\lambda_0}u}{\|P_{\lambda_0}u\|}\rangle\Big) \tag{11.11}$$

Problem 11.5. In the linguistic interpretation, the phrase:"post-measurement state" in the (A_2) is meaningless. Also, the above $(=(\mathrm{A}_1)+(\mathrm{A}_2))$ is equivalent to the simultaneous measurement $\mathsf{M}_{B(H)}(\mathsf{O}_F \times \mathsf{O}_P, S_{[\rho]})$, which does not exist in the case that O_P and O_F do not commute. Hence the (A_2) is meaningless in general. Therefore, we have the following problem:

(B) Instead of the $\mathsf{O}_F \times \mathsf{O}_P$ in $\mathsf{M}_{B(H)}(\mathsf{O}_F \times \mathsf{O}_P, S_{[\rho]})$, what observable should be chosen?

In the following section, I answer this problem within the framework of the linguistic interpretation.

11.2.2 The derivation of von Neumann-Lüders projection postulate in the linguistic interpretation

Consider two basic structure $[\mathcal{C}(H), B(H)]_{B(H)}$ and $[\mathcal{C}(H\otimes K), B(H\otimes K)]_{B(H\otimes K)}$. Let $\{P_\lambda \mid \lambda \in \Lambda\}$ be as in Section 2.1, and let $\{e_\lambda\}_{\lambda\in\Lambda}$ be a complete orthonormal system in a Hilbert space K. Define the predual Markov operator $\Psi_* : Tr(H) \to Tr(H \otimes K)$ by, for any $u \in H$,

$$\Psi_*(|u\rangle\langle u|) = |\sum_{\lambda\in\Lambda}(P_\lambda u \otimes e_\lambda)\rangle\langle\sum_{\lambda\in\Lambda}(P_\lambda u \otimes e_\lambda)| \tag{11.12}$$

or

$$\Psi_*(|u\rangle\langle u|) = \sum_{\lambda\in\Lambda} |P_\lambda u \otimes e_\lambda\rangle\langle P_\lambda u \otimes e_\lambda| \tag{11.13}$$

Thus the Markov operator $\Psi : B(H \otimes K) \to B(H)$ (in Axiom 2) is defined by $\Psi = (\Psi_*)^*$. Define the observable $\mathsf{O}_G = (\Lambda, 2^\Lambda, G)$ in $B(K)$ such that

$$G(\{\lambda\}) = |e_\lambda\rangle\langle e_\lambda| \qquad (\lambda \in \Lambda)$$

Let $\mathsf{O}_F = (X, \mathcal{F}, F)$ be arbitrary observable in $B(H)$. Thus, we have the tensor observable $\mathsf{O}_F \otimes \mathsf{O}_G = (X \times \Lambda, \mathcal{F} \boxtimes 2^\Lambda, F \otimes G)$ in $B(H \otimes K)$, where $\mathcal{F} \boxtimes 2^\Lambda$ is the product σ-field.

Fix a pure state $\rho = |u\rangle\langle u|$ $(u \in H, \|u\|_H = 1)$. Consider the measurement $\mathsf{M}_{B(H)}(\Psi(\mathsf{O}_F \otimes \mathsf{O}_G), S_{[\rho]})$. Then, we see that

(C) the probability that a measured value (x, λ) obtained by the measurement $\mathsf{M}_{B(H)}(\Psi(\mathsf{O}_F \otimes \mathsf{O}_G),\ S_{[\rho]})$ belongs to $\Xi \times \{\lambda_0\}$ is given by

$$\begin{aligned}
&\mathrm{Tr}_H[(|u\rangle\langle u|)\Psi(F(\Xi) \otimes G(\{\lambda_0\}))] = {}_{Tr(H)}\big(|u\rangle\langle u|, \Psi(F(\Xi) \otimes G(\{\lambda_0\}))\big)_{B(H)} \\
=&{}_{Tr(H\otimes K)}\big(\Psi_*(|u\rangle\langle u|), F(\Xi) \otimes G(\{\lambda_0\})\big)_{B(H\otimes K)} = \mathrm{Tr}_{H\otimes K}[(\Psi_*(|u\rangle\langle u|))(F(\Xi) \otimes G(\{\lambda_0\}))]
\end{aligned}$$

$$
\begin{aligned}
=&\mathrm{Tr}_{H\otimes K}[(|\sum_{\lambda\in\Lambda}(P_\lambda u\otimes e_\lambda)\rangle\langle\sum_{\lambda\in\Lambda}(P_\lambda u\otimes e_\lambda)|)(F(\Xi)\otimes|e_{\lambda_0}\rangle\langle e_{\lambda_0}|)]\\
=&\langle P_{\lambda_0}u, F(\Xi)P_{\lambda_0}u\rangle \quad (\forall\Xi\in\mathcal{F})
\end{aligned}
$$

(In a similar way, the same result is easily obtained in the case of (7)).
Thus, we see the following.

(D$_1$) if $\Xi = X$, then

$$
\mathrm{Tr}_H[(|u\rangle\langle u|)\Psi(F(X)\otimes G(\{\lambda_0\}))] = \langle P_{\lambda_0}u, P_{\lambda_0}u\rangle = \|P_{\lambda_0}u\|^2 \tag{11.14}
$$

(D$_2$) in case that a measured value (x,λ) belongs to $X\times\{\lambda_0\}$, the conditional probability such that $x\in\Xi$ is given by

$$
\frac{\langle P_{\lambda_0}u, F(\Xi)P_{\lambda_0}u\rangle}{\|P_{\lambda_0}u\|^2}\Big(= \langle\frac{P_{\lambda_0}u}{\|P_{\lambda_0}u\|}, F(\Xi)\frac{P_{\lambda_0}u}{\|P_{\lambda_0}u\|}\rangle\Big) \quad (\forall\Xi\in\mathcal{F}) \tag{11.15}
$$

where it should be recalled that O_F is arbitrary. Also note that the above (i.e., the projection postulate (D)) is a consequence of Axioms 1 and 2.

Considering the correspondence: (A) $\Leftrightarrow$ (D), that is,

$$
\mathsf{M}_{B(H)}(\mathsf{O}_P, S_{[\rho]})\ \Big(\text{or, meaningless } \mathsf{M}_{B(H)}(\mathsf{O}_F\times\mathsf{O}_P, S_{[\rho]})\ \Big) \Leftrightarrow \mathsf{M}_{B(H)}(\Psi(\mathsf{O}_F\otimes\mathsf{O}_G), S_{[\rho]}),
$$

namely,

$$
(11.9)\Leftrightarrow(11.14),\quad (11.11)\Leftrightarrow(11.15)
$$

there is a reason to assume that the true meaning of the (A) is just the (D). Also, note the taboo phrase "**post-measurement state**" is not used in (D$_2$) but in (A$_2$). Hence, we obtain the answer of Problem 1 (i.e., $\Psi(\mathsf{O}_F\otimes\mathsf{O}_G)$).

Postulate 11.6. [Projection postulate] In the sense of the (D$_2$), the statement (A$_2$) is often used. That is, we often say:

(E) When a measured value λ_0 ($\in\Lambda$) is obtained by the measurement $\mathsf{M}_{B(H)}(\mathsf{O}_P := (\Lambda, 2^\lambda, P), S_{[\rho]})$, the post-measurement state ρ_{post} is given by

$$
\rho_{\mathrm{post}} = \frac{P_{\lambda_0}|u\rangle\langle u|P_{\lambda_0}}{\|P_{\lambda_0}u\|^2} \tag{11.16}
$$

Remark 11.7. So called Copenhagen interpretation may admit the post-measurement state (*cf.* [20]). Thus, in this case, readers may think that the post-measurement state is equal to $\frac{P_{\lambda_0}|u\rangle\langle u|P_{\lambda_0}}{\|P_{\lambda_0}u\|^2}$, which is obtained by the (D_2) (since O_F is arbitrary). However, this idea would not be generally approaved. That is because, if the post-measurement state is admitted, a series of problems occur, that is, "When is a measurement taken?", or "When does the wave function collapse happen?", which is beyond Axioms 1 and 2. Hence, the projection postulate is usually regarded as "postulate". On the other hand, in the linguistic interpretation, the projection postulate is completely clarified, and therefore, it should be regarded as a theorem. Recall the Wittgenstein's words: *"The limits of my language mean the limits of my world"*, or *"What we cannot speak about we must pass over in silence. "*

11.3 de Broglie's paradox(non-locality=faster-than-light)

In this section, we explain de Broglie's paradox in $B(L^2(\mathbb{R}))$ (*cf.* §2.10:de Broglie's paradox in $B(\mathbb{C}^2)$).

Putting $\mathbf{q} = (q_1, q_2, q_3) \in \mathbb{R}^3$, and

$$\nabla^2 = \frac{\partial^2}{\partial q_1^2} + \frac{\partial^2}{\partial q_2^2} + \frac{\partial^2}{\partial q_3^2}$$

consider Schrödinger equation (concerning one particle):

$$i\hbar \frac{\partial}{\partial t}\psi(\mathbf{q}, t) = \Big[\frac{-\hbar^2}{2m}\nabla^2 + V(\mathbf{q}, t)\Big]\psi(\mathbf{q}, t) \tag{11.17}$$

where, m is the mass of the particle, V is a potential energy.

In order to demonstrate in the picture, regard $\mathbb{R}^3$ as $\mathbb{R}$. Therefore, consider the Hilbert space $H = L^2(\mathbb{R}, dq)$. Putting $H_t = H$ $(t \in \mathbb{R})$, consider the quantum basic structure:

$$[\mathcal{C}(H) \subseteq B(H) \subseteq B(H)]$$

Equation 11.8. [Schrödinger equation]. There is a particle P (with mass m) in the box (that is, the closed interval $[0, 2] (\subseteq \mathbb{R})$). Let $\rho_{t_0} = |\psi_{t_0}\rangle\langle\psi_{t_0}| \in \mathfrak{S}^p(\mathcal{C}(H)^*)$ be an initial state (at time t_0) of the particle P. Let $\rho_t = |\psi_t\rangle\langle\psi_t|$ $(t_0 \le t \le t_1)$ be a state at time t, where $\psi_t = \psi(\cdot, t) \in H = L^2(\mathbb{R}, dq)$ satisfies the following Schrödinger equation:

$$\begin{cases} \text{initial state:}\psi(\cdot, t_0) = \psi_{t_0} \\ i\hbar\frac{\partial}{\partial t}\psi(q, t) = \Big[\frac{-\hbar^2}{2m}\frac{\partial^2}{\partial q^2} + V(q, t)\Big]\psi(q, t) \end{cases} \tag{11.18}$$

Consider the same situation in §10.5, i.e., a particle with the mass m in the box (i.e., the closed interval $[0, 2]$) in the one dimensional space $\mathbb{R}$.

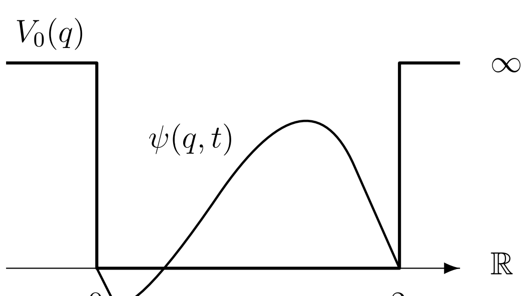


Figure 11.1(1)

Now let us partition the box $[0,2]]$ into $[0,1]]$ and $[1,2]$. That is, we change $V_0(q)$ to $V_1(q)$, where

$$V_1(q) = \begin{cases} 0 & (0 \le q < 1) \\ \infty & (q = 1) \\ 0 & (1 < q \le 2) \\ \infty & (\text{ otherwise }) \end{cases} \tag{11.19}$$

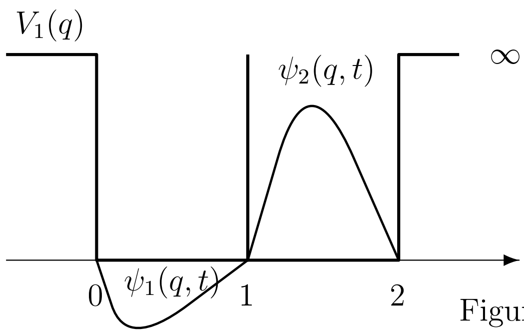


Figure 11.1(2)

Next, we carry the box $[0,1]$ $\big[$resp. the box $[1,2]\big]$ to New York (or, the earth) $\big[$resp. Tokyo (or, the polar star)$\big]$.

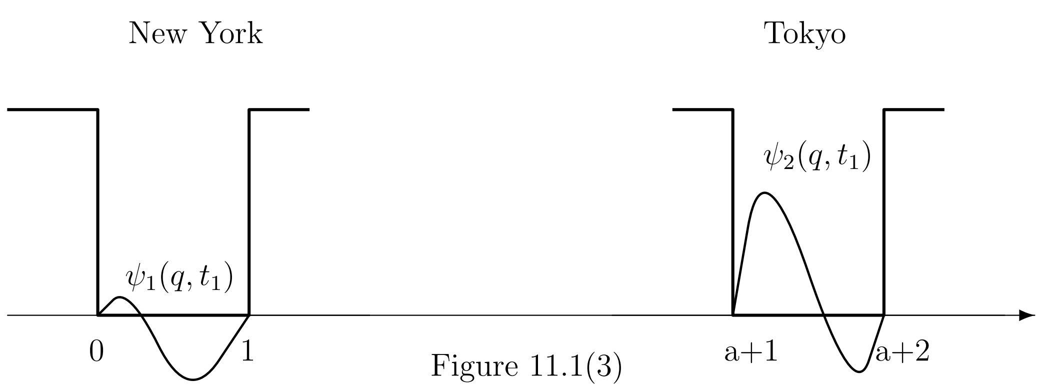


Figure 11.1(3)

Here, $1 \ll a$. Solving the Schrödinger equation (11.18), we see that

$$\psi_1(\cdot, t_1) + \psi_2(\cdot, t_1) = U_{t_0,t_1}\psi_{t_0}$$

where $U_{t_0,t_1} : L^2(\mathbb{R}_{t_1}) \to L^2(\mathbb{R}_{t_0})$ is the unitary operator. Define the causal operator $\Phi_{t_0,t_1} : B(L^2(\mathbb{R}_{t_2})) \to B(L^2(\mathbb{R}_{t_1}))$ by

$$\Phi_{t_0,t_1}(A) = U^*_{t_0,t_1} A U_{t_0,t_1} \qquad (\forall A \in B(L^2(\mathbb{R}_{t_2})))$$

Put $T = \{t_0, t_1\}$. And consider the observable $\mathsf{O} = (X = \{N, T.E\}, 2^X, F)$ in $B(L^2(\mathbb{R}_{t_1}))$ (where "N"=New York, "T"=Tokyo, "E"=elsewhere) such that

$$[F(\{N\})](q) = \begin{cases} 1 & 0 \le q < 1 \\ 0 & \text{elsewhere} \end{cases}, \quad [F(\{T\})](q) = \begin{cases} 1 & a+1 \le q < a+2 \\ 0 & \text{elsewhere} \end{cases},$$
$$[F(\{E\})](q) = 1 - [F(\{N\})](q) - [F(\{T\})](q).$$

Hence we have the measurement $\mathsf{M}_{B(L^2(\mathbb{R}_{t_0}))}\Big(\Phi_{t_0,t_1}\mathsf{O}, S_{[|\psi_{t_0}\rangle\langle\psi_{t_0}|]}\Big)$.

Conclusion 11.9.
In Heisenberg picture, we see, by Axiom 1 (measurement: §2.7), that

(A$_1$) the probability that a measured value $\begin{bmatrix} N \\ T \\ E \end{bmatrix}$ is obtained by the measurement $\mathsf{M}_{B(L^2(\mathbb{R}_{t_0}))}\Big(\Phi_{t_0,t_1}\mathsf{O}, S_{[|\psi_{t_0}\rangle\langle\psi_{t_0}|]}\Big)$ is given by

$$\begin{bmatrix} \langle u_{t_0}, \Phi_{t_0.t_1} F(\{N\}) u_{t_0}\rangle = \int_0^1 |\psi_1(q,t_1)|^2 dq \\ \langle u_{t_0}, \Phi_{t_0.t_1} F(\{T\}) u_{t_0}\rangle = \int_{a+1}^{a+2} |\psi_2(q,t_1)|^2 dq \\ \langle u_{t_0}, \Phi_{t_0.t_1} F(\{E\}) u_{t_0}\rangle = 0 \end{bmatrix}.$$

Also, In Schrödinger picture, we see Axiom 1 (measurement: §2.7), that

(A$_2$) the probability that a measured value $\begin{bmatrix} N \\ T \\ E \end{bmatrix}$ is obtained by the measurement $\mathsf{M}_{B(L^2(\mathbb{R}_{t_0}))}\Big(\mathsf{O}, S_{[\Phi^*_{t_0,t_1}(|\psi_{t_0}\rangle\langle\psi_{t_0}|)]}\Big)$ is given by

$$\begin{bmatrix} \mathrm{Tr}\Big(\Phi^*_{t_0,t_1}(|\psi_{t_0}\rangle\langle\psi_{t_0}|)\cdot F(\{N\})\Big) = \langle U_{t_0,t_1}\psi_{t_0}, F(\{N\}) U_{t_0,t_1}\psi_{t_0}\rangle = \int_0^1 |\psi_1(q,t_1)|^2 dq \\ \mathrm{Tr}\Big(\Phi^*_{t_0,t_1}(|\psi_{t_0}\rangle\langle\psi_{t_0}|)\cdot F(\{T\})\Big) = \langle U_{t_0,t_1}\psi_{t_0}, F(\{T\}) U_{t_0,t_1}\psi_{t_0}\rangle = \int_{a+1}^{a+2} |\psi_2(q,t_1)|^2 dq \\ \mathrm{Tr}\Big(\Phi^*_{t_0,t_1}(|\psi_{t_0}\rangle\langle\psi_{t_0}|)\cdot F(\{E\})\Big) = \langle U_{t_0,t_1}\psi_{t_0}, F(\{E\}) U_{t_0,t_1}\psi_{t_0}\rangle = 0 \end{bmatrix}$$

Note that the probability that we find the particle in the box $[0,1]$ $\big[$resp. the box $[a+1, a+2]\big]$ is given by $\int_{\mathbb{R}} |\psi_1(q,t_1)|^2 dq$ $\big[$resp. $\int_{\mathbb{R}} |\psi_2(q,t_1)|^2 dq\big]$. That is,

$$\textbf{(A}_1\textbf{)=(A}_2\textbf{)}$$

Remark 11.10. In the above, assume that we get a measured value "N", that is, we open the box $[0,1]$ at New York. And assume that we find the particle in the box $[0,1]$. Then, in the sense of Postulate 11.6, we say that at the moment the wave function ψ_2 vanishes. That is,

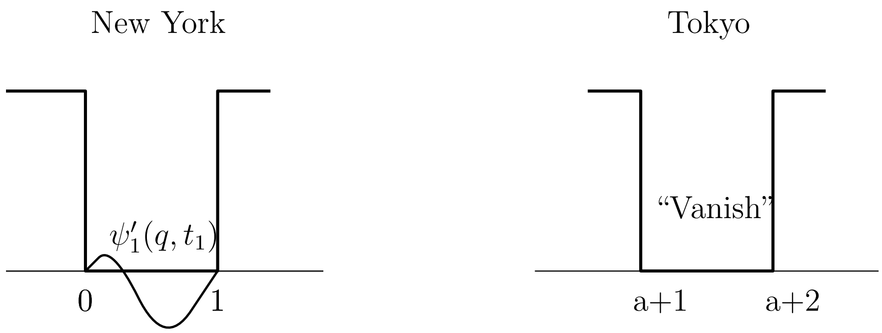


Figure 11.1(4) (The wave function after measurement)

where

$$\psi_1'(q,t_1) = \frac{\psi_1(q,t_1)}{\|\psi_1'(\cdot,t_1)\|}.$$

Thus, we may consider "the collapse of wave function" such as

$$\psi_1(\cdot,t_1) + \psi_2(\cdot,t_1) \xrightarrow[\text{the collapse of wave function}]{} \psi_1'(\cdot,t_1) \tag{11.20}$$

Also, note that New York [resp. Tokyo] may be the earth [resp. the polar star]. Thus,

- the above argument (in both cases (A_1) and (A_2)) implies that there is something faster than light.

This is called "the de Broglie paradox"(*cf.* [12, 68]). This is a true paradox, which is not clarified even in quantum language.

11.4 Quantum Zeno effect

This section is extracted from

- Ref. [39]: S. Ishikawa; Heisenberg uncertainty principle and quantum Zeno effects in the linguistic interpretation of quantum mechanics (arXiv:1308.5469 [quant-ph] 2014)

11.4.1 Quantum decoherence: non-deterministic sequential causal operator

Let us start from the review of Section 10.6.2 (quantum decoherence). Consider the quantum basic structure:

$$[\mathcal{C}(H) \subseteq B(H) \subseteq B(H)]$$

Let $\mathbb{P} = [P_n]_{n=1}^{\infty}$ be the spectrum decomposition in $B(H)$, that is,

$$P_n \text{ is a projection, and}, \sum_{n=1}^{\infty} P_n = I$$

Define the operator $(\Psi_{\mathbb{P}})_* : \mathcal{T}r(H) \to \mathcal{T}r(H)$ such that

$$(\Psi_{\mathbb{P}})_*(|u\rangle\langle u|) = \sum_{n=1}^{\infty} |P_n u\rangle\langle P_n u| \quad (\forall u \in H)$$

Clearly we see

$$\langle v, (\Psi_{\mathbb{P}})_*(|u\rangle\langle u|)v\rangle = \langle v, (\sum_{n=1}^{\infty} |P_n u\rangle\langle P_n u|)v\rangle = \sum_{n=1}^{\infty} |\langle v, |P_n u\rangle|^2 \geq 0 \qquad (\forall u, v \in H)$$

and,

$$\begin{aligned} & \mathrm{Tr}((\Psi_{\mathbb{P}})_*(|u\rangle\langle u|)) \\ =& \mathrm{Tr}(\sum_{n=1}^{\infty} |P_n u\rangle\langle P_n u|) = \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} |\langle e_k, P_n u\rangle|^2 = \sum_{n=1}^{\infty} \|P_n u\|^2 = \|u\|^2 \qquad (\forall u \in H) \end{aligned}$$

And so,

$$(\Psi_{\mathbb{P}})_*(\mathcal{T}r_{+1}^p(H)) \subseteq \mathcal{T}r_{+1}(H)$$

Therefore,

($\sharp$) $\Psi_{\mathbb{P}}(= ((\Psi_{\mathbb{P}})_*)^*) : B(H) \to B(H)$ is a causal operator, but it is not deterministic.

In this note, a non-deterministic (sequential) causal operator is called a **quantum decoherence**.

Example 11.11. [Quantum decoherence in quantum Zeno effect *cf.* [36]]. Further consider a causal operator $(\Psi_S^{\Delta t})_* : \mathcal{T}r(H) \to \mathcal{T}r(H)$ such that

$$(\Psi_S^{\Delta t})_*(|u\rangle\langle u|) = |e^{-\frac{i\mathcal{H}\Delta t}{\hbar}}u\rangle\langle e^{-\frac{i\mathcal{H}\Delta t}{\hbar}}u| \quad (\forall u \in H)$$

where the Hamiltonian $\mathcal{H}$ (*cf.* (10.22)) is, for example, defined by

$$\mathcal{H} = \Big[\frac{-\hbar^2}{2m}\frac{\partial^2}{\partial q^2} + V(q,t)\Big]$$

Let $\mathbb{P} = [P_n]_{n=1}^{\infty}$ be the spectrum decomposition in $B(H)$, that is, for each n, $P_n \in B(H)$ is a projection such that

$$\sum_{n=1}^{\infty} P_n = I$$

Define the $(\Psi_{\mathbb{P}})_* : \mathcal{T}r(H) \to \mathcal{T}r(H)$ such that

$$(\Psi_{\mathbb{P}})_*(|u\rangle\langle u|) = \sum_{n=1}^{\infty} |P_n u\rangle\langle P_n u| \quad (\forall u \in H)$$

Also, we define the Schrödinger time evolution $(\Psi_S^{\Delta t})_* : \mathcal{T}r(H) \to \mathcal{T}r(H)$ such that

$$(\Psi_S^{\Delta t})_*(|u\rangle\langle u|) = |e^{-\frac{i\mathcal{H}\Delta t}{\hbar}}u\rangle\langle e^{-\frac{i\mathcal{H}\Delta t}{\hbar}}u| \quad (\forall u \in H)$$

where $\mathcal{H}$ is the Hamiltonian (10.21). Consider $t = 0, 1$. Putting $\Delta t = \frac{1}{N}$, $H = H_0 = H_1$, we can define the $(\Phi_{0,1}^{(N)})_* : Tr(H_0) \to Tr(H_1)$ such that

$$(\Phi_{0,1}^{(N)})_* = ((\Psi_S^{1/N})_*(\Psi_{\mathbb{P}})_*)^N$$

which induces the Markov operator $\Phi_{0,1}^{(N)} : B(H_1) \to B(H_0)$ as the dual operator $\Phi_{0,1}^{(N)} = ((\Phi_{0,1}^{(N)})_*)^*$. Let $\rho = |\psi\rangle\langle\psi|$ be a state at time 0. Let $\mathsf{O}_1 :=(X, \mathcal{F}, F)$ be an observable in $B(H_1)$. Then, we see

$$\underset{\boxed{B(H_0)}}{\overset{\rho=|\psi\rangle\langle\psi|}{}} \xleftarrow[\Phi_{0,1}^{(N)}]{} \underset{\mathsf{O}_1 :=(X,\mathcal{F},F)}{\boxed{B(H_1)}}$$

Thus, we have a measurement:

$$\mathsf{M}_{B(H_0)}(\Phi_{0,1}^{(N)}\mathsf{O}_1, S_{[\rho]})$$

$\big($ or more precisely, $\mathsf{M}_{B(H_0)}(\Phi_{0,1}^{(N)}\mathsf{O} :=(X, \mathcal{F}, \Phi_{0,1}^{(N)}F),\ S_{[|\psi\rangle\langle\psi|]})$ $\big)$. Here, Axiom 1 (§2.7) says that

(A) the probability that the measured value obtained by the measurement belongs to $\Xi(\in \mathcal{F})$ is given by

$$\mathrm{Tr}(|\psi\rangle\langle\psi| \cdot \Phi_{0,1}^{(N)} F(\Xi)) \tag{11.21}$$

Now we shall explain "quantum Zeno effect" in the following example.

Example 11.12. [Quantum Zeno effect] Let $\psi \in H$ such that $\|\psi\| = 1$. Define the spectrum decomposition

$$\mathbb{P} = [P_1(= |\psi\rangle\langle\psi|), P_2(= I - P_1)] \tag{11.22}$$

And define the observable $\mathsf{O}_1 :=(X, \mathcal{F}, F)$ in $B(H_1)$ such that

$$X = \{x_1, x_2\}, \qquad \mathcal{F} = 2^X$$

and

$$F(\{x_1\}) = |\psi\rangle\langle\psi|(= P_1), F(\{x_2\}) = I - |\psi\rangle\langle\psi|(= P_2),$$

Now we can calculate (11.21)(i.e., the probability that a measured value x_1 is obtained) as follows.

$$\begin{aligned}(11.21) &= \langle\psi, ((\Psi_S^{1/N})_*(\Psi_{\mathbb{P}})_*)^N(|\psi\rangle\langle\psi|)\psi\rangle \\ &\geq |\langle\psi, e^{-\frac{i\mathcal{H}}{\hbar N}}\psi\rangle\langle\psi, e^{\frac{i\mathcal{H}}{\hbar N}}\psi\rangle|^N \\ &\approx \Big(1 - \frac{1}{N^2}\big(||(\frac{\mathcal{H}}{\hbar})\psi||^2 - |\langle\psi, (\frac{\mathcal{H}}{\hbar})\psi\rangle|^2\big)\Big)^N \to 1 \\ &\qquad (N \to \infty)\end{aligned} \tag{11.23}$$

Thus, if N is sufficiently large, we see that

$$\mathsf{M}_{B(H_0)}(\Phi_{0,1}^{(N)}\mathsf{O}_1, S_{[|\psi\rangle\langle\psi|]}) \approx \mathsf{M}_{B(H_0)}(\Phi_I\mathsf{O}_1, S_{[|\psi\rangle\langle\psi|]})$$

(where $\Phi_I : B(H_1) \to B(H_0)$ is the identity map)

$$= \mathsf{M}_{B(H_0)}(\mathsf{O}_1, S_{[|\psi\rangle\langle\psi|]})$$

Hence, we say, roughly speaking in terms of the Schrödinger picture, that

the state $|\psi\rangle\langle\psi|$ does not move.

Remark 11.13. The above argument is motivated by B. Misra and E.C.G. Sudarshan [60]. However, the title of their paper: "The Zeno's paradox in quantum theory" is not proper. That is because

(B) the spectrum decomposition $\mathbb{P}$ should not be regarded as an observable (or moreover, measurement).

The effect in Example 11.12 should be called "brake effect" and not "watched pot effect".

11.5 Schrödinger's cat, Wigner's friend and Laplace's demon

11.5.1 Schrödinger's cat and Wigner's friend

Let us explain Schrödinger's cat paradox in the Schrödinger picture.

Problem 11.14. [Schrödinger's cat]

(a) Suppose we put a cat in a cage with a radioactive atom, a Geiger counter, and a poison gas bottle; further suppose that the atom in the cage has a half-life of one hour, a fifty-fifty chance of decaying within the hour. If the atom decays, the Geiger counter will tick; the triggering of the counter will get the lid off the poison gas bottle, which will kill the cat. If the atom does not decay, none of the above things happen, and the cat will be alive.

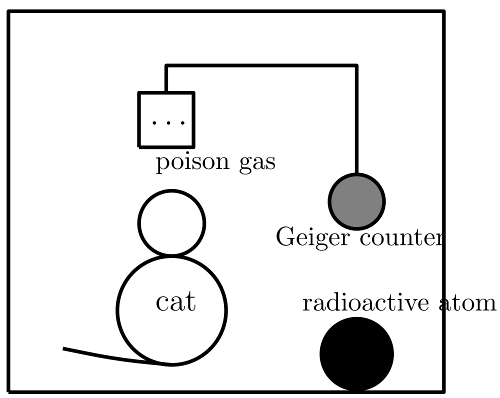


Figure 11.2: Schrödinger's cat

Here, we have the following question:

(b) Is the cat dead or alive after 1 hour ($= 60^{60}$ seconds) ?
Of course, we say that it is half-and-half whether the cat is alive. However, our problem is

Clarify the meaning of "half-and-half"

♠**Note 11.1.** [Wigner's friend]: Instead of the above (b), we consider as follows.

(b′) after one hour, Wigner's friend look at the inside of the box, and thus, he knows whether the cat is dead or alive after one hour. And further, after two hours, Wigner's friend informs you of the fact. How is the cat ?

This problem is not difficult. That is because the linguistic interpretation says that "the moment you measured" is out of quantum language. Recall the spirit of the linguistic world-view (i.e., Wittgenstein's words) such as

The limits of my language mean the limits of my world

and

What we cannot speak about we must pass over in silence.

11.5.2 The usual answer

Answer 11.15. [The first answer to Problem11.14(i.e., the pure state, projection postulate)].

Put $\mathbf{q} = (q_{11}, q_{12}, q_{13}, q_{21}, q_{22}, q_{23}, \ldots, q_{n1}, q_{n2}, q_{n3}) \in \mathbb{R}^{3n}$. And put

$$\nabla_i^2 = \frac{\partial^2}{\partial q_{i1}^2} + \frac{\partial^2}{\partial q_{i2}^2} + \frac{\partial^2}{\partial q_{i3}^2}$$

Consider the quantum system basic structure:

$$[\mathcal{C}(H) \subseteq B(H) \subseteq B(H)] \qquad (\text{ where, } H = L^2(\mathbb{R}^{3n}, d\mathbf{q}) \ )$$

And consider the Schrödinger equation (concerning n-particles system):

$$\begin{cases} i\hbar \frac{\partial}{\partial t}\psi(\mathbf{q}, t) = \Big[\sum_{i=1}^{n} \frac{-\hbar^2}{2m_i} \nabla_i^2 + V(\mathbf{q}, t) \Big] \psi(\mathbf{q}, t) \\ \\ \psi_0(\mathbf{q}) = \psi(\mathbf{q}, 0) : \text{initial condition} \end{cases} \tag{11.24}$$

where, m_i is the mass of a particle P_i, V is a potential energy.

If we believe in quantum mechanics, it suffices to solve this Schrödinger equation (11.24). That is,

(A$_1$) Assume that the wave function $\psi(\cdot, 60^2) = U_{0,60^2}\psi_0$ after one hour (i.e., 60^2 seconds) is calculated. Then, the state ρ_{60^2} $(\in \mathcal{T}r_{+1}^{p}(H))$ after 60^2 seconds is represented by

$$\rho_{60^2} = |\psi_{60^2}\rangle\langle\psi_{60^2}| \tag{11.25}$$

(where, $\psi_{60^2} = \psi(\cdot, 60^2)$).

Now, define the observable $\mathsf{O} = (X = \{\text{life}, \text{death}\}, 2^X, F)$ in $B(H)$ as follows.

(A$_2$) that is, putting

$$V_{\text{life}}(\subseteq H) = \Big\{ u \in H \mid \text{“ the state } \frac{|u\rangle\langle u|}{\|u\|^2} \text{”} \Leftrightarrow \text{“cat is alive”} \Big\}$$
$$V_{\text{death}}(\subseteq H) = \text{the orthogonal complement space of } V_{\text{life}}$$
$$= \{ u \in H \mid \langle u, v \rangle = 0 \ \ (\forall v \in V_{\text{life}}) \}$$

define $F(\{\text{life}\})(\in B(H))$ is the projection of the closed subspace V_{life} and $F(\{\text{death}\}) = I - F(\{\text{life}\})$,

Here,

(A$_3$) Consider the measurement $\mathsf{M}_{B(H)}(\mathsf{O} = (X, 2^X, F), S_{[\rho_{60^2}]})$. The probability that a measured value $\begin{bmatrix} \text{life} \\ \text{death} \end{bmatrix}$ is obtained is given by

$$\begin{bmatrix} \mathcal{T}r(H)\Big(\rho_{60^2}, F(\{\text{life}\})\Big)_{B(H)} = \langle \psi_{60^2}, F(\{\text{life}\})\psi_{60^2} \rangle = 0.5 \\ \mathcal{T}r(H)\Big(\rho_{60^2}, F(\{\text{death}\})\Big)_{B(H)} = \langle \psi_{60^2}, F(\{\text{death}\})\psi_{60^2} \rangle = 0.5 \end{bmatrix}$$

Therefore, we can assure that

$$\psi_{60^2} = \frac{1}{\sqrt{2}}(\psi_{\text{life}} + \psi_{\text{death}}) \tag{11.26}$$
$$(\text{where, } \ \psi_{\text{life}} \in V_{\text{life}}, \|\psi_{\text{life}}\| = 1 \quad \psi_{\text{death}} \in V_{\text{death}}, \|\psi_{\text{death}}\| = 1)$$

Hence. we can conclude that

(A$_4$) the state (or, wave function) of the cat (after one hour) is represented by (11.26), that is,

$$\frac{\text{“Fig.}(\sharp_1)\text{”} + \text{“Fig.}(\sharp_2)\text{”}}{\sqrt{2}}$$

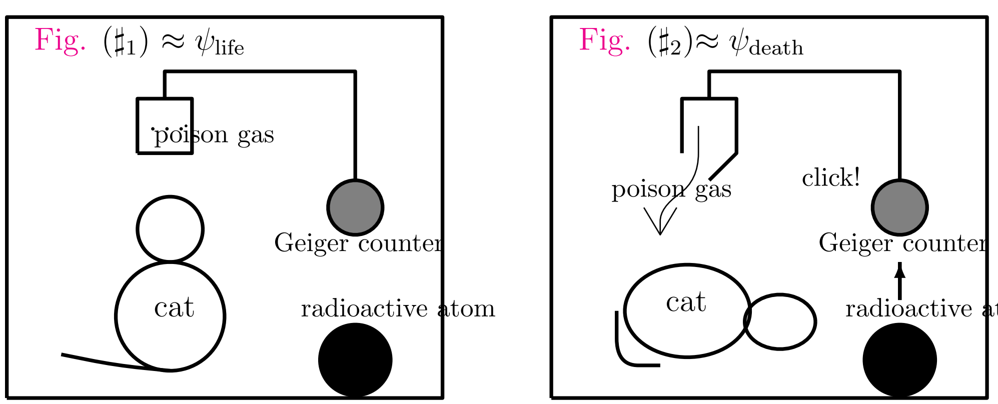


Figure 11.3: Schrödinger's cat(half and half)

And,

(A$_5$) After one hour (i.e, to the moment of opening a window), It is decided "the cat is dead" or "the cat is vigorously alive." That is,

$$\text{“half-dead”}\Big(= \frac{1}{2}(|\psi_{\text{life}} + \psi_{\text{death}}\rangle\langle\psi_{\text{life}} + \psi_{\text{death}}|)\Big)$$

in the sense of Postulate 11.6 (precisely speaking, by the misunderstanding of Postulate 11.6),

$$\xrightarrow[\text{the collapse of wave function}]{\text{to the moment of opening a window}} \begin{cases} \text{“alive”}(= |\psi_{\text{life}}\rangle\langle\psi_{\text{life}}|) \\ \\ \text{“dead”}(= |\psi_{\text{death}}\rangle\langle\psi_{\text{death}}|) \end{cases}$$

□

11.5.3 The answer by quantum decoherence

Answer 11.16. [The second answer to Problem11.14 (i.e., decoherence)].

In quantum language, the quantum decoherence is permitted. That is, we can assume that

(B$_1$) the state ρ'_{60^2} after one hour is represented by the following mixed state

$$\rho'_{60^2} = \frac{1}{2}\Big(|\psi_{\text{life}}\rangle\langle\psi_{\text{life}}| + |\psi_{\text{death}}\rangle\langle\psi_{\text{death}}|\Big)$$

That is, we can assume the decoherent causal operator $\Phi_{0,60^2} : B(H) \to B(H)$ such that

$$(\Phi_{0,60^2})_*(\rho_0) = \rho'_{60^2}$$

Here, consider the measurement $\mathsf{M}_{B(H)}(\mathsf{O} = (X, 2^X, F), S_{[\rho'_{60^2}]})$, or, its Heisenberg picture $\mathsf{M}_{B(H)}(\Phi_{0,60^2}\mathsf{O} = (X, 2^X, \Phi_{0,60^2}F), S_{[\rho'_0]})$. Of course we see:

(B$_2$) The probability that a measured value $\begin{bmatrix} \text{life} \\ \text{death} \end{bmatrix}$ is obtained by the measurement $\mathsf{M}_{B(H)}(\Phi_{0,60^2}\mathsf{O} = (X, 2^X, \Phi_{0,60^2}F), S_{[\rho'_0]})$ is given by

$$\begin{bmatrix} \mathfrak{Tr}(H)\Big(\rho_0, \Phi_{0,60^2}F(\{\text{life}\})\Big)_{B(H)} = \langle\psi'_{60^2}, F(\{\text{life}\})\psi_{60^2}\rangle = 0.5 \\ \mathfrak{Tr}(H)\Big(\rho_0, \Phi_{0,60^2}F(\{\text{death}\})\Big)_{B(H)} = \langle\psi'_{60^2}, F(\{\text{death}\})\psi_{60^2}\rangle = 0.5 \end{bmatrix}$$

Also, "the moment of measuring" and "the collapse of wave function" are prohibited in the linguistic interpretation, but the statement (B$_2$) is within quantum language. □

Summary 11.17. [**Schrödinger's cat in quantum language**]
Here, let us examine

Answer11.15 :(A_5) v.s. **Answer11.16 :(B_2)**

(C_1) the answer (A_5) may be unnatural, but it is an argument which cannot be confuted,

On the other hand,

(C_2) the answer (B_2) is natural. but the non-deterministic time evolution is used.

Since the non-deterministic causal operator (i.e., quantum decoherence) is permitted in quantum language, we conclude that

(C_3) **Answer11.16:(B_2)** is superior to **Answer11.15:(A_1)**

For the reason that the non-deterministic causal operator (i.e., quantum decoherence) is permitted in quantum language, we add the following.

- If Newtonian mechanics is applied to the whole universe, Laplace's demon appears. Also, if Newtonian mechanics is applied to the microworld, chaos appears. This kind of supremacy of physics is not natural, and thus, we consider that these are out of "the limit of Newtonian mechanics"

And,

- when we want to apply Newton mechanics to phenomena out of "the limit of Newtonian mechanics", we often use the stochastic differential equation (and Brownian motion). This approach is called "dynamical system theory", which is not physics but metaphysics.

$$\underset{\text{physics}}{\boxed{\text{Newtonian mechanics}}} \xrightarrow[\text{linguistic turn}]{\text{out of the limits}} \underset{\text{metaphysics}}{\boxed{\text{dynamical system theory; statistics}}}$$

In the same sense, we consider that quantum mechanics has "the limit". That is,

- Schrödinger's cat is out of quantum mechanics.

And thus,

- When we want to apply quantum mechanics to phenomena out of "the limit of quantum mechanics", we often use the quantum decoherence. Although this approach is not physics but metaphysics, it is quite powerful.

$$\underset{\text{physics}}{\boxed{\text{quantum mechanics}}} \xrightarrow[\text{linguistic turn}]{\text{out of the limits}} \underset{\text{metaphysics}}{\boxed{\text{quantum language}}}$$

♠**Note 11.2.** If we know the present state of the universe and the kinetic equation (=the theory of everything), and if we calculate it, we can know everything (from past to future). There may be a reason to believe this idea. This intellect is often referred to as **Laplace's demon**. Laplace's demon is sometimes discussed as the realistic-view over which the degree passed. Thus, we consider the following correspondence:

$$\underset{\text{Newtonian mechanics}}{\boxed{\text{Laplace's Demon}}} \underset{\text{correspondence}}{\longleftrightarrow} \underset{\text{quantum mechanics}}{\boxed{\text{Schrödinger's cat in Answer 11.15}}}$$

11.6 Wheeler's Delayed choice experiment: "Particle or wave?" is a foolish question

This section is extracted from

(♯) [44] S. Ishikawa, *The double-slit quantum eraser experiments and Hardy's paradox in the quantum linguistic interpretation,* arxiv:1407.5143[quantum-ph],(2014)

11.6.1 "Particle or wave?" is a foolish question

In the conventional quantum mechanics, the question: "particle or wave?" may frequently appear. However, this is a foolish question.

On the other hand, the argument about the "particle vs. wave" is clear in quantum language. As seen in the following table, this argument is traditional:

Table 11.1: Particle vs. Wave in several world-views (*cf.* Table 2.1, Table 3.1)

World-views \ P or W	Particle(=symbol)	Wave(= mathematical representation)
Aristotle	hyle	eidos
Newton mechanics	point mass	state (=(position, momentum))
Statistics	population	parameter
Quantum mechanics	particle	state (≈ wave function)
Quantum language	system (=measuring object)	state

In the table 11.1, Newtonian mechanics (i.e., mass point ↔ state) may be easiest to understand. Thus, "particle" and "wave" are not confrontation concepts.

Concerning "particle or wave", we have the following statements:

(A_1) "Particle or wave" is a foolish question.

(A_2) Wheeler's delayed choice experiment is related to the question "particle or wave"

If so, it may be interesting to answer the following:

(A_3) How is Wheeler's delayed choice experiment described in terms of quantum mechanics?

This is the purpose of this section. And we answer it in the conclusion (H).

11.6.2 Preparation

Let us start from the review of Section 2.10 (de Broglie paradox in $B(\mathbb{C}^2)$)

Let H be a two dimensional Hilbert space, i.e., $H = \mathbb{C}^2$. Consider the basic structure

$$[B(\mathbb{C}^2) \subseteq B(\mathbb{C}^2) \subseteq B(\mathbb{C}^2)]$$

Let $f_1, f_2 \in H$ such that

$$f_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad f_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Put

$$u = \frac{f_1 + f_2}{\sqrt{2}}$$

Thus, we have the state $\rho = |u\rangle\langle u|$ $(\in \mathfrak{S}^p(B(\mathbb{C}^2)))$.

Let $U(\in B(\mathbb{C}^2))$ be an unitary operator such that

$$U = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\pi/2} \end{bmatrix}$$

and let $\Phi : B(\mathbb{C}^2) \to B(\mathbb{C}^2)$ be the homomorphism such that

$$\Phi(F) = U^* F U \qquad (\forall F \in B(\mathbb{C}^2))$$

Consider two observable $\mathsf{O}_f = (\{1,2\}, 2^{\{1,2\}}, F)$ and $\mathsf{O}_g = (\{1,2\}, 2^{\{1,2\}}, G)$ in $B(\mathbb{C}^2)$ such that

$$F(\{1\}) = |f_1\rangle\langle f_1|, \quad F(\{2\}) = |f_2\rangle\langle f_2|$$

and

$$G(\{1\}) = |g_1\rangle\langle g_1|, \quad G(\{2\}) = |g_2\rangle\langle g_2|$$

where

$$g_1 = \frac{f_1 + f_2}{\sqrt{2}}, \qquad g_2 = \frac{f_1 - f_2}{\sqrt{2}}$$

11.6.3 de Broglie's paradox in $B(\mathbb{C}^2)$ (No interference)

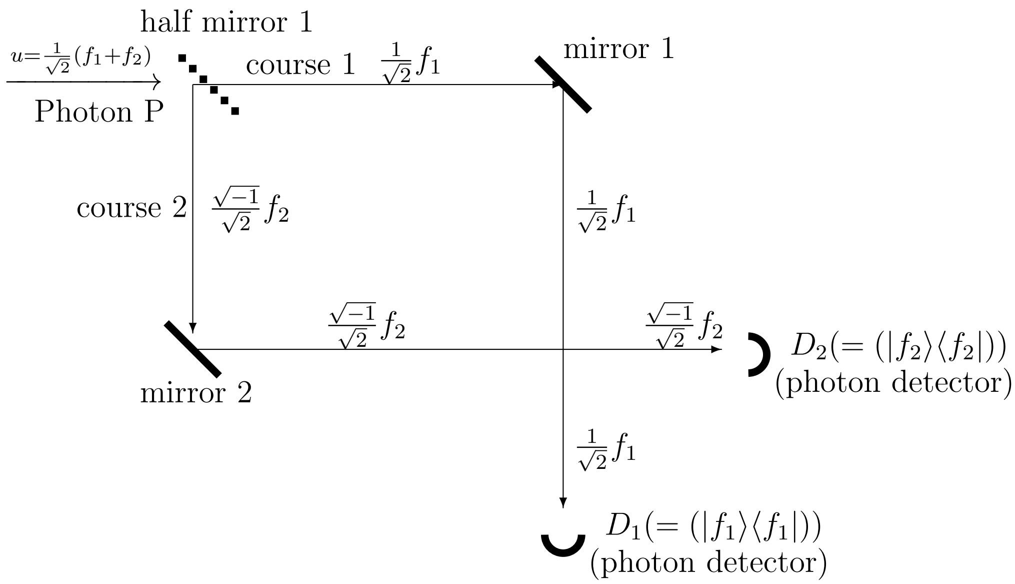


Figure 11.4(1). $[D_1 + D_2]$**=Observable**O_f

Now we shall explain, by the Schrödinger picture, Figure 11.4(1) as follows.

The photon P with the state $u = \frac{1}{\sqrt{2}}(f_1 + f_2)$ (precisely, $\rho = |u\rangle\langle u|$) rushed into the half-mirror 1,

(B$_1$) the f_1 part in $u = \frac{1}{\sqrt{2}}(f_1+f_2)$ passes through the half-mirror 1, and goes along the course 1. And it is reflected in the mirror 1, and goes to the photon detector D_1.

(B$_2$) the f_2 part in $u = \frac{1}{\sqrt{2}}(f_1 + f_2)$ rebounds on the half-mirror 1 (and strictly saying, the f_2 changes to $\sqrt{-1}f_2$, we are not concerned with it), and goes along the course 2. And it is reflected in the mirror 2, and goes to the photon detector D_2.

This is, by the Heisenberg picture, represented by the following measurement:

$$\mathsf{M}_{B(\mathbb{C}^2)}(\Phi\mathsf{O}_f, S_{[\rho]}) \tag{11.27}$$

Then, we see:

(C) the probability that $\begin{bmatrix} \text{a measured value 1} \\ \text{a measured value 2} \end{bmatrix}$ is obtained by $\mathsf{M}_{B(\mathbb{C}^2)}(\Phi\mathsf{O}_f, S_{[\rho]})$ is given by

$$\begin{bmatrix} \langle Uu, F(\{1\})Uu\rangle \\ \langle Uu, F(\{2\})Uu\rangle \end{bmatrix} = \begin{bmatrix} |\langle Uu, f_1\rangle|^2 \\ |\langle Uu, f_2\rangle|^2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \tag{11.28}$$

Remark 11.18. [Projection postulate] By the analogy of Section 11.2 (The projection postulate), Figure 11.4(1) is also described as follows. That is, putting $e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ $(\in \mathbb{C}^2)$, we have the observable $\mathsf{O}_E = (\{1,2\}, 2^{\{1,2\}}, E)$ in $B(\mathbb{C}^2)$ such that $E(\{1\}) = |e_1\rangle\langle e_1$ and $E(\{1\}) = |e_1\rangle\langle e_1$. Hence,

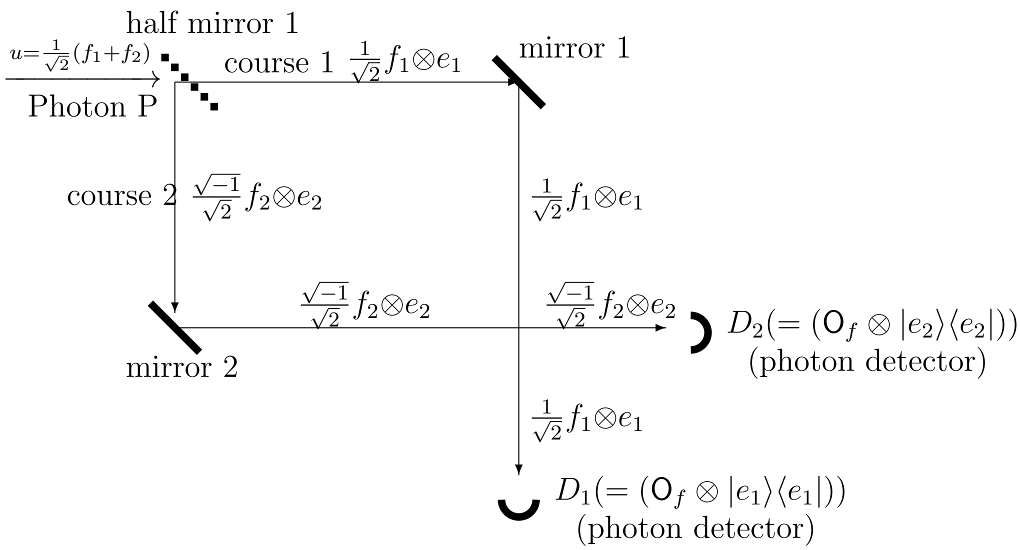


Figure 11.4(1′). $[D_1 + D_2]{=}\mathsf{O}_f \otimes \mathsf{O}_E$

Thus, using the Schrödinger picture, in the above figure we see:

$$u = \frac{1}{\sqrt{2}}(f_1 + f_2) \xrightarrow[\text{time evolution}]{} \frac{1}{\sqrt{2}} f_1{\otimes}e_1 + \frac{\sqrt{-1}}{\sqrt{2}} f_2{\otimes}e_2$$

which may imply that spacetime and quantum entanglement are related.

11.6.4 Mach-Zehnder interferometer (Interference)

Next, consider the following figure:

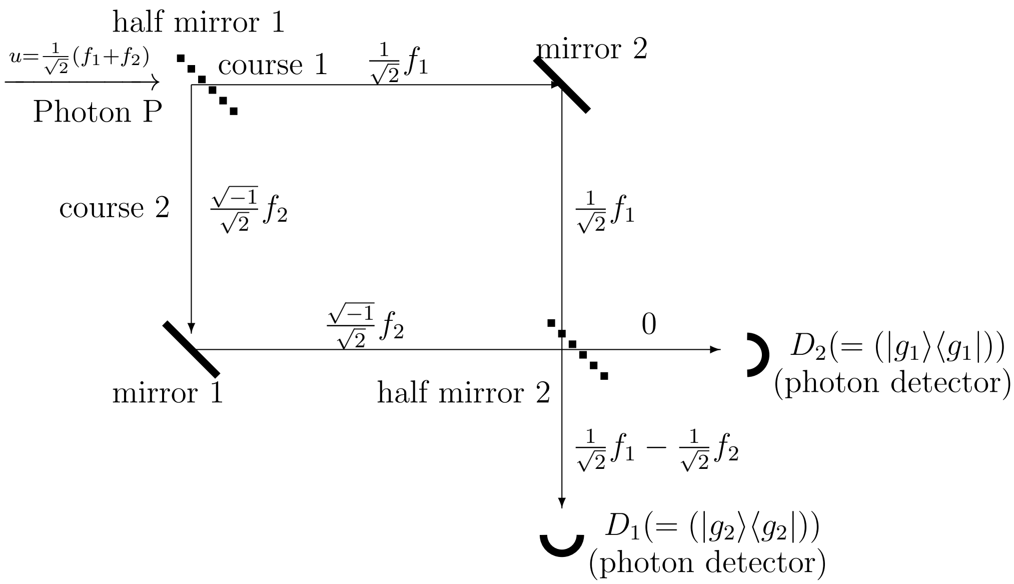


Figure 11.4(2). $[D_1+D_2]$**=Observable**O_g

Now we shall explain, by the Schrödinger picture, Figure 11.4(2) as follows.

The photon P with the state $u = \frac{1}{\sqrt{2}}(f_1 + f_2)$ (precisely, $\rho = |u\rangle\langle u|$) rushed into the half-mirror 1,

(D$_1$) the f_1 part in $u = \frac{1}{\sqrt{2}}(f_1 + f_2)$ passes through the half-mirror 1, and goes along the course 1. And it is reflected in the mirror 1, and passes through the half-mirror 2, and goes to the photon detector D_1.

(D$_2$) the f_2 part in $u = \frac{1}{\sqrt{2}}(f_1 + f_2)$ rebounds on the half-mirror 1 (and strictly saying, the f_2 changes to $\sqrt{-1}f_2$, we are not concerned with it), and goes along the course 2. And it is reflected in the mirror 2, and further reflected in the half-mirror 2, and goes to the photon detector D_2.

This is, by the Heisenberg picture, represented by the following measurement:

$$\mathsf{M}_{B(\mathbb{C}^2)}(\Phi^2\mathsf{O}_g, S_{[\rho]}) \tag{11.29}$$

Then, we see:

(E) the probability that $\begin{bmatrix} \text{a measured value 1} \\ \text{a measured value 2} \end{bmatrix}$ is obtained by $\mathsf{M}_{B(\mathbb{C}^2)}(\Phi^2\mathsf{O}_g, S_{[\rho]})$ is given by

$$\begin{bmatrix} \langle u, \Phi^2 G(\{1\})u\rangle \\ \langle u, \Phi^2 G(\{2\})u\rangle \end{bmatrix} = \begin{bmatrix} |\langle u, UUg_1\rangle|^2 \\ |\langle u, UUg_2\rangle|^2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

11.6.5 Another case

Consider the following Figure 11.4(3).

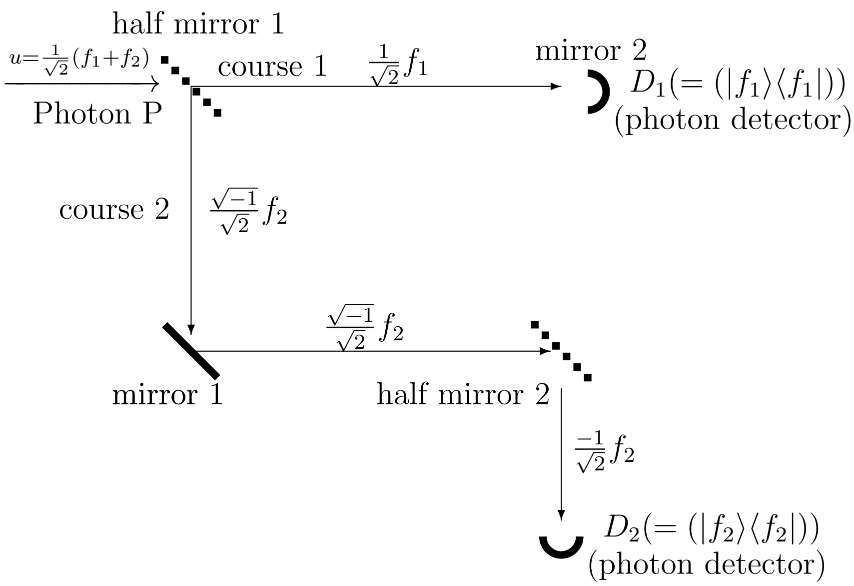


Figure 11.4(3). $[D_2 + D_1]$ =**Observable**O_f

Now we shall explain, by the Schrödinger picture, Figure 11.4(3) as follows.

The photon P with the state $u = \frac{1}{\sqrt{2}}(f_1 + f_2)$ (precisely, $\rho = |u\rangle\langle u|$) rushed into the half-mirror 1,

(F_1) the f_1 part in $u = \frac{1}{\sqrt{2}}(f_1 + f_2)$ passes through the half-mirror 1, and goes along the course 1. And it reaches to the photon detector D_1.

(F_2) the f_2 part in $u = \frac{1}{\sqrt{2}}(f_1 + f_2)$ rebounds on the half-mirror 1 (and strictly saying, the f_2 changes to $\sqrt{-1}f_2$, we are not concerned with it), and goes along the course 2. And it is again reflected in the mirror 1, and further reflected in the half-mirror 2, and goes to the photon detector D_2.

This is, by the Heisenberg picture, represented by the following measurement:

$$\mathsf{M}_{B(\mathbb{C}^2)}(\Phi^2\mathsf{O}_f, S_{[\rho]}) \tag{11.30}$$

Therefore, we see the following:

(G) The probability that $\begin{bmatrix}\text{measured value 1}\\ \text{measured value 2}\end{bmatrix}$ is obtained by the measurement $\mathsf{M}_{B(\mathbb{C}^2)}(\Phi^2\mathsf{O}_f, S_{[\rho]})$ is given by

$$\begin{bmatrix}\mathrm{Tr}(\rho\cdot\Phi^2F(\{1\}))\\ \mathrm{Tr}(\rho\cdot\Phi^2F(\{2\}))\end{bmatrix} = \begin{bmatrix}\langle UUu, F(\{1\})UUu\rangle\\ \langle UUu, F(\{2\})UUu\rangle\end{bmatrix} = \begin{bmatrix}|\langle UUu, f_1\rangle|^2\\ |\langle UUu, f_2\rangle|^2\end{bmatrix} = \begin{bmatrix}\frac{1}{2}\\ \frac{1}{2}\end{bmatrix}$$

Therefore, if the photon detector D_1 does not react, it is expected that the photon detector D_2 reacts.

11.6.6 Conclusion

The above argument is just Wheeler's delayed choice experiment. It should be noted that the difference among Examples in §11.5.3 (Figure 11.4(1))– §11.5 (Figure 11.4(3)) is that of the observables (= measuring instrument). That is,

$$\begin{cases} \text{§11.5.3 (Figure 11.4(1))} \xrightarrow[\text{Heisenberg picture}]{} \Phi \mathsf{O}_f \\ \text{§11.5.4 (Figure 11.4(2))} \xrightarrow[\text{Heisenberg picture}]{} \Phi^2 \mathsf{O}_g \\ \text{§11.5.5 (Figure 11.4(3))} \xrightarrow[\text{Heisenberg picture}]{} \Phi^2 \mathsf{O}_f \end{cases}$$

Hence, it should be noted that

(H) Wheeler's delayed choice experiment —"after the photon P passes through the half-mirror 1, one of Figure 11.4(1), Figure 11.4(2) and Figure 11.4(3) is chosen" — can not be described paradoxically in quantum language.

However, it should be noted that the non-locality paradox (i.e., "there is some thing faster than light") is not solved even in quantum language.

♠**Note 11.3.** What we want to assert in this book may be the following:

(♯) everything (except "there is some thing faster than light") can not be described paradoxically in terms of quantum language

11.7 Hardy's paradox

In this section, we shall introduce the Hardy's paradox (*cf.* ref.[16]) in terms of quantum language[1].

Let H be a two dimensional Hilbert space, i.e., $H = \mathbb{C}^2$. Let $f_1, f_2, g_1, g_2 \in H$ such that

$$f_1 = f_1' = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad f_2 = f_2' = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad g_1 = g_1' = \frac{f_1 + f_2}{\sqrt{2}}, \qquad g_2 = g_2' = \frac{f_1 - f_2}{\sqrt{2}}$$

Put

$$u = \frac{f_1 + f_2}{\sqrt{2}} \Big(= g_1 \Big)$$

Consider the tensor Hilbert space $H \otimes H = \mathbb{C}^2 \otimes \mathbb{C}^2$ and define the state $\widehat{\rho}$ such that

$$\widehat{u} = u \otimes u' = \frac{f_1 + f_2}{\sqrt{2}} \otimes \frac{f_1' + f_2'}{\sqrt{2}}, \qquad \widehat{\rho} = |u \otimes u'\rangle\langle u \otimes u'|$$

As shown in the next section (e.g., annihilation (i.e., $f_1 \otimes f_1 \mapsto 0$), etc.), define the operator $P : \mathbb{C}^2 \otimes \mathbb{C}^2 \to \mathbb{C}^2 \otimes \mathbb{C}^2$ such that

$$P(\alpha_{11} f_1 \otimes f_1 + \alpha_{12} f_1 \otimes f_2 + \alpha_{21} f_2 \otimes f_1 + \alpha_{22} f_2 \otimes f_2) = -\alpha_{12} f_1 \otimes f_2 - \alpha_{21} f_2 \otimes f_1 + \alpha_{22} f_2 \otimes f_2$$

Here, it is clear that

$$P^2(\alpha_{11} f_1 \otimes f_1 + \alpha_{12} f_1 \otimes f_2 + \alpha_{21} f_2 \otimes f_1 + \alpha_{22} f_2 \otimes f_2) = \alpha_{12} f_1 \otimes f_2 + \alpha_{21} f_2 \otimes f_1 + \alpha_{22} f_2 \otimes f_2$$

hence, we see that $P^2 : \mathbb{C}^2 \otimes \mathbb{C}^2 \to \mathbb{C}^2 \otimes \mathbb{C}^2$ is a projection.

Also, define the causal operator $\widehat{\Psi} : B(\mathbb{C}^2 \otimes \mathbb{C}^2) \to B(\mathbb{C}^2 \otimes \mathbb{C}^2)$ by

$$\widehat{\Psi}(\widehat{A}) = P\widehat{A}P \qquad (\widehat{A} \in B(\mathbb{C}^2 \otimes \mathbb{C}^2))$$

Here, it is easy to see that $\widehat{\Psi} : B(\mathbb{C}^2 \otimes \mathbb{C}^2) \to B(\mathbb{C}^2 \otimes \mathbb{C}^2)$ satisfies

(A_1) $\qquad \widehat{\Psi}(\widehat{A}^* \widehat{A}) \geq 0 \quad (\forall \widehat{A} \in B(\mathbb{C}^2 \otimes \mathbb{C}^2))$

(A_2) $\qquad \widehat{\Psi}(I) = P^2$

Since it is not always assured that $\widehat{\Psi}(I) = I$, strictly speaking, the $\widehat{\Psi} : B(\mathbb{C}^2 \otimes \mathbb{C}^2) \to B(\mathbb{C}^2 \otimes \mathbb{C}^2)$ is a causal operator in the wide sense.

---

[1]This section is extracted from

(♯) [44] S. Ishikawa, *The double-slit quantum eraser experiments and Hardy's paradox in the quantum linguistic interpretation,* arxiv:1407.5143[quantum-ph],(2014)

11.7.1 Observable $\mathsf{O}_g \otimes \mathsf{O}_g$

Consider the following figure

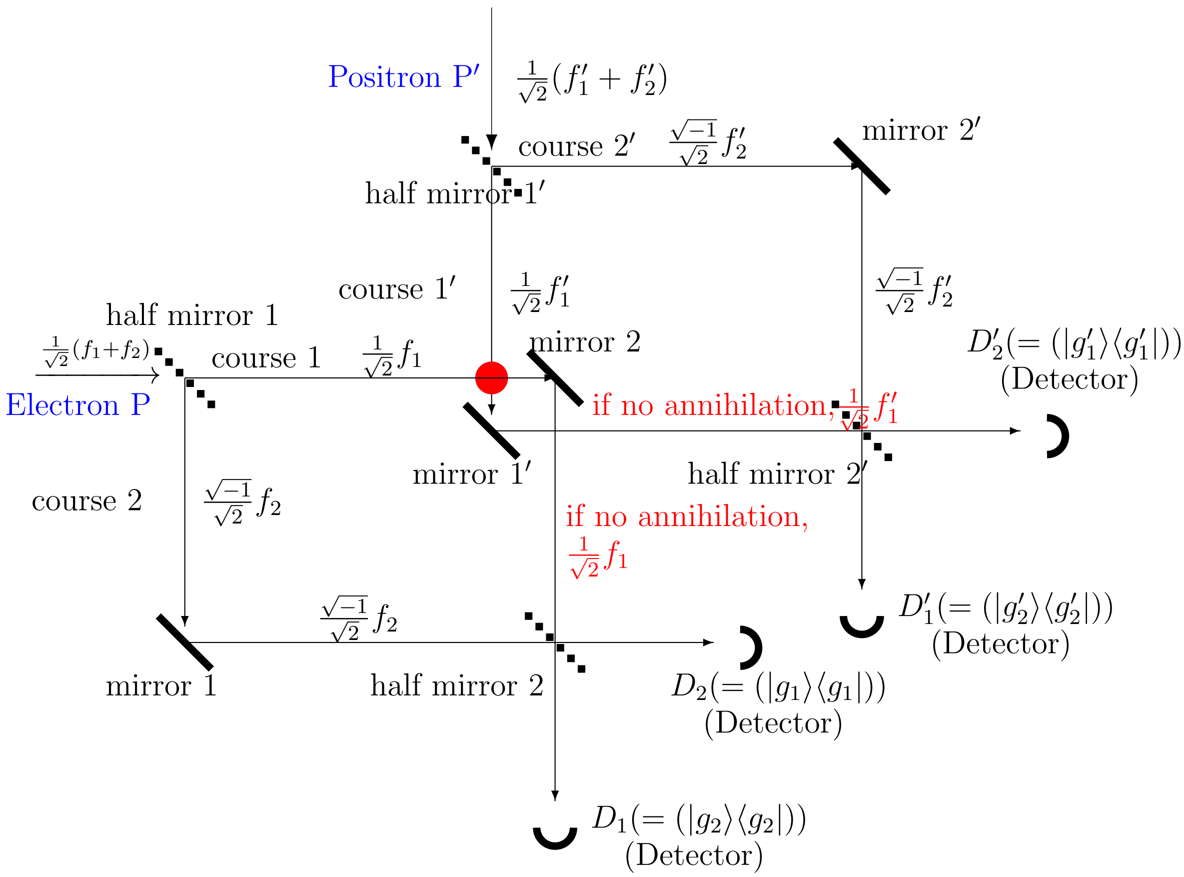


Figure 11.5(1). Electron P and Positron P′ are annihilated at ●

In the above, Electron P and Positron P' rush into the half-mirror 1 and the half-mirror 1′ respectively. Here, "half-mirror" has the following property:

$$\begin{bmatrix}1\\0\end{bmatrix} (= f_1 = f_1') \xrightarrow[\text{pass through half-mirror}]{} \begin{bmatrix}1\\0\end{bmatrix} (= f_1 = f_1')$$

$$\begin{bmatrix}0\\1\end{bmatrix} (= f_2 = f_2') \xrightarrow[\text{be reflected in half-mirror, and } \times\sqrt{-1}]{} \sqrt{-1}\begin{bmatrix}0\\1\end{bmatrix} (= f_2 = f_2')$$

Assume that the initial state of Electron P [resp. Positron P'] is $\beta_1 f_1 + \beta_2 f_2$ [resp. $\beta_1' f_1' + \beta_2' f_2'$]. Then, we see, by the Schrödinger picture, that

$$(\beta_1 f_1 + \beta_2 f_2) \otimes (\beta_1' f_1' + \beta_2' f_2') = \beta_1\beta_1' f_1 \otimes f_1' + \beta_1\beta_2' f_1 \otimes f_2' + \beta_2\beta_1' f_2 \otimes f_1' + \beta_2\beta_2' f_2 \otimes f_2'$$

$$\xrightarrow[\text{(half-mirror)}]{}$$

$$\beta_1\beta_1' f_1 \otimes f_1' + \sqrt{-1}\beta_1\beta_2' f_1 \otimes f_2' + \sqrt{-1}\beta_2\beta_1' f_2 \otimes f_1' - \beta_2\beta_2' f_2 \otimes f_2'$$

$$\xrightarrow[\text{(annihilation(i.e., } f_1 \otimes f_1' = 0))]{}$$

$$\sqrt{-1}\beta_1\beta_2' f_1 \otimes f_2' + \sqrt{-1}\beta_2\beta_1' f_2 \otimes f_1' - \beta_2\beta_2' f_2 \otimes f_2'$$

$$\xrightarrow[\text{(second half-mirror)}]{}$$

$$-\beta_1\beta_2' f_1 \otimes f_2' - \beta_2\beta_1' f_2 \otimes f_1' + \beta_2\beta_2' f_2 \otimes f_2'$$

The above is written by the Schrödinger picture $\widehat{\Psi}_* : \mathcal{T}r(\mathbb{C}^2 \otimes \mathbb{C}^2) \to \mathcal{T}r(\mathbb{C}^2 \otimes \mathbb{C}^2)$. Thus, we have the Heisenberg picture (i.e., the causal operator) $\widehat{\Psi} : B(\mathbb{C}^2 \otimes \mathbb{C}^2) \to B(\mathbb{C}^2 \otimes \mathbb{C}^2)$ by $\widehat{\Psi} = (\widehat{\Psi}_*)^*$.

Define the observable $\widehat{\mathsf{O}}_{gg} = (\{1,2\} \times \{1,2\}, 2^{\{1,2\}\times\{1,2\}}, \widehat{H}_{gg})$ in $B(\mathbb{C}^2 \otimes \mathbb{C}^2)$ by the tensor observable $\mathsf{O}_g \otimes \mathsf{O}_g$, that is,

$$\widehat{H}_{gg}(\{(1,1)\}) = |g_1 \otimes g_1\rangle\langle g_1 \otimes g_1|, \quad \widehat{H}_{gg}(\{(1,2)\}) = |g_1 \otimes g_2\rangle\langle g_1 \otimes g_2|,$$
$$\widehat{H}_{gg}(\{(2,1)\}) = |g_2 \otimes g_1\rangle\langle g_2 \otimes g_1|, \quad \widehat{H}_{gg}(\{(2,2)\}) = |g_2 \otimes g_2\rangle\langle g_2 \otimes g_2|$$

Consider the measurement:

$$\mathsf{M}_{B(\mathbb{C}^2\otimes\mathbb{C}^2)}(\widehat{\Psi}\widehat{\mathsf{O}}_{gg}, S_{[\widehat{\rho}]}) \tag{11.31}$$

Then, the probability that a measured value $(2,2)$ is obtained by $\mathsf{M}_{B(\mathbb{C}^2\otimes\mathbb{C}^2)}(\widehat{\Psi}\widehat{\mathsf{O}}, S_{[\widehat{\rho}]})$ is given by

$$\begin{aligned}
&\langle u \otimes u, P\widehat{H}_{gg}(\{(2,2)\})P(u \otimes u)\rangle \\
=&\frac{|\langle (f_1 - f_2) \otimes (f_1 - f_2), f_1 \otimes f_2 + f_2 \otimes f_1 + f_2 \otimes f_2\rangle|^2}{16} \\
=&\frac{|\langle f_1 \otimes f_1 - f_1 \otimes f_2 - f_2 \otimes f_1 + f_2 \otimes f_2,\ \ f_1 \otimes f_2 + f_2 \otimes f_1 + f_2 \otimes f_2\rangle|^2}{16} = \frac{1}{16}
\end{aligned}$$

Also, the probability that a measured value $(1,1)$ is obtained by $\mathsf{M}_{B(\mathbb{C}^2\otimes\mathbb{C}^2)}(\widehat{\Psi}\widehat{\mathsf{O}}_{gg}, S_{[\widehat{\rho}]})$ is given by

$$\begin{aligned}
&\langle u \otimes u, P\widehat{H}_{gg}(\{(1,1)\})P(u \otimes u)\rangle \\
=&\frac{|\langle (f_1 + f_2) \otimes (f_1 + f_2), f_1 \otimes f_2 + f_2 \otimes f_1 + f_2 \otimes f_2\rangle|^2}{16} \\
=&\frac{|\langle f_1 \otimes f_1 + f_1 \otimes f_2 + f_2 \otimes f_1 + f_2 \otimes f_2,\ \ f_1 \otimes f_2 + f_2 \otimes f_1 + f_2 \otimes f_2\rangle|^2}{16} = \frac{9}{16}
\end{aligned}$$

Further, the probability that a measured value $(1,2)$ is obtained by $\mathsf{M}_{B(\mathbb{C}^2\otimes\mathbb{C}^2)}(\widehat{\Psi}\widehat{\mathsf{O}}_{gg}, S_{[\widehat{\rho}]})$ is given by

$$\langle u \otimes u, P\widehat{H}_{gg}(\{(1,2)\})P(u \otimes u)\rangle$$

$$=\frac{|\langle (f_1+f_2)\otimes(f_1-f_2), f_1\otimes f_2+f_2\otimes f_1+f_2\otimes f_2\rangle|^2}{16}$$
$$=\frac{|\langle f_1\otimes f_1-f_1\otimes f_2+f_2\otimes f_1-f_2\otimes f_2,\ \ f_1\otimes f_2+f_2\otimes f_1+f_2\otimes f_2\rangle|^2}{16}=\frac{1}{16}$$

Similarly,

$$\langle u\otimes u, P\widehat{H}_{gg}(\{(2,1)\})P(u\otimes u)\rangle=\frac{1}{16}$$

Remark 11.19. Note that

$$\frac{1}{16}+\frac{9}{16}+\frac{1}{16}+\frac{1}{16}=\frac{3}{4}<1$$

which is due to the annihilation. Thus, the probability that no measured value is obtained by the measurement $\mathsf{M}_{B(\mathbb{C}^2\otimes\mathbb{C}^2)}(\widehat{\Psi}\widehat{\mathsf{O}}, S_{[\widehat{\rho}]})$ is equal to $\frac{1}{4}$.

11.7.2 The case that there is no half-mirror 2′

Consider the case that there is no half-mirror 2′, the case described in the following figure:

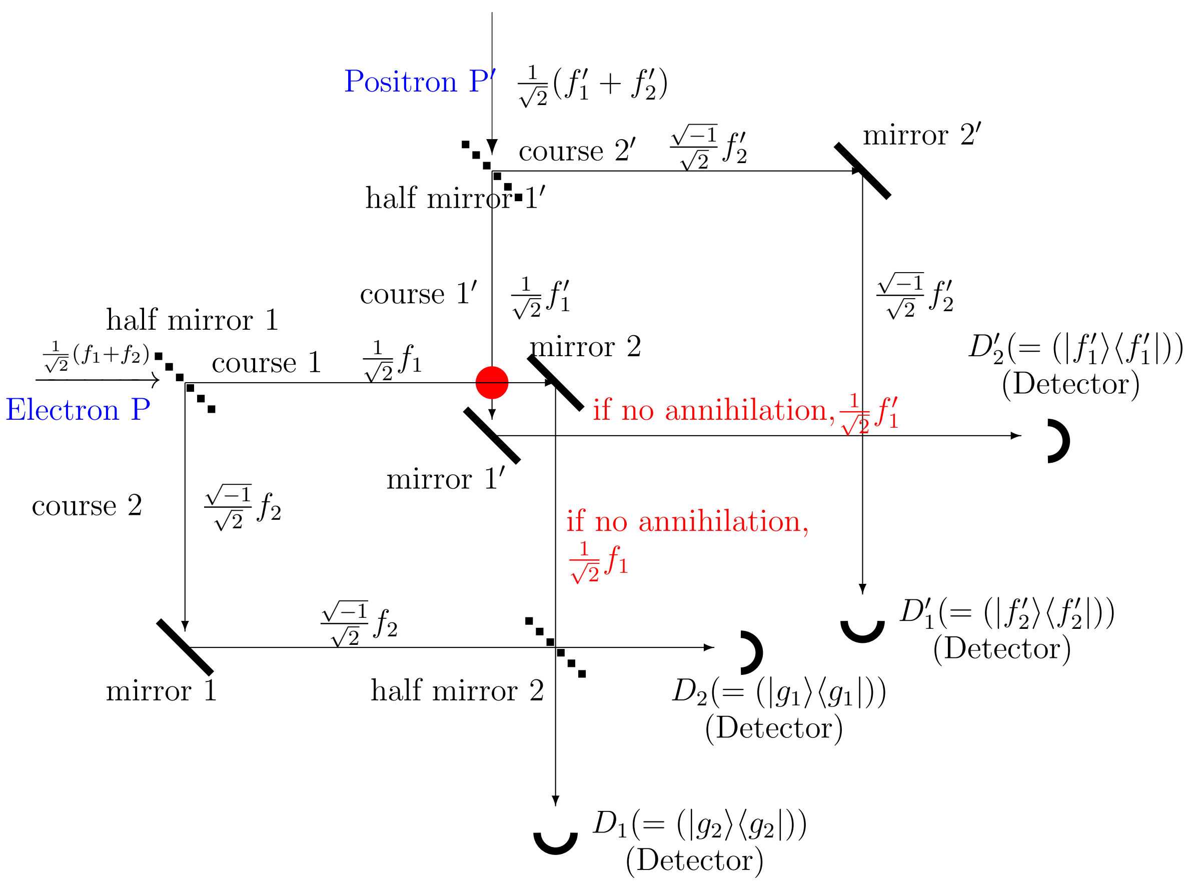


Figure 11.5(2). Electron P and Positron P′ are annihilated at ●

Define the observable $\widehat{\mathsf{O}}_{gf} = (\{1,2\} \times \{1,2\}, 2^{\{1,2\}\times\{1,2\}}, \widehat{H}_{gf})$ in $B(\mathbb{C}^2 \otimes \mathbb{C}^2)$ by the tensor observable $\mathsf{O}_g \otimes \mathsf{O}_f$, that is,

$$\widehat{H}_{gf}(\{(1,1)\}) = |g_1 \otimes f_1\rangle\langle g_1 \otimes f_1|, \quad \widehat{H}_{gf}(\{(1,2)\}) = |g_1 \otimes f_2\rangle\langle g_1 \otimes f_2|,$$
$$\widehat{H}_{gf}(\{(2,1)\}) = |g_2 \otimes f_1\rangle\langle g_2 \otimes f_1|, \quad \widehat{H}_{gf}(\{(2,2)\}) = |g_2 \otimes f_2\rangle\langle g_2 \otimes f_2|$$

Since the causal operator $\widehat{\Psi} : B(\mathbb{C}^2 \otimes \mathbb{C}^2) \to B(\mathbb{C}^2 \otimes \mathbb{C}^2)$ is the same, we get the measurement:

$$\mathsf{M}_{B(\mathbb{C}^2\otimes\mathbb{C}^2)}(\widehat{\Psi}\widehat{\mathsf{O}}_{gf}, S_{[\widehat{\rho}]}) \tag{11.32}$$

Then, the probability that a measured value $(2,2)$ is obtained by $\mathsf{M}_{B(\mathbb{C}^2\otimes\mathbb{C}^2)}(\widehat{\Psi}\widehat{\mathsf{O}}_{gf}, S_{[\widehat{\rho}]})$ is given by

$$\begin{aligned}&\langle u \otimes u, P\widehat{H}_{gf}(\{(2,2)\})P(u \otimes u)\rangle \\ =&\frac{|\langle (f_1 - f_2) \otimes f_2, f_1 \otimes f_2 + f_2 \otimes f_1 + f_2 \otimes f_2\rangle|^2}{8} = 0\end{aligned}$$

Also, the probability that a measured value $(1,1)$ is obtained by $\mathsf{M}_{B(\mathbb{C}^2\otimes\mathbb{C}^2)}(\widehat{\Psi}\widehat{\mathsf{O}}_{gf}, S_{[\widehat{\rho}]})$ is given by

$$\begin{aligned}&\langle u \otimes u, P\widehat{H}_{gf}(\{(1,1)\})P(u \otimes u)\rangle \\ =&\frac{|\langle (f_1 + f_2) \otimes f_1, f_1 \otimes f_2 + f_2 \otimes f_1 + f_2 \otimes f_2\rangle|^2}{8} = \frac{1}{8}\end{aligned}$$

Further, the probability that a measured value $(1,2)$ is obtained by $\mathsf{M}_{B(\mathbb{C}^2\otimes\mathbb{C}^2)}(\widehat{\Psi}\widehat{\mathsf{O}}_{gf}, S_{[\widehat{\rho}]})$ is given by

$$\begin{aligned}&\langle u \otimes u, P\widehat{H}_{gf}(\{(1,2)\})P(u \otimes u)\rangle \\ =&\frac{|\langle (f_1 + f_2) \otimes f_2, f_1 \otimes f_2 + f_2 \otimes f_1 + f_2 \otimes f_2\rangle|^2}{16} = \frac{4}{8}\end{aligned}$$

Similarly,

$$\begin{aligned}&\langle u \otimes u, P\widehat{H}_{gf}(\{(2,1)\})P(u \otimes u)\rangle \\ =&\frac{|\langle (f_1 - f_2) \otimes f_1, f_1 \otimes f_2 + f_2 \otimes f_1 + f_2 \otimes f_2\rangle|^2}{8} = \frac{1}{8}\end{aligned}$$

Remark 11.20. It is usual to consider that "Which way pass problem" is nonsense. It should be noted that, in the Heisenberg picture, the observable (= measuring instrument) does not only include detectors but also mirrors.

11.8 quantum eraser experiment

Let us explain quantum eraser experiment(*cf.* [71]). This section is extracted from

(♯) [44] S. Ishikawa, *The double-slit quantum eraser experiments and Hardy's paradox in the quantum linguistic interpretation,* arxiv:1407.5143[quantum-ph],(2014)

11.8.1 Tensor Hilbert space

Let $\mathbb{C}^2$ be the two dimensional Hilbert space, i,e., $\mathbb{C}^2 = \left\{ \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \mid z_1, z_2 \in \mathbb{C} \right\}$. And put

$$e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Here, define the observable $\mathsf{O}_x = (\{-1,1\}, 2^{\{-1,1\}}, F_x)$ in $B(\mathbb{C}^2)$ such that

$$F_x(\{1\}) = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad F_x(\{-1\}) = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix},$$

Here, note that

$$F_x(\{1\})e_1 = \frac{1}{2}(e_1 + e_2), \quad F_x(\{1\})e_2 = \frac{1}{2}(e_1 + e_2)$$
$$F_x(\{-1\})e_1 = \frac{1}{2}(e_1 - e_2), \quad F_x(\{-1\})e_2 = \frac{1}{2}(-e_1 + e_2)$$

Let H be a Hilbert space such that $L^2(\mathbb{R})$. And let $\mathsf{O} = (X, \mathcal{F}, F)$ be an observable in $B(H)$. For example, consider the position observable, that is, $X = \mathbb{R}$, $\mathcal{F} = \mathcal{B}_{\mathbb{R}}$, and

$$[F(\Xi)](q) = \begin{cases} 1 & (q \in \Xi \in \mathcal{F}) \\ 0 & (q \notin \Xi \in \mathcal{F}) \end{cases}$$

Let u_1 and u_2 ($\in H$) be orthonormal elements, i.e., $\|u_1\|_H = \|u_2\|_H = 1$ and $\langle u_1, u_2 \rangle = 0$. Put

$$u = \alpha_1 u_1 + \alpha_2 u_2$$

where $\alpha_i \in \mathbb{C}$ such that $|\alpha_1|^2 + |\alpha_2|^2 = 1$.

Further, define $\psi \in \mathbb{C}^2 \otimes H$ (the tensor Hilbert space of $\mathbb{C}^2$ and H) such that

$$\psi = \alpha_1 e_1 \otimes u_1 + \alpha_2 e_2 \otimes u_2$$

where $\alpha_i \in \mathbb{C}$ such that $|\alpha_1|^2 + |\alpha_2|^2 = 1$.

11.8.2 Interference

Consider the measurement:

$$\mathsf{M}_{B(\mathbb{C}^2 \otimes H)}(\mathsf{O}_x \otimes \mathsf{O}, S_{[|\psi\rangle\langle\psi|]}) \tag{11.33}$$

Then, we see:

(A$_1$) the probability that a measured value $(1,x)(\in \{-1,1\} \times X)$ belongs to $\{1\} \times \Xi$ is given by

$$\begin{aligned}
&\langle\psi, (F_x(\{1\}) \otimes F(\Xi))\psi\rangle \\
=&\langle \alpha_1 e_1 \otimes u_1 + \alpha_2 e_2 \otimes u_2, (F_x(\{1\} \otimes F(\Xi)))(\alpha_1 e_1 \otimes u_1 + \alpha_2 e_2 \otimes u_2)\rangle \\
=&\frac{1}{2}\langle \alpha_1 e_1 \otimes u_1 + \alpha_2 e_2 \otimes u_2, \alpha_1(e_1 + e_2) \otimes F(\Xi)u_1 + \alpha_2(e_1 + e_2) \otimes F(\Xi)u_2\rangle \\
=&\frac{1}{2}\Big(|\alpha_1|^2\langle u_1, F(\Xi)u_1\rangle + |\alpha_2|^2\langle u_2, F(\Xi)u_2\rangle + \overline{\alpha}_1\alpha_2\langle u_1, F(\Xi)u_2\rangle + \alpha_1\overline{\alpha}_2\langle u_2, F(\Xi)u_1\rangle\Big) \\
=&\frac{1}{2}\Big(|\alpha_1|^2\langle u_1, F(\Xi)u_1\rangle + |\alpha_2|^2\langle u_2, F(\Xi)u_2\rangle + 2[\text{Real part}](\overline{\alpha}_1\alpha_2\langle u_1, F(\Xi)u_2\rangle)\Big)
\end{aligned}$$

where the interference term (i.e., the third term) appears.

Define the probability density function p_1 by

$$\int_\Xi p_1(q)dq = \frac{\langle\psi, (F_x(\{1\}) \otimes F(\Xi))\psi\rangle}{\langle\psi, (F_x(\{1\}) \otimes I)\psi\rangle} \quad (\forall \Xi \in \mathcal{F})$$

Then, by the interference term (i.e., $2[\text{Real part}](\overline{\alpha}_1\alpha_2\langle u_1, F(\Xi)u_2\rangle)$), we get the following graph.

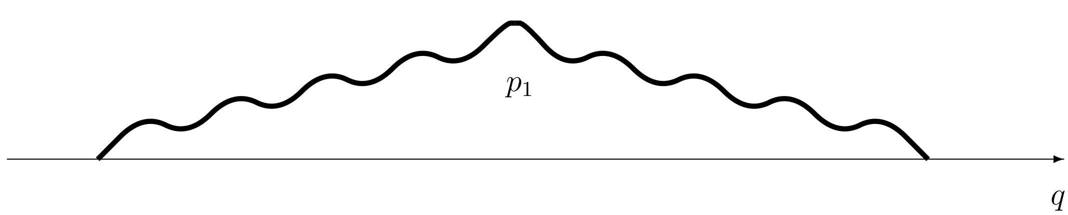


Figure 11.6(1): The graph of p_1

Also, we see:

(A$_2$) the probability that a measured value $(-1,x)(\in \{-1,1\} \times X)$ belongs to $\{-1\} \times \Xi$ is given by

$$\begin{aligned}
&\langle\psi, (F_x(\{-1\}) \otimes F(\Xi))\psi\rangle \\
=&\langle \alpha_1 e_1 \otimes u_1 + \alpha_2 e_2 \otimes u_2, (F_x(\{-1\} \otimes F(\Xi)))(\alpha_1 e_1 \otimes u_1 + \alpha_2 e_2 \otimes u_2)\rangle
\end{aligned}$$

$$
\begin{aligned}
=&\frac{1}{2}\langle \alpha_1 e_1 \otimes u_1 + \alpha_2 e_2 \otimes u_2, \alpha_1 (e_1 - e_2) \otimes F(\Xi) u_1 + \alpha_2(-e_1 + e_2) \otimes F(\Xi) u_2 \rangle \\
=&\frac{1}{2}\Big(|\alpha_1|^2 \langle u_1, F(\Xi) u_1 \rangle + |\alpha_2|^2 \langle u_2, F(\Xi) u_2 \rangle - \overline{\alpha}_1 \alpha_2 \langle u_1, F(\Xi) u_2 \rangle - \alpha_1 \overline{\alpha}_2 \langle u_2, F(\Xi) u_1 \rangle \Big) \\
=&\frac{1}{2}\Big(|\alpha_1|^2 \langle u_1, F(\Xi) u_1 \rangle + |\alpha_2|^2 \langle u_2, F(\Xi) u_2 \rangle - 2[\text{Real part}](\overline{\alpha}_1 \alpha_2 \langle u_1, F(\Xi) u_2 \rangle) \Big)
\end{aligned}
$$

where the interference term (i.e., the third term) appears.

Define the probability density function p_2 by

$$
\int_\Xi p_2(q)dq = \frac{\langle \psi, (F_x(\{-1\}) \otimes F(\Xi))\psi \rangle}{\langle \psi, (F_x(\{-1\}) \otimes I)\psi \rangle} \quad (\forall \Xi \in \mathcal{F})
$$

Then, by the interference term (i.e., $-2[\text{Real part}](\overline{\alpha}_1 \alpha_2 \langle u_1, F(\Xi) u_2 \rangle)$), we get the following graph.

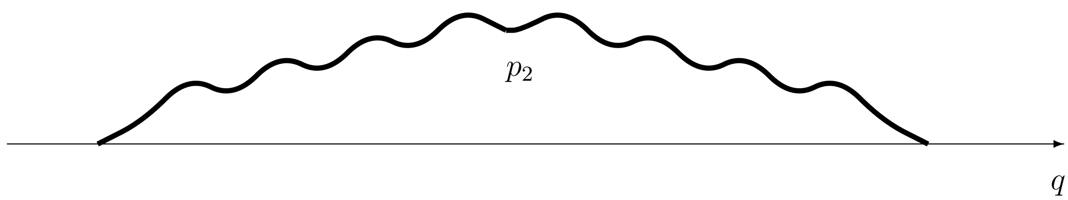


Figure 11.6(2): The graph of p_2

11.8.3 No interference

Consider the measurement:

$$
\mathsf{M}_{B(\mathbb{C}^2 \otimes H)}(\mathsf{O}_x \otimes \mathsf{O}, S_{[|\psi\rangle\langle\psi|]}) \tag{11.34}
$$

Then, we see

(A$_3$) the probability that a measured value $(u, x) (\in \{1, -1\} \times X)$ belongs to $\{1, -1\} \times \Xi$ is given by

$$
\begin{aligned}
&\langle \psi, (I \otimes F(\Xi))\psi \rangle \\
=&\langle \alpha_1 e_1 \otimes u_1 + \alpha_2 e_2 \otimes u_2, (I \otimes F(\Xi))(\alpha_1 e_1 \otimes u_1 + \alpha_2 e_2 \otimes u_2) \rangle \\
=&\langle \alpha_1 e_1 \otimes u_1 + \alpha_2 e_2 \otimes u_2, \alpha_1 e_1 \otimes F(\Xi) u_1 + \alpha_2 e_2 \otimes F(\Xi) u_2 \rangle \\
=&|\alpha_1|^2 \langle u_1, F(\Xi) u_1 \rangle + |\alpha_2|^2 \langle u_2, F(\Xi) u_2 \rangle
\end{aligned}
$$

where the interference term disappears.

Define the probability density function p_3 by

$$
\int_\Xi p_3(q)dq = \langle \psi, (I \otimes F(\Xi))\psi \rangle \quad (\forall \Xi \in \mathcal{F})
$$

Since there is no interference term, we get the following graph.

$$p_3 = p_1 + p_2$$

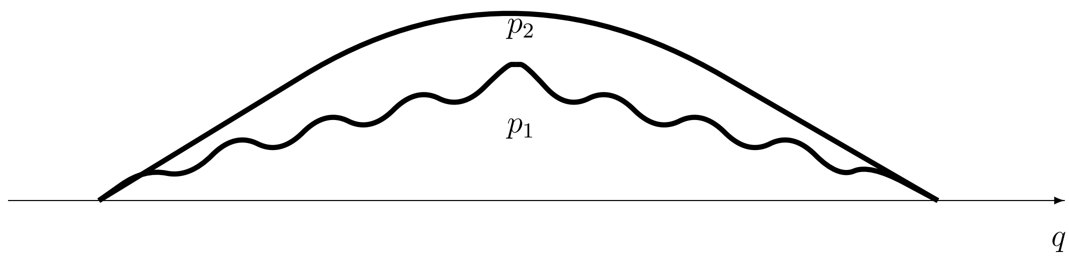


Figure 11.6(3): The graph of $p_3 = p_1 + p_2$

Remark 11.21. Note that

$$\underset{\text{no interference}}{\boxed{(\mathrm{A}_3)}} \quad = \quad \underset{\text{interferences are canceled}}{\boxed{(\mathrm{A}_1)+(\mathrm{A}_2)}}$$

This was experimentally examined in [71].

Chapter 12

Realized causal observable in general theory

Until the previous chapter, we studied all of quantum language, that is,

$$
(\sharp)\quad
\begin{cases}
(\sharp_1): \boxed{\text{pure measurement theory}} \\
\qquad\qquad \text{(=quantum language)} \\
:= \underbrace{\overset{\text{[(pure)Axiom 1]}}{\underset{(cf.\ \S 2.7)}{\boxed{\text{pure measurement}}}} + \overset{\text{[Axiom 2]}}{\underset{(cf.\ \S 10.3)}{\boxed{\text{Causality}}}}}_{\text{a kind of spell(a priori judgment)}} + \underbrace{\overset{\text{[quantum linguistic interpretation]}}{\underset{(cf.\ \S 3.1)}{\boxed{\text{Linguistic interpretation}}}}}_{\text{the manual to use spells}} \\
\\
(\sharp_2): \boxed{\text{mixed measurement theory}} \\
\qquad\qquad \text{(=quantum language)} \\
:= \underbrace{\overset{\text{[(mixed)Axiom}^{(m)}\text{ 1]}}{\underset{(cf.\ \S 9.1)}{\boxed{\text{mixed measurement}}}} + \overset{\text{[Axiom 2]}}{\underset{(cf.\ \S 10.3)}{\boxed{\text{Causality}}}}}_{\text{a kind of spell(a priori judgment)}} + \underbrace{\overset{\text{[quantum linguistic interpretation]}}{\underset{(cf.\ \S 3.1)}{\boxed{\text{Linguistic interpretation}}}}}_{\text{the manual to use spells}}
\end{cases}
$$

As mentioned in the previous chapter, what is important is

- **to exercise the relationship of measurement and causality**

In this chapter, we discuss the relationship more systematically.

12.1 Finite realized causal observable

In dualism (i.e., quantum language), Axiom 2 (Causality) is not used independently, but is always used with Axiom 1 (measurement), just as George Berkeley (A.D. 1685- A.D.1753) said :

(A_1) To be is to be perceived.

♠**Note 12.1.** Note that Berkeley's words is opposite to Einstein's words:

($\sharp_3$) *The moon is there whether one looks at it or not.*

in Einstein and Tagore's conversation.

In this chapter, we devote ourselves to finite realized causal observable. (For the infinite realized causal observable, see Chapter 14.) The readers should understand:

- "realized causal observable" is a direct consequence of the linguistic interpretation, that is,

Only one measurement is permitted.

Now we shall review the following theorem:

Theorem 12.1. [=Theorem 11.1:Causal operator and observable] Consider the basic structure:

$$[\mathcal{A}_k \subseteq \overline{\mathcal{A}}_k \subseteq B(H_k)] \qquad (k=1,2)$$

Let $\Phi_{1,2} : \overline{\mathcal{A}}_2 \to \overline{\mathcal{A}}_1$ be a causal operator, and let $\mathsf{O}_2 = (X, \mathcal{F}, F_2)$ be an observable in $\overline{\mathcal{A}}_2$. Then, $\Phi_{1,2}\mathsf{O}_2 = (X, \mathcal{F}, \Phi_{1,2}F_2)$ is an observable in $\overline{\mathcal{A}}_1$.

Proof. See the proof of Theorem 11.1 $\square$

In this section, we consider the case that the tree ordered set $T(t_0)$ is finite. Thus, putting $T(t_0) = \{t_0, t_1, \ldots, t_N\}$, consider the finite tree $(T(t_0), \leqq)$ with the root t_0, which is represented by $(T{=}\{t_0, t_1, \ldots, t_N\}, \pi : T \setminus \{t_0\} \to T)$ with the the parent map π. .

Definition 12.2. [(finite)sequential causal observable] Consider the basic structure:

$$[\mathcal{A}_k \subseteq \overline{\mathcal{A}}_k \subseteq B(H_k)] \qquad (t \in T(t_0) = \{t_0, t_1, \cdots, t_n\})$$

in which, we have a **sequential causal operator** $\{\Phi_{t_1,t_2} : \overline{\mathcal{A}}_{t_2} \to \overline{\mathcal{A}}_{t_1}\}_{(t_1,t_2)\in T^2_{\leqq}}$ (*cf.* Definition 10.10) such that

(i) for each $(t_1, t_2) \in T^2_{\leqq}$, a causal operator $\Phi_{t_1,t_2} : \overline{\mathcal{A}}_{t_2} \to \overline{\mathcal{A}}_{t_1}$ satisfies that $\Phi_{t_1,t_2}\Phi_{t_2,t_3} = \Phi_{t_1,t_3}$ $(\forall (t_1,t_2), \forall (t_2,t_3) \in T^2_{\leqq})$. Here, $\Phi_{t,t} : \overline{\mathcal{A}}_t \to \overline{\mathcal{A}}_t$ is the identity.

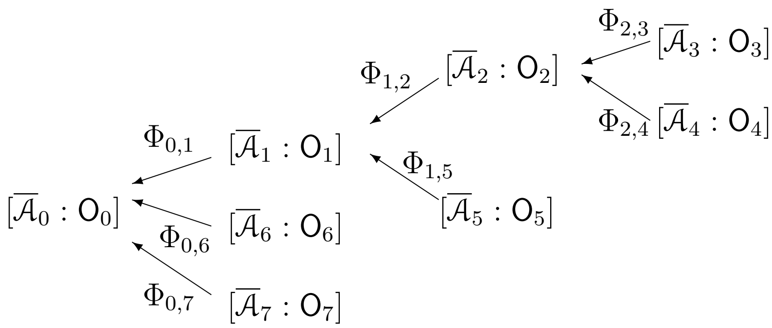


Figure 12.1 : Simple example of sequential causal observable

For each $t \in T$, consider an observable $\mathsf{O}_t{=}(X_t, \mathcal{F}_t, F_t)$ in $\overline{\mathcal{A}}_t$. The pair $[\{\mathsf{O}_t\}_{t\in T}, \{\Phi_{t_1,t_2} : \overline{\mathcal{A}}_{t_2} \to \overline{\mathcal{A}}_{t_1}\}_{(t_1,t_2)\in T^2_{\leqq}}]$ is called a **sequential causal observable**, denoted by $[\mathsf{O}_T]$ or $[\mathsf{O}_{T(t_0)}]$. That is, $[\mathsf{O}_T] = [\{\mathsf{O}_t\}_{t\in T}, \{\Phi_{t_1,t_2} : \overline{\mathcal{A}}_{t_2} \to \overline{\mathcal{A}}_{t_1}\}_{(t_1,t_2)\in T^2_{\leqq}}]$. Using the parent map $\pi : T \setminus \{t_0\} \to T$, $[\mathsf{O}_T]$ is also denoted by $[\mathsf{O}_T] = [\{\mathsf{O}_t\}_{t\in T}, \{\overline{\mathcal{A}}_t \xrightarrow{\Phi_{\pi(t),t}} \overline{\mathcal{A}}_{\pi(t)}\}_{t\in T\setminus\{t_0\}})]$.

Now we can show our present problem.

Problem 12.3. We want to formulate the measurement of a sequential causal observable$[\mathsf{O}_T] = [\{\mathsf{O}_t\}_{t\in T}, \{\Phi_{t_1,t_2} : \overline{\mathcal{A}}_{t_2} \to \overline{\mathcal{A}}_{t_1}\}_{(t_1,t_2)\in T^2_{\leqq}}]$ for a system S with an initial state $\rho_{t_0} (\in \mathfrak{S}^p(\mathcal{A}^*_{t_0}))$.

How do we formulate this measurement?

Now let us solve this problem as follows. Note that the linguistic interpretation says that

only one measurement (and thus, only one observable) is permitted

Thus, we have to combine many observables in a sequential causal observable$[\mathsf{O}_T] = [\{\mathsf{O}_t\}_{t\in T}, \{\Phi_{t_1,t_2} : \overline{\mathcal{A}}_{t_2} \to \overline{\mathcal{A}}_{t_1}\}_{(t_1,t_2)\in T^2_{\leqq}}]$. This is realized as follows.

Definition 12.4. [Realized causal observable]

Let $T(t_0) = \{t_0, t_1, \ldots, t_N\}$ be a finite tree. Let $[\mathsf{O}_{T(t_0)}] = [\{\mathsf{O}_t\}_{t\in T}, \{\Phi_{\pi(t),t} : \overline{\mathcal{A}}_t \xrightarrow{\Phi_{\pi(t),t}} \overline{\mathcal{A}}_{\pi(t)}\}_{t\in T\setminus\{t_0\}}]$ be a sequential causal observable.
For each s $(\in T)$, put $T_s = \{t \in T \mid t \geqq s\}$. Define the observable $\widehat{\mathsf{O}}_s{=}(\times_{t\in T_s} X_t, \ \boxtimes_{t\in T_s}\mathcal{F}_t, \widehat{F}_s)$ in $\overline{\mathcal{A}}_s$ such that

$$\widehat{\mathsf{O}}_s = \begin{cases} \mathsf{O}_s & (\text{ if } s \in T \setminus \pi(T)\ ) \\ \\ \mathsf{O}_s {\times} (\times_{t \in \pi^{-1}(\{s\})} \Phi_{\pi(t),t} \widehat{\mathsf{O}}_t) & (\text{ if } s \in \pi(T)\ ) \end{cases} \tag{12.1}$$

(In quantum case, the existence of $\widehat{\mathsf{O}}_s$ is not always guaranteed). And further, iteratively, we get the observable $\widehat{\mathsf{O}}_{t_0} = (\times_{t \in T} X_t,\ \boxtimes_{t \in T} \mathcal{F}_t,\ \widehat{F}_{t_0})$ in $\overline{\mathcal{A}}_{t_0}$. Put $\widehat{\mathsf{O}}_{t_0} = \widehat{\mathsf{O}}_{T(t_0)}$.

The observable $\widehat{\mathsf{O}}_{T(t_0)} = (\times_{t \in T} X_t,\ \boxtimes_{t \in T} \mathcal{F}_t,\ \widehat{F}_{t_0})$ is called the *(finite) realized causal observable* of the sequential causal observable$[\mathsf{O}_{T(t_0)}] = [\{\mathsf{O}_t\}_{t \in T}, \{\Phi_{\pi(t),t} : \overline{\mathcal{A}}_t \to \overline{\mathcal{A}}_{\pi(t)}\}_{t \in T \setminus \{t_0\}}\ ]$.

Summing up the above arguments, we have the following theorem:
In the classical case, the realized causal observable $\widehat{\mathsf{O}}_{T(t_0)} = (\times_{t \in T} X_t,\ \boxtimes_{t \in T} \mathcal{F}_t,\ \widehat{F}_{t_0})$ always exists.

♠**Note 12.2.** In the above (12.1), the product "$\times$" may be generalized as the quasi-product "$\overset{\text{qp}}{\times}$". However, in this note we are not concerned with such generalization.

Example 12.5. [A simple classical example] Suppose that a tree $(T \equiv \{0, 1, ..., 6, 7\}, \pi)$ has an ordered structure such that $\pi(1) = \pi(6) = \pi(7) = 0$, $\pi(2) = \pi(5) = 1$, $\pi(3) = \pi(4) = 2$.

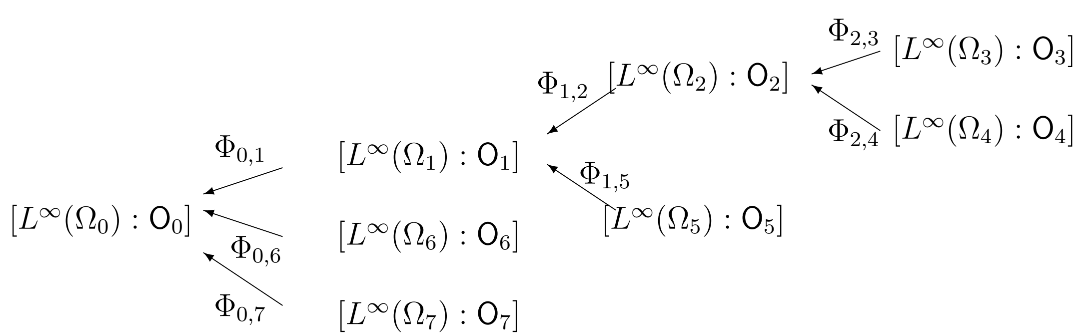


Figure 12.2 : Simple classical example of sequential causal observable

Consider a sequential causal observable $[\mathsf{O}_T] = [\{\mathsf{O}_t\}_{t \in T},\ \{L^\infty(\Omega_t) \overset{\Phi_{\pi(t),t}}{\to} L^\infty(\Omega_{\pi(t)})\}_{t \in T \setminus \{0\}}]$. Now, we shall construct its realized causal observable $\widehat{\mathsf{O}}_{T(t_0)} = (\times_{t \in T} X_t,\ \boxtimes_{t \in T} \mathcal{F}_t,\ \widehat{F}_{t_0})$ in what follows.

Put

$$\widehat{\mathsf{O}}_t = \mathsf{O}_t \quad \text{and thus} \quad \widehat{F}_t = F_t \quad (t = 3, 4, 5, 6, 7).$$

First we construct the product observable $\widehat{\mathsf{O}}_2$ in $L^\infty(\Omega_2)$ such as

$$\widehat{\mathsf{O}}_2 = (X_2 \times X_3 \times X_4, \mathcal{F}_2 \boxtimes \mathcal{F}_3 \boxtimes \mathcal{F}_4, \widehat{F}_2) \quad \text{where } \widehat{F}_2 = F_2 \times (\mathop{\times}_{t=3,4} \Phi_{2,t} \widehat{F}_t),$$

Iteratively, we construct the following:

$$
\begin{array}{ccccc}
L^\infty(\Omega_0) & \xleftarrow{\Phi_{0,1}} & L^\infty(\Omega_1)P & \xleftarrow{\Phi_{1,2}} & L^\infty(\Omega_2) \\
F_0 \times \Phi_{0,6}\widehat{F}_6 \times \Phi_{0,7}\widehat{F}_7 & & F_1 \times \Phi_{1,5}\widehat{F}_5 & & \\
\downarrow & & \downarrow & & \\
\underset{(F_0\times\Phi_{0,6}\widehat{F}_6\times\Phi_{0,7}\widehat{F}_7\times\Phi_{0,1}\widehat{F}_1)}{\widehat{F}_0} & \xleftarrow{\Phi_{0,1}} & \underset{(F_1\times\Phi_{1,5}\widehat{F}_5\times\Phi_{1,2}\widehat{F}_2)}{\widehat{F}_1} & \xleftarrow{\Phi_{1,2}} & \underset{(F_2\times\Phi_{2,3}\widehat{F}_3\times\Phi_{2,4}\widehat{F}_4)}{\widehat{F}_2}
\end{array}.
$$

That is, we get the product observable $\widehat{\mathsf{O}}_1 \equiv (\times_{t=1}^5 X_t,\ \boxtimes_{t=1}^5 \mathcal{F}_t, \widehat{F}_1)$ of O_1, $\Phi_{1,2}\widehat{\mathsf{O}}_2$ and $\Phi_{1,5}\widehat{\mathsf{O}}_5$, and finally, the product observable

$$
\widehat{\mathsf{O}}_0 \equiv (\times_{t=0}^7 X_t,\ \boxtimes_{t=0}^7 \mathcal{F}_t, \widehat{F}_0(= F_0 \times (\underset{t=1,6,7}{\times} \Phi_{0,t}\widehat{F}_t))
$$

of O_0, $\Phi_{0,1}\widehat{\mathsf{O}}_1$, $\Phi_{0,6}\widehat{\mathsf{O}}_6$ and $\Phi_{0,7}\widehat{\mathsf{O}}_7$. Then, we get the realization of a sequential causal observable $[\{\mathsf{O}_t\}_{t\in T}, \{L^\infty(\Omega_t) \overset{\Phi_{\pi(t),t}}{\to} L^\infty(\Omega_{\pi(t)})\}_{t\in T\setminus\{0\}}]$. For completeness, $\widehat{F}_0$ is represented by

$$
\begin{aligned}
&\widehat{F}_0(\Xi_0 \times \Xi_1 \times \Xi_2 \times \Xi_3 \times \Xi_4 \times \Xi_5 \times \Xi_6 \times \Xi_7)] \\
=&F_0(\Xi_0) \times \Phi_{0,1}\Big(F_1(\Xi_1) \times \Phi_{1,5}F_5(\Xi_5) \times \Phi_{1,2}\Big(F_2(\Xi_2) \times \Phi_{2,3}F_3(\Xi_3) \times \Phi_{2,4}F_4(\Xi_4)\Big)\Big) \\
&\qquad \times \Phi_{0,6}(F_6(\Xi_6)) \times \Phi_{0,7}(F_7(\Xi_7))
\end{aligned}
\tag{12.2}
$$

(In quantum case, the existence of $\widehat{\mathsf{O}}_0$ in not guaranteed). □

Remark 12.6. In the above example, consider the case that O_t $(t = 2, 6, 7)$ is not determined. In this case,it suffices to define O_t by the existence observable $\mathsf{O}_t^{(\mathrm{exi})}{=}(X_t, \{\emptyset, X_t\}, F_t^{(\mathrm{exi})})$. Then, we see that

$$
\begin{aligned}
&\widehat{F}_0(\Xi_0 \times \Xi_1 \times X_2 \times \Xi_3 \times \Xi_4 \times \Xi_5 \times X_6 \times X_7) \\
=&F_0(\Xi_0) \times \Phi_{0,1}\Big(F_1(\Xi_1) \times \Phi_{1,5}F_5(\Xi_5) \times \Phi_{1,2}\Big(\Phi_{2,3}F_3(\Xi_3) \times \Phi_{2,4}F_4(\Xi_4)\Big)\Big)
\end{aligned}
\tag{12.3}
$$

This is true. However, the following is not wrong. Putting $T' = \{0, 1, 3, 4, 5\}$, consider the $[\mathsf{O}_{T'}]$ $= [\{\mathsf{O}_t\}_{t\in T'}, \{\Phi_{t_1,t_2} : L^\infty(\Omega_{t_2}) \to L^\infty(\Omega_{t_1})\}_{(t_1,t_2)\in (T')^2_{\leqq}}\ ]$. Then, the realized causal observable $\widehat{\mathsf{O}}_{T'(0)} = (\times_{t\in T'} X_t,\ \boxtimes_{t\in T'}\mathcal{F}_t,\ \widehat{F}'_0)$ is defined by

$$
\begin{aligned}
&\widehat{F}'_0(\Xi_0 \times \Xi_1 \times \Xi_3 \times \Xi_4 \times \Xi_5) = F_0(\Xi_0) \\
&\times \Phi_{0,1}\Big(F_1(\Xi_1) \times \Phi_{1,5}F_5(\Xi_5) \times \Phi_{1,4}F_4(\Xi_4) \times \Phi_{1,3}F_3(\Xi_3) \times \Phi_{1,4}F_4(\Xi_4)\Big)
\end{aligned}
\tag{12.4}
$$

which is different from the true (12.2). We may sometimes omit "existence observable". However, if we do so, we omit it on the basis of careful cautions.

Thus, we can answer Problem 12.3 as follows.

Problem 12.7. [=Problem 12.3] (written again)
We want to formulate the measurement of a sequential causal observable$[\mathsf{O}_T] = [\{\mathsf{O}_t\}_{t \in T}, \{\Phi_{t_1,t_2} : \overline{\mathcal{A}}_{t_2} \to \overline{\mathcal{A}}_{t_1}\}_{(t_1,t_2) \in T^2_{\leqq}}]$ for a system S with an initial state $\rho_{t_0} (\in \mathfrak{S}^p(\mathcal{A}^*_{t_0}))$.

How do we formulate the measurement ?

Answer: If the realized causal observable $\widehat{\mathsf{O}}_{t_0}$ exists, the measurement is formulated by

$$\textit{measurement} \ \mathsf{M}_{\overline{\mathcal{A}}_{t_0}}(\widehat{\mathsf{O}}_{t_0}, S_{[\rho_{t_0}]})$$

Thus, according to Axiom 1 (measurement: §2.7), we see that

(A) The probability that a measured value $(x_t)_{t \in T}$ obtained by the measurement $\mathsf{M}_{\overline{\mathcal{A}}_{t_0}}(\widehat{\mathsf{O}}_T, S_{[\rho_{t_0}]})$ belongs to $\widehat{\Xi} (\in \boxtimes_{t \in T} \mathcal{F}_t)$ is given by

$$_{\mathcal{A}^*_0}\Big(\rho_{t_0}, \widehat{F}_{t_0}(\widehat{\Xi})\Big)_{\overline{\mathcal{A}}_{t_0}} \tag{12.5}$$

The following theorem, which holds in classical systems, is frequently used.

Theorem 12.8. [The realized causal observable of deterministic sequential causal observable in classical systems] Let $(T(t_0), \leqq)$ be a finite tree. For each $t \in T(t_0)$, consider the classical basic structure

$$[C_0(\Omega_t) \subseteq L^\infty(\Omega_t, \nu_t) \subseteq B(L^2(\Omega_t, \nu_t))]$$

Let $[\mathbb{O}_T] = [\{\mathsf{O}_t\}_{t \in T}, \{\Phi_{t_1,t_2} : L^\infty(\Omega_{t_2}) \to L^\infty(\Omega_{t_1})\}_{(t_1,t_2) \in T^2_{\leqq}}]$ be deterministic causal observable. Then, the realization $\widehat{\mathsf{O}}_{t_0} \equiv (\times_{t \in T} X_t, \boxtimes_{t \in T} \mathcal{F}_t, \widehat{F}_{t_0})$ is represented by

$$\widehat{\mathsf{O}}_{t_0} = \mathop{\times}_{t \in T} \Phi_{t_0,t} \mathsf{O}_t$$

That is, it holds that

$$[\widehat{F}_{t_0}(\mathop{\times}_{t \in T} \Xi_t)](\omega_{t_0}) = \mathop{\times}_{t \in T} [\Phi_{t_0,t} F_t(\Xi_t)](\omega_{t_0}) = \mathop{\times}_{t \in T} [F_t(\Xi_t)](\phi_{t_0,t}\omega_{t_0})$$

$$(\forall \omega_{t_0} \in \Omega_{t_0}, \forall \Xi_t \in \mathcal{F}_t)$$

Proof. It suffices to prove the simple classical case of Example 12.5. Using Theorem 10.6 repeatedly, we see that

$$
\begin{aligned}
&\widehat{F}_0 = F_0 \times \big(\mathop{\times}_{t=1,6,7} \Phi_{0,t}\widehat{F}_t \big) \\
=&F_0 \times (\Phi_{0,1}\widehat{F}_1 \times \Phi_{0,6}\widehat{F}_6 \times \Phi_{0,7}\widehat{F}_7) = F_0 \times (\Phi_{0,1}\widehat{F}_1 \times \Phi_{0,6}F_6 \times \Phi_{0,7}F_7) \\
=&\Big(\mathop{\times}_{t=0,6,7} \Phi_{0,t}F_t \Big) \times (\Phi_{0,1}\widehat{F}_1) = \Big(\mathop{\times}_{t=0,6,7} \Phi_{0,t}F_t \Big) \times \Phi_{0,1}(F_1 \times (\mathop{\times}_{t=2,5} \Phi_{1,t}\widehat{F}_t)) \\
=&\Big(\mathop{\times}_{t=0,1,6,7} \Phi_{0,t}F_t \Big) \times \Phi_{0,1}(\mathop{\times}_{t=2,5} \Phi_{1,t}\widehat{F}_t) = \Big(\mathop{\times}_{t=0,1,6,7} \Phi_{0,t}F_t \Big) \times \Phi_{0,1}(\Phi_{1,2}\widehat{F}_2 \times \Phi_{1,5}\widehat{F}_5) \\
=&\Big(\mathop{\times}_{t=0,1,5,6,7} \Phi_{0,t}F_t \Big) \times \Phi_{0,1}(\Phi_{1,2}\widehat{F}_2) = \Big(\mathop{\times}_{t=0,1,5,6,7} \Phi_{0,t}F_t \Big) \times \Phi_{0,1}(\Phi_{1,2}(F_2 \times (\mathop{\times}_{t=3,4} \Phi_{2,t}\widehat{F}_t))) \\
=& \mathop{\times}_{t=0}^{7} \Phi_{0,t}F_t
\end{aligned}
$$

This completes the proof. $\square$

12.2 Double-slit experiment

12.2.1 Interference

For each $t \in T = [0, \infty)$, define the quantum basic structure

$$[\mathcal{C}(H_t) \subseteq B(H_t) \subseteq B(H_t)],$$

where $H_t = L^2(\mathbb{R}^2)$ $(\forall t \in T)$.

Let $u_0 \in H_0 = L^2(\mathbb{R}^2)$ be an initial wave-function such that ($k_0 > 0$, small $\sigma > 0$):

$$u_0(x, y) \approx \psi_x(x, 0)\psi_y(y, 0) = \frac{1}{\sqrt{\pi^{1/2}\sigma}} \exp\Big(ik_0 x - \frac{x^2}{2\sigma^2}\Big) \cdot \frac{1}{\sqrt{\pi^{1/2}\sigma}} \exp\Big(-\frac{y^2}{2\sigma^2}\Big),$$

where the average momentum (p_1^0, p_2^0) is calculated by

$$(p_1^0, p_2^0) = \Big(\int_{\mathbb{R}} \overline{\psi}_x(x, 0) \cdot \frac{\hbar \partial \psi_x(x, 0)}{i \partial x} dx, \int_{\mathbb{R}} \overline{\psi}_y(y, 0) \cdot \frac{\hbar \partial \psi_y(y, 0)}{i \partial y} dy\Big) = (\hbar k_0, 0).$$

That is, we assume that the initial state of the particle P is equal to $|u_0\rangle\langle u_0|$.

Picture 12.9. $\mathsf{M}_{B(H_0)}(\Phi_{0,t_2}\mathsf{O}_2 = (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, \Phi_{0,t_2}F_2), S_{[|u_0\rangle\langle u_0|]})$

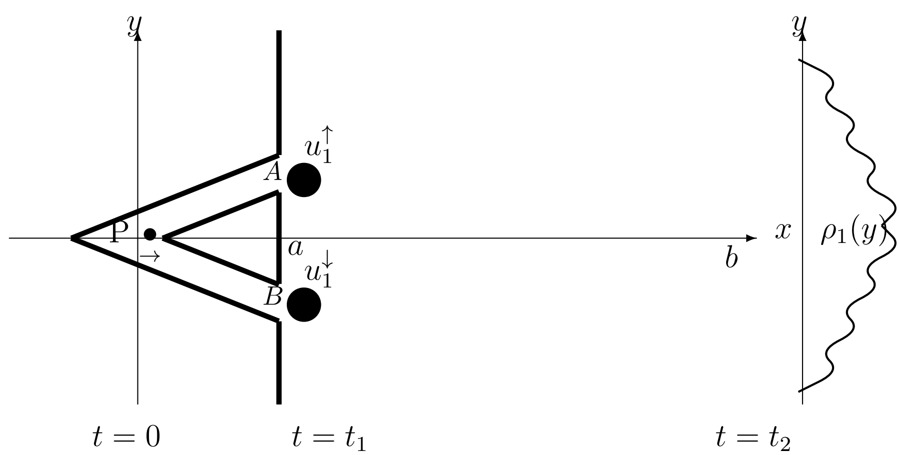


Figure 12.3(1) Potential $V(x, y) = \infty$ on the thick line, $= 0$ (elsewhere)

Thus, we have the following Schrödinger equation:

$$i\hbar \frac{\partial}{\partial t} u_t(x, y) = \mathcal{H} u_t(x, y), \qquad \mathcal{H} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} - \frac{\hbar^2}{2m}\frac{\partial^2}{\partial y^2} + V(x, y)$$

Let s, t be $0 < s < t < \infty$. Thus, we have the causal relation: $\{\Phi_{s,t} : B(H_t) \to B(H_s)\}_{0<s<t<\infty}$ where

$$\Phi_{s,t} A = e^{\frac{\mathcal{H}(t-s)}{i\hbar}} A e^{-\frac{\mathcal{H}(t-s)}{i\hbar}} \qquad (\forall A \in B(H_t) = B(L^2(\mathbb{R}^2)))$$

Thus, $(\Phi_{0,t_1})_*(u_0) = u_1^{\uparrow} + u_1^{\downarrow}$ in Picture 12.9.

Let $\mathsf{O}_2 = (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, F_2)$ be the position observable in $B(L^2(\mathbb{R}^2)$ such that

$$[F(\Xi)](x,y) = \chi_\Xi(y) = \begin{cases} 1 & (x,y) \in \mathbb{R} \times \Xi \\ \\ 0 & (x,y) \in \mathbb{R} \times \mathbb{R} \setminus \Xi \end{cases}$$

Hence, we have the measurement $\mathsf{M}_{B(H_0)}(\Phi_{0,t_2}\mathsf{O}_2 = (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, \Phi_{0,t_2}F_2), S_{[|u_0\rangle\langle u_0|]})$. Axiom 1 (measurement: §2.7) says that

(A) the probability that a measured value $a \in \mathbb{R}$ by $\mathsf{M}_{B(H_0)}(\Phi_{0,t_2}\mathsf{O}, S_{|u_0\rangle\langle u_0|})$ belongs to $(-\infty, y]$ is given by

$$\langle u_0, (\Phi_{0,t_2}F((-\infty,y]))u_0\rangle = \int_{-\infty}^{y} \rho_1(y)dy$$

♠**Note 12.3.** Precisely speaking, we say as follows. Let Δ, ϵ be small positive real numbers. For each $k \in \mathbb{Z} = \{k \mid k = 0, \pm 1, \pm 2, \pm 3, , , , , \}$, define the rectangle D_k such that

$$D_0 = \{(x,y) \in \mathbb{R}^2 \mid x < b\},$$
$$D_k = \{(x,y) \in \mathbb{R}^2 \mid b \le x, (k-1)\Delta < y \le k\Delta\}, \quad k = 1,2,3,...$$
$$D_k = \{(x,y) \in \mathbb{R}^2 \mid b \le x, k\Delta < y \le (k+1)\Delta\}, \quad k = -1,-2,-3,...$$

Thus we have the projection observable $\mathsf{O}_2^{\Delta} = (\mathbb{Z}, 2^{\mathbb{Z}}, F_2^{\Delta})$ in $L^2(\mathbb{R}^2)$ such that

$$[F(\{k\})](x,y) = 1 \ \ ((x,y) \in D_k), \quad = 0 \ \ ((x,y) \in \mathbb{R}^2 \setminus D_k) \qquad (k \in \mathbb{Z})$$

Then it suffices to consider

- for each time $t_n = t_2 + n\epsilon (n = 0,1,2,...)$, the projection observable O_2^{Δ} is measured in the sense of Projection Postulate 11.6.

12.2.2 Which-way path experiment

Picture 12.10. Which-way path experiment: A measured value by $\mathsf{M}_{B(L^2(\mathbb{R}^2))}(\Phi_{0,t_1}(\Psi(\mathsf{O}_G \otimes \Phi_{t_1,t_2}\mathsf{O}_2)), S_{[|u_0\rangle\langle u_0|]})$ belongs to $\{\uparrow\} \times (-\infty, y]$

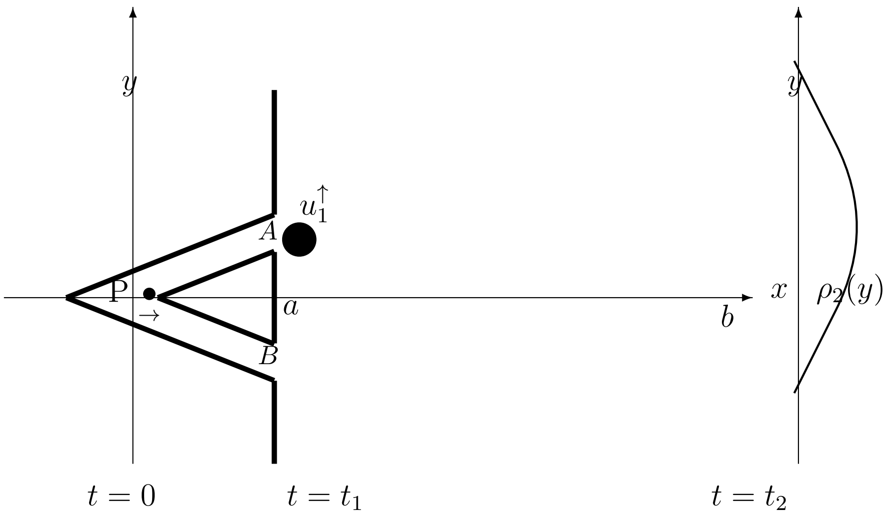


Figure 12.3(2) Potential $V(x,y)=\infty$ on the thick line, $=0$ (elsewhere)

Next, let us explain the above figure. Define the projection observable $\mathsf{O}_1=(\{\uparrow,\downarrow\},2^{\{\uparrow,\downarrow\}},F_1)$ in $B(L^2(\mathbb{R}^2))$ such that

$$
[F_1(\{\uparrow\})](x,y)=\begin{cases} 1 & y\geq 0 \\ 0 & y<0 \end{cases}
$$
$$
[F_1(\{\downarrow\})](x,y)=1-[F_1(\{\uparrow\})](x,y)
$$

According to Section 11.2 (Projection postulate), consider the CONS $\{e_1,e_2\}$ $(\in\mathbb{C}^2)$. Define the predual operator $\Psi_*:Tr(L^2(\mathbb{R}^2))\to Tr(\mathbb{C}^2\otimes L^2(\mathbb{R}^2))$ such that

$$
\Psi_*(|u\rangle\langle u|)=|(e_1\otimes F_1(\{\uparrow\})u)+(e_2\otimes F_1(\{\downarrow\})u)\rangle\langle(e_1\otimes F_1(\{\uparrow\})u)+(e_2\otimes F_1(\{\downarrow\})u)|
$$

Then we have the causal operator $\Psi:B(\mathbb{C}^2\otimes L^2(\mathbb{R}^2))\to L^2(\mathbb{R}^2)$ such that $\Psi=(\Psi_*)^*$. Define the observable $\mathsf{O}_G=(\{\uparrow,\downarrow\},2^{\{\uparrow,\downarrow\}},G)$ in $B(\mathbb{C}^2)$ such that

$$
G(\{\uparrow\})=|e_1\rangle\langle e_1|,\qquad G(\{\downarrow\})=|e_2\rangle\langle e_2|
$$

Hence we have the tensor observable $\mathsf{O}_G\otimes\Phi_{t_1,t_2}\mathsf{O}_2$ in $B(\mathbb{C}^2\otimes L^2(\mathbb{R}^2))$, and hence, the measurement $\mathsf{M}_{B(L^2(\mathbb{R}^2))}(\Phi_{0,t_1}(\Psi(\mathsf{O}_G\otimes\Phi_{t_1,t_2}\mathsf{O}_2)),S_{[|u_0\rangle\langle u_0|]})$. Then, Axiom 1 (measurement: §2.7) says that

(B) the probability that a measured value $(\lambda,y)\in\{\uparrow,\downarrow\}\times\mathbb{R}$ by $\mathsf{M}_{B(L^2(\mathbb{R}^2))}(\Phi_{0,t_1}(\Psi(\mathsf{O}_G\otimes\Phi_{t_1,t_2}\mathsf{O}_2)),S_{[|u_0\rangle\langle u_0|]})$ belongs to $\{\uparrow\}\times(-\infty,y]$ is given by

$$
\langle u_1^{\uparrow},(\Phi_{t_1,t_2}F_2((-\infty,y]))u_l^{\uparrow}\rangle=\frac{1}{2}\int_{-\infty}^{y}\rho_2(y)dy
$$

♠**Note 12.4.** Precisely speaking, in the above case, it suffices to consider the following procedure (1) and (ii):

(i) for time t_1, the projection observable O_1 is measured in the sense of Projection Postulate 11.6

(ii) for each time $t_n = t_2 + n\epsilon (n = 0, 1, 2, ...)$, the projection observable $\mathsf{O}_2^{\triangle}$ is measured in the sense of Projection Postulate 11.6.

12.3 Wilson cloud chamber in double slit experiment

In this section, we shall analyze a discrete trajectory of a quantum particle, which is assumed one of the models of the Wilson cloud chamber (i.e., a particle detector used for detecting ionizing radiation). The main idea is due to. [23, 24, (1991, 1994, S. Ishikawa, *et al.*)].

12.3.1 Trajectory of a particle is non-sense

We shall consider a particle P in the one-dimensional real line $\mathbb{R}$, whose initial state function is $u(x) \in H = L^2(\mathbb{R})$. Since our purpose is to analyze the discrete trajectory of the particle in the double-slit experiment, we choose the state $u(x)$ as follows:

$$u(x) = \begin{cases} l/\sqrt{2}, x \in (-3/2, -1/2) \cup (1/2, 3/2) \\ 0, \text{ otherwise} \end{cases} \tag{12.6}$$

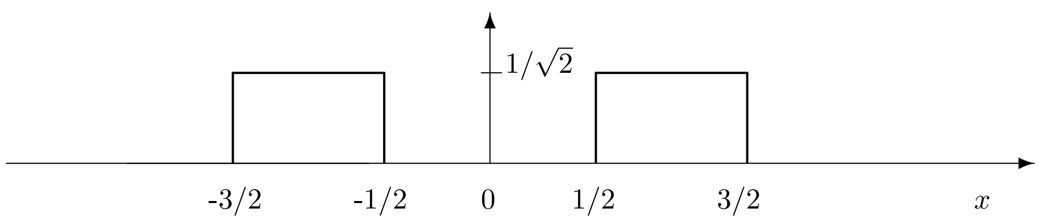


Figure 12.4 The initial wave function $u(x)$

Let A_0 be a position observable in H, that is,

$$(A_0 v)(x) = xv(x) \qquad (\forall x \in \mathbb{R}, \qquad (\text{ for } v \in H = L^2(\mathbb{R})$$

which is identified with the observable $\mathsf{O} = (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, E_{A_0})$ defined by the spectral representation: $A_0 = \int_{\mathbb{R}} x E_{A_0}(dx)$.

We treat the following Heisenberg's kinetic equation of the time evolution of the observable A, $(-\infty < t < \infty)$ in a Hilbert space H with a Hamiltonian $\mathcal{H}$ such that $\mathcal{H} = -(\hbar^2/2m)\partial^2/\partial x^2$ (i.e., the potential $V(x) = 0$), that is,

$$-i\hbar \frac{dA_t}{dt} = \mathcal{H}A_t - A_t\mathcal{H}, \quad -\infty < t < \infty, \text{ where } A_0 = A \tag{12.7}$$

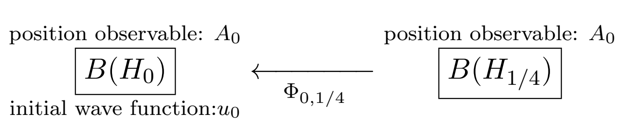
The one-parameter unitary group U_t is defined by $\exp(-itA)$. An easy calculation shows that

$$A_t = U_t^* A U_t = U_t^* x U_t = x + \frac{\hbar t}{im}\frac{d}{dx} \tag{12.8}$$

Put $t = 1/4$, $\hbar/m = 1$. And put

$$A = A_0 (= x), \qquad B = A_{1/4} (= x + \frac{1}{4i}\frac{d}{dx}) = U_{1/4}^* A_0 U_{1/4} = \Phi_{0,1/4} A_0$$

Thus, we have the sequential causal observable



KSTS/RR-16/001
January 8, 2016

However, $A_0(=A)$ and $\Phi_{0,1/4}A_0(=B)$ do not commute, that is, we see:

$$AB - BA = x(x + \frac{1}{4i}\frac{d}{dx}) - (x + \frac{1}{4i}\frac{d}{dx})x = i/4 \neq 0$$

Therefore, **the realized causal observable does not exist**. In this sense,

the trajectory of a particle is non-sense

12.3.2 Approximate measurement of trajectories of a particle

In spite of this fact, we want to consider "trajectories" as follows. That is, we consider the approximate simultaneous measurement of self-adjoint operators $\{A, B\}$ for a particle P with an initial state $u(x)$.

Recall Definition 4.13, that is,

Definition 12.11. (=Definition 4.13). The quartet $(K, s, \widehat{A}, \widehat{B})$ is called an approximately simultaneous observable of A and B, if it satisfied that

(A$_1$) K is a Hilbert space. $s \in K$, $\|s\|_K = 1$, $\widehat{A}$ and $\widehat{B}$ are commutative self-adjoint operators on a tensor Hilbert space $H \otimes K$ that satisfy the average value coincidence condition, that is,

$$\langle u \otimes s, \widehat{A}(u \otimes s)\rangle = \langle u, Au\rangle, \qquad \langle u \otimes s, \widehat{B}(u \otimes s)\rangle = \langle u, Bu\rangle \qquad (12.9)$$
$$(\forall u \in H, \|u\|_H = 1)$$

Also, the measurement $\mathsf{M}_{B(H\otimes K)}(\mathsf{O}_{\widehat{A}} \times \mathsf{O}_{\widehat{B}}, S_{[\widehat{\rho}_{us}]})$ is called the approximately simultaneous measurement of $\mathsf{M}_{B(H)}(\mathsf{O}_A, S_{[\rho_u]})$ and $\mathsf{M}_{B(H)}(\mathsf{O}_B, S_{[\rho_u]})$, where

$$\widehat{\rho}_{us} = |u \otimes s\rangle\langle u \otimes s| \qquad (\|s\}_K = 1)$$

And we define that

(A$_2$) $\Delta^{\widehat{\rho}_{us}}_{\widehat{N}_1}$ $(= \|(\widehat{A} - A \otimes I)(u \otimes s)\|)$ and $\Delta^{\widehat{\rho}_{us}}_{\widehat{N}_2}$ $(= \|(\widehat{B} - B \otimes I)(u \otimes s)\|)$ are called errors of the approximate simultaneous measurement measurement $\mathsf{M}_{B(H\otimes K)}(\mathsf{O}_{\widehat{A}} \times \mathsf{O}_{\widehat{B}}, S_{[\widehat{\rho}_{us}]})$

Now, let us constitute the approximately observable $(K, s, \widehat{A}, \widehat{B})$ as follows.

Put

$$K = L^2(\mathbb{R}_y), \qquad s(y) == \Big(\frac{\omega_1}{\pi}\Big)^{1/4} \exp\Big(-\frac{\omega_1|y|^2}{2}\Big)$$

where ω_1 is assumed to be $\omega_1 = 4,\ 16,\ 64$ later. It is easy to show that $\|s\|_{L^2(\mathbb{R}_y)} = 1$ (i.e., $\|s\|_K = 1$) and

$$\langle s, As\rangle = \langle s, Bs\rangle = 0 \qquad (12.10)$$

And further, put

$$\widehat{A} = A \otimes I + 2I \otimes A$$
$$\widehat{B} = B \otimes I - \frac{1}{2}I \otimes B$$

Note that the two commute (i.e., $\widehat{A}\widehat{B} = \widehat{B}\widehat{A}$). Also, we see, by (12.10),

$$\langle u \otimes s, \widehat{A}(u \otimes s) \rangle = \langle u \otimes s, (A \otimes I + 2I \otimes A)(u \otimes s) \rangle = \langle u, Au \rangle \tag{12.11}$$

$$\langle u \otimes s, \widehat{A}(u \otimes s) \rangle = \langle u \otimes s, (B \otimes I - 2I \otimes A)(u \otimes s) \rangle = \langle u, Bu \rangle \tag{12.12}$$

$$(\forall u \in H, i = 1, 2)$$

Thus, we have the approximately simultaneous measurement $\mathsf{M}_{B(H \otimes K)}(\mathsf{O}_{\widehat{A}} \times \mathsf{O}_{\widehat{B}}, S_{[\widehat{\rho}_{us}]})$, and the errors are calculated as follows:

$$\delta_0 = \Delta_{\widehat{N}_1}^{\widehat{\rho}_{us}} = \|(\widehat{A} - A \otimes I)(u \otimes s)\| = \|2(I \otimes A)(u \otimes s)\| = 2\|As\| \tag{12.13}$$

$$\delta_{1/4} = \Delta_{\widehat{N}_2}^{\widehat{\rho}_{us}} = \|(\widehat{B} - B \otimes I)(u \otimes s)\| = (1/2)\|(I \otimes B)(u \otimes s)\| = (1/2)\|Bs\| \tag{12.14}$$

By the parallel measurement $\bigotimes_{k=1}^{N} \mathsf{M}_{B(H\otimes K)}(\mathsf{O}_{\widehat{A}} \times \mathsf{O}_{\widehat{B}}, S_{[\widehat{\rho}_{us}]})$, assume that a measured value:

$$\Big((x_1, x_1'), (x_2, x_2'), \cdots, (x_N, x_N')\Big)$$

is obtained. This is numerically calculated as follows.

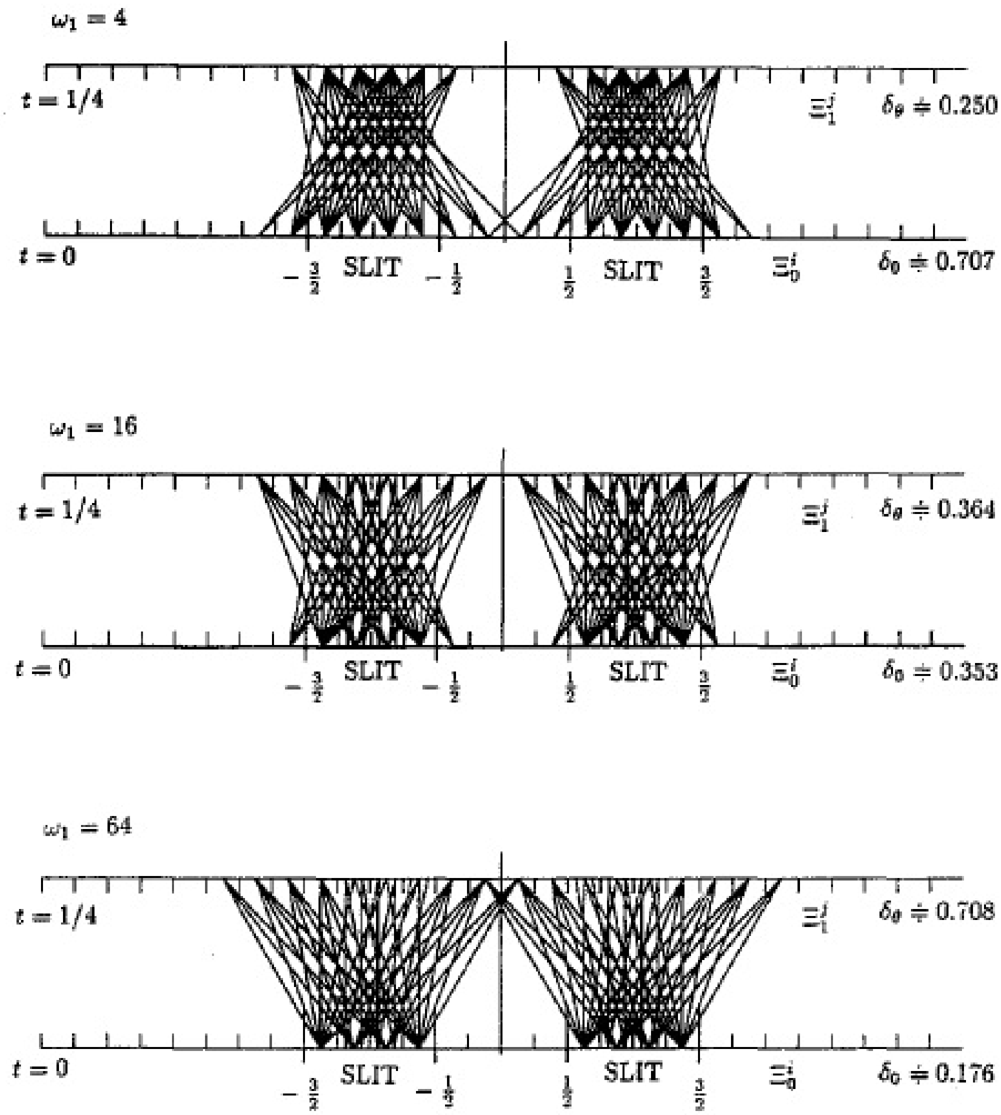


Figure 12.5: The lines connecting two points (i.e., x_k and x_k') ($k = 1, 2, ...$)

Here, note that $\delta_\theta (= \delta_{1/4})$ and δ_0 are depend on ω_1.

♠**Note 12.5.** For the further arguments, see the following refs.

($\sharp_1$) [23]: S. Ishikawa, *Uncertainties and an interpretation of nonrelativistic quantum theory,* International Journal of Theoretical Physics 30, 401–417 (1991)
doi: 10.1007/BF00670793

($\sharp_2$) [24]: Ishikawa, S., Arai, T. and Kawai, T. *Numerical Analysis of Trajectories of a Quantum Particle in Two-slit Experiment,* International Journal of Theoretical Physics, Vol. 33, No. 6, 1265-1274, 1994
doi: 10.1007/BF00670793

12.4 Two kinds of absurdness — idealism and dualism

This section is extracted from ref. [38].

Measurement theory (= quantum language) has two kinds of absurdness. That is,

$$
(\sharp)\quad \text{Two kinds of absurdness}\begin{cases} \text{idealism}\cdots\text{linguistic world-view} \\ \quad \text{The limits of my language mean the limits of my world} \\ \\ \text{dualism}\cdots\text{Descartes=Kant philosophy} \\ \quad \text{The dualistic description for monistic phenomenon} \end{cases}
$$

In what follows, we explain these.

12.4.1 The linguistic interpretation — A spectator does not go up to the stage

Problem 12.12. [A spectator does not go up to the stage]

Consider the elementary problem with two steps (a) and (b):

(a) Consider an urn, in which 3 white balls and 2 black balls are. And consider the following trial:

- Pick out one ball from the urn. If it is black, you return it in the urn If it is white, you do not return it and have it. Assume that you take three trials.

.

(b) Then, calculate the probability that you have 2 white ball after (a)(i.e., three trials).

Answer Put $\mathbb{N}_0 = \{0, 1, 2, \ldots\}$ with the counting measure. Assume that there are m white balls and n black balls in the urn. This situation is represented by a state $(m, n) \in \mathbb{N}_0^2$. We can define the dual causal operator $\Phi^* : \mathcal{M}_{+1}(\mathbb{N}_0^2) \to \mathcal{M}_{+1}(\mathbb{N}_0^2)$ such that

$$
\Phi^*(\delta_{(m,n)}) = \begin{cases} \frac{m}{m+n}\delta_{(m-1,n)} + \frac{n}{m+n}\delta_{(m,n)} & (\text{when } \ m \neq 0\ ) \\ \delta_{(0,n)} & (\text{when } m = 0\ ). \end{cases} \tag{12.15}
$$

where $\delta_{(\cdot)}$ is the point measure.

Let $T = \{0, 1, 2, 3\}$ be discrete time. For each $t \in T$, put $\Omega_t = \mathbb{N}_0^2$. Thus, we see:

$$
\begin{aligned}
&[\Phi^*]^3(\delta_{(3,2)}) = [\Phi^*]^2\left(\frac{3}{5}\delta_{(2,2)} + \frac{2}{5}\delta_{(3,2)}\right) \\
=&\Phi^*\left((\frac{3}{5}(\frac{2}{4}\delta_{(1,2)} + \frac{2}{4}\delta_{(2,2)}) + \frac{2}{5}(\frac{3}{5}\delta_{(2,2)} + \frac{2}{5}\delta_{(3,2)})\right) \\
=&\Phi^*\left(\frac{3}{10}\delta_{(1,2)} + \frac{27}{50}\delta_{(2,2)} + \frac{4}{25}\delta_{(3,2)}\right) \\
=&\frac{3}{10}(\frac{1}{3}\delta_{(0,2)} + \frac{2}{3}\delta_{(1,2)}) + \frac{27}{50}(\frac{2}{4}\delta_{(1,2)} + \frac{2}{4}\delta_{(2,2)}) + \frac{4}{25}(\frac{3}{5}\delta_{(2,2)} + \frac{2}{5}\delta_{(3,2)}) \\
=&\frac{1}{10}\delta_{(0,2)} + \frac{47}{100}\delta_{(1,2)} + \frac{183}{500}\delta_{(2,2)} + \frac{8}{125}\delta_{(3,2)}
\end{aligned} \tag{12.16}
$$

Define the observable $\mathsf{O} = (\mathbb{N}_0, 2^{\mathbb{N}_0}, F)$ in $L^\infty(\Omega_3)$ such that

$$
[F(\Xi)](m, n) = \begin{cases} 1 & (m, n) \in \Xi \times \mathbb{N}_0 \subseteq \Omega_3 \\ 0 & (m, n) \notin \Xi \times \mathbb{N}_0 \subseteq \Omega_3 \end{cases}
$$

Therefore, the probability that a measured value "2" is obtained by the measurement $\mathsf{M}_{L^\infty(\mathbb{N}_0^2)}(\Phi^3\mathsf{O}, S_{[(3,2)]})$ is given by

$$[\Phi^3(F(\{2\}))](3,2) = \int_{\Omega_3} [F(\{2\})](\omega)([\Phi^*]^3(\delta_{(3,2)}))(d\omega) = \frac{183}{500} \tag{12.17}$$

□

The above may be easy, but we should note that

(c) the part (a) is related to causality, and the part (b) is related to measurement.

Thus, the observer is not in the (a). Figuratively speaking, we say:

A spectator does not go up to the stage

Thus, someone in the (a) should be regard as "robot".

♠**Note 12.6.** The part (a) is not related to "probability". That is because The spirit of measurement theory says that

there is no probability without measurements.

although something like "probability" in the (a) is called "Markov probability".

12.4.2 In the beginning was the words—Fit feet to shoes

Remark 12.13. [The confusion between measurement and causality (Continued from Example2.31)] Recall Example2.31 [The measurement of "cold or hot" for water]. Consider the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{ch}, S_{[\omega]})$ where $\omega = 5(\,^\circ\mathrm{C})$. Then we say that

(a) By the **measurement** $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{ch}, S_{[\omega(=5)]})$, the probability that a **measured value**

$$x(\in X = \{c,h\}) \text{ belongs to a set } \begin{bmatrix} \emptyset(=\text{empty set}) \\ \{\mathrm{c}\} \\ \{\mathrm{h}\} \\ \{\mathrm{c},\mathrm{h}\} \end{bmatrix} \text{ is equal to } \begin{bmatrix} 0 \\ [F(\{c\})](5)=1 \\ [F(\{h\})](5)=0 \\ 1 \end{bmatrix}$$

Here, we should not think:

"5 °C" is the cause and "cold" is a result.

That is, we **never** consider that

(b) $$\underset{\text{(cause)}}{\boxed{5\ \ ^\circ\mathrm{C}}} \longrightarrow \underset{\text{(result)}}{\boxed{\text{cold}}}$$

That is because Axiom 2 (causality; §10.3) is not used in (a), though the (a) may be sometimes regarded as the causality (b) in ordinary language.

♠**Note 12.7.** However, from the different point of view, the above (b) can be justified as follows. Define the dual causal operator $\Phi^* : \mathcal{M}([0,100]) \to \mathcal{M}(\{c,h\})$ by

$$[\Phi^* \delta_\omega](D) = f_c(\omega) \cdot \delta_{\rm C}(D) + f_h(\omega) \cdot \delta_{\rm H}(D) \qquad (\forall \omega \in [0,100], \ \ \forall D \subseteq \{c,h\})$$

Then, the (b) can be regarded as "causality". That is,

($\sharp$) **"measurement or causality" depends on how to describe a phenomenon.**

This is the linguistic world-description method.

Remark 12.14. [Mixed measurement and causality] Reconsider Problem 9.5(urn problem:mixed measurement). That is, consider a state space $\Omega = \{\omega_1, \omega_2\}$, and define the observable $\mathsf{O} = (\{w,b\}, 2^{\{w,b\}}, F)$ in $L^\infty(\Omega)$ in Problem 9.5. Define the mixed state by $\rho^m = p\delta_{\omega_1} + (1-p)\delta_{\omega_2}$. Then the probability that a measured value x ($\in \{w,b\}$) is obtained by the mixed measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[*]}(\rho^m))$ is, by (9.3), given by

$$\begin{aligned} P(\{x\}) &= \int_\Omega [F(\{x\})](\omega)\rho^m(d\omega) = p[F(\{x\})](\omega_1) + (1-p)[F(\{x\})](\omega_2) \\ &= \begin{cases} 0.8p + 0.4(1-p) & \text{(when } x = w \text{)} \\ 0.2p + 0.6(1-p)) & \text{(when } x = b \text{)} \end{cases} \end{aligned} \tag{12.18}$$

Now, define a new state space Ω_0 by $\Omega_0 = \{\omega_0\}$. And define the dual (non-deterministic) causal operator $\Phi^* : \mathcal{M}_{+1}(\Omega_0) \to \mathcal{M}_{+1}(\Omega)$ by $\Phi^*(\delta_{\omega_0}) = p\delta_{\omega_1} + (1-p)\delta_{\omega_2}$. Thus, we have the (non-deterministic) causal operator $\Phi : L^\infty(\Omega) \to L^\infty(\Omega_0)$. Here, consider a pure measurement $\mathsf{M}_{L^\infty(\Omega_0)}(\Phi\mathsf{O}, S_{[\omega_0]})$. Then, the probability that a measured value x ($\in \{w,b\}$) is obtained by the measurement is given by

$$\begin{aligned} P(\{x\}) &= [\Phi(F(\{x\}))](\omega_0) = \int_\Omega [F(\{x\})](\omega)\rho^m(d\omega) \\ &= \begin{cases} 0.8p + 0.4(1-p) & \text{(when } x = w \text{)} \\ 0.2p + 0.6(1-p)) & \text{(when } x = b \text{)} \end{cases} \end{aligned}$$

which is equal to the (12.18). Therefore, the mixed measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[*]}(\nu_0))$ can be regarded as the pure measurement $\mathsf{M}_{L^\infty(\Omega_0)}(\Phi\mathsf{O}, S_{[\omega_0]})$.

♠**Note 12.8.** In the above arguments, we see that

($\sharp$) **Concept depends on the description**

This is the linguistic world-description method. As mentioned frequently, we are not concerned with the question "what is ◯◯?". The reason is due to this ($\sharp$). "Measurement or Causality" depends on the description. Some may recall Nietzsche's famous saying:

There are no facts, only interpretations.

This is just the linguistic world-description method with the spirit: "Fit feet (=world) to shoes (language)".

♠**Note 12.9.** In the book "The astonishing hypothesis" ([10] by F. Click (the most noted for being a co-discoverer of the structure of the DNA molecule in 1953 with James Watson)), Dr. Click said that

(a) *You, your joys and your sorrows, your memories and your ambitions,your sense of personal identity and free will,are in fact no more than the behavior of a vast assembly of nerve cells and their associated molecules.*

It should be note that this (a) and the dualism do not contradict. That is because quantum language says:

(b) **Describe any monistic phenomenon by the dualistic language (= quantum language)!**

Also, if the above (a) is due to David Hume, he was a scientist rather than a philosopher.

Chapter 13

Fisher statistics (II)

Measurement theory (= quantum language) is formulated as follows.

$$\bullet \quad \underset{\text{(=quantum language)}}{\boxed{\text{measurement theory}}} := \underbrace{\overset{\text{[Axiom 1]}}{\underset{(cf.\ \S 2.7)}{\boxed{\text{Measurement}}}} + \overset{\text{[Axiom 2]}}{\underset{(cf.\ \S 10.3)}{\boxed{\text{Causality}}}}}_{\text{a kind of spell(a priori judgment)}} + \underbrace{\overset{\text{[quantum linguistic interpretation]}}{\underset{(cf.\ \S 3.1)}{\boxed{\text{Linguistic interpretation}}}}}_{\text{manual to use spells}}$$

In Chapter 5 (Fisher statistics (I)), we discuss "inference" in the relation of "measurement". In this chapter, we discuss "inference" in the relation of "measurement" and "causality". Thus, we devote ourselves to regression analysis. This chapter is extracted from the following:

($\sharp$) Ref. [29]: S. Ishikawa, "Mathematical Foundations of Measurement Theory," Keio University Press Inc. 2006.

13.1 "Inference" = "Control"

It is usually considered that

$$\begin{cases} \bullet \text{ statistics is closely related to inference} \\ \bullet \text{ dynamical system theory is closely related to control} \end{cases}$$

However, in this chapter, we show that

"inference" = "control"

In this sense, we conclude that statistics and dynamical system theory are essentially the same.

13.1.1 Inference problem(statistics)

Problem 13.1. [Inference problem and regression analysis]

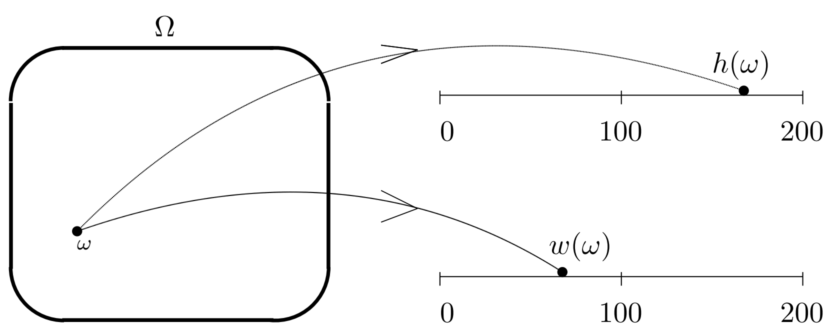
Let $\Omega \equiv \{\omega_1, \omega_2, ..., \omega_{100}\}$ be a set of all students of a certain high school. Define $h : \Omega \to [0, 200]$ and $w : \Omega \to [0, 200]$ such that:

$$
\begin{aligned}
h(\omega_n) = & \text{ "the height of a student } \omega_n\text{"} \quad (n = 1, 2, ..., 100) \\
w(\omega_n) = & \text{ "the weight of a student } \omega_n\text{"} \quad (n = 1, 2, ..., 100)
\end{aligned} \tag{13.1}
$$

For simplicity, put, $N = 5$. For example, see Table 13.1.

Table 13.1: Height and weight

Height· Weight ╲ Student	ω_1	ω_2	ω_3	ω_4	ω_5
Height ($h(\omega)$ cm)	150	160	165	170	175
Weight($w(\omega)$ kg)	65	55	75	60	65



Assume that:

(a$_1$) The principal of this high school knows the both functions h and w. That is, he knows the exact data of the height and weight concerning all students.

Also, assume that:

(a$_2$) Some day, a certain student helped a drowned girl. But, he left without reporting the name. Thus, all information that the principal knows is as follows:

(i) he is a student of his high school.

(ii) his height [resp. weight] is about 170 cm [resp. about 80 kg].

Now we have the following question:

(b) Under the above assumption (a$_1$) and (a$_2$), how does the principal infer who is he?

This will be answered in Answer 13.5.

13.1.2 Control problem(dynamical system theory)

Adding the measurement equation $g : \mathbb{R}^3 \to \mathbb{R}$ to the state equation, we have dynamical system theory(13.2). That is,

$$\boxed{\text{dynamical system theory}} = \begin{cases} \text{(i)} : \frac{d\omega(t)}{dt} = v(\omega(t), t, e_1(t), \beta) & \cdots (\text{ state equation}) \\ \qquad\qquad (\text{initial}_{\omega(0)=\alpha}) & \\ \\ \text{(ii)} : x(t) = g(\omega(t), t, e_2(t)) & \cdots (\text{ measurement}) \end{cases} \tag{13.2}$$

where α, β are parameters, $e_1(t)$ is noise, $e_2(t)$ is measurement error.

The following example is the simplest problem concerning inference.

Problem 13.2. [Control problem and regression analysis] We have a rectangular water tank filled with water.

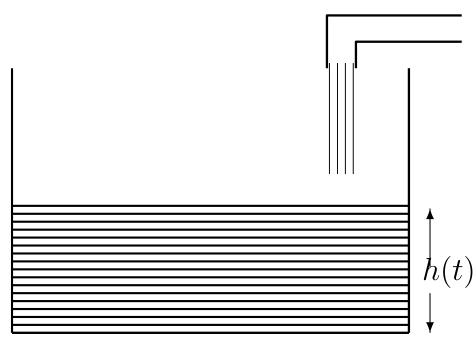


Figure 13.1: Water tank

Assume that the height of water at time t is given by the following function $h(t)$:

$$\frac{dh}{dt} = \beta_0, \text{ then } h(t) = \alpha_0 + \beta_0 t, \tag{13.3}$$

where α_0 and β_0 are unknown fixed parameters such that α_0 is the height of water filling the tank at the beginning and β_0 is the increasing height of water per unit time. The measured height $h_m(t)$ of water at time t is assumed to be represented by

$$h_m(t) = \alpha_0 + \beta_0 t + e(t),$$

where $e(t)$ represents a noise (or more precisely, a measurement error) with some suitable conditions. And assume that we obtained the measured data of the heights of water at $t = 1, 2, 3$ as follows:

$$h_m(1) = 1.9, \quad h_m(2) = 3.0, \quad h_m(3) = 4.7. \tag{13.4}$$

Under this setting, we consider the following problem:

(c_1) **[Control]:** Settle the state (α_0, β_0) such that measured data (13.4) will be obtained.

or, equivalently,

(c_2) **[Inference]:** when measured data (13.4) is obtained, infer the unknown state (α_0, β_0).

This will be answered in Answer 13.6.

Note that

$$(\mathrm{c}_1)=(\mathrm{c}_2)$$

from the theoretical point of view. Thus we consider that

(d) **Inference problem and control problem are the same problem. And these are characterized as the reverse problem of measurements.**

Remark 13.3. [Remark on dynamical system theory (cf. [29])] Again recall the formulation (13.2) of dynamical system theory, in which

($\sharp$) the noise $e_1(t)$ and the measurement error $e_2(t)$ have the same mathematical structure (i.e., stochastic processes).

This is a weak point of dynamical system theory. Since the noise and the measurement error are different, I think that the mathematical formulations should be different. In fact, the confusion between the noise and the measurement error frequently occur. This weakness is clarified in quantum language, as shown in Answer 13.6.

13.2 Regression analysis

According to Fisher's maximum likelihood method (Theorem5.6) and the existence theorem of the realized causal observable, we have the following theorem:

Theorem 13.4. [Regression analysis (cf. [29])] Let $(T{=}\{t_0, t_1, \ldots, t_N\}, \pi : T \setminus \{t_0\} \to T)$ be a tree. Let $\widehat{\mathsf{O}}_T = (\times_{t\in T} X_t, \ \boxtimes_{t\in T}\mathcal{F}_t, \ \widehat{F}_{t_0})$ be the realized causal observable of a sequential causal observable $[\{\mathsf{O}_t\}_{t\in T}, \{\Phi_{\pi(t),t} : L^\infty(\Omega_t) \to L^\infty(\Omega_{\pi(t)})\}_{t\in T\setminus\{t_0\}}]$. Consider a measurement

$$\mathsf{M}_{L^\infty(\Omega_{t_0})}(\widehat{\mathsf{O}}_T{=}(\mathop{\times}_{t\in T} X_t, \ \boxtimes_{t\in T}\mathcal{F}_t, \widehat{F}_{t_0}), S_{[*]})$$

Assume that a measured value obtained by the measurement belongs to $\widehat{\Xi}$ $(\in \ \boxtimes_{t\in T}\mathcal{F}_t)$. Then, there is a reason to infer that

$$[\,*\,] = \omega_{t_0}$$

where ω_{t_0} $(\in \Omega_{t_0})$ is defined by

$$[\widehat{F}_{t_0}(\widehat{\Xi})](\omega_{t_0}) = \max_{\omega\in\Omega_{t_0}} [\widehat{F}_{t_0}(\widehat{\Xi})](\omega)$$

The poof is a direct consequence of Axiom 2 (causality; §10.3) and Fisher maximum likelihood method (Theorem 5.6). Thus, we omit it.
It should be noted that

(♯) **regression analysis is related to Axiom 1 (measurement; §2.7) and Axiom 2 (causality; §10.3)**

Now we shall answer Problem13.1 in terms of quantum language, that is, in terms of regression analysis (Theorem13.4).

Answer 13.5. [(Continued from Problem13.1(Inference problem))Regression analysis] Let $(T{=}\{0,1,2\}, \pi : T \setminus \{0\} \to T)$ be the parent map representation of a tree, where it is assumed that

$$\pi(1) = \pi(2) = 0$$

Put $\Omega_0 = \{\omega_1, \omega_2, \ldots, \omega_5\}$, $\Omega_1 = \text{interval}[100, 200]$, $\Omega_2 = \text{interval}[30, 110]$. Here, we consider that

$$\Omega_0 \ni \omega_n \cdots\cdots \text{a state such that "the girl is helped by a student } \omega_n\text{"} \qquad (n = 1, 2, ..., 5)$$

For each t $(\in \{1, 2\})$, the deterministic map $\phi_{0,t} : \Omega_0 \to \Omega_t$ is defined by $\phi_{0,1} = h$(height function), $\phi_{0,2} = w$(weight function). Thus, for each t $(\in \{1, 2\})$, the deterministic causal operator $\Phi_{0,t} : L^\infty(\Omega_t) \to L^\infty(\Omega_0)$ is defined by

$$[\Phi_{0,t} f_t](\omega) = f_t(\phi_{0,t}(\omega)) \qquad (\forall\omega \in \Omega_0, \forall f_t \in L^\infty(\Omega_t))$$

$$
\begin{tikzcd}
 & L^\infty(\Omega_1) \arrow[dl, "\Phi_{0,1}"'] \\
L^\infty(\Omega_0) & \\
 & L^\infty(\Omega_2) \arrow[ul, "\Phi_{0,2}"]
\end{tikzcd}
$$

For each $t = 1,2$, let $\mathsf{O}_{G_{\sigma_t}}$=$(\mathbb{R}, \mathcal{B}_{\mathbb{R}}, G_{\sigma_t})$ be the normal observable with a standard deviation $\sigma_t > 0$ in $L^\infty(\Omega_t)$. That is,

$$
[G_{\sigma_t}(\Xi)](\omega) = \frac{1}{\sqrt{2\pi\sigma_t^2}} \int_\Xi e^{-\frac{(x-\omega)^2}{2\sigma_t^2}} dx \quad (\forall \Xi \in \mathcal{B}_{\mathbb{R}}, \forall \omega \in \Omega_t)
$$

Thus, we have a deterministic sequence observable $[\{\mathsf{O}_{G_{\sigma_t}}\}_{t=1,2}, \{\Phi_{0,t} : L^\infty(\Omega_t) \to L^\infty(\Omega_0)\}_{t=1,2}]$. Its realization $\widehat{\mathbb{O}}_T = (\mathbb{R}^2, \mathcal{F}_{\mathbb{R}^2}, \widehat{F}_0)$ is defined by

$$
[\widehat{F}_0(\Xi_1 \times \Xi_2)](\omega) = [\Phi_{0,1} G_{\sigma_1}](\omega) \cdot [\Phi_{0,2} G_{\sigma_2}](\omega) = [G_{\sigma_1}(\Xi_1)](\phi_{0,1}(\omega)) \cdot [G_{\sigma_2}(\Xi_2)](\phi_{0,2}(\omega))
$$

$$
(\forall \Xi_1, \Xi_2 \in \mathcal{B}_{\mathbb{R}}, \ \forall \omega \in \Omega_0 = \{\omega_1, \omega_2, \ldots, \omega_5\})
$$

Let N be sufficiently large. Define intervals $\Xi_1, \Xi_2 \subset \mathbb{R}$ by

$$
\Xi_1 = \Big[165 - \frac{1}{N}, 165 + \frac{1}{N} \Big], \qquad \Xi_2 = \Big[65 - \frac{1}{N}, 65 + \frac{1}{N} \Big]
$$

The measured data obtained by a measurement $\mathsf{M}_{L^\infty(\Omega_0)}(\widehat{\mathbb{O}}_T, S_{[*]})$ is

$$
(165, 65) \ (\in \mathbb{R}^2)
$$

Thus, measured value belongs to $\Xi_1 \times \Xi_2$. Using regression analysis (Theorem 13.4) is characterized as follows:

($\sharp$) Find ω_0 ($\in \Omega_0$) such as

$$
[\widehat{F}_0(\{\Xi_1 \times \Xi_2)](\omega_0) = \max_{\omega \in \Omega} [\widehat{F}_0(\{\Xi_1 \times \Xi_2)](\omega)
$$

Since N is sufficiently large,

$$
\begin{aligned}
(\sharp) &\Longrightarrow \max_{\omega \in \Omega_0} \frac{1}{\sqrt{(2\pi)^2 \sigma_1^2 \sigma_2^2}} \iint\limits_{\Xi_1 \times \Xi_2} \exp\big[- \frac{(x_1 - h(\omega))^2}{2\sigma_1^2} - \frac{(x_2 - w(\omega))^2}{2\sigma_2^2} \big] dx_1 dx_2 \\
&\Longrightarrow \max_{\omega \in \Omega_0} \exp\big[- \frac{(165 - h(\omega))^2}{2\sigma_1^2} - \frac{(65 - w(\omega))^2}{2\sigma_2^2} \big] \\
&\Longrightarrow \min_{\omega \in \Omega_0} \big[\frac{(165 - h(\omega))^2}{2\sigma_1^2} + \frac{(65 - w(\omega))^2}{2\sigma_2^2} \big] \qquad (\text{ for simplicity, assume that } \sigma_1 = \sigma_2)
\end{aligned}
$$

$$\Longrightarrow \text{When } \omega_4, \text{ minimum value } \frac{(165-170)^2+(65-60)^2}{2\sigma_1^2} \text{is obtained}$$

$$\Longrightarrow \text{The student is } \omega_4$$

Therefore, we can infer that the student who helps the girl is ω_4. $\square$

Now, let us answer Problem 13.2 in terms of quantum language (or, by using regression analysis (Theorem13.4)).

Answer 13.6. [(Continued from Problem 13.2(Control problem))Regression analysis] In Problem 13.2, it is natural to consider that the tree $T = \{0, 1, 2, 3\}$ is discrete time, that is, the linear ordered set with the parent map $\pi : T \setminus \{0\} \to T$ such that $\pi(t) = t - 1 \quad (t = 1, 2, 3)$. For example, put

$$\Omega_0 = [0,\ 1] \times [0,\ 2],\ \Omega_1 = [0,\ 4] \times [0,\ 2],\ \ \Omega_2 = [0,,\ 6] \times [0,\ 2],\ \ \Omega_3 = [0,\ 8] \times [0,\ 2]$$

For each $t = 1, 2, 3$, define the deterministic causal map $\phi_{\pi(t),t} : \Omega_{\pi(t)} \to \Omega_t$ by (13.3), that is,

$$\begin{aligned}
&\phi_{0,1}(\omega_0) = (\alpha+\beta, \beta) && (\forall \omega_0 = (\alpha, \beta) \in \Omega_0 = [0,\ 1] \times [0,\ 2]) \\
&\phi_{1,2}(\omega_1) = (\alpha+\beta, \beta) && (\forall \omega_1 = (\alpha, \beta) \in \Omega_1 = [0,\ 4] \times [0,\ 2]) \\
&\phi_{2,3}(\omega_2) = (\alpha+\beta, \beta) && (\forall \omega_2 = (\alpha, \beta) \in \Omega_2 = [0,\ 6] \times [0,\ 2])
\end{aligned}$$

Thus, we get the deterministic sequence causal map $\{\phi_{\pi(t),t} : \Omega_{\pi(t)} \to \Omega_t\}_{t \in \{1,2,3\}}$, and the deterministic sequence causal operator $\{\Phi_{\pi(t),t} : L^\infty(\Omega_t) \to L^\infty(\Omega_{\pi(t)})\}_{t \in \{1,2,3\}}$. That is,

$$\begin{aligned}
&(\Phi_{0,1} f_1)(\omega_0) = f_1(\phi_{0,1}(\omega_0)) && (\forall f_1 \in L^\infty(\Omega_1), \forall \omega_0 \in \Omega_0) \\
&(\Phi_{1,2} f_2)(\omega_1) = f_2(\phi_{1,2}(\omega_1)) && (\forall f_2 \in L^\infty(\Omega_2), \forall \omega_1 \in \Omega_1) \\
&(\Phi_{2,3} f_3)(\omega_2) = f_3(\phi_{2,3}(\omega_2)) && (\forall f_3 \in L^\infty(\Omega_3), \forall \omega_1 \in \Omega_2).
\end{aligned}$$

Illustrating by the diagram, we see

$$L^\infty(\Omega_0) \xleftarrow{\Phi_{0,1}} L^\infty(\Omega_1) \xleftarrow{\Phi_{1,2}} L^\infty(\Omega_2) \xleftarrow{\Phi_{2,3}} L^\infty(\Omega_3)$$

And thus, $\phi_{0,2}(\omega_0) = \phi_{1,2}(\phi_{0,1}(\omega_0))$, $\phi_{0,3}(\omega_0) = \phi_{2,3}(\phi_{1,2}(\phi_{0,1}(\omega_0)))$, Therefore, note that $\Phi_{0,2} = \Phi_{0,1} \cdot \Phi_{1,2}, \quad \Phi_{0,3} = \Phi_{0,1} \cdot \Phi_{1,2} \cdot \Phi_{2,3}$.

$$\begin{tikzcd}
 & L^\infty(\Omega_1) \arrow[dl, "\Phi_{0,1}"'] \\
L^\infty(\Omega_0) & L^\infty(\Omega_2) \arrow[l, "\Phi_{0,2}"'] \\
 & L^\infty(\Omega_3) \arrow[ul, "\Phi_{0,3}"]
\end{tikzcd}$$

Let $\mathbb{R}$ be the set of real numbers. Fix $\sigma > 0$. For each $t = 0, 1, 2$, define the *normal observable* $\mathsf{O}_t \equiv (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, G_\sigma)$ in $L^\infty(\Omega_t)$ such that

$$[G_\sigma(\Xi)](\omega_t) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_\Xi \exp(-\frac{(x-\alpha)^2}{2\sigma^2})dx$$
$$(\forall \Xi \in \mathcal{B}_{\mathbb{R}}, \forall \omega_t = (\alpha, \beta) \in \Omega_t = [0,\ 2t+2] \times [0,\ 2]).$$

Thus, we have the deterministic sequential causal observable $[\{\mathsf{O}_t\}_{t=1,2,3}, \{\Phi_{\pi(t),t} : L^\infty(\Omega_t) \to L^\infty(\Omega_{\pi(t)})\}_{t\in\{1,2,3\}}]$.

And thus, we have the realized causal observable $\widehat{\mathsf{O}}_T = (\mathbb{R}^3, \mathcal{F}_{\mathbb{R}^3}, \widehat{F}_0)$ in $L^\infty(\Omega_0)$ such that (using Theorem 12.8)

$$[\widehat{F}_0(\Xi_1 \times \Xi_2 \times \Xi_3)](\omega_0) = \big[\Phi_{0,1}\big(G_\sigma(\Xi_1)\Phi_{1,2}(G_\sigma(\Xi_2)\Phi_{2,3}(G_\sigma(\Xi_3)))\big)\big](\omega_0)$$
$$=[\Phi_{0,1}G_\sigma(\Xi_1)](\omega_0) \cdot [\Phi_{0,2}G_\sigma(\Xi_2)](\omega_0) \cdot [\Phi_{0,3}G_\sigma(\Xi_3)](\omega_0)$$
$$=[G_\sigma(\Xi_1)](\phi_{0,1}(\omega_0)) \cdot [G_\sigma(\Xi_2)](\phi_{0,2}(\omega_0)) \cdot [G_\sigma(\Xi_3)](\phi_{0,3}(\omega_0))$$
$$(\forall \Xi_1, \Xi_2, \Xi_3 \in \mathcal{B}_{\mathbb{R}},\ \forall \omega_0 = (\alpha, \beta) \in \Omega_0 = [0,\ 1] \times [0,\ 2])$$

Our problem (i.e., Problem 13.2) is as follows,

($\sharp_1$) Determine the parameter (α, β) such that the measured value of $\mathsf{M}_{L^\infty(\Omega_0)}(\ \widehat{\mathsf{O}}_T,\ S_{[*]})$ is equal to $(1.9, 3.0, 4.7)$

For a sufficiently large natural number N, put

$$\Xi_1 = \left[1.9 - \frac{1}{N}, 1.9 + \frac{1}{N}\right], \Xi_2 = \left[3.0 - \frac{1}{N}, 3.0 + \frac{1}{N}\right], \Xi_3 = \left[4.7 - \frac{1}{N}, 4.7 + \frac{1}{N}\right]$$

Fisher's maximum likelihood method (Theorem 5.6)) says that the above ($\sharp_1$) is equivalent to the following problem

($\sharp_2$) Find (α, β) $(= \omega_0 \in \Omega_0)$ such that

$$[\widehat{F}_0(\Xi_1 \times \Xi_2 \times \Xi_3)](\alpha, \beta) = \max_{(\alpha,\beta)} [\widehat{F}_0(\Xi_1 \times \Xi_2 \times \Xi_3)]$$

Since N is assumed to be sufficiently large, we see

$$(\sharp_2) \Longrightarrow \max_{(\alpha,\beta)\in\Omega_0} [\widehat{F}_0(\Xi_1 \times \Xi_2 \times \Xi_3)](\alpha, \beta)$$
$$\Longrightarrow \max_{(\alpha,\beta)\in\Omega_0} \frac{1}{\sqrt{2\pi\sigma^2}^3} \iiint_{\Xi_1\times\Xi_2\times\Xi_3} e^{[-\frac{(x_1-(\alpha+\beta))^2+(x_2-(\alpha+2\beta))^2+(x_3-(\alpha+3\beta))^2}{2\sigma^2}]}$$

$$
\begin{aligned}
&\times dx_1 dx_2 dx_3 \\
\Longrightarrow & \max_{(\alpha,\beta)\in\Omega_0} \exp(-J/(2\sigma^2)) \\
\Longrightarrow & \min_{(\alpha,\beta)\in\Omega_0} J
\end{aligned}
$$

where

$$J = (1.9 - (\alpha+\beta))^2 + (3.0 - (\alpha+2\beta))^2 + (4.7 - (\alpha+3\beta))^2$$

($\frac{\partial}{\partial\alpha}\{\cdots\} = 0, \frac{\partial}{\partial\beta}\{\cdots\} = 0$ and thus,)

$$
\begin{aligned}
&\Longrightarrow \begin{cases} (1.9 - (\alpha+\beta)) + (3.0 - (\alpha+2\beta)) + (4.7 - (\alpha+3\beta)) = 0 \\ (1.9 - (\alpha+\beta)) + 2(3.0 - (\alpha+2\beta)) + 3(4.7 - (\alpha+3\beta)) = 0 \end{cases} \\
&\Longrightarrow \quad (\alpha,\beta) = (0.4, 1.4)
\end{aligned}
$$

Therefore, in order to obtain a measured value $(1.9,\ 3.0,\ 4.7)$, it suffices to put

$$(\alpha,\beta) = (0.4,\ 1.4)$$

Remark 13.7. For completeness, note that,

- From the theoretical point of view,

$$\textbf{“inference”} = \textbf{“control”}$$

Thus, we conclude that statistics and dynamical system theory are essentially the same.

Chapter 14

Realized causal observable in classical systems

As mentioned in the previous chapters, what is important is

- **to exercise the relationship of measurement and causality**

In this chapter, we discuss the relationship more systematically. That is, we add the further argument concerning the realized causal observable. This field is too vast, thus, we mainly concentrate our interest to classical systems, particularly, Zeno's paradox. That is,

(♭) to describe the flying arrow (the best work in Zeno's paradoxes) in terms of quantum language (*cf.* refs.[36, 38])[1]

We believe that this is the final answer to Zeno's paradox.

14.1 Infinite realized causal observable in classical systems

In what follows, we shall generalize the argument (concerning the finite realized causal observable in Chapter 12) to infinite case. In the case of infinite trees, it is impossible to discuss quantum system deeply. thus, in this chapter,

we devote ourselves to classical systems

[1] This chapter is extracted from

[36]: S. Ishikawa, *"Zeno's paradoxes in the Mechanical World View,"* arXiv:1205.1290v1 [physics.hist-ph], (2012)

[38]: S. Ishikawa, *Measurement Theory in the Philosophy of Science,* arXiv:1209.3483 [physics.hist-ph] 2012, (177 pages)

Let $(T, \leq)$ be an **infinite tree**, i.e., an infinite tree like semi-ordered set such that

$$\text{“}t_1 \leqq t_3 \text{ and } t_2 \leqq t_3\text{”} \Longrightarrow \text{“}t_1 \leqq t_2 \text{ or } t_2 \leqq t_1\text{”}$$

Put $T^2_{\leq} = \{(t_1, t_2) \in T^2 : t_1 \leq t_2\}$. An element $t_0 \in T$ is called a *root* if $t_0 \leq t$ $(\forall t \in T)$ holds. If T has the root t_0, we sometimes denote T by $T(t_0)$. $T'(\subseteq T)$ is called *lower bounded* if there exists an element $t_i(\in T)$ such that $t_i \leqq t$ $(\forall t \in T')$. Therefore, if T has the root, any $T'(\subseteq T)$ is lower bounded. We always assume that T is complete, that is, for any $T'(\subseteq T)$ which is lower bounded, there exists an element $\mathrm{Inf}_T(T')(\in T)$ that satisfies the following (i) and (ii):

(i) $\mathrm{Inf}_T(T') \leqq t \qquad (\forall t \in T')$

(ii) If $s \leqq t$ $(\forall t \in T')$, then it holds that $s \leqq \mathrm{Inf}_T(T')$

///

Let $(T(t_0), \leqq)$ be an infinite tree with the root t_0. For each $t \in T$, consider the classical basic structure:

$$[C_0(\Omega_t) \subseteq L^\infty(\Omega_t, \nu_t) \subseteq B(L^2(\Omega_t, \nu_t))]$$

Also, for each $t \in T$, define the separable complete metric space X_t, and the Borel field $\mathcal{B}_{X_t}$, and further, define the observable $\mathsf{O}_t{=}(X_t, \mathcal{F}_t, F_t)$ in $L^\infty(\Omega_t, \nu_t)$. That is, we have a **sequential causal observable:**

$$[\mathbb{O}_{T(t_0)}] = [\{\mathsf{O}_t\}_{t \in T}, \{\Phi_{t_1,t_2} : L^\infty(\Omega_{t_2}, \nu_{t_2}) \to L^\infty(\Omega_{t_1}, \nu_{t_1})\}_{(t_1,t_2) \in T^2_{\leqq}}]$$

Now let us construct the realized causal observable in what follows:

Here, define, $\overline{\mathcal{P}}_0(T)$ $(= \overline{\mathcal{P}}_0(T(t_0)) \subseteq \mathcal{P}(T))$ such that

$$\begin{aligned}&\overline{\mathcal{P}}_0(T(t_0))\\ =&\{T' \subseteq T \mid T' \text{ is finite}, t_0 \in T' \text{ and satisfies } \mathrm{Inf}_{T'}S = \mathrm{Inf}_T S \ \ (\forall S \subseteq T')\}\end{aligned}$$

Let $T'(t_0) \in \overline{\mathcal{P}}_0(T(t_0))$. Since $(T'(t_0), \leqq)$ is finite, we can put $(T'{=}\{t_0, t_1, \ldots, t_N\}, \pi : T' \setminus \{t_0\} \to T')$, where π is a parent map.

Review 14.1. [The review of Definition 12.4]. Let $T'(= T'(t_0)) \in \overline{\mathcal{P}}_0(T)$. Consider the sequential causal observable $[\{\mathsf{O}_t\}_{t \in T'}, \{\Phi_{\pi(t),t} : L^\infty(\Omega_t, \nu_t) \to L^\infty(\Omega_{\pi(t)}, \nu_{\pi(t)})\}_{t \in T' \setminus \{t_0\}}]$. For each s $(\in T')$, putting $T_s = \{t \in T' \mid t \geqq s\}$, define the observable $\widehat{\mathsf{O}}_s{=}(\times_{t \in T_s} X_t, \ \times_{t \in T_s} \mathcal{F}_t, \widehat{F}_s)$ in

$L^\infty(\Omega_t, \nu_t)$ such that

$$\widehat{\mathsf{O}}_s = \begin{cases} \mathsf{O}_s & (s \in T' \setminus \pi(T') \quad \text{and }) \\ \\ \mathsf{O}_s \times (\displaystyle\bigtimes_{t \in \pi^{-1}(\{s\})} \Phi_{\pi(t),t} \widehat{\mathsf{O}}_t) & (s \in \pi(T') \quad \text{and }) \end{cases} \tag{14.1}$$

And further, iteratively, we get $\widehat{\mathsf{O}}_{t_0} = (\bigtimes_{t \in T'} X_t, \ \bigtimes_{t \in T'} \mathcal{F}_t, \widehat{F}_{t_0})$, which is also denoted by $\widehat{\mathsf{O}}_{T'} = (\bigtimes_{t \in T'} X_t, \ \bigtimes_{t \in T'} \mathcal{F}_t, \widehat{F}_{T'})$.
$\Big($ In classical cases, the existence is guaranteed by Definition 12.4 $\Big)$

For any subsets $T_1 \subseteq T_2 (\subseteq T)$, define the natural map $\pi_{T_1,T_2} : \bigtimes_{t \in T_2} X_t \longrightarrow \bigtimes_{t \in T_1} X_t$ by

$$\bigtimes_{t \in T_2} X_t \ni (x_t)_{t \in T_2} \mapsto (x_t)_{t \in T_1} \in \bigtimes_{t \in T_1} X_t$$

It is clear that the observables $\Big\{ \widehat{\mathsf{O}}_{T'} = (\bigtimes_{t \in T'} X_t, \ \bigtimes_{t \in T'} \mathcal{F}_t, \widehat{F}_{T'}) \ \ | \ \ T' \in \overline{\mathcal{P}}_0(T) \Big\}$ in $L^\infty(\Omega_{t_0}, \nu_{t_0})$ satisfy the following **consistency condition**, that is,

- for any T_1, T_2 $(\in \overline{\mathcal{P}}_0(T))$ such that $T_1 \subseteq T_2$, it holds that

$$\widehat{F}_{T_2}\big(\pi^{-1}_{T_1,T_2}(\Xi_{T_1})\big) = \widehat{F}_{T_1}\big(\Xi_{T_1}\big) \quad (\forall \Xi_{T_1} \in \bigtimes_{t \in T_1} \mathcal{F}_t)$$

Then, by Theorem 4.1[Kolmogorov extension theorem in measurement theory], there uniquely exists the observable $\widehat{\mathsf{O}}_T = \big(\bigtimes_{t \in T} X_t, \ \boxtimes_{t \in T} \mathcal{F}_t, \ \widehat{F}_T\big)$ in $L^\infty(\Omega_{t_0}, \nu_{t_0})$ such that:

$$\widehat{F}_T\big(\pi^{-1}_{T',T}(\Xi_{T'})\big) = \widehat{F}_{T'}\big(\Xi_{T'}\big) \quad (\forall \Xi_{T'} \in \boxtimes_{t \in T'} \mathcal{F}_t, \ \forall T' \in \overline{\mathcal{P}}_0(T))$$

This observable $\widehat{\mathsf{O}}_T = (\bigtimes_{t \in T} X_t, \ \boxtimes_{t \in T} \mathcal{F}_t, \ \widehat{F}_T)$ is called the realization of the sequential causal observable $[\mathbb{O}_{T(t_0)}] = [\{\mathsf{O}_t\}_{t \in T}, \{\Phi_{t_1,t_2} : L^\infty(\Omega_{t_2}, \nu_{t_2}) \to L^\infty(\Omega_{t_1}, \nu_{t_1})\}_{(t_1,t_2) \in T^2_{\leqq}}]$.

Summing up the above argument, we have the following theorem in classical systems. This is the infinite version of Definition 12.4.

Theorem 14.2. [The existence theorem of an infinite realized causal observable in classical systems] Let T be an infinite tree with the root t_0. For each $t \in T$, consider the basic structure:

$$[C_0(\Omega_t) \subseteq L^\infty(\Omega_t, \nu_t) \subseteq B(L^2(\Omega_t, \nu_t))]$$

Also, for each $t \in T$, define the separable complete metric space X_t, the Borel field $(X_t, \mathcal{F}_t)$ and an observable $\mathsf{O}_t = (X_t, \mathcal{F}_t, F_t)$ in $L^\infty(\Omega_t, \nu_t)$. And, consider the sequential causal observable$[\mathsf{O}_{T(t_0)}] = [\{\mathsf{O}_t\}_{t \in T}, \{\Phi_{t_1,t_2} : L^\infty(\Omega_{t_2}, \nu_{t_2}) \to L^\infty(\Omega_{t_1}, \nu_{t_1})\}_{(t_1,t_2) \in T^2_{\leqq}}]$. Then, there uniquely exists the realized causal observable $\widehat{\mathsf{O}}_T = \big(\times_{t \in T} X_t, \ \boxtimes_{t \in T} \mathcal{F}_t, \ \widehat{F}_T \big)$ in $L^\infty(\Omega_{t_0}, \nu_{t_0})$, that is, it satisfies that

$$\widehat{F}_T\big(\pi^{-1}_{T',T}(\Xi_{T'})\big) = \widehat{F}_{T'}\big(\Xi_{T'}\big) \quad (\forall \Xi_{T'} \in \boxtimes_{t \in T'} \mathcal{F}_t, \ \forall T' \in \overline{\mathcal{P}}_0(T)) \tag{14.2}$$

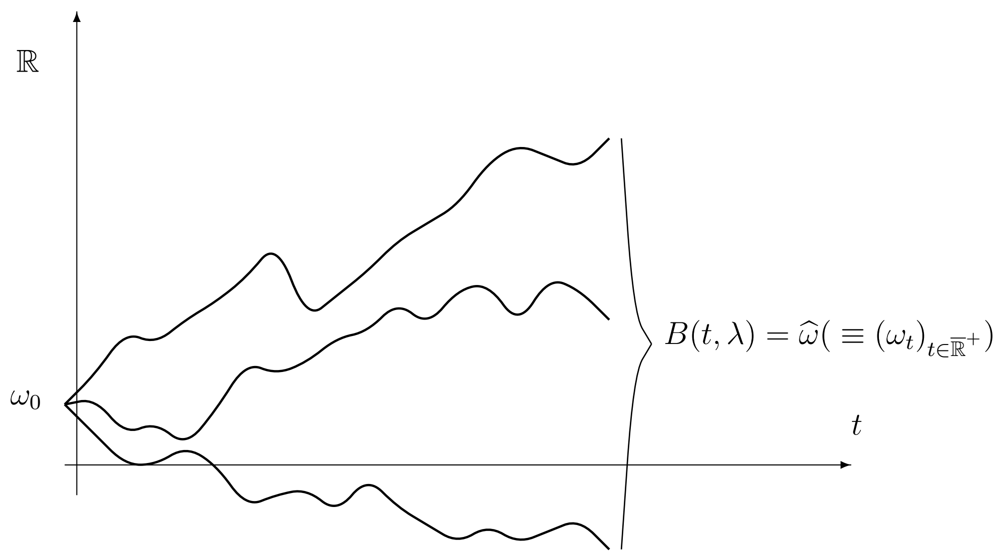
14.2 Is Brownian motion a motion?

14.2.1 Brownian motion in probability theory

There is a reason to consider that

(A) Brownian motion should be understood in measurement theory.

That is because Brownian motion is not in Newtonian mechanics. As one of applications of Theorem 14.2, we discuss the Brown motion in quantum language.



Let us explain the above figure as follows.

Definition 14.3. [The review of Brownian motion in probability theory [54]].
Let $(\Lambda, \mathcal{F}_\Lambda, P)$ be a probability space. For each $\lambda \in \Lambda$, define the real-valued continuous function $B(\cdot, \lambda) : T(=[0,\infty)) \to \mathbb{R}$ such that, for any $t_0 = 0 < t_1 < t_2 < \cdots < t_n$,

$$
\begin{aligned}
& P(\{\lambda \in \Lambda \mid B(t_k, \lambda) \in \Xi_k \in \mathcal{B}_{\mathbb{R}} \;\; (k=1,2,\ldots,n)\}) \\
= & \int_{\Xi_1} \Big(\cdots (\int_{\Xi_{t_{n-1}}} (\int_{\Xi_{t_n}} \mathop{\times}_{k=1}^{n} G_{\sqrt{t_k - t_{k-1}}}(\omega_k - \omega_{k-1}) d\omega_n) d\omega_{n-1}) \cdots \Big) d\omega_1
\end{aligned}
\tag{14.3}
$$

where, $\omega_0 \in \mathbb{R}$, $d\omega_k$ is the Lebesgue measure on $\mathbb{R}$, and $G_{\sqrt{t}}(q) = \frac{1}{\sqrt{2\pi t}} \exp \Big[- \frac{q^2}{2t} \Big]$.
The $B(\cdot, \lambda) : T(=[0,\infty)) \to \mathbb{R}$ is called the **Brownian motion**.

14.2.2 Brownian motion in quantum language

Now consider the diffusion equation:

$$\frac{\partial \rho_t(q)}{\partial t} = \frac{\partial^2 \rho_t(q)}{\partial q^2}, \qquad (\forall q \in \mathbb{R}, \forall t \in T{=}\overline{\mathbb{R}}_+ = [t_0 = 0, \infty)\ )$$

By the solution ρ_t, we get predual operator $\{[\Phi_{t_1,t_2}]_* : L^1(\mathbb{R}, dq) \to L^1(\mathbb{R}, dq)\}$ as follows. That is, for each $\rho_{t_1} \in L^1(\mathbb{R}, m)$, define

$$\big([\Phi_{t_1,t_2}]_*(\rho_{t_1})\big)(q) = \rho_{t_2}(q) = \int_{-\infty}^{\infty} \rho_{t_1}(y) G_{\sqrt{t_2-t_1}}(q-y) m(dy) \ \ (\forall q \in \mathbb{R}, \forall (t_1, t_2) \in T^2_{\leqq})$$

For simplicity, we put $(\Omega_t.\mathcal{B}_{\Omega_t}, d\omega_t) = (\Omega, \mathcal{B}, d\omega) = (\mathbb{R}_q, \mathcal{B}_{\mathbb{R}_q}, dq)$. And thus, for each $t \in T$, consider the classical basic structure:

$$[C_0(\Omega_t) \subseteq L^\infty(\Omega_t, d\omega_t) \subseteq B(L^2(\Omega_t, d\omega_t))]$$

Putting $\Phi_{t_1,t_2} = ([\Phi_{t_1,t_2}]_*)^*$, we get the sequential causal operator

$$\{\Phi_{t_1,t_2} : L^\infty(\Omega_{t_2}, d\omega_{t_2}) \to L^\infty(\Omega_{t_1}, d\omega_{t_1}) \mid (t_1, t_2) \in T^2_{\leq}\}$$

For each $t \in T$, consider the exact observable $\mathsf{O}_t^{(\text{exa})} = (\Omega, \mathcal{B}_\Omega, F^{(\text{exa})})$ in $L^\infty(\Omega, d\omega)$. Thus, we get the sequential causal exact observable $[\mathbb{O}_T] = [\{\mathsf{O}_t^{(\text{exa})}\}_{t\in T};\ \{\Phi_{t_1,t_2} \mid (t_1, t_2) \in T^2_{\leq}\}]$. The existence theorem of the infinite classical realized causal observable (Theorem 14.2) says that $\mathbb{O}_T$ has the realized causal observable $\widehat{\mathsf{O}}_{t_0} = (\Omega^T, \mathcal{B}(\Omega^T), \widehat{F}_{t_0})$ in $L^\infty(\Omega, d\omega)$.

Assume that

(B) a measured value $\widehat{\omega}$ $(= (\omega_t)_{t\in T} \in \Omega^T)$ is obtained by $\mathsf{M}_{L^\infty(\Omega)}(\widehat{\mathsf{O}}_{t_0}, S_{[\delta_{\omega_0}]})$.

Let $T' = \{t_0, t_1, t_2, \cdots, t_n\}$ be a finite subset of T, where $t_0 = 0 < t_1 < t_2 < \cdots < t_n$. Put $\widehat{\Xi} = \times_{t\in T}^{T'} \Xi_t \ \big(\in \mathcal{B}^{\overline{\mathbb{R}}^+}\big)$ where $\Xi_t = \Omega$ $(\forall t \notin T')$. Then, by Axiom 1 (measurement; §2.7) , we see

> the probability that $\widehat{\omega}(= (\omega_t)_{t\in T})$ belongs to the set $\widehat{\Xi} \equiv \times_{t\in T}^{T'} \Xi_t$ is given by $[\widehat{F}_{t_0}(\times_{t\in T}^{T'} \Xi_t)](\omega_0)$

where

$$\begin{aligned}
&[\widehat{F}_{t_0}(\times_{t\in T}^{T'} \Xi_t)](\omega_0) \\
=&\Big(F(\Xi_0)\Phi_{0,t_1}\Big(F(\Xi_{t_1})\cdots\Phi_{t_{n-2},t_{n-1}}\Big(F(\Xi_{t_{n-1}})\big(\Phi_{t_{n-1},t_n}F(\Xi_{t_n})\big)\Big)\cdots\Big)(\omega_0) \\
=&\int_{\Xi_1}\Big(\cdots\Big(\int_{\Xi_{t_{n-1}}}\Big(\int_{\Xi_{t_n}} \times_{k=1}^n G_{\sqrt{t_k - t_{k-1}}}(\omega_k - \omega_{k-1}) d\omega_n\Big) d\omega_{n-1}\Big)\cdots\Big) d\omega_1
\end{aligned} \tag{14.4}$$

which is equal to the (14.3).

Thus, we see that

$$\underset{\text{Brownian motion}}{\overset{\text{probability theory}}{\boxed{\Big(B(t,\cdot) \Big)_{t \in T}}}} = \underset{\text{measured value}}{\overset{\text{quantum language}}{\boxed{\Big(\widehat{\omega}_t \Big)_{t \in T}}}}$$

♠**Note 14.1.** Thus, the following assertion has a reason in some sense:

- The Brownian motion $B(t, \lambda)$ is not a motion but a measured value. Some may recall Parmenides' saying:

  ($\sharp$) *There are no "plurality", but only "one". And therefore, there is no movement.*

  which is the same as the essence of the linguistic interpretation.

That is, the spirit of quantum language says that

($\sharp$) *Describe "plurality" as if only "one".*

($\sharp$) *Describe moving one as if not moving.*

14.3 The Schrödinger picture of the sequential deterministic causal operator

14.3.1 The preparation of the next section (§14.4: Zeno's paradox)

The linguistic interpretation (§3.1) says that

a state does no move,

which is called the Heisenberg picture (i.e., a state does not move, and, an observable moves). This is formal. On the other hand, we sometimes use the Schrödinger picture (i.e., a state moves, and, an observable does not move), which is handy and makeshift.

In this section, we explain something about the Schrödinger picture in classical deterministic systems.

This section is the preparation of the next section (Zeno's paradoxes).

Let $(T(t_0), \leqq)$ be an infinite tree with the root t_0. For each $t \in T$, consider the classical basic structure:

$$[C_0(\Omega_t) \subseteq L^\infty(\Omega_t, \nu_t) \subseteq B(L^2(\Omega_t, \nu_t))]$$

Definition 14.4. [State changes — the Schrödinger picture] Let $\{\Phi_{t_1,t_2} : L^\infty(\Omega_{t_2}, \nu_{t_2}) \to L^\infty(\Omega_{t_1}), \nu_{t_1})\}_{(t_1,t_2)\in T^2_{\leqq}}$ be a deterministic causal relation with the deterministic causal maps $\phi_{t_1,t_2} : \Omega_{t_1} \to \Omega_{t_2}$ $(\forall (t_1, t_2) \in T^2_{\leqq})$. Let $\omega_{t_0} \in \Omega_{t_0}$ be an initial state. Then, the $\{\phi_{t_0,t}(\omega_{t_0})\}_{t\in T}$ (or, $\{\delta_{\phi_{t_0,t}(\omega_{t_0})}\}_{t\in T}$ is called the Schrödinger picture representation.

The following is the infinite version of Theorem12.8.

Theorem 14.5. [Deterministic sequential causal operator and realized causal observable] Let $(T(t_0), \leqq)$ be an infinite tree with the root t_0. Let $[\mathbb{O}_T] = [\{\mathsf{O}_t\}_{t\in T}, \{\Phi_{t_1,t_2} : L^\infty(\Omega_{t_2}, \nu_{t_2}) \to L^\infty(\Omega_{t_1}, \nu_{t_1})\}_{(t_1,t_2)\in T^2_{\leqq}}]$ be a deterministic sequential causal observable. Then, the realization $\widehat{\mathsf{O}}_{t_0} \equiv (\times_{t\in T} X_t, \boxtimes_{t\in T} \mathcal{F}_t, \widehat{F}_{t_0})$ is represented by

$$\widehat{\mathsf{O}}_{t_0} = \mathop{\times}_{t\in T} \Phi_{t_0,t}\mathsf{O}_t$$

That is, it holds that

$$[\widehat{F}_{t_0}(\mathop{\times}_{t\in T} \Xi_t)](\omega_{t_0}) = \mathop{\times}_{t\in T} [\Phi_{t_0,t}F_t(\Xi_t)](\omega_{t_0}) = \mathop{\times}_{t\in T} [F_t(\Xi_t)](\phi_{t_0,t}(\omega_{t_0}))$$

KSTS/RR-16/001
January 8, 2016

$$(\forall \omega_{t_0} \in \Omega_{t_0}, \forall \Xi_t \in \mathcal{F}_t)$$

Proof. The proof is similar to that of Theorem12.8 □

Theorem 14.6. Let $[\mathbb{O}_{T(t_0)}] = [\{\mathsf{O}_t^{(\text{exa})}\}_{t \in T}, \{\Phi_{t_1,t_2} : L^\infty(\Omega_{t_2}, \nu_{t_2}) \to L^\infty(\Omega_{t_1}, \nu_{t_1})\}_{(t_1,t_2) \in T^2_{\leqq}}]$ be a deterministic sequential causal exact observable, which has the deterministic causal maps $\phi_{t_1,t_2} : \Omega_{t_1} \to \Omega_{t_2}$ $(\forall (t_1, t_2) \in T^2_{\leqq})$. And let $\widehat{\mathsf{O}}_{t_0} = (\times_{t \in T} X_t, \times_{t \in T} \mathcal{F}_t, \widehat{F}_T)$ be its realized causal observable in $L^\infty(\Omega_{t_0}, \nu_{t_0})$. Assume that the measured value $(x_t)_{t \in T}$ is obtained by $\mathsf{M}_{L^\infty(\Omega_{t_0})}(\widehat{\mathbb{O}}_T = (\times_{t \in T} X_t, \times_{t \in T} \mathcal{F}_t, \widehat{F}_0), S_{[\omega_{t_0}]})$. Then, we surely believe that

$$x_t = \phi_{t_0,t}(\omega_{t_0}) \qquad (\forall t \in T)$$

Thus, we say that, as far as a deterministic sequential causal observable,

(a) exact measured value $(x_t)_{t \in T}$ = the Schrödinger picture representation $(\phi_{t_0,t}(\omega_{t_0}))_{t \in T}$

Proof. Let $D = \{t_1, t_2, \ldots, t_n\} (\subseteq T)$ be any finite subset of T. Put $\widehat{\Xi} = \times^D_{t \in T} \Xi_t = (\times_{t \in D} \Xi_t) \times (\times_{t \in T \setminus D} X_t)$, where $\Xi_t \subseteq X_t (= \Omega_t)$ is an open set such that $\phi_{t_0,t}(\omega_{t_0}) \in \Xi_t$ $(\forall t \in D)$. Then, we see that

(b) the probability that the measured value $(x_t)_{t \in T}$ belongs to $\widehat{\Xi} = \times^D_{t \in T} \Xi_t$ is equal to 1.

That is because Theorem 14.5 says that

$$\begin{aligned} &\big(\widehat{F}_T(\widehat{\Xi})\big)(\omega_{t_0}) = \Big(\mathop{\times}_{k=1}^{n} \big(\Phi_{t_0,t_k} F^{(\text{exa})}(\Xi_{t_k})\big)\Big)(\omega_{t_0}) \\ =&\Big(\mathop{\times}_{k=1}^{n} F^{(\text{exa})}(\phi^{-1}_{t_0,t_k}(\Xi_{t_k}))\Big)(\omega_{t_0}) = \mathop{\times}_{k=1}^{n} \chi_{\Xi_{t_k}}(\phi_{t_0,t_k}(\omega_{t_0})) = 1 \end{aligned}$$

Thus, from the arbitrariness of Ξ_t, we surely believe that

(c) $(x_t)_{t \in T} = \phi_{t_0,t}(\omega_{t_0}) \qquad (\forall t \in T)$

□

♠**Note 14.2.** Note that "(b) ⇔(c)" in the above. That is, (b) is the definition of (c).

Thus, we have the following corollary, which is the generalization of Theorem 3.15.

Corollary 14.7. [System quantity and exact observable]. For each $t \in T(t_0)$, consider the exact observable $\mathsf{O}_t^{(\text{exa})} = (X, \mathcal{F}_t, F^{(\text{exa})})(= (\Omega_t, \mathcal{B}_t, \chi))$ in $L^\infty(\Omega_t, \nu_t)$ and a system quantity $g_t : \Omega_t \to \mathbb{R}$ on Ω_t. Let $\mathsf{O}'_t = (\mathbb{R}, \mathcal{B}_\mathbb{R}, G_t)$ be the observable representation of the quantity g_t in $L^\infty(\Omega_t)$. Assuming the simultaneous observable $\mathsf{O}_t^{(\text{exa})} \times \mathsf{O}'_t$, define the sequential deterministic causal observable:

$$[\mathbb{O}_{T(t_0)}] = [\{\mathsf{O}_t^{(\text{exa})} \times \mathsf{O}'_t\}_{t \in T}, \{\Phi_{t_1,t_2} : L^\infty(\Omega_{t_2}, \nu_{t_2}) \to L^\infty(\Omega_{t_1}, \nu_{t_1})\}_{(t_1,t_2) \in T^2_{\leqq}}]$$

Let $\phi_{t_1,t_2} : \Omega_{t_1} \to \Omega_{t_2}$ $(\forall (t_1, t_2) \in T^2_{\leqq})$ be the deterministic causal map. Let $\widehat{\mathsf{O}}_{t_0} = \big(\times_{t \in T}(X_t \times \mathbb{R}), \boxtimes_{t \in T}(\mathcal{F}_t \boxtimes \mathcal{B}_\mathbb{R}), \widehat{F}_{t_0}\big)$ be the realized causal observable. Thus, we have the measurement $\mathsf{M}_{L^\infty(\Omega_{t_0})}(\widehat{\mathsf{O}}_{t_0}, S_{[\omega_{t_0}]})$. Let $(x_t, y_t)_{t \in T}$ be the measured value obtained by the measurement $\mathsf{M}_{L^\infty(\Omega_{t_0})}(\widehat{\mathsf{O}}_{t_0}, S_{[\omega_{t_0}]})$. Then, we can surely believe that

$$x_t = \phi_{t_0,t}(\omega_{t_0}) \quad \text{and} \quad y_t = g_t(\phi_{t_0,t}(\omega_{t_0})) \qquad (\forall t \in T)$$

Remark 14.8. [Why doesn't Newtonian mechanics have measurement?]. Newtonian mechanics and quantum mechanics are formulated as follows:

$$(\sharp)\left\{\begin{array}{llll} \boxed{\text{Newtoinan mechanics}} & = & \text{Nothing} & + \underset{\text{(Newtonian equation)}}{\boxed{\text{Causality}}} \\ \boxed{\text{quantum mechanics}} & = & \underset{\text{(Born's quantum measurement)}}{\boxed{\text{Measurement}}} & + \underset{\text{(Heisenberg (and Schrödinger) equation)}}{\boxed{\text{Causality}}} \end{array}\right.$$

Thus, the following question is natural:

($\sharp_2$) Why doesn't Newtonian mechanics have measurement ?

Some may think that the reason is due to Theorem 14.6 (or, Corollary 14.7), which says that we need only $\phi_{t_0,t}(\omega_{t_0})$ and not x_t. However, this answer is superficial. The question ($\sharp_2$) is significant in the light of Einstein's words:

($\sharp_3$) *The moon is there whether one looks at it or not.*

in Einstein and Tagore's conversation. This should be compared with Berkley's words "To be is to be perceived". We believe that the ($\sharp_3$) is the same as ($\sharp_4$) (= ($\sharp_5$)):

($\sharp_4$) *Physics should exist without measurement*

($\sharp_5$) *The concept of "measurement" is metaphysical and not physical*

14.4 Zeno's paradoxes—Flying arrow is at rest

First we explain *what Zeno's paradox means*, one of the oldest paradoxes in science.

14.4.1 What is Zeno's paradox?

Although Zeno's paradox has some types (i.e., "flying arrow", "Achilles and a tortoise", "dichotomy", "stadium", etc.), I think that **these are essentially the same problem**. And I think that the flying arrow expresses the essence of the problem exactly and is the first masterpiece in Zeno's paradoxes. However, since "Achilles and the tortoise" may be more famous, I will also describe this as follows.

Paradox 14.9. [Zeno's paradox]
[Flying arrow is at rest]

- Consider a flying arrow. In any one instant of time, the arrow is not moving. Therefore, If the arrow is motionless at every instant, and time is entirely composed of instants, then motion is impossible.

[Achilles and a tortoise]

- I consider competition of Achilles and a tortoise. Let the start point of a tortoise (a late runner) be the front from the starting point of Achilles (a quick runner). Suppose that both started simultaneously. If Achilles tries to pass a tortoise, Achilles has to go to the place in which a tortoise is present now. However, then, the tortoise should have gone ahead more. Achilles has to go to the place in which a tortoise is present now further. Even Achilles continues this infinite, he can never catch up with a tortoise.

In order to explain

"What is Zeno's paradox?"

we have to start from the following Figure. That is, we assert that

Zeno's paradox can not be understood without the following figure:

Figure 14.10. [=Figure 1.1: The location of quantum language in the history of world-description (*cf.* ref.[31])]

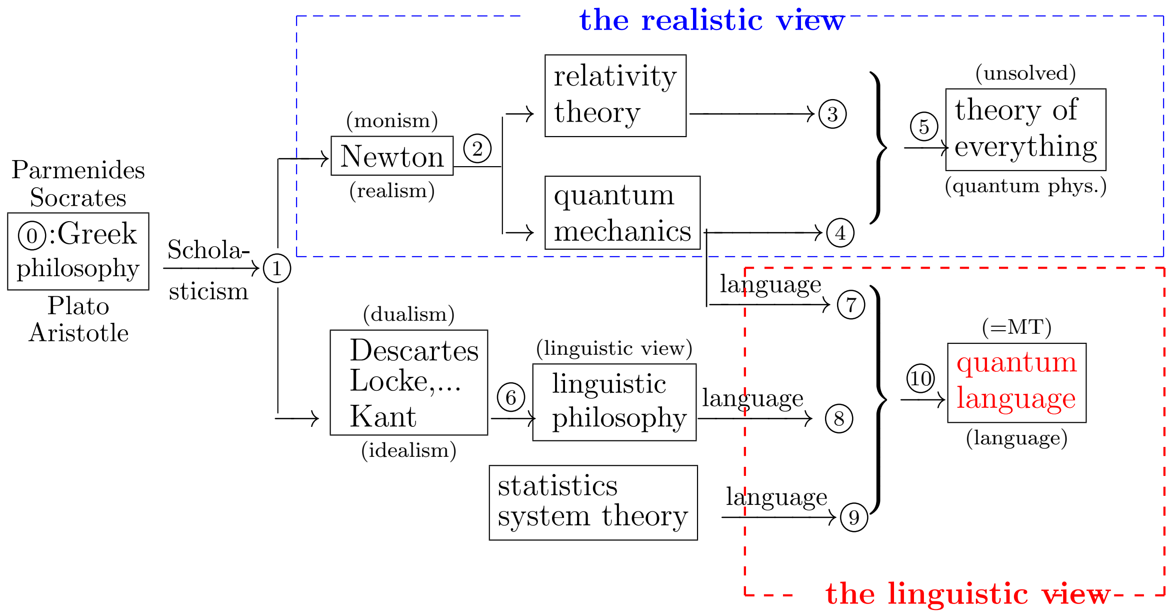


Figure 1.1: The history of the world-view

It is clear that

(A) Descartes=Kant philosophy and the philosophy of language have no power to describe Zeno's paradox 14.9.

However, we have the following problems:

(B_1) How do we describe Zeno's paradox 14.9 in terms of Newtonian mechanics?

(B_2) How do we describe Zeno's paradox 14.9 in terms of quantum mechanics?

(B_3) How do we describe Zeno's paradox 14.9 in terms of the theory of relativity?

(B_4) How do we describe Zeno's paradox 14.9 in terms of statistics (i.e., the dynamical system theory) ?

(B_5) How do we describe Zeno's paradox 14.9 in terms of quantum language?

And, finally, we have

(C) What is the most proper world description for Zeno's paradox 14.9?

We assert that

(D) "to solve Zeno's paradox 14.9" $\Longleftrightarrow$ "to answer the above (C)"

and conclude that

(E) The answer of the above (C) is just quantum language

Therefore, it suffices to answer the above **(B_5)**, that is,

Problem 14.11. [The meaning of Zeno's paradox]

Describe "flying arrow" and "Achilles an a tortoise" in (classical) quantum language!

14.4.2 The answer to (B_4): the dynamical system theoretical answer to Zeno's paradox

Before the answer of Problem 14.11, we give the answer to the Problem (B_4), i.e., the dynamical system theoretical answer. However, in order to do it, we have to start from the formulation of dynamical system theory in what follows

.

14.4.2.1 The formulation of dynamical system theory

Although statistics and dynamical system theory have no clear formulations, as mentioned in Chapter 13, we have the opinion that statistics and dynamical system theory are the same things. At least, the following formulation (i.e., the formulation of dynamical system theory in the narrow sense) should belong to statistics.

Formulation 14.12. [The formulation of dynamical system theory in the narrow sense]

Dynamical system theory is formulated as follows.

$$\boxed{\text{Dynamical system theory}} = \boxed{\text{①:State equation}} + \boxed{\text{②:Measurement equation}} \tag{14.5}$$

①: $\boxed{\text{State equation}}$ is as follows. Let $T = \mathbb{R}$ be the time axis. For each $t (\in T)$, consider the state space $\Omega_t = \mathbb{R}^n$ (n-dimensional real space). The **state equation** (Chap. 13(13.2)) is defined by the following simultaneous ordinary differential equation of the first order

$$\boxed{\text{State equation}} = \begin{cases} \frac{d\omega_1}{dt}(t) = v_1(\omega_1(t), \omega_2(t), \ldots, \omega_n(t), \epsilon_1(t), t) \\ \frac{d\omega_2}{dt}(t) = v_2(\omega_1(t), \omega_2(t), \ldots, \omega_n(t), \epsilon_2(t), t) \\ \cdots\cdots \\ \frac{d\omega_n}{dt}(t) = v_n(\omega_1(t), \omega_2(t), \ldots, \omega_n(t), \epsilon_n(t), t) \end{cases} \tag{14.6}$$

where $\epsilon_k(t)$ is a noise $(k = 1, 2, \cdots, n)$.

②: $\boxed{\text{Measurement equation}}$ is as follows. Consider the measured value space $X = \mathbb{R}^m$ (m-dimensional real space). The **measurement equation** (Chap. 13(13.2)) is defined by

$$\boxed{\text{Measurement equation}} = \begin{cases} x_1(t) = g_1(\omega_1(t), \omega_2(t), \ldots, \omega_n(t), \eta_1(t), t) \\ x_2(t) = g_2(\omega_1(t), \omega_2(t), \ldots, \eta_n(t), \eta_2(t), t) \\ \cdots\cdots \\ x_m(t) = g_m(\omega_1(t), \omega_2(t), \ldots, \eta_n(t), \eta_n(t), t) \end{cases} \tag{14.7}$$

where $g (= (g_1, g_2, \cdots, g_n)) : \Omega \times \mathbb{R}^2 \to X$ is the system quantity and $\eta_k(t)$ is a noise $(k = 1, 2, \cdots, m)$. Here, $x(t) (= (x_1(t), x_2(t), \cdots, x_n(t)))$ is called a motion function.

14.4.2.2 The dynamical system theoretical answer to Zeno's paradox

Answer 14.13. [The dynamical system theoretical answer to "flying arrow (in Paradox 14.9)"]

Let $q(t)$ be the position of the flying arrow at time t. That is, consider the motion function $q(t)$.

- Note that the following logic (i.e., Zeno's logic) is wrong:
  - for each time t, the position $q(t)$ of the flying arrow is determined.
    $\Longrightarrow$
    the motion function q is a constant function

Thus, Zeno's logic is wrong.

[The dynamical system theoretical answer to "Achilles and a tortoise (in Paradox 14.9)"] For example, assume that the velocity v_q [resp. v_s] of the quickest [resp. slowest] runner is equal to $v(>0)$ [resp. γv $(0<\gamma<1)$]. And further, assume that the position of the quickest [resp. slowest] runner at time $t=0$ is equal to 0 [resp. a (>0)]. Thus, we can assume that the position $\xi(t)$ of the quickest runner and the position $\eta(t)$ of the slowest runner at time t (≥ 0) is respectively represented by

$$\begin{cases} \xi(t) = vt \\ \eta(t) = \gamma vt + a \end{cases} \tag{14.8}$$

- **Calculations**

The formula (14.8) can be calculated as follows (i.e., (i) or (ii)):
[(i): Algebraic calculation of (14.8)]:
Solving $\xi(s_0)=\eta(s_0)$, that is,

$$vs_0 = \gamma vs_0 + a$$

we get $s_0 = \frac{a}{(1-\gamma)v}$. That is, at time $s_0 = \frac{a}{(1-\gamma)v}$, the fast runner catches up with the slow runner.
[(ii): Iterative calculation of (14.8)]:
Define t_k $(k=0,1,...)$ such that, $t_0=0$ and

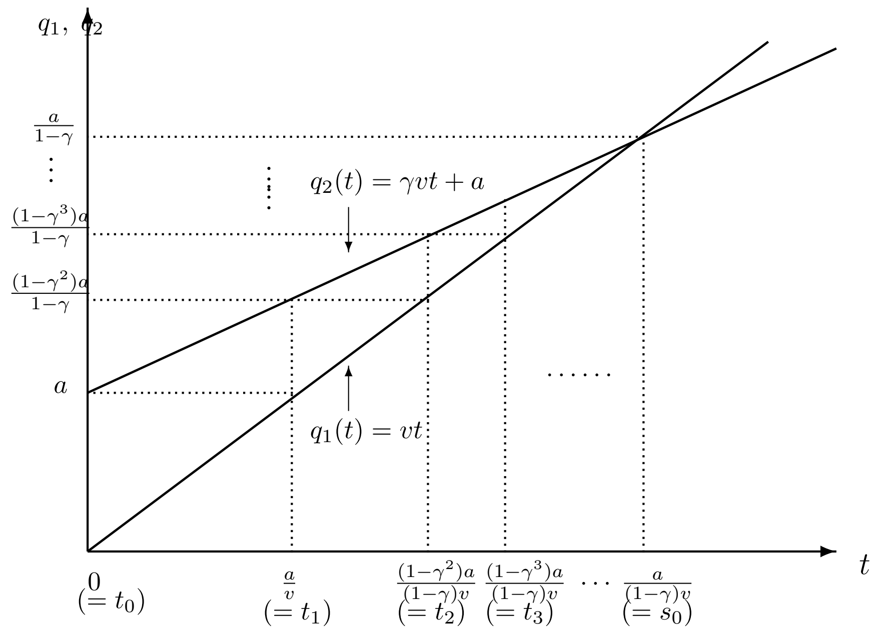
$$t_{k+1} = \gamma vt_k + a \quad (k=0,1,2,...)$$

Thus, we see that $t_k = \frac{(1-\gamma^k)a}{(1-\gamma)v}$ $(k=0,1,...)$. Then, we have that

$$\begin{aligned} \big(\xi(t_k),\eta(t_k)\big) &= \big(\frac{(1-\gamma^k)a}{1-\gamma}, \frac{(1-\gamma^{k+1})a}{1-\gamma}\big) \\ &\to \big(\frac{a}{1-\gamma}, \frac{a}{1-\gamma}\big) \end{aligned} \tag{14.9}$$

as $k\to\infty$. Therefore, the quickest runner catches up with the slowest at time $s_0 = \frac{a}{(1-\gamma)v}$.
[(iii): Conclusion]: After all, by the above (i) or (ii), we can conclude that

($\sharp$) the quickest runner can overtake the slowest at time $s_0 = \frac{a}{(1-\gamma)v}$.



The graph of $\mathbf{q_1(t) = vt,\ q_2(t) = \gamma vt + a}$

14.4.2.3 Why isn't the Answer 14.13 authorized?

We believe that the Answer 14.13 is not the wrong answer of Zeno's paradox. If so, we have to answer the following question:

(F) **Why isn't the Answer 14.13 accepted as the final answer of Zeno's paradox?**

We of course believe that

(G_1) **the reason is due to the fact that statistics (=dynamical system theory) is not accepted as the world-view in Figure 14.10.**

Or equivalently,

(G_1) **the linguistic world-view is not accepted as the world-view in Figure 14.10.**

If so, the readers note that

(H) **the purpose of this note is to assert that the linguistic world view should be authorized in Figure 14.10.**

14.4.3 Quantum linguistic answer to Zeno's paradoxes

Before reading Answer 14.14 (Zeno's paradox(flying arrow)), confirm our spirit:

(I) The theory described in ordinary language should be described in a certain world description. That is because almost ambiguous problems are due to the lack of "the world-description method".

Therefore,

(J) it suffices to describe "**motion function** $q(t)$ in Answer 14.13 (flying arrow)" in terms of quantum language. Here, the motion function should be a **measured value**, in which the **causality** is concealed.

This will be done as follows.

Answer 14.14. [The answer to Problem14.11] or [Answer to Problem 14.9: Zeno's paradox(flying arrow) (*cf.* ref. [36, 38])] In Corollary 14.7, putting

$$q(t) = y_t (= g_t(\phi_{t_0,t}(\omega_{t_0})))$$

we get the time-position function $q(t)$.

Although there may be several opinions, we consider that the followings (i.e., (K_1) and (K_2)) are equivalent:

(K_1) to accept Figure 14.10:[The history of the world-view]

(K_2) to believe in Answer 14.14 as the final answer of Zeno's paradox

♠**Note 14.3.** I think that "the flying arrow" is Zeno's best work. If readers agree to the above answer, they can easily answer the other Zeno's paradoxes. Also, it should be noted that Zeno of Elea (BC. 490-430) was a Greek philosopher (about 2500 years ago). Hence, we are not concerned with the historical aspect of Zeno's paradoxes. Therefore, we think that

($\sharp$) "How did Zeno think Zeno's paradoxes?" is not important from the scientific point of view.

and

($\sharp$) What is important is "How do we think Zeno's paradoxes?"

Also, for the quantum linguistic space-time, see §10.7 (Leibniz=Clarke correspondence). I doubt great philosophers' opinions concerning Zeno's paradoxes.

KSTS/RR-16/001
January 8, 2016

Chapter 15

Least-squares method and Regression analysis

Although regression analysis has a great history, we consider that it has always continued being confused. For example, the fundamental terms in regression analysis (e.g., "regression", "least-squares method", "explanatory variable", "response variable", etc.) seem to be historically conventional, that is, these words do not express the essence of regression analysis. In this chapter, we show that the least squares method acquires a quantum linguistic story as follows.

$$\underset{\text{(Section 15.1)}}{\boxed{\text{The least squares method}}} \xrightarrow[\text{quantum language}]{\text{describe by}} \underset{\text{(Section 15.2)}}{\boxed{\text{Regression analysis}}}$$
$$\xrightarrow[\text{generalization}]{\text{natural}} \underset{\text{(Section 15.4)}}{\boxed{\text{Generalized linear model}}} \tag{$\sharp$}$$

In this story, the terms "explanatory variable" and "response variable" are clarified in terms of quantum language. As the general theory of regression analysis, it suffices to devote ourselves to Theorem 13.4. However, from the practical point of view, we have to add the above story $(\sharp)$[1].

15.1 The least squares method

Let us start from the simple explanation of the least-squares method. Let $\{(a_i, x_i)\}_{i=1}^n$ be a sequence in the two dimensional real space $\mathbb{R}^2$. Let $\phi^{(\beta_1,\beta_2)} : \mathbb{R} \to \mathbb{R}$ be the simple function such that

$$\mathbb{R} \ni a \mapsto x = \phi^{(\beta_1,\beta_2)}(a) = \beta_1 a + \beta_0 \in \mathbb{R} \tag{15.1}$$

---

[1] This chapter is extracted from

- Ref. [42]: S. Ishikawa; Regression analysis in quantum language (arxiv:1403.0060[math.ST],(2014)) where the pair $(\beta_1, \beta_2)(\in \mathbb{R}^2)$ is assumed to be unknown. Define the error σ by

$$\sigma^2(\beta_1, \beta_2) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \phi^{(\beta_1,\beta_2)}(a_i))^2 \Big(= \frac{1}{n}\sum_{i=1}^{n}(x_i - (\beta_1 a_i + \beta_0))^2 \Big) \tag{15.2}$$

Then, we have the following minimization problem:

Problem 15.1. [The least squares method].

Let $\{(a_i, x_i)\}_{i=1}^n$ be a sequence in the two dimensional real space $\mathbb{R}^2$.
Find the $(\hat{\beta}_0, \hat{\beta}_1)$ $(\in \mathbb{R}^2)$ such that

$$\sigma^2(\hat{\beta}_0, \hat{\beta}_1) = \min_{(\beta_1,\beta_2)\in\mathbb{R}^2} \sigma^2(\beta_1, \beta_2) \Big(= \min_{(\beta_1,\beta_2)\in\mathbb{R}^2} \frac{1}{n}\sum_{i=1}^{n}(x_i - (\beta_1 a_i + \beta_0))^2 \Big) \tag{15.3}$$

where $(\hat{\beta}_0, \hat{\beta}_1)$ is called "sample regression coefficients".

This is easily solved as follows. Taking partial derivatives with respect to β_0, β_1, and equating the results to zero, gives the equations (i.e., "likelihood equations"),

$$\frac{\partial \sigma^2(\beta_1, \beta_2)}{\partial \beta_0} = \sum_{i=1}^{n}(x_i - \beta_0 - \beta_1 a_i) = 0, \quad (i = 1, ..., n) \tag{15.4}$$

$$\frac{\partial \sigma^2(\beta_1, \beta_2)}{\partial \beta_1} = \sum_{i=1}^{n}(x_i - \beta_0 - \beta_1 a_i)a_i = 0, \quad (i = 1, ..., n) \tag{15.5}$$

Solving it, we get that

$$\hat{\beta}_1 = \frac{s_{ax}}{s_{aa}}, \quad \hat{\beta}_0 = \overline{x} - \frac{s_{ax}}{s_{aa}}\overline{a}, \quad \hat{\sigma}^2 \Big(= \frac{1}{n}\sum_{i=1}^{n}(x_i - (\hat{\beta}_1 a_i + \hat{\beta}_0))^2 \Big) = s_{xx} - \frac{s_{ax}^2}{s_{aa}} \tag{15.6}$$

where

$$\bar{a} = \frac{a_1 + \cdots + a_n}{n}, \qquad \bar{x} = \frac{x_1 + \cdots + x_n}{n}, \tag{15.7}$$

$$s_{aa} = \frac{(a_1 - \bar{a})^2 + \cdots + (a_n - \bar{a})^2}{n}, \quad s_{xx} = \frac{(x_1 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n}, \tag{15.8}$$

$$s_{ax} = \frac{(a_1 - \bar{a})(x_1 - \bar{x}) + \cdots + (a_n - \bar{a})(x_n - \bar{x})}{n}. \tag{15.9}$$

Remark 15.2. [Applied mathematics]. Note that the above result is in (applied) mathematics, that is,

- the above is neither in statistics nor in quantum language.

The purpose of this chapter is to add a quantum linguistic story to Problem 15.1 (i.e., the least-squares method) in the framework of quantum language.

15.2 Regression analysis in quantum language

Put $T = \{0, 1, 2, \cdots, i, \cdots, n\}$. And let $(T, \tau : T \setminus \{0\} \to T)$ be the parallel tree such that

$$\tau(i) = 0 \qquad (\forall i = 1, 2, \cdots, n) \tag{15.10}$$

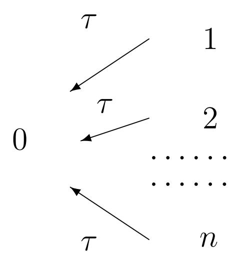


Figure 15.1: Parallel structure

♠**Note 15.1.** In regression analysis, we usually devote ourselves to "classical deterministic causal relation". Thus, Theorem 12.8 is important, which says that it suffices to consider only the parallel structure.

For each $i \in T$, define a locally compact space Ω_i such that

$$\Omega_0 = \mathbb{R}^2 = \left\{ \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} \; : \; \beta_0, \beta_1 \in \mathbb{R} \right\} \tag{15.11}$$

$$\Omega_i = \mathbb{R} = \left\{ \mu_i \; : \; \mu_i \in \mathbb{R} \right\} \quad (i = 1, 2, \cdots, n) \tag{15.12}$$

where the Lebesgue measures m_i are assumed.

Assume that

$$a_i \in \mathbb{R} \qquad (i = 1, 2, \cdots, n), \tag{15.13}$$

which are called *explanatory variables* in the conventional statistics. Consider the deterministic causal map $\psi_{a_i} : \Omega_0 (= \mathbb{R}^2) \to \Omega_i (= \mathbb{R})$ such that

$$\Omega_0 = \mathbb{R}^2 \ni \beta = (\beta_0, \beta_1) \mapsto \psi_{a_i}(\beta_0, \beta_1) = \beta_0 + \beta_1 a_i = \mu_i \in \Omega_i = \mathbb{R} \tag{15.14}$$

which is equivalent to the deterministic causal operator $\Psi_{a_i} : L^\infty(\Omega_i) \to L^\infty(\Omega_0)$ such that

$$[\Psi_{a_i}(f_i)](\omega_0) = f_i(\psi_{a_i}(\omega_0)) \quad (\forall f_i \in L^\infty(\Omega_i), \;\; \forall \omega_0 \in \Omega_0, \forall i \in 1, 2, \cdots, n) \tag{15.15}$$

$$
\begin{tikzcd}
 & L^\infty(\Omega_1(\equiv \mathbb{R})) \arrow[dl, "\Psi_{a_1}"'] \\
L^\infty(\Omega_0(\equiv \mathbb{R}^2)) & L^\infty(\Omega_2(\equiv \mathbb{R})) \arrow[l, "\Psi_{a_2}"'] \\
 & \cdots\cdots \\
 & L^\infty(\Omega_n(\equiv \mathbb{R})) \arrow[uul, "\Psi_{a_n}"]
\end{tikzcd}
$$

Figure 15.2: Parallel structure (Causal relation Ψ_{a_i})

Thus, under the identification: $a_i \Leftrightarrow \Psi_{a_i}$, the term *"explanatory variable"* means a kind of causal relation Ψ_{a_i}.

For each $i = 1, 2, \cdots, n$, define the *normal observable* $\mathsf{O}_i {\equiv} (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, G_\sigma)$ in $L^\infty(\Omega_i(\equiv \mathbb{R}))$ such that

$$
[G_\sigma(\Xi)](\mu) = \frac{1}{(\sqrt{2\pi\sigma^2})} \int_\Xi \exp\Big[-\frac{(x-\mu)^2}{2\sigma^2} \Big] dx \qquad (\forall \Xi \in \mathcal{B}_{\mathbb{R}}, \forall \mu \in \Omega_i(\equiv \mathbb{R})) \tag{15.16}
$$

where σ is a positive constant.
Thus, we have the observable $\mathsf{O}_0^{a_i} {\equiv} (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, \Psi_{a_i} G_\sigma)$ in $L^\infty(\Omega_0(\equiv \mathbb{R}^2))$ such that

$$
[\Psi_{a_i}(G_\sigma(\Xi))](\beta) = [(G_\sigma(\Xi))](\psi_{a_i}(\beta)) = \frac{1}{(\sqrt{2\pi\sigma^2})} \int_\Xi \exp\Big[-\frac{(x-(\beta_0 + a_i\beta_1))^2}{2\sigma^2} \Big] dx \tag{15.17}
$$
$$
(\forall \Xi \in \mathcal{B}_{\mathbb{R}}, \forall \beta = (\beta_0, \beta_1) \in \Omega_0(\equiv \mathbb{R}^2))
$$

Hence, we have the simultaneous observable $\times_{i=1}^n \mathsf{O}_0^{a_i} {\equiv} (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}, \times_{i=1}^n \Psi_{a_i} G_\sigma)$ in $L^\infty(\Omega_0(\equiv \mathbb{R}^2))$ such that

$$
\begin{aligned}
& [(\mathop{\times}_{i=1}^n \Psi_{a_i} G_\sigma)(\mathop{\times}_{i=1}^n \Xi_i)](\beta) = \mathop{\times}_{i=1}^n \Big([(\Psi_{a_i} G_\sigma)(\Xi_i)](\beta) \Big) \\
= & \frac{1}{(\sqrt{2\pi\sigma^2})^n} \int \cdots \int_{\times_{i=1}^n \Xi_i} \exp\Big[-\frac{\sum_{i=1}^n (x_i - (\beta_0 + a_i\beta_1))^2}{2\sigma^2} \Big] dx_1 \cdots dx_n \\
= & \int \cdots \int_{\times_{i=1}^n \Xi_i} p_{(\beta_0, \beta_1, \sigma)}(x_1, x_2, \cdots, x_n) dx_1 \cdots dx_n
\end{aligned} \tag{15.18}
$$
$$
(\forall \mathop{\times}_{i=1}^n \Xi_i \in \mathcal{B}_{\mathbb{R}^n}, \forall \beta = (\beta_0, \beta_1) \in \Omega_0(\equiv \mathbb{R}^2))
$$

Assuming that σ is variable, we have the observable $\mathsf{O} = \Big(\mathbb{R}^n(=X), \mathcal{B}_{\mathbb{R}^n}(=\mathcal{F}), F\Big)$ in $L^\infty(\Omega_0 \times \mathbb{R}_+)$ such that

$$
[F(\mathop{\times}_{i=1}^n \Xi_i)](\beta, \sigma) = [(\mathop{\times}_{i=1}^n \Psi_{a_i} G_\sigma)(\mathop{\times}_{i=1}^n \Xi_i)](\beta) \quad (\forall \Xi_i \in \mathcal{B}_{\mathbb{R}}, \forall (\beta, \sigma) \in \mathbb{R}^2(\equiv \Omega_0) \times \mathbb{R}_+) \tag{15.19}
$$

Problem 15.3. [Regression analysis in quantum language]

Assume that a measured value $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in X = \mathbb{R}^n$ is obtained by the measurement $\mathsf{M}_{L^\infty(\Omega_0 \times \mathbb{R}_+)}(\mathsf{O} \equiv (X, \mathcal{F}, F), S_{[(\beta_0, \beta_1, \sigma)]})$. (The measured value is also called a *response variable*.) And assume that we do not know the state $(\beta_0, \beta_1, \sigma^2)$.
Then,

- from the measured value $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, infer the β_0, β_1, σ!

That is, represent the $(\beta_0, \beta_1, \sigma)$ by $(\hat{\beta}_0(x), \hat{\beta}_1(x), \hat{\sigma}(x))$ (i.e., the functions of x).

Answer.

Taking partial derivatives with respect to β_0, β_1, σ^2, and equating the results to zero, gives the log-likelihood equations. That is, putting

$$L(\beta_0, \beta_1, \sigma^2, x_1, x_2, \cdots, x_n) = \log \Big(p_{(\beta_0, \beta_1, \sigma)}(x_1, x_2, \cdots, x_n) \Big),$$

(where "log" is not essential), we see that

$$\frac{\partial L}{\partial \beta_0} = 0 \quad \Longrightarrow \quad \sum_{i=1}^n (x_i - (\beta_0 + a_i \beta_1)) = 0 \tag{15.20}$$

$$\frac{\partial L}{\partial \beta_1} = 0 \quad \Longrightarrow \quad \sum_{i=1}^n a_i (x_i - (\beta_0 + a_i \beta_1)) = 0 \tag{15.21}$$

$$\frac{\partial L}{\partial \sigma^2} = 0 \quad \Longrightarrow -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n (x_i - \beta_0 - \beta_1 a_i)^2 = 0 \tag{15.22}$$

Therefore, using the notations (15.7)-(15.9), we obtain that

$$\hat{\beta}_0(x) = \overline{x} - \hat{\beta}_1(x)\overline{a} = \overline{x} - \frac{s_{ax}}{s_{aa}}\overline{a}, \quad \hat{\beta}_1(x) = \frac{s_{ax}}{s_{aa}} \tag{15.23}$$

and

$$\begin{aligned} (\hat{\sigma}(x))^2 = & \frac{\sum_{i=1}^n \Big(x_i - (\hat{\beta}_0(x) + a_i \hat{\beta}_1(x)) \Big)^2}{n} \\ = & \frac{\sum_{i=1}^n \Big(x_i - (\overline{x} - \frac{s_{ax}}{s_{aa}}\overline{a}) - a_i \frac{s_{ax}}{s_{aa}} \Big)^2}{n} = \frac{\sum_{i=1}^n \Big((x_i - \overline{x}) + (\overline{a} - a_i) \frac{s_{ax}}{s_{aa}} \Big)^2}{n} \\ = & s_{xx} - 2 s_{ax} \frac{s_{ax}}{s_{aa}} + s_{aa} (\frac{s_{ax}}{s_{aa}})^2 = s_{xx} - \frac{s_{ax}^2}{s_{aa}} \end{aligned} \tag{15.24}$$

Note that the above (15.23) and (15.24) are the same as (15.6). Therefore, Problem 15.3 (i.e., regression analysis in quantum language) is a quantum linguistic story of the least squares method (Problem 15.1).

Remark 15.4. Again, note that

(A) the least squares method (15.6) and the regression analysis (15.23) and (15.24) are the same.

Therefore, a small mathematical technique (the least squares method) can be understood in a grand story (regression analysis in quantum language). The readers may think that

(B) **Why do we choose "complicated (Problem 15.3)" rather than "simple (Problem 15.1)"?**

Of course, such a reason is unnecessary for quantum language! That is because

(C) **the spirit of quantum language says that**

"Everything should be described by quantum language"

However, this may not be a kind answer. The reason is that the grand story has a merit such that statistical methods (i.e., the confidence interval method and the statistical hypothesis testing) can be applicable. This will be mentioned in the following section.

15.3 Regression analysis(distribution , confidence interval and statistical hypothesis testing)

As mentioned in Problem 15.3 (regression analysis), consider the measurement $\mathsf{M}_{L^\infty(\Omega_0 \times \mathbb{R}_+)}(\mathsf{O} \equiv (X(=\mathbb{R}^n), \mathcal{F}, F), \ S_{[(\beta_0,\beta_1,\sigma)]})$

For each $(\beta,\sigma) \in \mathbb{R}^2 \times \mathbb{R}_+$, define the sample probability space $(X, \mathcal{F}, P_{(\beta,\sigma)})$, where

$$P_{(\beta,\sigma)}(\Xi) = [F(\Xi)](\beta_0, \beta_1, \sigma) \qquad (\forall \Xi \in \mathcal{F})$$

Define $L^2(X, P_{(\beta,\sigma)})$ (or in short, $L^2(X)$) by

$$L^2(X) = \{\text{measurable function } f : X \to \mathbb{R} \ \mid \ [\int_X |f(x)|^2 P_{(\beta,\sigma)}(dx)]^{1/2} < \infty\}. \tag{15.25}$$

Further, for each $f, g \in L^2(X)$, define $E(f)$ and $V(f)$ such that

$$E(f) = \int_X f(x) P_{(\beta,\sigma)}(dx), \quad V(f) = \int_X |f(x) - E(f)|^2 P_{(\beta,\sigma)}(dx). \tag{15.26}$$

Our main assertion is to mention Problem 15.3 (i.e., regression analysis in quantum language). This section should be regarded as an easy consequence of Problem 15.3 (regression analysis). For the detailed proof of Lemma 15.5, see standard books of statistics (e.g., ref. [8]).

Lemma 15.5. Consider the measurement $\mathsf{M}_{L^\infty(\Omega_0 \times \mathbb{R}_+)}(\mathsf{O} \equiv (X, \mathcal{F}, F), S_{[(\beta_0,\beta_1,\sigma)]})$ in Problem 15.3 (regression analysis). And assume the above notations. Then, we see:

(A$_1$) (1): $V(\hat{\beta}_0) = \frac{\sigma^2}{n}(1 + \frac{\overline{a}^2}{s_{aa}})$, (2): $V(\hat{\beta}_1) = \frac{\sigma^2}{n}\frac{1}{s_{aa}}$,

(A$_2$) [Studentization]. Motivated by the (A$_1$), we see:

$$T_{\beta_0} := \frac{\sqrt{n}(\hat{\beta}_0 - \beta_0)}{\sqrt{\hat{\sigma}^2(1 + \overline{a}^2/s_{aa})}} \sim t_{n-2}, \qquad T_{\beta_1} := \frac{\sqrt{n}(\hat{\beta}_1 - \beta_1)}{\sqrt{\hat{\sigma}^2/s_{aa}}} \sim t_{n-2} \tag{15.27}$$

where t_{n-2} is the student's distribution with $n-2$ degrees of freedom.

For the proof. see ref. [8].

Let $\mathsf{M}_{L^\infty(\Omega_0(=\mathbb{R}^2) \times \mathbb{R}_+)}(\mathsf{O} \equiv (X(=\mathbb{R}^n), \mathcal{F}, F), S_{[(\beta_0,\beta_1,\sigma)]})$ be the measurement in Problem 15.3 (regression analysis). For each $k = 0, 1$, define the estimator $\widehat{E}_k : X(=\mathbb{R}^n) \to \Theta_k(=\mathbb{R})$ and the quantity $\pi_k : \Omega(=\mathbb{R}^2 \times \mathbb{R}_+) \to \Theta_k(=\mathbb{R})$ as follows.

$$\widehat{E}_0(x)(=\hat{\beta}_0(x)) = \overline{x} - \frac{s_{ax}}{s_{aa}}\overline{a}, \quad \widehat{E}_1(x)(=\hat{\beta}_1(x)) = \frac{s_{ax}}{s_{aa}}, \quad \pi_0(\beta_0, \beta_1, \sigma) = \beta_0. \quad \pi_1(\beta_0, \beta_1, \sigma) = \beta_1, \tag{15.28}$$

$$(\forall (\beta_0, \beta_1, \sigma) \in \mathbb{R}^2 \times \mathbb{R}_+)$$

Let α be a real number such that $0 < \alpha \ll 1$, for example, $\alpha = 0.05$. For any state $\omega = (\beta, \sigma)(\in \Omega = \mathbb{R}^2 \times \mathbb{R}_+)$, define the positive number $\eta^\alpha_{\omega,k}$ (> 0) by (6.9), (6.15), that is,

$$\eta^\alpha_{\omega,k}(=\delta^{1-\alpha}_{\omega,k}) = \inf\{\eta > 0 : [F(\{x \in X \ : \ d^x_{\Theta_k}(\widehat{E}_k(x), \pi_k(\omega)) \geq \eta\})](\omega) \leq \alpha\} \tag{15.29}$$

where, for each $\theta^0_k, \theta^1_k (\in \Theta_k)$, the semi-distance $d^x_{\Theta_k}$ in Θ_k is defined by

$$d^x_{\Theta_k}(\theta^0_k, \theta^1_k) = \begin{cases} \frac{\sqrt{n}|\theta^0_0 - \theta^1_0|}{\sqrt{\hat{\sigma}^2(1+\overline{a}^2/s_{aa})}} & (\text{if } k=0) \\ \\ \frac{\sqrt{n}|\theta^0_1 - \theta^1_1|}{\sqrt{\hat{\sigma}^2/s_{aa}}} & (\text{if } k=1) \end{cases} \tag{15.30}$$

Therefore, we see, by Lemma 15.5, that

$$\eta^\alpha_{\omega,k} = \begin{cases} \inf\{\eta > 0 : [F(\{x \in X \ : \ \frac{\sqrt{n}|\hat{\beta}_0(x) - \beta_0|}{\sqrt{\hat{\sigma}^2(1+\overline{a}^2/s_{aa})}} \geq \eta\})](\omega) \leq \alpha\} & (\text{if } k=0) \\ \\ \inf\{\eta > 0 : [F(\{x \in X \ : \ \frac{\sqrt{n}|\hat{\beta}_1(x) - \beta_1|}{\sqrt{\hat{\sigma}^2(x)/s_{aa}}} \geq \eta\})](\omega) \leq \alpha\} & (\text{if } k=1) \end{cases} \tag{15.31}$$

$$= t_{n-2}(\alpha/2) \tag{15.32}$$

Summing up the above arguments, we have the following proposition:

Proposition 15.6. [confidence interval]. Assume that a measured value $x \in X$ is obtained by the measurement $\mathsf{M}_{L^\infty(\Omega_0 \times \mathbb{R}_+)}(\mathsf{O} \equiv (X, \mathcal{F}, F), S_{[(\beta_0, \beta_1, \sigma)]})$. Here, the state $(\beta_0, \beta_1, \sigma)$ is assumed to be unknown. Then, we have the $(1-\alpha)$-confidence interval $I^{1-\alpha}_{x,k}$ in Corollary 6.6 as follows.

$$I^{1-\alpha}_{x,k} = \{\pi_k(\omega)(\in \Theta_k) : d^x_{\Theta_k}(\widehat{E}_k(x), \pi_k(\omega)) < \eta^{1-\alpha}_{\omega,k}\}$$

$$= \begin{cases} I^{1-\alpha}_{x,0} = \Big\{\beta_0 = \pi_0(\omega)(\in \Theta_0) \ : \ \frac{|\hat{\beta}_0(x) - \beta_0|}{\sqrt{\frac{\hat{\sigma}^2(x)}{n}(1+\overline{a}^2/s_{aa})}} \leq t_{n-2}(\alpha/2)\Big\} & (\text{if } k=0) \\ \\ I^{1-\alpha}_{x,1} = \Big\{\beta_1 = \pi_1(\omega)(\in \Theta_1) : \frac{|\hat{\beta}_1(x) - \beta_1|}{\sqrt{\frac{\hat{\sigma}^2(x)}{n}(1/s_{aa})}} \leq t_{n-2}(\alpha/2)\Big\} & (\text{if } k=1) \end{cases} \tag{15.33}$$

Proposition 15.7. [Statistical hypothesis testing]. [Hypothesis test]. Consider the measurement $\mathsf{M}_{L^\infty(\Omega_0 \times \mathbb{R}_+)}(\mathsf{O} \equiv (X, \mathcal{F}, F), S_{[(\beta_0, \beta_1, \sigma)]})$. Here, the state $(\beta_0, \beta_1, \sigma)$ is assumed to be unknown. Then, according to Corollary 6.6, we say:

(B$_1$) Assume the null hypothesis $H_N = \{\beta_0\}(\subseteq \Theta_0 = \mathbb{R})$. Then, the rejection region is as follows:

$$
\begin{aligned}
\widehat{R}_{H_N}^{\alpha;X} &= \widehat{E}_0^{-1}(\widehat{R}_{H_N}^{\alpha;\Theta_0}) = \bigcap_{\omega \in \Omega \text{ such that } \pi_0(\omega) \in H_N} \{x (\in X) : d_{\Theta_0}^x(\widehat{E}_0(x), \pi_0(\omega)) \geq \eta_\omega^\alpha\} \\
&= \Big\{ x \in X \ : \ \frac{|\hat{\beta}_0(x) - \beta_0|}{\sqrt{\frac{\hat{\sigma}^2(x)}{n}(1 + \overline{a}^2/s_{aa})}} \geq t_{n-2}(\alpha/2) \Big\}
\end{aligned}
\tag{15.34}
$$

(B$_2$) Assume the null hypothesis $H_N = \{\beta_1\}(\subseteq \Theta_1 = \mathbb{R})$. Then, the rejection region is as follows:

$$
\begin{aligned}
\widehat{R}_{H_N}^{\alpha;X} &= \widehat{E}_1^{-1}(\widehat{R}_{H_N}^{\alpha;\Theta_1}) = \bigcap_{\omega \in \Omega \text{ such that } \pi_1(\omega) \in H_N} \{x (\in X) : d_{\Theta_1}^x(\widehat{E}_1(x), \pi_1(\omega)) \geq \eta_\omega^\alpha\} \\
&= \Big\{ x \in X \ : \ \frac{|\hat{\beta}_1(x) - \beta_1|}{\sqrt{\frac{\hat{\sigma}^2(x)}{n}(1/s_{aa})}} \geq t_{n-2}(\alpha/2) \Big\}
\end{aligned}
\tag{15.35}
$$

15.4 Generalized linear model

Put $T = \{0, 1, 2, \cdots, i, \cdots, n\}$, which is the same as the tree (15.10), that is,

$$\tau(i) = 0 \qquad (\forall i = 1, 2, \cdots, n) \tag{15.36}$$

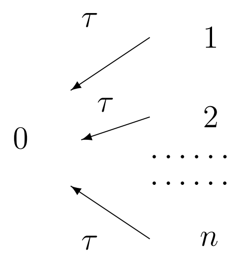


Figure 15.3: Parallel structure

For each $i \in T$, define a locally compact space Ω_i such that

$$\Omega_0 = \mathbb{R}^{m+1} = \left\{ \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_m \end{bmatrix} \; : \; \beta_0, \beta_1, \cdots, \beta_m \in \mathbb{R} \right\} \tag{15.37}$$

$$\Omega_i = \mathbb{R} = \left\{ \mu_i \; : \; \mu_i \in \mathbb{R} \right\} \quad (i = 1, 2, \cdots, n) \tag{15.38}$$

Assume that

$$a_{ij} \in \mathbb{R} \qquad (i = 1, 2, \cdots, n, \;\; j = 1, 2, \cdots, m, (m+1 \le n)) \tag{15.39}$$

which are called *explanatory variables* in the conventional statistics. Consider the deterministic causal map $\psi_{a_{i\bullet}} : \Omega_0 (= \mathbb{R}^{m+1}) \to \Omega_i (= \mathbb{R})$ such that

$$\Omega_0 = \mathbb{R}^{m+1} \ni \beta = (\beta_0, \beta_1, \cdots, \beta_m) \mapsto \psi_{a_{i\bullet}}(\beta_0, \beta_1, \cdots, \beta_m) = \beta_0 + \sum_{j=1}^{m} \beta_j a_{ij} = \mu_i \in \Omega_i = \mathbb{R} \tag{15.40}$$

$$(i = 1, 2, \cdots, n)$$

Summing up, we see

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix} \mapsto \begin{bmatrix} \psi_{a_{1\bullet}}(\beta_0, \beta_1, \cdots, \beta_m) \\ \psi_{a_{2\bullet}}(\beta_0, \beta_1, \cdots, \beta_m) \\ \psi_{a_{3\bullet}}(\beta_0, \beta_1, \cdots, \beta_m) \\ \vdots \\ \psi_{a_{n\bullet}}(\beta_0, \beta_1, \cdots, \beta_m) \end{bmatrix} = \begin{bmatrix} 1 & a_{11} & a_{12} & \cdots & a_{1m} \\ 1 & a_{21} & a_{22} & \cdots & a_{2m} \\ 1 & a_{31} & a_{32} & \cdots & a_{3m} \\ 1 & a_{41} & a_{42} & \cdots & a_{4m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} \cdot \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix} \tag{15.41}$$

which is equivalent to the deterministic Markov operator $\Psi_{a_{i\bullet}} : L^\infty(\Omega_i) \to L^\infty(\Omega_0)$ such that

$$[\Psi_{a_{i\bullet}}(f_i)](\omega_0) = f_i(\psi_{a_{i\bullet}}(\omega_0)) \quad (\forall f_i \in L^\infty(\Omega_i), \ \ \forall \omega_0 \in \Omega_0, \forall i \in 1, 2, \cdots, n) \tag{15.42}$$

Thus, under the identification: $a_{ij} \Leftrightarrow \Psi_{a_{i\bullet}}$, the term "explanatory variable" means a kind of causality.

$$\begin{array}{ccc} & \overset{\Psi_{a_{1\bullet}}}{\swarrow} & L^\infty(\Omega_1(\equiv \mathbb{R})) \\ & \overset{\Psi_{a_{2\bullet}}}{\swarrow} & L^\infty(\Omega_2(\equiv \mathbb{R})) \\ L^\infty(\Omega_0(\equiv \mathbb{R}^{m+1})) & \cdots\cdots & \\ & \cdots\cdots & \\ & \nwarrow & \\ & \Psi_{a_{n\bullet}} & L^\infty(\Omega_n(\equiv \mathbb{R})) \end{array}$$

Figure 15.4: Parallel structure(Causal relation $\Psi_{a_{i\bullet}}$)

Therefore, we have the observable $\mathsf{O}_0^{a_{i\bullet}} {\equiv} (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, \Psi_{a_{i\bullet}} G_\sigma)$ in $L^\infty(\Omega_0(\equiv \mathbb{R}^{m+1}))$ such that

$$[\Psi_{a_{i\bullet}}(G_\sigma(\Xi))](\beta) = [(G_\sigma(\Xi))](\psi_{a_{i\bullet}}(\beta)) = \frac{1}{(\sqrt{2\pi\sigma^2})} \int_\Xi \exp\Big[- \frac{(x - (\beta_0 + \sum_{j=1}^m a_{ij}\beta_j))^2}{2\sigma^2} \Big] dx \tag{15.43}$$

$$(\forall \Xi \in \mathcal{B}_{\mathbb{R}}, \forall \beta = (\beta_0, \beta_1, \cdots, \beta_m) \in \Omega_0(\equiv \mathbb{R}^{m+1}))$$

Hence, we have the simultaneous observable $\times_{i=1}^n \mathsf{O}_0^{a_{i\bullet}} {\equiv} (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}, \times_{i=1}^n \Psi_{a_{i\bullet}} G_\sigma)$ in $L^\infty(\Omega_0(\equiv \mathbb{R}^{m+1}))$ such that

$$\begin{aligned} &[(\mathop{\times}_{i=1}^n \Psi_{a_{i\bullet}} G_\sigma)(\mathop{\times}_{i=1}^n \Xi_i)](\beta) = \mathop{\times}_{i=1}^n \Big([(\Psi_{a_{i\bullet}} G_\sigma)(\Xi_i)](\beta) \Big) \\ =& \frac{1}{(\sqrt{2\pi\sigma^2})^n} \int \cdots \int_{\times_{i=1}^n \Xi_i} \exp\Big[- \frac{\sum_{i=1}^n (x_i - (\beta_0 + \sum_{j=1}^m a_{ij}\beta_j))^2}{2\sigma^2} \Big] dx_1 \cdots dx_n \\ & (\forall \mathop{\times}_{i=1}^n \Xi_i \in \mathcal{B}_{\mathbb{R}^n}, \forall \beta = (\beta_0, \beta_1, \cdots, \beta_m) \in \Omega_0(\equiv \mathbb{R}^{m+1})) \end{aligned} \tag{15.44}$$

Assuming that σ is variable, we have the observable $\mathsf{O} = \Big(\mathbb{R}^n(= X), \mathcal{B}_{\mathbb{R}^n}(= \mathcal{F}), F \Big)$ in $L^\infty(\Omega_0 \times \mathbb{R}_+)$ such that

$$[F(\mathop{\times}_{i=1}^n \Xi_i)](\beta, \sigma) = [(\mathop{\times}_{i=1}^n \Psi_{a_{i\bullet}} G_\sigma)(\mathop{\times}_{i=1}^n \Xi_i)](\beta) \quad (\forall \mathop{\times}_{i=1}^n \Xi_i \in \mathcal{B}_{\mathbb{R}^n}, \forall (\beta, \sigma) \in \mathbb{R}^{m+1}(\equiv \Omega_0) \times \mathbb{R}_+) \tag{15.45}$$

Thus, we have the following problem.

Problem 15.8. [Generalized linear model in quantum language]

Assume that a measured value $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in X = \mathbb{R}^n$ is obtained by the measurement $\mathsf{M}_{L^\infty(\Omega_0 \times \mathbb{R}_+)}(\mathsf{O} \equiv (X, \mathcal{F}, F), S_{[(\beta_0, \beta_1, \cdots, \beta_m, \sigma)]})$. (The measured value is also called a *response variable*.) And assume that we do not know the state $(\beta_0, \beta_1, \cdots, \beta_m, \sigma^2)$.
Then,

- from the measured value $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, infer the $\beta_0, \beta_1, \cdots, \beta_m, \sigma$!

That is, represent the $(\beta_0, \beta_1, \cdots, \beta_m, \sigma)$ by $(\hat{\beta}_0(x), \hat{\beta}_1(x), \cdots, \beta_m(x), \hat{\sigma}(x))$ (i.e., the functions of x).

The answer is easy, since it is a slight generalization of Problem 15.3. Also, it suffices to follow ref. [8]. However, note that the purpose of this chapter is to propose Problem 15.8 (i.e, the quantum linguistic formulation of the generalized linear model) and not to give the answer to Problem 15.8.

Remark 15.9. As a generalization of regression analysis, we also see measurement error model (*cf.* §5.5 (117 page) in ref. [29]), That is, we have two different generalizations such as

$$\boxed{\text{Regression analysis}} \xrightarrow[\text{generalization}]{} \begin{cases} ① : \boxed{\text{generalized linear model}} \\ ② : \boxed{\text{measurement error model}} \end{cases} \tag{15.46}$$

However, we believe that the ① is the main street.

Chapter 16

Kalman filter (calculation)

The Kalman filter [52, 56] is located as in the following ($\sharp$):

$$
(\sharp) : \text{Statistics} \begin{cases} \text{Fisher's maximum likelihood method} & \xrightarrow[\text{usually deterministic}]{+\text{ causality}} \text{regression analysis} \\ \\ \text{Bayes' method} & \xrightarrow[\text{non-deterministic}]{+\text{ causality}} \text{Kalman filter} \end{cases}
$$

Thus, I can not emphasize too much the importance of the Kalman filter. Though Kalman filter belongs to Bayes' statistics, this fact may not be a common sense. This present state is due to the confusion between Fisher's statistics and Bayes' statistics. I hope that such confusion should be clarified by the above ($\sharp$) (based on quantum language). This chapter is extracted from the following paper:

- S. Ishikawa, K. Kikuchi: *Kalman filter in quantum language,* arXiv:1404.2664 [math.ST] 2014.

16.1 Bayes=Kalman method (in $L^\infty(\Omega, m)$)

Recall Theorem 9.11(Bayes' theorem), particularly, the Bayes operator (9.5). This will be generalized as Bayes=Kalman operator as follows.

Let t_0 be the root of a tree T. For each $t \in T$, consider the classical basic structure:

$$
[C_0(\Omega_t) \subseteq L^\infty(\Omega_t, m_t) \subseteq B(L^2(\Omega_t, m_t))]
$$

Let $[\mathbb{O}_T] = [\{\mathsf{O}_t(\equiv (X_t, \mathcal{F}_t, F_t))\}_{t\in T}, \{\Phi^{t_1,t_2} : L^\infty(\Omega_{t_2}) \to L^\infty(\Omega_{t_1})\}_{(t_1,t_2)\in T^2_{\leqq}}]$ be a sequential causal observable with the realization $\widehat{\mathsf{O}}_{t_0} \equiv (\times_{t\in T} X_t, \ \boxtimes_{t\in T}\mathcal{G}_t, \ \widehat{F}_{t_0})$ in $L^\infty(\Omega_{t_0})$.

For example,

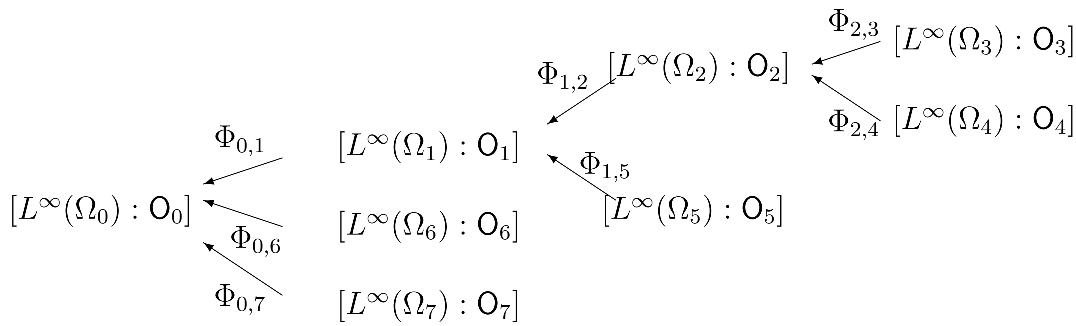


Figure 16.1 : Simple classical example of sequential causal observable

For each $t \in T$, consider another observable $\mathsf{O}_t' = (Y_t, \mathcal{G}_t, G_t)$ in $L^\infty(\Omega_t, m_t)$, and the simultaneous observable $\mathsf{O} \times \mathsf{O}_t' = (X_t \times Y_t, \mathcal{F}_t \boxtimes \mathcal{G}_t, F_t \times G_t)$ in $L^\infty(\Omega_t, m_t)$. And let $[\mathbb{O}_T^\times] = [\{\mathsf{O}_t^\times (\equiv (X_t \times Y_t, \mathcal{F}_t \boxtimes \mathcal{G}_t, F_t \times G_t))\}_{t \in T}, \{\Phi^{t_1,t_2} : L^\infty(\Omega_{t_2}) \to L^\infty(\Omega_{t_1})\}_{(t_1,t_2) \in T^2_{\leq}}\ ]$ be a sequential causal observable with the realization $\widehat{\mathsf{O}}_{t_0}^\times \equiv (\times_{t \in T}(X_t \times Y_t),\ \boxtimes_{t \in T}(\mathcal{F}_t \boxtimes \mathcal{G}_t),\ \widehat{H}_{t_0})$ in $L^\infty(\Omega_{t_0})$.

For example,

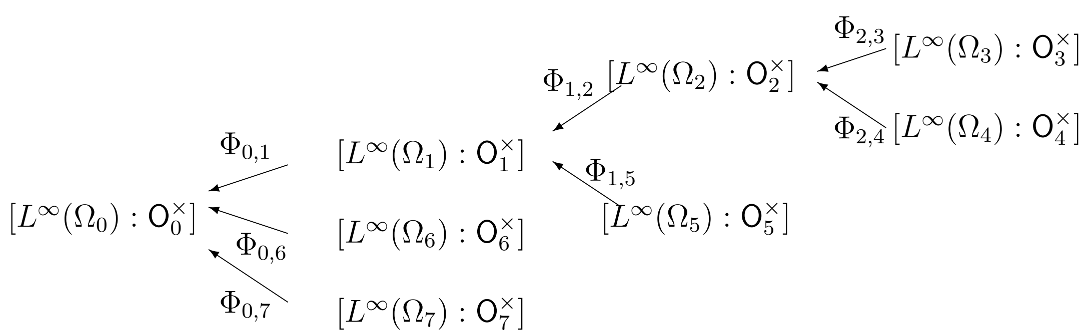


Figure 16.2 : Simple classical example of sequential causal observable

Thus we have the mixed measurement $\mathsf{M}_{L^\infty(\Omega_{t_0})}(\widehat{\mathsf{O}}_{t_0}^\times, S_{[*]}(z_0))$, where $z_0 \in L^1_{+1}(\Omega_{t_0})$. Assume that we know that the measured value $(x, y)\ (= ((x_t)_{t \in T}, (y_t)_{t \in T},) \in (\times_{t \in T} X_t) \times (\times_{t \in T} Y_t))$ obtained by the measurement $\mathsf{M}_{L^\infty(\Omega_{t_0})}(\widehat{\mathsf{O}}_{t_0}^\times, S_{[*]}(z_0))$ belongs to $(\times_{t \in T} \Xi_t) \times\ (\times_{t \in T} Y_t)\ \ (\in (\boxtimes_{t \in T}\mathcal{F}_t) \boxtimes (\boxtimes_{t \in T}\mathcal{G}_t))$. Then, by Axiom$^{(\mathrm{m})}$ 1(§9.1), we can infer that

(A) the probability $P_{\times_{t \in T}\Xi_t}((G_t(\Gamma_t))_{t \in T})$ that y belongs to $\times_{t \in T} \Gamma_t (\in \boxtimes_{t \in T}\mathcal{G}_t)$ is given by

$$
\begin{aligned}
& P_{\times_{t \in T}\Xi_t}((G_t(\Gamma_t))_{t \in T}) \\
= & \frac{\int_{\Omega_0} [\widehat{H}_{t_0}((\times_{t \in T} \Xi_t) \times (\times_{t \in T} \Gamma_t))](\omega_0)\ z_0(\omega_0)\ m_0(d\omega_0)}{\int_{\Omega_0} [\widehat{H}_{t_0}(\times_{t \in T} \Xi_t) \times (\times_{t \in T} Y_t)](\omega_0)\ z_0(\omega_0)\ m_0(d\omega_0)} \\
& (\forall \Gamma_t \in \mathcal{G}_t, t \in T).
\end{aligned}
\tag{16.1}
$$

Let $s \in T$ be fixed. Assume that

$$\Gamma_t = Y_t \quad (\forall t \in T \text{ such that } t \neq s)$$

Thus, putting $\widehat{P}_{\times_{t\in T}\Xi_t}(G_s(\Gamma_s)) = P_{\times_{t\in T}\Xi_t}((G_t(\Gamma_t))_{t\in T})$, we see that $\widehat{P}_{\times_{t\in T}\Xi_t} \in L^1_{+1}(\Omega_s, m_s)$. That is, there uniquely exists $z_s^a \in L^1_{+1}(\Omega_s, m_s)$ such that

$$\widehat{P}_{\times_{t\in T}\Xi_t}((G_s(\Gamma_s)) = {}_{L^1(\Omega_s)}\langle z_s^a, G_s(\Gamma_s)\rangle_{L^\infty(\Omega_s)} = \int_{\Omega_s} [G_s(\Gamma_s)](\omega_s) z_s^a(\omega_s) m_s(d\omega_s)$$

for any observable $(Y_s, \mathcal{G}_s, G_s)$ in $L^\infty(\Omega_s)$. That is because the linear functional $\widehat{P}_{\times_{t\in T}\Xi_t} : L^\infty(\Omega_s) \to \mathbb{C}$ (complex numbers) is weak* continuous. After all,

(B) we can define the **Bayes-Kalman operator** $[B^s_{\widehat{\mathsf{O}}_{t_0}}(\times_{t\in T}\Xi_t)] : L^1_{+1}(\Omega_{t_0}) \to L^1_{+1}(\Omega_s)$ such that

$$\underset{(\in L^1_{+1}(\Omega_{t_0}))}{\overset{\text{(pretest state)}}{\boxed{z_0}}} \xrightarrow[\text{Bayes-Kalman operator}]{[B^s_{\widehat{\mathsf{O}}_{t_0}}(\times_{t\in T}\Xi_t)]} \underset{(\in L^1_{+1}(\Omega_s))}{\overset{\text{(posttest state)}}{\boxed{z_s^a}}} \tag{16.2}$$

which is the generalization of the Bayes operator (9.5).

Remark 16.1. We have frequently discussed the Bayes=Kalman filter, for example, in [29, 32]. However, these arguments are too theoretical. In this chapter, we devote ourselves to the numerical aspect of the Kalman filter.

16.2 Problem establishment (concrete calculation)

In the previous section, we study the general theory of Kalman filter. In this section, we devote ourselves to the calculation of Kalman filter in the case of a linear ordered tree $T = \{0, 1, 2, \cdots, n\}$ such that the parent map $\pi : T \setminus \{0\} \to T$ is defined by $\pi(k) = k - 1$:

$$0 \xleftarrow{\pi} 1 \xleftarrow{\pi} 2 \xleftarrow{\pi} \cdots \xleftarrow{\pi} n-1 \xleftarrow{\pi} n$$

Figure 16.3: Linear ordered tree

For each $k \in T$, consider the classical basic structure:

$$[C_0(\Omega_k) \subseteq L^\infty(\Omega_k, m_k) \subseteq B(L^\infty(\Omega_k, m_k))] \Big(= [C_0(\mathbb{R}) \subseteq L^\infty(\mathbb{R}, d\omega) \subseteq B(L^2(\mathbb{R}, d\omega))] \Big)$$

where $d\omega$ is the Lebesgue measure on $\mathbb{R}$.

Consider the sequential causal observable $[\mathbb{O}_T] = [\{\mathsf{O}_t\}_{t \in T}, \ \{\Phi^{t-1,t} : L^\infty(\Omega_t) \to L^\infty(\Omega_{t-1})\}_{T=1,2,\cdots,n}]$, and assume the initial state $z_0 \in L^1_{+1}(\Omega_0, m_0)$.
Thus, we have the following situation:

initial state z_0

$$\underset{\mathsf{O}_0=(X_0,\mathcal{F}_0 F_0)}{\boxed{L^\infty(\Omega_0, m_0)}} \xleftarrow{\Phi^{0,1}} \underset{\mathsf{O}_1=(X_1,\mathcal{F}_1 F_1)}{\boxed{L^\infty(\Omega_1, m_1)}} \xleftarrow{\Phi^{1,2}} \cdots \xleftarrow{\Phi^{s-1,s}} \underset{\mathsf{O}_s=(X_s,\mathcal{F}_s F_s)}{\boxed{L^\infty(\Omega_s, m_s)}} \xleftarrow{\Phi^{s,s+1}} \cdots \xleftarrow{\Phi^{n-1,n}} \underset{\mathsf{O}_n=(X_n,\mathcal{F}_n F_n)}{\boxed{L^\infty(\Omega_n, m_n)}}$$

or, equivalently,

initial state z_0

$$\underset{\mathsf{O}_0=(X_0,\mathcal{F}_0,F_0)}{\boxed{L^1(\Omega_0, m_0)}} \xrightarrow{\Phi_*^{0,1}} \underset{\mathsf{O}_1=(X_1,\mathcal{F}_1,F_1)}{\boxed{L^1(\Omega_1, m_1)}} \xrightarrow{\Phi_*^{1,2}} \cdots \xrightarrow{\Phi_*^{s-1,s}} \underset{\mathsf{O}_s=(X_s,\mathcal{F}_s,F_s)}{\boxed{L^1(\Omega_s, m_s)}} \xrightarrow{\Phi_*^{s,s+1}} \cdots \xrightarrow{\Phi_*^{n-1,n}} \underset{\mathsf{O}_n=(X_n,\mathcal{F}_n,F_n)}{\boxed{L^1(\Omega_n, m_n)}}$$

In the above, the **initial state** $z_0 (\in L^1_{+1}(\Omega_0, m_0))$ is defined by

$$z_0(\omega_0) = \frac{1}{\sqrt{2\pi}\sigma_0} \exp[-\frac{(\omega_0 - \mu_0)^2}{2\sigma_0^2}] \qquad (\forall \omega_0 \in \Omega_0) \tag{16.3}$$

where it is assumed that μ_0 and σ_0 are known.

Also, for each $t \in T = \{0, 1, \cdots, n\}$, consider the **observable** $\mathsf{O}_t = (X_t, \mathcal{F}_t, F_t) = (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, F_t)$ in $L^\infty(\Omega_t, m_t)$ such that

$$[F_t(\Xi_t)](\omega_t) = \int_{\Xi_t} \frac{1}{\sqrt{2\pi}q_t} \exp[-\frac{(x_t - c_t\omega_t - d_t)^2}{2q_t^2}] dx_t \equiv \int_{\Xi_t} f_{x_t}(\omega_t) dx_t \quad (\forall \Xi_t \in \mathcal{F}_t, \ \ \forall \omega_t \in \Omega_t) \tag{16.4}$$

where it is assumed that c_t, d_t and q_t are known $(t \in T)$.

And further, the **causal operator** $\Phi^{t-1.t} : L^\infty(\Omega_t) \to L^\infty(\Omega_{t-1})$ is defined by

$$[\Phi^{t-1,t} \widetilde{f}_{x_t}](\omega_{t-1}) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}r_t} \exp[-\frac{(\omega_t - a_t\omega_{t-1} - b_t)^2}{2r_t^2}] \widetilde{f}_{x_t}) d\omega_t \equiv f_{t-1}(\omega_{t-1}) \tag{16.5}$$

$$(\forall \widetilde{f}_{x_t} \in L^\infty(\Omega_t, m_t), \quad \forall \omega_{t-1} \in \Omega_{t-1})$$

where it is assumed that a_t, b_t and r_t are known $(t \in T)$.
Or, equivalently, the **pre-dual causal operator** $\Phi_*^{t-1.t} : L^1_{+1}(\Omega_{t-1}) \to L^1_{+1}(\Omega_t)$ is defined by

$$[\Phi_*^{t-1,t} \widetilde{z}_{t-1}](\omega_t) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi} r_t} \exp[-\frac{(\omega_t - a_t \omega_{t-1} - b_t)^2}{2 r_t^2}] \widetilde{z}_{t-1}(\omega_{t-1}) d\omega_{t-1} \tag{16.6}$$
$$(\forall \widetilde{z}_{t-1} \in L^1_{+1}(\Omega_{t-1}, m_{t-1}), \ \forall \omega_t \in \Omega_t)$$

Now we have the sequential causal observable

$$[\mathbb{O}_T] = [\{\mathsf{O}_t\}_{t\in T}, \{\Phi^{t-1,t} : L^\infty(\Omega_t) \to L^\infty(\Omega_{t-1})\}_{T=1,2,\cdots,n}$$

Let $\widehat{\mathsf{O}}_0$ $(\times_{t=0}^n X_t, \boxtimes_{t=0}^n \mathcal{F}_t, \widehat{F})$ be its realization. Then we have the following problem:

Problem 16.2. [Kalman filter; calculation]

Assume that a measured value $(x_0, x_2, \cdots, x_n)$ $(\in \times_{t=0}^n X_t)$ is obtained by the measurement $\mathsf{M}_{L^\infty(\Omega_0)}$ $(\widehat{\mathsf{O}}_0, S_{[*]}(z_0))$. Let $s (\in T)$ be fixed. Then, calculate the Bayes-Kalman operator $[B^s_{\widehat{\mathsf{O}}_0}(\times_{t\in T}\{x_t\})](z_0)$ in (16.2), where

$$[B^s_{\widehat{\mathsf{O}}_0}(\times_{t\in T}\{x_t\})](z_0) = z_s^a = \lim_{\Xi_t \to x_t \ (t\in T)} [B^s_{\widehat{\mathsf{O}}_0}(\times_{t\in T}\Xi_t)](z_0)$$

That is,

$$L^1_{+1}(\Omega_0) \ni z_0 \xrightarrow[B^s_{\widehat{\mathsf{O}}_0}(\times_{t\in T}\{x_t\})]{\text{measured value:}(x_0,x_1,\ldots,x_n)} z_s^a \in L^1_{+1}(\Omega_s)$$

16.3 Bayes=Kalman operator $B^s_{\widehat{\mathsf{O}}_0}(\times_{t\in T}\{x_t\})$

In what follows, we solve Problem 16.2. For this, it suffices to find the $z_s \in L^1_{+1}(\Omega_s)$ such that

$$\lim_{\Xi_t \to x_t \ (t\in T)} \frac{\int_{\Omega_0}[\widehat{F}_0((\times_{t=0}^n \Xi_t)\times\Gamma_s)](\omega_0) \ \ z_0(\omega_0)d\omega_0}{\int_{\Omega_0}[\widehat{F}_0(\times_{t=0}^n \Xi_t)](\omega_0) \ \ z_0(\omega_0)d\omega_0} = \int_{\Omega_s}[G_s(\Gamma_s)](\omega_s) \ \ z_s(\omega_s)d\omega_s \quad (\forall \Gamma_s \in \mathcal{F}_s)$$

Let us calculate $z_s = [B^s_{\widehat{\mathsf{O}}_0}(\times_{t\in T}\{x_t\})](z_0)$ as follows.

$$\begin{aligned}
&\int_{\Omega_0}[\widehat{F}_0((\mathop{\times}_{t=0}^{n}\Xi_t)\times\Gamma_s)](\omega_0) \ \ z_0(\omega_0)d\omega_0\\
=&{}_{L^1(\Omega_0)}\langle z_0, \widehat{F}_0((\mathop{\times}_{t=0}^{n}\Xi_t)\times\Gamma_s)\rangle_{L^\infty(\Omega_0)}\\
=&{}_{L^1(\Omega_1)}\langle \Phi_*^{0,1}(F_0(\Xi_0)z_0), \widehat{F}_1((\mathop{\times}_{t=1}^{n}\Xi_t)\times\Gamma_s)\rangle_{L^\infty(\Omega_1)}
\end{aligned} \tag{16.7}$$

(A) and, putting $\widetilde{z}_0 = F_0(\Xi_0)z_0$ (or, exactly, its normalization, i.e., $\widetilde{z}_0 = \lim_{\Xi_0\to x_0}\frac{F_0(\Xi_0)z_0}{\int_{\Omega_0}F_0(\Xi_0)z_0 d\omega_0}$) , $\widetilde{z}_1 = F_1(\Xi_1)\Phi_*^{0,1}(\widetilde{z}_0)$, $\widetilde{z}_2 = F_2(\Xi_2)\Phi_*^{1,2}(\widetilde{z}_1)$, $\cdots$, $\widetilde{z}_{s-1} = F_{s-1}(\Xi_{s-1})\Phi_*^{s-2,s-1}(\widetilde{z}_{s-2})$, we see that

$$\begin{aligned}
(16.7) =&{}_{L^1(\Omega_1)}\langle \Phi_*^{0,1}(\widetilde{z}_0), \widehat{F}_1((\mathop{\times}_{t=1}^{n}\Xi_t)\times\Gamma_s)\rangle_{L^\infty(\Omega_1)}\\
=&{}_{L^1(\Omega_2)}\langle \Phi_*^{1,2}(\widetilde{z}_1), \widehat{F}_2((\mathop{\times}_{t=2}^{n}\Xi_t)\times\Gamma_s)\rangle_{L^\infty(\Omega_2)}\\
&\cdots\cdots\\
=&{}_{L^1(\Omega_{s+1})}\langle \Phi_*^{s,s+1}(\widetilde{z}_s), \widehat{F}_{s+1}((\mathop{\times}_{t=s+1}^{n}\Xi_t)\times\Gamma_s)\rangle_{L^\infty(\Omega_{s+1})}\\
=&{}_{L^1(\Omega_s)}\langle \Phi_*^{s-1,s}(\widetilde{z}_{s-1}), \widehat{F}_s((\mathop{\times}_{t=s}^{n}\Xi_t)\times\Gamma_s)\rangle_{L^\infty(\Omega_s)}\\
=&{}_{L^1(\Omega_s)}\langle \Phi_*^{s-1,s}(\widetilde{z}_{s-1}), F_s(\Xi_s)G_s(\Gamma_s)\Phi^{s,s+1}\widehat{F}_{s+1}(\mathop{\times}_{t=s+1}^{n}\Xi_t)\rangle_{L^\infty(\Omega_s)}\\
=&{}_{L^1(\Omega_s)}\langle \Big(F_s(\Xi_s)\Phi^{s,s+1}\widehat{F}_{s+1}(\mathop{\times}_{t=s+1}^{n}\Xi_t)\Big)\Big(\Phi_*^{s-1,s}(\widetilde{z}_{s-1})\Big), G_s(\Gamma_s)\rangle_{L^\infty(\Omega_s)}
\end{aligned} \tag{16.8}$$

Thus, we see

$$[B^s_{\widehat{\mathsf{O}}_0}(\mathop{\times}_{t\in T}\{x_t\})](z_0) = \lim_{\Xi_t\to x_t \ (t\in T)} \frac{\Big(F_s(\Xi_s)\Phi^{s,s+1}\widehat{F}_{s+1}(\times_{t=s+1}^{n}\Xi_t)\Big)\times\Big(\Phi_*^{s-1,s}\widetilde{z}_{s-1})\Big)}{\int_{\Omega_0}[\widehat{F}_0(\times_{t=0}^{n}\Xi_t)](\omega_0) \ \ z_0(\omega_0)d\omega_0} \tag{16.9}$$

16.4 Calculation: prediction part

16.4.1 Calculation: $z_s = \Phi_*^{s-1,s}(\widetilde{z}_{s-1})$ in (16.9)

We prepare the following lemma.

Lemma 16.3. It holds that

(B$_1$) $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}A} \exp[-\frac{(x-By)^2}{2A^2}] \frac{1}{\sqrt{2\pi}C} \exp[-\frac{(y-D)^2}{2C^2}]dy = \frac{1}{\sqrt{2\pi}\sqrt{A^2+B^2C^2}} \exp[-\frac{(x-BD)^2}{2(A^2+B^2C^2)}]$

(B$_2$) $\exp[-\frac{(A\omega-B)^2}{2E^2}] \exp[-\frac{(C\omega-D)^2}{2F^2}] \approx \exp[-\frac{1}{2}(\frac{A^2F^2+C^2E^2}{E^2F^2})\Big(\omega - \frac{(ABF^2+CDE^2)}{(A^2F^2+C^2E^2)}\Big)^2]$

where the notation "$\approx$" means as follows:

"$f(\omega) \approx g(\omega)$" $\Longleftrightarrow$ "there exists a positive K such that $f(\omega) = Kg(\omega)$ $(\forall \omega \in \Omega)$"

Proof. It is easy, thus we omit the proof.

We see, by (16.3) and (A), that

$$\begin{aligned}\widetilde{z}_0(\omega_0) =& \lim_{\Xi_0 \to x_0} \frac{F(\Xi_0)z_0}{\int_{\mathbb{R}} F(\Xi_0)z_0 d\omega_0} \\ \approx& \frac{1}{\sqrt{2\pi}q_0} \exp[-\frac{(x_0 - c_0\omega_0 - d_0)^2}{2q_0^2}] \frac{1}{\sqrt{2\pi}\sigma_0} \exp[-\frac{(\omega_0-\mu_0)^2}{2\sigma_0^2}] \\ \approx& \frac{1}{\sqrt{2\pi}\widetilde{\sigma}_0} \exp[-\frac{(\omega_0 - \widetilde{\mu}_0)^2}{2\widetilde{\sigma}_0^2}] \end{aligned} \tag{16.10}$$

where

$$\widetilde{\sigma}_0^2 = \frac{q_0^2\sigma_0^2}{q_0^2 + c_0^2\sigma_0^2}, \quad \widetilde{\mu}_0 = \mu_0 + \widetilde{\sigma}_0^2(\frac{c_0}{q_0^2})(x_0 - d_0 - c_0\mu_0) \tag{16.11}$$

Further, the (B$_1$) in Lemma 16.3 and (16.6) imply that

$$\begin{aligned} z_1(\omega_1) =& [\Phi_*^{0,1}\widetilde{z}_0](\omega_1) \\ =& \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}r_1} \exp[-\frac{(\omega_1 - a_1\omega_0 - b_1)^2}{2r_1^2}] \frac{1}{\sqrt{2\pi}\widetilde{\sigma}_0} \exp[-\frac{(\omega_0 - \widetilde{\mu}_0)^2}{2\widetilde{\sigma}_0^2}]d\omega_0 \\ =& \frac{1}{\sqrt{2\pi}\sigma_1} \exp[-\frac{(\omega_1 - \mu_1)^2}{2{\sigma_1}^2}] \end{aligned} \tag{16.12}$$

where

$$\sigma_1^2 = \ a_1^2\widetilde{\sigma}_0^2 + r_1^2, \quad \mu_1 = a_1\widetilde{\mu}_0 + b_1 \tag{16.13}$$

Thus, we see, by (B$_2$) in Lemma 16.3, that

$$\widetilde{z}_{t-1}(\omega_{t-1}) = \lim_{\Xi_{t-1} \to x_{t-1}} \frac{F(\Xi_{t-1})z_{t-1}}{\int_{\mathbb{R}} F(\Xi_{t-1})z_{t-1}d\omega_{t-1}}$$

$$
\begin{aligned}
&\approx \frac{1}{\sqrt{2\pi}q_{t-1}}\exp[-\frac{(x_{t-1}-c_{t-1}\omega_{t-1}-d_{t-1})^2}{2q_{t-1}^2}]\frac{1}{\sqrt{2\pi}\sigma_{t-1}}\exp[-\frac{(\omega_{t-1}-\mu_{t-1})^2}{2\sigma_{t-1}^2}]\\
&\approx \frac{1}{\sqrt{2\pi}\widetilde{\sigma}_{t-1}}\exp[-\frac{(\omega_{t-1}-\widetilde{\mu}_{t-1})^2}{2\widetilde{\sigma}_{t-1}^2}] \qquad (16.14)
\end{aligned}
$$

where

$$
\begin{aligned}
\widetilde{\sigma}_{t-1}^2 &= \frac{q_{t-1}^2\sigma_{t-1}^2}{q_{t-1}^2+c_{t-1}^2\sigma_{t-1}^2} = \sigma_{t-1}^2\frac{q_{t-1}^2+c_{t-1}^2\sigma_{t-1}^2+q_{t-1}^2-q_{t-1}^2-c_{t-1}^2\sigma_{t-1}^2}{q_{t-1}^2+c_{t-1}^2\sigma_{t-1}^2}\\
&= \sigma_{t-1}^2(1-\frac{c_{t-1}^2\sigma_{t-1}^2}{q_{t-1}^2+c_{t-1}^2\sigma_{t-1}^2})\\
\widetilde{\mu}_{t-1} &= \mu_{t-1}+\widetilde{\sigma}_{t-1}^2(\frac{c_{t-1}}{q_{t-1}^2})(x_{t-1}-c_{t-1}\mu_{t-1}) \qquad (16.15)
\end{aligned}
$$

Further, we see, by (B_1) in Lemma 16.3, that

$$
\begin{aligned}
z_t(\omega_t) &= [\Phi_*^{t-1,t}\widetilde{z}_{t-1}](\omega_t)\\
&\approx \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}r_t}\exp[-\frac{(\omega_t-a_t\omega_{t-1}-b_t)^2}{2r_t^2}]\frac{1}{\sqrt{2\pi}\widetilde{\sigma}_{t-1}}\exp[-\frac{(\omega_{t-1}-\widetilde{\mu}_{t-1})^2}{2\widetilde{\sigma}_{t-1}^2}]d\omega_{t-1}\\
&\approx \frac{1}{\sqrt{2\pi}\sigma_t}\exp[-\frac{(\omega_t-\mu_t)^2}{2{\sigma_t}^2}] \qquad (16.16)
\end{aligned}
$$

where

$$
\sigma_t^2 = a_t^2\widetilde{\sigma}_{t-1}^2+r_t^2, \quad \mu_t = a_t\widetilde{\mu}_{t-1}+b_t \qquad (16.17)
$$

Summing up the above (16.10)–(16.17), we see:

$$
\underset{\mu_0,\sigma_0}{\boxed{z_0}} \xrightarrow[(16.11)]{x_0} \underset{\widetilde{\mu}_0,\widetilde{\sigma}_0}{\boxed{\widetilde{z}_0}} \xrightarrow[(16.13)]{\Phi_*^{0,1}} \underset{\mu_1,\sigma_1}{\boxed{z_1}} \xrightarrow{x_1} \cdots \xrightarrow{\Phi_*^{t-2,t-1}} \underset{\mu_{t-1},\sigma_{t-1}}{\boxed{z_{t-1}}} \xrightarrow[(16.15)]{x_{t-1}} \underset{\widetilde{\mu}_{t-1},\widetilde{\sigma}_{t-1}}{\boxed{\widetilde{z}_{t-1}}} \xrightarrow[(16.17)]{\Phi_*^{t-1,t}} \underset{\mu_t,\sigma_t}{\boxed{z_t}} \xrightarrow{x_{t+1}} \cdots \xrightarrow{\Phi_*^{s-1,s}} \underset{\mu_s,\sigma_s}{\boxed{z_s}}
$$

And thus, we get

$$
z_s = \Phi_*^{s-1,s}(\widetilde{z}_{s-1}) \qquad (16.18)
$$

in (16.9).

16.5 Calculation: Smoothing part

16.5.1 Calculation: $\Big(F_s(\Xi_s)\Phi^{s,s+1}\widehat{F}_{s+1}(\times_{t=s+1}^{n}\Xi_t)\Big)$ in (16.9)

Put

$$\begin{aligned}\widetilde{f}_{x_n}(\omega_n) &= \frac{1}{\sqrt{2\pi}q_n}\exp[-\frac{(x_n-c_n\omega_n-d_n)^2}{2q_n^2}]\\ &\approx \exp[-\frac{(c_n\omega_n-(x_n-d_n))^2}{2q_n^2}] \equiv \exp[-\frac{1}{2}\Big(\widetilde{u}_n\omega_n-\widetilde{v}_n\Big)^2]\end{aligned}\tag{16.19}$$

where it is assumed that c_n, d_n and q_n are known ($t\in T$). And thus, put

$$\widetilde{u}_n=\frac{c_n}{q_n},\quad \widetilde{v}_n=\frac{x_n-d_n}{q_n}\tag{16.20}$$

And further, Lemma 16.3 implies that the causal operator $\Phi^{t-1.t}: L^\infty(\Omega_t)\to L^\infty(\Omega_{t-1})$ is defined by

$$\begin{aligned}&f_{t-1}(\omega_{t-1})=[\Phi^{t-1,t}\widetilde{f}_{x_t}](\omega_{t-1})\\ \approx&\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}r_t}\exp[-\frac{(\omega_t-a_t\omega_{t-1}-b_t)^2}{2r_t^2}]\exp[-\frac{(\widetilde{u}_t\omega_t-\widetilde{v}_t)^2}{2}]d\omega_t\\ \approx&\exp[-\frac{1}{2}\Big(\frac{\widetilde{v}_t}{\sqrt{1+r_t^2\widetilde{u}_t^2}}-\frac{\widetilde{u}_t(a_t\omega_{t-1}+b_t)}{\sqrt{1+r_t^2\widetilde{u}_t^2}}\Big)^2]\approx\exp[-\frac{1}{2}\Big(u_{t-1}\omega_{t-1}-v_{t-1}\Big)^2]\end{aligned}\tag{16.21}$$

where

$$u_{t-1}=-\frac{a_t\widetilde{u}_t}{\sqrt{1+r_t^2\widetilde{u}_t^2}},\quad v_{t-1}=\frac{b_t\widetilde{u}_t-\widetilde{v}_t}{\sqrt{1+r_t^2\widetilde{u}_t^2}}\tag{16.22}$$

And also, Lemma 16.3 implies that

$$\begin{aligned}\widetilde{f}_{x_{t-1}}(\omega_{t-1})&=\exp[-\frac{(c_{t-1}\omega_{t-1}+d_{t-1}-x_{t-1})^2}{2q_{t-1}^2}]\exp[-\frac{(u_{t-1}\omega_{t-1}-v_{t-1})^2}{2}]\\ &\approx\exp[-\frac{1}{2}(\frac{c_{t-1}^2+u_{t-1}^2q_{t-1}^2}{q_{t-1}^2})\Big(\omega_{t-1}-\frac{c_{t-1}(d_{t-1}-t_{t-1})+u_{t-1}v_{t-1}q_{t-1}^2}{c_{t-1}^2+u_{t-1}^2q_{t-1}^2}\Big)^2]\\ &\approx\exp[-\frac{1}{2}\Big(\widetilde{u}_{t-1}\omega_{t-1}-\widetilde{v}_{t-1}\Big)^2]\end{aligned}\tag{16.23}$$

where

$$\widetilde{u}_{t-1}=\frac{\sqrt{c_{t-1}^2+u_{t-1}^2q_{t-1}^2}}{q_{t-1}},\quad \widetilde{v}_{t-1}=\frac{c_{t-1}(d_{t-1}-t_{t-1})+u_{t-1}v_{t-1}q_{t-1}^2}{q_{t-1}\sqrt{c_{t-1}^2+u_{t-1}^2q_{t-1}^2}}\tag{16.24}$$

Summing up the above (16.19)-(16.24), we see:

$$\underset{\widetilde{w}_s}{\overset{\widetilde{u}_s,\widetilde{v}_s}{\boxed{\widetilde{f}_{x_s}}}}\xleftarrow{x_s}\cdots\xleftarrow{\Phi^{t-2,t-1}}\underset{\widetilde{w}_{t-1}}{\overset{\widetilde{u}_{t-1},\widetilde{v}_{t-1}}{\boxed{\widetilde{f}_{x_{t-1}}}}}\xleftarrow[(16.24)]{x_{t-1}}\underset{w_{t-1}}{\overset{u_{t-1},v_{t-1}}{\boxed{f_{t-1}}}}\xleftarrow[(16.22)]{\Phi^{t-1,t}}\underset{\widetilde{w}_t}{\overset{\widetilde{u}_t,\widetilde{v}_t}{\boxed{\widetilde{f}_{x_t}}}}\xleftarrow{x_t}\cdots\xleftarrow{x_{n-1}}\underset{w_{n-1}}{\overset{u_{n-1},v_{n-1}}{\boxed{f_{n-1}}}}\xleftarrow{\Phi^{n-1,n}}\underset{\widetilde{w}_n}{\overset{\widetilde{u}_n\widetilde{v}_n}{\boxed{\widetilde{f}_{x_n}=(16.19)}}}$$

And thus, we get

$$\widetilde{f}_{x_s} \approx \lim_{\Xi_t \to x_t \ (t \in \{s.s+1,\cdots,n\})} \frac{\Big(F_s(\Xi_s)\Phi^{s,s+1}\widehat{F}_{s+1}(\times_{t=s+1}^{n} \Xi_t)\Big)}{\|F_s(\Xi_s)\Phi^{s,s+1}\widehat{F}_{s+1}(\times_{t=s+1}^{n} \Xi_t)\Big)\|_{L^\infty(\Omega_s)}} \tag{16.25}$$

in (16.9)

After all, we solve Problem16.2(Kalman Filter), that is,

Answer 16.4. [The answer to Problem16.2(Kalman Filter)]

(A) Assume that a measured value $(x_0, x_2, \cdots, x_n)$ $(\in \times_{t=0}^{n} X_t)$ is obtained by the measurement $\mathsf{M}_{L^\infty(\Omega_0)}$ $(\widehat{\mathsf{O}}_{t_0}, S_{[*]}(z_0))$. Let $s(\in T)$ be fixed. Then, we get the Bayes-Kalman operator $[B^s_{\widehat{\mathsf{O}}_{t_0}}(\times_{t\in T}\{x_t\})](z_0)$, that is,

$$\Big([B^s_{\widehat{\mathsf{O}}_{t_0}}(\underset{t\in T}{\times}\{x_t\})]z_0\Big)(\omega_s) = \frac{\widetilde{f}_{x_s}(\omega_s)\cdot z_s(\omega_s)}{\int_{-\infty}^{\infty}\widetilde{f}_{x_s}(\omega_s)\cdot z_s(\omega_s)d\omega_s} = z_s^a(\omega_s) \quad (\forall \omega_s \in \Omega_s)$$

where z_s in (16.18) and $\widetilde{f}_{x_s}$ in (16.25) can be iteratively calculated as mentioned in this section.

Remark 16.5. The following classification is usual

(B_1) Smoothing: in the case that $0 \le s < n$

(B_2) Filter: in the case that $s = n$

(B_3) Prediction: in the case that $s = n$ and, for any m such that $n_0 \le m < n$, the existence observable $(X_m, \mathcal{F}_m, F_m) = (\{1\}, \{\emptyset, \{1\}\}, F_m)$ is defined by $F_m(\emptyset) \equiv 0$, $F_m(\{1\}) \equiv 1$,

Chapter 17

Equilibrium statistical mechanics

In this chapter, we study and answer the following fundamental problems concerning classical equilibrium statistical mechanics:

(A) Is the principle of equal a priori probabilities indispensable for equilibrium statistical mechanics?

(B) Is the ergodic hypothesis related to equilibrium statistical mechanics?

(C) Why and where does the concept of "probability" appear in equilibrium statistical mechanics?

Note that there are several opinions for the formulation of equilibrium statistical mechanics. In this sense, the above problems are not yet answered. Thus we propose the measurement theoretical foundation of equilibrium statistical mechanics, and clarify the confusion between two aspects (i.e., probabilistic and kinetic aspects in equilibrium statistical mechanics), that is, we discuss

$$\begin{cases} \text{the kinetic aspect (i.e, causality)} & \cdots \text{ in Section 17.1} \\ \text{the probabilistic aspect (i.e., measurement)} & \cdots \text{ in Section 17.2} \end{cases}$$

And we answer the above (A) and (B), that is, we conclude that

(A) is "No", but, (B) is "Yes".

and further, we can understand the problem (C).

This chapter is extracted from the following: [34] S. Ishikawa, "Ergodic Hypothesis and Equilibrium Statistical Mechanics in the Quantum Mechanical World View," *World Journal of Mechanics*, Vol. 2, No. 2, 2012, pp. 125-130. doi: 10.4236/wim.2012.22014.

17.1 Equilibrium statistical mechanical phenomena concerning Axiom 2 (causality)

17.1.1 Equilibrium statistical mechanical phenomena

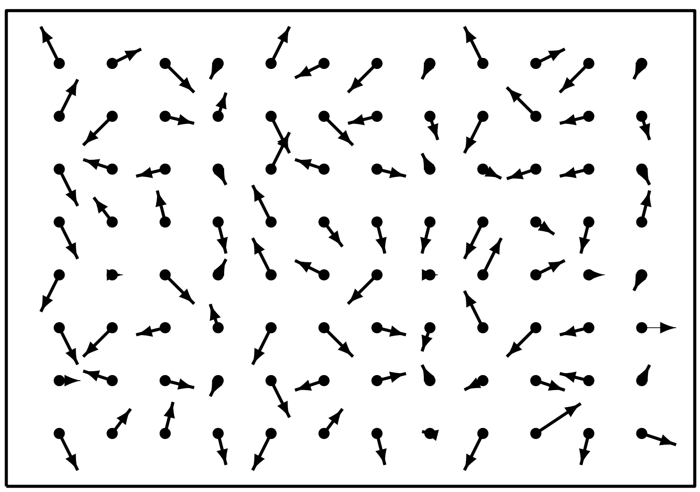
Hypothesis 17.1. [Equilibrium statistical mechanical hypothesis]. Assume that about $N(\approx 10^{24} \approx 6.02 \times 10^{23} \approx$ "the Avogadro constant") particles (for example, hydrogen molecules) move in a box with about 20 liters. It is natural to assume the following phenomena ① − ④:

① Every particle obeys Newtonian mechanics.

② Every particle moves uniformly in the box. For example, a particle does not halt in a corner.

③ Every particle moves with the same statistical behavior concerning time.

④ The motions of particles are (approximately) independent of each other.



(17.1)

In what follows we shall devote ourselves to the problem:

(D) how to describe the above equilibrium statistical mechanical phenomena ① − ④ in terms of quantum language (=measurement theory).

17.1.2 About ① in Hypothesis 17.1

In Newtonian mechanics, any state of a system composed of $N(\approx 10^{24})$ particles is represented by a point (q,p) $\big(\equiv$ (position, momentum) $= (q_{1n}, q_{2n}, q_{3n}, p_{1n}, p_{2n}, p_{3n})_{n=1}^N \big)$ in a phase (or state) space $\mathbb{R}^{6N}$. Let $\mathcal{H} : \mathbb{R}^{6N} \to \mathbb{R}$ be a Hamiltonian such that

$$\mathcal{H}\big((q_{1n}, q_{2n}, q_{3n}, p_{1n}, p_{2n}, p_{3n})_{n=1}^N\big) = \text{momentum energy} + \text{potential energy}$$

$$=[\sum_{n=1}^{N}\sum_{k=1,2,3}\frac{(p_{kn})^2}{2\times \text{particle's mass}}]+U((q_{1n},q_{2n},q_{3n})_{n=1}^{N}). \tag{17.2}$$

Fix a positive $E>0$. And define the measure $\nu_{_E}$ on the energy surface $\Omega_{_E}$ ($\equiv \{(q,p)\in \mathbb{R}^{6N} \mid \mathcal{H}(q,p)=E\}$) such that

$$\nu_{_E}(B)=\int_B |\nabla \mathcal{H}(q,p)|^{-1} dm_{6N-1} \qquad (\forall B \in \mathcal{B}_{\Omega_{_E}},\ \text{the Borel field of } \Omega_{_E})$$

where

$$|\nabla \mathcal{H}(q,p)|=[\sum_{n=1}^{N}\sum_{k=1,2,3}\{(\frac{\partial \mathcal{H}}{\partial p_{kn}})^2+(\frac{\partial \mathcal{H}}{\partial q_{kn}})^2\}]^{1/2}$$

and dm_{6N-1} is the usual surface Lebesgue measure on $\Omega_{_E}$. Let $\{\psi_t^{E}\}_{-\infty<t<\infty}$ be the flow on the energy surface $\Omega_{_E}$ induced by the Newton equation with the Hamiltonian $\mathcal{H}$, or equivalently, Hamilton's canonical equation:

$$\frac{dq_{kn}}{dt}=\frac{\partial \mathcal{H}}{\partial p_{kn}}, \quad \frac{dp_{kn}}{dt}=-\frac{\partial \mathcal{H}}{\partial q_{kn}}, \tag{17.3}$$
$$(k=1,2,3, \quad n=1,2,\ldots,N).$$

Liouville's theorem (*cf.*[55]) says that the measure $\nu_{_E}$ is invariant concerning the flow $\{\psi_t^{E}\}_{-\infty<t<\infty}$. Defining the normalized measure $\overline{\nu}_{_E}$ such that $\overline{\nu}_{_E}=\frac{\nu_{_E}}{\nu_{_E}(\Omega_{_E})}$, we have the normalized measure space $(\Omega_{_E},\mathcal{B}_{\Omega_{_E}},\overline{\nu}_{_E})$.

Putting $\mathcal{A}=C_0(\Omega_{_E})=C(\Omega_{_E})$ (from the compactness of $\Omega_{_E}$), we have the classical basic structure:

$$[C(\Omega_{_E})\subseteq L^{\infty}(\Omega_{_E},\nu_{_E})\subseteq B(L^2(\Omega_{_E},\nu_{_E}))]$$

Thus, putting $T=\mathbb{R}$, and solving the (17.4), we get $\omega_t=(q(t),p(t))$, $\phi_{t_1.t_2}=\psi^{E}_{t_2-t_1}$, $\Phi^*_{t_1.t_2}\delta_{\omega_{t_1}}=\delta_{\phi_{t_1.t_2}(\omega_{t_1})}$ ($\forall \omega_{t_1}\in\Omega_{_E}$), and further we define the sequential deterministic causal operator $\{\Phi_{t_1,t_2}:L^{\infty}(\Omega_{_E})\to L^{\infty}(\Omega_{_E})\}_{(t_1.t_2)\in T^2_{\leq}}$ (*cf.* Definition 10.4).

17.1.3 About ② in Hypothesis 17.1

Now let us begin with the well-known ergodic theorem (*cf.* [55]). For example, consider one particle P_1. Put

$$S_{P_1}=\{\omega\in\Omega_{_E} \mid \text{ a state } \omega \text{ such that the particle } P_1 \text{ stays around a corner of the box }\}$$

Clearly, it holds that $S_{P_1}\subsetneqq \Omega_{_E}$. Also, if $\psi_t^{E}(S_{P_1})\subseteq S_{P_1}$ ($0\leqq \forall t<\infty$), then the particle P_1 must always stay a corner. This contradicts ②. Therefore, ② means the following:

② $'$ [Ergodic property]: If a compact set $S(\subseteq \Omega_{_E}, S \neq \emptyset)$ satisfies $\psi_t^{_E}(S) \subseteq S$ $(0 \leqq \forall t < \infty)$, then it holds that $S = \Omega_{_E}$.

The ergodic theorem (*cf.* [55]) says that the above ②$'$ is equivalent to the following equality:

$$\underbrace{\int_{\Omega_{_E}} f(\omega)\overline{\nu}_{_E}(d\omega)}_{\text{((state) space average)}} = \lim_{T\to\infty} \underbrace{\frac{1}{T}\int_{\alpha}^{\alpha+T} f(\psi_t^{_E}(\omega_0))dt}_{\text{(time average)}} \tag{17.4}$$
$$(\forall \alpha \in \mathbb{R}, \forall f \in C(\Omega_{_E}), \quad \forall \omega_0 \in \Omega_{_E})$$

After all, the ergodic property ②$'$ ($\Leftrightarrow$ (17.4)) says that if T is sufficiently large, it holds that

$$\int_{\Omega_{_E}} f(\omega)\overline{\nu}_{_E}(d\omega) \approx \frac{1}{T}\int_{\alpha}^{\alpha+T} f(\psi_t^{_E}(\omega_0))dt. \tag{17.5}$$

Put $\overline{m}_{_T}(dt) = \frac{dt}{T}$. The probability space $([\alpha, \alpha{+}T], \mathcal{B}_{[\alpha,\alpha+T]}, \overline{m}_{_T})$ (or equivalently, $([0, T], \mathcal{B}_{[0,T]}, \overline{m}_{_T})$) is called a (normalized) *first staying time space*, also, the probability space $(\Omega_{_E}, \mathcal{B}_{\Omega_{_E}}, \overline{\nu}_{_E})$ is called a (normalized)*second staying time space*. Note that these mathematical probability spaces are not related to "probability" (Recall the linguistic interpretation (§3.1) :*there is no probability without measurement*).

17.1.4 About ③ and ④ in Hypothesis 17.1

Put $K_N = \{1, 2, \ldots, N(\approx 10^{24})\}$. For each k ($\in K_N$), define the coordinate map $\pi_k : \Omega_{_E}(\subset \mathbb{R}^{6N}) \to \mathbb{R}^6$ such that

$$\begin{aligned}\pi_k(\omega) = \pi_k(q,p) =&\pi_k((q_{1n}, q_{2n}, q_{3n}, p_{1n}, p_{2n}, p_{3n})_{n=1}^N)\\ =&(q_{1k}, q_{2k}, q_{3k}, p_{1k}, p_{2k}, p_{3k})\end{aligned} \tag{17.6}$$

for all $\omega = (q,p) = (q_{1n},\ q_{2n},\ q_{3n},\ p_{1n},\ p_{2n}, p_{3n})_{n=1}^N \in \Omega_{_E}(\subset \mathbb{R}^{6N})$.

Also, for any subset K ($\subseteq K_N = \{1, 2,\ \ldots, N\ (\approx 10^{24})\}$), define the distribution map $D_K^{(\cdot)} : \Omega_{_E}\ (\subset \mathbb{R}^{6N}) \to \mathcal{M}_{+1}^m(\mathbb{R}^6)$ such that

$$D_K^{(q,p)} = \frac{1}{\sharp[K]}\sum_{k\in K}\delta_{\pi_k(q,p)} \quad (\forall (q,p) \in \Omega_{_E}(\subset \mathbb{R}^{6N}))$$

where $\sharp[K]$ is the number of the elements of the set K.

Let $\omega_0 (\in \Omega_{_E})$ be a state. For each n ($\in K_N$), we define the map $X_n^{\omega_0} : [0,T] \to \mathbb{R}^6$ such that

$$X_n^{\omega_0}(t) = \pi_n(\psi_t^{_E}(\omega_0)) \qquad (\forall t \in [0,T]). \tag{17.7}$$

And, we regard $\{X_n^{\omega_0}\}_{n=1}^N$ as random variables (i.e., measurable functions) on the probability space $([0,T], \mathcal{B}_{[0,T]}, \overline{m}_T)$. Then, ③ and ④ respectively means

③′ $\{X_n^{\omega_0}\}_{n=1}^N$ is a *sequence with the approximately identical distribution concerning time.* In other words, there exists a normalized measure ρ_E on $\mathbb{R}^6$ (i.e., $\rho_E \in \mathcal{M}_{+1}^m(\mathbb{R}^6)$) such that:

$$\overline{m}_T(\{t \in [0,T] \ : \ X_n^{\omega_0}(t) \in \Xi\}) \approx \rho_E(\Xi) \qquad (\forall \Xi \in \mathcal{B}_{\mathbb{R}^6}, n = 1, 2, \ldots, N) \tag{17.8}$$

④′ $\{X_n^{\omega_0}\}_{n=1}^N$ is *approximately independent*, in the sense that, for any $K_0 \subset \{1, 2, \ldots, N(\approx 10^{24})\}$ such that $1 \leqq \sharp[K_0] \ll N$ (that is, $\frac{\sharp[K_0]}{N} \approx 0$), it holds that

$$\overline{m}_T(\{t \in [0,T] : X_k^{\omega_0}(t) \in \Xi_k(\in \mathcal{B}_{\mathbb{R}^6}), k \in K_0\}) \approx \mathop{\times}_{k \in K_0} \overline{m}_T(\{t \in [0,T] : X_k^{\omega_0}(t) \in \Xi_k(\in \mathcal{B}_{\mathbb{R}^6})\}).$$

Here, we can assert the advantage of our method in comparison with Ruelle's method (*cf.*[66]) as follows.

Remark 17.2. [About the time interval $[0,T]$]. For example, as one of typical cases, consider the motion of 10^{24} particles in a cubic box (whose long side is 0.3m). It is usual to consider that "averaging velocity"$=5 \times 10^2$m/s, "mean free path"$=10^{-7}$m. And therefore, the collisions rarely happen among $\sharp[K_0]$ particles in the time interval $[0,T]$, and therefore, the motion is "almost independent". For example, putting $\sharp[K_0] = 10^{10}$, we can calculate the number of times a certain particle collides with K_0-particles in [0,T] as $(10^{-7} \times \frac{10^{24}}{10^{10}})^{-1} \times (5 \times 10^2) \times T \approx 5 \times 10^{-5} \times T$. Hence, in order to expect that ③′ and ④′ hold, it suffices to consider that $T \approx 5$ seconds. ///

Also, we see, by (17.7) and (17.5), that, for $K_0 (\subseteq K_N)$ such that $1 \leq \sharp[K_0] \ll N$,

$$\begin{aligned}
&\overline{m}_T(\{t \in [0,T] \ : \ X_k^{\omega_0}(t) \in \Xi_k(\in \mathcal{B}_{\mathbb{R}^6}), k \in K_0\}) \\
=&\overline{m}_T(\{t \in [0,T] : \pi_k(\psi_t^E(\omega_0) \in \Xi_k(\in \mathcal{B}_{\mathbb{R}^6}), k \in K_0\}) \\
=&\overline{m}_T(\{t \in [0,T] : \psi_t^E(\omega_0) \in ((\pi_k)_{k \in K_0})^{-1}(\mathop{\times}_{k \in K_0} \Xi_k)\}) \\
\approx& \overline{\nu}_E\big(((\pi_k)_{k \in K_0})^{-1}(\mathop{\times}_{k \in K_0} \Xi_k)\big) \\
\equiv&\big(\overline{\nu}_E \circ ((\pi_k)_{k \in K_0})^{-1}\big)(\mathop{\times}_{k \in K_0} \Xi_k).
\end{aligned} \tag{17.9}$$

Particularly, putting $K_0 = \{k\}$, we see:

$$\overline{m}_T(\{t \in [0,T] \; : \; X_k^{\omega_0}(t) \in \Xi\}) \approx (\overline{\nu}_E \circ \pi_k^{-1})(\Xi) \qquad (\forall \Xi \in \mathcal{B}_{\mathbb{R}^6}). \tag{17.10}$$

Hence, we can describe the ③ and ④ in terms of $\{\pi_k\}$ in what follows.

Hypothesis 17.3. [③ and ④]. Put $K_N = \{1, 2, \ldots, N(\approx 10^{24})\}$. Let $\mathcal{H}$, E, ν_E, $\overline{\nu}_E$, $\pi_k : \Omega_E \to \mathbb{R}^6$ be as in the above. Then, summing up ③ and ④, by (17.9) we have:

(E) $\{\pi_k : \Omega_E \to \mathbb{R}^6\}_{k=1}^N$ is approximately independent random variables with the identical distribution in the sense that there exists ρ_E $(\in \mathcal{M}_{+1}^m(\mathbb{R}^6))$ such that

$$\bigotimes_{k \in K_0} \rho_E (= \text{"product measure"}) \approx \overline{\nu}_E \circ ((\pi_k)_{k \in K_0})^{-1}. \tag{17.11}$$

for all $K_0 \subset K_N$ and $1 \leqq \sharp[K_0] \ll N$.

Also, a state $(q,p)(\in \Omega_E)$ is called an *equilibrium state* if it satisfies $D_{K_N}^{(q,p)} \approx \rho_E$.

17.1.5 Ergodic Hypothesis

Now, we have the following theorem (*cf.*[34]):

Theorem 17.4. [Ergodic hypothesis]. Assume Hypothesis 17.3 (or equivalently, ③ and ④). Then, for any $\omega_0 = (q(0), p(0)) \in \Omega_E$, it holds that

$$[D_{K_N}^{(q(t),p(t))}](\Xi) \approx \overline{m}_T(\{t \in [0,T] \; : \; X_k^{\omega_0}(t) \in \Xi\}) \qquad (\forall \Xi \in \mathcal{B}_{\mathbb{R}^6}, k = 1, 2, \ldots, N(\approx 10^{24})) \tag{17.12}$$

for almost all t. That is, $0 \leqq \overline{m}_T(\{t \in [0,T] : (17.12) \text{ does not hold}\}) \ll 1$.

Proof. Let $K_0 \subset K_N$ such that $1 \ll \sharp[K_0] \equiv N_0 \ll N$ (that is, $\frac{1}{\sharp[K_0]} \approx 0 \approx \frac{\sharp[K_0]}{N}$). Then, from Hypothesis A, the law of large numbers (*cf.* [54]) says that

$$D_{K_0}^{(q(t),p(t))} \approx \overline{\nu}_E \circ \pi_k^{-1} \; (\approx \rho_E \;) \tag{17.13}$$

for almost all time t. Consider the decomposition $K_N = \{K_{(1)}, K_{(2)}, \ldots, K_{(L)}\}$. (i.e., $K_N = \bigcup_{l=1}^L K_{(l)}$, $K_{(l)} \cap K_{(l')} = \emptyset \;\; (l \neq l')$), where $\sharp[K_{(l)}] \approx N_0$ $(l = 1, 2, \ldots, L)$. From (7.13), it holds that, for each k $(= 1, 2, \ldots, N \; (\approx 10^{24}))$,

$$D_{K_N}^{(q(t),p(t))} = \frac{1}{N} \sum_{l=1}^{L} [\sharp[K_{(l)}] \times D_{K_{(l)}}^{(q(t),p(t))}]$$

$$\approx \frac{1}{N}\sum_{l=1}^{L}[\sharp[K_{(l)}] \times \rho_{_E}] \approx \overline{\nu}_{_E} \circ \pi_k^{-1} \; (\approx \rho_{_E} \;), \tag{17.14}$$

for almost all time t. Thus, by (17.10), we get (17.12). Hence, the proof is completed.

We believe that Theorem 17.4 is just what should be represented by the *"ergodic hypothesis"* such that

$$\begin{aligned} &\text{"population average of } N \text{ particles at each } t\text{"} \\ =&\text{"time average of one particle".} \end{aligned}$$

Thus, we can assert that the ergodic hypothesis is related to equilibrium statistical mechanics (*cf.* the (B) in the abstract). Here, the ergodic property ②′ (or equivalently, equality (17.5)) and the above ergodic hypothesis should not be confused. Also, it should be noted that the ergodic hypothesis does not hold if the box (containing particles) is too large.

Remark 17.5. [The law of increasing entropy]. The entropy $H(q,p)$ of a state $(q,p)(\in \Omega_{_E})$ is defined by

$$H(q,p) = k \log[\nu_{_E}(\{(q',p') \in \Omega_{_E} : D_{K_N}^{(q,p)} \approx D_{K_N}^{(q',p')})\})]$$

where

$$k = [\text{Boltzmann constant}]/([\text{Plank constant}]^{3N} N!)$$

Since almost every state in $\Omega_{_E}$ is equilibrium, the entropy of almost every state is equal $k \log \nu_{_E}(\Omega_{_E})$. Therefore, it is natural to assume that the law of increasing entropy holds.

17.2 Equilibrium statistical mechanical phenomena concerning Axiom 1 (Measurement)

In this section we shall study the probabilistic aspects of equilibrium statistical mechanics. For completeness, note that

(F) the argument in the previous section is not related to “probability”

since Axiom 1 (measurement; §2.7) does not appear in Section 17.1. Also, Recall the linguistic interpretation (§3.1) : *there is no probability without measurement.*

Note that the (17.12) implies that the equilibrium statistical mechanical system at almost all time t can be regarded as:

(G) a box including about 10^{24} particles such as the number of the particles whose states belong to Ξ ($\in \mathcal{B}_{\mathbb{R}^6}$) is given by $\rho_{_E}(\Xi) \times 10^{24}$.

Thus, it is natural to assume as follows.

(H) if we, at random, choose a particle from 10^{24} particles in the box at time t, then the probability that the state $(q_1, q_2, q_3, p_1, p_2, p_3)$ $(\in \mathbb{R}^6)$ of the particle belongs to Ξ ($\in \mathcal{B}_{\mathbb{R}^6}$) is given by $\rho_{_E}(\Xi)$.

In what follows, we shall represent this (H) in terms of measurements. Define the observable $\mathsf{O}_0 = (\mathbb{R}^6, \mathcal{B}_{\mathbb{R}^6}, F_0)$ in $L^\infty(\Omega_{_E})$ such that

$$
[F_0(\Xi)](q,p) = [D_{K_N}^{(q,p)}](\Xi)\Big(\equiv \frac{\sharp[\{k \mid \pi_k(q,p) \in \Xi\}]}{\sharp[K_N]}\Big)
$$
$$
(\forall \Xi \in \mathcal{B}_{\mathbb{R}^6}, \forall (q,p) \in \Omega_{_E}(\subset \mathbb{R}^{6N})). \tag{17.15}
$$

Thus, we have the measurement $\mathsf{M}_{L^\infty(\Omega_E)}(\mathsf{O}_0 := (\mathbb{R}^6, \mathcal{B}_{\mathbb{R}^6}, F_0), S_{[\delta_{\psi_t(q_0,p_0)}]})$. Then we say, by Axiom 1 (measurement; §2.7) , that

(I) the probability that the measured value obtained by the measurement $\mathsf{M}_{L^\infty(\Omega_E)}(\mathsf{O}_0 := (\mathbb{R}^6, \mathcal{B}_{\mathbb{R}^6}, F_0), S_{[\delta_{\psi_t(q_0,p_0)}]})$ belongs to $\Xi(\in \mathcal{B}_{\mathbb{R}^6})$ is given by $\rho_{_E}(\Xi)$. That is because Theorem A says that $[F_0(\Xi)](\psi_t(q_{_0}, p_{_0})) \approx \rho_{_E}(\Xi)$ (almost every time t).

Also, let $\Psi_t^E : L^\infty(\Omega_{_E}) \to L^\infty(\Omega_{_E})$ be a deterministic Markov operator determined by the continuous map $\psi_t^E : \Omega_{_E} \to \Omega_{_E}$ (*cf.* Section 17.1.2). Then, it clearly holds $\Psi_t^E \mathsf{O}_0 = \mathsf{O}_0$. And, we must take a $\mathsf{M}_{L^\infty(\Omega_{_E})}(\mathsf{O}_0, S_{[(q(t_k),p(t_k))]})$ for each time $t_1, t_2, \ldots, t_k, \ldots, t_n$. However, the linguistic interpretation (§3.1) :(*there is no probability without measurement*) says that it suffices to take the simultaneous measurement $\mathsf{M}_{C(\Omega_{_E})}(\times_{k=1}^n \mathsf{O}_0,\ S_{[\delta_{(q(0),p(0))}]})$.

Remark 17.6. [The principle of equal a priori probabilities]. The (H) (or equivalently, (I)) says "choose a particle from N particles in box", and not "choose a state from the state space Ω_E". Thus, as mentioned in the abstract of this chapter, the principle of equal (a priori) probability is not related to our method. If we try to describe Ruele's method [66] in terms of measurement theory, we must use mixed measurement theory (*cf.* Chapter 9). However, this trial will end in failure.

17.3 Conclusions

Our concern in this chapter may be regarded as the problem: "What is the classical mechanical world view?" Concretely speaking, we are concerned with the problem:

"our method" vs. "Ruele's method [66] (which has been authorized for a long time)"

And, we assert the superiority of our method to Ruele's method in Remarks 17.2, 17.5, 17.6.

Chapter 18

Reliability in psychological tests

In this chapter, we shall introduce a measurement theoretical approach to a problem of analyzing scores of tests for students. The obtained score is assumed to be a sum of a true value and a measurement error. It is also subject to a systematic error (=noise) depending on his/her health or psychological condition at the test. In such cases, statistical measurements are convenient since these two errors (i.e., measurement error and systematic error) in measurement theory can be characterized in different mathematical structures. As a result, we show that

$$\text{“reliability coefficient”} = \text{“correlation coefficient”}$$

in a clear formulation.

This chapter is extracted from the following.

[50] K. Kikuchi, S. Ishikawa, “Psychological tests in Measurement Theory,” Far east journal of theoretical statistics, 32(1) 81-99, (2010) ISSN: 0972-0863

18.1 Reliability in psychological tests

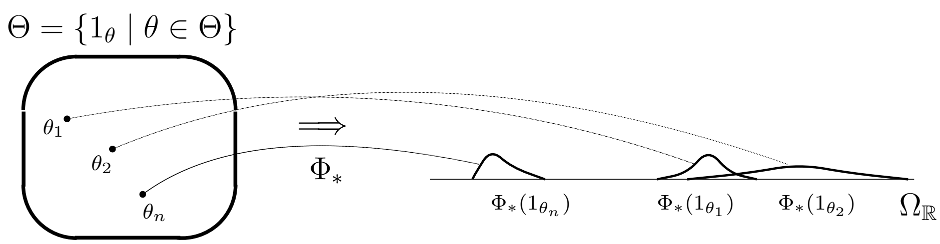
18.1.1 Preparation

In this section, let us consider reliability of psychological tests for a group of students. We discuss examples from measurement theoretical characterization of tests to measure mathematical ability of students.

Let $\Theta := \{\theta_1, \theta_2, \ldots, \theta_n\}$ be a set of students, say, there are n students $\theta_1, \theta_2, \ldots, \theta_n$. Define the counting measure ν_c on Θ such that $\nu_c(\{\theta_i\}) = 1$ $(i = 1, 2, \ldots, n)$. The Θ will be regarded as a state. For each θ_i $(\in \Theta)$, we define 1_{θ_i} $(\in L^1_{+1}(\Theta, \nu_c))$ by $1_{\theta_i}(\theta) = 1$ (if $\theta = \theta_i$), $=$ 0 (if $\theta \neq \theta_i$). Recall that Θ can be identified with the $\{1_{\theta_i} \mid \theta_i \in \Theta\}$ under the identification: $\Theta \ni \theta_i \leftrightarrow 1_{\theta_i} \in \{1_\theta \mid \theta \in \Theta\}$.

For simplicity, we shall begin with the test for one student θ_i $(\in \Theta)$. Let $(\Omega_{\mathbb{R}}, \mathcal{F}_{\Omega_{\mathbb{R}}}, d\omega)$ be the Lebesgue measure space where $\Omega_{\mathbb{R}} = \mathbb{R}$.

Example 18.1. (test in mathematics for a student θ_i) Let $\Theta := \{\theta_1, \theta_2, \ldots, \theta_n\}$ be a state space which is identified with the set of the students. The mathematical ability of the student θ_i $(\in \Theta)$ is assumed to be represented by a statistical state $\Phi_*(1_{\theta_i})$ $(\in L^1_{+1}(\Omega_{\mathbb{R}}, d\omega))$ $(i = 1, 2, \ldots, n)$ where $\Phi_* : L^1(\Theta, \nu_c) \to L^1(\Omega_{\mathbb{R}}, d\omega)$ is a pre-dual Markov causal operator of $\Phi : L^\infty(\Omega_{\mathbb{R}}, d\omega) \to L^\infty(\Theta, \nu_c)$.



Let $\mathsf{O} := (X_{\mathbb{R}}, \mathcal{F}_{X_{\mathbb{R}}}, F)$ be an observable in $L^\infty(\Omega_{\mathbb{R}}, d\omega)$. Axiom$^{(\mathrm{m})}$ 1 (§9.1) asserts that

(A) the probability that the score (measured value) of the student θ_i $(\in \Theta)$ obtained by the statistical measurement $\mathsf{M}_{L^\infty(\Omega_{\mathbb{R}}, d\omega)}(\mathsf{O}, S_{[*]}(\Phi_*(1_{\theta_i})))$ belongs to a set Ξ $(\in \mathcal{F}_{X_{\mathbb{R}}})$ is given by

$$ {}_{L^1(\Omega_{\mathbb{R}}, d\omega)}\langle \Phi_*(1_{\theta_i}), F(\Xi) \rangle_{L^\infty(\Omega_{\mathbb{R}}, d\omega)} \Big(= \int_{\Omega_{\mathbb{R}}} [F(\Xi)](\omega) \, [\Phi_*(1_{\theta_i})](\omega) \, d\omega \Big). $$

Remark 18.2. In the above, readers may have a question

(B) What is the unknown pure state $[*]$ in $S_{[*]}$?

Imaging the deterministic causal map $\psi : \Theta \to \Omega_{\mathbb{R}}$, we may consider that

$$ [*] = \psi(\theta_i) = \int_{\Omega_{\mathbb{R}}} \omega [\Phi_*(1_{\theta_i})](\omega) \, d\omega. $$

Also, note that the $[*]$ does not play an important role in this chapter since Bayes' theorem 9.11 is not used.

Remark 18.3. It should be kept in mind that the variance σ_i^2 of the ability of θ_i $(\in \Theta)$ $(i = 1, 2, \ldots, n)$ is not constant, that is to say, we do not assume that $\sigma_i^2 = \sigma_j^2$ $(\forall i, \forall j)$:

$$\sigma_i^2 := \int_{\Omega_{\mathbb{R}}} (\omega - \mu_i)^2 \, [\Phi_*(1_{\theta_i})](\omega) \, d\omega \quad (i = 1, 2, \ldots, n), \tag{18.1}$$

where μ_i is an expectation of $\Phi_*(1_{\theta_i})$:

$$\mu_i := \int_{\Omega_{\mathbb{R}}} \omega \, [\Phi_*(1_{\theta_i})](\omega) \, d\omega \quad (i = 1, 2, \ldots, n). \tag{18.2}$$

18.1.2 Group measurement (= parallel measurement)

The above example is the test for a student θ_i $(\in \Theta)$. Keeping this in mind, we will next consider the test for a group of n students. Let $\Omega_{\mathbb{R}}^n = \mathbb{R}^n$, and let $(\Omega_{\mathbb{R}}^n, \mathcal{F}_{\Omega_{\mathbb{R}}^n}, d\omega^n)$ be a n-dimensional Lebesgue measure space. Furthermore, let $\mathsf{O} := (X_{\mathbb{R}}, \mathcal{F}_{X_{\mathbb{R}}}, F)$ and $\mathsf{M}_{L^\infty(\Omega_{\mathbb{R}}, d\omega)}(\mathsf{O}, S_{[*]}(\Phi_*(1_{\theta_i})))$ $(i = 1, 2, \ldots, n)$ be as in above example. Here, we consider a parallel measurement $\mathsf{M}_{L^\infty(\Omega_{\mathbb{R}}^n, d\omega^n)}(\widehat{\mathsf{O}}, S_{[*]}(\widehat{\rho}))$ where $\widehat{\mathsf{O}} := (X_{\mathbb{R}}^n, \mathcal{F}_{X_{\mathbb{R}}^n}, \widehat{F})$ is an observable in $L^\infty(\Omega_{\mathbb{R}}^n, d\omega^n)$. If

$$[\widehat{F}(\Xi_1 \times \Xi_2 \times \cdots \times \Xi_n)](\omega_1, \omega_2, \ldots, \omega_n) = [F(\Xi_1)](\omega_1) \cdot [F(\Xi_2)](\omega_2) \cdots [F(\Xi_n)](\omega_n),$$

and

$$\widehat{\rho}(\omega_1, \omega_2, \ldots, \omega_n) = [\Phi_*(1_{\theta_1})](\omega_1) \cdot [\Phi_*(1_{\theta_2})](\omega_2) \cdots [\Phi_*(1_{\theta_n})](\omega_n),$$

then, the parallel measurement $\mathsf{M}_{L^\infty(\Omega_{\mathbb{R}}^n, d\omega^n)}(\widehat{\mathsf{O}}, S_{[*]}(\widehat{\rho}))$ is denoted by

$$\otimes_{\theta_i \in \Theta} \mathsf{M}_{L^\infty(\Omega_{\mathbb{R}}, d\omega)}(\mathsf{O}, S_{[*]}(\Phi_*(1_{\theta_i}))).$$

In addition, we introduce the following notations concerning tensor product:

$$\otimes_{k=1}^n L^\infty(\Omega_{\mathbb{R}}, d\omega) = L^\infty(\Omega_{\mathbb{R}}^n, d\omega^n) \quad \text{and} \quad \otimes_{k=1}^n L^1(\Omega_{\mathbb{R}}, d\omega) = L^1(\Omega_{\mathbb{R}}^n, d\omega^n).$$

By the way, we introduce the test observable.

Definition 18.4. [Test observable] The $\mathsf{O}_\tau = (X_{\mathbb{R}}, \mathcal{F}_{X_{\mathbb{R}}}, F_\tau)$ is called a *test observable* in $L^\infty(\Omega_{\mathbb{R}}, d\omega)$, if F_τ satisfies the following no-bias condition:

$$\int_{X_{\mathbb{R}}} x \, [F_\tau(dx)](\omega) = \omega \quad (\forall \omega \in \Omega_{\mathbb{R}}). \tag{18.3}$$

Recall that the normal observable (*cf.* Example 2.24) and the exact observable (*cf.* Example 2.25).

For each θ_i $(\in \Theta)$, we use the notation $\mathsf{M}^{(i)}_{\mathsf{O}_\tau}$ to the test for θ_i $(\in \Theta)$ (the measurement of the test observable O_τ for the statistical state $\Phi_*(1_{\theta_i})$):

$$\mathsf{M}^{(i)}_{\mathsf{O}_\tau} := \mathsf{M}_{L^\infty(\Omega_{\mathbb{R}},d\omega)}(\mathsf{O}_\tau, S_{[*]}(\Phi_*(1_{\theta_i}))). \tag{18.4}$$

Now we are ready to consider the test for a set of the n students in our measurement theory.

Definition 18.5. [Test, Group test] Let $\Theta := \{\theta_1, \theta_2, \ldots, \theta_n\}$, $X_{\mathbb{R}} = \Omega_{\mathbb{R}} = \mathbb{R}$ and $\Phi_* : L^1_{+1}(\Theta, \nu_c) \to L^1_{+1}(\Omega_{\mathbb{R}}, d\omega)$ be as in Example 18.1. Let $\mathsf{O}_\tau := (X_{\mathbb{R}}, \mathcal{F}_{X_{\mathbb{R}}}, F_\tau)$ be a test observable in $L^\infty(\Omega_{\mathbb{R}}, d\omega)$. The measurement $\mathsf{M}_{L^\infty(\Omega_{\mathbb{R}},d\omega)}(\mathsf{O}_\tau, S_{[*]}(\Phi_*(1_{\theta_i})))$ is called a *test for a student* θ_i $(\in \Theta)$ and symbolized by $\mathsf{M}^{(i)}_{\mathsf{O}_\tau}$ for short. And the measurement

$$\otimes_{\theta_i \in \Theta} \mathsf{M}_{L^\infty(\Omega_{\mathbb{R}},d\omega)}(\mathsf{O}_\tau, S_{[*]}(\Phi_*(1_{\theta_i}))) \quad (\text{or in short, } \otimes_{\theta_i \in \Theta} \mathsf{M}^{(i)}_{\mathsf{O}_\tau}), \tag{18.5}$$

is called a *group test* and symbolized by $\mathsf{M}^{\otimes}_{\mathsf{O}_\tau}$ for short.

Axiom$^{(\mathrm{m})}$ 1 (§9.1) says that

(C) the probability that the score $(x_1, x_2, \ldots, x_n)$ $(\in X^n_{\mathbb{R}})$ obtained by the group test $\otimes_{\theta_i \in \Theta} \mathsf{M}_{L^\infty(\Omega_{\mathbb{R}},d\omega)}$ $(\mathsf{O}_\tau, S_{[*]}(\Phi_*(1_{\theta_i})))$ (or in short, $\mathsf{M}^{\otimes}_{\mathsf{O}_\tau}$) belongs to the set $\times_{i=1}^n \Xi_i$ $(\in \mathcal{F}_{X^n_{\mathbb{R}}})$ is given by

$$\mathop{\times}_{\theta_i \in \Theta} {}_{L^1(\Omega_{\mathbb{R}},d\omega)}\langle \Phi_*(1_{\theta_i}), F_\tau(\Xi_i) \rangle_{L^\infty(\Omega_{\mathbb{R}},d\omega)} \Big(=: \widehat{P}_1(\mathop{\times}_{i=1}^n \Xi_i) = \mathop{\times}_{i=1}^n P_i(\Xi_i) \Big). \tag{18.6}$$

Here, $(X_{\mathbb{R}}, \mathcal{F}_{X_{\mathbb{R}}}, P_i)$ is a sample probability space of $\mathsf{M}^{(i)}_{\mathsf{O}_\tau}$.

Let $W : X^n_{\mathbb{R}} \to \mathbb{R}$ be a statistics (i.e., measurable function). Then, $\mathcal{E}_{\mathsf{M}^{\otimes}_{\mathsf{O}_\tau}}[W]$, the expectation of W, is defined by

$$\mathcal{E}_{\mathsf{M}^{\otimes}_{\mathsf{O}_\tau}}[W] = \int_{X_{\mathbb{R}}} \cdots \int_{X_{\mathbb{R}}} W(x_1, x_2, \ldots, x_n)\, \widehat{P}_1(dx_1\, dx_2 \cdots dx_n).$$

Definition 18.6. Let $\mathsf{O}_\tau := (X_{\mathbb{R}}, \mathcal{F}_{X_{\mathbb{R}}}, F_\tau)$ be a test observable in $L^\infty(\Omega_{\mathbb{R}}, d\omega)$.
(i: Score of θ_i) Let $\mathsf{M}_{L^\infty(\Omega_{\mathbb{R}},d\omega)}(\mathsf{O}_\tau, S_{[*]}(\Phi_*(1_{\theta_i})))$ (or in short, $\mathsf{M}^{(i)}_{\mathsf{O}_\tau}$) be a test for a student θ_i $(\in \Theta)$. Here, we consider the expectation of x_i $(\in X_{\mathbb{R}})$ and its variance.

1. $\mathrm{Av}[\mathsf{M}^{(i)}_{\mathsf{O}_\tau}] := \mathcal{E}_{\mathsf{M}^{(i)}_{\mathsf{O}_\tau}}[x_i]$,
2. $\mathrm{Var}[\mathsf{M}^{(i)}_{\mathsf{O}_\tau}] := \mathcal{E}_{\mathsf{M}^{(i)}_{\mathsf{O}_\tau}}\Big[(x_i - \mathrm{Av}[\mathsf{M}^{(i)}_{\mathsf{O}_\tau}])^2\Big]$.

(ii: Scores of n students) Let $\otimes_{\theta_i\in\Theta}\mathsf{M}_{L^\infty(\Omega_\mathbb{R},d\omega)}(\mathsf{O}_\tau, S_{[*]}(\Phi_*(1_{\theta_i})))$ (or in short, $\mathsf{M}^{\otimes}_{\mathsf{O}_\tau}$) be a group test. Here, we consider the expectation of $\frac{1}{n}(x_1+x_2+\cdots+x_n)$ and its variance.

1. $\mathrm{Av}[\mathbf{M}^{\otimes}_{\mathsf{O}_\tau}] := \mathcal{E}_{\mathsf{M}^{\otimes}_{\mathsf{O}_\tau}}\Big[\dfrac{1}{n}(x_1+x_2+\cdots+x_n)\Big]$,
2. $\mathrm{Var}[\mathsf{M}^{\otimes}_{\mathsf{O}_\tau}] := \mathcal{E}_{\mathsf{M}^{\otimes}_{\mathsf{O}_\tau}}\Big[\dfrac{1}{n}\sum_{k=1}^n(x_k - \mathrm{Av}[\mathsf{M}^{\otimes}_{\mathsf{O}_\tau}])^2\Big]$.

From the no-bias condition (18.3), we get

$$\mathrm{Av}[\mathsf{M}^{(i)}_{\mathsf{O}_\tau}] = \mathrm{Av}[\mathsf{M}^{(i)}_{\mathsf{O}_E}] = \int_{\Omega_\mathbb{R}} \omega\,[\Phi_*(1_{\theta_i})](\omega)\,d\omega = \mu_i, \tag{18.7}$$

$$\mathrm{Av}[\mathsf{M}^{\otimes}_{\mathsf{O}_\tau}] = \frac{1}{n}\sum_{i=1}^n \mathrm{Av}[\mathsf{M}^{(i)}_{\mathsf{O}_\tau}] = \mathrm{Av}[\mathsf{M}^{\otimes}_{\mathsf{O}_E}] = \frac{1}{n}\sum_{i=1}^n \mathrm{Av}[\mathsf{M}^{(i)}_{\mathsf{O}_E}] = \frac{1}{n}\sum_{i=1}^n \mu_i =: \overline{\mu}, \tag{18.8}$$

where $\mathsf{O}_E := (X_\mathbb{R}, \mathcal{F}_{X_\mathbb{R}}, E)$ is an exact observable in $L^\infty(\Omega_\mathbb{R}, d\omega)$.

18.1.3 Reliability coefficient

When we suppose the group test, we can consider the reliability coefficient which can be represented by a proportion of variance of mathematical abilities to obtained variance.

Definition 18.7. [Reliability coefficient] Let $\mathsf{O}_\tau := (X_\mathbb{R}, \mathcal{F}_{X_\mathbb{R}}, F_\tau)$ [resp. $\mathsf{O}_E := (X_\mathbb{R}, \mathcal{F}_{X_\mathbb{R}}, E)$] be a test observable [resp. an exact observable] in $L^\infty(\Omega_\mathbb{R}, d\omega)$. And, let

$$\mathsf{M}^{\otimes}_{\mathsf{O}_\tau} := \otimes_{\theta_i\in\Theta}\mathsf{M}_{L^\infty(\Omega_\mathbb{R},d\omega)}(\mathsf{O}_\tau, S_{[*]}(\Phi_*(1_{\theta_i})))$$

be a group test. The *reliability coefficient* $\mathrm{RC}[\mathsf{M}^{\otimes}_{\mathsf{O}_\tau}]$ of the group test $\mathsf{M}^{\otimes}_{\mathsf{O}_\tau}$ is defined by

$$\mathrm{RC}[\mathsf{M}^{\otimes}_{\mathsf{O}_\tau}] = \frac{\mathrm{Var}[\mathsf{M}^{\otimes}_{\mathsf{O}_E}]}{\mathrm{Var}[\mathsf{M}^{\otimes}_{\mathsf{O}_\tau}]}.$$

Now let us consider the measurement error. First, when the ability (true value) is ω $(\in\Omega)$, the measurement error Δ_ω is as follows:

$$\Delta_\omega := \Big(\int_{X_\mathbb{R}} (x-\omega)^2\,[F_\tau(dx)](\omega)\Big)^{1/2} \quad (\forall\omega\in\Omega). \tag{18.9}$$

Note that the error Δ_ω ($\forall \omega \in \Omega$) depends on ω ($\in \Omega$) in general, that is, we do not assume that $\Delta_\omega = \Delta_{\omega'}$ ($\forall \omega, \forall \omega' \in \Omega$). Next, for each θ_i ($\in \Theta$), the error Δ_i for the student θ_i ($\in \Theta$) is as follows:

$$\begin{aligned}\Delta_i &:= \Big(\int_{X_{\mathbb{R}}} \Delta_\omega \, [\Phi_*(1_{\theta_i})](\omega) \, d\omega \Big)^{1/2} \\ &= \Big(\int_{\Omega_{\mathbb{R}}} \Big(\int_{X_{\mathbb{R}}} (x-\omega)^2 \, [F_\tau(dx)](\omega) \Big) [\Phi_*(1_{\theta_i})](\omega) \, d\omega \Big)^{1/2} \quad (i = 1,2,\ldots,n). \end{aligned} \tag{18.10}$$

Finally, the group average of the student θ_i's error Δ_i $(i = 1,2,\ldots,n)$ is as follows:

$$\Delta_g := \Big(\frac{1}{n} \sum_{i=1}^n \Delta_i^2 \Big)^{1/2}. \tag{18.11}$$

From what we have seen, we can get the following theorem.

Theorem 18.8. (i: The variance $\mathrm{Var}[\mathsf{M}^{(i)}_{\mathsf{O}_\tau}]$) Let $\mathsf{M}^{(i)}_{\mathsf{O}_\tau} := \mathsf{M}_{L^\infty(\Omega_{\mathbb{R}}, d\omega)}(\mathsf{O}_\tau, S_{[*]}(\Phi_*(1_{\theta_i})))$ be the measurement of test observable O_τ for the statistical state $\Phi_*(1_{\theta_i})$. Then, we see

$$\mathrm{Var}[\mathsf{M}^{(i)}_{\mathsf{O}_\tau}] = \mathrm{Var}[\mathsf{M}^{(i)}_{\mathsf{O}_E}] + \Delta_i^2. \tag{18.12}$$

(ii: The variance $\mathrm{Var}[\mathsf{M}^{\otimes}_{\mathsf{O}_\tau}]$) We consider the group test $\mathsf{M}^{\otimes}_{\mathsf{O}_\tau} := \otimes_{\theta_i \in \Theta} \mathsf{M}^{(i)}_{\mathsf{O}_\tau} = \otimes_{\theta_i \in \Theta} \mathsf{M}_{L^\infty(\Omega_{\mathbb{R}}, d\omega)}(\mathsf{O}_\tau, \, S_{[*]}(\Phi_*(1_{\theta_i})))$. And, we obtain the following:

$$\mathrm{Var}[\mathsf{M}^{\otimes}_{\mathsf{O}_\tau}] = \mathrm{Var}[\mathsf{M}^{\otimes}_{\mathsf{O}_E}] + \Delta_g^2. \tag{18.13}$$

Proof. Let μ_i be an expectation of $\Phi_*(1_{\theta_i})$. Then, we see

$$\begin{aligned}\mathrm{Var}[\mathsf{M}^{(i)}_{\mathsf{O}_\tau}] &= \int_{\Omega_{\mathbb{R}}} \Big(\int_{X_{\mathbb{R}}} (x-\mu_i)^2 \, [F_\tau(dx)](\omega) \Big) [\Phi_*(1_{\theta_i})](\omega) \, d\omega \\ &= \int_{\Omega_{\mathbb{R}}} (\omega-\mu_i)^2 \, [\Phi_*(1_{\theta_i})](\omega) \, d\omega + \int_{\Omega_{\mathbb{R}}} \Big(\int_{X_{\mathbb{R}}} (x-\omega)^2 \, [F_\tau(dx)](\omega) \Big) [\Phi_*(1_{\theta_i})](\omega) \, d\omega \\ &\quad + \int_{\Omega_{\mathbb{R}}} \Big(\int_{X_{\mathbb{R}}} 2(x-\omega)(\omega-\mu_i) \, [F_\tau(dx)](\omega) \Big) [\Phi_*(1_{\theta_i})](\omega) \, d\omega \\ &= \mathrm{Var}[\mathsf{M}^{(i)}_{\mathsf{O}_E}] + \Delta_i^2. \end{aligned}$$

From the above formula, it follows that the group average of $\mathrm{Var}[\mathsf{M}^{(i)}_{\mathsf{O}_\tau}]$ becomes

$$\begin{aligned}\mathrm{Var}[\mathsf{M}^{\otimes}_{\mathsf{O}_\tau}] &= \int_{\Omega_{\mathbb{R}}} \cdots \int_{\Omega_{\mathbb{R}}} \Big(\int_{X_{\mathbb{R}}} \cdots \int_{X_{\mathbb{R}}} \frac{1}{n} \sum_{i=1}^n (x_i - \overline{\mu})^2 \mathop{\times}_{i=1}^n [F_\tau(dx_i)](\omega_i) \Big) \mathop{\times}_{i=1}^n [\Phi_*(1_{\theta_i})](\omega_i) \, d\omega_i \\ &= \frac{1}{n} \sum_{i=1}^n \int_{\Omega_{\mathbb{R}}} \Big(\int_{X_{\mathbb{R}}} (\omega - \overline{\mu} + x - \omega)^2 \, [F_\tau(dx)](\omega) \Big) [\Phi_*(1_{\theta_i})](\omega) \, d\omega \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{n}\sum_{i=1}^{n}\int_{\Omega_{\mathbb{R}}} (\omega-\overline{\mu})^2\,[\Phi_*(1_{\theta_i})](\omega)\,d\omega \\
&\quad + \frac{1}{n}\sum_{i=1}^{n}\int_{\Omega_{\mathbb{R}}} \Big(\int_{X_{\mathbb{R}}} (x-\omega)^2\,[F_\tau(dx)](\omega)\Big)[\Phi_*(1_{\theta_i})](\omega)\,d\omega \\
&\quad + \frac{1}{n}\sum_{i=1}^{n}\int_{\Omega_{\mathbb{R}}} \Big(\int_{X_{\mathbb{R}}} 2(x-\omega)(\omega-\overline{\mu})\,[F_\tau(dx)](\omega)\Big)[\Phi_*(1_{\theta_i})](\omega)\,d\omega \\
&= \int_{\Omega_{\mathbb{R}}}\cdots\int_{\Omega_{\mathbb{R}}} \frac{1}{n}\sum_{i=1}^{n}(\omega_i-\overline{\mu})^2 \mathop{\times}_{i=1}^{n}[\Phi_*(1_{\theta_i})](\omega_i)\,d\omega_i + \frac{1}{n}\sum_{i=1}^{n}\Delta_i^2 \\
&= \mathrm{Var}[\mathsf{M}^{\otimes}_{\mathsf{O}_E}] + \Delta_g^2. \qquad \square
\end{aligned}
$$

18.2 Correlation coefficient: How to calculate the reliability coefficient

In the previous section, we define the reliability coefficient $\mathrm{RC}[\mathsf{M}^{\otimes}_{\mathsf{O}_\tau}] := \frac{\mathrm{Var}[\mathsf{M}^{\otimes}_{\mathsf{O}_E}]}{\mathrm{Var}[\mathsf{M}^{\otimes}_{\mathsf{O}_\tau}]}$. However, from the measured data $(x_1, x_2, \ldots, x_n)$ $(\in X_{\mathbb{R}}^n)$, we can not get the variance of mathematical abilities of n students $\mathrm{Var}[\mathsf{M}^{\otimes}_{\mathsf{O}_E}]$ directly (though we can calculate the $\mathrm{Var}[\mathsf{M}^{\otimes}_{\mathsf{O}_\tau}]$). Thus, we focus on the problem how to estimate the reliability coefficient. Here we consider one typical method, say the split-half method.

Split-half method: This method is appropriate where the testing procedure may in some fashion be divided into two halves and two scores obtained. These may be correlated. With psychological tests, a common procedure is to obtain scores on the odd and even items.

Now we introduce the measurement theoretical characterizations of the split-half method.

Definition 18.9. [Group simultaneous test] Let $\Theta := \{\theta_1, \theta_2, \ldots, \theta_n\}$, $X_{\mathbb{R}} = \Omega_{\mathbb{R}} = \mathbb{R}$ and $\Phi_* : L^1_{+1}(\Theta, \nu_c) \to L^1_{+1}(\Omega_{\mathbb{R}}, d\omega)$ be as in Example 18.1. Let $\mathsf{O}_{\tau_1} := (X_{\mathbb{R}}, \mathcal{F}_{X_{\mathbb{R}}}, F_{\tau_1})$ and $\mathsf{O}_{\tau_2} := (X_{\mathbb{R}}, \mathcal{F}_{X_{\mathbb{R}}}, F_{\tau_2})$ be test observables in $L^\infty(\Omega_{\mathbb{R}}, d\omega)$. The measurement

$$\otimes_{\theta_i\in\Theta}\mathsf{M}_{L^\infty(\Omega_{\mathbb{R}},d\omega)}(\mathsf{O}_{\tau_1}\times\mathsf{O}_{\tau_2}, S_{[*]}(\Phi_*(1_{\theta_i}))),$$

is called a *group simultaneous test* of O_{τ_1} and O_{τ_2} and it is symbolized by $\mathsf{M}^{\otimes}_{\mathsf{O}_{\tau_1}\times\mathsf{O}_{\tau_2}}$ for short.

Axiom$^{(\mathrm{m})}$ 1 (§9.1) says that

(A) the probability that the score $((x_1^1, x_1^2), (x_2^1 . x_2^2), \ldots, (x_n^1, x_n^2))$ $(\in X_{\mathbb{R}}^{2n})$ obtained by the group simultaneous test $\otimes_{\theta_i \in \Theta} \mathsf{M}_{L^\infty(\Omega_{\mathbb{R}}, d\omega)}(\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}, S_{[*]}(\Phi_*(1_{\theta_i})))$ (or in short, $\mathsf{M}^{\otimes}_{\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}}$) belongs to the set $\times_{i=1}^n (\Xi_i^1 \times \Xi_i^2)$ $(\in \mathcal{F}_{X_{\mathbb{R}}^{2n}})$ is given by

$$\mathop{\times}_{\theta_i \in \Theta} {}_{L^1(\Omega_{\mathbb{R}}, d\omega)} \langle \Phi_*(1_{\theta_i}), (F_{\tau_1} \times F_{\tau_2})(\Xi_i^1 \times \Xi_i^2) \rangle_{L^\infty(\Omega_{\mathbb{R}}, d\omega)} \Big(=: \widehat{P}_2 (\mathop{\times}_{i=1}^{n} (\Xi_i^1 \times \Xi_i^2)) \Big). \tag{18.14}$$

Here note that $(X_{\mathbb{R}}^{2n}, \mathcal{F}_{X_{\mathbb{R}}^{2n}}, \widehat{P}_2)$ is a sample probability space.

Let $W_2 : X_{\mathbb{R}}^{2n} \to \mathbb{R}$ be a statistics (i.e., measurable function). Then, $\mathcal{E}_{\mathsf{M}^{\otimes}_{\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}}}[W_2]$, the expectation of W_2, is defined by

$$\mathcal{E}_{\mathbf{M}^{\otimes}_{\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}}}[W_2] = \int_{X_{\mathbb{R}}^n} W(x_1^1, x_1^2, x_2^1, x_2^2, \ldots, x_n^1, x_n^2)\, \widehat{P}_2(dx_1^1\, dx_1^2\, dx_2^1\, dx_2^2 \cdots dx_n^1\, dx_n^2).$$

We use the following notations:

$$\text{(i)}\ \ \mathrm{Av}^{(k)}[\mathsf{M}^{\otimes}_{\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}}] := \mathcal{E}_{\mathsf{M}^{\otimes}_{\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}}}\Big[\frac{1}{n}\sum_{i=1}^n x_i^k\Big] \qquad (k=1,2),$$

$$\text{(ii)}\ \ \mathrm{Var}^{(k)}[\mathsf{M}^{\otimes}_{\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}}] := \mathcal{E}_{\mathsf{M}^{\otimes}_{\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}}}\Big[\frac{1}{n}\sum_{i=1}^n (x_i^k - \mathrm{Av}^{(k)}[\mathsf{M}^{\otimes}_{\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}}])^2\Big] \qquad (k=1,2),$$

$$\begin{aligned}\text{(iii)}\ \ \mathrm{Cov}[\mathbf{M}^{\otimes}_{\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}}] := \mathcal{E}_{\mathbf{M}^{\otimes}_{\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}}}\Big[\frac{1}{n}\sum_{i=1}^n &(x_i^1 - \mathrm{Av}^{(1)}[\mathbf{M}^{\otimes}_{\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}}]) \\ &\times (x_i^2 - \mathrm{Av}^{(2)}[\mathbf{M}^{\otimes}_{\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}}])\Big].\end{aligned}$$

It is clear that $\mathrm{Av}^{(k)}[\mathsf{M}^{\otimes}_{\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}}] = \mathrm{Av}[\mathsf{M}^{\otimes}_{\mathsf{O}_{\tau_k}}] = \mathrm{Av}[\mathsf{M}^{\otimes}_{\mathsf{O}_E}]$ $(k=1,2)$.

Definition 18.10. [Equivalency of test observables] We call that test observables $\mathsf{O}_{\tau_1} := (X_{\mathbb{R}}, \mathcal{F}_{X_{\mathbb{R}}}, F_{\tau_1})$ and $\mathsf{O}_{\tau_2} := (X_{\mathbb{R}}, \mathcal{F}_{X_{\mathbb{R}}}, F_{\tau_2})$ in $L^\infty(\Omega_{\mathbb{R}}, d\omega)$ are *equivalent* if it holds

$$\Delta_\omega^{(1)} = \Delta_\omega^{(2)} \quad (\forall \omega \in \Omega_{\mathbb{R}}), \tag{18.15}$$

where $\Delta_\omega^{(k)} := (\int_{X_{\mathbb{R}}} (x-\omega)^2 \,[F_{\tau_k}(dx)](\omega))^{1/2}$ (see (18.9)).

In case that test observables $\mathsf{O}_{\tau_1} := (X_{\mathbb{R}}, \mathcal{F}_{X_{\mathbb{R}}}, F_{\tau_1})$ and $\mathsf{O}_{\tau_2} := (X_{\mathbb{R}}, \mathcal{F}_{X_{\mathbb{R}}}, F_{\tau_2})$ in $L^\infty(\Omega_{\mathbb{R}}, d\omega)$ are equivalent and $\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}$ is a product test observable in $L^\infty(\Omega_{\mathbb{R}}, d\omega)$, it holds that

$$\mathrm{Var}[\mathsf{M}^{\otimes}_{\mathsf{O}_{\tau_1}}] = \mathrm{Var}^{(1)}[\mathsf{M}^{\otimes}_{\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}}] = \mathrm{Var}^{(2)}[\mathsf{M}^{\otimes}_{\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}}] = \mathrm{Var}[\mathsf{M}^{\otimes}_{\mathsf{O}_{\tau_2}}]. \tag{18.16}$$

In consequence of these properties, we introduce the correlation coefficient of the measured values $(x_1^1, x_2^1, \ldots, x_n^1)$ $(\in X_{\mathbb{R}}^n)$ and $(x_1^2, x_2^2, \ldots, x_n^2)$ $(\in X_{\mathbb{R}}^n)$ which are obtained by the group simultaneous test $\mathsf{M}^{\otimes}_{\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}}$.

Theorem 18.11. [The reliability coefficient and the correlation coefficient in group simultaneous tests] Let O_{τ_1} and O_{τ_2} be equivalent test observables in $L^\infty(\Omega_{\mathbb{R}}, d\omega)$. And let $\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}$ be a product test observable in $L^\infty(\Omega_{\mathbb{R}}, d\omega)$. Let $\mathsf{M}^{\otimes}_{\mathsf{O}_{\tau_k}} := \otimes_{\theta_i \in \Theta} \mathsf{M}_{L^\infty(\Omega_{\mathbb{R}}, d\omega)}(\mathsf{O}_{\tau_k}, S_{[*]}(\Phi_*(1_{\theta_i})))$ $(k = 1, 2)$ and $\mathsf{M}^{\otimes}_{\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}} := \otimes_{\theta_i \in \Theta} \mathbf{M}(\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}, S_{[*]}(\Phi_*(1_{\theta_i})))$ be group tests as above notations. Then we see that

$$\mathrm{RC}[\mathbf{M}^{\otimes}_{\mathsf{O}_{\tau_1}}] = \mathrm{RC}[\mathbf{M}^{\otimes}_{\mathsf{O}_{\tau_2}}] = \frac{\mathrm{Cov}[\mathbf{M}^{\otimes}_{\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}}]}{\sqrt{\mathrm{Var}[\mathbf{M}^{\otimes}_{\mathsf{O}_{\tau_1}}]} \cdot \sqrt{\mathrm{Var}[\mathbf{M}^{\otimes}_{\mathsf{O}_{\tau_2}}]}}. \tag{18.17}$$

Proof. From the (18.3), we get the following:

$$\begin{aligned}
&\mathrm{Cov}[\mathbf{M}^{\otimes}_{\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}}] := \mathcal{E}_{\mathsf{M}^{\otimes}_{\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}}} \Big[\frac{1}{n} \sum_{i=1}^n (x_i^1 - \mathrm{Av}^{(1)}[\mathsf{M}^{\otimes}_{\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}}])(x_i^2 - \mathrm{Av}^{(2)}[\mathsf{M}^{\otimes}_{\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}}]) \Big] \\
&= \int_{\Omega_{\mathbb{R}}} \cdots \int_{\Omega_{\mathbb{R}}} \Big(\int_{X_{\mathbb{R}}} \cdots \int_{X_{\mathbb{R}}} \frac{1}{n} \sum_{i=1}^n (x_i^1 - \mathrm{Av}^{(1)}[\mathsf{M}^{\otimes}_{\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}}])(x_i^2 - \mathrm{Av}^{(2)}[\mathsf{M}^{\otimes}_{\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}}]) \\
&\quad \times \mathop{\times}_{i=1}^n [F_{\tau_1}(dx_i^1)\, F_{\tau_2}(dx_i^2)](\omega_i) \Big) \mathop{\times}_{i=1}^n [\Phi_*(1_{\theta_i})](\omega_i)\, d\omega_i \\
&= \frac{1}{n} \sum_{i=1}^n \Big(\int_{\Omega_{\mathbb{R}}} \Big(\int_{X_{\mathbb{R}}} \int_{X_{\mathbb{R}}} (x_i^1 - \mathrm{Av}[\mathsf{M}^{\otimes}_{\mathsf{O}_E}])(x_i^2 - \mathrm{Av}[\mathsf{M}^{\otimes}_{\mathsf{O}_E}]) \\
&\qquad\qquad \times [F_{\tau_1}(dx_i^1)](\omega)\, [F_{\tau_2}(dx_i^2)](\omega) \Big) [\Phi_*(1_{\theta_i})](\omega)\, d\omega \Big) \\
&= \frac{1}{n} \sum_{i=1}^n \Big(\int_{\Omega_{\mathbb{R}}} \Big(\int_{X_{\mathbb{R}}} (x_i^1 - \mathrm{Av}[\mathsf{M}^{\otimes}_{\mathsf{O}_E}])\, [F_{\tau_1}(dx_i^1)](\omega) \\
&\qquad\qquad \times \int_{X_{\mathbb{R}}} (x_i^2 - \mathrm{Av}[\mathsf{M}^{\otimes}_{\mathsf{O}_E}])\, [F_{\tau_2}(dx_i^2)](\omega) \Big) [\Phi_*(1_{\theta_i})](\omega)\, d\omega \Big) \\
&= \frac{1}{n} \sum_{i=1}^n \int_{\Omega_{\mathbb{R}}} (\omega - \mathrm{Av}[\mathsf{M}^{\otimes}_{\mathsf{O}_E}])^2\, [\Phi_*(1_{\theta_i})](\omega)\, d\omega = \mathrm{Var}[\mathsf{M}^{\otimes}_{\mathsf{O}_E}].
\end{aligned} \tag{18.18}$$

Then, we see that

$$\frac{\mathrm{Cov}[\mathbf{M}^{\otimes}_{\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}}]}{\sqrt{\mathrm{Var}[\mathbf{M}^{\otimes}_{\mathsf{O}_{\tau_1}}]} \cdot \sqrt{\mathrm{Var}[\mathbf{M}^{\otimes}_{\mathsf{O}_{\tau_2}}]}} = \frac{\mathrm{Var}[\mathsf{M}^{\otimes}_{\mathsf{O}_E}]}{\mathrm{Var}^{(1)}[\mathbf{M}^{\otimes}_{\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}}]} = \frac{\mathrm{Var}[\mathsf{M}^{\otimes}_{\mathsf{O}_E}]}{\mathrm{Var}^{(2)}[\mathbf{M}^{\otimes}_{\mathsf{O}_{\tau_1} \times \mathsf{O}_{\tau_2}}]}. \tag{18.19}$$

□

18.3 Conclusions

In this chapter, we introduce the measurement theoretical understanding of psychological test and the split-half method which estimate reliability. Measurement theoretical approach show

KSTS/RR-16/001
January 8, 2016

the following correspondences:

$$\text{split-half method} \longleftrightarrow \begin{array}{c}\text{group simultaneous test.}\\ \mathsf{M}^{\otimes}_{\mathsf{O}_{\tau_1}\times\mathsf{O}_{\tau_2}} := \otimes_{\theta_i\in\Theta}\mathsf{M}_{L^\infty(\Omega_{\mathbb{R}},d\omega)}(\mathsf{O}_{\tau_1}\times\mathsf{O}_{\tau_2}, S_{[*]}(\Phi_*(1_{\theta_i})))\end{array}$$

And further, we show the well-known theorem:

$$\text{“reliability coefficient”} = \text{“correlation coefficient”}$$

in Theorem 18.11.

Chapter 19

How to describe "belief"

Recall the spirit of quantum language (i.e., the spirit of the quantum mechanical world view), that is,

($\sharp$) every phenomenon should be described in quantum language.

Thus, we consider that even "belief" should be described in quantum language. For this, it suffices to consider the identification:

$$\text{"belief"} = \text{"odds by bookmaker"}$$

This approach has a great merit such that the principle of equal weight holds.

This chapter is extracted from Chapter 8 in

Ref. [29]: S. Ishikawa, "Mathematical Foundations of Measurement Theory," Keio University Press Inc. 2006.

19.1 Belief, probability and odds

For instance, we want to formulate the following "probability":

(A) the "probability" that Japan will win the victory in the next FIFA World Cup.

This is possible (*cf.* [29]), if "parimutuel betting (or, odds in bookmaker)" is formulated by Axiom$^{(m)}$ 1 (mixed measurement). The purpose of this chapter is to show it, and further, to propose the principle of equal weight, that is,

(B) *the principle that, in the absence of any reason to expect one event rather than another, all the possible events should be assigned the same probability.*

whose validity has not been proven yet. It is one of the most important unsolved problems in statistics.

In Chapter 9, we studied the mixed measurement: that is,

$$
\underset{\text{(=quantum language)}}{\boxed{\text{mixed measurement theory}}} := \underbrace{\overset{[\text{(mixed)Axiom}^{(m)}\ 1]}{\underset{(cf.\ \S 9.1\ )}{\boxed{\text{mixed measurement}}}} + \overset{[\text{Axiom 2}]}{\underset{(cf.\ \S 10.3)}{\boxed{\text{Causality}}}}}_{\text{a kind of spells (a priori judgment)}}
$$
$$
+ \underbrace{\overset{[\text{quantum linguistic interpretation}]}{\underset{(cf.\ \S 3.1)}{\boxed{\text{Linguistic interpretation}}}}}_{\text{manual to use spells}} \tag{19.1}
$$

The purpose of this chapter is to characterize "belief" as a kind of mixed measurement.

19.1.1 A simple example; how to describe "belief" in quantum language

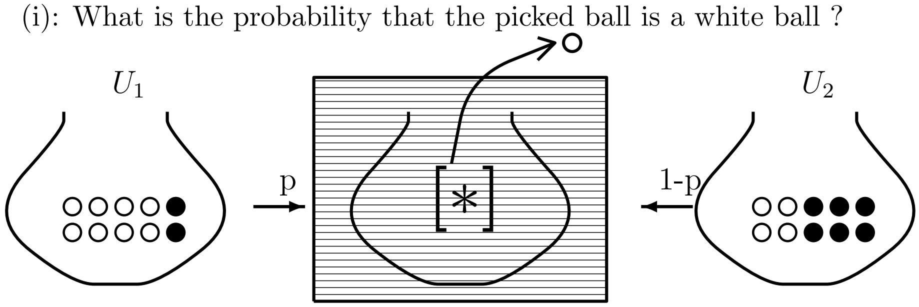
We begin with a simplest example (*cf.* Problem 9.5) as follows.

Problem 19.1. [= Problem 9.5; Bayes' method] Assume the following situation:

(C) You do not know which the urn behind the curtain is, U_1 or U_2, but the "probability": p and $1-p$.

Here, consider the following problem:

Assume that you pick up a ball from the urn behind the curtain.
(i): What is the probability that the picked ball is a white ball ?



(ii): If the picked ball is white, what is the probability that the urn behind the curtain is U_1 ?

Figure 19.1:(Mixed measurement)

Answer 19.2. (=Answer 9.13)
Put $\Omega = \{\omega_1, \omega_2\}$ with the discrete metric and the counting measure ν_c, thus, note that

$C_0(\Omega) = C(\Omega) = L^\infty(\Omega, \nu)$. Thus, in this chapter, we devote ourselves to the C^*-algebraic formulation: Define the observables $\mathsf{O} = (\{W,B\}, 2^{\{W,B\}}, F)$ and $\mathsf{O}_U = (\{U_1,U_2\}, 2^{\{U_1,U_2\}}, G_U)$ in $C(\Omega)$ by

$$F(\{W\})(\omega_1) = 0.8,\ F(\{B\})(\omega_1) = 0.2,\ F(\{W\})(\omega_2) = 0.4,\ F(\{B\})(\omega_2) = 0.6$$
$$G_U(\{U_1\})(\omega_1) = 1, G_U(\{U_2\})(\omega_1) = 0, G_U(\{U_1\})(\omega_2) = 0, G_U(\{U_2\})(\omega_2) = 1$$

Here "W" and "B" means "white" and "black" respectively. Under the identification: $U_1 \approx \omega_1$ and $U_2 \approx \omega_2$, the above situation is represented by the mixed state $\rho^{(p)}_{\mathrm{prior}} (\in \mathcal{M}_{+1}(\Omega))$ such that

$$\rho^{(p)}_{\mathrm{prior}} = p\delta_{\omega_1} + (1-p)\delta_{\omega_2},$$

where δ_ω is the point measure at ω. Thus, we have the mixed measurement:

$$\mathsf{M}_{C(\Omega)}(\mathsf{O} \times \mathsf{O}_U := (\{W,B\} \times \{U_1, U_2\}, 2^{\{W,B\}\times\{U_1,U_2\}}, F \times G_U), S_{[*]}(\rho^{(p)}_{\mathrm{prior}})). \tag{19.2}$$

Axiom$^{(m)}$ 1 gives the answer to the (i) in Problem 19.1 as follows.

(D) the probability that a measured value (x, y) obtained by the mixed measurement $\mathsf{M}_{C(\Omega)}(\mathsf{O} \times \mathsf{O}_U, S_{[*]}(\rho^{(p)}_{\mathrm{prior}}))$ belongs to $\{W\} \times \{U_1, U_2\}$ is given by

$$_{\mathcal{M}(\Omega)}(\rho^{(p)}_{\mathrm{prior}}, F(\{W\}))_{C(\Omega)} = 0.8p + 0.4(1-p).$$

Since a white ball is obtained, Answer 9.13 (=Bayes' theorem) says that a new mixed state $\rho^{(p)}_{\mathrm{post}} (\in \mathcal{M}_{+1}(\Omega))$ is given by

$$\rho^{(p)}_{\mathrm{post}} = \frac{F(\{W\})\rho^{(p)}_{\mathrm{prior}}}{\int_\Omega [F(\{W\})](\omega)\rho^{(p)}_{\mathrm{prior}}(d\omega)} = \frac{0.8p}{0.8p + 0.4(1-p)}\delta_{\omega_1} + \frac{0.4(1-p)}{0.8p + 0.4(1-p)}\delta_{\omega_2} \tag{19.3}$$

Hence, the answer of the (ii) is given by

$$_{\mathcal{M}(\Omega)}(\rho^{(p)}_{\mathrm{post}}, G_U(\{U_1\}))_{C(\Omega)} = \frac{0.8p}{0.8p + 0.4(1-p).}$$

By an analogy of the above Problem 19.1 (for simplicity, we put: $p = 1/4$), we consider as follows.

Assume that there are 100 people. And moreover assume the following situation (E) such that, for some reasons,

$$\text{(E)} \begin{cases} \text{25 people believe (or vote) that } [*] = U_1 \text{ (i.e., } U_1 \text{ is behind the curtain)} \\ \text{75 people believe (or vote) that } [*] = U_2 \text{ (i.e., } U_2 \text{ is behind the curtain)} \end{cases}$$

That is, we have the following picture instead of Figure 19.1:

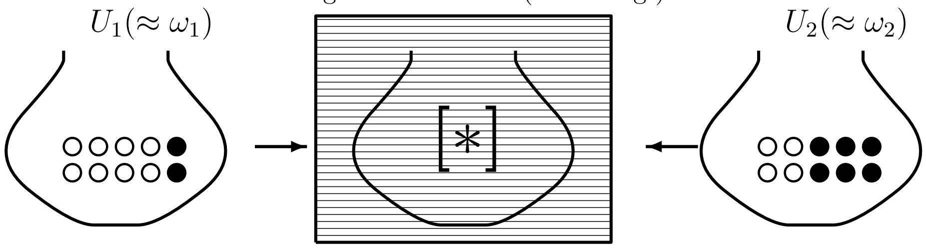


Figure 19.2: Belief (or voting)

25 people believe that $[*] = U_1$, 75 people believe that $[*] = U_2$.

Now, we have the following problem:

Problem 19.3. Consider Situation (E) and Situation (C) ($p = 1/4, \ \ 1-p = 3/4$). Then,

($\mathbf{F}_1$) Can Situation (E) be understood like Situation (C) ?

or, in the same sense,

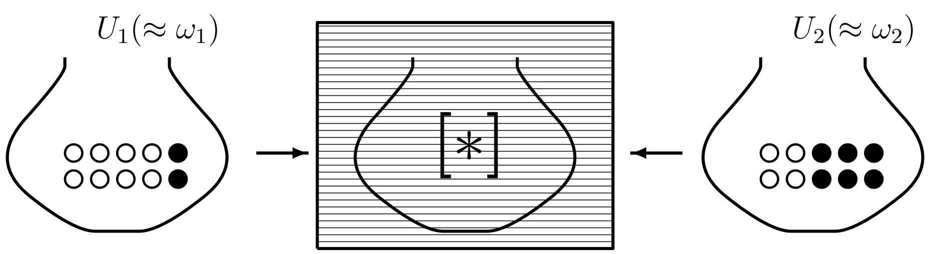
($\mathbf{F}_2$) Can Situation (E) be formulated in mixed measurement (i.e., Axiom$^{(m)}$ 1)? That is, can Situation (E) be described in quantum language ?

19.1.2 The affirmative answer to Problem 19.3

Since 100 people know the situation of the urn (i.e., Figure 19.2, the assumption (E)) implies (G)(=Figure 19.3), that is,

$$
\textbf{(G)} \left\{ \begin{array}{l}
\text{25 people (in 100 people) believe that } [*] = U_1 \\
\quad \Longrightarrow \left\{ \begin{array}{l} \text{(G}_1\text{): 20 people guess (or bet) that a white ball will be picked} \\ \text{(G}_2\text{): 5 people guess (or bet) that a black ball will be picked} \end{array} \right. \\
\text{75 people (in 100 people) believe that } [*] = U_2 \\
\quad \Longrightarrow \left\{ \begin{array}{l} \text{(G}_3\text{): 30 people guess (or bet) that a white ball will be picked} \\ \text{(G}_4\text{): 45 people guess (or bet) that a black ball will be picked} \end{array} \right.
\end{array} \right.
$$

Figure 19.3: The odds in bookmaker



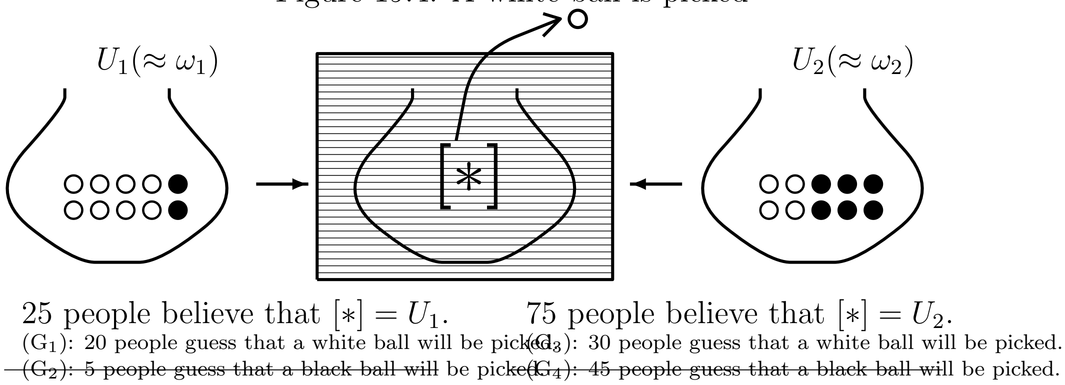
25 people believe that $[*] = U_1$. 75 people believe that $[*] = U_2$.

(G_1): 20 people guess that a white ball will be picked. (G_3): 30 people guess that a white ball will be picked.

(G_2): 5 people guess that a black ball will be picked. (G_4): 45 people guess that a black ball will be picked.

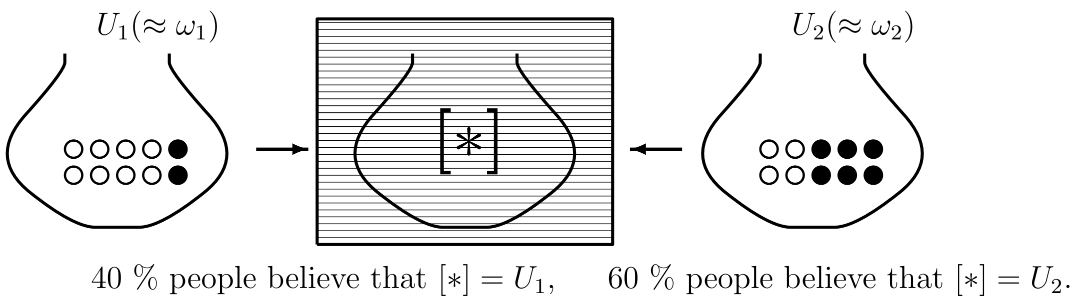
Assume that a white ball is picked in the above figure. Then, the above (G_2) and (G_4) are vanished as follows.

Figure 19.4: A white ball is picked



After all, we get the following figure:

Figure 19.5: After all, we get the new odds



Thus we see that

$$\underset{\frac{1}{4}\delta_{\omega_1}+\frac{3}{4}\delta_{\omega_2}}{\overset{\text{(prior state)}}{\boxed{\text{Fig. 19.3}}}} \longrightarrow \overset{\text{(a white ball is picked)}}{\boxed{\text{Fig. 19.4}}} \longrightarrow \underset{\frac{2}{5}\delta_{\omega_1}+\frac{3}{5}\delta_{\omega_2}}{\overset{\text{(post state)}}{\boxed{\text{Fig. 19.5}}}} \tag{19.4}$$

Considering the mixed measurement (i.e., the (19.2) in the case that $p = 1/4$):

$$\mathsf{M}_{C(\Omega)}(\mathsf{O} \times \mathsf{O}_U = (\{W,B\} \times \{U_1, U_2\}, 2^{\{W,B\}\times\{U_1,U_2\}}, F \times G_U), S_{[*]}(\rho^{(1/4)}_{\text{prior}})) \tag{19.5}$$

we see that the above (19.4) is the same as the Bayesian result (19.3).

Note that the measurement (19.5) is interpreted as

(H) choose one person from the 100 people at random, and ask him/her "Do you guess that a white ball (or, a black ball) will be picked from the urn behind the curtain, and its urn is U_1 or U_2 ?"

In what follows, let us explain it. Consider the product observable $\widehat{\mathsf{O}}\times\widehat{\mathsf{O}}_U$ of $\widehat{\mathsf{O}} = (\{W, B\}, 2^{\{W,B\}}, \widehat{F})$ and $\widehat{\mathsf{O}}_U = (\{U_1, U_2\}, 2^{\{U_1,U_2\}}, \widehat{G}_U)$ in $C(\Theta)$ (where $\Theta = \{\theta_1, \theta_2, ..., \theta_{100}\}$) such that

$$\begin{aligned} &[\widehat{F}(\{W\})](\theta_k) = 4/5, \quad [\widehat{F}(\{B\})](\theta_k) = 1/5, \quad (k = 1, 2, ..., 25) \\ &[\widehat{F}(\{W\})](\theta_k) = 2/5, \quad [\widehat{F}(\{B\})](\theta_k) = 3/5, \quad (k = 26, 27, ..., 100) \end{aligned} \tag{19.6}$$

$$
\begin{aligned}
&[\widehat{G}_U(\{U_1\})](\theta_k)=1, \quad [\widehat{G}_U(\{U_2\})](\theta_k)=0, \quad (k=1,2,...,25)\\
&[\widehat{G}_U(\{U_1\})](\theta_k)=0, \quad [\widehat{G}_U(\{U_2\})](\theta_k)=1, \quad (k=26,27,...,100)
\end{aligned}
\tag{19.7}
$$

And put $\nu_0 = (1/100)\sum_{k=1}^{100}\delta_{\theta_k}(\in \mathcal{M}_{+1}(\Theta))$. Then, the above measurement (H) is formulated by

$$
\mathsf{M}_{C(\Theta)}(\widehat{\mathsf{O}} \times \widehat{\mathsf{O}}_U = (\{W,B\}\times\{U_1,U_2\}, 2^{\{W,B\}\times\{U_1,U_2\}}, \widehat{F}\times\widehat{G}_U), S_{[*]}(\nu_0)) \tag{19.8}
$$

which is identified with the measurement (19.5) under the deterministic causal operator $\Phi : C(\Omega) \to C(\Theta)$ such that $\Phi^*(\delta_{\theta_k}) = \delta_{\omega_1}$ $(k=1,2,...,25)$, $=\delta_{\omega_2}$ $(k=26,27,...,100)$. That is, we see, symbolically,

$$
\boxed{\text{(H)=(19.8): the Heisenberg picture}} \xleftarrow[\text{identification}]{\Phi} \boxed{\text{(19.5): the Schrödinger picture}}
$$

Thus, as a particular case of the above arguments, we can answer Problem 19.3 such that

(I$_1$) Situation (E) can be understood like Situation (C).

That is,

(I$_2$) Situation (E) can be formulated in mixed measurement (i.e., Axiom$^{(m)}$ 1). In the same sense, Situation (E) can be described in quantum language.

19.2 The principle of equal odds weight

From the above arguments, we see that

Proclaim 19.4. [The principle of equal weight] Consider a finite state space Ω with the discrete metric, that is, $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$. Let $\mathsf{O} = (X, \mathcal{F}, F)$ be an observable in $C(\Omega)$. Consider a measurement $\mathsf{M}_{C(\Omega)}(\mathsf{O}, S_{[*]})$. If the observer has no information for the unknown state $[*]$, there is a reason to assume that this measurement is also represented by the mixed measurement $\mathsf{M}_{C(\Omega)}(\mathsf{O}, S_{[*]}(\rho_{\text{prior}}))$, where

$$
\rho_{\text{prior}} = \frac{1}{n}\sum_{k=1}^{n}\delta_{\omega_k}. \tag{19.9}
$$

Explanation. In betting, it is certain that everybody wants to choose an unpopular ω_k. Thus, I believe that everybody agrees with Proclaim 19.4. Also, it should be noted that

(J) the term "probability" can be freely used within the rule of Axiom 1 or Axiom$^{(m)}$ 1.

The reason that the justice of the (B: the principle of equal weight) is not assured yet is due to the lack of the understanding of the (J).

♠**Note 19.1.** In this book, we dealt with the following three kinds:

($\sharp_1$) the principle of equal weight in Remark 5.19

($\sharp_2$) the principle of equal weight in Theorem 9.18

($\sharp_3$) the principle of equal weight in Proclaim 19.4

which are essentially the same.

In order to promote the readers' understanding of the difference between Theorem 9.18 and Proclaim 19.4, we show the following example, which should be compared with Problem 5.14 and Problem 9.17

Problem 19.5. [Monty Hall problem (=Problem 5.14; The principle of equal weight)]

You are on a game show and you are given a choice of three doors. Behind one door is a car, and behind the other two are goats. You choose, say, door 1, and the host, who knows where the car is, opens another door, behind which is a goat. For example, the host says that

($\flat$) the door 3 has a goat.

And further, he now gives you a choice of sticking to door 1 or switching to door 2 ? *What should you do ?*

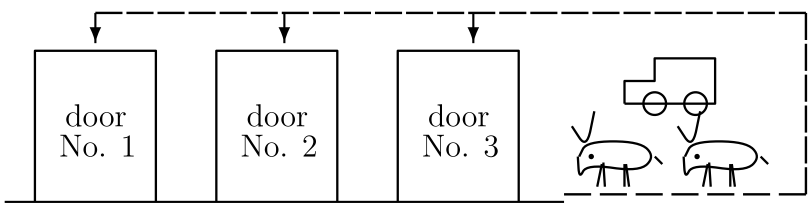


Figure 19.6: Monty Hall problem

Proof. It should be noted that the above is completely the same as Problem 5.14. However, the proof is different. That is, it suffices to use Proclaim 19.4 and Bayes theorem (B_2). That is, the proof is similar to Problem 9.16 . □

Chapter 20

Postscript

20.1 Two kinds of (realistic and linguistic) world-views

In this lecture note, we assert the following figure:

Figure 20.1. [=Figure 1.1: The location of quantum language in the history of world-description (*cf.* ref.[31])]

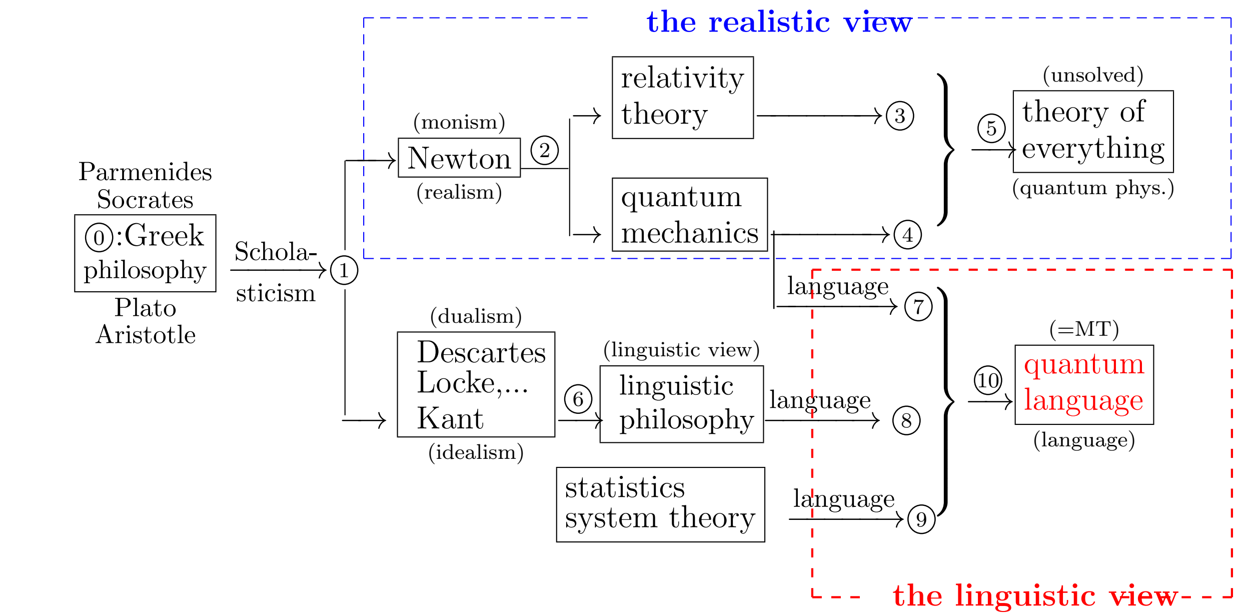


Figure 20.1: The history of the world-view

Most physicists feel that

(A_1) quantum mechanics has both realistic aspect and metaphysical aspect.

And they want to unify the two aspects. However, quantum language asserts that

(A_2) Two aspects are separated, and they develop in the respectively different directions ⑤ and ⑩ in Figure 20.1.

20.2 The summary of quantum language

20.2.1 The big-picture view of quantum language

The big-picture view of quantum language

Measurement theory (= quantum language) is classified as follows.

$$
\text{(B)}\ \underset{\text{(=quantum language)}}{\text{measurement theory}}
\begin{cases}
\underset{(\mathrm{B}_1)}{\text{pure type}} \begin{cases} \text{classical system : Fisher statistics} \\ \text{quantum system : usual quantum mechanics} \end{cases} \\
\\
\underset{(\mathrm{B}_2)}{\text{mixed type}} \begin{cases} \text{classical system : including Bayesian statistics, Kalman filter} \\ \text{quantum system : quantum decoherence} \end{cases}
\end{cases}
$$

And the structure is as follows.

$$
\text{(C)}
\begin{cases}
(\mathrm{C}_1): \underset{\text{(=quantum language)}}{\boxed{\text{pure measurement theory}}} \\
\quad := \underbrace{\overset{\text{[(pure)Axiom 1]}}{\underset{(cf.\ \S 2.7)}{\boxed{\text{pure measurement}}}} + \overset{\text{[Axiom 2]}}{\underset{(cf.\ \S 10.3)}{\boxed{\text{Causality}}}}}_{\text{a kind of spell(a priori judgment)}} + \underbrace{\overset{\text{[quantum linguistic interpretation]}}{\underset{(cf.\ \S 3.1)}{\boxed{\text{Linguistic interpretation}}}}}_{\text{the manual to use spells}} \\
\\
(\mathrm{C}_2): \underset{\text{(=quantum language)}}{\boxed{\text{mixed measurement theory}}} \\
\quad := \underbrace{\overset{\text{[(mixed)Axiom}^{(m)}\text{ 1]}}{\underset{(cf.\ \S 9.1)}{\boxed{\text{mixed measurement}}}} + \overset{\text{[Axiom 2]}}{\underset{(cf.\ \S 10.3)}{\boxed{\text{Causality}}}}}_{\text{a kind of spell(a priori judgment)}} + \underbrace{\overset{\text{[quantum linguistic interpretation]}}{\underset{(cf.\ \S 3.1)}{\boxed{\text{Linguistic interpretation}}}}}_{\text{the manual to use spells}}
\end{cases}
$$

In the above,

(D_1) **Axioms 1 and 2 (i.e., kinds of spells) are essential**

On the other hand, the linguistic interpretation (i.e., the manual to use Axioms 1 and 2) may not be indispensable. However,

(D_2) if we would like to make speed of acquisition of a quantum language as quick as possible, we may want the good manual to use the axioms.

In this sense, this note is a manual book (=cookbook). Although all written in this note can be regarded as a part of the linguistic interpretation, the most important statement is

Only one measurement is permitted

Also, since we assert that quantum language is the final goal of dualistic idealism (= Descartes=Kant philosophy) in Figure20.1, we think that

(E) Many philosophers' maxims and thoughts constitute a part of the linguistic interpretation

20.2.2 The characteristic of quantum language

Also, we see:

The characteristic of quantum language

(F_1) **Non-reality (metaphysics)**: Quantum language is metaphysics (= language), which asserts the linguistic world-view.

(F_2) **The collapse of wave function does not occur**: According to the linguistic interpretation (i.e., only one measurement is permitted), we can not get information after the measurement. That is, the collapse of wave function can not be found. However, the projection postulate holds in the sense of Postulate 11.6.

(F_3) **Non-deterministic**: Since we usually consider non-deterministic processes in classical system, it is natural to assume non-deterministic processes (i.e., quantum decoherence) in quantum language.

(F_4) **Dualism**: The two concepts: "measurement" and "dualism" are non-separable. Thus, quantum language says

($\sharp$) describe any monistic phenomenon in the dualistic language !

(F_5) **Non-locality, faster-than-light**: Quantum language accepts "non-locality". This is the only one paradox in quantum language.

(F_6) **Many paradoxes and unsolved problems are clarified**:

(a) Paradoxes and unsolved problems due to a lack of quantum language:
What is probability (causality, space-time) ? Zeno's paradox, the principle of equal probability, classical syllogizm, classical Bell's inequlity

(b) Paradoxes and unsolved problems solved by descriptive power of quantum language:
Schrödinger's cat

(c) What we cannot speak about we must pass over in silence:
Heisenberg's uncertainty principle (due to the thought experiment by γ-ray microscope), Cogit proposition, Wigner's friend, delayed choice experiment

(d) Everything should be spoken by quantum language:
Several problems in statistics (Fisher's maximum likelihood method, Bayes method, semi-distance (confidence interval, statistical hypothesis, ANOVA), regression analysis, Kalman filter)

20.3 Quantum language is located at the center of science

Dr. Hawking said in his best seller book [17]:

(G) *Philosophers reduced the scope of their inquiries so much that Wittgenstein the most famous philosopher this century, said "The sole remaining task for philosophy is the analysis of language." What a comedown from the great tradition of philosophy from Aristotle to Kant!*

I think that this is not only his opinion but also most scientists' opinion. And moreover, I mostly agree with him. However, I believe that it is worth reconsidering the series in the linguistic world view (①–⑥–⑧–⑩ in Figure 20.1).

It is a matter of course that quantum language is different from pure mathematics. Hence, in spite of Lord Kelvin's saying: *Mathematics is the only good metaphysics* , I assert that

(H_1) **quantum language is located at the center of science**

That is, I believe, from the pure theoretical point of view, that quantum language will replace statistics.

Since quantum language is not physics but language (= metaphysics), quantum language (= the linguistic interpretation of quantum mechanics) is completely different from other interpretations. In this sense, I am convinced that

(H_2) **quantum language is forever**,

even if someone discovers the "final" interpretation of quantum mechanics in the realistic view (i.e., ⑤ in Figure 20.1).

I hope that my proposal will be examined from various view-points.

Shiro ISHIKAWA

January in 2016

References ([ ]⋆ is fundamental)

[1] Alexander, H. G., ed. *The Leibniz-Clarke Correspondence,* Manchester University Press, 1956.

[2] Arthurs, E. and Kelly, J.L.,Jr. *On the simultaneous measurement of a pair of conjugate observables,* Bell System Tech. J. **44**, 725-729 (1965)

[3] Aspect, A, Dallibard, J. and Roger, G. *Experimental test of Bell inequalities time-varying analysis,* Physical Review Letters 49, 1804–1807 (1982)

[4] Bell, J.S. *On the Einstein-Podolosky-Rosen Paradox,* Physics 1, 195–200 (1966)

[5] Bohr, N. *Can quantum-mechanical description of physical reality be considered complete?,* Phys. Rev. (48) 696-702 1935

[6] Born, M. *Zur Quantenmechanik der Stoßprozesse (Vorläufige Mitteilung),* Z. Phys. (37) 863–867 1926

[7] Busch, P. *Indeterminacy relations and simultaneous measurements in quantum theory,* International J. Theor. Phys. **24**, 63-92 (1985)

[8] G. Caella, R.L. Berger, *Statistical Inference,* Wadsworth and Brooks, 1999.

[9] D.J. Chalmers, *The St. Petersburg Two-Envelope Paradox,* Analysis, Vol.62, 155-157, 2002.

[10] F. Click, *The Astonishing Hypothesis: The Scientific Search For The Soul,* New York: Charles Scribner's Sons., 1994.

[11] Davies, E.B. *Quantum theory of open systems,* Academic Press 1976

[12] de Broglie, L. *L'interpretation de la mecanique ondulatoire,* Journ. Phys. Rad. 20, 963 (1959)

[13] Einstein, A., Podolosky, B. and Rosen, N. *Can quantum-mechanical description of reality be considered completely?* Physical Review Ser 2(47) 777–780 (1935)

[14] R. P. Feynman *The Feynman lectures on Physics; Quantum mechanics* Addison-Wesley Publishing Company, 1965

[15] G.A. Ferguson, Y. Takane, *Statistical analysis in psychology and education* (Sixth edition). New York: McGraw-Hill. (1989)

[16] L. Hardy, *Quantum mechanics, local realistic theories, and Lorentz-invariant realistic theories,* Physical Review Letters 68 (20): 2981-2984 1992

[17] Hawking, Stephen *A brief History of Time,* Bantam Dell Publishing Group 1988

[18] Heisenberg, W. *Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik,* Z. Phys. 43, 172–198 (1927)

[19] Holevo, A.S. *Probabilistic and statistical aspects of quantum theory,* North-Holland publishing company (1982)

[20] D.Howard, Who invented the “Copenhagen Interpretation”? A study in mythology, Philosophy of Science, 71 2004, 669-682

[21] Isaac, R. *The pleasures of probability,* Springer-Verlag (Undergraduate texts in mathematics) 1995

[22]* S. Ishikawa, *Uncertainty relation in simultaneous measurements for arbitrary observables,* Rep. Math. Phys., **9**, 257-273, 1991
doi: 10.1016/0034-4877(91)90046-P

[23] Ishikawa, S. *Uncertainties and an interpretation of nonrelativistic quantum theory,* International Journal of Theoretical Physics 30 401–417 (1991) doi: 10.1007/BF00670793

[24] Ishikawa, S., Arai, T. and Kawai, T. *Numerical Analysis of Trajectories of a Quantum Particle in Two-slit Experiment,* International Journal of Theoretical Physics, Vol. 33, No. 6, 1265-1274, 1994
doi: 10.1007/BF00670793

[25]* Ishikawa,S. *Fuzzy inferences by algebraic method,* Fuzzy Sets and Systems 87, 181–200 (1997) doi:10.1016/S0165-0114(96)00035-8

[26]* S. Ishikawa, *A Quantum Mechanical Approach to Fuzzy Theory,* Fuzzy Sets and Systems, Vol. 90, No. 3, 277-306, 1997, doi: 10.1016/S0165-0114(96)00114-5

[27] S. Ishikawa, T. Arai, T. Takamura, *A dynamical system theoretical approach to Newtonian mechanics*, Far east journal of dynamical systems 1, 1-34 (1999)
(http://www.pphmj.com/abstract/191.htm)

[28]* S. Ishikawa, *Statistics in measurements*, Fuzzy sets and systems, Vol. 116, No. 2, 141-154, 2000
doi:10.1016/S0165-0114(98)00280-2

[29]* S. Ishikawa, *Mathematical Foundations of Measurement Theory,* Keio University Press Inc. 335pages, 2006, (http://www.keio-up.co.jp/kup/mfomt/)

[30]* S. Ishikawa, *A New Interpretation of Quantum Mechanics,* Journal of quantum information science, Vol. 1, No. 2, 35-42, 2011, doi: 10.4236/jqis.2011.12005
(http://www.scirp.org/journal/PaperInformation.aspx?paperID=7610)

[31]* S. Ishikawa, *Quantum Mechanics and the Philosophy of Language: Reconsideration of traditional philosophies,* Journal of quantum information science, Vol. 2, No. 1, 2-9, 2012
doi: 10.4236/jqis.2012.21002
(http://www.scirp.org/journal/PaperInformation.aspx?paperID=18194)

[32] S. Ishikawa, *A Measurement Theoretical Foundation of Statistics,* Applied Mathematics, Vol. 3, No. 3, 283-292, 2012, doi: 10.4236/am.2012.33044
(http://www.scirp.org/journal/PaperInformation.aspx?paperID=18109&)

[33] S. Ishikawa, *Monty Hall Problem and the Principle of Equal Probability in Measurement Theory, Applied Mathematics*, Vol. 3 No. 7, 2012, pp. 788-794, doi: 10.4236/am.2012.37117.
(http://www.scirp.org/journal/PaperInformation.aspx?PaperID=19884)

[34] S. Ishikawa, *Ergodic Hypothesis and Equilibrium Statistical Mechanics in the Quantum Mechanical World View, World Journal of Mechanics*, Vol. 2, No. 2, 2012, pp. 125-130. doi: 10.4236/wim.2012.22014.
(http://www.scirp.org/journal/PaperInformation.aspx?PaperID=18861#.VKevmiusWap)

[35]* S. Ishikawa, *The linguistic interpretation of quantum mechanics,*
arXiv:1204.3892v1[physics.hist-ph],(2012) (http://arxiv.org/abs/1204.3892)

[36] S. Ishikawa, *Zeno's paradoxes in the Mechanical World View,* arXiv:1205.1290v1 [physics.hist-ph], (2012)

[37] S. Ishikawa, *What is Statistics?; The Answer by Quantum Language,* arXiv:1207.0407 [physics.data-an] 2012. (http://arxiv.org/abs/1207.0407)

[38]* S. Ishikawa, *Measurement Theory in the Philosophy of Science,* arXiv:1209.3483 [physics.hist-ph] 2012. (http://arxiv.org/abs/1209.3483)

[39] S. Ishikawa, *Heisenberg uncertainty principle and quantum Zeno effects in the linguistic interpretation of quantum mechanics,* arxiv:1308.5469[quant-ph],(2013)

[40] S. Ishikawa, *A quantum linguistic characterization of the reverse relation between confidence interval and hypothesis testing,* arxiv:1401.2709[math.ST],(2014)

[41] S. Ishikawa, *ANOVA (analysis of variance) in the quantum linguistic formulation of statistics,* arxiv:1402.0606[math.ST],(2014)

[42] S. Ishikawa, *Regression analysis in quantum language,* arxiv:1403.0060[math.ST],(2014)

[43] S. Ishikawa, K. Kikuchi: *Kalman filter in quantum language,* arXiv:1404.2664 [math.ST] 2014. (http://arxiv.org/abs/1404.2664)

[44] S. Ishikawa, *The double-slit quantum eraser experiments and Hardy's paradox in the quantum linguistic interpretation,* arxiv:1407.5143[quantum-ph],(2014)

[45] S. Ishikawa, *The Final Solutions of Monty Hall Problem and Three Prisoners Problem,* arXiv:1408.0963 [stat.OT] 20 14. (http://arxiv.org/abs/1408.0963)

[46] S. Ishikawa, *Two envelopes paradox in Bayesian and non-Bayesian statistics* arXiv:1408.4916v4 [stat.OT] 2014. (http://arxiv.org/abs/1408.4916)

[47]* S. Ishikawa, Linguistic interpretation of quantum mechanics; Quantum Language, Research Report KSTS/RR-15/001, Dept. Math. Keio University, 416 p.
(http://www.math.keio.ac.jp/en/academic/research.html)
or, (http://www.math.keio.ac.jp/academic/research_pdf/report/2015/15001.pdf)

[48]* S. Ishikawa, Linguistic Interpretation of Quantum Mechanics –Towards World-Description in Quantum Language– Shiho-Shuppan Publisher, 409 p. (2015)
(http://www.shiho-shuppan.com/index.php?LIQM)

[49]* S. Ishikawa, *Linguistic interpretation of quantum mechanics; Projection Postulate,* Journal of quantum information science, Vol. 5, No.4 , 150-155, 2015, DOI: 10.4236/jqis.2015.54017
(http://www.scirp.org/Journal/PaperInformation.aspx?PaperID=62464)

[50] K. Kikuchi, S. Ishikawa, *Psychological tests in Measurement Theory, Far east journal of theoretical statistics,* 32(1) 81-99, (2010) ISSN: 0972-0863

[51] K. Kikuchi,, *Axiomatic approach to Fisher's maximum likelihood method,* Non-linear studies, 18(2) 255-262, (2011)

[52] Kalman, R. E. *A new approach to linear filtering and prediction problems,* Trans. ASME, J. Basic Eng. 82, 35 (1960)

[53] I. Kant, *Critique of Pure Reason (Edited by P. Guyer, A. W. Wood),* Cambridge University Press, 1999

[54] A. Kolmogorov, *Foundations of the Theory of Probability (Translation),* Chelsea Pub Co. Second Edition, New York, 1960,

[55] U. Krengel, "Ergodic Theorems," Walter de Gruyter. Berlin, New York, 1985.

[56] Lee, R. C. K. *Optimal Estimation, Identification, and Control,* M.I.T. Press 1964

[57] G. Lüders, Über die Zustandsänderung durch den Messprozess, Ann. Phys. (Leipzig) (6)8,322-328, 1951

[58] J. M. E. McTaggart, *The Unreality of Time,* Mind (A Quarterly Review of Psychology and Philosophy), Vol. 17, 457-474, 1908

[59] G. Martin, *Aha! Gotcha: Paradoxes to Puzzle and Delight* Freeman and Company, 1982

[60] B. Misra and E. C. G. Sudarshan, *The Zeno's paradox in quantum theory,* Journal of Mathematical Physics 18 (4): 756-763 (1977)

[61] N.D. Mermin, *Boojums all the way through, Communicating Science in a Prosaic Age,* Cambridge university press, 1994.

[62] Ozawa, M. *Quantum limits of measurements and uncertainty principle,* in Quantum Aspects of Operational Communication edited by Bendjaballah et all. Springer, Berlin, 3–17, (1991)

[63] M. Ozawa, *Universally valid reformation of the Heisenberg uncertainty principle on noise and disturbance in measurement,* Physical Review A, Vol. 67, pp. 042105-1–042105-6, 2003,

[64] Prugovečki, E. *Quantum mechanics in Hilbert space,* Academic Press, New York. (1981).

[65] Robertson, H.P. *The uncertainty principle,* Phys. Rev. 34, 163 (1929)

[66] D. Ruelle, "Statistical Mechanics, Rigorous Results," World Scientific, Singapore, 1969.

[67] Sakai, S. *C^*-algebras and W^*-algebras,* Ergebnisse der Mathematik und ihrer Grenzgebiete (Band 60), Springer-Verlag, Berlin, Heidelberg, New York 1971

[68] Selleri, F. *Die Debatte um die Quantentheorie,* Friedr. Vieweg&Sohn Verlagsgesellscvhaft MBH, Braunschweig (1983)

[69] Shannon, C.E., Weaver. W *A mathematical theory of communication,* Bell Syst. Tech.J. 27 379–423, 623–656, (1948)

[70] von Neumann, J. *Mathematical foundations of quantum mechanics* Springer Verlag, Berlin (1932)

[71] S. P.Walborn, et al. *"Double-Slit Quantum Eraser,"* Phys.Rev.A 65, (3), 2002

[72] J. A. Wheeler, *The 'Past' and the 'Delayed-Choice Double-Slit Experiment',* pp 9-48, in A.R. Marlow, editor, Mathematical Foundations of Quantum Theory, Academic Press (1978)

[73] Wittgenstein, L *Tractatus Logico-philosophicus,* Oxford: Routledge and Kegan Paul, 1921

[74] Yosida, K. *Functional analysis,* Springer-Verlag (Sixth Edition) 1980

Index

Notation

Department of Mathematics
Faculty of Science and Technology
Keio University

Research Report

2015

[15/001] Shiro Ishikawa,
Linguistic interpretation of quantum mechanics: Quantum Language,
KSTS/RR-15/001, January 22, 2015

[15/002] Takuji Arai, Ryoichi Suzuki,
Local risk-minimization for Lévy markets,
KSTS/RR-15/002, February 2, 2015

[15/003] Kazuma Teramoto, Takashi Nodera,
Lanczos type method for computing PageRank,
KSTS/RR-15/003, March 9, 2015

[15/004] Yoichi Matsuo, Takashi Nodera,
Block symplectic Gram-Schmidt method,
KSTS/RR-15/004, March 9, 2015

[15/005] Yuto Yokota, Takashi Nodera,
The L-BFGS method for nonlinear GMRES acceleration,
KSTS/RR-15/005, March 9, 2015

[15/006] Takatoshi Nakamura, Takashi Nodera,
The flexible incomplete LU preconditioner for large nonsymmetric linear systems,
KSTS/RR-15/006, April 13, 2015

[15/007] Takuro Kutsukake, Takashi Nodera,
The deflated flexible GMRES with an approximate inverse preconditioner,
KSTS/RR-15/007, April 15, 2015

[15/008] Dai Togashi, Takashi Nodera,
The GKB-GCV method for solving the general form of the Tikhonov regularization,
KSTS/RR-15/008, September 29, 2015

[15/009] Shiro Ishikawa,
The projection postulate in the linguistic interpretation of quantum mechanics,
KSTS/RR-15/009, November 8, 2015

2016

[16/001] Shiro Ishikawa,
Linguistic interpretation of quantum mechanics: Quantum Language [*Ver. 2*],
KSTS/RR-16/001, January 8, 2016