Linguistic Copenhagen interpretation of quantum mechanics: Quantum Language [Ver. 7]

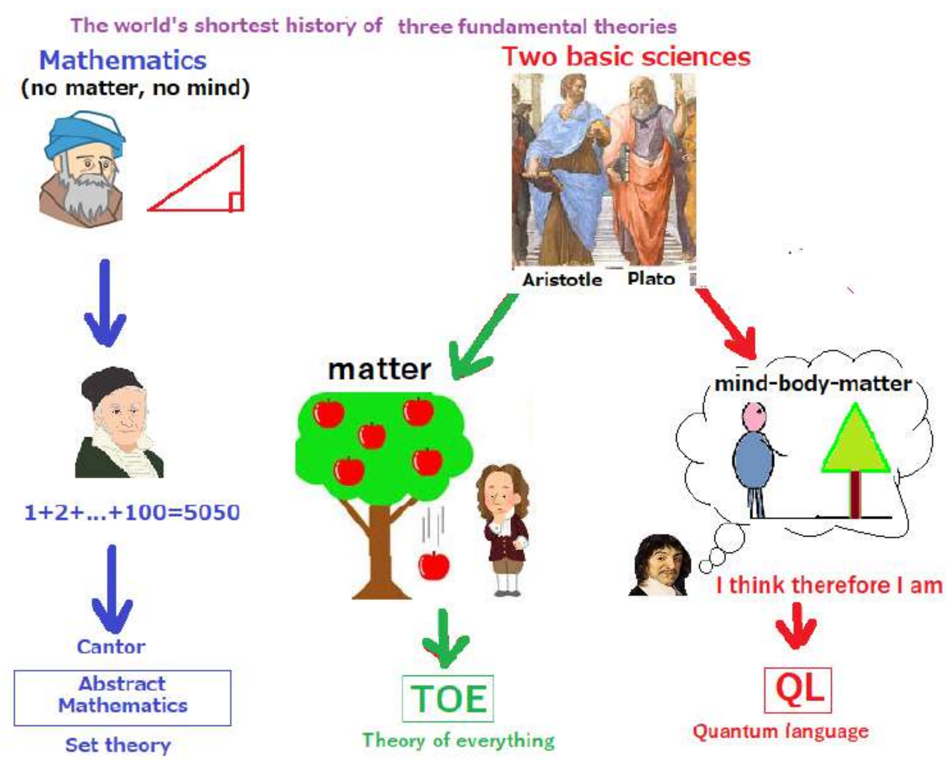
Shiro ISHIKAWA (ishikawa@math.keio.ac.jp) [20 July 2024]

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

Abstract Recently I proposed "QL (=quantum language)" (or,"the linguistic Copenhagen interpretation of quantum theory"), which was not only characterized as the metaphysical and linguistic turn of quantum mechanics but also as the scientific understanding of Descartes=Kant epistemology. Namely, quantum language is the scientific final goal of dualistic idealism. It has a great power to describe classical systems as well as quantum systems. In this research report, quantum language is seen as a fundamental theory of statistics and reveals the true nature of statistics. I hope the readers to enjoy the beautiful world of dualistic idealism.

///

The following diagram sums up my point:



I would like you to read this preprint with this figure in mind

Preface QL (=Quantum language), a language for talking science

QL (=Quantum language) is a mathematical abstraction of the language of quantum mechanics. I argue throughout this book that quantum language is the most powerful language of science, that is, it is not only the language of quantum mechanics but also the language of classical systems (i.e., everyday science). This language is located as illustrated in the following figure. This implies that quantum language is the scientific destination of dualistic idealism, and also, from a scientific perspective, the history of western philosophy can be almost regarded as the history of the pursuit of scientific language (i.e., Socrates' absolutism, *cf.* ref. [80]).

0.1 Two aspects (QM Ⓔ, QLⓁ) of quantum theory

Figure 0.1 [The locations of statistics and (classical) QL in the history of scientific world-description (*cf.* refs.[42, 63, 76, 80])]

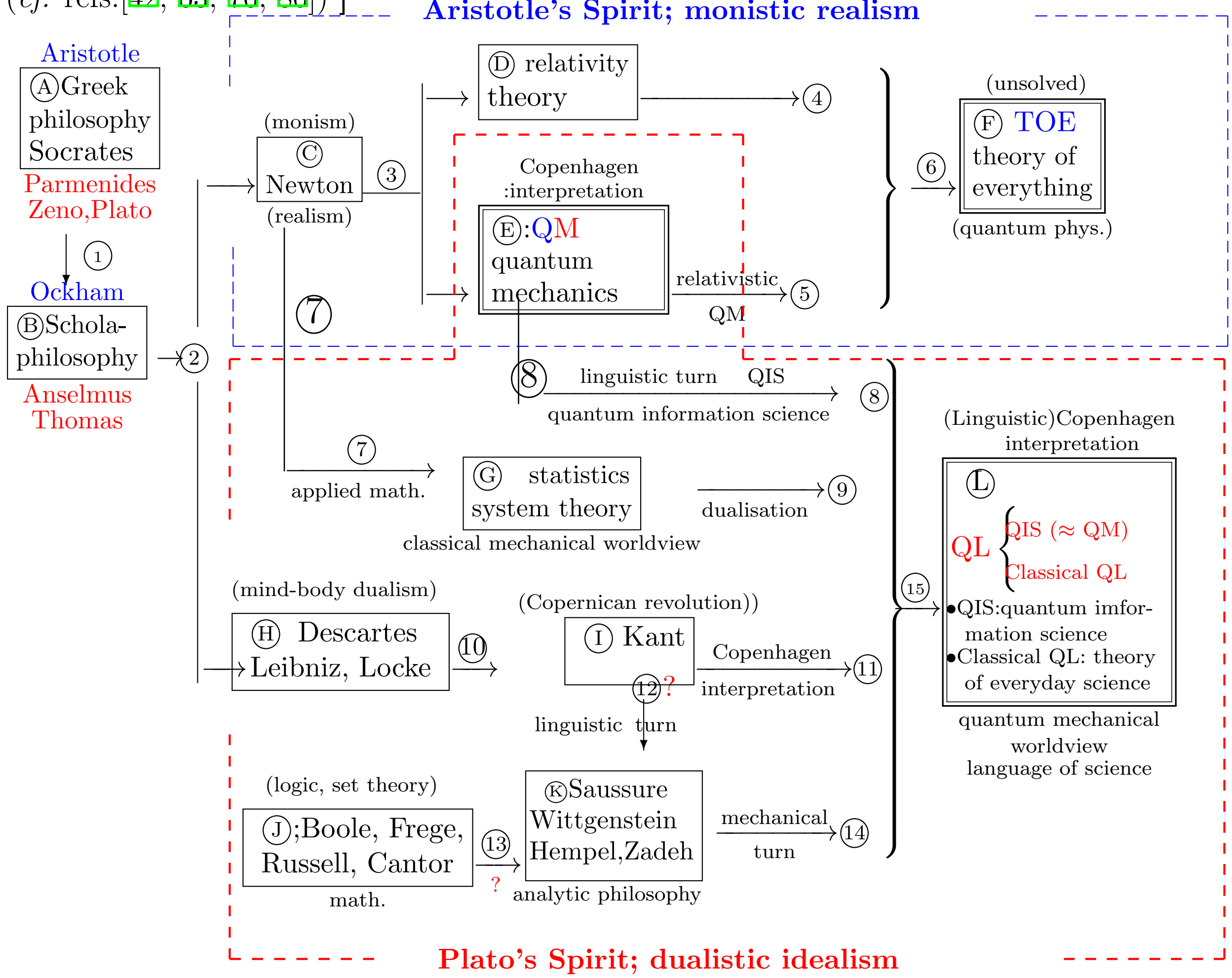


Figure 0.1 (in Preface); The history of the scientific world-descriptions

(*cf.* ref. [42, 63, 76, 80]): **Philosophy (≈ dualistic idealism) has progressed toward QL**

(i.e., the main stream of western philosophy history: Ⓐ→ Ⓑ→ Ⓗ→ Ⓘ→ Ⓚ→ Ⓛ)[1]

[1] For [linguistic turn ⑫?] and [$\xrightarrow[?]{⑬}$], see Note 0.1 (vi).

For further information see my homepage

I concluded that

(A) from a scientific perspective (i.e. from the standpoint of the perfection of Socrates' absolutism (*cf.* ref. [63, 76, 80])), 'progress' can be introduced into the history of Western philosophy.

That is,

(A') If "to make progress" is defined by "to come near quantum language" (i.e., "becoming more and more like quantum language")† we can say that the time series [⓪ - ① - ② - ⑩- ⑫- ⑭] can be regarded as progress, that is,

Or, almost equivalently, this means that Socrates' dream has come true by QL (*cf.* ref. [80])

Note 0.1 Some additional information on Figure 0.1 is provided below. Here, QM: quantum mechanics, QL: quantum language, RQM: relativistic quantum mechanics, QIS: quantum information science,

(i) The main theme of this book is the following:

$$\underset{\text{the classical mechanical wordview}}{\boxed{\text{statistics}}} \xrightarrow[\text{dualisation}]{⑧,⑨,⑮} \underset{\text{the quantum mechanical wordview}}{\boxed{\text{QL}}}$$

(ii) For a detailed discussion of the main stream of western philosophy history [⓪ - ① - ② - ⑩- ⑫- ⑭], see ref. [76].

(iii) Roughly speaking, I (not a philosopher) think

- realistic: 'thing' first, 'theory (≈ language)' later. (e.g., Newtonian mechanics)
- idealistic: 'theory (≈ language)' first, 'thing' later. (e.g., statistics)

However, when we normally use Newtonian mechanics or statistics, we are not particularly aware of the above difference. Therefore, this book does not emphasize the difference between 'idealism' and 'realism' too strongly (see Note 2.7). For example, quantum language is idealism and quantum mechanics is realism, but if one knows quantum language, one can use quantum mechanics freely.

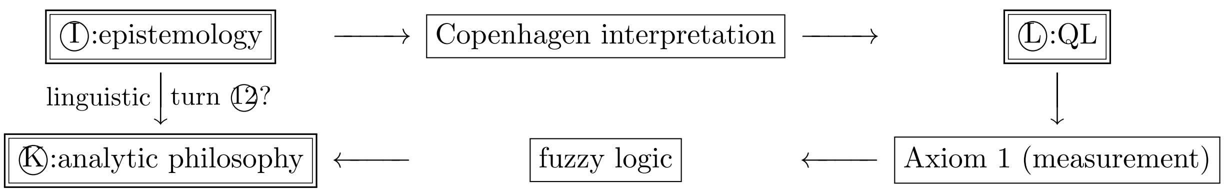
(iv) I think that

- $\overset{\text{the Copenhagen interpretation}}{\boxed{\text{QM in Ⓔ}}} \approx \overset{\text{the linguistic Copenhagen interpretation}}{\boxed{\text{QIS in Ⓛ}}}$

- $\overset{\text{the linguistic Copenhagen interpretation}}{\boxed{\text{QL in Ⓛ}}} = \overbrace{\boxed{\text{QSI in Ⓛ}} + \boxed{\text{classical QL in Ⓛ}}}^{\text{the linguistic Copenhagen interpretation}}$

Thus, we can use quantum mechanics if we know the linguistic Copenhagen interpretation without knowing the Copenhagen interpretation. Rather, we consider the linguistic Copenhagen interpretation to be the true Copenhagen interpretation. That is, I consider that there was no so-called Copenhagen interpretation[2].

(v) The linguistic turn [⑫?] in Figure 0.1 does not mean that Kantian philosophy (≈ "Copenhagen interpretation") influenced analytic philosophy (Wittgenstein). QL clarified the relation between Kantian epistemology and analytic philosophy such as



(*cf.* ref. [80]).

(vi) If we close our eyes to the historical background and think about it from a purely theoretical point of view, I don't think mathematical logic and analytic philosophy are completely related (*cf.* ref. [76]). That is because I believe that no worldview can come from mathematics. On the question of whether the most important key word in analytic philosophy is 'logical' or 'scientific', I take side of 'scientific'. (I think that mathematical logic is a language of mathematics, not science). I think that

- $\begin{cases} \text{mathematics} \cdots \text{logic} \\ \text{science} \begin{cases} \text{classical mechanical worldview} \cdots \text{causality} \\ \text{quantum mechanical worldview} \cdots \text{measurement + causality} \end{cases} \end{cases}$

Since my interest is science, not mathematics, I have marked '?' as in $\underset{?}{\overset{⑬}{\longrightarrow}}$.

If QL is seen as a philosophy, its slogan is "From 'Be logical!' to 'Be scientific!'". Through the problems of the flagpole and Hempel's ravens, Hempel cast doubt on 'Be logical!' (*cf.* ref. [42, 63, 76]).

(vii) I am not familiar with RQM (=relativistic quantum mechanics) and $\boxed{\boxed{\text{TOE in ⓑ}}}$.

0.2 Socrates's dream come true!

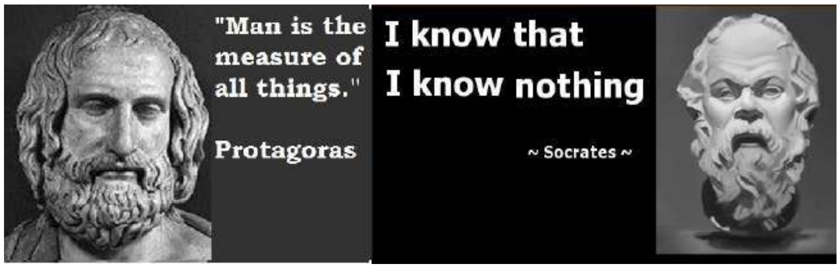
Let me say a few words about Socrates' absolutism (*cf.* ref. [80]).

[2]As discussed in Sec. 3.1.1, I think that there are a lot of 'so-called Copenhagen interpretations', that is, the established Copenhagen interpretation does not exist. On the other hand, the linguistic Copenhagen interpretation is expected to be uniquely determined.

For further information see my homepage

We adopt the general convention of considering Socrates as the founder of philosophy. Therefore, we have:

Socrates (absolutism: pursuit of truth) vs. Protagoras (relativism: mastering rhetoric)



In ancient Athens, it was customary for citizens to gather in the agora, a public square, to freely debate. So how did one "win an argument"?

- Protagoras, the relativist, responded to this question by saying "improve your rhetoric skills"
- Socrates, the absolutist, said "speak the truth"

If I were in the agora, I would probably agree with Protagoras, but that's not where philosophy begins. So Socrates' disciples pursued the question, "What is absolute truth?" This pursuit has continued through

- Plato, Aristotle, Augustine, Anselm, Thomas Aquinas, Descartes, Kant, and Wittgenstein,

and has formed the mainstream of Western philosophical history. However, despite being pursued by the most brilliant geniuses of every era for the past 2,500 years, no clear answer has yet been found. That's why some people, like Rorty (the flag bearer of neo-pragmatism), say, "Let's give up on the pursuit of truth here." If Rorty says something like that, I would think that Rorty's opinion is correct, but still, the stubborn pursuit of true If so, I think everyone would agree with the following:

- The most important problem in Western philosophy is the completion of Socratic absolutism, i.e., the end of the mainstream (Plato, Aristotle, Augustine, Anselm, Thomas Aquinas, Descartes, Kant, Wittgenstein).

And my answer is as follows.

- Next for Wittgenstein is QL, which is the perfection of Socratic absolutism.

Some readers might be thinking:

- Why are there no names like Spinoza, Hegel, Nietzsche, Husserl, Heidegger, Sartre etc?

The reason is simple: their achievements are not scientific (=QL). That does not mean that they are philosophically inferior to QL philosophers. However, the spirit that permeates the mainstream of Western philosophy is "scientific."

Thanks to QL we can say 'what Wittgenstein wanted to say' as follows.

- **What we cannot speak about in QL, we must pass over in silence.**

0.3 The purpose of this paper [3]

Note that Figure 0.1 says that QL is structured as shown in the following diagram.

Figure 0.2: Several fields of QL (i.e., quantum mechanical worldview

The four disciplines (Analytic philosophy, Descartes-Kant epistemology, quantum mechanics and statistic) are not separate disciplines but four aspects of quantum language as follows.

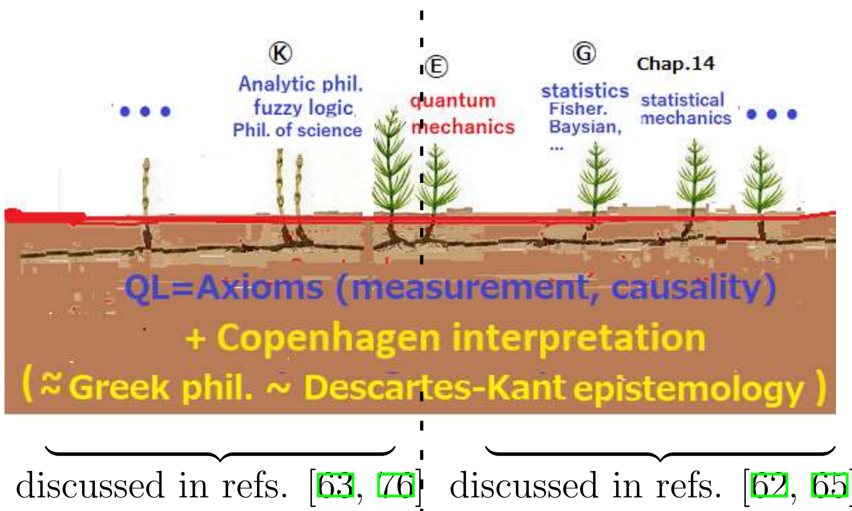


Figure 0.2: The relation among Analytic philosophy, Descartes-Kant epistemology, quantum mechanics and statistic

And

Figure 0.3 : [QL] = [QM(=QIS)] $\bigcup$ [classical QL(=everyday science)]

[3]This book was written under the assumption that it is a 3nd edition of my book [64]. However, the name "linguistic Copenhagen interpretation" is more often used these days than the name "linguistic interpretation". Therefore, I have used "linguistic Copenhagen interpretation" as the title of this book. With the change of title, it is no longer possible to add [3nd edition] to the book title. Also, ref. [62] is used as a draft of this book.

i.e.,

$$\text{QL}\begin{cases}\text{QM}(=\text{QIS}) \\ \\ \text{classical QL } (= \text{everyday science})\end{cases}$$

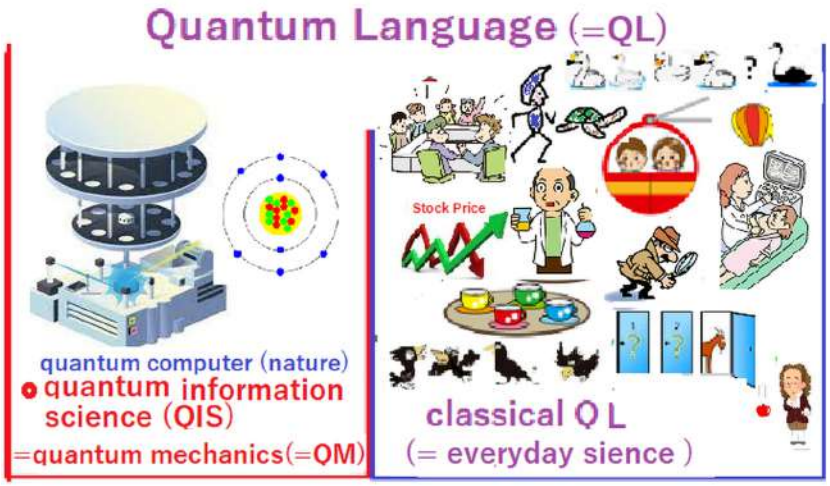


Figure 0.3: [QL] = [QIS(=QM)] $\bigcup$ [classical QL]

Thus, the purpose of this book is to assert that

($\sharp_1$) Statistics can be formulated in QL (= measurement theory)

Or, more precisely,

($\sharp_2$) When thinking in quantum language, the results of statistics can be used for the computational part.[4]

And thus, I would like to assert that

- For the question 'Why does statistics work in our world?', I would like to answer 'That is because QL works in our world'.

or equivalently,

- classical QL is the theory of everyday science.

Our argument is not common sense. Common sense would dictate that "the fundamental spirit of science is a mechanistic worldview". But our claim is that

(A_1) "the fundamental spirit of science is a quantum-mechanical worldview".

or equivalently,

(A_2) "QL is a language of science".

where we mean that 'science=non-relativistic science'.

[4] As statistics is a vast discipline, it is impossible to achieve this objective with this book alone. Therefore, my real aim is to convince readers that "statistics can be formulated in QL". And to have each reader write papers showing that various methods of statistics can be described in quantum language. Many readers may be more familiar with statistics than I am, so they may have found Chapter 6, for example, insufficient. If so, I would like to see this deepened further. As this is an area of 'quick wins', I think readers could write many papers on it.

Contents

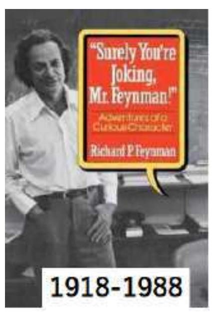
For further information see my homepage

For further information see my homepage

For further information see my homepage

Chapter 1

Nobody understands quantum mechanics (by R. Feynman)



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.

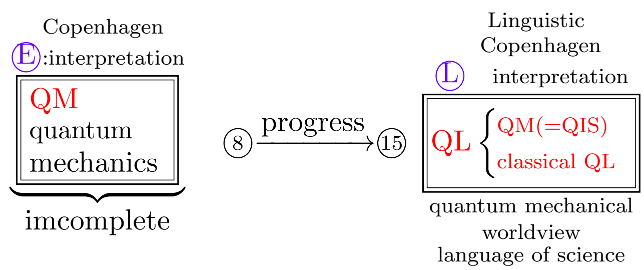
and

($\sharp_2$) 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.

As Feynman says, the 'lofty essence' of quantum mechanics may have to be left to the geniuses of the future.

[1]The importance of the two ($\sharp_1$) and ($\sharp_2$) was emphasized in Mermin's book [94].

However, there are many aspects of quantum mechanics. In particular, the perspective of viewing quantum mechanics as a fundamental theory of ‘everyday science’ can double the range of applications of quantum mechanics such as ⑧ $\xrightarrow{\text{progress}}$ ⑮ in Figure 0.1 in Preface, that is,

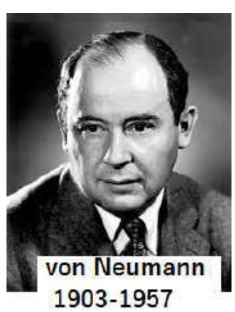


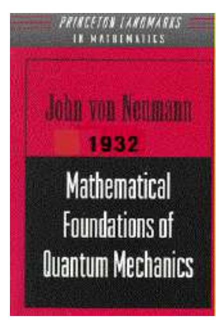
As will be discussed in Sec. 3.1.1, I think that there are a lot of ‘so-called Copenhagen interpretations’, that is, the established Copenhagen interpretation does not exist. In this sense, I think QM is incomplete, and the linguistic Copenhagen interpretation in QL [Ⓛ] is the only correct Copenhagen interpretation. This is precisely what we are trying to do in this publication.

1.1 Outline of quantum language

The quantum language is a mathematical abstraction of the language of quantum mechanics. I argue throughout this book that quantum language is the most powerful language of science, that is, it is not only the language of quantum mechanics but also the language of everyday science.

1.1.1 Von Neumann’s quantum theory





Various ‘interpretations ’ of quantum mechanics have been proposed. Examples include the Copenhagen interpretation and the many-worlds interpretation. Furthermore, there are various ‘versions ’ of the Copenhagen interpretation.

The ‘linguistic Copenhagen interpretation ’ of this book is a kind o ‘Copenhagen interpretation derived from von Neumann ’ s formulation of quantum mechanics on Hilbert spaces (*cf.* ref. [110]). Throughout this book, I argue that the linguistic Copenhagen interpretation is the true Copenhagen interpretation.

Von Neumann had the widest coverage of any mathematician of his time, integrating pure and applied sciences and making major contributions to many fields, including mathematics, physics, economics, computing, and statistics. He was not a genius who specialized only in mathematics and physics, but an all-round genius.

From this fact we are tempted to expect the following.

($\sharp_1$) the 'quantum theory' generated from the linguistic Copenhagen interpretation (which is called 'quantum language') is a very large theory that includes not only quantum mechanics of physics but also classical statistics.

Or, more generally, we may say

($\sharp_2$) Quantum language is the scientific realization of the dualistic idealism of philosophy.

In this book, I devote myself to proving ($\sharp_1$). (For ($\sharp_2$), see my previous book [76]). Quantum language consists of two axioms (measurement and causality) and the linguistic Copenhagen interpretation. I first prove 'von Neumann-Lüders projection postulate' in QL. This is a solution in QL, and whether it is a physical solution is undecided, but the theorem allows quantum language to be discussed without being plagued by various paradoxes (e.g., Schrödinger's cat, etc.).

Also, recall that there are no axioms in statistics. This fact means that we do not yet have 'theoretical statistics'. However, if we consider that 'quantum language for classical systems = theoretical statistics', we can then introduce an 'elegant understanding ' into statistics. The aim of this book can therefore be seen as a proposal for theoretical statistics.

Throughout this book, I assert that

- **Von Neumann's formulation of quantum mechanics should not be confined in physics, but should be regarded as a fundamental theory of science.**

1.1.2 Classification of quantum language

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

(A) measurement theory (=quantum language)
- pure type (A_1)
  - classical system : Fisher statistics
  - quantum system : usual quantum mechanics
- mixed type (A_2)
  - classical system : including Bayesian statistics and Kalman filter
  - quantum system : quantum decoherence

Here, we have two kinds of quantum languages, i.e., pure measurement theory and mixed measurement theory (or, statistical measurement theory). The former is formulated as

(A_1) [pure measurement theory] (=quantum language) := [(pure)Axiom 1] [pure measurement] (*cf.* §2.7) + [Axiom 2] [(deterministic) Causality] (*cf.* §9.3) — a kind of spells (a priori judgment)

+ [Linguistic Copenhagen interpretation] (*cf.* § 3.1) — manual to use spells

and the latter as

(A_2) [mixed measurement theory] (=quantum language) := [(mixed)Axiom$^{(m)}$ 1] [mixed measurement] (*cf.* §8.1) + [Axiom 2] [Causality] (*cf.* §9.3) — a kind of spells (a priori judgment)

+ [Linguistic Copenhagen interpretation] (*cf.* §3.1) — manual to use spells

1.1.3 Axiom 1 (measurement) and Axiom 2 (causality) in (A_1)

Let us sketch what is implied in the most fundamental case of pure measurement theory (A_1). This scheme involves two axioms, Axiom 1 for measurement and Axiom 2 for causality.

(B):Axiom 1 (measurement) pure type

(This can be read in §2.7)

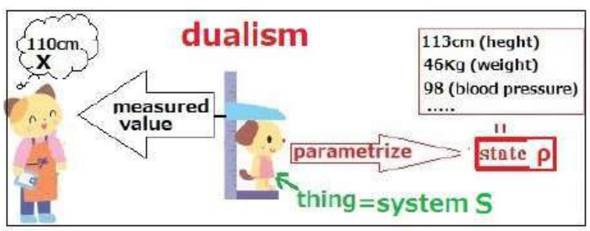
With any system S, a basic structure $[\mathcal{A} \subseteq \overline{\mathcal{A}}]_{B(H)}$ can be associated in which measurement theory of the system can be formulated. In $[\mathcal{A} \subseteq \overline{\mathcal{A}}]_{B(H)}$, consider a *$W^*$-measurement* $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}{=}(X,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

* a W^*-measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}, S_{[\rho]})$ (or, C^*-measurement $\mathsf{M}_{\mathcal{A}}(\mathsf{O}{=}(X, \mathcal{F}, F), S_{[\rho]})$) 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 (2.55) in Definition 2.18).

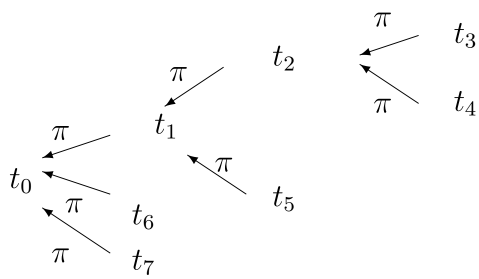


Max Born (1882–1970)

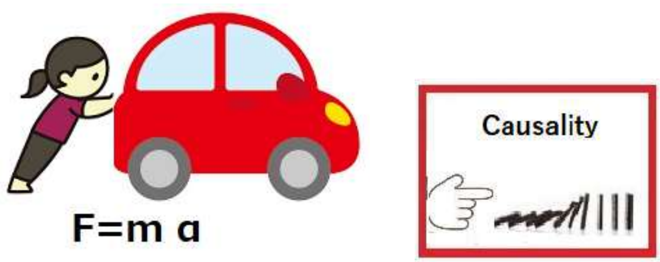
(C): Axiom 2 (causality)

(This can be read in §9.3)

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* $\{\Phi_{t_1,t_2} : \mathcal{A}_{t_2} \to \mathcal{A}_{t_1}\}_{(t_1,t_2)\in T^2_{\leqq}}$ $\Big)$



Later Figure 9.2: Tree: $(T = \{t_0, t_1, ..., t_7\}, \pi : T \setminus \{t_0\} \to T)$



Note that

(D$_1$) *the two axioms are a kind of spells (i.e., incantation, magic words, metaphysical statements), so that it is impossible to verify them experimentally.*

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" requires us to understand mathematical definitions of the 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 1.1.** 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 1.2.** 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.

William Thomson (=Lord Kelvin), was a British mathematician, mathematical physicist. Absolute temperatures are stated in units of kelvin in his honor.

1.1.4 The linguistic Copenhagen interpretation

Many theories have the following form.

$$\boxed{\text{Theory}} = \boxed{\text{Axiom (=Principle)}} + \boxed{\text{Interpretation}}$$

For example, in our society, too, it is not enough for a law to have a text alone; the law only works when there is a set of interpretations of it.

Axioms 1 and 2 are the most fundamental. But they are not all. That is,

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 define as follows.

$$\text{the linguistic Copenhagen interpretation} \underset{\text{Def(1)}}{:=} \text{the manual how to use Axioms 1 and 2.}$$ [2]

[2] Also, in Chap. 3, we introduce another definition:

$$\text{the linguistic Copenhagen interpretation} \underset{\text{Def(2)}}{:=} \text{common knowledge in the world of dualistic idealism}$$

I prefer it to Def(1). However, I am devoted to Def(1) here, since Def(1) is understandable.

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

(E):The linguistic Copenhagen interpretation)

(This will be explained in §3.1)

Only one measurement is permitted.

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

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

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

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

♠**Note 1.3.** You may have the following questions.

($\sharp_0$) Why is Newtonian mechanics (or statistics) without a stated 'interpretation'?

This question is profound. I think as follows.

($\sharp_1$) in the case of Newtonian mechanics, the interpretation is almost self-evident. That is, in Newtonian mechanics, we can do well without explicit interpretation.

($\sharp_2$) In the case of statistics, the distinction between 'axiom' or 'mathematics' is blurred. And thus, the boundary between 'Axiom' and 'Interpretation' is not clear. Thus, in statistics, the term 'interpretation' is not usually used.

If so, the following problem is the most fundamental and important in the field of theoretical statistics:

($\flat$) Rewrite statistics in the following format:

$$\boxed{\text{Statistics}} = \boxed{\text{Axiom (=Principle)}} + \boxed{\text{Interpretation}}$$

This problem is important. I think that statistics is generally regarded as a kind of applied mathematics, because the problem ($\flat$) is not yet answered. In this book, this problem ($\flat$) will be automatically solved since "statistics $\subset$ QL" will be studied.

1.1.5 Remarks

Let's take some precautions.

Remark 1.1. (i): It is easier to think that use of two phrases ('the Copenhagen interpretation' and 'linguistic Copenhagen interpretation') should be defined as follows.

($\sharp_1$) the phrase 'so-called Copenhagen interpretation' is used in QM (quantum mechanics) in ⓐ in Figure 0.1 (in Preface)

($\sharp_2$) the phrase 'the linguistic Copenhagen interpretation' is used in QL (quantum language) in ⓒ in Figure 0.1 (in Preface)

However, I think that the linguistic Copenhagen interpretation is the true Copenhagen interpretation. Therefore, in this book, I frequently use the term 'Copenhagen interpretation' in the sense of the linguistic Copenhagen interpretation.
(ii): As mentioned in Preface, my purpose is to propose the theory of everyday science (= classical QL) such as

$$\text{QL}\begin{cases}\text{QM(=QIS)} \\ \\ \text{classical QL (= everyday science)}\end{cases}$$

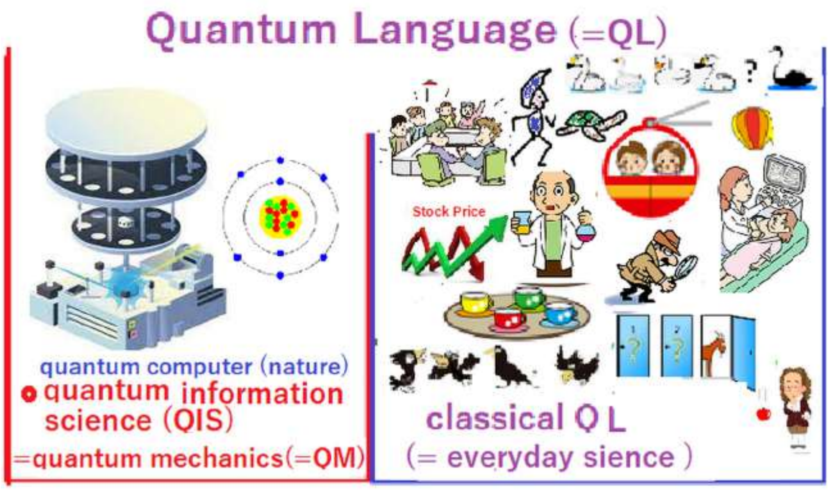


Figure 1.1(=Figure 0.3) : [QL] = [QM(=QIS)] $\bigcup$ [classical QL(everyday science)]

Therefore, my true purpose may be to introduce 'Copenhagen interpretation' to classical QL. To do this, it is necessary to have a good knowledge of the Copenhagen interpretation of quantum mechanics.

Remark 1.2. QL (i.e., the linguistic Copenhagen interpretation) has various advantages, two of which are mentioned here.

($\sharp_1$) About classical QL (i.e., $\mathcal{A}$ is commutative, especially, statistics):
Statistics is a very useful theory with a wide range of applications, but it lacks beauty, as the theory is regarded as 'a piece of applied mathematics'. But, seeing [statistics $\subseteq$ classical QL], we can regard statistics as the theory of dualism.
That is, I believe that

statistics such as 'statistics ⊆ classical QL' is 'true statistics'

Therefore, we can completely solve the problem (♭) in Note 1.3

(♯$_2$) About quantum QL:
Von Neumann-Lüders projection postulate can be proved in quantum QL. I believe that this implies that the linguistic Copenhagen interpretation is the true interpretation of quantum mechanics. That is, I believe that

the linguistic Copenhagen interpretation is 'true Copenhagen interpretation'

1.2 Example: Bald man paradox

Axioms 1 and 2 mentioned later may be too abstract to use quantum language now. So, let me show a simple example. The following example may promote your understanding of QL without the knowledge of Axioms 1 and 2.

♠**Note 1.4.** Readers may ask the following question:

(♯$_1$) Where does the (linguistic) Copenhagen interpretation lie in the above?

This question is similar to

(♯$_2$) Where does the (linguistic) Copenhagen interpretation lie in statistics (and theory of probability)?

In other words, the (linguistic) Copenhagen interpretation is quite difficult to find in classical QL. In most cases of classical systems, without the (linguistic) Copenhagen interpretation, we can do well. However, somewhat difficult problems (e.g. the Monty Hall problem, Zeno's paradox, Kolmogorov's extension theorem (*cf.* Sec. 4.1) cannot be solved without the Copenhagen interpretation. For example, the next question is quite educational for the current reader.

(♯) What measurement is assumed in the Monty Hall problem? (*cf.* Sec. 5.5)

Example 1.3. [Bald man paradox]
For simplicity, consider the basic structure $C(\Omega)$, where the state space Ω = the closed interval $[0,1] (\subset \mathbb{R}$: real line).

Let's assume that the maximum number of hairs on the head of adult men is 150 000(= 1.5×10^5). Let M be a set of all adult men. For any $m_i (\in M)$, define his 'bald rate $\omega(m_i)$' by

$$\omega(m_i) = 1 - \frac{\text{Number of hairs on Mr. } m_i\text{'s head.}}{1.5 \times 10^5}$$

Put $\Omega = [0,1]$. Define the 'bald observable' $\mathsf{O}_B = (\{Y, N\}, 2^{\{Y,N\}}, F_B)$ in $C(\Omega)$ such that

$$[F_B(\{Y\})](\omega) = \begin{cases} 0 & (0 \le \omega \le 0.3) \\ \frac{5}{2}\omega - \frac{3}{4} & (0.3 \le \omega \le 0.7) \\ 1 & (0.7 \le \omega \le 1.0) \end{cases}$$

$$[F_B(\{N\})](\omega) = 1 - [F_B(\{Y\})](\omega) \qquad (0 \le \omega \le 1)$$

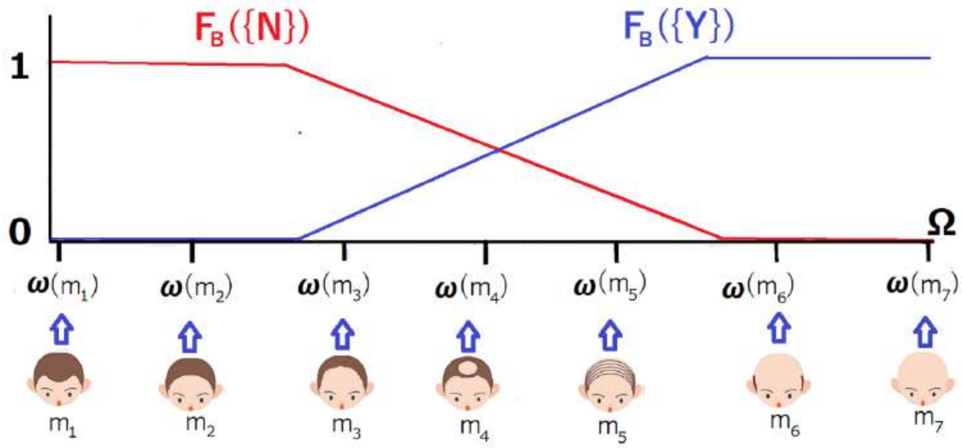


Figure 2: Bald observable $\mathsf{O}_B = (\{Y, N\}, 2^{\{Y,N\}}, F_B)$ in $C([0,1])$

Further, suppose that there are 100 respondents, and furthermore, the following question is asked to them

(G$_1$) Is Mr. m_i (with the bald rate $\omega(m_i)$) bald or not?

Assume that the results of the responses are as follows.

$$
(\mathrm{G}_2) \begin{cases} 100[F_B(\{Y\})](\omega(m_i)) \text{ respondents say "Yes, Mr. } m_i \text{ is bald"} \\ \\ 100[1 - F_B(\{Y\})](\omega(m_i)) \text{ respondents say "No, Mr. } m_i \text{ is not bald"} \end{cases}
$$

This can be probabilistically interpreted as follows.

(G$_3$) When a respondent is *randomly* selected out of 100, the probability that this respondent answer "yes" to question (D$_1$) is $p_1 = [F_B(\{Y\})](\omega(m_i)))$.
(Here, note that 'probability' can be created by 'ratio' + 'at random')

which is equivalent to

(G$_4$) Consider the measurement $\mathsf{M}_{C(\Omega)}(\mathsf{O} = (\{Y, N\}, 2^{\{Y,N\}}, F_B), S_{[\omega(m_i)]})$. Then, the probability that a measured value Y is obtained is given by $[F_B(\{Y\})](\omega(m_i))$.

Remark 1.4. I think the above argument is almost identical to 'Zadeh's Fuzzy set' argument (*cf.* ref. [115];(1965)), which is one of the most cited papers of the 20th century. I therefore believe that the 'bald man paradox', unresolved for 2500 years, has been resolved by Zadeh. Zadeh's late paper [116](2008) is written some criticisms of fuzzy theory by Kalman and others fairly. Obviously, his fuzzy theory can never beat statistics. For fuzzy theory to be generally accepted, it must be formulated within QL (a theory more powerful and beautiful than statistics). This has been my policy since the beginning when I proposed QL (*cf.* [29, 30, 31, 33]). However, these papers of mine did not move public opinion, but they did not give up. In this paper, I present the following illustration several times and hope that my arguments will be accepted by the reader.

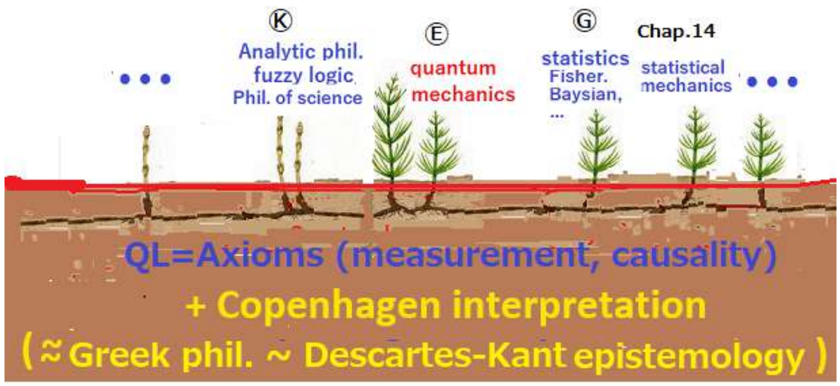


Figure 0.2: Several fields of QL (i.e., quantum mechanical worldview

Chapter 2

Axiom 1 – measurement

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

- $\underset{\text{(=quantum language)}}{\boxed{\text{measurement theory}}} := \underbrace{\underset{(cf.\ \S2.7)}{\overset{\text{[Axiom 1]}}{\boxed{\text{Measurement}}}} + \underset{(cf.\ \S9.3)}{\overset{\text{[Axiom 2]}}{\boxed{\text{(deterministic) Causality}}}}}_{\text{a kind of spells (a priori judgment)}}$

  $+ \underbrace{\underset{(cf.\ \S3.1)}{\overset{\text{[quantum linguistic Copenhagen interpretation]}}{\boxed{\text{Linguistic Copenhagen interpretation}}}}}_{\text{manual to use spells}}$

Measurement theory says :

- Describe every phenomenon based on Axioms 1 and 2 through the linguistic Copenhagen interpretation !

In this chapter, we introduce Axiom 1 for measurement. Axiom 2 for causality will be explained in Chapter 9.

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.* ref. [110]). In this section, we introduce the mathematical results concerning the Hilbert space without proofs. For the proofs, see, for example, ref. [106].

2.1.1 Hilbert space and operator algebra

Let $\mathbb{C}$ be the set of all complex numbers. Let H be a complex Hilbert space with an inner product $\langle \cdot, \cdot \rangle$, where the inner product $\langle \cdot, \cdot \rangle : H \times H \to \mathbb{C}$ satisfies that

(i) $\langle u, u \rangle \geq 0$ $(\forall u \in H)$, (ii) $\langle u, u \rangle = 0 \Leftrightarrow u = 0$,

(iii) $\langle u, \alpha_1 u_1 + \alpha_2 u_2 \rangle = \alpha_1 \langle u, u_1 \rangle + \alpha_2 \langle u, u_2 \rangle$ $(\forall u, u_1, u_2 \in H, \forall \alpha_1, \alpha_2 \in \mathbb{C})$,

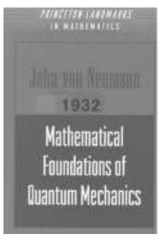
(iv) $\langle u_1, u_2 \rangle = \overline{\langle u_2, u_1 \rangle}$ (i.e.,, conjugate complex) $(\forall u_1, u_2 \in H)$

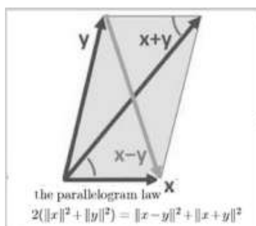
And, defining the norm $\|u\|$ (or, $\|u\|_H$) by $\|u\| = |\langle u, u \rangle|^{1/2}$, we get a Banach space $(H, \| \cdot \|)$. It is well known that the parallelogram law (i.e., $2(\|x\|^2 + \|y\|^2) = \|x - y\|^2 + \|x + y\|^2$) holds. Define $B(H)$ by

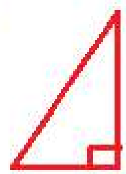
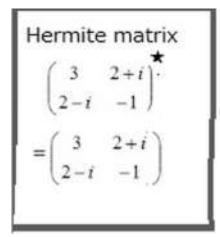
$$B(H) = \{T : H \to H \mid T \text{ is a continuous linear operator}\}. \tag{2.1}$$

$B(H)$ is regarded as a 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.

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

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

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

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 in the sense of the (operator) norm topology to T, that is, $n - \lim_{n \to \infty} T_n = T$, if

$$\lim_{n \to \infty} ||T_n - T||_{B(H)} = 0.$$

Also, 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].

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

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

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

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:

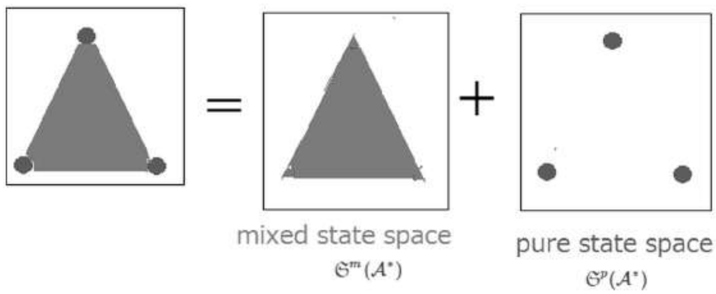
$$[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)].$$

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 $\mathfrak{S}^p(\mathcal{A}^*)$ are respectively defined by

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

(b) $\mathfrak{S}^p(\mathcal{A}^*) = \{\rho \in \mathfrak{S}^m(\mathcal{A}^*) \mid \rho$ 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 compact spaces (*cf.* ref.[114]).

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) $\overline{\mathfrak{S}}^m(\overline{\mathcal{A}}_*) = \{\rho \in \overline{\mathcal{A}}_* \mid \|\rho\|_{\overline{\mathcal{A}}_*} = 1, \rho \geq 0$ (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 diagram:

(C): General basic structure and State spaces

$$\underset{C^*\text{-pure state}}{\mathfrak{S}^p(\mathcal{A}^*)} \subset \underset{C^*\text{-mixed state}}{\mathfrak{S}^m(\mathcal{A}^*)} \subset \mathcal{A}^*$$

$$\uparrow \text{dual}$$

$$\boxed{\mathcal{A}} \xrightarrow[\text{subalgebra}\cdot\text{weak-closure}]{\subseteq} \boxed{\overline{\mathcal{A}}} \xrightarrow[\text{subalgebra}]{\subseteq} \boxed{B(H)} \qquad (2.6)$$

$$\downarrow \text{pre-dual}$$

$$\underset{W^*\text{-mixed state}}{\overline{\mathfrak{S}}^m(\overline{\mathcal{A}}_*)} \subset \overline{\mathcal{A}}_*$$

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

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

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

Let me show you a concluding classification in advance concerning quantum and classical state spaces as follows.

(A)

$$\boxed{\text{General basic structure}[\mathcal{A} \subseteq \overline{\mathcal{A}}]_{B(H)}}$$
$$\begin{matrix}\text{pure state space } \mathfrak{S}^p(\mathcal{A}^*) \\ C^*\text{-mixed state space } \mathfrak{S}^m(\mathcal{A}^*) \\ W^*\text{-mixed state } space\ \overline{\mathfrak{S}}^m(\overline{\mathcal{A}}_*)\end{matrix}$$

$$\Longrightarrow \begin{cases} \boxed{(\text{A}_1)\text{:Quantum basic structure}[\mathcal{C}(H) \subseteq B(H)]_{B(H)}} \\ \quad \text{pure state space } \mathfrak{S}^p(\mathcal{T}r(H)(\approx H)) \\ \quad C^*\text{-mixed state space } \mathfrak{S}^m(\mathcal{T}r(H))(=\mathcal{T}r_{+1}(H)) \\ \quad W^*\text{-mixed state space } \mathfrak{S}^m(\mathcal{T}r(H))(=\mathcal{T}r_{+1}(H)) \\ \\ \boxed{(\text{A}_2)\text{:Classical basic structure}[C_0(\Omega) \subseteq L^\infty(\Omega,\nu)]_{B(L^2(\Omega,\nu))}} \\ \quad \text{pure state space } \Omega \\ \quad C^*\text{-mixed state space } \mathcal{M}_{+1}(\Omega) \\ \quad W^*\text{-mixed state space } L^1_{+1}(\Omega,\nu) \end{cases}$$

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 systems, 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:

For further information see my homepage

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

$$
\begin{array}{ccccc}
\mathcal{T}r(H) & & & & \\
\uparrow \text{dual} & & & & \\
\boxed{\mathcal{C}(H)} & \xrightarrow[\text{subalgebra}\cdot\text{weak-closure}]{\subseteq} & \boxed{B(H)} & \xrightarrow[\text{subalgebra}]{\subseteq} & \boxed{B(H)} \\
 & & \downarrow \text{pre-dual} & & \\
 & & \mathcal{T}r(H) & &
\end{array}
\tag{2.8}
$$

Before we explain "compact operators class $\mathcal{C}(H)$" and "trace class $\mathcal{T}r(H)$", 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"$\big]$.

$$u = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}, \quad v = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

$$|u\rangle\langle v| = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \cdot [\overline{\beta}_1, \overline{\beta}_2] = \begin{bmatrix} \alpha_1\overline{\beta}_1 & \alpha_1\overline{\beta}_2 \\ \alpha_2\overline{\beta}_1 & \alpha_2\overline{\beta}_2 \end{bmatrix}$$

$$\langle v||u\rangle = \langle v, u\rangle = \bar{\beta}_1\alpha_1 + \bar{\beta}_1\alpha_1$$

[(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, "$(\mathrm{C}_1)\leftrightarrow$ (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 (D_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

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

For further information see my homepage

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

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

and thus,

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

However, it should be noted that the norms are different as mentioned in (C_3) and (D_3).

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

Consider the quantum basic structure:

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

and see the following diagram:

(E): Quantum basic structure and State space

$$\underset{C^*\text{-pure state}}{\mathfrak{S}^p(\mathfrak{Tr}(H))} \subset \underset{C^*\text{-mixed state}}{\mathfrak{S}^m(\mathfrak{Tr}(H))} \subset \mathfrak{Tr}(H)$$

$$\uparrow \text{dual}$$

$$\boxed{\mathfrak{C}(H)} \xrightarrow[\substack{\text{subalgebra}\\ \text{weak-closure}}]{\subseteq} \boxed{B(H)} \xrightarrow[\text{subalgebra}]{\subseteq} \boxed{B(H)}$$

$$\downarrow \text{pre-dual}$$

$$\underset{W^*\text{-mixed state}}{\overline{\mathfrak{S}}^m(\mathfrak{Tr}(H))} \subset \mathfrak{Tr}(H) \tag{2.17}$$

In what follows, we shall explain the above diagram.
Firstly, we note that

$$\mathfrak{C}(H)^* = \mathfrak{Tr}(H), \qquad \mathfrak{Tr}(H)^* = B(H) \tag{2.18}$$

and

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

Also, concerning the pure state space, we see:

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

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

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

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

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

$$T = \begin{bmatrix} -2i & -1-i & -3i \\ 2-i & -6+i & -3i \\ 5+2i & 2-i & 3 \end{bmatrix}$$

$$\begin{aligned}\mathrm{Tr}(T) &= \sum_{k=1}^{3} \langle e_k, T e_k \rangle \\ &= -2i + (-6+i) + 3 \\ &= -3 - i\end{aligned}$$

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

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

For further information see my homepage

and further, under the following notations:

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

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

we see

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

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

For further information see my homepage

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

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

In what follows, we shall explain this diagram.

2.3.1.1 Commutative C^*-algebra $C_0(\Omega)$

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

$$\begin{array}{l} \mathbb{R}(=\text{the real line}),\ \ \mathbb{R}^2(=\text{plane}),\ \ \mathbb{R}^n(=n\text{-dimensional Euclidean space}), \\ [a,b](=\text{interval}),\ \ \underset{(\text{with discrete metric } d_D)}{\text{finite set } \Omega(=\{\omega_1,...,\omega_n\})} \end{array}$$

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 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

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

For further information see my homepage

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. Furthermore, 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

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

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.

Riesz theorem (*cf.* ref. [114]) 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”.

2.3.1.2 Commutative W^*-algebra $L^\infty(\Omega,\nu)$

Furthermore, 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

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

Therefore, we see

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

Hence,

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

Then, we have the classical basic structure:

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

For further information see my homepage

Theorem 2.12. [Gelfand theorem (*cf.* ref. [106])] 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

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

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

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

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

$$
\begin{aligned}
&\mathfrak{S}^p(C_0(\Omega)^*) \\
=&\{\rho = \delta_{\omega_0} \in \mathfrak{S}^p(C_0(\Omega)^*) \ : \ \delta_{\omega_0} \text{ is the point measure at } \omega_0(\in \Omega), \omega_0 \in \Omega\}
\end{aligned}
$$

$$\equiv \mathcal{M}^p_{+1}(\Omega). \tag{2.41}$$

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,

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

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

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

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

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

and

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

In this sense, we have the following identification:

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

 For further information see my homepage

After all, we have the following identification:

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

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

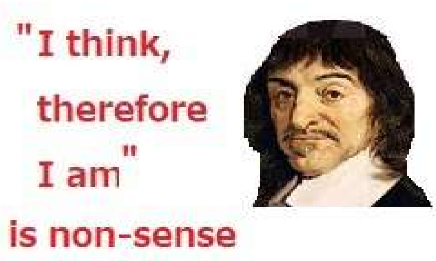
$$\|z\|_{C_0(\Omega)} = \max_{k=1,2,\ldots,n} |z_k| \qquad \forall z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{C}^n, \tag{2.45}$$

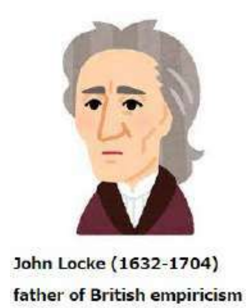
and the norm $\|\cdot\|_{\mathcal{M}(\Omega)}$ in the latter by

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

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

2.4.1 Mind-matter dualism (= mind-body dualism), Descartes, John Locke





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

(A): Axiom 1 (pure measurement)

(This can be read in §2.7)

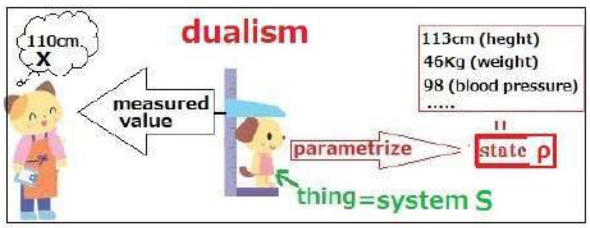
With any system S, a basic structure $[\mathcal{A} \subseteq \overline{\mathcal{A}}]_{B(H)}$ can be associated in which measurement theory of the system can be formulated. In $[\mathcal{A} \subseteq \overline{\mathcal{A}}]_{B(H)}$, consider a *$W^*$-measurement* $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}{=}(X,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

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

$$\rho(F(\Xi))(\equiv {}_{\mathcal{A}^*}(\rho, F(\Xi))_{\overline{\mathcal{A}}}) \tag{2.47}$$

(if $F(\Xi)$ is essentially continuous at ρ, or see (2.55) in Definition 2.18).




Max Born (1882–1970)

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 Enlightenment thinkers and known as the “Father of Classical Liberalism”. We think the following correspondence:

$$\begin{cases} \text{[state]} & \longleftrightarrow \text{[the primary quality]} \\ \text{[observable]} & \longleftrightarrow \text{[the secondary quality]} \end{cases} \tag{2.48}$$

And thus, we think

- “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 :Dualism and monism in world-views

dualism \ key-words	[A](= **mind**)	[B](Mediating of A and C) **(body)**	[C](= **matter**)
Plato (philosophical dualism)	actual world	Idea	Idea world
Aristotle (philosophical monism)	/	/	hyle [eidos]
Descartes (philosophical dualism)	mind	body	matter
Newton (scientific monism)	/	/	particle(point mass) [state] $\omega(\in \Omega)$
Locke (philosophical dualism)	brain	secondary quality	primary quality
quantum mechanics QL (scientific dualism)	observer [measured value] $[x(\in X)]$	measuring instrument [observable] $[\mathsf{O} = (X, \mathcal{F}, F)]$	particle (system) [state] $\rho(\in \mathfrak{S}^p(\mathcal{A}^*))$
classical QL (scientific dualism)	observer [measured value] $[x(\in X)]$	measuring instrument [observable] $[\mathsf{O} = (X, \mathcal{F}, F)]$	particle (system) [state] $\delta_\omega \approx \omega(\in \Omega)$
statistics† (incomplete dualism)	person to try [sample] $[x(\in X)]$	trial / /	population [parameter] $\omega(\in \Omega)$

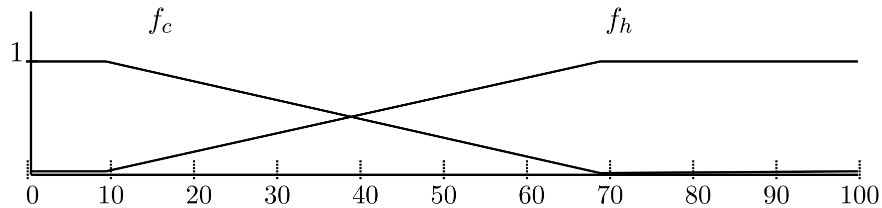
†:Statistics (supposedly dualistic) was formulated in the form of monism under the influence of Newtonian mechanics. In the end, statistics is neither monism nor dualism, but is regarded as a type of applied mathematics, which it is to this day (*cf.* Sec. 5.1).

Although I am not familiar with ”ontology”, I want to consider that ”keyword” exists in each world-view.

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

$$\text{“observable”} = \text{“the partition of word”} = \text{“the secondary quality”} \tag{2.49}$$

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



[Figure 1.2 in Chapter 1]: cold or hot?

Note that "measuring instrument" is the instrument that chooses a word among words. In this sense, we consider that "observable"= "measuring instrument". Note also that John Locke's words *"primary quality* (e.g., length, weight, etc.)*"* and *"secondary quality* (e.g., sweet, dark, cold, etc.)*"* come from dualism.

2.4.2 Essentially continuous

In §2.1.2, we introduced the following diagram:

(E):General basic structure and state space

$$
\begin{array}{ccccccc}
\underset{C^*-purestate}{\mathfrak{S}^p(\mathcal{A}^*)} & \subset & \underset{C^*\text{-mixed state}}{\mathfrak{S}^m(\mathcal{A}^*)} & \subset & \mathcal{A}^* & & & & \\
 & & & & \uparrow \text{dual} & & & & \\
 & & & & \boxed{\mathcal{A}} & \xrightarrow[\text{subalgebra weak-closure}]{\subseteq} & \boxed{\overline{\mathcal{A}}} & \xrightarrow[\text{subalgebra}]{\subseteq} & \boxed{B(H)} \\
 & & & & & & \downarrow \text{pre-dual} & & \\
 & & & & & & \underset{W^*\text{-mixed state}}{\overline{\mathfrak{S}}^m(\overline{\mathcal{A}}_*)} \subset \overline{\mathcal{A}}_* & &
\end{array}
\tag{2.50}
$$

In the above diagram, we introduce the following definition.

Definition 2.14. [Essentially continuous (*cf.* ref. [41])] 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 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 the case that $\rho_0(\in \mathfrak{S}^p(\mathcal{A}^*))$.

For further information see my homepage

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

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

Hence, we see that "essentially continuous" $\Leftrightarrow$ "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, $F(\omega_0) = \alpha$.

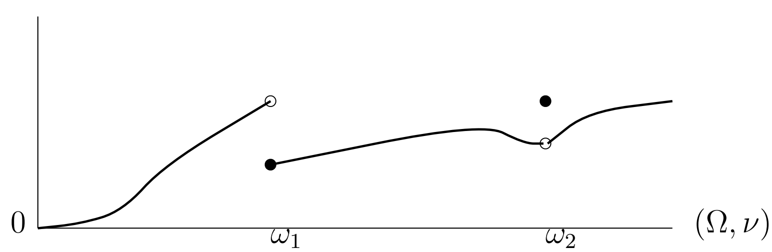


Figure 2.1: not essentially continuous at ω_1, essentially continuous at ω_2

2.4.3 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

(a) : $\emptyset$(=“empty set”) $\in \mathcal{F}$,

(b) : $\Xi_i \in \mathcal{F} \quad (i=1,2,\ldots) \Longrightarrow \bigcup_{i=1}^{n} \Xi_i \in \mathcal{F}, \quad \bigcap_{i=1}^{n} \Xi_i \in \mathcal{F}$

(c) : $\Xi_1, \Xi_2 \in \mathcal{F} \Longrightarrow \Xi_1 \setminus \Xi_2 \in \mathcal{F} \quad$ (where $\Xi_1 \setminus \Xi_2 = \{x \mid x \in \Xi_1, x \notin \Xi_2\}$)

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 (due to Davies, E.B. (*cf.* ref. [12])) 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 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) $0 \leqq F(\Xi) \quad (\forall \Xi \in \mathcal{F})$, $F(\emptyset) = 0$, $F(X) = I$

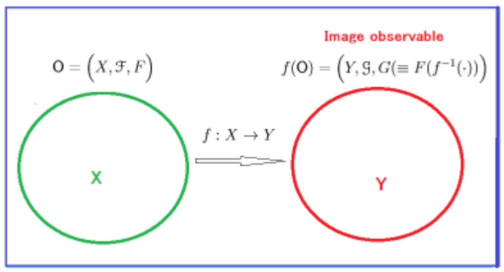
(b) for any $\overline{\rho}(\in \overline{\mathfrak{S}}^m(\overline{\mathcal{A}}_*))$, there exists a probability space $(X, \mathcal{F}, P_{\overline{\rho}})$ such that

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

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

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

Also, an *image observable* of O is defined by



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. 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 11.1).

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 $k = 1, 2, ...$

For further information see my homepage

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

" the resolution of the identity I"

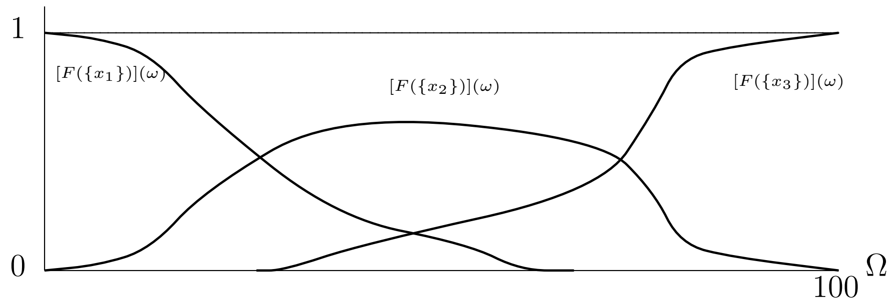


Figure 2.2: $\mathsf{O} \equiv (\{x_1, x_2, x_3\},\ 2^{\{x_1,x_2,x_3\}},\ F)$

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

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

 For further information see my homepage

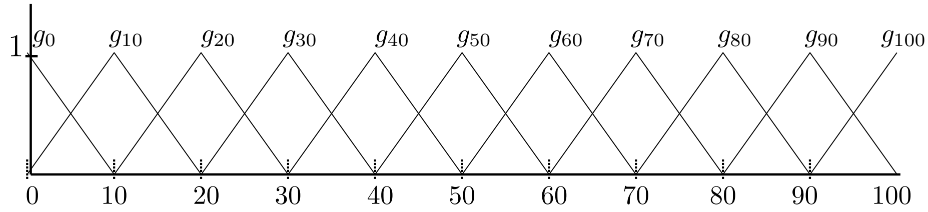


Figure 2.3: Triangle observable

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] Consider a classical basic structure $[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))]$. Here, $\Omega = \mathbb{R} \times \mathbb{R}^+$, where $\mathbb{R} = \{\mu : \mu \in \mathbb{R}\}$, $\mathbb{R}^+ = \{\sigma \in \mathbb{R} : \sigma > 0\}$. $\Omega = \mathbb{R} \times \mathbb{R}^+$ is assumed to have Lebesgue measure $\nu(d\omega)(= d\mu \times d\sigma)$. The *normal observable* O=$(\mathbb{R}, \mathcal{B}_{\mathbb{R}}, G)$ in $L^\infty(\Omega, \nu)$ is defined by

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

This is the most fundamental observable in statistics.

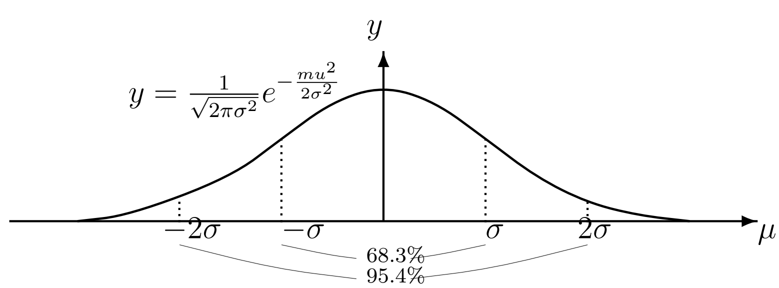


Figure 2.4: Error function

For further information see my homepage

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 narrower 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 $[F^{(\mathrm{exa})}(\Xi)](\omega) = \chi_\Xi(\omega)$ $(\Xi \in \mathcal{B}_\Omega, \omega \in \Omega)$. The triplet $\mathsf{O}^{(\mathrm{exa})} = (\Omega, \mathcal{B}_\Omega, F^{(\mathrm{exa})})$ is called the *exact observable* in $L^\infty(\Omega,\nu)$. This is the W^*-observable and not C^*-observable, since $[F^{(\mathrm{exa})}(\Xi)](\omega)$ is not always continuous. For the argument about the sample probability space (*cf.* Hypothesis 2.19), 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

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

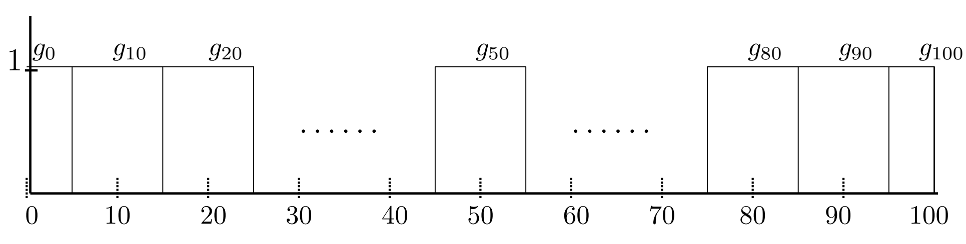


Figure 2.4: Round observable

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

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

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 a 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

(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:

$$\begin{aligned} &(\sharp_1) : \widetilde{q} : \Omega \to \mathbb{R}, && \widetilde{q}(q,p) = q \quad (\forall (q,p) \in \Omega) \\ &(\sharp_2) : \widetilde{p} : \Omega \to \mathbb{R}, && \widetilde{p}(q,p) = p \quad (\forall (q,p) \in \Omega) \\ &(\sharp_3) : \widetilde{e} : \Omega \to \mathbb{R}, && \widetilde{e}(q,p) = [\text{potential energy }] + [\text{kinetic energy }] \\ & && \qquad\quad = \underset{\text{(Hamiltonian)}}{U(q) + \frac{p^2}{2m}} \qquad (\forall (q,p) \in \Omega) \end{aligned}$$

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

For further information see my homepage

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]:

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

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 :

($\sharp$) 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 $\mathrm{Tr}(\rho \cdot F_A(\{x\}))$ $(= \langle \omega, F_A(\{x\})\omega \rangle)$,

[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 $\mathsf{O}_A = (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, F_A)$.

(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 $(\delta^\omega_A)^2$ is given by

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

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.* ref. [114]) 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

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

For further information see my homepage

2.7 Axiom 1 – No science without measurement

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

- measurement theory (=quantum language) := Measurement [Axiom 1] (*cf.* §2.7) + (deterministic) Causality [Axiom 2] (*cf.* §9.3)
  (Measurement + Causality: a kind of spells (a priori judgment))
  + Linguistic Copenhagen interpretation [quantum linguistic Copenhagen interpretation] (*cf.* §3.1)
  (Linguistic Copenhagen interpretation: 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

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

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

Here we introduce the following axiom.

(B): Axiom 1 (measurement) pure type

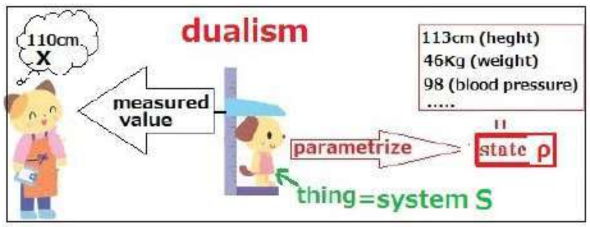
With any system S, a basic structure $[\mathcal{A} \subseteq \overline{\mathcal{A}}]_{B(H)}$ can be associated in which measurement theory of the system can be formulated. In $[\mathcal{A} \subseteq \overline{\mathcal{A}}]_{B(H)}$, consider a *$W^*$-measurement* $\mathsf{M}_{\overline{\mathcal{A}}}\big(\mathsf{O}{=}(X,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

- a W^*-measurement $\mathsf{M}_{\overline{\mathcal{A}}}\big(\mathsf{O}, S_{[\rho]}\big)$ (or, C^*-measurement $\mathsf{M}_{\mathcal{A}}\big(\mathsf{O}{=}(X, \mathcal{F}, F), S_{[\rho]}\big)$) 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 $\mathsf{M}_{\overline{\mathcal{A}}}\big(\mathsf{O}, S_{[\rho]}\big)$ (or C^*-measurement $\mathsf{M}_{\mathcal{A}}\big(\mathsf{O}{=}(X, \mathcal{F}, F), S_{[\rho]}\big)$) belongs to Ξ $(\in \mathcal{F})$ is given by

$$\rho(F(\Xi))(\equiv {}_{\mathcal{A}^*}(\rho, F(\Xi))_{\overline{\mathcal{A}}}) \tag{2.68}$$

(if $F(\Xi)$ is essentially continuous at ρ, or see (2.55) in Definition 2.18).



Max Born (1882–1970)

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)
quantum mechanics (Born's quantum measurement) —linguistic turn→
(physics)

(a kind of spell)
measurement theory(Axiom 1) (2.69)
(metaphysics, language)

♠**Note 2.4.** Recall a part of Table 2.1 as follows.

a part of Table 2.1

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

dualism \ key-words	[A](= **mind**)	[B](Mediating of A and C) **(body)**	[C](= **matter**)
quantum mechanics QL (scientific dualism)	observer [measured value] $[x(\in X)]$	measuring instrument [observable] $[\mathsf{O} = (X, \mathcal{F}, F)]$	particle (system) [state] $\rho(\in \mathfrak{S}^p(\mathcal{A}^*))$
classical QL (scientific dualism)	observer [measured value] $[x(\in X)]$	measuring instrument [observable] $[\mathsf{O} = (X, \mathcal{F}, F)]$	particle (system) [state] $\delta_\omega \approx \omega(\in \Omega)$
statistics (incomplete dualism)	person to try [sample] $[x(\in X)]$	trial / /	population [parameter] $\omega(\in \Omega)$

In the above, let's compare classical QL and statistics as follows. The classical QL has a measurement $\mathsf{M}(\mathsf{O}{=}(X, \mathcal{F}, F), S_{[\delta_\omega]})$, on the other hand, statistics has no measurement, but it is usually assumed that statistics starts from $(X, \mathcal{F}, P_\omega)$, i.e., the sample probability space with a parameter $\delta_\omega \approx \omega(\in \Omega)$. That is,

$$\underset{\text{scientific dualism}}{\overset{\text{quantum language}}{\boxed{\mathsf{M}(\mathsf{O}{=}(X, \mathcal{F}, F), S_{[\delta_\omega]})\ )}}} \xrightarrow[\text{Elimination of observable } \mathsf{O}]{} \underset{\text{applied math}}{\overset{\text{statistics}}{\boxed{(X, \mathcal{F}, P_\omega(\cdot))}}}$$

where $(X, \mathcal{F}, P_\omega(\cdot))$ is a probability space with a parameter $\omega(\in \Omega)$
The elimination of an observable O implies the elimination of dualism.

♠**Note 2.5.** The above axiom is due to Max Born (1926) (*cf.* ref.[6]). 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.

♠**Note 2.6.** In Chaps. 11 and 12 of ref. [76], I discussed the following.

The question ’What is a proposition?’ is always the most important question for languages. In mathematics, the followings are equivalent

($\sharp_1$) What is a proposition in mathematics?

($\sharp_2$) What is a set?

This is answered in axiomatic set theory (e.g., Zermelo-Fraenkel set theory. Thus, a proposition in mathematics is always an analytic proposition.

Our interest is the problem: "what is a proposition in QL?" In ref. [76], I answered as follows

($\flat_1$) a proposition in QL which is defined by a sentence like Axiom 1 is a QL proposition (i.e., a proposition in QL).

That is because a sentence like Axiom 1 can be judged true or false by experimentation (*cf.* Sec. 4.2: the law of large numbers). Thus, I can say that

($\flat_2$) In QL, Popper's falsifiability is not needed. That is, Popper's falsifiability is automatically included in QL propositions

2.7.2 A simplest example

Example 2.31. [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

$$f_c(\omega) = \frac{g_c(\omega)}{\text{the numbers of testees}}$$
$$f_h(\omega) = \frac{g_h(\omega)}{\text{the numbers of testees}} \tag{2.70}$$

And

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

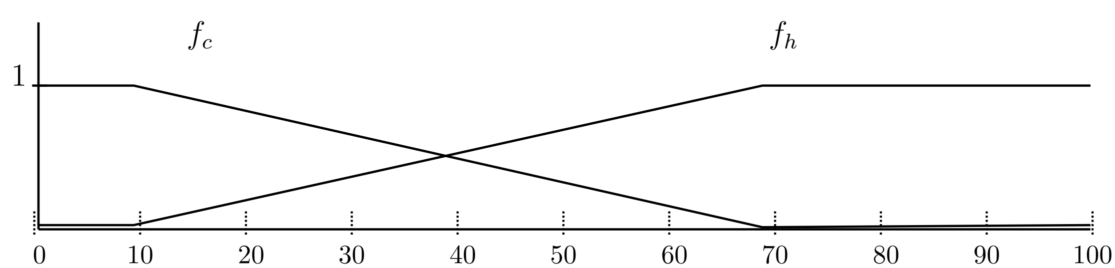


Figure 1.2: Cold or hot?

Therefore, for example,

(C_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 (C_1) in terms of quantum language (i.e., Axiom 1).

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

(C_2) for water with ω °C, [C-H]-thermometer presents $\begin{bmatrix} \text{c} \\ \text{h} \end{bmatrix}$ 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:

(C_3) $\qquad (C_1) \Longleftrightarrow (C_2)$

Therefore, the statement (C_1) in ordinary language can be represented in terms of measurement theory as follows.

(C_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 $\underset{(\textbf{state}(= \omega \in \Omega))}{[55\ °\text{C}]}$, the probability that **measured value** $\begin{bmatrix} \text{c} \\ \text{h} \end{bmatrix}$ 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 (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

$$\text{(a)} \begin{cases} (\mathrm{a}_1)\text{: Use the language defined by Axiom 1 (§2.7)} \\ \\ (\mathrm{a}_2)\text{: Trust in man's linguistic competence} \end{cases}$$

Thus, some readers may have a question:

(b) Is it science ?

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

♠**Note 2.7.** [**Realistic worldview vs. Idealistic (=linguistic) worldview**] I am not a philosopher, thus, my use of the terms 'realistic worldview' and 'idealistic worldview' may differ from their use in philosophy. Generally, it is said:

$$\begin{cases} \text{Realistic worldview} & \cdots \text{Object first, theory second.} \\ \text{Idealistic worldview} & \cdots \text{Theory first, object second.} \end{cases}$$

In this book, I think as follows.

($\sharp_1$) Realistic worldview:
There is a one-to-one correspondence between theory and object.

$$\boxed{\text{realistic theory}} \longleftrightarrow \boxed{\text{object}}$$

e.g., Newtonian mechanics, theory of relativity, quantum mechanics,...

($\sharp_2$) Idealistic (=linguistic) worldview
idealistic theory is applicable to many objects

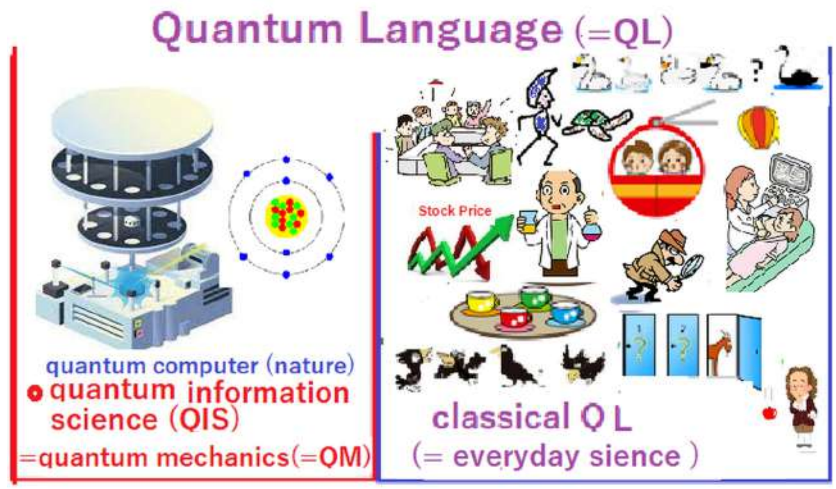
$$\boxed{\text{idealistic theory}} \longrightarrow \begin{cases} \longrightarrow \boxed{\text{object}_1} \\ \longrightarrow \boxed{\text{object}_2} \\ \longrightarrow \quad \vdots \end{cases}$$

e.g., statistics (= dynamical system theory), which is applicable to economics, medicine, etc.

Consequently, philosophical arguments can be avoided if we understand 'realistic≈precise' and 'idealistic≈rough' though this may be misleading. If so, readers may find it fruitless to expect much from scientific idealism. However, the theme of this book is 'classical QL' in 'scientific idealism'.
Recall the following Figure 0.3 in Preface:

For further information see my homepage

Figure 0.3 in Preface: [QL] = [QIS] $\bigcup$ [classical QL]



2.8.2 Elementary examples – urn problem, etc.

Since measurement theory (= QL) 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 Example 2.22 [triangle observable])] Consider the classical basic structure:

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

where Ω = "the closed interval $[0, 100]$" with the Lebesgue measure ν.
Let testees check water at various temperature ω °C ($0 \leqq \omega \leqq 100$). And you ask them "What is the approximate temperature (°C) of this water ?" Gather the data, (for example, $h_n(\omega)$ persons say n °C ($n = 0, 10, 20, \ldots, 90, 100$), and normalize them to get 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

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

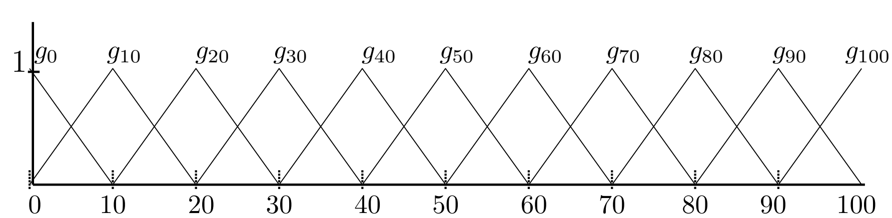


Figure 2.5: Triangle observable

(a) You choose one person from the testees, and you ask him/her "What is the approximate temperature (°C) of 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

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

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 have the following translation:

$$
\underset{\text{(ordinary language)}}{\boxed{\text{statement (a)}}} \xrightarrow[\text{translation}]{} \underset{\text{(quantum language)}}{\boxed{\text{statement (b)}}} \tag{2.71}
$$

///

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 $\mathsf{O}^{(\text{exa})} = (X(= \Omega), \mathcal{F}(= \mathcal{B}_\Omega), F^{(\text{exa})})$ in $L^\infty(\Omega, \nu)$ such that

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

Let $\delta_{\omega_0} \approx \omega_0 (\in \Omega)$. Consider the exact measurement $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O}^{(\text{exa})}, S_{[\delta_{\omega_0}]})$. 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 $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O}^{(\text{exa})}, S_{[\delta_{\omega_0}]})$ 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 $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O}^{(\mathrm{exa})}, S_{[\delta_{\omega_0}]})$.

Furthermore, 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]

Table 2.2: urn problem

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

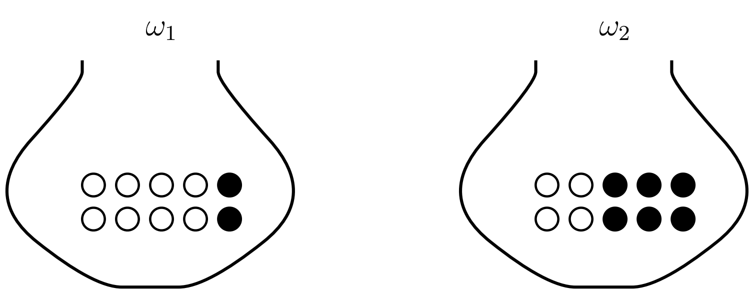


Figure 2.6: Urn problem

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.

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

$$\begin{aligned} U_1 &\quad \cdots \quad \text{"the urn with the state } \omega_1\text{"} \\ U_2 &\quad \cdots \quad \text{"the urn with the state } \omega_2\text{"} \end{aligned}$$

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 $F(\{w\})(\omega_2) = 0.4$.

Therefore, we see:

$$\underset{\text{(ordinary language)}}{\boxed{\text{statement (a)}}} \xrightarrow[\text{translation}]{} \underset{\text{(quantum language)}}{\boxed{\text{statement (b)}}} \tag{2.72}$$

♠**Note 2.8.** [$L^\infty(\Omega, \nu)$, 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 necessarily indispensable. For example, even if we assume that $\nu(\{\omega_1\}) = 2$ and $\nu(\{\omega_2\}) = 1/3$, we can obtain the same conclusion. Thus, in this book, $L^\infty(\Omega, \nu)$. is often abbreviated to $L^\infty(\Omega)$

♠**Note 2.9.** 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 $|b - a|$.

For further information see my homepage

The common sense – "probability" can not be derived from mathematics itself – is well known as Bertrand's paradox (*cf.* §9.12). 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 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

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

That is, we consider the following identification: (Therefore, image Figure 2.7):

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

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

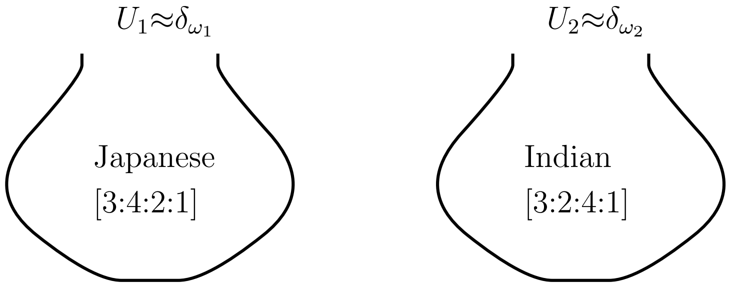


Figure 2.7: Population(=system)≈urn

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

$$
\begin{aligned}
&[F_{\mathrm{BT}}(\{O\})](\omega_1) = 0.3, \qquad && [F_{\mathrm{BT}}(\{A\})](\omega_1) = 0.4, \\
&[F_{\mathrm{BT}}(\{B\})](\omega_1) = 0.2, \qquad && [F_{\mathrm{BT}}(\{AB\})](\omega_1) = 0.1,
\end{aligned} \tag{2.73}
$$

and

$$
\begin{aligned}
&[F_{\mathrm{BT}}(\{O\})](\omega_2) = 0.3, \qquad && [F_{\mathrm{BT}}(\{A\})](\omega_2) = 0.2, \\
&[F_{\mathrm{BT}}(\{B\})](\omega_2) = 0.4, \qquad && [F_{\mathrm{BT}}(\{AB\})](\omega_2) = 0.1.
\end{aligned} \tag{2.74}
$$

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 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 2.10.** Readers may feel that Example 2.32–Example 2.35 are too easy. However, as mentioned in (a) of Sec. 2.8.1, what we can do is

$$
\bullet \begin{cases} \text{to be faithful to the Axioms} \\ \text{to trust in man's linguistic competence} \end{cases}
$$

If some find another language that is more powerful than quantum language, it will be praised as the greatest discovery in the history of science. That is because the discovery allows us to go beyond quantum mechanics.

2.9 Simple quantum examples (Stern=Gerlach experiment)

2.9.1 Stern=Gerlach experiment

Example 2.36. [Quantum measurement (Stern–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 upward or downward at the ratio of 50:50. See Figure 2.10.

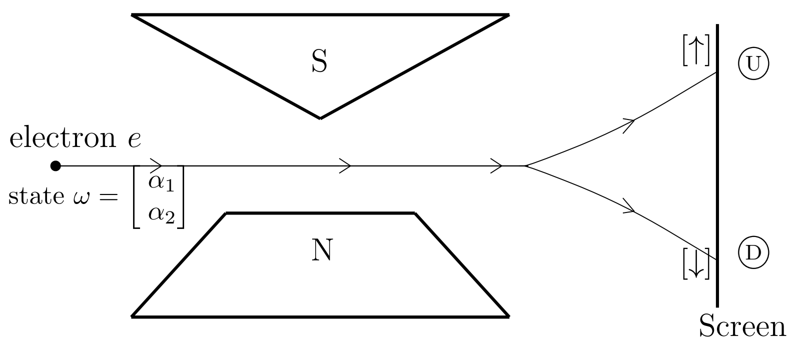


Figure 2.8: Stern–Gerlach experiment (1922)

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 $\|\omega\|_{\mathbb{C}^2} = 1$. Put $\omega = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$ (where $||\omega||^2_{\mathbb{C}^2} = |\alpha_1|^2 + |\alpha_2|^2 = 1$). Define $\mathsf{O}_z \equiv (Z, 2^Z, F_z)$, the spin observable concerning the z-axis, such that, $Z = \{\uparrow, \downarrow\}$ and

$$F_z(\{\uparrow\}) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad F_z(\{\downarrow\}) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad F_z(\emptyset) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad F_z(\{\uparrow, \downarrow\}) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{2.75}$$

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

For further information see my homepage

(♯) 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 $\omega = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$, we say :

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

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

Remark 2.37. We can define $\mathsf{O}^x \equiv (X, 2^X, F^x)$, the spin observable concerning the x-axis, such that, $X = \{\uparrow_x, \downarrow_x\}$ and

$$F^x(\{\uparrow_x\}) = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}, \quad F^x(\{\downarrow_x\}) = \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix}. \tag{2.76}$$

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

$$F^y(\{\uparrow_y\}) = \begin{bmatrix} 1/2 & i/2 \\ -i/2 & 1/2 \end{bmatrix}, \quad F^y(\{\downarrow_y\}) = \begin{bmatrix} 1/2 & -i/2 \\ i/2 & 1/2 \end{bmatrix}, \tag{2.77}$$

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:

$$\hat{S}_y\hat{S}_z - \hat{S}_z\hat{S}_y = 2i\hat{S}_x, \quad \hat{S}_z\hat{S}_x - \hat{S}_x\hat{S}_z = 2i\hat{S}_y, \quad \hat{S}_x\hat{S}_y - \hat{S}_y\hat{S}_x = 2i\hat{S}_z. \tag{2.78}$$

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

For further information see my homepage

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

Axiom 1 (measurement) includes the 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.* §10.3, and refs.[13, 107]). Also, it should be noted that the argument below is essentially same as the one for the Stern=Gerlach experiment.

Example 2.38. [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.

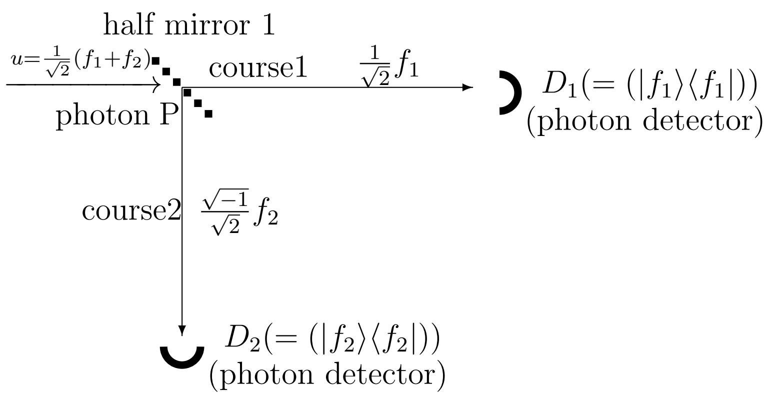


Figure 2.9: $[D_2 + D_1] =$ observable O

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

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

For further information see my homepage

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 2.9. 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 $\sqrt{-1}f_2$), and goes along the course 2 to the photon detector D_2.

Thus, we have the measurement:

$$\mathsf{M}_{B(\mathbb{C}^2)}(\Phi\mathsf{O}_f, S_{[\rho]}). \tag{2.79}$$

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 is significantly separated from detector D_2.

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.

For further information see my homepage

This is the de Broglie paradox (*cf.* refs. [13, 107]). From a viewpoint of quantum language, we give up to solve the paradox, that is, we declare

Stop to be bothered !

♠**Note 2.11.** 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 (Stern-Gerlach experiment). I think that

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

The readers will find that the other paradoxes (see "paradox" in the index of this lecture note) in quantum mechanics are solved in this note.

For further information see my homepage

Chapter 3

Linguistic Copenhagen interpretation (dualism and idealism)

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

$$
\bullet \quad \underset{\text{(=quantum language)}}{\boxed{\text{measurement theory}}} := \underbrace{\overset{\text{[Axiom 1]}}{\underset{(cf.\ \S 2.7)}{\boxed{\text{Measurement}}}} + \overset{\text{[Axiom 2]}}{\underset{(cf.\ \S 9.3)}{\boxed{\text{Causality}}}}}_{\text{a kind of spell(a priori judgment)}}
$$

$$
+ \underbrace{\overset{\text{[linguistic Copenhagen interpretation]}}{\underset{(cf.\ \S 3.1)}{\boxed{\text{Linguistic Copenhagen interpretation}}}}}_{\text{manual to use spells}}
$$

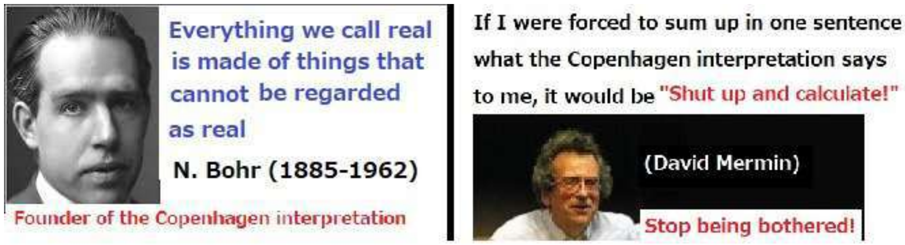
QL says that

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

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

3.1 Linguistic Copenhagen interpretation

This section was written with reference to ref. [76].



3.1.1 What is the linguistic Copenhagen interpretation?

In the previous section, an overview of quantum language [Axiom 1 (measurement) and Axiom 2 (causality)] was outlined.
(A)

$$\underset{\text{(=language of science)}}{\overset{\text{(=measurement theory(=MT))}}{\boxed{\text{quantum language(=QL)}}}} = \overset{\text{[Axiom 1]}}{\boxed{\text{measurement}}} + \overset{\text{[Axiom 2]}}{\boxed{\text{causal relation}}}$$

$$+ \underset{\text{[the manual to use Axioms 1 and 2]}}{\boxed{\text{(linguistic) Copenhagen interpretation}}} \tag{3.0}$$

In this section, the "Copenhagen interpretation ((linguistic) Copenhagen interpretation)" will be explained. Of course, as stated in Sec. 1.1.1, I believe that the linguistic Copenhagen interpretation is the true Copenhagen interpretation.

Before doing so, let us reiterate the following.

(B_1) Axioms are a kind of incantation (spell, magic word, metaphysical statement) and cannot be experimentally verified

Further,

(B_2) Quantum language is a language, and you may not be able to use it well at first. You can only acquire the ability to use it through practice and trial and error.

♠**Note 3.1.** (i): In Mermin's book [94], he said

- If I were forced to sum up in one sentence what the Copenhagen interpretation says to me, it would be **"Shut up and calculate"**
- Stop being bothered!

Also, D. Howard said in ref. [24]:

- Even within the Copenhagen School, there was a wide range of opinion on the Copenhagen interpretation. For example, there was disagreement about “wave function collapse” which is supposed to be the central theme of the Copenhagen interpretation. (See ref. [59] (or, Sec. 10.2 in this book) for my opinion on “wave function collapse.”)

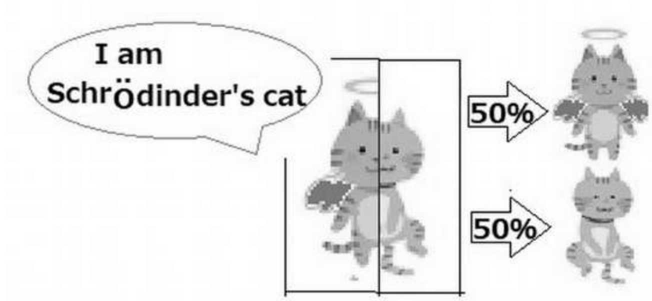


Figure 3.1: Schrödinger’s cat

This means that

“What is the Copenhagen interpretation?” has not yet been resolved

We believe that this is one of the most important unsolved problems in science. Thus, I can say that one of the purposes of this book is to answer the unsolved problem: “What is the Copenhagen interpretation?”.

(ii): Among the different schools of thought on the ’Copenhagen Interpretation’, the following is interesting:

- ‘Copenhagen interpretation’ is a manual on how to use quantum mechanics formulated in the von Neumann style (i.e., the Hilbert space formulation of quantum mechanics (*cf.* ref. [110]).

Our ‘linguistic Copenhagen interpretation’ is a mathematical generalization of this (*cf.* ref. [35, 41]). I assert that the linguistic Copenhagen interpretation is the true Copenhagen interpretation. That is, we assert that the Copenhagen interpretation is justified in philosophy (i.e., language) and not in physics.

Thus. in this book, “Copenhagen interpretation” is identified with “linguistic Copenhagen interpretation”.

(iii): Saying the same thing over and over again, my opinion is as follows.

- as mentioned in Note 0.1 (in Preface), I want to consider that QM in ⓐ and QIS in ⓒ (in Figure 0.1 in Preface) are essentially the same.

Thus, I think that the Copenhagen interpretation in QM of ⓐ is not necessary. The linguistic Copenhagen interpretation is all that is needed.

////

(C) It is essential to acquire a habitual thinking to master the axioms (Axioms 1 and 2). For this, as Mermin says, it may be sufficient to just 'Shut up and calculate'. But in order to master the quantum language as quickly as possible, you will need a good manual for mastering the axioms (Axioms 1 and 2).

Thus, we get the following definition,

Definition 3.1. [Linguistic Copenhagen Interpretation (=Copenhagen Interpretation)] We have two definitions as follows:

(C_1) **Linguistic Copenhagen Interpretation** $\underset{\text{def.}}{:=}$ **Manual for using spells (= Axioms 1 and 2)**

However, there is another way of thinking about it. In the case that we do not know Axioms 1 and 2, the Copenhagen Interpretation may have to be considered. Thus we have another definition as follows.

(C_2) **Linguistic Copenhagen Interpretation** $\underset{\text{def.}}{:=}$ **common knowledge in the world of dualistic idealism**

(To be more specific, it is a memo that records things that are obvious in the world of dualistic idealism, but not obvious to our normal senses.)

Although (C_1) is easy to understand, I rather prefer (C_2); therefore, in this paper, I would like to consider (C_2) as the main one. If (C_2) cannot be used alone, then (C_1) is used as an auxiliary. ////

♠**Note 3.2.** (i) I believe that our Copenhagen interpretation is more closely related to dualistic idealism (=mind-body dualism) than to quantum mechanics. And I am convinced that this Copenhagen interpretation is the true Copenhagen interpretation. In the above, note that we have two definitions of the Copenhagen interpretations in QL such that

$$\boxed{\text{QL}} = \boxed{\text{two Axioms}} + \boxed{\text{Copenhagen interpretation}}$$

That is, "Which comes first, $\boxed{\text{two Axioms}}$ or $\boxed{\text{Copenhagen interpretation}}$?". It is clear that ($C_1$) is due to the assumption that $\boxed{\text{two Axioms}}$ comes first. On the other hand, (C_2) is due to the assumption that $\boxed{\text{Copenhagen interpretation}}$ comes first. Surprisingly, as seen in the following section (e.g., Parmenides, Descartes, etc.), most of the rules in the Copenhagen Interpretation were discovered before the discovery of quantum mechanics. I therefore prefer the latter definition, but it may not be a matter determined by my preferences.

(ii) I believe that

For further information see my homepage

($\sharp_1$) main objective of the philosophy of science = to create a language of science (i.e., quantum language).

If so, then the closest (non-physical and idealistic) theory that has so far come to the aims of the philosophy of science is statistics. However, statistics is not regarded as a major area of philosophy of science. The reason for this are described in Note 1.1. That is,

($\sharp_2$) in statistics, the concept of 'Idea' (= observable) has been erased (*cf.* Note 2.4).

3.1.2 Descartes figure

Now, let's go on to explain the (linguistic) Copenhagen Interpretation.

Since Axiom 1 includes the term "measurement", the concept of "measurement" should be, at first, understood in dualism (i.e., "observer" and "measuring object") as illustrated in Figure 3.2.

Figure 3.2. [Descartes Figure] Image of "measurement(=ⓐ+ⓑ)" in dualism

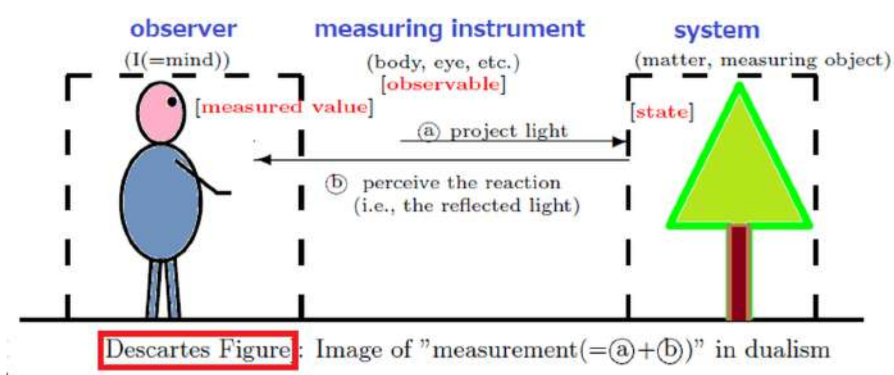


Figure 3.2: Descartes Figure

In the figure, "measurement" is characterized as interaction between "observer" and "system"(matter or object to be measured, measuring object), composed of

(D_1) ⓐ projection of light onto the object (i.e., someone, not necessarily an observer, shines the light.)
ⓑ perception of the reaction of the object (i.e., the observer receives the reaction.)

However, I want to emphasize that the interaction cannot be represented by kinetic equations. Therefore,

(D_2) in measurement theory (= quantum language), we use the term "measurement" instead of "interaction". Therefore, we won't say the above (D_1) outright.

After all, we think that

(D_3) **there is no measured value without observer.**

Thus, measurement theory is composed of three keywords:

$$\underset{\text{(observer, brain, mind)}}{\boxed{\text{measured value}}}, \quad \underset{\text{(telescope, thermometer, eye, ear, body, polar star)}}{\boxed{\begin{array}{c}\text{observable}\\ (\approx \text{measuring instrument})\end{array}}}, \quad \underset{\text{(matter, measuring object)}}{\boxed{\text{state}}}.$$

In view of the above figure, it might be called "ternary relation (or, trialism)" instead of "dualism". But, following the convention, we use *"dualism"* throughout this book.

♠**Note 3.3.**

(i) Descartes' dualistic idealism has the following form:

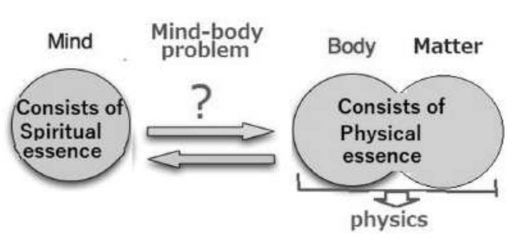
$$[\text{A}](\textbf{mind}) \quad \longleftarrow \underset{\textbf{(medium)}}{[\text{B}(\textbf{body, sensory organ})]} \longrightarrow \quad [\text{C}](\textbf{matter})$$

The following is a part of Table 2.1

dualism \ key-words	[A](= **mind**)	[B](Mediating of A and C) **(body)**	[C](= **matter**)
quantum mechanics QL (scientific dualism)	observer [measured value] $[x(\in X)]$	measuring instrument [observable] $[\mathsf{O} = (X, \mathcal{F}, F)]$	particle (system) [state] $\rho(\in \mathfrak{S}^p(\mathcal{A}^*))$
classical QL (scientific dualism)	observer [measured value] $[x(\in X)]$	measuring instrument [observable] $[\mathsf{O} = (X, \mathcal{F}, F)]$	particle (system) [state] $\delta_\omega \approx \omega(\in \Omega)$

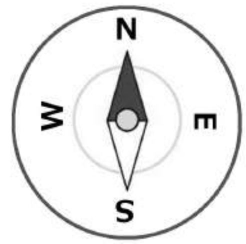
(ii)The most important issue in philosophy is said to be the mind-body problem. That is,

(♯) Clarify the relationship between 'mind', 'body' and 'matter'?



I assert that this problem can be completely solved by Axiom 1. That is because Axiom 1 says the relationship between 'mind (≈measured value)', 'body (≈observable)' and 'matter(≈system)'. Or see Sec. 12.8 in ref. [76].

(iii) The concept of "observable" (which can be identified with "measuring instrument") is not easy. For example, telescopes, glasses and eyes are a type of measuring instrument. A directional magnet is, of course, a measuring instrument. If so, then the polar star is also a type of measuring instrument.



////

3.1.3 The linguistic Copenhagen interpretation [(E_0)-(E_7)]

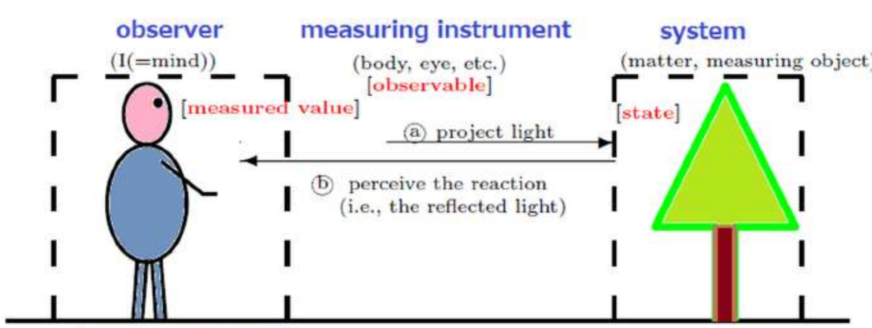
The (linguistic) Copenhagen interpretation is "a manual for using Axiom 1 (measurement) and Axiom 2 (causality). If that were the case (if it were a manual), wouldn't we have to list all sorts of miscellaneous things and "there would be no end to the explanations"? Even car driving manuals are endless in detail. There is no such thing as a complete rulebook for baseball or soccer, either. The author believes that the (linguistic) Copenhagen interpretation may have such a fear (*cf.* Wittgenstein's paradox in ref. [76]). However, I think that a Copenhagen Interpretation that covers the problems we are likely to encounter in practice is possible.

Now, below [(E_0)–(E_7)], I will briefly explain the Copenhagen interpretation. The most important of these, and especially important, is,

(E_4) Only one measurement is permitted.

(E) The linguistic Copenhagen interpretation

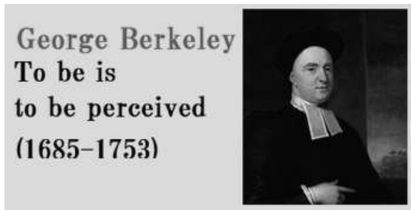
With Descartes figure below and the following (E_0)-(E_7) in mind, describe every phenomenon in terms of Axioms 1 and 2!

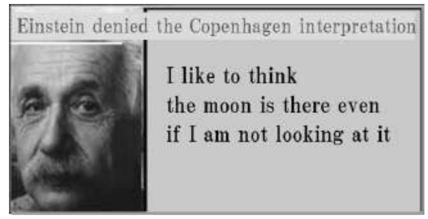


Descartes Figure: Image of "measurement(=ⓐ+ⓑ)" in dualism

Descartes figure 3.2

(E$_0$) (i) If you don't measure it, you don't know anything. Don't talk about things that cannot be measured. This seems to be Berkeley's saying "To be is to be perceived." On the other hand, Einstein, a monistic realist of the anti-Copenhagen interpretation, said "The moon is there even when we are not looking".

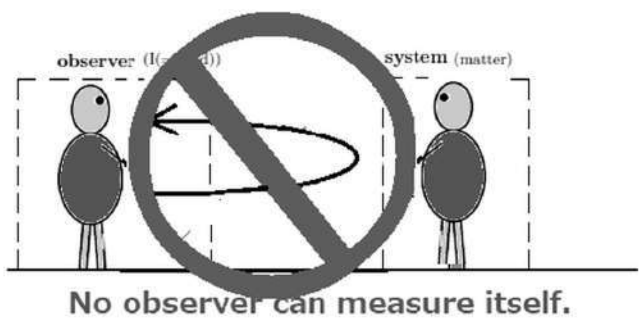




(ii) [Popper's Falsifiability in the linguistic interpretation; (*cf.* Sec 12.4 in ref. [76]]

Popper's Falsifiability is usually explained as follows. In order to guarantee the objectivity of a scientific theory, there must be a possibility that the hypothesis will be disproved by experiment or observation. That is, truth must always be subjected to experiments that deny its truth. And if the denying experiment is confirmed, then the truth must be rejected. As mentioned in Note , recall that "QL proposition" = "measurement". Therefore, the importance of Popper's Falsifiability cannot be over-emphasised in QL.

(E$_1$) Consider the **dualism** composed of "observer" and "matter (= object to be measured)", where "observer" and "matter (= measuring object)" must be absolutely separated. Figuratively speaking, "Audience should not go on stage", or



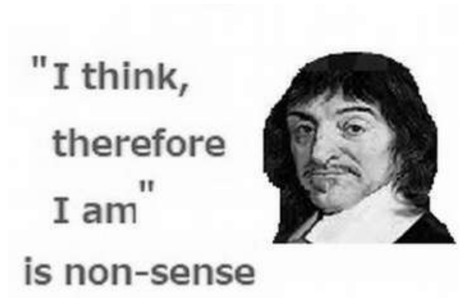
"the observer cannot measure the observer himself"

or

"The measurement is not dependent on the observer"

That is, the following qualia problem is non-sense.

For further information see my homepage



To be more specific, the words "I", "Here", "Now" are forbidden . Hence, "I think, therefore I am" is non-sense.

♠**Note 3.4.** Consider the followings:

($\sharp_1$) I measure my body temperature with a thermometer.

($\sharp_2$) I feel my body feverish.

and

($\flat_1$) The doctor measures my body temperature with a thermometer.

($\flat_2$) The doctor feels my body feverish.

In terms of measurement, ($\sharp_1$) and ($\flat_1$) are the same. On the other hand, ($\sharp_2$)and ($\flat_2$) are different. Thus, in the strictest sense, we consider that ($\sharp_2$) cannot be regarded as a measurement. However, the ($\flat_2$) seems to be a measurement. This example will help you understand that cogito proposition "I think, therefore I am" in Chapter 8.

(E_2) **Space and time are not the most basic words in QL (i.e., in science).**

QL agrees to Leibniz's relationalism concerning space-time (Sec.9.7). That is,

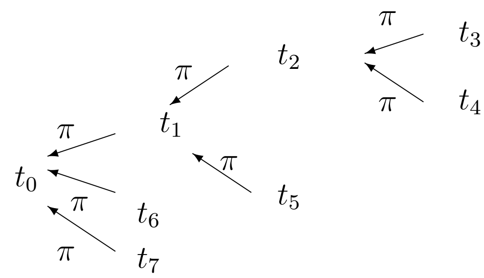
($\sharp$) [The metaphysical space-time]
I think that **Leibniz's relationalism** says that

($\sharp_1$) Space $\mathcal{S}$ is a kind of state space $\mathfrak{S}^p(\mathcal{A}^*)$ (Recall Axiom 1 in Sec, 1.1.3) a parameter is regarded as a state (*cf.* Sec. 12.2)

($\sharp_2$) Time $\mathcal{T}$ is an order of occurring in succession which changes one after another. That is, $\mathcal{T}$ is a kind of tree T(i.e., semi-ordered tree structure). (Recall Axiom 2 in Sec, 1.1.3).

($\sharp_3$) "Causality precedes time"(*cf.* Note 9.4)

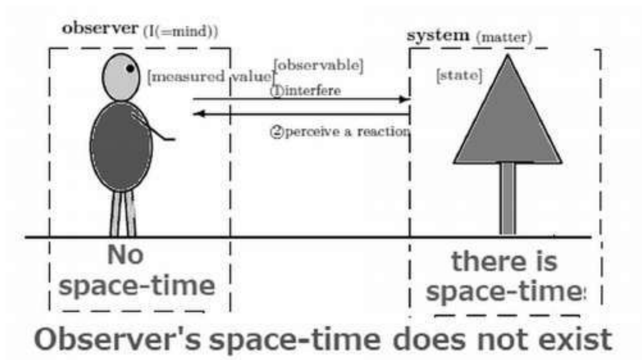
Therefore, if " thing " does not exits, the space-time does not exist.



Also, QL (i.e., Axiom 1 and 2 in Sec, 1.1.3) says nothing about observer's time and place. Therefore, observer's space-time does not exist.

there is no tense in QL.

If QL is seen as a mind-matter dualism, then space-time can be considered to belong to 'matter (=thing)'. That is, we see:



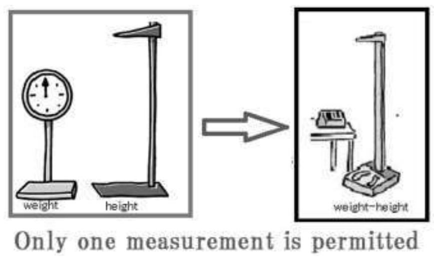
Thus, the question: "When, where and by whom was the measured value obtained?" is out of the scope of QL. Thus, words such as "now," "here," and "I" should not be used in a scientific proposition. If you are going to use it, you need to be very careful.

The "tense" is a treasure trove of word play (*cf.* Augustinus "Only the present exists", McTaggart's paradox, Russell's five-minute hypothesis in ref. [76]).

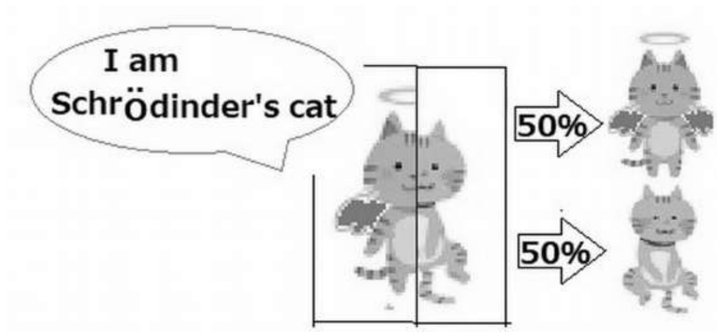
(E_3) In measurement theory, "observable(=measuring instrument ≈ body)" is the most important than "measured value(≈mind)" and "state(≈matter)" in (D_3). The prototype of observables is Plato's Idea. Also, statistics is not philosophical because it does not have "observables". I would like to remind you of the following written in Note 2.4.

$$\underset{\text{dualistic science}}{\overset{\text{quantum language}}{\boxed{\mathsf{M}(\mathsf{O}{=}(X, 2^X, F), S_{[\rho]})\ )}}} \xrightarrow[\text{Elimination of observable } \mathsf{O}]{} \underset{\text{applied math}}{\overset{\text{statistics}}{\boxed{(X, 2^X, P_\rho(\cdot))}}}$$

(E_4) **Only one measurement is permitted**. The post-measurement state (as it is disturbed by the measurement) is not meaningful. Therefore, only one measurement can be made. I like to think that this was discovered by Parmenides and Kolmogorov (*cf.* Chap. 2 in ref.[76]).



♠**Note 3.5.** This is particularly essential in quantum measurements. In classical measurements where the measurement object is large, (E_4) can sometimes be neglected, considering that the influence of the measurement is small. In principle, however, (E_4) is common to both classical and quantum systems.



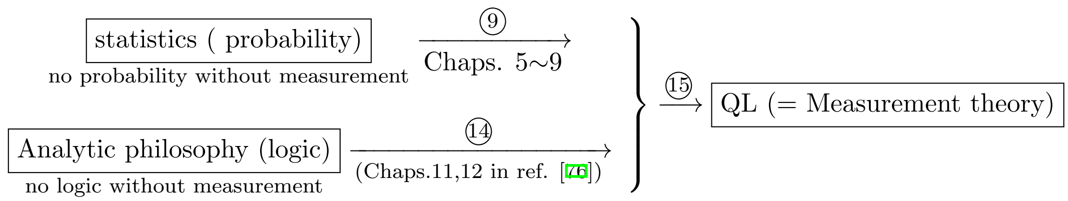
For the virtual wave function collapse, see Sec. 10.2, or

- ref.[59] S. Ishikawa, Linguistic Copenhagen interpretation of quantum mechanics; Projection Postulate, JQIS, Vol. 5, No.4, 150-155, 2015, DOI: 10.4236/jqis.2015.54017 (`http://www.scirp.org/Journal/PaperInformation.aspx?PaperID=62464`)

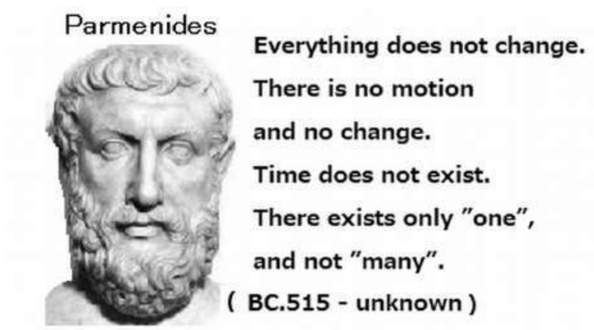
♠**Note 3.6.** This virtual wave function collapse in ref.[59] is powerful as follows. The Schrödinger cat is the most famous paradox in quantum mechanics. However, we are not bothered by this paradox since the state after measurement is not described in quantum language.

(E_5) There is no probability without measurement. Also, the measurement cannot be measured.

There is no logic without measurement. See Figure 0.1 in Preface such as



(E_6) There is one state and it never moves. Therefore, there is no time (time is just an ordered parameter (*cf.* Axiom 2 in Sec.9.3)). Therefore, we always use the Heisenberg picture (basically we do not use the Schrödinger picture), etc. It is still surprising that Parmenides mentioned almost all of the Copenhagen interpretations 2500 years ago (cf. Sec. 2.3 in ref. [76]).

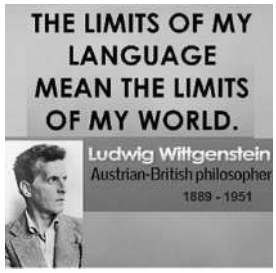


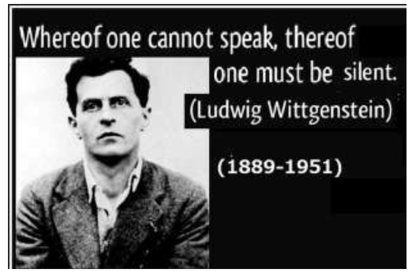
and so on.

If we believe that quantum language is the final destination of dualistic idealism (cf. ⑪ and ⑮ in Figure 0.1 (in Preface)), it seems natural to think as follows

(E_7) Explanations of the (linguistic) Copenhagen interpretation (E_0) to (E_6) are not sufficient (*cf.* Wittgenstein's paradox in Sec. 12.2 of ref. [76]). As with national laws and sports rules, the Copenhagen interpretation cannot be described completely. They must be amended whenever inadequacies are exposed. Many philosophers' aphorisms (especially dualistic idealism) can be seen as expressions of the Copenhagen interpretation. For example, the following Wittgenstein sayings can be regarded as Copenhagen interpretations.

- What we cannot speak about in QL, we must pass over in silence.
- The limits of QL means the Limits of our world

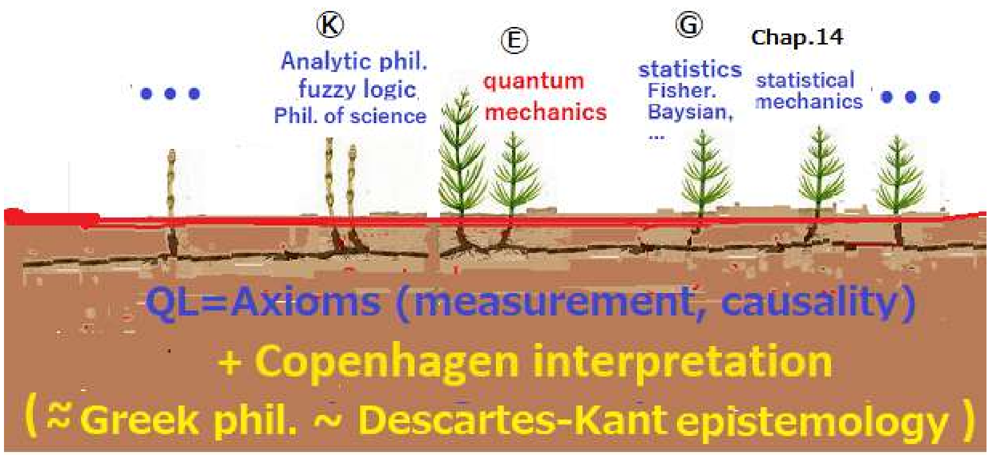




though these may be more appropriately described as the 'spirit of QL' rather than the 'Copenhagen Interpretation'.[1]

(E_8) As we saw above, there is a strong affinity between the Copenhagen interpretation and "quotes from philosophers". As we saw above, there is a strong affinity between the Copenhagen Interpretation and the "quotes of the (epistemological) philosophers". This is not surprising, since the goal of both was to establish the "doctrine of dualistic idealism."Without knowing the above diagram, it is not surprising that some philosophers have dismissed epistemology as metaphysics, as Wittgenstein did.

Thus, we think that



Also, I think that there is no 'perfect Copenhagen Interpretation', in the same sense that there is 'no perfect manual'.

♠**Note 3.7.** — (i): Historically, the Copenhagen interpretation is closely related to the 'projection postulate' (i.e., 'the problem of wave-function collapse'). Thus we must solve the following Problem:

(a) Why does the wave function contract after a measurement?

If I answer "by the Copenhagen interpretation, the post-measurement state is meaningless", the reader will be disappointed. And thus, the reader should then cautiously ask the following question.

[1] As mentioned in ref. [76], I think the only thing Wittgenstein said in 'TLP (= ref. [113])' was the spirit of QL. Since he is a philosopher, it is natural for him to talk about "spirit."

(b) Why does the wave function appear to contract after a measurement?

This will be answered in Sec. 10.2.

(ii): Readers may ask:

(♯) Is there a perfect 'linguistic Copenhagen interpretation'?

I cannot say for sure either way, however, I say that it is possible to offer 'the linguistic Copenhagen interpretation' that is satisfactory from a practical point of view. The various extraordinary situations discussed in the philosophy of mind are useful in refining the Copenhagen interpretation. However, it is not "precision" but "ease of use" that is important to the manual. Again see Figure 0.1 in Preface, and confirm that we are not in physics but in dualistic idealism.

(iii): Some may say that the Heisenberg cut should be added in the linguistic Copenhagen interpretation. At present, I am reluctant to make this suggestion, since it is closely connected to Axiom 1. But it may not be a matter determined by my preferences.

(iv): The projection postulate does not belong to the linguistic Copenhagen interpretation. This is proved from the linguistic Copenhagen interpretation (*cf.* Sec.10.2).

♠**Note 3.8.** Kolmogorov's probability theory (*cf.* ref. [85]) starts from the following spell:

($\sharp_1$) Let $(X, \mathcal{F}, P)$ be a probability space. Then, the probability that an event $\Xi (\in \mathcal{F})$ happens is given by $P(\Xi)$.

Through trial and error, Kolmogorov found his extension theorem, whose spirit says

($\sharp_2$) *Only one probability space is permitted.*

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

(the most fundamental theorem)		(the linguistic Copenhagen interpretation)
Probability theory	(correspondence) $\longleftrightarrow$	Quantum language
(Only one probability space is permitted)		(Only one measurement is permitted)

In this sense, we want to say

($\sharp_4$) Kolmogorov is one of the main discoverers of the linguistic Copenhagen interpretation.[2]

Therefore, I am optimistic to believe that the linguistic Copenhagen interpretation "Only one measurement is permitted" can be acquired, through trial and error, if we start from Axioms 1 and 2. In fact, I myself acquired skill of linguistic Copenhagen interpretation with this method. So, I consider, as mentioned in (E_1), that we can theoretically do well without the linguistic Copenhagen interpretation. Also, one of our purposes may be to assert the superiority of Axioms 1 and 2 to the above spell ($\sharp_1$).

[2]Since the mainstream of philosophy is dualistic idealism, it is not surprising that many philosophers have stated something similar to the 'linguistic Copenhagen interpretation'. However, it is surprising that the mathematician Kolmogorov said something similar: "Only one measurement is possible."

3.2 Tensor operator algebra

3.2.1 Tensor product of Hilbert space

Recall that the linguistic Copenhagen interpretation 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":

$$H \otimes K = \Big\{ g = \sum_{(m,n)\in\mathbb{N}^2} \alpha_{m,n} e_m \otimes f_n \;\Big|\; ||g||_{H\otimes K} \equiv [\sum_{(m,n)\in\mathbb{N}^2} |\alpha_{m,m}|^2]^{1/2} < \infty \Big\} \tag{3.1}$$

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

$$\begin{aligned} \langle e_{m_1} \otimes f_{n_1}, e_{m_2} \otimes f_{n_2} \rangle_{H\otimes K} &\equiv \langle e_{m_1}, e_{m_2} \rangle_H \cdot \langle f_{n_1}, f_{n_2} \rangle_K \\ &= \begin{cases} 1 & (m_1, n_1) = (m_2, n_2) \\ 0 & (m_1, n_1) \neq (m_2, n_2) \end{cases} \end{aligned} \tag{3.2}$$

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 3.3. [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

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

Thus, we see that

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

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

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

This Example 3.3 can be easily generalized as follows.

Theorem 3.4. [Finite tensor Hilbert space]

$$\mathbb{C}^{m_1} \otimes \mathbb{C}^{m_2} \otimes \cdots \otimes \otimes \mathbb{C}^{m_n} = \mathbb{C}^{\sum_{k=1}^{n} m_k} \tag{3.3}$$

Theorem 3.5. [Concrete tensor Hilbert space]

$$L^2(\Omega_1, \nu_1) \otimes L^2(\Omega_2, \nu_2) = L^2(\Omega_1 \times \Omega_2, \nu_1 \otimes \nu_2) \tag{3.4}$$

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

Definition 3.6. [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:

$$\Big\{ \bigotimes_{k=1}^{\infty} e_k^{b(k)} \;\Big|\; b : \mathbb{N} \to \mathbb{N} \text{ is a map} \Big\} \tag{3.5}$$

Hence we can define the infinite Hilbert space $\bigotimes_{k=1}^{\infty} H_k$ such that it has the CONS (3.5).

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 3.7. [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 3.8. [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:

$$[\mathcal{C}(H_1) \subseteq B(H_1) \subseteq B(H_1)] \otimes [\mathcal{C}(H_2) \subseteq B(H_2) \subseteq B(H_2)]$$

For further information see my homepage

$$=[\mathfrak{C}(H_1 \otimes H_2) \subseteq B(H_1 \otimes H_2) \subseteq B(H_1 \otimes H_2)]$$

[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 3.9. 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

$$\bigotimes_{k=1}^\infty B(H_k) = B(\bigotimes_{k=1}^\infty H_k) \tag{3.6}$$

Theorem 3.10. The followings hold:

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

♠**Note 3.9.** 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 Exercise — Only one measurement is permitted

In this section, we examine the linguistic Copenhagen 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:

$$[\text{only one measurement}] \Longrightarrow \begin{cases} \text{only one observable (=measuring instrument)} \\ \\ \text{only one state} \end{cases} \tag{3.7}$$

♠**Note 3.10.** 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

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

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

3.3.1 "Observable is only one" and simultaneous measurement

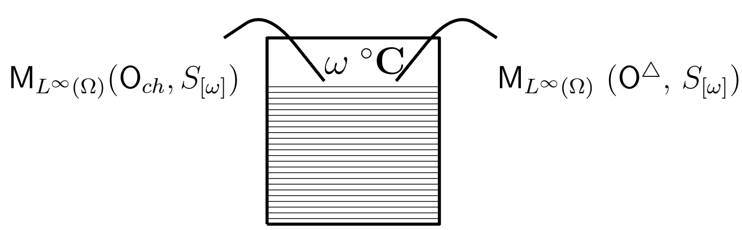
Recall the measurement Example 2.31 (Cold or hot?) and Example 2.32 (Approximate temperature), and consider the following situation:

(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

$$\begin{cases} (\sharp_1)\text{: } \mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{ch}{=}(\{c,h\}, 2^{\{c,h\}}, F_{ch}), S_{[\omega]}) \text{ in Example 2.31} \\ \\ (\sharp_2) : \mathsf{M}_{L^\infty(\Omega)} \; (\mathsf{O}^{\triangle} = (\mathbb{N}_{10}^{100},\, 2^{\mathbb{N}_{10}^{100}}, G^{\triangle}),\, S_{[\omega]}) \text{ in Example 2.32} \end{cases}$$



However, as mentioned in the linguistic Copenhagen interpretation,

$$\textbf{"only one measurement"} \Longrightarrow \textbf{"only one observable"}$$

Thus, we have the following problem.

Problem 3.11. 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 3.12. [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 3.13. [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:

$$\widehat{F}(\Xi_1 \times \Xi_2 \times \cdots \times \Xi_n) = F_1(\Xi_1) \cdot F_2(\Xi_2) \cdots F_n(\Xi_n) \qquad (3.8)$$
$$(\ \forall \Xi_k \in \mathcal{F}_k \ (k = 1, 2, \ldots, n))$$

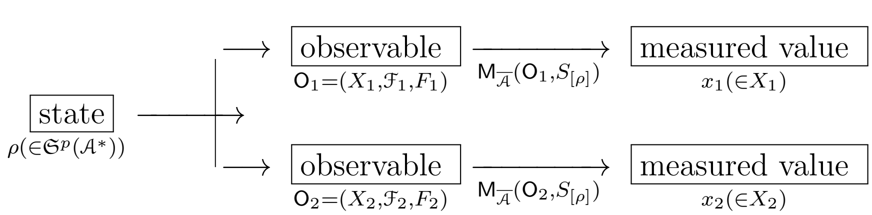
$\widehat{\mathsf{O}}$ is also denoted by $\times_{k=1}^n \mathsf{O}_k$, $\widehat{F} = \times_{k=1}^n F_k$. Also, the measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\times_{k=1}^n \mathsf{O}_k, S_{[\rho]})$ is called the simultaneous measurement. Here, it should be noted that

- 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:

(b)


However, according to the linguistic Copenhagen 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.

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

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

Answer 3.14. [The answer to Problem 3.11] 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 Example 2.31) and triangle observable $\mathsf{O}^{\triangle} = (Y(=\mathbb{N}_{10}^{100}), 2^Y, G^{\triangle})$ (in Example 2.32). 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

$$\text{that the measured value} \begin{bmatrix} (\text{c}, \text{about } 50\ {}^\circ\text{C}) \\ (\text{c}, \text{about } 60\ {}^\circ\text{C}) \\ (\text{h}, \text{about } 50\ {}^\circ\text{C}) \\ (\text{h}, \text{about } 60\ {}^\circ\text{C}) \end{bmatrix} \text{is obtained is given by} \begin{bmatrix} 0.125 \\ 0.125 \\ 0.375 \\ 0.375 \end{bmatrix} \tag{3.9}$$

That is because

$$\begin{aligned} &[(F_{ch} \times G^{\triangle})(\{(\text{c}, \text{about } 50\ {}^\circ\text{C})\})](55) \\ =&[F_{ch}(\{\text{c}\})](55) \cdot [G^{\triangle}(\{\text{about } 50\ {}^\circ\text{C}\})](55) = 0.25 \cdot 0.5 = 0.125 \end{aligned}$$

and similarly,

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

♠**Note 3.11.** 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 3.15 and Heisenberg's uncertainty principle in §4.4).

Example 3.15. [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 3.16. [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 2.25(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 $y = \widetilde{g}(\omega)$. □

For further information see my homepage

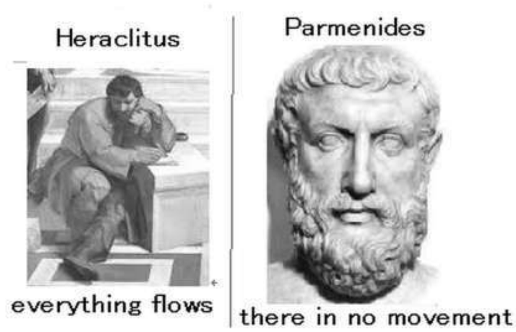
3.3.2 "State does not move" and quasi-product observable

We consider that

$$\textbf{"only one measurement"} \Longrightarrow \textbf{"state does not move"}$$

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"



For Heraclitus and Parmenides, see Sec. 9.1 or, more precisely, ref. [76].

Review 3.17. [= 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 3.3).

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

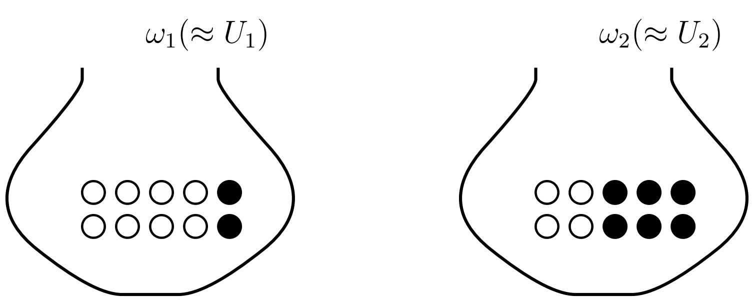


Figure 3.3: Urn problem

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.

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

$$\begin{aligned} U_1 &\quad \cdots \quad \text{"the urn with the state } \omega_1\text{"} \\ U_2 &\quad \cdots \quad \text{"the urn with the state } \omega_2\text{"} \end{aligned}$$

 For further information see my homepage

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 $\mathsf{O}_{wb}\big(\equiv (X \equiv \{w, b\}, 2^{\{w,b\}}, F_{wb})\big)$ in $L^\infty(\Omega)$ by

$$\begin{aligned} &[F_{wb}(\{w\})](\omega_1) = 0.8, \qquad && [F_{wb}(\{b\})](\omega_1) = 0.2, \\ &[F_{wb}(\{w\})](\omega_2) = 0.4, \qquad && [F_{wb}(\{b\})](\omega_2) = 0.6. \end{aligned} \tag{3.10}$$

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 3.18. (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.

$$(w, w) \;\; (w, b) \;\; (b, w) \;\; (b, b)$$

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 It 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 $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{wb} \times \mathsf{O}_{wb}, S_{[\delta_{\omega_2}]})$ 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 3.19. (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.

$$(w,w) \ \ (w,b) \ \ (b,w) \ \ (b,b)$$

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 (Definition3.13) 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

$$\widehat{F}(\Xi_1 \times \Xi_2 \times \cdots \times \Xi_n) = F_1(\Xi_1)F_2(\Xi_2)\cdots F_n(\Xi_n)$$

For further information see my homepage

$$(\forall \Xi_k \in \mathcal{F}_k, \forall k = 1, 2, \ldots, n)$$

The following definition ("quasi-product observable") is a kind of simultaneous observable:

Definition 3.20. [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

$$F_{12\ldots n}(X_1 \times \cdots \times X_{k-1} \times \Xi_k \times X_{k+1} \times \cdots \times X_n) = F_k(\Xi_k) \tag{3.11}$$
$$(\forall \Xi_k \in \mathcal{F}_k, \forall k = 1, 2, \ldots, n)$$

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

$$\overset{\text{qp}}{\underset{k=1,2,\ldots,n}{\times}} \mathsf{O}_k = (\overset{n}{\underset{k=1}{\times}} X_k, \boxtimes_{k=1}^n \mathcal{F}_k, \overset{\text{qp}}{\underset{k=1,2,\ldots,n}{\times}} F_k)$$

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.

Answer 3.21. [The answer to Problem 3.18] Define the quasi-product observable $\mathsf{O}_{wb} \overset{\text{qp}}{\times} \mathsf{O}_{wb} = (\{w, b\} \times \{w, b\}, 2^{\{w,b\}\times\{w,b\}}, F_{12}(= F_{wb} \overset{\text{qp}}{\times} F_{wb}))$ 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 by $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{wb} \overset{\text{qp}}{\times} \mathsf{O}_{wb}, S_{[\delta_{\omega_2}]})$ is given by

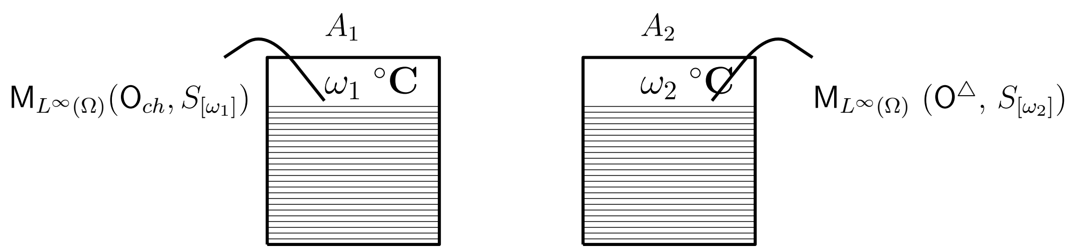
$$\begin{bmatrix} [F_{12}(\{(w,w)\})](\omega_2) = \frac{4\times3}{90} \\ [F_{12}(\{(w,b)\})](\omega_2) = \frac{4\times6}{90} \\ [F_{12}(\{(b,w)\})](\omega_2) = \frac{4\times6}{90} \\ [F_{12}(\{(b,b)\})](\omega_2) = \frac{6\times5}{90} \end{bmatrix}$$

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

$$\begin{cases} (\sharp_1)\text{: } \mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{ch}{=}(\{c,h\},2^{\{c,h\}},F_{ch}),S_{[\omega_1]}) \text{ in Example 2.31} \\ \\ (\sharp_2) : \mathsf{M}_{L^\infty(\Omega)} \ (\mathsf{O}^{\triangle} = (\mathbb{N}_{10}^{100},\ 2^{\mathbb{N}_{10}^{100}}, G^{\triangle}),\ S_{[\omega_2]}) \text{ in Example 2.32} \end{cases}$$



However, as mentioned in the above,

"only one state" must be demanded.

Thus, we have the following problem.

Problem 3.22. 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 3.23. [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

$$\begin{aligned} &\widetilde{F}(\Xi_1 \times \Xi_2 \times \cdots \times \Xi_n) = F_1(\Xi_1) \otimes F_2(\Xi_2) \otimes \cdots \otimes F_n(\Xi_n) \\ &\qquad \forall \Xi_k \in \mathcal{F}_k \ (k=1,2,\ldots,n) \end{aligned} \tag{3.12}$$

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:

$$
\text{(b)} \quad \begin{cases} \underset{\rho_1(\in \mathfrak{S}^p(\mathcal{A}_1^*))}{\boxed{\text{state}}} \longrightarrow \underset{\mathsf{O}_1}{\boxed{\text{observable}}} \xrightarrow[\mathsf{M}_{\overline{\mathcal{A}}_1}(\mathsf{O}_1, S_{[\rho_1]})]{} \underset{x_1(\in X_1)}{\boxed{\text{measured value}}} \\ \\ \underset{\rho_2(\in \mathfrak{S}^p(\mathcal{A}_2^*))}{\boxed{\text{state}}} \longrightarrow \underset{\mathsf{O}_2}{\boxed{\text{observable}}} \xrightarrow[\mathsf{M}_{\overline{\mathcal{A}}_2}(\mathsf{O}_2, S_{[\rho_2]})]{} \underset{x_2(\in X_2)}{\boxed{\text{measured value}}} \end{cases}
$$

However, according to the linguistic Copenhagen 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 3.24. [The answer to Problem 3.22] 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

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

For further information see my homepage

That is because

$$
\begin{aligned}
&[(F_{ch} \otimes G^{\triangle})(\{(\mathrm{c}, \text{about } 50\ {}^\circ\mathrm{C})\})](25, 55) \\
=&[F_{ch}(\{\mathrm{c}\})](25) \cdot [G^{\triangle}(\{\text{about } 50\ {}^\circ\mathrm{C}\})](55) = 0.75 \cdot 0.5 = 0.375
\end{aligned}
$$

Thus, similarly,

$$
\begin{aligned}
&[(F_{ch} \otimes G^{\triangle})(\{(\mathrm{c}, \text{about } 60\ {}^\circ\mathrm{C})\})](25, 55) = 0.75 \cdot 0.5 = 0.375 \\
&[(F_{ch} \otimes G^{\triangle})(\{(\mathrm{h}, \text{about } 50\ {}^\circ\mathrm{C})\})](25, 55) = 0.25 \cdot 0.5 = 0.125 \\
&[(F_{ch} \otimes G^{\triangle})(\{(\mathrm{h}, \text{about } 60\ {}^\circ\mathrm{C})\})](25, 55) = 0.25 \cdot 0.5 = 0.125
\end{aligned}
$$

Remark 3.25. Also, for example, putting $(\omega_1, \omega_2) = (55, 55)$, we see:

(e) the probability that a measured value $\begin{bmatrix} (\mathrm{c}, \text{about } 50\ {}^\circ\mathrm{C}) \\ (\mathrm{c}, \text{about } 60\ {}^\circ\mathrm{C}) \\ (\mathrm{h}, \text{about } 50\ {}^\circ\mathrm{C}) \\ (\mathrm{h}, \text{about } 60\ {}^\circ\mathrm{C}) \end{bmatrix}$ 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

$$
\begin{cases}
[F_{ch}(\{\mathrm{c}\})](55) \cdot [G^{\triangle}(\{\text{about } 50\ {}^\circ\mathrm{C}\})](55) = 0.25 \cdot 0.5 = 0.125 \\
[F_{ch}(\{\mathrm{c}\})](55) \cdot [G^{\triangle}(\{\text{about } 60\ {}^\circ\mathrm{C}\})](55) = 0.25 \cdot 0.5 = 0.125 \\
[F_{ch}(\{\mathrm{h}\})](55) \cdot [G^{\triangle}(\{\text{about } 50\ {}^\circ\mathrm{C}\})](55) = 0.75 \cdot 0.5 = 0.375 \\
[F_{ch}(\{\mathrm{h}\})](55) \cdot [G^{\triangle}(\{\text{about } 60\ {}^\circ\mathrm{C}\})](55) = 0.75 \cdot 0.5 = 0.375
\end{cases}
\tag{3.13}
$$

Note that this is the same as Answer 3.14 (cf. Note 3.12 later).

The following theorem is clear. But, the assertion is significant.

Theorem 3.26. [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

$$
(\mathop{\times}_{k=1}^{n} X_k,\ \boxtimes_{k=1}^{n} \mathcal{F}_k, \bigotimes_{k=1}^{n} P_k^\omega)
\tag{3.14}
$$

For further information see my homepage

Proof. It is clear, and thus we omit the proof. (Also, see Note 3.12 later.) □

Example 3.27. [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

$$\begin{bmatrix} \langle u, F^z(\{\uparrow\})u\rangle\langle u, F^x(\{\uparrow\})u\rangle = p_1 p_2 \\ \langle u, F^z(\{\uparrow\})u\rangle\langle u, F^x(\{\downarrow\})u\rangle = p_1(1-p_2) \\ \langle u, F^z(\{\downarrow\})u\rangle\langle u, F^x(\{\uparrow\})u\rangle = (1-p_1)p_2 \\ \langle u, F^z(\{\downarrow\})u\rangle\langle u, F^x(\{\downarrow\})u\rangle = (1-p_1)(1-p_2) \end{bmatrix}$$

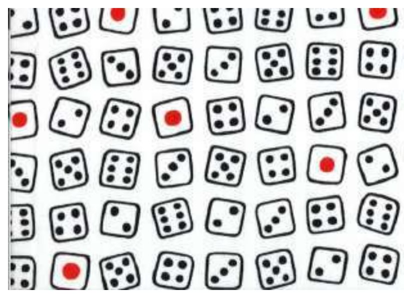
where $p_1 = |\alpha_1|^2, \quad p_2 = \frac{1}{2}(|\alpha_1|^2 + \widehat{\alpha}_1\alpha_2 + \alpha_1\widehat{\alpha}_2 + |\alpha_2|^2)$

For further information see my homepage

♠**Note 3.12.** Theorem 3.26 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 Copenhagen interpretation of quantum systems

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

- $$\underset{\text{(=quantum language)}}{\boxed{\text{measurement theory}}} := \underbrace{\overset{[\text{Axiom 1}]}{\underset{(cf.\ \S 2.7)}{\boxed{\text{Measurement}}}} + \overset{[\text{Axiom 2}]}{\underset{(cf.\ \S 9.3)}{\boxed{\text{Causality}}}}}_{\text{a kind of spells (a priori judgment)}}$$
$$+ \underbrace{\overset{[\text{quantum linguistic Copenhagen interpretation}]}{\underset{(cf.\ \S 3.1)}{\boxed{\text{Linguistic Copenhagen interpretation}}}}}_{\text{manual to use Axioms, or common knowledge in dualistic idealism world}}$$

Measurement theory says :

- Describe every phenomenon based on Axioms 1 and 2 through linguistic Copenhagen interpretation !

In this chapter, we discuss the linguistic Copenhagen interpretation (§3.1) generally, including quantum systems. I believe that the linguistic Copenhagen interpretation is common to both classical and quantum systems. The previous chapter and this chapter should not have been separated, but they were separated due to page numbers.

4.1 Kolmogorov's extension theorem and the linguistic Copenhagen interpretation

Kolmogorov's probability theory (*cf.* ref. [85]) starts from the following spell:

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

And, through trial and error, Kolmogorov found his extension theorem, whose spirit says

($\sharp_1$) *"Only one probability space is permitted"*

which surely corresponds to

($\sharp_2$) *"Only one measurement is permitted" in linguistic Copenhagen interpretation* (§3.1)

Therefore, we want to say that

($\sharp_3$) Parmenides (born around BC. 515) and Kolmogorov (1903-1987) said about the same thing.

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

$$\mathop{\times}_{\lambda\in\Lambda_2} X_\lambda \ni (x_\lambda)_{\lambda\in\Lambda_2} \mapsto (x_\lambda)_{\lambda\in\Lambda_1} \in \mathop{\times}_{\lambda\in\Lambda_1} X_\lambda \tag{4.1}$$

Especially, put $\pi_\Lambda = \pi_{\Lambda,\widehat{\Lambda}}$.

The following theorem guarantees the existence and uniqueness of the observable. It should be noted that this is due to the linguistic Copenhagen interpretation (§3.1), i.e., "only one measurement is permitted".

Theorem 4.1. [Kolmogorov extension theorem in measurement theory (cf.ref. [32])] Consider the basic structure

$$[\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)]. \tag{4.2}$$

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 $\{ \mathsf{O}_\Lambda \equiv (\times_{\lambda\in\Lambda} X_\lambda, \times_{\lambda\in\Lambda} \mathcal{F}_\lambda, F_\Lambda) \mid \Lambda \in \mathcal{P}_0(\widehat{\Lambda}) \}$ in $\overline{\mathcal{A}}$ satisfies the following "*consistency condition*":

- for any $\Lambda_1, \Lambda_2 \in \mathcal{P}_0(\widehat{\Lambda})$ such that $\Lambda_1 \subseteq \Lambda_2$,

$$F_{\Lambda_2}\big(\pi^{-1}_{\Lambda_1,\Lambda_2}(\Xi_{\Lambda_1})\big) = F_{\Lambda_1}\big(\Xi_{\Lambda_1}\big) \quad (\forall \Xi_{\Lambda_1} \in \bigtimes_{\lambda\in\Lambda_1} \mathcal{F}_\lambda). \tag{4.3}$$

Then, there uniquely exists an observable $\widetilde{\mathsf{O}}_{\widehat{\Lambda}} \equiv \big(\bigtimes_{\lambda\in\widehat{\Lambda}} X_\lambda,\ \bigtimes_{\lambda\in\widehat{\Lambda}} \mathcal{F}_\lambda,\ \widetilde{F}_{\widehat{\Lambda}}\big)$ in $\overline{\mathcal{A}}$ such that

$$\widetilde{F}_{\widehat{\Lambda}}\big(\pi^{-1}_{\Lambda}(\Xi_\Lambda)\big) = F_\Lambda\big(\Xi_\Lambda\big) \quad (\forall \Xi_\Lambda \in \bigtimes_{\lambda\in\Lambda} \mathcal{F}_\lambda,\ \forall \Lambda \in \mathcal{P}_0(\widehat{\Lambda})). \tag{4.4}$$

Proof. For the proof, see ref.[32]. □

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 $\widehat{\mathsf{O}} = (\bigtimes_{\lambda\in\widetilde{\Lambda}} X_\lambda,\ \boxtimes_{\lambda\in\widetilde{\Lambda}}\mathcal{F}_\lambda,\ \widehat{F} = \bigtimes_{\lambda\in\widetilde{\Lambda}} F_\lambda)$ uniquely exists. That is, for any finite set $\Lambda_0 (\subseteq \widetilde{\Lambda})$, it holds that

$$\widehat{F}\big((\bigtimes_{\lambda\in\Lambda_0} \Xi_\lambda) \times (\bigtimes_{\lambda\in\widetilde{\Lambda}\setminus\Lambda_0} X_\lambda)\big) = \bigtimes_{\lambda\in\Lambda_0} F_\lambda(\Xi_\lambda) \qquad (\forall \Xi_\lambda \in \mathcal{F}_\lambda, \forall \lambda \in \Lambda_0).$$

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:

(Linguistic Copenhagen interpretation)
Only one measurement is permitted

(Kolmogorov's extension theorem 4.1 in quantum language)
$\longrightarrow$ The existence of measurement

(Kolmogorov's extension theorem)
$\longrightarrow$ 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 a 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 (Corollary 4.2). 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

$$P_{\bigotimes_{k=1}^{\infty} \rho}(\Xi_1 \times \Xi_2 \times \cdots \times \Xi_n \times (\bigtimes_{k=n+1}^{\infty} X)) = \bigtimes_{k=1}^{n} {}_{\mathcal{T}r(H)}\Big(\rho, F(\Xi_k)\Big)_{B(H)} \tag{4.7}$$
$$(\ \forall \Xi_k \in \mathcal{F} = \mathcal{F}_{\rho}, (\ k = 1, 2, \ldots, n), n = 1, 2, 3 \cdots)$$

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

$$\mathfrak{S}^p(C_0(\times_{k=1}^{\infty}\Omega)^*)\Big(\approx \bigtimes_{k=1}^{\infty} \Omega\Big). \tag{4.9}$$

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

$$P_{\bigotimes_{k=1}^\infty \rho}(\Xi_1 \times \Xi_2 \times \cdots \times \Xi_n \times (\mathop{\times}_{k=n+1}^{\infty} X)) = \mathop{\times}_{k=1}^{n} [F(\Xi_k)](\omega) \tag{4.10}$$

$$(\forall \Xi_k \in \mathcal{F} = \mathcal{F}_\rho, (k = 1, 2, \ldots, n), n = 1, 2, 3 \cdots)$$

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 (originally due to J. Bernoulli)] 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 (Corollary 4.2), 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 $(X^{\mathbb{N}}, \boxtimes_{k=1}^\infty \mathcal{F}, \bigotimes_{k=1}^\infty P_\rho)$. Furthermore, we see

(A) for any $f \in L^1(X, P_\rho)$, put

$$D_f = \Big\{ (x_1, x_2, \ldots) \in X^{\mathbb{N}} \mid \lim_{n \to \infty} \frac{f(x_1) + f(x_2) + \cdots + f(x_n)}{n} = E(f) \Big\} \tag{4.11}$$
$$(\text{ where } \quad E(f) = \int_X f(x) P_\rho(dx) \ )$$

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:

$$(\sharp_2) \quad \begin{cases} \text{G. Galileo} & \cdots \text{the originator of the realistic world view} \\ \\ \text{J. Bernoulli} & \cdots \text{the originator of the linguistic world view} \end{cases}$$

the law of large numbers
J. Bernoulli (1654-1705)

4.2.2 Mean, variance, unbiased variance

Definition 4.7. [population mean, population variance, sample mean, sample 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:

$$\textit{population mean } (\mu_{\mathsf{O}}^{\rho}) : E[\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O} = (\mathbb{R}, \mathcal{B}_{\mathbb{R}}F), S_{[\rho]})] = \int_{\mathbb{R}} x P_\rho(dx) (= \mu) \tag{4.16}$$

$$\textit{population variance } ((\sigma_{\mathsf{O}}^{\rho})^2) : \quad V[\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O} = (\mathbb{R}, \mathcal{B}_{\mathbb{R}}F), S_{[\rho]})] = \int_{\mathbb{R}} (x-\mu)^2 P_\rho(dx) \tag{4.17}$$

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

$$\textit{sample distribution } (\nu_n) : \nu_n = (1/n) \sum_{i=1}^n \delta_{x_i} \in \mathcal{M}_{+1}(X)$$

$$\textit{sample mean } (\overline{\mu}_n) : \overline{E}[\otimes_{k=1}^n \mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}, S_{[\rho]})] = (1/n) \sum_{i=1}^n x_i (= \overline{\mu}) = \int_{\mathbb{R}} x \nu_n(dx)$$

$$\textit{sample variance } (s_n^2) : \overline{V}[\otimes_{k=1}^n \mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}, S_{[\rho]})] = (1/n) \sum_{i=1}^n (x_i - \overline{\mu})^2$$

$$= \int_{\mathbb{R}} (x - \overline{\mu})^2 \nu_n(dx)$$

$$\textit{unbiased variance } (u_n^2) : \overline{U}[\otimes_{k=1}^n \mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}, S_{[\rho]})] = (1/(n-1)) \sum_{i=1}^n (x_i - \overline{\mu})^2$$
$$= \frac{n}{n-1} \int_{\mathbb{R}} (x - \overline{\mu})^2 \nu_n(dx)$$

///

Under the above preparation, we have:

Theorem 4.8. [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 (Theorem 4.5) says that

$$(4.16) = \text{population mean } (\mu_{\mathsf{O}}^\rho) = \lim_{n\to\infty} (1/n) \sum_{i=1}^n x_i =: \overline{\mu} = \text{sample mean}$$

$$(4.17) = \text{population variance } (\sigma_{\mathsf{O}}^\rho) = \lim_{n\to\infty} (1/n) \sum_{i=1}^n (x_i - \mu_{\mathsf{O}}^\rho)^2$$
$$= \lim_{n\to\infty} (1/n) \sum_{i=1}^n (x_i - \overline{\mu})^2 =: \text{sample variance}$$

Example 4.9. [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

$$\text{population mean}(\mu_{\mathsf{O}_A}^{\rho_u}) : E[\mathsf{M}_{B(H)}(\mathsf{O}_A, S_{[|u\rangle\langle u|]})] = \int_{\mathbb{R}} \lambda \langle u, F_A(d\lambda)u\rangle = \langle u, Au\rangle \tag{4.18}$$

$$\text{population variance}((\sigma_{\mathsf{O}_A}^{\rho_u})^2) : \quad V[\mathsf{M}_{B(H)}(\mathsf{O}_A, S_{[|u\rangle\langle u|]})]$$
$$= \int_{\mathbb{R}} (\lambda - \langle u, Au\rangle)^2 \langle u, F_A(d\lambda)u\rangle$$
$$= \|(A - \langle u, Au\rangle)u\|^2 \tag{4.19}$$

4.2.3 Robertson's uncertainty principle

Now we can introduce Robertson's uncertainty principle as follows.

Theorem 4.10. [Robertson's uncertainty principle (parallel measurement) (*cf.* ref. [102])] 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

$$\sigma_{A_1}^{\rho_u} \cdot \sigma_{A_2}^{\rho_u} \geqq \frac{1}{2}|\langle u, (A_1A_2 - A_2A_1)u\rangle| \qquad (\forall |u\rangle\langle u| = \rho_u, \ \ \|u\|_H = 1),$$

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 $[A_1, A_2] \equiv A_1A_2 - A_2A_1$, 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,

$$\begin{aligned} & |\langle u, [A_1, A_2]u\rangle| = |\langle u, (A_1A_2 - A_2A_1)u\rangle| \\ =& \Big|\Big\langle u, \Big((A_1 - \langle u, A_1 u\rangle)(A_2 - \langle u, A_2 u\rangle) - (A_2 - \langle u, A_2 u\rangle)(A_1 - \langle u, A_1 u\rangle)\Big)u\Big\rangle\Big| \\ \leq& 2\|(A_1 - \langle u, A_1 u\rangle)u\| \cdot \|(A_2 - \langle u, A_2 u\rangle)u\|. \end{aligned}$$

□

4.3 Heisenberg's uncertainty principle

For further information see my homepage

4.3.1 Why is Heisenberg's uncertainty principle famous ?

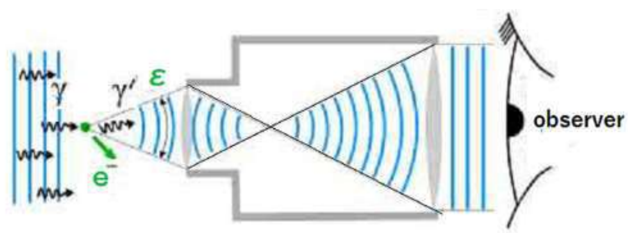
Heisenberg's uncertainty principle is as follows.

Proposition 4.11. [Heisenberg's uncertainty principle (*cf.* ref. [20]:1927)]
(This will be justified as Heisenberg's uncertainty relation (= Theorem 4.16))

(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.

$$\Delta_x \cdot \Delta_p \doteqdot \hbar (= \text{Plank constant}/2\pi \doteqdot 1.5547 \times 10^{-34} Js). \tag{4.21}$$



Heisenberg's thought experiment
with γ-ray microscope

It is generally said that the above were discovered by the following Heisenberg's thought experiment with γ-ray microscope.

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 4.11) is meaningless. That is because, for example,

(♯) The approximate measurement and "error" in Proposition 4.11 are not defined.

This will be improved in Theorem 4.16 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

If it be so, why is Heisenberg's uncertainty principle (Proposition 4.11) famous ?

For further information see my homepage

I think

> Heisenberg's uncertainty principle (Proposition 4.11) was used as a 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, the cogito proposition due to Descartes

I think, therefore I am.

is also meaningless (*cf.* ref. [76]). However, it is certain that the cogito proposition built the foundation of modern science.

♠**Note 4.2.** Heisenberg's uncertainty principle (Proposition 4.11) may include contradiction (*cf.* ref. [26]), if we think as follows

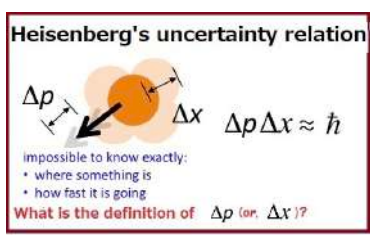
(♯) it is "natural" to consider that

$$\Delta_x = |x - \widetilde{x}|, \quad \Delta_p = |p - \widetilde{p}|,$$

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

It was long believed that Robertson's uncertainty relation (= Theorem 4.10) was the mathematical expression of Heisenberg's uncertainty principle (= Proposition 4.11). However, with Theorem 4.16 later, the true nature of Heisenberg's uncertainty principle was revealed at once.

4.3.2.1 Preparation

In this section, we will propose the mathematical formulation (Theorem 4.16) of Heisenberg's uncertainty principle (Proposition 4.11). 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:

(B$_1$) $\mathsf{M}_{B(H)}(\mathsf{O}_{A_1} :=(\mathbb{R}, \mathcal{B}, F_{A_1}), S_{[\rho_u]}) \xrightarrow[\text{expectation}]{\text{by (4.18)}} \langle u, A_1 u\rangle$

(B$_2$) $\mathsf{M}_{B(H)}(\mathsf{O}_{A_2} :=(\mathbb{R}, \mathcal{B}, F_{A_2}), S_{[\rho_u]}) \xrightarrow[\text{expectation}]{\text{by (4.18)}} \langle u, A_2 u\rangle$

$$(\forall \rho_u = |u\rangle\langle u| \in \mathfrak{S}^p(\mathcal{C}(H)^*))$$

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 a 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

$$\textit{the tensor state } \widehat{\rho}_{us} = |u \otimes s\rangle\langle u \otimes s|.$$

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 4.12. 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

$$\begin{aligned}\Delta^{u \otimes s}_{\widehat{N}_i} &= \|\widehat{N}_i(u \otimes s)\| = \|(\widehat{A}_i - A_i \otimes I)(u \otimes s)\| \\ \overline{\Delta}^{u \otimes s}_{\widehat{N}_i} &= \|(\widehat{N}_i - \langle u \otimes s, \widehat{N}_i(u \otimes s)\rangle)(u \otimes s)\| \\ &= \|((\widehat{A}_i - A_i \otimes I) - \langle u \otimes s, (\widehat{A}_i - A_i \otimes I)(u \otimes s)\rangle)(u \otimes s)\|.\end{aligned} \tag{4.24}$$

Here the following inequality:

$$\Delta^{\widehat{\rho}_{us}}_{\widehat{N}_i} \geq \overline{\Delta}^{\widehat{\rho}_{us}}_{\widehat{N}_i} \tag{4.25}$$

is well-known. By the commutative condition (4.22), (4.23) implies that

$$[\widehat{N}_1, \widehat{N}_2] + [\widehat{N}_1, A_2 \otimes I] + [A_1 \otimes I, \widehat{N}_2] = -[A_1 \otimes I, A_2 \otimes I]. \tag{4.26}$$

For further information see my homepage

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.* Theorem 4.10), estimated as follows:

$$2\overline{\Delta}_{\widehat{N}_1}^{\widehat{\rho}_{us}} \cdot \overline{\Delta}_{\widehat{N}_2}^{\widehat{\rho}_{us}} \geq |\langle u\otimes s, [\widehat{N}_1, \widehat{N}_2](u\otimes s)\rangle|. \tag{4.27}$$

4.3.2.2 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 4.13. [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 4.14. [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 mea-

surement 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

$$(\mathrm{D}_1) \text{ and } (\mathrm{D}_2) \text{ as substitutes of } (\mathrm{C}_1) \text{ and } (\mathrm{C}_2).$$

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 $\mathsf{M}_{B(H\otimes K)}(\mathsf{O}_{\widehat{A}_1} \times \mathsf{O}_{\widehat{A}_2}, S_{[\widehat{\rho}_{us}]})$

Lemma 4.15. 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

$$\Delta^{\widehat{\rho}_{us}}_{\widehat{N}_i} = \overline{\Delta}^{\widehat{\rho}_{us}}_{\widehat{N}_i} \tag{4.31}$$

$$\langle u \otimes s, [\widehat{N}_1, A_2 \otimes I](u \otimes s)\rangle = 0 \qquad (\forall u \in H) \tag{4.32}$$

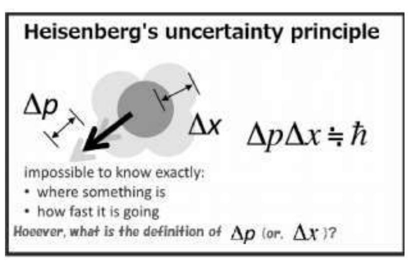
$$\langle u \otimes s, [A_1 \otimes I, \widehat{N}_2](u \otimes s)\rangle = 0 \quad (\forall u \in H) \tag{4.33}$$

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 4.16. [The mathematical formulation of Heisenberg's uncertainty principle] Let A_1 and A_2 be (unbounded) self-adjoint operators on a Hilbert space H. In general, the simultaneous measurement of A_1 and A_2does not exists. However. 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.

For further information see my homepage

(ii) And further, we have the following inequality (i.e., Heisenberg's uncertainty principle).

$$\Delta^{\widehat{\rho}_{us}}_{\widehat{N}_1} \cdot \Delta^{\widehat{\rho}_{us}}_{\widehat{N}_2} (= \overline{\Delta}^{\widehat{\rho}_{us}}_{\widehat{N}_1} \cdot \overline{\Delta}^{\widehat{\rho}_{us}}_{\widehat{N}_2}) = \|(\widehat{A}_1 - A_1 \otimes I)(u \otimes s)\| \cdot \|(\widehat{A}_2 - A_2 \otimes I)(u \otimes s)\|$$
$$\geq \frac{1}{2} |\langle u, [A_1, A_2] u \rangle| \quad (\forall u \in H \text{ such that } ||u|| = 1). \tag{4.35}$$

(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. [26]: S. Ishikawa, Rep. Math. Phys. Vol.29(3), 1991, pp.257–273,

As shown in the above (4.33), the proof (ii) is easy (*cf.* refs. [35, 95]), but the proof (i) is not easy (*cf.* refs. [7, 35]).

4.3.3 Without the average value coincidence condition

Now we have the complete form of Heisenberg's uncertainty relation as Theorem 4.16. To be compared with Theorem 4.16, we should note that the conventional Heisenberg's uncertainty relation (= Proposition 4.11) is ambiguous. Wrong conclusions are sometimes derived from the ambiguous statement (= Proposition 4.11). For example, in some books of physics, it is concluded that EPR-experiment (Einstein, Podolsky and Rosen [14], 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, some may consider that

[II] EPR-experiment says that the position and the momentum of a certain "particle" can be measured simultaneously and exactly.

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 refs. [26, 35]. 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 $A_1 : L^2(\mathbb{R}^2_{(q_1,q_2)}) \to L^2(\mathbb{R}^2_{(q_1,q_2)})$ and $A_2 : L^2(\mathbb{R}^2_{(q_1,q_2)}) \to L^2(\mathbb{R}^2_{(q_1,q_2)})$ 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 4.16 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 4.16) 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

For further information see my homepage

(*cf.* Remark 3 in ref.[26], or p.316 in ref. [35]).
Also, we add the following remark.

Remark 4.17. 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 (4.25), we get, by Robertson's uncertainty principle (4.20),

$$2\overline{\Delta}^{\widehat{\rho}_{us}}_{\widehat{N}_1} \cdot \sigma(A_2; u) \geq |\langle u \otimes s, [\widehat{N}_1, A_2 \otimes I](u \otimes s)\rangle| \tag{4.44}$$

$$2\overline{\Delta}^{\widehat{\rho}_{us}}_{\widehat{N}_2} \cdot \sigma(A_1; u) \geq |\langle u \otimes s, [A_{\otimes} I, \widehat{N}_2](u \otimes s)\rangle| \tag{4.45}$$

$$(\forall u \in H \text{ such that } ||u|| = 1)$$

and, from (4.25), (4.26), (4.44),(4.45), we can get the following inequality

$$\begin{aligned} &\Delta^{\widehat{\rho}_{us}}_{\widehat{N}_1} \cdot \Delta^{\widehat{\rho}_{us}}_{\widehat{N}_2} + \Delta^{\widehat{\rho}_{us}}_{\widehat{N}_2} \cdot \sigma(A_1; u) + \Delta^{\widehat{\rho}_{us}}_{\widehat{N}_1} \cdot \sigma(A_2; u) \\ \geq &\overline{\Delta}^{\widehat{\rho}_{us}}_{\widehat{N}_1} \cdot \overline{\Delta}^{\widehat{\rho}_{us}}_{\widehat{N}_2} + \overline{\Delta}^{\widehat{\rho}_{us}}_{\widehat{N}_2} \cdot \sigma(A_1; u) + \overline{\Delta}^{\widehat{\rho}_{us}}_{\widehat{N}_1} \cdot \sigma(A_2; u) \\ \geq &\frac{1}{2}|\langle u, [A_1, A_2]u\rangle| \quad (\forall u \in H \text{ such that } ||u|| = 1). \end{aligned} \tag{4.46}$$

Since we do not assume the average value coincidence condition, it is a matter of course that this (4.46) is rougher 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:$\eta(A_2,$ 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.* ref. [96]). He asserted that this inequality is a faithful description of Heisenberg's thought experiment (due to γ-ray microscope).

♠**Note 4.3.** After my discovery of Heisenberg's uncertainty relation, I shifted my research focus to fuzzy logic (*cf.* Chap. 7). That is because I think that the theoretical study of EPR-paradox (in the next section) will not make progress for the next 50 years. At the time, there were signs that research on quantum computers was popular. However, the research on quantum computers is rather technological, so I have no advantage because I cannot conduct experiments. See Sec. 12.1.2 [Zadeh and Kalman; The problem of universals in 20th century] in ref. [76] for why I thought 'fuzzy logic' was promising.

4.4 EPR-paradox (1935) and faster-than-light

4.4.1 EPR-paradox

Next, let us explain EPR-paradox (Einstein–Podolsky–Rosen: refs. [14, 107]). 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 away from each other.

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

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

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

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

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

Thus, we get the measurement $\mathsf{M}_{B(\mathbb{C}^2 \otimes \mathbb{C}^2)}(\mathsf{O} \otimes \mathsf{O}, S_{[\rho]})$. Born's quantum measurement theory says :

For further information see my homepage

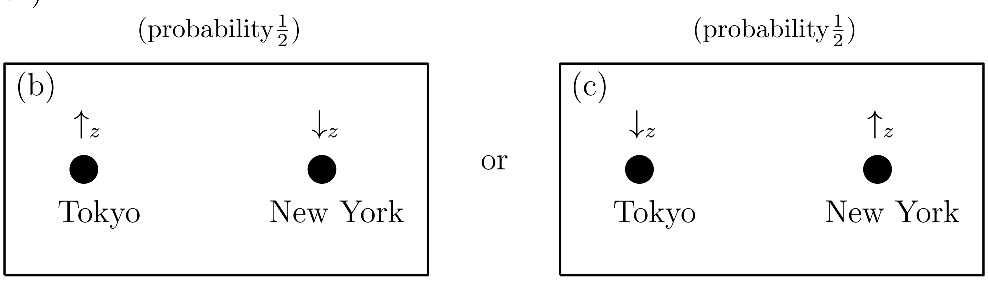
When the parallel measurement $\mathsf{M}_{B(\mathbb{C}^2\otimes\mathbb{C}^2)}(\mathsf{O}\otimes\mathsf{O}, S_{[s]})$ is taken,
the probability that the measured value $\begin{bmatrix} (\uparrow,\uparrow) \\ (\downarrow,\uparrow) \\ (\uparrow,\downarrow) \\ (\downarrow,\downarrow) \end{bmatrix}$ is obtained 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 x_1 and 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).



After measurement ?

This fact is, figuratively speaking, explained as follows:

- Immediately after the particle (or, wave function) in Tokyo is measured and the measured value $\uparrow_z$ [resp. $\downarrow_z$] is observed, the particle (or, wave function) in Tokyo informs the particle (or, wave function) 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.* ref. [107]). 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").

For further information see my homepage

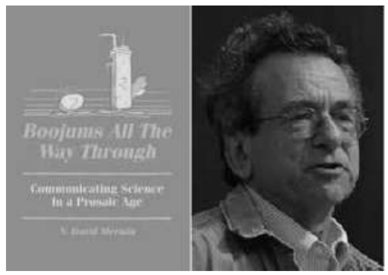
♠**Note 4.4.** [Shut up and calculate]. The above argument may suggest that there is something faster than light. However, when faster-than-light appears, our standing point is

Stop being bothered

This is not only our opinion but also most physicists'. In fact, in Mermin's book [94], 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 Copenhagen interpretation (§3.1) is the true colors of "the Copenhagen interpretation". That is because I also consider that

(c) As mentioned before, the Copenhagen interpretation is a manual to use Axioms 1 and 2. The word "manual" means that you can learn how to use [Axioms 1 and 2] by trial and error without looking at a manual, and Dr. Mermin's famous "Shut up and calculate!" can be considered a similar definition of the Copenhagen interpretation.

♠**Note 4.5.** It is difficult to actually perform EPR-experiment exactly in this form. Using the singlet state $\rho_0 = |\psi_s\rangle\langle\psi_s| \ \big(\in \mathfrak{S}^p(B(\mathbb{C}^2 \otimes \mathbb{C}^2)^*)\big)$, where

$$\psi_s = (e_1 \otimes e_2 - e_2 \otimes e_1)/\sqrt{2}$$

In 1966, J.S.Bell proposed Bell's inequality (which makes EPR paradox considerably easier to verify experimentally). In 1982, Aspect, A. et al. actually carried out experimental verification and showed that 'there is something faster than light', earning them the Nobel Prize in Physics for 2022 .

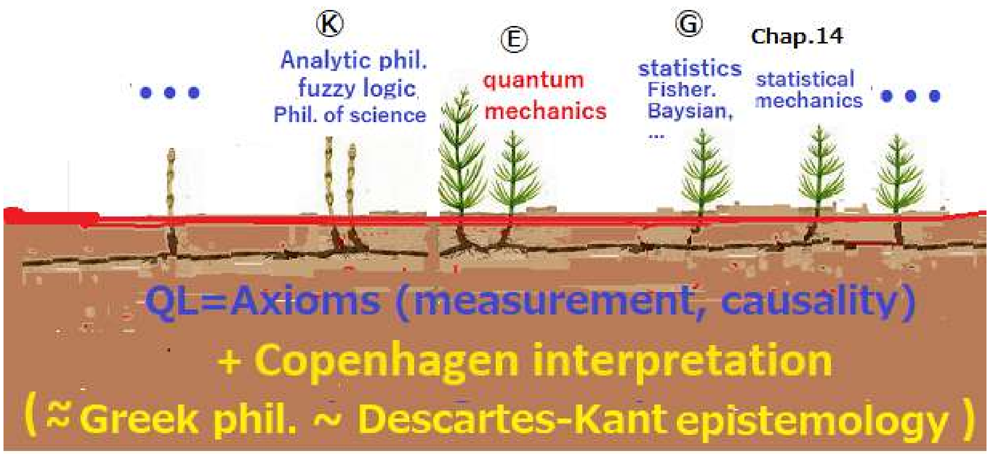
John Steward Bell (1928-1990) (Wiki)

More than half a century has passed since Bell's discovery, but there has not been a single step forward in this time. We are still waiting for the emergence of a genius of Einstein's level.

Chapter 5

Why does statistics work? : Fisher statistics (I)

Recall the following figure (Figure 0.2 in preface):



The following two problems are one of the most fundamental in science.

($\sharp_1$) Why does statistics work in our world?

($\sharp_2$) Why does fuzzy logic work in our world?

These two are answered by ⑨ and ⑭ in the Figure above such as

($\flat$) both statistics and fuzzy logic hold since QL holds in our world.

Especially, the problem ($\sharp_1$) was, for the first time, solved in ref.[33]. In this chapter (and Chaps 6 an 7), I review ($\flat$) for statistics.

Also, it should be noted that

- theoretically, statistics is to be formulated within quantum language. However, the way in which statistics can be understood using probability theory (= theory of random variables) is practical and not to be dismissed.

Though in this section, we devote ourselves to discuss "Statistics vs. Quantum language", the outcome of the loser and winner is easily predictable. That is because

"Statistics (with no axioms) vs. Quantum language (with axioms)"

5.1 Statistics is, after all, urn problems

5.1.1 Population (=system) ↔ parameter (=state)

Let us start with the following Note (i.e. QL and statistics).

♠**Note 5.1.**

The following is a part of Table 2.1:

dualism \ key-words	[A](= **mind**)	[B](Mediating of A and C) **(body)**	[C](= **matter**)
quantum mechanics QL (scientific dualism)	observer [measured value] $[x(\in X)]$	measuring instrument [observable] $[\mathsf{O} = (X, \mathcal{F}, F)]$	particle (system) [state] $\rho(\in \mathfrak{S}^p(\mathcal{A}^*))$
classical QL (scientific dualism)	observer [measured value] $[x(\in X)]$	measuring instrument [observable] $[\mathsf{O} = (X, \mathcal{F}, F)]$	particle (system) [state] $\delta_\omega \approx \omega(\in \Omega)$
statistics (incomplete dualism)	person to try [sample] $[x(\in X)]$	trial / /	population [parameter] $\omega(\in \Omega)$

Axiom 1 (in classical quantum language) says that

($\sharp_1$) the probability that a measured $x(\in X)$ obtained by a measurement $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O} = (X, \mathcal{F}, F), S_{[\delta_{\omega_0}]})$ belongs to $\Xi(\in \mathcal{F})$ is given by $[F(\Xi)](\omega_0)$.

Also, statistic say that

($\sharp_2$) the probability that a sample $x(\in X)$ obtained from a population with a parameter $\omega_0(\in \Omega)$ is given by $P_\omega(\Xi)$, if it holds $P_\omega(\Xi) = [F(\Xi)](\omega_0)$ $(\forall \omega \in \Omega, \forall \Xi \in \mathcal{F})$

Thus, in statistics, the concept of 'observable $\mathsf{O} = (X, \mathcal{F}, F)$' does not appear on the surface. In this sense, statistics does not belong to the class of dualism.

////

For further information see my homepage

Example 5.1. The density functions of the Japanese male's height and the American male's height are denoted by f_J and f_A, respectively. That is,

$$\int_\alpha^\beta f_J(x)dx = \frac{\text{number of Japanese males whose heights are from } \alpha \text{ to } \beta}{\text{total number of Japanese males}}$$
$$\int_\alpha^\beta f_A(x)dx = \frac{\text{number of American males whose heights are from } \alpha \text{ to } \beta}{\text{total number of American males}}$$

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 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 $\nu(\{\omega_J\}) = a, \;\; \nu(\{\omega_A\}) = b \;\; (a, b > 0)\Big)$. 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

$$\begin{array}{lll} \delta_{\omega_J} & \cdots & \text{"the state of the set } U_1 \text{ of all Japanese males"}, \\ \delta_{\omega_A} & \cdots & \text{"the state of the set } U_2 \text{ of all American males"}, \end{array}$$

and thus, we have the following identification (that is, Figure 5.1):

$$U_1 \approx \delta_{\omega_J}, \qquad U_2 \approx \delta_{\omega_A}$$

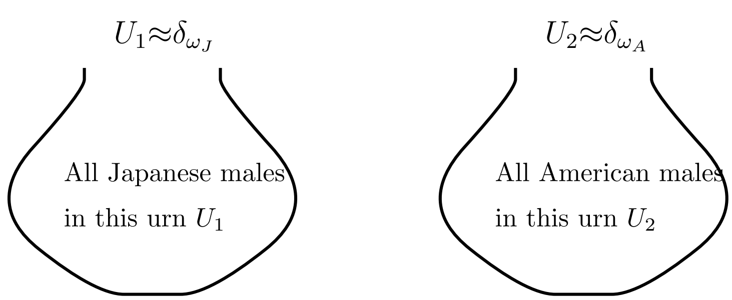


Figure 5.1: Population $\approx$ urn ($\leftrightarrow$ state)

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 quantum language by

(B) The probability that a measured value obtained by the measurement $\begin{bmatrix} \mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_h, S_{[\omega_J]}) \\ \mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_h, S_{[\omega_A]}) \end{bmatrix}$ belongs to an interval $[\alpha, \beta)$ is given by

$$\begin{bmatrix} {}_{C_0(\Omega)^*}\Big(\delta_{\omega_J}, F_h([\alpha,\beta))\Big)_{L^\infty(\Omega,\nu)} = [F_h([\alpha,\beta))](\omega_J) \\ {}_{C_0(\Omega)^*}\Big(\delta_{\omega_A}, F_h([\alpha,\beta))\Big)_{L^\infty(\Omega,\nu)} = [F_h([\alpha,\beta))](\omega_A) \end{bmatrix}.$$

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

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

$$[G_\sigma(\Xi)](\omega) = \frac{1}{\sqrt{2\pi}\sigma} \int_\Xi \exp\Big[-\frac{1}{2\sigma^2}(x-\omega)^2 \Big] dx \qquad (\forall \Xi \in \mathcal{B}_X(=\mathcal{B}_{\mathbb{R}}),\ \forall \omega \in \Omega(=\mathbb{R})) \tag{5.1}$$

where $\mathcal{B}_{\mathbb{R}}$ is the Borel field. For example,

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\sigma}^{\sigma} e^{-\frac{x^2}{2\sigma^2}} dx = 0.683..., \qquad \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-2\sigma}^{2\sigma} e^{-\frac{x^2}{2\sigma^2}} dx = 0.954...,$$

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-1.96\sigma}^{1.96\sigma} e^{-\frac{x^2}{2\sigma^2}} dx \doteqdot 0.95 \qquad \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-2.58\sigma}^{2.58\sigma} e^{-\frac{x^2}{2\sigma^2}} dx \doteqdot 0.99$$

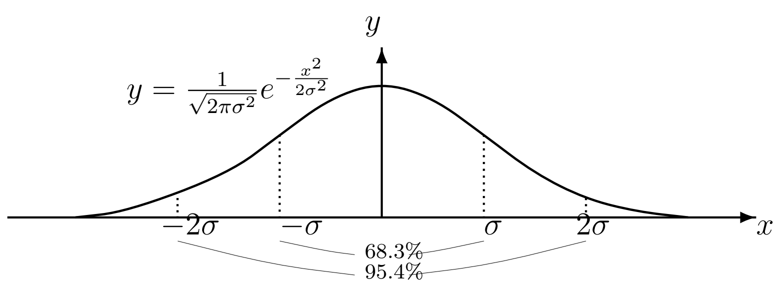


Figure 5.2: Error function

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

$$K = \{(\omega, \omega, \ldots, \omega) \in \Omega^n \mid \omega \in \Omega\} (\subseteq \Omega^n). \tag{5.2}$$

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 $L^\infty(\Omega)$. That is,

$$\begin{aligned}
&[(\mathop{\times}_{k=1}^{n} G_\sigma)(\Xi_1 \times \Xi_2 \times \cdots \times \Xi_n)](\omega) = \mathop{\times}_{k=1}^{n} [G_\sigma(\Xi_k)](\omega) \\
=& \mathop{\times}_{k=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \int_{\Xi_k} \exp\Big[- \frac{1}{2\sigma^2}(x_k - \omega)^2 \Big] dx_k \\
&(\forall \Xi_k \in \mathcal{B}_X (= \mathcal{B}_{\mathbb{R}}), \ \forall \omega \in \Omega (= \mathbb{R}))
\end{aligned} \tag{5.3}$$

Then, for each $(x_1, x_2, \cdots, x_n) \in X^n (= \mathbb{R}^n)$, define

$$\begin{aligned}
\overline{x}_n &= \frac{x_1 + x_2 + \cdots + x_n}{n} \\
U_n^2 &= \frac{(x_1 - \overline{x}_n)^2 + (x_2 - \overline{x}_n)^2 + \cdots + (x_n - \overline{x}_n)^2}{n-1},
\end{aligned}$$

5.2 The reverse relation between Fisher and Born

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 (=Example 2.34), 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].

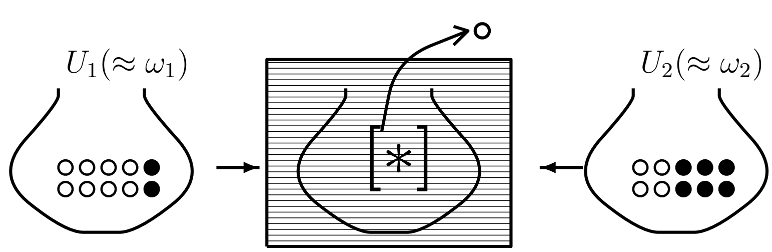


Figure 5.3: Pure measurement (Fisher's maximum likelihood method)

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. However, though easy, 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) $\mathsf{M}_{\overline{\mathcal{A}}}$ $(\mathsf{O}{=}(X, \mathcal{F}, F),\ S_{[*]})$

Furthermore, 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

$$\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}, S_{[*]}) = \mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}, S_{[*]}(\!(\mathfrak{S}^p(\mathcal{A}^*))\!)).$$

For further information see my homepage

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,

$$(\mathrm{B}_1) \qquad (\text{observable } [\mathsf{O}], \text{ state } [\omega (\in \Omega)]) \xrightarrow[\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[\omega]})]{\text{measurement}} \text{measured value } [x (\in X)]$$

On the other hand, the inference is "the view from the back", that is,

$$(\mathrm{B}_2) \qquad (\text{observable } [\mathsf{O}], \text{ measured value } [x \in \Xi (\in \mathcal{F})]) \xrightarrow[\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[*]})]{\text{inference}} \text{ state } [\omega (\in \Omega)]$$

In this sense, we say that

the inference problem is the reverse problem of measurement.

Therefore, it suffices to image Fig. 5.4.

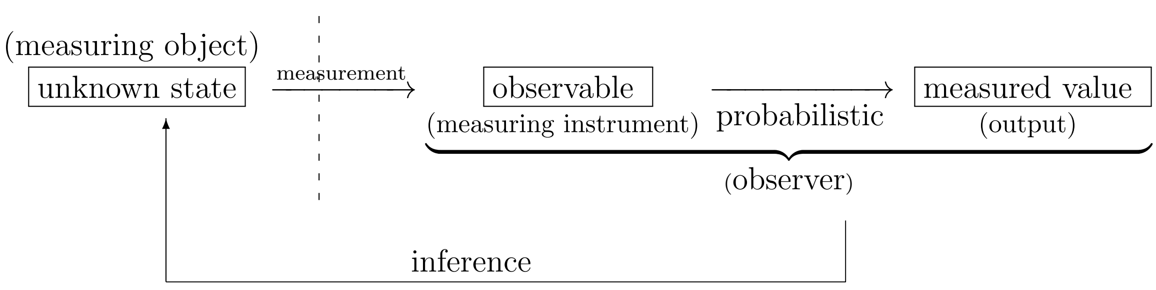


Figure 5.4: The image of inference

In order to answer the above problem 5.4, we shall describe Fisher maximum likelihood method in 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.

$$\rho_0(H(\Xi \times Y)) = \max_{\rho \in K} \rho(H(\Xi \times Y)). \tag{5.4}$$

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 less likely than (ii)". Thus, there is a reason to infer that $[*] = \omega_2$. Therefore, the ρ_0 in (5.4) 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 $(\in \mathcal{F})$.

(a) Then, there is a reason to infer that the unknown state state $[*]$ is ω_0 $(\in \Omega)$ such that

$$[F(\Xi)](\omega_0) = \max_{\omega \in \Omega} [F(\Xi)](\omega). \tag{5.5}$$

Or more generally,

(b) if it holds that $[F(\Xi)](\omega_1) < [F(\Xi)](\omega_2)$, then ω_2 should be chosen.

(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

$$f(x,\omega) = \inf_{\omega_1 \in K} \Big[\lim_{\Xi \ni x, [F(\Xi)](\omega_1) \neq 0, \Xi \to \{x\}} \frac{[F(\Xi)](\omega)}{[F(\Xi)](\omega_1)} \Big]. \tag{5.6}$$

Then, there is a reason to infer that $[*] = \omega_0 (\in K)$ such that $f(x_0, \omega_0) = 1$.

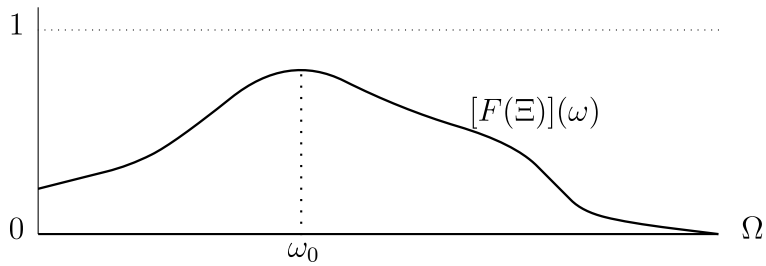


Figure 5.5: Fisher maximum likelihood method

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

$$P(\Gamma) = \frac{[H(\Xi)](\omega_0) \times [G(\Gamma)](\omega_0)}{[H(\Xi)](\omega_0) \times [G(Y)](\omega_0)} = [G(\Gamma)](\omega_0) \quad (\forall \Gamma \in \mathcal{G}). \tag{5.7}$$

And, from the arbitrariness of O_2, there is a reason to infer that

$$[*] = \delta_{\omega_0} (\underset{\text{identification}}{\approx} \omega_0).$$

□

♠**Note 5.2.** The linguistic Copenhagen interpretation says that the state after measurement is nonsense. In this sense, the readers may consider that

($\sharp_1$) Theorem 5.6 is also nonsense

However, we say that

($\sharp_2$) in the sense of (5.7), Theorem 5.6 should be accepted.

or

($\sharp_3$) as far as classical systems are concerned, it suffices to believe in Theorem 5.6

However, in the quantum case, the above discussion is related to the famous paradox concerning the Schrödinger cat. This is solved in Sec. 10.2 'the wavefunction collapse', which is one of the most important results in this book.

Answer 5.7. [The answer to Problem 5.2 by Fisher's maximum likelihood method]
You do not know the urn behind the curtain is. Assume that you pick up a white ball from the urn. Which urn do you think is more likely, U_1 or U_2 ?

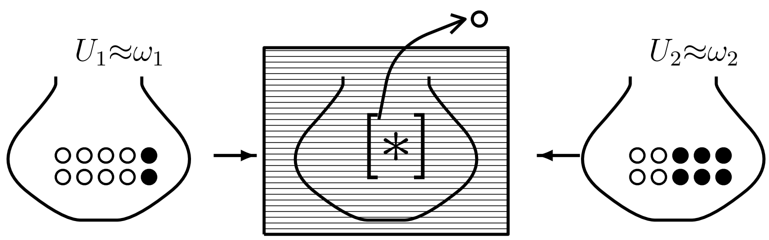


Figure 5.6: Pure measurement (Fisher's maximum likelihood method)

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

$$\begin{aligned} &[F_{wb}(\{w\})](\omega_1) = 0.8, \qquad && [F_{wb}(\{b\})](\omega_1) = 0.2 \\ &[F_{wb}(\{w\})](\omega_2) = 0.4, && [F_{wb}(\{b\})](\omega_2) = 0.6 \end{aligned} \tag{5.8}$$

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 5.3.** 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 ref. [6] (1926)

(b) Fisher's great book *"Statistical Methods for Research Workers"* (1925)

For further information see my homepage

Thus, it is surprising that Fisher and Born investigated the same thing in the different fields in the same age. Throughout this book, I emphasize that Fisher's maximum likelihood method is the most fundamental method in statistics. In quantum mechanics books, Born is always given a fair assessment.

However, I find it disappointing that Fisher's maximum likelihood method is sometimes not given its due credit in books on statistics.

♠**Note 5.4.** Note 3.1.3 says that

($\sharp_1$) a statement like Axiom 1 is a (comprehensive) proposition.

Now we have the following question:

($\sharp_2$) Is a statement like Fisher's maximum likelihood method a (comprehensive) proposition?

I think that it is a comprehensive proposition though I do not have a clear explanation. Also, see Sec. 7.1.

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 the consequences 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 balls 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 ?

Furthermore,

(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 ?

For further information see my homepage

In what follows, we shall answer the above problems (i) and (ii) in 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. Furthermore, define the observable $\mathsf{O}=(\{w,b\},2^{\{w,b\}},F)$ in $C(\Omega)$ such that

$$\begin{array}{lll} F(\{w\})(\omega_1)=0.8, & F(\{w\})(\omega_2)=0.4, & F(\{w\})(\omega_3)=0.1 \\ F(\{b\})(\omega_1)=0.2, & F(\{b\})(\omega_2)=0.6, & F(\{b\})(\omega_3)=0.9. \end{array}$$

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 (Theorem 5.6) 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 (Theorem 5.6), 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:

$$[C_0(\Omega) \subseteq L^\infty(\Omega,\nu) \subseteq B(L^2(\Omega,\nu))] \qquad (\text{where} \quad \Omega=\mathbb{R}) \ .$$

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

$$[G_\sigma(\Xi)](\mu)=\frac{1}{\sqrt{2\pi}\sigma}\int_{\Xi}\exp[-\frac{1}{2\sigma^2}(x-\mu)^2]dx$$

For further information see my homepage

$$(\forall \Xi \in \mathcal{B}_{\mathbb{R}}, \quad \forall \mu \in \Omega = \mathbb{R})$$

Thus, the simultaneous observable $\times_{k=1}^3 \mathsf{O}_{G_\sigma}$ (in short, $\mathsf{O}_{G_\sigma}^3$) $= (\mathbb{R}^3, \mathcal{B}_{\mathbb{R}^3}, G_\sigma^3)$ in $L^\infty(\mathbb{R})$ is defined by

$$\begin{aligned}
&[G_\sigma^3(\Xi_1 \times \Xi_2 \times \Xi_3)](\mu) = [G_\sigma(\Xi_1)](\mu) \cdot [G_\sigma(\Xi_2)](\mu) \cdot [G_\sigma(\Xi_3)](\mu) \\
=&\frac{1}{(\sqrt{2\pi}\sigma)^3} \iiint_{\Xi_1 \times \Xi_2 \times \Xi_3} \exp[-\frac{(x_1-\mu)^2+(x_2-\mu)^2+(x_3-\mu)^2}{2\sigma^2}] \\
&\times dx_1 dx_2 dx_3 \\
&(\forall \Xi_k \in \mathcal{B}_{\mathbb{R}}, k=1,2,3, \quad \forall \mu \in \Omega = \mathbb{R})
\end{aligned}$$

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:

$$[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))] \qquad (\text{where} \quad \Omega = \mathbb{R} \times \mathbb{R}_+)$$

and consider the case:

For further information see my homepage

- we know that the length of the pencil μ satisfies that $10 \le \mu \le 30$.

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 (Theorem 5.6) says that the unknownstate $[\,*\,] = (\mu_0, \sigma_0)$ such that

$$\begin{aligned}
&\frac{1}{(\sqrt{2\pi}\sigma_0)^3} \exp[-\frac{(x_1^0-\mu_0)^2 + (x_2^0-\mu_0)^2 + (x_3^0-\mu_0)^2}{2\sigma_0^2}] \\
=&\max_{(\mu,\sigma)\in[10,30]\times\mathbb{R}_+} \Big\{\frac{1}{(\sqrt{2\pi}\sigma)^3} \exp[-\frac{(x_1^0-\mu)^2 + (x_2^0-\mu)^2 + (x_3^0-\mu)^2}{2\sigma^2}]\Big\}
\end{aligned} \tag{5.9}$$

Thus, solving $\frac{\partial}{\partial \mu}\{\cdots\} = 0$, $\frac{\partial}{\partial \sigma}\{\cdots\} = 0$ we see

$$\mu_0 = \begin{cases} 10 & (\text{when } (x_1^0 + x_2^0 + x_3^0)/3 < 10\ ) \\ (x_1^0 + x_2^0 + x_3^0)/3 & (\text{when } 10 \leqq (x_1^0 + x_2^0 + x_3^0)/3 \leqq 30\ ) \\ 30 & (\text{when } 30 < (x_1^0 + x_2^0 + x_3^0)/3\ ) \end{cases} \tag{5.10}$$

$$\sigma_0 = \sqrt{\{(x_1^0 - \widetilde{\mu})^2 + (x_2^0 - \widetilde{\mu})^2 + (x_3^0 - \widetilde{\mu})^2\}/3}$$

where

$$\widetilde{\mu} = (x_1^0 + x_2^0 + x_3^0)/3. \qquad \square$$

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 (5.3)). This is essentially the same as the simultaneous observable $\mathsf{O}^n = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}^n}, \times_{k=1}^n G)$ in $L^\infty(\mathbb{R} \times \mathbb{R}_+)$. That is,

$$[(\mathop{\times}_{k=1}^{n} G)(\Xi_1 \times \Xi_2 \times \cdots \times \Xi_n)](\omega) = \mathop{\times}_{k=1}^{n} [G(\Xi_k)](\omega)$$
$$= \mathop{\times}_{k=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \int_{\Xi_k} \exp\left[- \frac{1}{2\sigma^2}(x_k - \mu)^2 \right] dx_k$$
$$(\forall \Xi_k \in \mathcal{B}_X (= \mathcal{B}_{\mathbb{R}}), \ \forall \omega = (\mu, \sigma) \in \Omega (= \mathbb{R} \times \mathbb{R}_+))$$

Assume that a measured value $x = (x_1, x_2, \ldots, x_n) (\in \mathbb{R}^n)$ is obtained by the measurement $\mathsf{M}_{L^\infty(\mathbb{R} \times \mathbb{R}_+)}(\mathsf{O}^n = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n, G^n), S_{[*]})$. The likelihood function $L_x(\mu, \sigma) (= L(x, (\mu, \sigma))$ is equal to

$$L_x(\mu, \sigma) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp[- \frac{\sum_{k=1}^n (x_k - \mu)^2}{2\sigma^2}],$$

or, in the sense of (5.6),

$$L_x(\mu, \sigma) = \frac{\frac{1}{(\sqrt{2\pi}\sigma)^n} \exp[- \frac{\sum_{k=1}^n (x_k - \mu)^2}{2\sigma^2}]}{\frac{1}{(\sqrt{2\pi}\overline{\sigma}(x))^n} \exp[- \frac{\sum_{k=1}^n (x_k - \overline{\mu}(x))^2}{2\overline{\sigma}(x)^2}]}. \tag{5.11}$$
$$(\forall x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n, \quad \forall \omega = (\mu, \sigma) \in \Omega = \mathbb{R} \times \mathbb{R}_+)$$

Therefore, we get the following likelihood equation:

$$\frac{\partial L_x(\mu, \sigma)}{\partial \mu} = 0, \quad \frac{\partial L_x(\mu, \sigma)}{\partial \sigma} = 0 \tag{5.12}$$

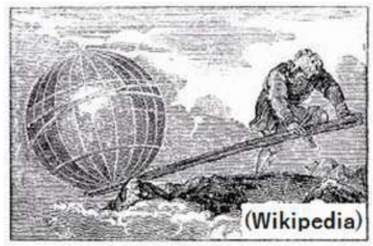
which is easily solved. That is, Fisher's maximum likelihood method (Theorem 5.6) says that the unknown state $[*] = (\mu, \sigma)$ $(\in \mathbb{R} \times \mathbb{R}_+)$ is inferred as follows.

$$\mu = \overline{\mu}(x) = \frac{x_1 + x_2 + \ldots + x_n}{n}, \tag{5.13}$$

$$\sigma = \overline{\sigma}(x) = \sqrt{\frac{\sum_{k=1}^n (x_k - \overline{\mu}(x))^2}{n}}. \tag{5.14}$$

For further information see my homepage

5.4 Moment method: useful but artificial



Let us explain the moment method (*cf.* ref. [35]) which is used as frequently as Fisher's maximum likelihood method. Consider the measurement $\mathsf{M}_{\mathcal{A}}\big(\mathsf{O} \equiv (X, \mathcal{F}, F),\ S_{[\rho]}\big)$, and its parallel measurement $\otimes_{k=1}^{n} \mathsf{M}_{\mathcal{A}}\big(\mathsf{O} \equiv (X, \mathcal{F}, F),\ S_{[\rho]}\big)$ $\big(= \mathsf{M}_{\otimes \mathcal{A}}\big(\bigotimes_{k=1}^{n} \mathsf{O} := (X^n, \mathcal{F}^n, \bigotimes_{k=1}^{n} F), S_{[\otimes_{k=1}^{n} \rho]}\big)$. 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

$$\mathcal{M}_{+1}(X) \ni \nu_n\Big(\equiv \frac{\delta_{x_1} + \delta_{x_2} + \cdots + \delta_{x_n}}{n}\Big) \doteqdot \rho(F(\cdot)) \in \mathcal{M}_{+1}(X). \tag{5.15}$$

Thus,

(A) in order to infer the unknown state $\rho(\in \mathfrak{S}^p(\mathcal{A}^*))$, it suffices to solve the equation (5.15)

For example, we have several methods to solve the equation (5.15) as follows.

(B_1) Solve the following equation:

$$\|\nu_n(\cdot) - \rho(F(\cdot))\|_{\mathcal{M}(X)} = \min\{\|\nu_n(\cdot) - \rho_1(F(\cdot))\|_{\mathcal{M}(X)} \mid \rho_1(\in \mathfrak{S}^p(\mathcal{A}^*))\} \tag{5.16}$$

(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 case 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

$$0 = \sum_{k=1}^{n} \Big| \frac{f_k(x_1) + f_k(x_2) + \cdots + f_k(x_n)}{n} - \int_X f_k(\xi)[F(d\xi)](\omega)\Big|, \tag{5.17}$$

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 called *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. [=Problem 5.2: Urn problem: by the moment 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 ?

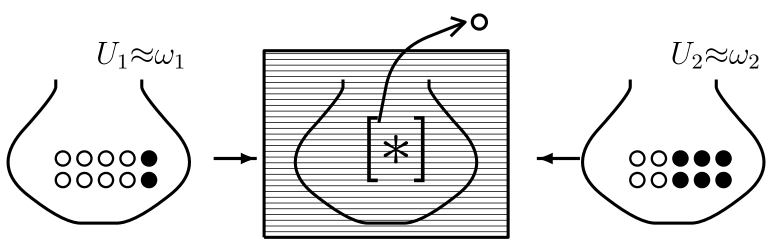


Figure 5.7: Inference(by moment method)

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

$$\nu_1(\{w\}) = 1, \quad \nu_1(\{b\}) = 0.$$

For further information see my homepage

[when the unknown state [$*$] is ω_1]

$$(5.16) = |1 - 0.8| + |0 - 0.2| = 0.4.$$

[when the unknown state [$*$] is ω_2]

$$(5.16) = |1 - 0.4| + |0 - 0.6| = 1.2.$$

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 urn is behind the curtain, 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]

$$(5.16) = |3/7 - 0.8| + |4/7 - 0.2| = 52/70.$$

[when [$*$] is ω_2]

$$(5.16) = |3/7 - 0.4| + |4/7 - 0.6| = 10/70.$$

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

$$\nu_3 = \frac{\delta_{x_1} + \delta_{x_2} + \delta_{x_3}}{3} \in \mathcal{M}_{+1}(\mathbb{R}).$$

For further information see my homepage

Put $f_1(\xi)=\xi, f_2(\xi)=\xi^2$. Then, by the moment method (5.17), 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 (5.10) concerning the normal measurement.

♠**Note 5.5.** (i): Consider the measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}{=}(X,2^X,\ F),\ S_{[*]})$, where $X=\{x_1,x_2,\dots,\ x_n\}$ is finite. Then, we see that

"Fisher's maximum likelihood method" = "moment method"

.

Proof : 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.

(ii): If we did not know Axiom 1 (in Sec. 1.2), we would not be able to answer the question, "Which is more essential, Fisher's maximum likelihood method or the moment method?"

5.5 Monty Hall problem in Fisher's maximum likelihood method

Monty Hall problem is as follows[1].

Problem 5.14. [Monty Hall problem; High school puzzle]

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

(♭) 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 ?*

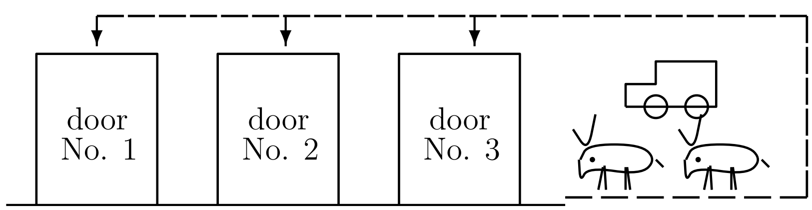
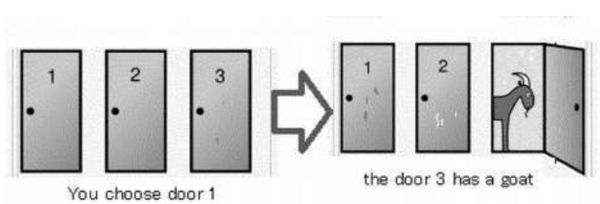


Figure 5.8: Monty Hall problem



Answer: (I believe that this answer is new, and the most fundamental in several answers of Monty Hall problem. See (ix) in Sec. 16.2).

---

[1]This section is extracted from the followings:

(a) Ref. [35]: S. Ishikawa, "Mathematical Foundations of Measurement Theory," Keio University Press Inc. 2006.

(b) Ref. [44]: 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.

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

$$\begin{aligned}
&[F_1(\{1\})](\omega_1) = 0.0, \qquad [F_1(\{2\})](\omega_1) = 0.5, \qquad [F_1(\{3\})](\omega_1) = 0.5,\\
&[F_1(\{1\})](\omega_2) = 0.0, \qquad [F_1(\{2\})](\omega_2) = 0.0, \qquad [F_1(\{3\})](\omega_2) = 1.0,\\
&[F_1(\{1\})](\omega_3) = 0.0, \qquad [F_1(\{2\})](\omega_3) = 1.0, \qquad [F_1(\{3\})](\omega_3) = 0.0,
\end{aligned} \tag{5.18}$$

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 the unknown state $[*]$ is equal to δ_{ω_2}. Thus, you should switch to door 2. This is the first answer to Problem 5.14 (Monty-Hall problem). □

♠**Note 5.6.** Examining the above example, the readers should understand that the problem "What is measurement ?" is an unreasonable demand. Thus,

we have to abandon the realistic approach, and accept the metaphysical approach.

In other words, we assert that

the concept of measurement is metaphysical.

Also, for a Bayesian approach to Monty Hall problem, see Chapter 8 and Chapter 16.

For further information see my homepage

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]

$$(5.16) = |0-0| + |0-0.5| + |1-0.5| = 1,$$

[when the unknown $[*]$ is ω_2]

$$(5.16) = |0-0| + |0-0| + |1-1| = 0,$$

[when the unknown $[*]$ is ω_3]

$$(5.16) = |0-0| + |0-1| + |1-0| = 2.$$

Thus, we can infer that $[*] = \omega_2$. That is, you should change to the Door 2. □

5.6 The two envelopes problem – High school student puzzle

This section is extracted from the following:

> Ref. [58]: 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 envelopes problem, see Chapter 8.

5.6.1 Problem (the two envelopes problem)

The following problem is the famous "two envelopes problem(*cf.* ref. [90])".

Problem 5.16. [The two envelopes 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

$$\overline{x} = \begin{cases} V_2, & (\text{ if } x = V_1), \\ V_1 & (\text{ if } x = V_2) \end{cases}$$

Assume that

(b) 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 whether $\overline{x}_1 = 2x_1$ 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 envelopes 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

$$E_{\text{other}}(\alpha) = (1/2)(\alpha/2) + (1/2)(2\alpha) = 1.25\alpha \tag{5.19}$$

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 envelopes 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 $\mathsf{O}_k = (X(= \overline{\mathbb{R}}_+), \mathcal{F}(= \mathcal{B}_{\overline{\mathbb{R}}_+}$: the Borel field$), F_k)$ in $L^\infty(\Omega, \nu)$ such that

$$[F_k(\Xi)](\omega) = \begin{cases} 1 & (\text{ if } V_k(\omega) \in \Xi) \\ 0 & (\text{ if } V_k(\omega) \notin \Xi) \end{cases}$$

$$(\forall \omega \in \Omega, \forall \Xi \in \mathcal{F} = \mathcal{B}_{\overline{\mathbb{R}}_+} \text{ i.e., the Bore field in } X (= \overline{\mathbb{R}}_+) \ )$$

Furthermore, by the hypothesis (b), define the observable $\mathsf{O} = (X, \mathcal{F}, F)$ in $L^\infty(\Omega, \nu)$ such that

$$F(\Xi) = \frac{1}{2}\Big(F_1(\Xi) + F_2(\Xi)\Big) \quad (\forall \Xi \in \mathcal{F}). \tag{5.20}$$

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 envelopes 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:

$$\Omega \ni \omega \underset{\text{(identification)}}{\longleftrightarrow} (\omega, 2\omega) \in \widehat{\Omega} \tag{5.21}$$

Furthermore, 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 $\left\{ \begin{array}{c} x = V_1(\omega) = \omega \\ x = V_2(\omega) = 2\omega \end{array} \right\}$ is obtained by $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O} = (X, \mathcal{F}, F), S_{[\omega]})$ is given by $\left\{ \begin{array}{c} 1/2 \\ 1/2 \end{array} \right\}$.

If you switch to $\left\{ \begin{array}{c} V_2(\omega) \\ V_1(\omega) \end{array} \right\}$, your gain is $\left\{ \begin{array}{c} V_2(\omega) - V_1(\omega) \\ V_1(\omega) - V_2(\omega) \end{array} \right\}$. 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 envelopes 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 investigate 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

$$f(\alpha, \omega) \equiv \inf_{\omega_1 \in \Omega} \left[\lim_{\Xi \to \{x\}, [F(\Xi)](\omega_1) \neq 0} \frac{[F(\Xi)](\omega)}{[F(\Xi)](\omega_1)} \right] = \begin{cases} 1 & (\omega = \alpha/2 \text{ or } \alpha) \\ 0 & (\text{ elsewhere }) \end{cases}$$

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

$$(B_2) \begin{cases} \text{"the probability that } [*]=\alpha/2\text{"}=1/2 \\ \text{"the probability that } [*]=\alpha\text{"}=1/2 \\ \text{"the probability that } [*] \text{ is otherwise"}=0 \end{cases}$$

Therefore, we can not calculate as (5.19):

$$(\alpha/2-\alpha)\times\frac{1}{2}+(2\alpha-\alpha)\times\frac{1}{2}=1.25\alpha$$

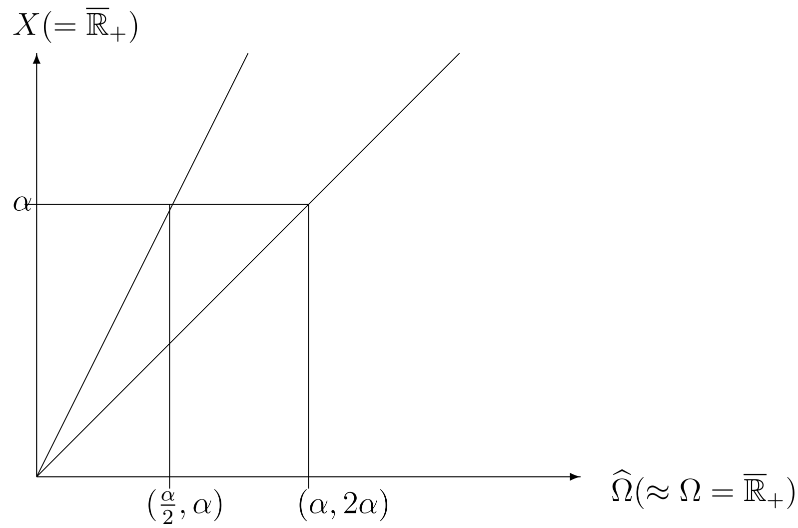


Figure 5.10: Two envelopes problem

(C_1) Thus, the sentence "with probability 1/2" in [(P1):Why is it paradoxical ?] is wrong.

Hence, we can conclude :

(C_2) *Fisher's maximum likelihood method is invalid for Problem 5.16.*

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 define 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*

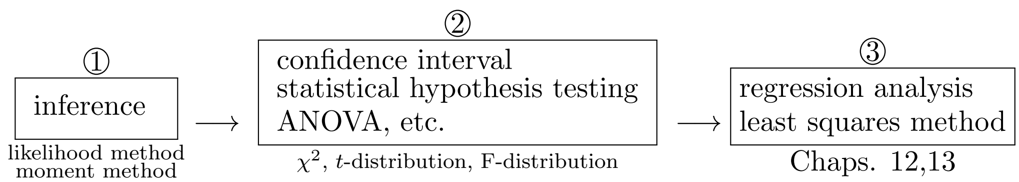
Under this hypothesis, the condition (b) may be often omitted. Also, we will again discuss the principle of equal weight in Chapters 8 and 15. Also, see Note 8.5.

For further information see my homepage

Chapter 6

Confidence interval and hypothesis testing

The following is the standard teaching schedule for university statistics courses.



In the previous chapter, we are concerned with ① (inference) in quantum language. In this chapter, we discuss ② (confidence interval and statistical hypothesis testing). This chapter is an extract from papers (refs. [51, 52, 53], etc.). As mentioned in Preface, the main purpose of this book is to assert that

(♯) Statistics is the part you write on the calculation paper when you think in quantum language.[1]

However, this field (e.g., ②) is far from my area of expertise, and moreover, I have done no more than the above-mentioned "arxiv thesis (non-peer-reviewed)". As statistics is a vast discipline, it is impossible to achieve this objective with this book alone. Therefore, my real aim is to convince readers that "from the pure theoretical point of view, statistics should be formulated in QL". And to have each reader write papers showing that various methods of statistics can be described in quantum language. If you are an expert in this field (a graduate student), you have an overwhelming advantage over me. This chapter emphasizes that in QL, it is exposed that ① and ② above are more closely related than in traditional statistics.

6.1 Review; Estimation and testing problems in conventional statistics

In this section, conventional statistical methods (confidence intervals with random variables, tests) are reviewed. And, in the next section 6.2, these are described in terms of quantum language. I assert, from the theoretical point of view, that statistics should be described in quantum language and the style of using random variables is seen as a type of powerful computational technique.

[1] I don't mean it in a negative nuance. I consider Einstein and Fisher to be the true geniuses.

♠**Note 6.1.** I think that

($\sharp_1$) the most surprising thing for mathematics students when they learn about quantum mechanics is that the concept of 'random variables' does not appear in quantum mechanics, even though probability appears frequently in quantum mechanics.

And

($\sharp_2$) the most surprising thing to physics students when they learn about probability theory in mathematics is the emphasis on 'probability = measure ($\approx$ area)' and, moreover, the invisible concept of 'random variables'.

This chapter should help the reader to somewhat bridge the gap between the two ($\sharp_2$) and ($\sharp_2$) above. After reading this chapter, readers will be convinced that the following.

($\flat_1$) physics is better at answering the question "What is probability?"
Born proposed Axiom 1 (§2.7) as the answer of "What is measurement?"

On the other hand,

($\flat_2$) mathematics (i.e., stochastic variable method) is the superior method for calculating probabilities.
In fact, if there had been no random variables method, Kolmogorov would not have been able to perform such an enormous calculation.

There may be arguments for both positions, but this book is in position ($\flat_1$) ($\approx$the quantum mechanical worldview). Of course, the most demanding thing for researchers is to develop the ability to freely cross-translate between Axiom 1 (§2.7) and random variables method. Readers are encouraged to acquire the ability to cross-translate after reading this chapter.

6.1.1 The theory of random variables

Let a triplet $(S, \mathcal{B}_S, P)$ be a probability space (i.e., $P(S) = 1$). A measurable function $X : S \to \mathbb{R}$ is called a random variable. And, let $\{X_i\}_{i=1}^{\infty}$ be independent and identically distributed random variables on S such that $\int_S |X_i(s)|^2 P(ds) < \infty$ $(i = 1, 2, ...)$.

Definition 6.1. [population mean, population variance, sample mean, sample variance][2]:
Define the population mean μ and the population variance σ^2 (or, standard deviation σ) by

$$\mu = \int_S X_i(s) P(ds) \qquad (i = 1, 2, ...), \qquad \text{(population mean)}$$
$$\sigma^2 = \int_S (X_i(s) - \mu)^2 P(ds) \qquad (i = 1, 2, ...), \qquad \text{(population variance)}$$

which are usually assumed to be unknown Further, define

$$\overline{X}_n(s) = \frac{X_1(s) + X_2(s) + ... + X_n(s)}{n} \quad \text{(sample mean)}$$

[2]This should be compared to Definition 4.7

$$SS_n(s) = (X_1(s) - \overline{X}(s))^2 + (X_2(s) - \overline{X}(s))^2 + ... + (X_n(s) - \overline{X}_n(s))^2$$
$$\frac{SS_n(s)}{n} : \quad \text{(sample variance)}$$
$$\frac{SS_n(s)}{n-1} : \quad \text{(unbiased sample variance)}$$

////

It is well-known that the law of large numbers (*cf.* Sec. 4.2) says that,

$$\mu = \lim_{n\to\infty} \frac{X_1(s) + X_2(s) + ... + X_n(s)}{n} = \lim_{n\to\infty} \overline{X}_n(s), \tag{6.1}$$

$$\begin{aligned}\sigma^2 &= \lim_{n\to\infty} \frac{(X_1(s) - \overline{X}_n)^2 + (X_2(s) - \overline{X}_n)^2 + \cdots + (X_n(s) - \overline{X}_n)^2}{n-1} \\ &= \lim_{n\to\infty} \frac{SS_n(s)}{n-1} = \lim_{n\to\infty} \frac{SS_n(s)}{n}\end{aligned} \tag{6.2}$$

6.1.2 Normal distribution

Our aim is to study formulas (6.1) and (6.2) for a not very large n. To do so, we start by summarizing our knowledge of the normal distribution as follows.

♠**Note 6.2.** In this chapter, we devote ourselves to the normal distributions. Thus, we think as follows (*cf.* Note 2.4):

- population ≈ system
  (statistics) (QL)
- parameter (=(population mean μ, standard deviation σ)) ≈ state
  (statistics) (QL)

Review 6.2. Normal distribution $N(\mu, \sigma^2)$:
Let $X : S \to \mathbb{R}$ be a random variable with normal distribution (with 'population mean' μ, 'population variance' σ^2, i.e., $N(\mu, \sigma^2)$), that is, $X : S \to \mathbb{R}$ has the following distribution: it holds that

$$[G(\Xi)](\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \int_{\Xi} \exp[-\frac{(u-\mu)^2}{2\sigma^2}] du \tag{6.3}$$
$$(\forall \Xi \in \mathcal{B}_{\mathbb{R}}, \forall \omega = (\mu, \sigma) \in \Omega = \mathbb{R} \times \mathbb{R}_+, i = 1, 2, ...)$$

Also,

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\sigma}^{\sigma} e^{-\frac{x^2}{2\sigma^2}} dx = 0.683..., \qquad \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-2\sigma}^{2\sigma} e^{-\frac{x^2}{2\sigma^2}} dx = 0.954...,$$

$$\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-1.96\sigma}^{1.96\sigma} e^{-\frac{x^2}{2\sigma^2}}dx \fallingdotseq 0.95 \qquad \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-2.58\sigma}^{2.58\sigma} e^{-\frac{x^2}{2\sigma^2}}dx \fallingdotseq 0.99$$

$$\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-Z(\alpha)\sigma}^{Z(\alpha)\sigma} e^{-\frac{x^2}{2\sigma^2}}dx \fallingdotseq 1-2\alpha$$

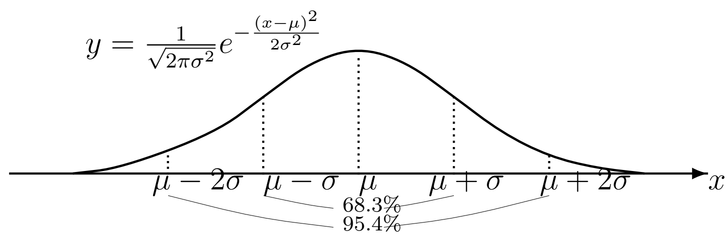


Figure 6.1:Normal distribution $N(\mu, \sigma)$

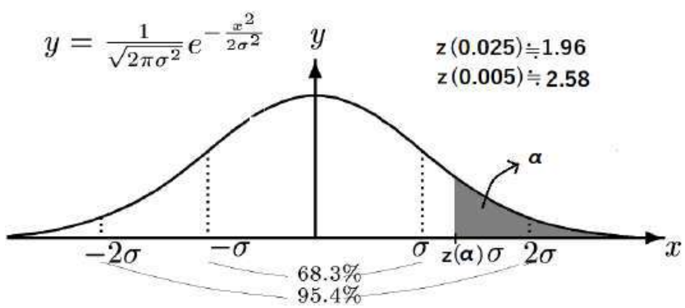


Figure 6.2:Normal distribution $N(0, \sigma)$

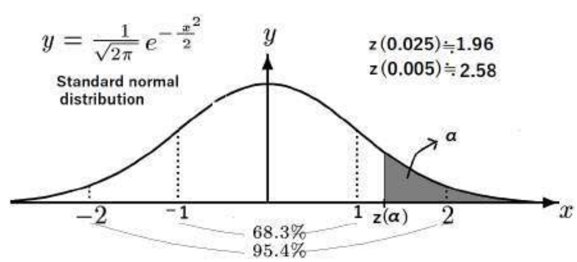


Figure 6.3:Standard normal distribution $N(0, 1)$

Therefore, from a statistical point of view, what we need to do is to answer the following problem.

Problem 6.3. In statistics, we are interested in the case that $\{X_i\}_{i=1}^{\infty}$ is a sequence of independent random variables with the normal distribution. And we focus on the following problems:

($\sharp_1$) Population mean (Confidence interval and Hypothesis testing)
• Study the statistical meaning of "$\mu \approx \overline{X}_n(s)$ (for a not very large n)" in (6.1) !
(Or, approximate μ using $\{X_1(s), X_2(s), ..., X_2(s)\}$!)

For further information see my homepage

($\sharp_2$) Population variance (Confidence interval and Hypothesis Testing)
- Study the statistical meaning of "$\sigma^2 \approx \frac{SS_n(s)}{n-1}$ (for a not very large n)" in (6.2) !
(Or, approximate σ using $\{X_1(s), X_2(s), ..., X_2(s)\}$!)

This will be done in the next subsection. To discuss ($\sharp_1$) and ($\sharp_2$) in detail, we consider that $\{X_i\}_{i=1}^{\infty}$ is a sequence of independent random variables with the normal distribution (with 'population mean' μ, 'population variance' σ^2).

6.1.3 (Student) t-distribution, χ^2-distribution

Review 6.4. [Student's t-distribution $p_n^{(t)}$ with n degrees of freedom (precisely, probability density function $p_n^{(t)}$)]
The Student's t-distribution $p_n^{(t)}$ with n degrees of freedom is defined by

$$p_n^{(t)}(x) = \frac{\Gamma((n+1)/2)}{\sqrt{n\pi}\Gamma(n/2)}\Big(1 + \frac{x^2}{n}\Big)^{-(n+1)/2} \tag{6.4}$$

(Γ is Gamma function, i.e., $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$)

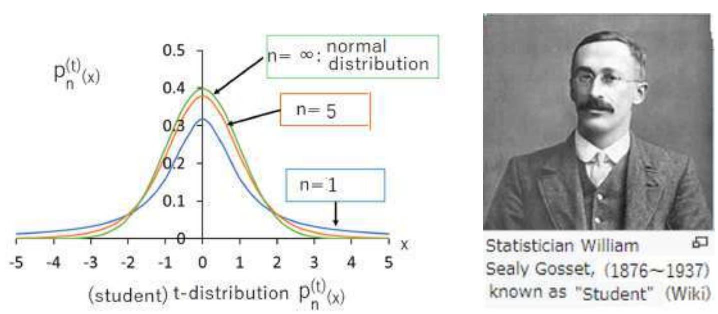


Figure 6.4: Student's t-distribution $p_n^{(t)}$ with n degrees of freedom

Also note that

$$\lim_{n\to\infty} p_n^{(t)}(x) = \lim_{n\to\infty} \frac{\Gamma((n+1)/2)}{\sqrt{n\pi}\Gamma(n/2)}\Big(1 + \frac{x^2}{n}\Big)^{-(n+1)/2}$$
$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},$$

thus, if $n \geq 30$, it can be regarded as the normal distribution $N(0,1)$ with mean 0 and the standard deviation 1.

Also, define the map $t_n : [0,1] \to [0,\infty)$, $n = 1, 2, ...,$ such that

$$\int_{t_n(\alpha)}^{\infty} p_n^{(t)}(x)dx = \alpha$$

For further information see my homepage

For example, we see,

$$t_5(0.025) = 2.571, \quad t_5(0.005) = 4.032$$
$$t_6(0.025) = 2.447, \quad t_6(0.005) = 3.707 \tag{6.5}$$

Review 6.5. The χ^2-distribution ($\approx \chi^2$-probability density function) with n degree of freedom is defined by

$$p_n^{\chi^2}(x) = \frac{x^{n/2-1}e^{-x/2}}{2^{n/2}\Gamma(n/2)} \quad (x > 0), \tag{6.6}$$

where Γ is the Gamma function.

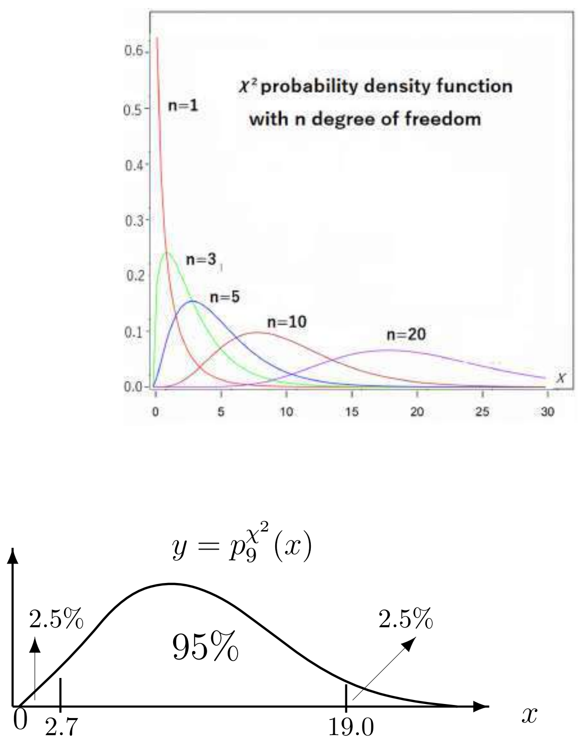


Figure 6.5:χ^2 distribution $p_n^{\chi^2}(x)$ and $y = p_9^{\chi^2}(x)$

The following Lemma is fundamental.

Lemma 6.6. Let $X_1, X_2, ..., X_n$ be independent random variables (on a probability space $(S, \mathcal{B}_S, P)$) with the normal distribution $N(\mu, \sigma^2)$. Also, recall the notations $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$, $SS_n = \sum_{i=1}^n (X_i - \overline{X}_n)^2$, $U = \sqrt{\frac{SS_n}{n-1}}$.

(i) (we want to know μ when σ is known)

Define the random variable $Z : S \to \mathbb{R}$ such that $Z = \frac{\overline{X}_n-\mu}{\sigma/\sqrt{n}}$. Then it holds that

$$Z = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

where $N(0,1)$ is the standard normal distribution.

(ii) (we want to know μ when σ is unknown)

Define the random variable $T : S \to \mathbb{R}$ such that $T = \frac{\overline{X}_n-\mu}{U/\sqrt{n}}$, where $U = \sqrt{\frac{SS_n}{n-1}}$. Then, it holds that

$$T = \frac{\overline{X}_n - \mu}{U/\sqrt{n}} \sim p^{(t)}_{n-1}$$

where $p^{(t)}_{n-1}$ is the Student's t-distribution with $n-1$ degrees of freedom.

(iii) (we want to know σ)

Define the random variables $K_i : S \to \mathbb{R}$ $(i = 1,2)$ such that $K_1 = \sum_{i=1}^n (\frac{X_i-\mu}{\sigma})^2$ and $K_2 = \sum_{i=1}^n (\frac{X_i-\overline{X}_n}{\sigma})^2$. Then, we see

$$K_1 = \sum_{i=1}^n \Big(\frac{X_i - \mu}{\sigma}\Big)^2 \sim p^{\chi^2}_n, \qquad K_2 = \sum_{i=1}^n \Big(\frac{X_i - \overline{X}}{\sigma}\Big)^2 \sim p^{\chi^2}_{n-1}$$

(when we know μ) (when we do not know μ)

where $p^{\chi^2}_n$ is the χ^2-distribution with n degrees of freedom.

Proof. See ref. [109]. ///

♠**Note 6.3.** The above is the most important theorem in statistics. It should therefore be called a 'theorem' in common sense. The reason we call it a 'Lemma' in this book is that I will use it in the proof of Theorem 6.9, which is one of the most important theorems in QL.

6.1.4 Answer to Problem 6.3 about "$\mu \approx \overline{X}_n(s)$"; Confidence interval and Hypothesis Testing

6.1.4.1 (when σ is known)

Recall our problem (i.e., Problem 6.3 $(\sharp_1)$):

$(\sharp_1)$ Confidence interval and Hypothesis Testing
• Study the statistical meaning of "$\mu \approx \overline{X}_n(s)$"!

Fix $\alpha = 0.0025$ and thus, $z(0.0025) = 1.96$ (*cf.* Figure 6.3). Then, Lemma 6.6 (i) says that

(A) the probability that a sample $(X_1(s), X_2(s), ..., X_n(s))$ satisfies that $|\frac{\overline{X}_n(s)-\mu}{\sigma/\sqrt{n}}| \le z(0.0025) = 1.96$ is given by 0.95

That is,

(B) [95%-Confidence interval]
the probability that μ belongs to the (confidence) interval $[\overline{X}_n(s) - 1.96\sigma/\sqrt{n}, \overline{X}_n(s) + 1.96\frac{\sigma}{\sqrt{n}}]$ is 0.95, that is,

$$\overline{X}_n(s) - 1.96\frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X}_n(s) + 1.96\frac{\sigma}{\sqrt{n}}$$

6.1.4.2 (when σ is unknown)

Recall Lemma 6.6 (ii). Fix $\alpha = 0.0025$, $n = 6$, thus $t_5(0.0025) = 2.571$ (*cf.* (*cf.* (formula (6.5))) and $U = \sqrt{\frac{SS_6}{5}} = \sqrt{\frac{\sum_{i=1}^{6}(X_i - \overline{X}_6)^2}{5}}$, $\overline{X}_6 = \frac{1}{6}\sum_{i=1}^{6} X_i$. Lemma 6.6 (ii) says that

(C) the probability that a sample $(X_1(s), X_2(s), ..., X_6(s))$ satisfies that $|\frac{\overline{X}_6(s)-\mu}{U/\sqrt{6}}| \leq t_6(0.0025) \approx 2.571$ is given by 0.95 (*cf.* formula (6.5)).

That is,

(D) [95%-Confidence interval]
the probability that μ belongs to the (confidence) interval $[\overline{X}_6(s) - 2.571U/\sqrt{6}, \overline{X}_6(s) + 2.571\frac{U}{\sqrt{6}}]$ is 0.95, that is,

$$\overline{X}_6(s) - 2.571\frac{U}{\sqrt{6}} \leq \mu \leq \overline{X}_6(s) + 2.571\frac{U}{\sqrt{6}}$$

Let's think about the next.

(E) [95%-Statistical hypothesis testing]
Coco (your dog's name) said that

($\flat_1$) $\overline{X}_n(s) \approx \mu_0$ (called Null hypothesis).

However, you believe the ($\sharp_2$) to be wrong. How can you convince Coco that the above ($\flat_1$) is wrong?

[Answer]: Assume the ($\flat_1$), which is called the null hypothesis. Let $\{X_1, X_2, ..., X_6\}$ be the sample (e.g., n=6). Then you can check the following.

$$|\frac{\sum_{i=1}^{6} X_i}{6} - \mu_0| \leq 2.571\frac{\sigma}{\sqrt{6}}$$

Then, Lemma 6.6 says that

(D) If it is true, there is a possibility that Coco is true. However, it is not true, as this would be a very rare occurrence, ($\flat_1$) should be considered wrong.

6.1.5 Answer to Problem 6.3 "$\sigma \approx \frac{SS_n(s)}{n-1}$"; Hypothesis Testing

Next we study the statistical understanding of "$\sigma \approx \frac{SS_n(s)}{n-1}$" in Problem 6.3. Of course, σ is unknown. Recall Lemma 6.6 (iii), which says that

($\sharp$) Let $X_1, X_2, ..., X_n$ be independent random variables with the normal distribution $N(\mu, \sigma^2)$. Then, it holds that

$$\sum_{i=1}^{n} \Big(\frac{X_i - \overline{X}}{\sigma} \Big)^2 \sim p_{n-1}^{\chi^2}$$

For example, assume the following data:

(F) $\qquad n = 10, \quad \overline{X} = 9.90, \quad U^2 = 0.250$

Then, Lemma 6.6 (iii) and Figure 6.5 say that

$$2.70 \le \frac{(n-1)U^2}{\sigma^2} \le 19.0,$$

A simple calculation says that

$$0.118 \le \sigma^2 \le 0.833$$

Thus, we can estimate the population variance σ^2 such as

(G) the probability that it holds that $0.118 \le \sigma^2 \le 0.833$ is given by 0.95

6.2 Confidence and testing problem in QL terms

This section concentrates on rewriting the 'conventional statistical methods described in the previous section' in the language of quantum language.

I belonged to a mathematics department and was somewhat familiar with probability theory (=theory of random variables). However, when I learned about quantum mechanics, I was surprised to find out that quantum mechanics understands probability without using random variables. It is hoped that readers reading this section will experience the same surprise that the author once experienced.

6.2.1 Review of Fisher's maximal likelihood method

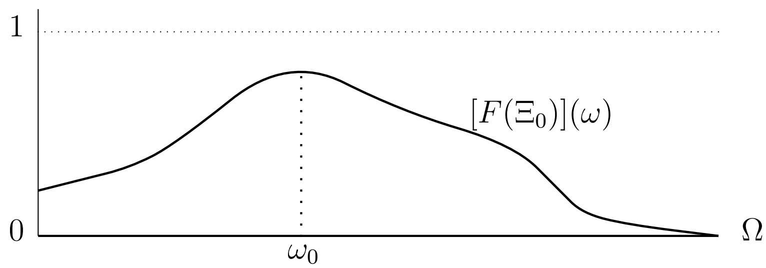
Consider the classical basic structure:

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

Consider a classical measurement $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O} = (X, \mathcal{F}, F), S_{[\omega_0]})$. It is usual to consider that the state ω_0 is unknown. And, we can usually estimate the unknown state ω_0 by a measured value as follows.

[Fisher's maximal likelihood method (*cf.* Sec. 5.2)]:
Consider a classical measurement $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O} = (X, \mathcal{F}, F), S_{[\omega_0]})$. Assume that you know a measured value belongs to $\Xi_0 (\subseteq \mathcal{F}, \max\{[F(\Xi_0)](\omega) | \omega \in \Omega\} \neq 0)$. Then, Fisher's maximal likelihood method says that the state ω_0 is predicted to satisfy the following

(A) $\qquad [F(\Xi_0)](\omega_0) = \max_{\omega \in \Omega} [F(\Xi_0)](\omega)|$



Fisher maximum likelihood method (*cf.* Figure5.5)

This is the most fundamental result in inferential statistics. However, as mentioned in the previous section, the most applicable result in inferential statistics is the theory of random variables. This section therefore attempts to rewrite the inference problem with random variables in QL terms.

6.2.2 Confidence interval and testing problems by QL

Definition 6.7. [Normal observable]. Define the state space $\Omega = \mathbb{R} \times \mathbb{R}_+$ with the Lebesgue measure ν. Consider the classical basic structure:

$$[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))] \qquad \text{(where } \Omega = \mathbb{R} \times \mathbb{R}_+)$$

The normal observable $\mathsf{O}_G = (\mathbb{R}, \mathcal{B}_\mathbb{R}, G)$ $(= (X, \mathcal{F}, G))$ in $L^\infty(\Omega (\equiv \mathbb{R} \times \mathbb{R}_+))$ is defined by

$$[G(\Xi)](\omega) = [G(\Xi)](\mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \int_\Xi \exp[-\frac{(x-\mu)^2}{2\sigma^2}] dx. \tag{6.7}$$

$$(\forall \Xi \in \mathcal{B}_\mathbb{R} (= \text{the Borel field in } \mathbb{R})), \quad \forall \omega = (\mu, \sigma) \in \Omega = \mathbb{R} \times \mathbb{R}_+)$$

For further information see my homepage

Definition 6.8. [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 $\mathsf{O}_G^n = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n, G^n)$ $(= (X^n, \mathcal{F}^n, G^n))$ in $L^\infty(\mathbb{R} \times \mathbb{R}_+)$ such that

$$
\begin{aligned}
&[G^n(\times_{k=1}^n \Xi_k)](\omega) = \times_{k=1}^n [G(\Xi_k)](\omega) \\
=&\frac{1}{(\sqrt{2\pi}\sigma)^n} \underset{\times_{k=1}^n \Xi_k}{\int \cdots \int} \exp[-\frac{\sum_{k=1}^n (x_k - \mu)^2}{2\sigma^2}] dx_1 dx_2 \cdots dx_n. \qquad (6.8)\\
&(\forall \Xi_k \in \mathcal{B}_{\mathbb{R}} (k = 1, 2, \ldots, n), \quad \forall \omega = (\mu, \sigma) \in \Omega = \mathbb{R} \times \mathbb{R}_+)
\end{aligned}
$$

///

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 $\overline{\mu} : \mathbb{R}^n \to \mathbb{R}$, $ss_n : \mathbb{R}^n \to \mathbb{R}$ and $\overline{\sigma} : \mathbb{R}^n \to \mathbb{R}$ such that

$$
\overline{\mu}(x) = \overline{\mu}(x_1, x_2, \ldots, x_n) = \frac{x_1 + x_2 + \cdots + x_n}{n} \quad (\forall x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n) \qquad (6.9)
$$

$$
ss_n(x) = ss_n(x_1, x_2, \ldots, x_n) = \sum_{k=1}^n (x_k - \overline{\mu}(x))^2 \quad (\forall x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n) \qquad (6.10)
$$

$$
\overline{\sigma}(x) = \overline{\sigma}(x_1, x_2, \ldots, x_n) = \sqrt{\frac{\sum_{k=1}^n (x_k - \overline{\mu}(x))^2}{n-1}} \quad (\forall x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n) \qquad (6.11)
$$

The following Theorem is fundamental.

Theorem 6.9. Consider the normal simultaneous measurement $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}(\times_{i=1}^n \mathsf{O}_G = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n, G^n),\ S_{[(\mu_0,\sigma_0)]})$. Also, we use the notations: $x = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$, $\overline{x}_n = \frac{1}{n}\sum_{i=1}^n x_i$, $ss_n = \sum_{i=1}^n (x_i - \overline{x}_n)^2$, $u = \sqrt{\frac{ss_n}{n-1}}$.

(i) (we want to know μ_0 when σ_0 is known)

Define the map $z : \mathbb{R}^n \to \mathbb{R}$ such that $z = \frac{\overline{x}_n - \mu_0}{\sigma_0/\sqrt{n}}$. Then it holds that

$$
\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}(z(\mathsf{O}_G) = (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, G^n([z]^{-1}(\cdot))), S_{[(\mu_0,\sigma_0)]}) \sim N(0,1)
$$

where $N(0,1)$ is the standard normal distribution.

(ii) (we want to know μ_0 when σ_0 is unknown)

Define the map $t : \mathbb{R}^n \to \mathbb{R}$ such that $t = \frac{\overline{x}_n - \mu_0}{u/\sqrt{n}}$. Then it holds that

$$
\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}(t(\mathsf{O}_G) = (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, G^n([t]^{-1}(\cdot))), S_{[(\mu_0,\sigma_0)]}) \sim p_{n-1}^{\chi^2}
$$

where $p_{n-1}^{\chi^2}$ is the χ^2-distribution with n degrees of freedom.

(iii) (we want to know σ_0)

Define the maps $k_i : \mathbb{R}^n \to \mathbb{R}$ $(i = 1, 2)$ such that $k_1 = \sum_{i=1}^n (\frac{x_i - \mu_0}{\sigma_0})^2$ and $k_2 = \sum_{i=1}^n (\frac{x_i - \overline{x}_n}{\sigma_0})^2$.

Then, we see

- (when we know μ_0)
  $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}(k_1(\mathsf{O}_G) = (\mathbb{R}, \mathcal{B}_\mathbb{R}, G^n([k_1]^{-1}(\cdot))), S_{[(\mu_0,\sigma_0)]}) \sim p_n^{\chi^2}$,
- (when we do not know μ_0)
  $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}(k_2(\mathsf{O}_G) = (\mathbb{R}, \mathcal{B}_\mathbb{R}, G^n([k_2]^{-1}(\cdot))), S_{[(\mu_0,\sigma_0)]}) \sim p_{n-1}^{\chi^2}$

where $p_n^{\chi^2}$ is the χ^2-distribution with n degrees of freedom.

Proof. This is a direct consequence of Lemma 6.6. ///

6.2.3 Measurement theoretical answer to Problem 6.3 "$\mu \approx \overline{X}_n(s)$"; Confidence interval and Hypothesis Testing

6.2.3.1 (when σ is unknown)

In this section, [Answer to Problem 6.3 "$\mu \approx \overline{X}_n(s)$" in Sec. 6.1.4] will be rewrote in terms of QL (using Theorem 6.9). Consider the normal simultaneous measurement $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}(\times_{i=1}^n \mathsf{O}_G = (\mathbb{R}^n, \mathcal{B}_\mathbb{R}^n, G^n), S_{[(\mu_0,\sigma_0)]})$. Also, we use the notations: $x = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$, $\overline{x}_n = \frac{1}{n}\sum_{i=1}^n x_i$, $ss_n = \sum_{i=1}^n (x_i - \overline{x}_n)^2$, $u = \sqrt{\frac{ss_n}{n-1}}$. Recall Theorem 6.9 (ii). Fix $\alpha = 0.0025$, $n = 6$, thus $t_5(0.0025) = 2.571$ (*cf.* Figure 6.4) and $u = \sqrt{\frac{ss_6}{5}} = \sqrt{\frac{\sum_{i=1}^6 (x_i - \overline{x}_6)^2}{5}}$, $\overline{x}_6 = \frac{1}{6}\sum_{i=1}^6 x_i$.

(B) (we want to know μ_0 when σ_0 is unknown)
Define the map $t : \mathbb{R}^n \to \mathbb{R}$ such that $t = \frac{\overline{x}_n - \mu_0}{u/\sqrt{n}}$. Then, Theorem 6.9 (ii) says that

$$\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}(t(\mathsf{O}_G) = (\mathbb{R}, \mathcal{B}_\mathbb{R}, G^n([t]^{-1}(\cdot))), S_{[(\mu_0,\sigma_0)]}) \sim p_5^{\chi^2}$$

where $p_5^{\chi^2}$ is the χ^2-distribution with 5 degrees of freedom.

This implies that

(C) the probability that a measured value $(x_1, x_2, ..., x_6)$ by $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}(\times_{i=1}^n \mathsf{O}_G = (\mathbb{R}^n, \mathcal{B}_\mathbb{R}^n, G^n), S_{[(\mu_0,\sigma_0)]})$ satisfies that $|\frac{\overline{x}_6(s) - \mu_0}{U/\sqrt{6}}| \le t_6(0.0025) \approx 2.571$ is given by 0.95

That is,

(D) [95%-Confidence interval]
the probability that μ_0 belongs to the (confidence) interval $[\overline{x}_6(s) - 2.571u/\sqrt{6}, \overline{x}_6(s) + 2.571\frac{u}{\sqrt{6}}]$ is 0.95, that is,

$$\overline{x}_6(s) - 2.571\frac{u}{\sqrt{6}} \le \mu \le \overline{x}_6(s) + 2.571\frac{u}{\sqrt{6}} \tag{6.12}$$

Let's think about the next.

(E) [95%-Statistical hypothesis testing]

Coco (your dog's name) said that

($\flat_2$) $\overline{x}_n \approx \mu_0$ (called Null hypothesis).

However, you believe the ($\flat_2$) to be wrong. How can you convince Coco that ($\flat_2$) is wrong?

[Answer]: Assume the ($\flat_2$), which is called the null hypothesis. Let $\{x_1, x_2, ..., x_6\}$ be the measured value (e.g., n=6). Then you can check the following.

$$|\frac{\sum_{i=1}^{6} x_i}{6} - \mu_0| \leq 2.571 \frac{\sigma}{\sqrt{6}}$$

Then, Theorem 6.9 says that

(F) If it is true, there is a possibility that Coco is true. However, it is not true, as this would be a very rare occurrence, ($\flat_1$) should be considered wrong.

6.3 Random valuable vs. measurement

In this chapter, I discussed the relation among following three:

($\sharp_1$) $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O} = (\mathbb{R}, \mathcal{B}_\mathbb{R}, G), S_{[(\mu,\sigma)]})$: normal measurement, $\mathsf{O} = (\mathbb{R}, \mathcal{B}_\mathbb{R}, G)$: observable, (μ, σ): state,
multidimension$\to$ $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O}^n = (\mathbb{R}^n, \mathcal{B}^n_\mathbb{R}, G^n), S_{[\mu,\sigma]})$

($\sharp_2$) $\big(\mathbb{R}, \mathcal{B}_\mathbb{R}, [G(\cdot)](\mu, \sigma)\big)$: normal sample space (= normal distribution) with a parameter (μ, σ)
multidimension$\to$ $\big(\mathbb{R}^n, \mathcal{B}^n_\mathbb{R}, [G^n(\cdot)](\mu, \sigma)\big)$

($\sharp_3$) $X_{\mu,\sigma} : S \to \mathbb{R}$: random variable such that $P(\{s \in S : \alpha_1 \leq X_{\mu,\sigma}(s) \leq \alpha_2\}) = [G([\alpha_1, \alpha_2])](\mu, \sigma)$
multidimension$\to$ Consider $X^i_{\mu,\sigma} : S \to \mathbb{R}$ $(i = 1, 2, ..., n)$ are independent

In Sec. 6.1 (the arguments in statistics), we devote ourselves to ($\sharp_2$) and ($\sharp_3$). And in Sec. 6.2 (the arguments in measurement), we devote ourselves to ($\sharp_1$) and ($\sharp_3$). The above is illustrated as follows

 For further information see my homepage

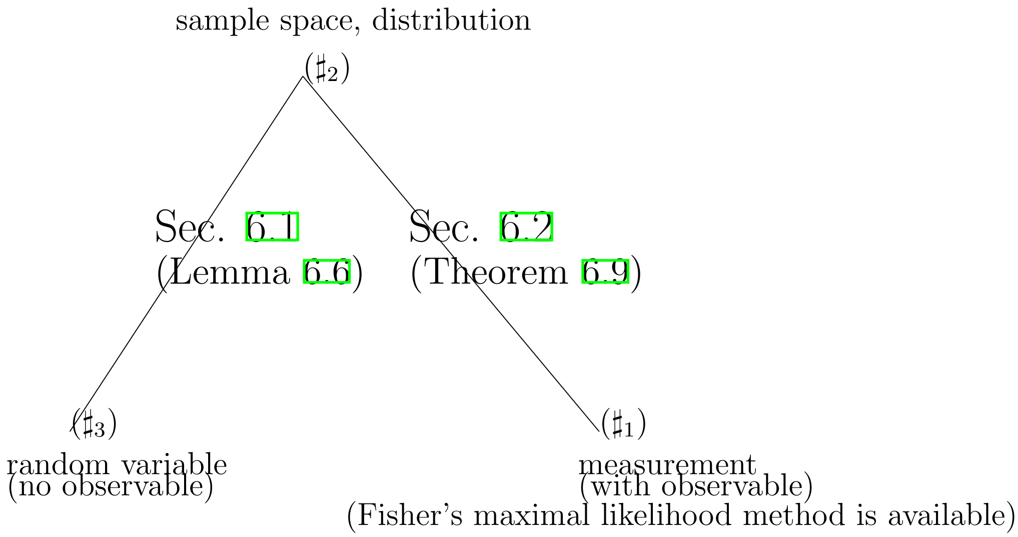


Figure 6.6 Random variable and measurement
(Compare Definition 4.7 to Definition 6.1)

The above says that

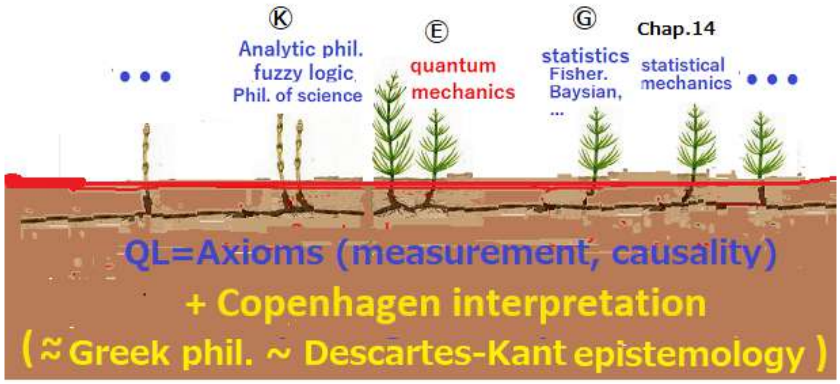
(♭) statistics is the part you write on the calculation paper when you think in quantum language

Looking above, one might think, from the theoretical point of view, that measurement theory is superior to traditional statistics. For example, the random variable method is impotent for Fisher' maximum likelihood method. However, note that the random variable method is handy in this chapter, and thus, Theorem 6.9 is proved by Lemma 6.6. As mentioned in Note 6.3, QL and random variable method are not in a rivalry relationship. The most demanding thing for researchers is to develop the ability to freely cross-translate between Axiom 1 (§2.7) and random variables method, and the ability to use these differently.

Remark 6.10. (i): Tests on two or more types of measurements can be done in the same way (using the F distribution). Namely, it suffices to start from

$$\mathsf{M}_{L^\infty(\Omega_1,\nu_1)}(\mathsf{O}_1^{n_1} = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n, G_1^{n_1}), S_{[\mu_1,\sigma_1]}) \bigotimes \mathsf{M}_{L^\infty(\Omega_2,\nu_2)}(\mathsf{O}_2^{n_2} = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n, G_2^{n_2}), S_{[\mu_2,\sigma_2]})$$

(ii): Just to be clear, I am not rejecting the 'random variables method'. I believe that the 'random variables method' is as important as ever, even with the formulation of statistics by measurement theory. As I have said many times, my argument is the following.



For this, the formulation of statistics by QL is needed.

♠**Note 6.4.** (i): See Note 2.4. That is,

- population ≈ system,
- parameter (=(population mean μ, standard deviation σ)) ≈ state

This illustrates the difficulty of using the term 'population'.
(ii): If the test is carried out several times in succession, errors are said to add up and multiplicity issues occur. In measurement theory, the linguistic Copenhagen interpretation says "Only one measurement is possible". Therefore, in measurement theory, multiplicity issue is a matter of principle.and thus, it is recommended that multiple testing is not carried out. I am a layman and don't know all the details, but I believe that computers can help us get around multiplicity issues, since the linguistic Copenhagen interpretation does not require an analytical solution.

(iii): As illustrated in Figure 6.6, the discussion of the random variable method can automatically be replaced by a discussion of measurement theory. Therefore, the discussion of analysis of variance (F-distribution) should be left as an exercise for the reader.

6.4 Regression analysis

6.4.1 Simple regression analysis

Put $\mathbb{R} = \Omega = \Omega_0 = \Omega_1 = \Omega_2 = \Omega_3 = ... = \Omega_n$) and $t_1 < t_2 < t_3 < ... < t_n$ (in $\mathbb{R}$) (for simplicity, put $n = 3$, and $t_1 = 1, t_2 = 2, t_3 = 3$).

Assume that $\mathsf{O}_{G_\sigma}(= \mathsf{O}_1 = \mathsf{O}_2 = \mathsf{O}_3)$ is the normal observable with a standard deviation σ, i.e., O_{G_σ}=$(\mathbb{R}, \mathcal{B}_{\mathbb{R}}, G_\sigma)$ where

$$[G_\sigma(\Xi)](\omega) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_\Xi e^{-\frac{(x-\omega)^2}{2\sigma^2}} dx \quad (\forall \Xi \in \mathcal{B}_{\mathbb{R}}, \forall \omega \in \Omega_k).$$

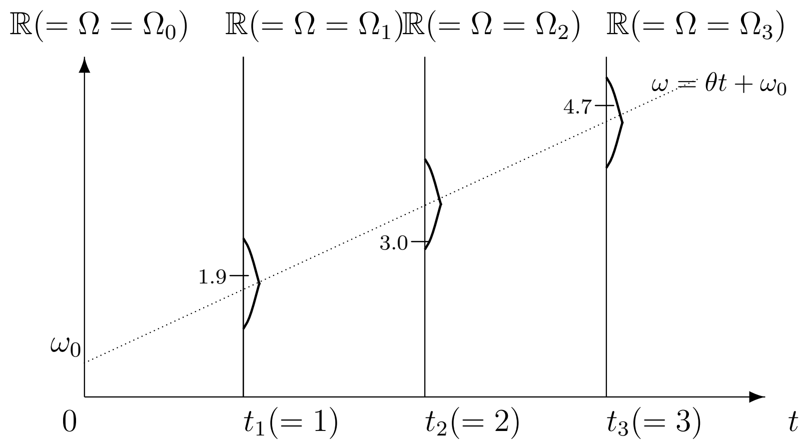


Figure 12.2 Problem: Solve unknown θ and ω_0 in the equation $\omega=\theta t+\omega_0$ of the dashed line

Define the observable $\widehat{\mathsf{O}}_T=(\mathbb{R}^3, \mathcal{F}_{\mathbb{R}^3}, \widehat{F}_0)$ in $L^\infty(\Omega_0\times\Theta)$ such that

$$
\begin{aligned}
&[\widehat{F}_0(\Xi_1\times\Xi_2\times\Xi_3)](\omega_0,\theta)\\
=&[G_\sigma(\Xi_1)](\omega_0+\theta))\cdot[G_\sigma(\Xi_2)](\omega_0+2\theta))\cdot[G_\sigma(\Xi_3)](\omega_0+3\theta))\\
&(\forall\Xi_1,\Xi_2,\Xi_3\in\mathcal{B}_{\mathbb{R}},\ \forall(\omega_0,\theta)\in\Omega_0\times\Theta)
\end{aligned}
$$

Our problem is as follows:

Problem 6.11. Consider the measurement $\mathsf{M}_{L^\infty(\Omega_0\times\Theta)}(\widehat{\mathsf{O}}_T, S_{[\omega_0]})$

($\sharp_1$) Find the parameter ($\approx$ state) $(\omega_0,\theta)\in\Theta\times\mathbb{R}$ that is most likely to yield the measured value $(1.9, 3.0, 4.7)$.

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 (ω_0,θ) ($\in\Omega_0\times\Theta$) such that

$$
[\widehat{F}_0(\Xi_1\times\Xi_2\times\Xi_3)](\omega_0,\theta)=\max_{(\omega_0,\theta)}[\widehat{F}_0(\Xi_1\times\Xi_2\times\Xi_3)].
$$

Since N is assumed to be sufficiently large, we see

$$
\begin{aligned}
(\sharp_2)\Longrightarrow&\max_{(\omega_0,\theta)\in\Omega_0}[\widehat{F}_0(\Xi_1\times\Xi_2\times\Xi_3)](\omega_0,\theta)\\
\Longrightarrow&\max_{(\omega_0,\theta)\in\Omega_0}\frac{1}{\sqrt{2\pi\sigma^2}^3}\iiint_{\Xi_1\times\Xi_2\times\Xi_3} e^{[-\frac{(x_1-(\omega_0+\theta))^2+(x_2-(\omega_0+2\theta))^2+(x_3-(\omega_0+3\theta))^2}{2\sigma^2}]}
\end{aligned}
$$

For further information see my homepage

$$
\begin{aligned}
& \times dx_1 dx_2 dx_3 \\
\Longrightarrow & \max_{(\omega_0,\theta)\in\Omega_0} \exp(-J/(2\sigma^2)) \\
\Longrightarrow & \min_{(\omega_0,\theta)\in\Omega_0} J
\end{aligned}
$$

where

$$J = (1.9 - (\omega_0 + \theta))^2 + (3.0 - (\omega_0 + 2\theta))^2 + (4.7 - (\omega_0 + 3\theta))^2.$$

($\frac{\partial}{\partial \omega_0}\{\cdots\} = 0, \frac{\partial}{\partial \theta}\{\cdots\} = 0$)

$$
\begin{aligned}
&\Longrightarrow \begin{cases} (1.9 - (\omega_0 + \theta)) + (3.0 - (\omega_0 + 2\theta)) + (4.7 - (\omega_0 + 3\theta)) = 0 \\ (1.9 - (\omega_0 + \theta)) + 2(3.0 - (\omega_0 + 2\theta)) + 3(4.7 - (\omega_0 + 3\theta)) = 0 \end{cases} \\
&\Longrightarrow \quad (\omega_0, \theta) = (0.4, 1.4)
\end{aligned}
$$

Therefore, in order to obtain a measured value (1.9, 3.0, 4.7), it suffices to put

$$(\omega_0, \theta) = (0.4,\ 1.4).$$

(Regression analysis will be discussed again in Chapter 13.) □

Remark 6.12. I am not denying the use of the terms “dependent variable” and “explanatory variable.” However, the reader should ask the following questions：

(♯) Why does statistics use terms like ”dependent variable” or ”explanatory variable”?

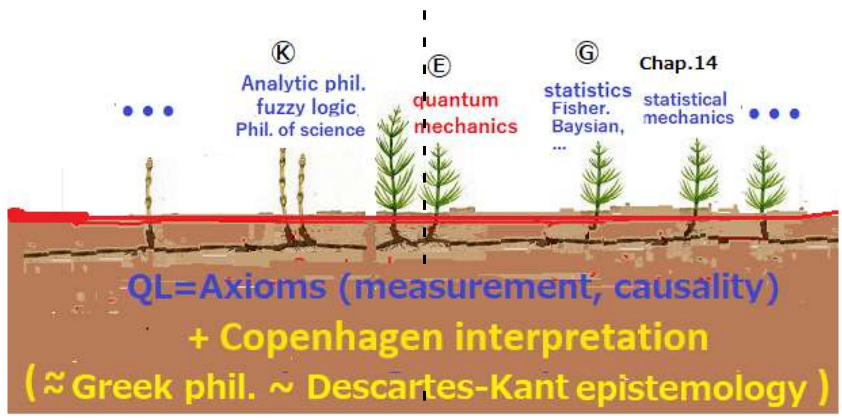
The reason, I believe, is that statistics is applied mathematics and does not have scientific axioms. On the other hand, QL starts from axioms, so OL are expected to use words within the axioms as much as possible.

For further information see my homepage

Chapter 7

Practical logic

Recall the following:



discussed in refs. [63, 76] discussed in refs. [62, 65]

The relation among Analytic philosophy, Descartes-Kant epistemology, quantum mechanics and statistic

As mentioned in Preface 0.1, I think that

$$\bullet \begin{cases} \text{mathematics} \cdots \text{logic} \\ \text{science} \begin{cases} \text{classical mechanical worldview} \cdots \text{causality} \\ \text{quantum mechanical worldview} \cdots \text{measurement + causality} \end{cases} \end{cases}$$

And our interest is focused on the quantum mechanical worldview. Thus, 'practical logic' (i.e., 'logic' in science) must be created in 'measurement + causality'. In our work, 'practical logic' is defined by the logic defined in QL (and not the mathematical logic defined by mathematical axioms). That is, mathematicians do not necessarily make excellent scientists or philosophers.
Concerning "practical logic" , I believe I have completed it in the next.

(♯) ref. [75]. Ishikawa, S., (2021) Fuzzy Logic in the Quantum Mechanical Worldview ; Related to Zadeh, Wittgenstein, Moore, Saussure, Quine, Lewis Carroll, etc. Journal of applied mathematics and physics, Vol. 9, No.7, 1583-1610, DOI:10.4236/jamp.2021.97108 (https://www.scirp.org/journal/paperinformation.aspx?paperid=110830)
Or, see ref. [76] Chap. 11.

In this chapter, I show my old result (in refs. [29, 30]) concerning 'fuzzy logic', which is not satisfactory. This work is memorable for me because the 1990s was the time when I changed my research focus from quantum mechanics to fuzzy logic.

By the time I had finished writing these papers [29, 30, 33], I was convinced that the 'quantum mechanical worldview' had surpassed the 'mechanistic worldview'.

The logical aspects of quantum languages are mainly discussed in ref. [76]. Thus, readers may skip Sections 7.2∼7.6, which are not related to the others in this book.

7.1 My recent opinion

When we study Newtonian mechanics, I consider that we speak 'Newtonian mechanical language'. Similarly, when we study QL (=quantum language), I consider that we speak QL. In this sense, I can say that this book is written in quantum language (or, this book is a textbook for learning quantum language).

Of course, QL has a lot of sentences. For example, consider the following two sentences:

(a) when the position of a particle m is measured, the measured value is x.

(b) when the momentum of a particle m is measured, the measured values is p.

But we says, by Theorem 4.16 [The mathematical formulation of Heisenberg's uncertainty principle], that

(c) '(a) $\wedge$ (b)' is not a sentence.

However, we say the following Proposition:

Proposition 7.1. Let Ω be a compact space. And fix $\omega_0 (\in \Omega)$. Put

$$\mathsf{CFL} \equiv \{ \mathsf{M}_{C(\Omega)}((\{T,F\}, 2^{\{T,F\}}, G), S_{[\delta_{\omega_0}]}) \quad | \quad G(\{T\}) \in C(\Omega), \\ 0 \le [G(\{T\})](\omega) \le 1 \;\; (\forall \omega \in \Omega) \}$$

Then, CFL is the class in which fuzzy logic holds, if we define

- $[\neg G(\{T\})](\omega) = [1 - G(\{T\})](\omega)$
- $[(G_1 \wedge G_2)(\{T\})](\omega) = \min \Big\{ [G_1(\{T\})](\omega), [G_2(\{T\})](\omega)] \Big\}$,
- $[(G_1 \vee G_2)(\{T\})](\omega) = \max \Big\{ [G_1(\{T\})](\omega), [G_2(\{T\})](\omega)] \Big\}$.

///

For more precise arguments, see [76].

I consider the directions described above to be promising. For example, I think the study of analytic and comprehensive propositions within QL was insufficient in ref. [76] (Sec.12.1). Also see ref. [100] and Note 5.4.

♠**Note 7.1.** I am not a philosopher, but I will state some philosophy. I consider the followings.

($\sharp_1$) analytic philosophy claims "Be logical!" (*cf.* ref. [113])

($\sharp_2$) QL's claim is "Be scientific!"

since 'logic' can be derived from QL as shown in the above. In fact, most scientists know statistics but not mathematical logic. Thus, I think that ($\sharp_1$) is wrong. However, QL and analytic philosophy are not entirely different. As we saw in Figure 0.1 in Preface (or, ref. [76]), QL is an evolution of analytic philosophy, since the relationship is that QL solves the problems posed in analytic philosophy. If QL is seen as a philosophy, its slogan is "From 'Be logical!' to 'Be scientific!'". Through the problems of the flagpole and Hempel's ravens, Hempel cast doubt on 'Be logical!' (*cf.* ref. [76, 21, 22]).

7.2 Marginal observable and quasi-product observable

Definition 7.2. [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

$$F_{12\ldots n}(X_1 \times \cdots \times X_{k-1} \times \Xi_k \times X_{k+1} \times \cdots \times X_n) = F_k(\Xi_k). \qquad (7.1)$$
$$(\forall \Xi_k \in \mathcal{F}_k, \forall k = 1, 2, \ldots, n)$$

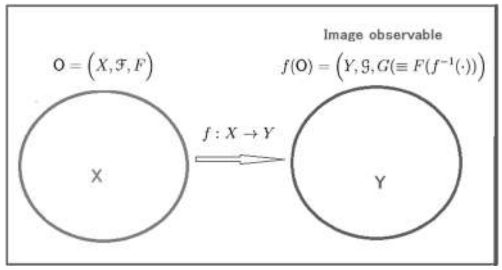
The observable $\mathsf{O}_{12\ldots n} = (\times_{k=1}^n X_k, \boxtimes_{k=1}^n \mathcal{F}_k, F_{12\ldots n})$ is called a *quasi-product observable* of $\{\mathsf{O}_k \mid k = 1, 2, \ldots, n\}$, and denoted by

$$\mathop{\boldsymbol{\times}}_{k=1,2,\ldots,n}^{\text{qp}} \mathsf{O}_k = (\mathop{\times}_{k=1}^n X_k, \boxtimes_{k=1}^n \mathcal{F}_k, \mathop{\boldsymbol{\times}}_{k=1,2,\ldots,n}^{\text{qp}} F_k).$$

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.

Definition 7.3. [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 $\mathsf{O}_{12...n} = (\times_{k=1}^n X_k, \ \boxtimes_{k=1}^n \mathcal{F}_k, \ F_{12...n})$ in $\overline{\mathcal{A}}$. For any natural number j such that $1 \leqq j \leqq n$, define $F_{12...n}^{(j)}$ such that

$$F_{12...n}^{(j)}(\Xi_j) = F_{12...n}(X_1 \times \cdots \times X_{j-1} \times \Xi_j \times X_{j+1} \times \cdots \times X_n) \quad (\forall \Xi_j \in \mathcal{F}_j).$$

Then we have the observable $\mathsf{O}_{12...n}^{(j)} = (X_j, \mathcal{F}_j, F_{12...n}^{(j)})$ in $\overline{\mathcal{A}}$. The $\mathsf{O}_{12...n}^{(j)}$ is called a marginal observable of $\mathsf{O}_{12...n}$ (or, precisely, (j)-marginal observable). Consider a map $P_j : \times_{k=1}^n X_k \to X_j$ such that

$$\mathop{\times}_{k=1}^{n} \ni (x_1, x_2, ..., x_j, ..., x_n) \mapsto x_j \in X_j.$$

Then, the marginal observable $\mathsf{O}_{12...n}^{(j)}$ is characterized as the image observable $P_j(\mathsf{O}_{12...n})$.
The above can be easily generalized as follows. For example, define $\mathsf{O}_{12...n}^{(12)} = (X_1 \times X_2, \ \mathcal{F}_1 \boxtimes \mathcal{F}_2, \ F_{12...n}^{(12)})$ such that

$$F_{12...n}^{(12)}(\Xi_1 \times \Xi_2) = F_{12...n}^{(12)}(\Xi_1 \times \Xi_2 \times X_3 \times \cdots \times X_n) \quad (\forall \Xi_1 \in \mathcal{F}_1, \forall \Xi_2 \in \mathcal{F}_2).$$

Then, we have the (12)-marginal observable $\mathsf{O}_{12...n}^{(12)} = (X_1 \times X_2, \ \mathcal{F}_1 \boxtimes \mathcal{F}_2, \ F_{12...n}^{(12)})$. Of course, we also see that $F_{12...n} = F_{12...n}^{(12...n)}$.

The following theorem is often used:

Theorem 7.4. Consider the basic structure

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

Let $\mathsf{O}_1 \equiv (X_1, \mathcal{F}_1, F_1)$ and $\mathsf{O}_2 \equiv (X_2, \mathcal{F}_2, F_2)$ 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 O_2 is projective, i.e., $F_2 = (F_2)^2 \Big)$. Then, the following statements (i) and (ii) are equivalent:

(i) There exists a quasi-product observable $\mathsf{O}_{12} \equiv (X_1 \times X_2, \mathcal{F}_1 \times \mathcal{F}_2, F_1 \overset{\text{qp}}{\times} F_2)$ with marginal observables O_1 and O_2.

(ii) O_1 and 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 O_{12} of O_1 and O_2 is guaranteed.

Proof. See refs. [12, 29, 35].

7.3 Properties of quasi-product observables

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

$$\mathrm{Rep}_\rho^{\Xi_1\times\Xi_2}[\mathsf{O}_{12}] = \begin{bmatrix} {}_{\mathcal{A}^*}\big(\rho, F_{12}(\Xi_1\times\Xi_2)\big)_{\overline{\mathcal{A}}} & {}_{\mathcal{A}^*}\big(\rho, F_{12}(\Xi_1\times\Xi_2^c)\big)_{\overline{\mathcal{A}}} \\ {}_{\mathcal{A}^*}\big(\rho, F_{12}(\Xi_1^c\times\Xi_2)\big)_{\overline{\mathcal{A}}} & {}_{\mathcal{A}^*}\big(\rho, F_{12}(\Xi_1^c\times\Xi_2^c)\big)_{\overline{\mathcal{A}}} \end{bmatrix}$$
$$(\forall \Xi_1 \in \mathcal{F}_1, \forall \Xi_2 \in \mathcal{F}_2)$$

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 7.5. [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

$$\begin{aligned}
&\mathrm{Rep}_\rho^{\Xi_1\times\Xi_2}[\mathsf{O}_{12}] = \begin{bmatrix} {}_{\mathcal{A}^*}\big(\rho, F_{12}(\Xi_1\times\Xi_2)\big)_{\overline{\mathcal{A}}} & {}_{\mathcal{A}^*}\big(\rho, F_{12}(\Xi_1\times\Xi_2^c)\big)_{\overline{\mathcal{A}}} \\ {}_{\mathcal{A}^*}\big(\rho, F_{12}(\Xi_1^c\times\Xi_2)\big)_{\overline{\mathcal{A}}} & {}_{\mathcal{A}^*}\big(\rho, F_{12}(\Xi_1^c\times\Xi_2^c)\big)_{\overline{\mathcal{A}}} \end{bmatrix} \\
&= \begin{bmatrix} \alpha_\rho^{\Xi_1\times\Xi_2} & \rho(F_1(\Xi_1)) - \alpha_\rho^{\Xi_1\times\Xi_2} \\ \rho(F_2(\Xi_2)) - \alpha_\rho^{\Xi_1\times\Xi_2} & 1 + \alpha_\rho^{\Xi_1\times\Xi_2} - \rho(F_1(\Xi_1)) - \rho(F_2(\Xi_2)) \end{bmatrix}
\end{aligned} \tag{7.2}$$

and

$$\max\{0, \rho(F_1(\Xi_1)) + \rho(F_2(\Xi_2)) - 1\} \leqq \alpha_\rho^{\Xi_1\times\Xi_2} \leqq \min\{\rho(F_1(\Xi_1)),\ \rho(F_2(\Xi_2))\}$$

For further information see my homepage

$$(\forall \Xi_1 \in \mathcal{F}_1, \forall \Xi_2 \in \mathcal{F}_2, \forall \rho \in \mathfrak{S}^p(\mathcal{A}^*)). \tag{7.3}$$

Conversely, for any $\alpha_\rho^{\Xi_1 \times \Xi_2}$ satisfying (7.3), the observable O_{12} defined by (7.2) is a quasi-product observable of O_1 and O_2. Also, it holds that

$$\begin{aligned} \rho(F(\Xi_1 \times \Xi_2^c)) = 0 \iff & \alpha_\rho^{\Xi_1 \times \Xi_2} = \rho(F_1(\Xi_1)) \\ \Longrightarrow & \rho(F_1(\Xi_1)) \leqq \rho(F_2(\Xi_2)). \end{aligned} \tag{7.4}$$

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 (7.2) that

$$\begin{aligned} & 0 \leqq \alpha_\rho^{\Xi_1 \times \Xi_2} \leqq 1 \\ & 0 \leqq 1 + \alpha_\rho^{\Xi_1 \times \Xi_2} - \rho(F_1(\Xi_1)) - \rho(F_2(\Xi_2)) \leqq 1 \\ & 0 \leqq \rho(F_2(\Xi_2)) - \alpha_\rho^{\Xi_1 \times \Xi_2} \leqq 1 \\ & 0 \leqq \rho(F_1(\Xi_1)) - \alpha_\rho^{\Xi_1 \times \Xi_2} \leqq 1 \end{aligned}$$

which clearly implies (7.3). Conversely if α satisfies (7.3), then we easily see (7.2). Also, (7.4) is obvious. This completes the proof. □

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 (7.3), estimated as follows.

$$\begin{aligned} & \frac{\max\{0, \rho(F_1(\Xi_1)) + \rho(F_2(\Xi_2)) - 1\}}{\rho(F_{12}(\Xi_1 \times \Xi_2)) + \rho(F_{12}(\Xi_1 \times \Xi_2^c))} \leqq P \leqq \\ & \qquad \frac{\min\{\rho(F_1(\Xi_1)),\ \rho(F_2(\Xi_2))\}}{\rho(F_{12}(\Xi_1 \times \Xi_2)) + \rho(F_{12}(\Xi_1 \times \Xi_2^c))}. \end{aligned}$$

Example 7.6. [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

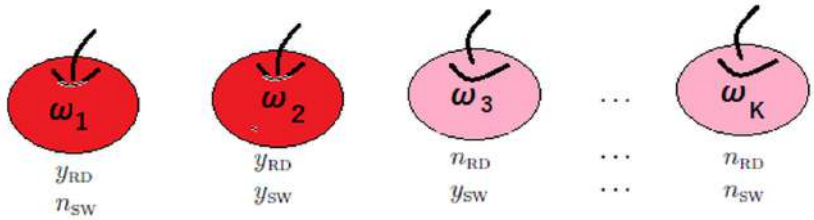
$$[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))].$$

Consider yes-no observables $\mathsf{O}_{\mathrm{RD}} \equiv (X_{\mathrm{RD}}, 2^{X_{\mathrm{RD}}}, F_{\mathrm{RD}})$ and $\mathsf{O}_{\mathrm{SW}} \equiv (X_{\mathrm{SW}}, 2^{X_{\mathrm{SW}}}, F_{\mathrm{SW}})$ in $C(\Omega)$ such that

$$X_{\mathrm{RD}} = \{y_{\mathrm{RD}}, n_{\mathrm{RD}}\} \text{ and } X_{\mathrm{SW}} = \{y_{\mathrm{SW}}, n_{\mathrm{SW}}\},$$

where we consider that "y_{RD}" and "n_{RD}" respectively mean "RED" and "NOT RED". Similarly, "y_{SW}" and "n_{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.

For further information see my homepage



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,

$$\mathrm{Rep}_{\omega_{\rm k}}^{\{(y_{\rm RD}, y_{\rm SW})\}}[\mathsf{O}_{12}] = \begin{bmatrix} [F(\{(y_{\rm RD}, y_{\rm SW})\})](\omega_k) & [F(\{(y_{\rm RD}, n_{\rm SW})\})](\omega_k) \\ [F(\{(n_{\rm RD}, y_{\rm SW})\})](\omega_k) & [F(\{(n_{\rm RD}, n_{\rm SW})\})](\omega_k) \end{bmatrix}$$

$$= \begin{bmatrix} \alpha_{\{(y_{\rm RD}, y_{\rm SW})\}} & [F_{\rm RD}(\{y_{\rm RD}\})] - \alpha_{\{(y_{\rm RD}, y_{\rm SW})\}} \\ [F_{\rm SW}(\{y_{\rm SW}\})] - \alpha_{\{(y_{\rm RD}, y_{\rm SW})\}} & 1 + \alpha_{\{(y_{\rm RD}, y_{\rm SW})\}} - [F_{\rm RD}(\{y_{\rm RD}\})] - [F_{\rm SW}(\{y_{\rm SW}\})] \end{bmatrix},$$

where $\alpha_{\{(y_{\rm RD}, y_{\rm SW})\}}(\omega_k)$ satisfies the (7.3). When we know that a tomato ω_k is red, the probability P that the tomato ω_k is sweet is given by

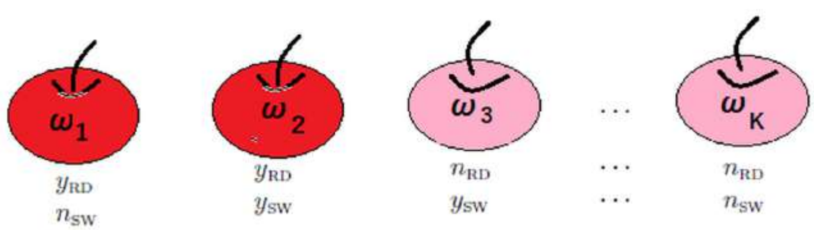
$$P = \frac{[F(\{(y_{\rm RD}, y_{\rm SW})\})](\omega_k)}{[F(\{(y_{\rm RD}, y_{\rm SW})\})](\omega_k) + [F(\{(y_{\rm RD}, n_{\rm SW})\})](\omega_k)} = \frac{[F(\{(y_{\rm RD}, y_{\rm SW})\})](\omega_k)}{[F_{\rm RD}(\{y_{\rm RD}\})](\omega_k)}.$$

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

$$\frac{\max\{0, [F_1(\{y_{\rm RD}\})](\omega_k) + [F_2(\{y_{\rm SW}\})](\omega_k) - 1\}}{[F_{\rm RD}(\{y_{\rm RD}\})](\omega_k)} \leqq P$$
$$\leqq \frac{\min[F_1(\{y_{\rm SW}\})](\omega_k),\ [F_2(\{y_{\rm SW}\})](\omega_k)\}}{[F_{\rm RD}(\{y_{\rm RD}\})](\omega_k)}.$$

7.4 Implication – the definition of "⇒"

7.4.1 Implication and contraposition



Tomatoes (Red or Sweet?)

In Example 7.6, 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 7.7. [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-product 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

$$[\mathsf{O}_{12}^{(1)}; \Xi_1] \underset{\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}_{12}, S_{[\rho]})}{\Longrightarrow} [\mathsf{O}_{12}^{(2)}; \Xi_2] \tag{7.5}$$

Of course, this (7.5) 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 $\mathsf{O}_{12\ldots n} = (\times_{k=1}^n X_k,\ \boxtimes_{k=1}^n \mathcal{F}_k,\ F_{12\ldots n} = \overset{\rm qp}{\underset{k=1,2,\ldots,n}{\times}} F_k)$ 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

$$_{\mathcal{A}^*}\Big(\rho, F_{12\ldots n}^{(ij)}(\Xi_i \times (\Xi_j^c))\Big)_{\overline{\mathcal{A}}} = 0$$

(where $\Xi^c = X \setminus \Xi$) is denoted by

$$[\mathsf{O}_{12\ldots n}^{(i)}; \Xi_i] \underset{\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}_{12\ldots n}, S_{[\rho]})}{\Longrightarrow} [\mathsf{O}_{12\ldots n}^{(j)}; \Xi_j] \tag{7.6}$$

Theorem 7.8. [Contraposition] Let $\mathsf{O}_{12} = (X_1 \times X_2,\ \mathcal{F}_1 \times \mathcal{F}_2,\ F_{12}{=}F_1 \overset{\rm qp}{\times} F_2)$ 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

$$[\mathsf{O}_{12}^{(1)}; \Xi_1] \underset{\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}_{12}, S_{[\rho]})}{\Longrightarrow} [\mathsf{O}_{12}^{(2)}; \Xi_2], \tag{7.7}$$

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 (7.7). 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

$$[\mathsf{O}_{12}^{(1)};\Xi_1^c] \underset{\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}_{12},S_{[\rho]})}{\Longleftarrow} [\mathsf{O}_{12}^{(2)};\Xi_2^c]$$

□

7.5 Combined observable – Only one measurement is permitted

7.5.1 Combined observable – only one observable

The linguistic interpretation says

"Only one measurement is permitted"
⇒ "only one observable" ⇒ "the necessity of the combined observable"

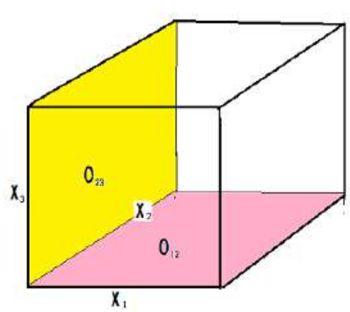
Thus, we prepare the following theorem.

Theorem 7.9. [The existence theorem of classical combined observables](*cf.*refs.[29, 35]) 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, and that $\mathcal{F}_i = 2^{X_i}$. Further assume that

$$\mathsf{O}_{12}^{(2)} = \mathsf{O}_{23}^{(2)} \quad (\text{That is,} \quad F_{12}(X_1 \times \Xi_2) = F_{23}(\Xi_2 \times X_3) \quad (\forall \Xi_2 \in 2^{X_2})).$$



For further information see my homepage

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,

$$F_{123}^{(12)}(\Xi_1 \times \Xi_2 \times X_3) = F_{12}(\Xi_1 \times \Xi_2), \;\; F_{123}^{(23)}(X_1 \times \Xi_2 \times \Xi_3) = F_{23}(\Xi_2 \times \Xi_3). \tag{7.8}$$
$$(\forall \Xi_1 \in \mathcal{F}_1, \forall \Xi_2 \in \mathcal{F}_2, \forall \Xi_3 \in \mathcal{F}_3))$$

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

$$[F_{123}(\{(x_1, x_2, x_3)\})](\omega)$$
$$= \begin{cases} \dfrac{[F_{12}(\{(x_1, x_2)\})](\omega) \cdot [F_{23}(\{(x_2, x_3)\})](\omega)}{[F_{12}(X_1 \times \{x_2\})](\omega)} & \\ & ([F_{12}(X_1 \times \{x_2\})](\omega) \;\neq 0 \text{ and }) \\ 0 & \\ & ([F_{12}(X_1 \times \{x_2\})](\omega) = 0 \text{ and }) \end{cases}$$
$$(\forall \omega \in \Omega, \forall (x_1, x_2, x_3) \in X_1 \times X_2 \times X_3)$$

This clearly satisfies (7.8). □

Counter example 7.10. [Counter example in quantum systems] Theorem 7.9 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

$$T_1 T_2 = T_2 T_1, \quad T_2 T_3 = T_3 T_2, \quad T_1 T_3 \neq T_3 T_1. \tag{7.9}$$

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

$$T_k = \int_{X_k} x_k F_k(dx_k), \tag{7.10}$$

where $X_k = \mathbb{R}, \mathcal{F}_k = \mathcal{B}_{\mathbb{R}}$. From the commutativity, we have the simultaneous observables

$$\mathsf{O}_{12}\text{=}\mathsf{O}_1 \times \mathsf{O}_2 = (X_1 \times X_2, \mathcal{F}_1 \boxtimes \mathcal{F}_2, F_{12} = F_1 \times F_2)$$

and

$$\mathsf{O}_{23}\text{=}\mathsf{O}_2 \times \mathsf{O}_3 = (X_2 \times X_3, \mathcal{F}_2 \boxtimes \mathcal{F}_3, F_{23} = F_2 \times F_3).$$

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 7.4 says that O_1 and O_3 commute, and it is in contradiction with the (7.9). Therefore, the *combined observable* O_{123} of O_{12} and O_{23} does not exist.

7.6 Syllogism and its variants

7.6.1 Syllogism and its variations: Classical systems

Next, we shall discuss practical syllogism (i.e., measurement theoretical theorem concerning implication (Definition7.7)). Before the discussion, we note

(♯) Since Theorem 7.9 (The existence of the combined observable) does not hold in quantum system, (*cf.* Counter Example7.10), 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 7.11. [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 $\mathsf{O}_{123} = (X_1 \times X_2 \times X_3,\ \mathcal{F}_1 \times \mathcal{F}_2 \times \mathcal{F}_3,\ F_{123}=\overset{\text{qp}}{\times}_{k=1,2,3} F_k)$ 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:

$$[\mathsf{O}_{123}^{(1)}; \Xi_1] \underset{\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{123}, S_{[\omega]})}{\Longrightarrow} [\mathsf{O}_{123}^{(3)}; \Xi_3]. \tag{7.11}$$

(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

$$\max\{[F_{123}^{(2)}(\Xi_2)](\omega), [F_{123}^{(1)}(\Xi_1)](\omega) + [F_{123}^{(3)}(\Xi_3)](\omega) - 1\}$$
$$\leqq \alpha_{\Xi_1\times\Xi_3}(\omega) \leqq \min\{[F_{123}^{(1)}(\Xi_1)](\omega), [F_{123}^{(3)}(\Xi_3)](\omega)\}. \tag{7.12}$$

(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,

$$[F_{123}^{(13)}(\Xi_1\times\Xi_3^c)](\omega) = [F_{123}(\Xi_1\times\Xi_2\times\Xi_3^c)](\omega) + [F_{123}^{(13)}(\Xi_1\times\Xi_2^c\times\Xi_3^c)](\omega) = 0.$$

For further information see my homepage

Thus, we get, (7.11).
For the proof of (ii) and (iii), see refs. [29, 35]. □

Example 7.12. [Continued from Example 7.6] 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 7.6. 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 observables:

$$
\begin{aligned}
\mathsf{O}_{12} &= (X_{\rm SW} \times X_{\rm RP}, 2^{X_{\rm SW} \times X_{\rm RP}}, F_{12} = F_{\rm SW} \overset{\rm qp}{\times} F_{\rm RP}) \\
\mathsf{O}_{23} &= (X_{\rm RP} \times X_{\rm RD}, 2^{X_{\rm RP} \times X_{\rm RD}}, F_{23} = F_{\rm RP} \overset{\rm qp}{\times} F_{\rm RD}).
\end{aligned}
$$

Let $\omega_k \in \Omega$. And assume that

$$
\begin{aligned}
&[\mathsf{O}_{123}^{(1)}; \{y_{\rm SW}\}] \underset{\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{123}, S_{[\omega_k]})}{\Longrightarrow} [\mathsf{O}_{123}^{(2)}; \{y_{\rm RP}\}], \\
&[\mathsf{O}_{123}^{(2)}; \{y_{\rm RP}\}] \underset{\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{123}, S_{[\omega_k]})}{\Longrightarrow} [\mathsf{O}_{123}^{(3)}; \{y_{\rm RD}\}].
\end{aligned} \tag{7.13}
$$

Then, by Theorem 7.11(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

$$
\frac{[F_{13}(\{y_{\rm SW}\} \times \{y_{\rm RD}\})](\omega_k)}{[F_{13}(\{y_{\rm SW}\} \times \{y_{\rm RD}\})](\omega_k) + [F_{13}(\{y_{\rm SW}\} \times \{n_{\rm RD}\})](\omega_k)} = \frac{[F_{\rm RD}(\{y_{\rm RD}\})](\omega_k)}{[F_{\rm RD}(\{y_{\rm RD}\})](\omega_k)} = 1. \tag{7.14}
$$

Of course, (7.13) means

$$
\text{"Sweet"} \Longrightarrow \text{"Ripe"} \qquad\qquad \text{"Ripe"} \Longrightarrow \text{"Red"}
$$

Therefore, by (7.11), we get the following conclusion.

$$
\text{"Sweet"} \Longrightarrow \text{"Red"}
$$

However, this result is not useful in the market. We want a statement like

$$
\text{"Red"} \Longrightarrow \text{"Sweet"}
$$

This will be discussed in the following example.

Example 7.13. [Continued from Example 7.6] Instead of (7.13), assume that

$$\mathsf{O}_1^{\{y_1\}} \underset{\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{12}, S_{[\delta_{\omega_n}]})}{\Longleftarrow} \mathsf{O}_2^{\{y_2\}}, \qquad \mathsf{O}_2^{\{y_2\}} \underset{\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{23}, S_{[\delta_{\omega_n}]})}{\Longrightarrow} \mathsf{O}_3^{\{y_3\}}. \tag{7.15}$$

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

$$Q = \frac{[F_{13}(\{y_{\rm SW}\} \times \{y_{\rm RD}\})](\omega_n)}{[F_{13}(\{y_{\rm SW}\} \times \{y_{\rm RD}\})](\omega_n) + [F_{13}(\{n_{\rm SW}\} \times \{y_{\rm RD}\})](\omega_n)}$$

which is, by (7.3), estimated as follows:

$$\max\left\{ \frac{[F_{\rm RP}(\{y_{\rm RP}\})](\omega_n)}{[F_{\rm RD}(\{y_{\rm RD}\})](\omega_n)}, \frac{[F_{\rm SW}(\{y_{\rm SW}\})] + [F_{\rm RD}(\{y_{\rm RD}\})] - 1}{[F_{\rm RD}(\{y_{\rm RD}\})](\omega_n)} \right\} \le Q$$
$$\le \min\{ \frac{[F_{\rm SW}(\{y_{\rm SW}\})](\omega_n)}{[F_{\rm RD}(\{y_{\rm RD}\})](\omega_n)}, 1\}. \tag{7.16}$$

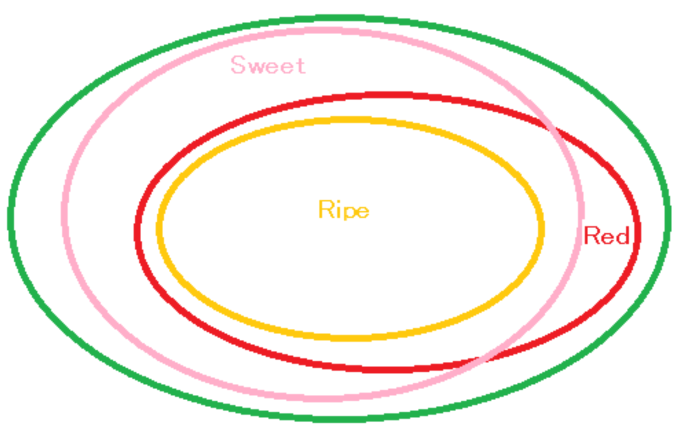
Note that (7.15) implies (and is implied by)

$$\text{"Ripe"} \Longrightarrow \text{"Sweet"} \qquad \text{and} \qquad \text{"Ripe"} \Longrightarrow \text{"Red"}$$

And note that the conclusion (7.16) is somewhat like

$$\text{"Red"} \Longrightarrow \text{"Sweet"}$$

Therefore, the estimation (7.16) may be useful in markets. ///



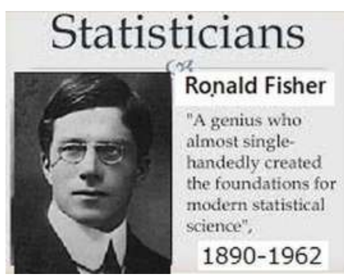
Chapter 8

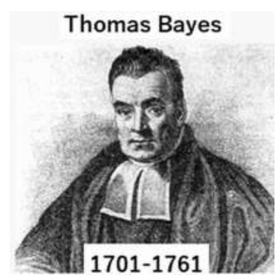
Bayesian statistics (⊂ Mixed measurement theory)

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

$$
(\sharp) \;\; \underset{\text{(=quantum language)}}{\text{measurement theory}} \begin{cases} \underset{(\sharp_1)}{\text{pure type}} \begin{cases} \text{classical system : Fisher statistics} \\ \text{quantum system : usual quantum mechanics} \end{cases} \\ \\ \underset{(\sharp_2)}{\text{mixed type}} \begin{cases} \text{classical system : including Bayesian statistics} \\ \qquad\qquad\text{and Kalman filter} \\ \text{quantum system : quantum decoherence} \end{cases} \end{cases}
$$

In this chapter, we study mixed measurement theory, which includes Bayesian statistics.





8.1 Mixed measurement theory

8.1.1 Axiom $^{(m)}$ 1 (mixed measurement)

In the previous chapters, we studied Axiom 1 (pure measurement: §2.7), that is,

$$
\underset{\text{(=quantum language)}}{\boxed{\text{pure measurement theory}}} := \underbrace{\overset{[\text{(pure)Axiom 1}]}{\underset{(cf.\ \S 2.7)}{\boxed{\text{pure measurement}}}} + \overset{[\text{Axiom 2}]}{\underset{(cf.\ \S 9.3)}{\boxed{\text{Causality}}}}}_{\text{a kind of spells (a priori judgment)}} + \underbrace{\underset{(cf.\ \S 3.1)}{\boxed{\text{Linguistic Copenhagen interpretation}}}}_{\text{manual to use spells}} \tag{8.1}
$$

In this chapter, we shall study "Axiom$^{(m)}$ 1 (mixed measurement)" in mixed measurement theory, that is,

$$\underset{\text{(=quantum language)}}{\boxed{\text{mixed measurement theory}}} := \underbrace{\overset{[(\text{mixed})\text{Axiom}^{(m)}\ 1]}{\underset{(cf.\ \S 8.1)}{\boxed{\text{mixed measurement}}}} + \overset{[\text{Axiom 2}]}{\underset{(cf.\ \S 9.3)}{\boxed{\text{Causality}}}}}_{\text{a kind of spells (a priori judgment)}} + \underbrace{\underset{(cf.\ \S 3.1)}{\boxed{\text{Linguistic Copenhagen interpretation}}}}_{\text{manual to use spells}} \tag{8.2}$$

In the previous chapters, we mainly discussed pure measurements listed in Review 9.1, especially W^*-measurement (A$_1$).

Review 8.1. [=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

$$\text{"measurement}^{W^*}\ \mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}.\ S_{[\rho]})\text{" , or \quad "measurement } \mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}.\ S_{[\rho]})\text{" ,}$$

(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

$$\text{"measurement}^{C^*}\ \mathsf{M}_{\mathcal{A}}(\mathsf{O}.\ S_{[\rho]})\text{" , or \quad "measurement } \mathsf{M}_{\mathcal{A}}(\mathsf{O}.\ S_{[\rho]})\text{" .}$$

In this chapter, we introduce four "mixed measurements" as follows.

Preparation 8.2.

(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 $\mathsf{M}_{\mathcal{A}}(\mathsf{O}{=}\ (X, \mathcal{F}, F),\ S_{[*]}(\rho_0))$, 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 $\mathsf{M}_{\mathcal{A}}(\mathsf{O},\ S_{[*]}(\rho_0))$" is also denoted by

"C^*-mixed measurement$^{C^*}$ $\mathsf{M}_{\mathcal{A}}(\mathsf{O}.\ \overline{S}_{[*]}(\rho_0))$", or "mixed measurement $\mathsf{M}_{\mathcal{A}}(\mathsf{O}.S_{[*]}(\rho_0))$"

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 a W^*-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 $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}{=}\ (X, \mathcal{F}, F),\ \overline{S}_{[*]}(w_0))$ 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 $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}{=}\ (X, \mathcal{F}, F),\ S_{[*]}(\rho_0))$ 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 8.3. 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

$$ \mathsf{M}_{L^\infty(\Omega.\nu)}\big(\mathsf{O}{=}(X, \mathcal{F}, F), \overline{S}_{[*]}(w_0)\big) = \mathsf{M}_{L^\infty(\Omega.\nu)}\big(\mathsf{O}{=}(X, \mathcal{F}, F), S_{[*]}(\rho_0)\big). $$

Hence, (C_1) and (C_2) are not so different. In order to avoid confusions, we use the following notation:

$$\begin{cases} W^*\text{-mixed state } w_0 \ (\in \overline{\mathfrak{S}}^m(\overline{\mathcal{A}}_*) \text{ is written by the Roman alphabet (e.g., } w_0, w, v, ...) \\ \\ C^*\text{-mixed state } \rho_0 \ (\in \mathfrak{S}^m(\mathcal{A}^*) \text{ is written by the Greek alphabet (e.g., } \rho_0, \rho, ...) \end{cases}$$

///

8.2 Simple examples in mixed measurement theory

Recall the following wise sayings:

Experience is the best teacher, or *Custom makes all things.*

Review 8.4. [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 ?

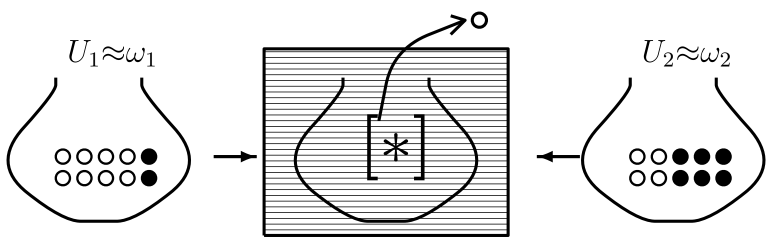


Figure 8.1 (= Figure 5.6:): Pure measurement (Fisher's maximum likelihood method)

Answer Consider the state space $\Omega = \{\omega_1, \omega_2\}$ with the discrete topology and the measure ν such that

$$\nu(\{\omega_1\}) = 1, \qquad \nu(\{\omega_2\}) = 1 \tag{8.3}$$

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

$$\begin{aligned} &[F_{WB}(\{W\})](\omega_1) = 0.8, && [F_{WB}(\{B\})](\omega_1) = 0.2 \\ &[F_{WB}(\{W\})](\omega_2) = 0.4, && [F_{WB}(\{B\})](\omega_2) = 0.6. \end{aligned} \tag{8.4}$$

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 8.5. [mixed measurement $\mathsf{M}_{L^\infty(\Omega,\nu)}(\mathsf{O} = (X, \mathcal{F}, F), S_{[*]}(w))$]

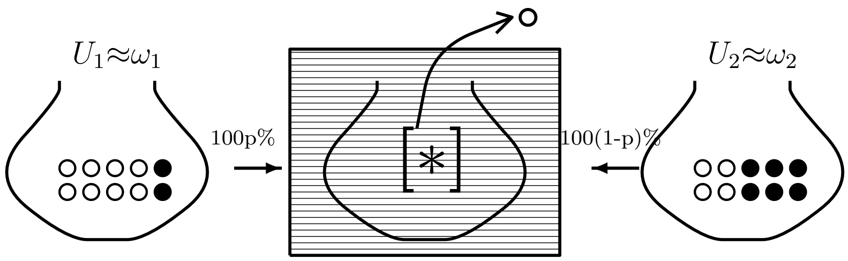


Figure 8.2: Mixed measurement (Urn problem)

($\sharp_1$) 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,

$$w(\omega) = \begin{cases} p & (\text{ if } \omega = \omega_1 \text{ }) \\ 1-p & (\text{ if } \omega = \omega_2 \text{ }) \end{cases}$$

($\sharp_2$) 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:

$$\nu(\{\omega_1\}) = 1, \qquad \nu(\{\omega_2\}) = 1. \tag{8.5}$$

Thus, we start from the classical basic structure:

$$[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))], \tag{8.6}$$

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

$$\begin{aligned} &[F_{WB}(\{W\})](\omega_1) = 0.8, && [F_{WB}(\{B\})](\omega_1) = 0.2 \\ &[F_{WB}(\{W\})](\omega_2) = 0.4, && [F_{WB}(\{B\})](\omega_2) = 0.6. \end{aligned} \tag{8.7}$$

Also, the mixed state $w_0 \in L^1_{+1}(\Omega, \nu)$ is defined by

$$w_0(\omega_1) = p, \qquad w_0(\omega_2) = 1 - p. \tag{8.8}$$

Then, by Axiom$^{(m)}$ 1, we see

(a): the probability that a measured value x $(\in \{W, B\})$ 7.2is obtained by $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}{=}\ (\{W,\ B\},\ 2^{\{W,B\}},\ F), S_{[*]}(w))$ is given by

$$\begin{aligned} P(\{x\}) &= {}_{L^1(\Omega)}\big(w_0, F(\{x\})\big)_{L^\infty(\Omega)} = \int_\Omega [F(\{x\})](\omega) \cdot w_0(\omega)\nu(d\omega) \\ &= p[F(\{x\})](\omega_1) + (1-p)[F(\{x\})](\omega_2) \\ &= \begin{cases} 0.8p + 0.4(1-p) & (\text{when} \ \ x = W) \\ 0.2p + 0.6(1-p) & (\text{when} \ \ x = B) \end{cases} \end{aligned} \tag{8.9}$$

The question (b) will be answered in Answer 8.13. □

♠**Note 8.1.** The following question is natural. That is,

($\flat_1$) In the above ($\sharp_1$) in Problem8.5, why is “the possibility that $[\ast] = \omega_1$ is $100p\%$ $\cdots$” replaced by “the probability that $[\ast] = \omega_1$ is $100p\%$ $\cdots$” ?

However, the linguistic Copenhagen interpretation says that

($\flat_2$) there is no probability without measurements.

This is the reason why the term “probability” is not used in (i). However, from a practical point of view, we are not sensitive to the difference between “probability” and “possibility”.

For further information see my homepage

Example 8.6. [Mixed spin measurement $\mathsf{M}_{B(\mathbb{C}^2)}(\mathsf{O} = (X = \{\uparrow, \downarrow\}, 2^X, F^z), \overline{S}_{[*]}(w))$] Consider the quantum basic structure:

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

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 $\left\{ \begin{array}{c} p \\ 1-p \end{array} \right\}$

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 $\mathsf{M}_{B(\mathbb{C}^2)}(\mathsf{O}_z = (X, 2^X, F^z), \overline{S}_{[*]}(w))$ 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 $\mathsf{M}_{B(\mathbb{C}^2)}(\mathsf{O}_z = (X, 2^X, F^z), \overline{S}_{[*]}(w))$ 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 8.7. As seen in the above, we say

For further information see my homepage

(a) Pure measurement theory is fundamental. Adding the concept of "mixed state", we can construct mixed measurement theory as follows.

$$\underset{\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O},\overline{S}_{[*]}(w))}{\boxed{\text{mixed measurement theory}}} := \underset{\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O},S_{[*]})}{\boxed{\text{pure measurement theory}}} + \underset{w}{\boxed{\text{mixed state}}}$$

That is, we usually devote ourselves to the case that

(b) *there is no mixed measurement without pure measurement.*

Hence, in quantum language, there is no confrontation between "frequency probability" and "subjective probability". The reason that a coin-tossing is used in Problem 8.5 is to emphasize that the naming of "subjective probability" is improper.

8.3 St. Petersburg two envelopes problem

This section is extracted from the following:

Ref. [58]: 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 envelopes problem (*cf.* ref. [9][1]).

Problem 8.8. [The St. Petersburg two envelopes 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 8.3.: Two envelopes 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 are both infinite. 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").

8.3.1 (P2): St. Petersburg two envelopes 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

$$X = \Omega, \quad \mathcal{F} = 2^X \equiv \{\Xi \mid \Xi \subseteq X\}$$

$$[F(\Xi)](\omega) = \chi_\Xi(\omega) \equiv \begin{cases} 1 & (\omega \in \Xi) \\ 0 & (\omega \notin \Xi) \end{cases} \qquad (\forall \Xi \in \mathcal{F}, \forall \omega \in \Omega)$$

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) (§8.1) 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 8.9. 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"*.

[1] D.J. Chalmers, "The St. Petersburg Two-Envelope Paradox," Analysis,Vol.62, 155-157, (2002)

♠**Note 8.2.** There are various criteria other than expectations. For example, consider the criterion such that

(♯) "the probability that the switching is disadvantageous" $< \frac{1}{2}$

Under this criterion, it is reasonable to judge that

$$\begin{cases} m = 1 & \Longrightarrow \text{switching to } B \\ m = 2, 3, ... & \Longrightarrow \text{keeping } A \end{cases}$$

8.4 Bayesian statistics is to use Bayes theorem in mixed measurement theory

Although there may be several opinions about the question "What is Bayesian statistics ?", we think that

Bayesian statistics is to use Bayes' theorem in mixed measurement.

Thus,

let us start from Bayes' theorem.

Recall Remark 8.7, i.e., "there is no mixed measurement without pure measurement", thus, we says

"there is no Bayesian statistics without pure measurement".

Or,

(♯) pure measurements (Fisher's statistics) are more fundamental than Bayesian statistics. (Or, Bayesian statistics is a variant of Fisher statistics.)

The following is clear.

Theorem 8.10. [The conditional probability]. Consider the mixed measurement $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}= (X \times Y, \mathcal{F} \boxtimes \mathcal{G}, H), \overline{S}_{[*]}(w))$, 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 $\mathsf{M}_{\overline{\mathcal{A}}}(\mathsf{O}= (X \times Y, \mathcal{F} \boxtimes \mathcal{G}, H), \overline{S}_{[*]}(w))$ 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 well-known property of conditional probability. □

In the classical case, this is rewritten as follows.

Theorem 8.11. [Bayes' Theorem (in classical mixed measurement)]. Consider the simultaneous measurement $\mathsf{M}_{\overline{\mathcal{A}}}\big(\mathsf{O}{=}\ (X\times Y, \mathcal{F}\boxtimes\mathcal{G}, F\times G),\ S_{[*]}(w_0)\big)$ 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,

$$(F\times G)(\Xi\times\Gamma)=F(\Xi)\cdot G(\Gamma)\qquad(\forall\Xi\in\mathcal{F},\forall\Gamma\in\mathcal{G}). \tag{8.10}$$

Assume that

(a) a measured value (x,y) $(\in X\times Y)$ obtained by the mixed measurement $\mathsf{M}_{L^\infty(\Omega)}\big(\mathsf{O}_{12}{=}\ (X\times Y,\mathcal{F}\boxtimes\mathcal{G},F\times G),\ S_{[*]}(w_0)\big)$ belongs to $\Xi\times Y$ (where, $\Xi\in\mathcal{F}$).

Then, the probability such that "$y\in\Gamma$" is given by

$$\frac{{}_{L^1(\Omega)}(w_0, H(\Xi\times\Gamma))_{L^\infty(\Omega)}}{{}_{L^1(\Omega)}(w_0, H(\Xi\times Y))_{L^\infty(\Omega)}}\Big(=\frac{\int_\Omega[F(\Xi)](\omega)\cdot[G(\Gamma)](\omega)\cdot w_0(\omega)\nu(d\omega)}{\int_\Omega[F(\Xi)](\omega)\cdot w_0(\omega)\nu(d\omega)}\Big). \tag{8.11}$$

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:

$$(8.11)=\int_\Omega[G(\Gamma)](\omega)w_{\rm new}(\omega)\nu(d\omega)\quad(\forall\Gamma\in\mathcal{G}). \tag{8.12}$$

Remark 8.12. [How to understand Bayes' Theorem] Bayes' theorem 8.11 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 Copenhagen interpretation, which says

A state never moves.

In this sense, the above (b) or (b′) (i.e., Bayes' theorem) is convenient and makeshift.

Answer 8.13. [About Bayes' Theorem (=Problem 8.5: (a) and (b))]

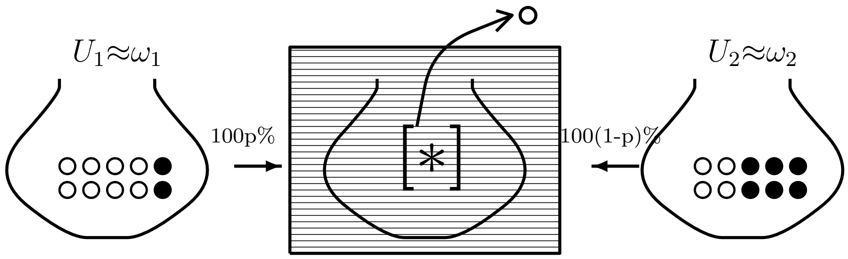


Figure 8.2 (in Problem 8.5): Mixed measurement (Urn problem)

Assume that the state space $\Omega = \{\omega_1, \omega_2\}$ is defined by the discrete metric with the following measure ν:

$$\nu(\{\omega_1\}) = 1, \qquad \nu(\{\omega_2\}) = 1. \tag{8.13}$$

Thus, we start from the classical basic structure:

$$[C_0(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^2(\Omega, \nu))], \tag{8.14}$$

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

$$\begin{aligned} &[F_{WB}(\{W\})](\omega_1) = 0.8, \qquad && [F_{WB}(\{B\})](\omega_1) = 0.2, \\ &[F_{WB}(\{W\})](\omega_2) = 0.4, && [F_{WB}(\{B\})](\omega_2) = 0.6. \end{aligned} \tag{8.15}$$

Also, the mixed state $w_0 \in L^1_{+1}(\Omega, \nu)$ is defined by

$$w_0(\omega_1) = p, \qquad w_0(\omega_2) = 1 - p. \tag{8.16}$$

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

$$\begin{aligned} P(\{x\}) &= {}_{L^1(\Omega)}\big(w_0, F(\{x\})\big)_{L^\infty(\Omega)} = \int_\Omega [F(\{x\})](\omega) \cdot w_0(\omega) \nu(d\omega) \\ &= p[F(\{x\})](\omega_1) + (1-p)[F(\{x\})](\omega_2) \\ &= \begin{cases} 0.8p + 0.4(1-p) & (\text{when } \ x = W) \\ 0.2p + 0.6(1-p) & (\text{when } \ x = B) \end{cases} \end{aligned} \tag{8.17}$$

(b)

• [W^*-algebraic answer to Problem 8.5(b)]

For further information see my homepage

Since "white ball" is obtained by a mixed measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[*]}(w_0))$, a new mixed state $w_{\rm new}(\in L^1_{+1}(\Omega))$ is given by

$$w_{\rm new}(\omega) = \frac{[F(\{W\})](\omega) w_0(\omega)}{\int_\Omega [F(\{W\})](\omega) w_0(\omega) \nu(d\omega)} = \begin{cases} \dfrac{0.8p}{0.8p + 0.2(1-p)} & (\text{when } \omega = \omega_1) \\ \\ \dfrac{0.2(1-p)}{0.8p + 0.2(1-p)} & (\text{when } \omega = \omega_2) \end{cases}$$

• [C^*-algebraic answer to Problem 8.5(b)]
Since "white ball" is obtained by a mixed measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[*]}(\rho_0))$, a new mixed state $\rho_{\rm new}(\in \mathcal{M}_{+1}(\Omega))$ is given by

$$\rho_{\rm new} = \frac{F(\{W\})\rho_0}{\int_\Omega [F(\{W\})](\omega) \rho_0(d\omega)} = \frac{0.8p}{0.8p + 0.2(1-p)} \delta_{\omega_1} + \frac{0.2(1-p)}{0.8p + 0.2(1-p)} \delta_{\omega_2}.$$

8.5 Two envelopes problem (Bayes' method)

This section is extracted from the following:

> ref. [58]: S. Ishikawa; The two envelopes paradox in non-Bayesian and Bayesian statistics (arXiv:1408.4916v4 [stat.OT] 2014)

Problem 8.14. [**(=Problem 5.16): the two envelopes 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) $\quad \frac{V_1}{V_2} = 1/2$ 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 whether $\overline{x}_1 = 2x_1$ 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 8.4: Two envelopes problem (in Bayesian statistics)

[(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

$$E_{\text{other}}(\alpha) = (1/2)(\alpha/2) + (1/2)(2\alpha) = 1.25\alpha. \tag{8.18}$$

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").

8.5.1 (P1): Bayesian approach to the two envelopes 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:

$$\Omega \ni \omega \underset{\text{(identification)}}{\longleftrightarrow} (\omega, 2\omega) \in \widehat{\Omega}. \tag{8.19}$$

Furthermore, 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

$$[F(\Xi)](\omega) = \begin{cases} 1 & (\text{ if } \omega \in \Xi, \ \ 2\omega \in \Xi) \\ 1/2 & (\text{ if } \omega \in \Xi, \ \ 2\omega \notin \Xi) \\ 1/2 & (\text{ if } \omega \notin \Xi, \ \ 2\omega \in \Xi) \\ 0 & (\text{ if } \omega \notin \Xi, \ \ 2\omega \notin \Xi) \end{cases} \qquad (\forall \omega \in \Omega, \forall \Xi \in \mathcal{F})$$

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 $\rho_0(\in \mathcal{M}^m(\Omega(= \overline{\mathbb{R}}_+)))$ has the probability density function w_0. Axiom$^{(m)}$ 1 (§8.1) says

(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

$$\begin{aligned} P(\Xi) &= \int_\Omega [F(\Xi)](\omega)\rho_0(d\omega) = \int_\Omega [F(\Xi)](\omega) w_0(\omega) d\omega \\ &= \int_\Xi \frac{w_0(x/2)}{4} + \frac{w_0(x)}{2} \;\; dx \quad (\forall \Xi \in \mathcal{B}_{\overline{\mathbb{R}}_+}). \end{aligned} \tag{8.20}$$

Therefore, the expectation is given by

$$\int_{\overline{\mathbb{R}}_+} xP(dx) = \frac{1}{2}\int_0^\infty x \cdot \Big(w_0(x/2)/2 + w_0(x)\Big) dx = \frac{3}{2}\int_{\overline{\mathbb{R}}_+} x w_0(x) dx. \tag{8.21}$$

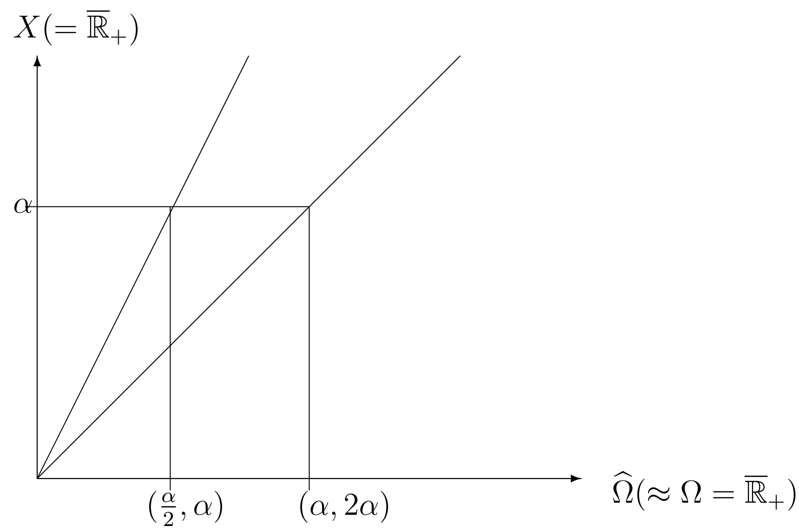


Figure 8.5(=Figure 5.10) : likelihood function

Furthermore, Theorem 8.11 (Bayes' theorem) says

(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 $\rho_{\rm post}(\in \mathcal{M}^m(\Omega))$ is given by

$$\rho_{\rm post}^\alpha = \frac{\frac{w_0(\alpha/2)}{2}}{\frac{h(\alpha/2)}{2} + w_0(\alpha)} \delta_{(\frac{\alpha}{2},\alpha)} + \frac{w_0(\alpha)}{\frac{w_0(\alpha/2)}{2} + w_0(\alpha)} \delta_{(\alpha,2\alpha)}. \tag{8.22}$$

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:

$$\int_{\overline{\mathbb{R}}_+} \Big((-\frac{\alpha}{2}) \frac{\frac{w_0(\alpha/2)}{2}}{\frac{w_0(\alpha/2)}{2} + w_0(\alpha)} + \alpha \frac{w_0(\alpha)}{\frac{w_0(\alpha/2)}{2} + w_0(\alpha)} \Big) P(d\alpha)$$
$$= \int_{\overline{\mathbb{R}}_+} (-\frac{\alpha}{2}) \frac{w_0(\alpha/2)}{4} + \alpha \cdot \frac{w_0(\alpha)}{2} \;\; d\alpha = 0. \tag{8.23}$$

Therefore, we see that the swapping is even, i.e., no advantage and no disadvantage.

8.6 Monty Hall problem (The Bayesian approach)

8.6.1 The review of Problem 5.14 (Monty Hall problem in pure measurement)

Problem 8.15. [**= Problem 5.14; Monty Hall problem; High school puzzle**] (The answer by Fisher's maximum likelihood method)

> 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
>
> (♭) 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 ?*

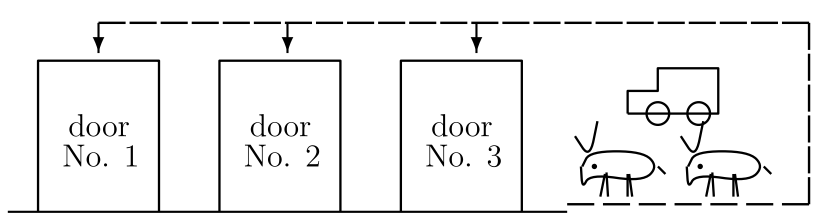
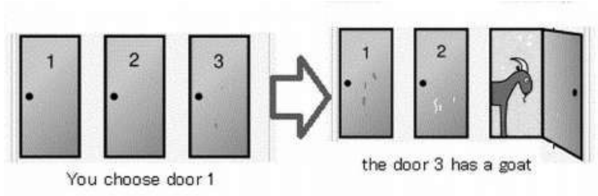


Figure 8.6: Monty Hall problem



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

$$\begin{aligned} &[F_1(\{1\})](\omega_1) = 0.0, && [F_1(\{2\})](\omega_1) = 0.5, && [F_1(\{3\})](\omega_1) = 0.5, \\ &[F_1(\{1\})](\omega_2) = 0.0, && [F_1(\{2\})](\omega_2) = 0.0, && [F_1(\{3\})](\omega_2) = 1.0, \\ &[F_1(\{1\})](\omega_3) = 0.0, && [F_1(\{2\})](\omega_3) = 1.0, && [F_1(\{3\})](\omega_3) = 0.0, \end{aligned} \tag{8.24}$$

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 :

a) "a measured value 1 is obtained $\Leftrightarrow$ The host says "Door 1 has a goat"

b) "measured value 2 is obtained $\Leftrightarrow$ The host says "Door 2 has a goat"

c) "measured value 3 is obtained $\Leftrightarrow$ 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. $\square$

8.6.2 Monty Hall problem in mixed measurement (=Bayesian measurement)

Next, let us study Monty Hall problem in mixed measurement theory (particularly, Bayesian statistics).

For further information see my homepage

Problem 8.16. [Monty Hall problem (The answer by Bayes' method)]

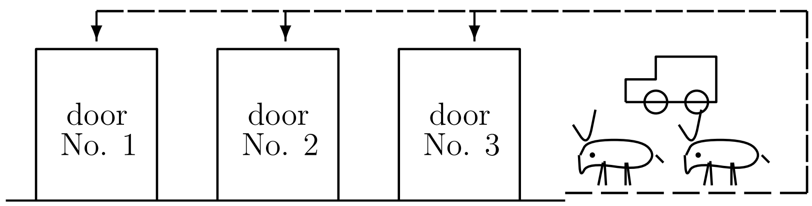
Suppose you are on a game show, and you are given a choice of three doors (i.e., "number 1," "number 2," "number 3"). Behind one door is a car, behind the others, goats. You pick a door, say number 1. Then, the host, who set a car behind a certain door, says

($\sharp_1$) 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 8.15 (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 8.11 (Bayes' theorem) says that the posterior state ν_{post} ($\in \mathcal{M}_{+1}(\Omega)$) is given by

$$\nu_{\text{post}} = \frac{F_1(\{3\}) \times \nu_0}{\langle \nu_0, F_1(\{3\}) \rangle}.$$

That is,

$$\nu_{\rm post}(\{\omega_1\}) = \frac{\frac{p_1}{2}}{\frac{p_1}{2}+p_2}, \quad \nu_{\rm post}(\{\omega_2\}) = \frac{p_2}{\frac{p_1}{2}+p_2}, \quad \nu_{\rm post}(\{\omega_3\}) = 0.$$

Particularly, we see that

- if $p_1 = p_2 = p_3 = 1/3$, then it holds that $\nu_{\rm post}(\{\omega_1\}) = 1/3$, $\nu_{\rm post}(\{\omega_2\}) = 2/3$, $\nu_{\rm post}(\{\omega_3\}) = 0$, *and thus, you should pick Door 2.* □

♠**Note 8.3.** It is not natural to assume the rule ($\sharp_1$) in Problem 8.16. That is because the host may intentionally set the car behind a certain door. Thus, we think that Problem 8.16 is temporary. For our formal assertion, see Problem 8.17 latter.

8.7 Monty Hall problem (The principle of equal weight)

8.7.1 The principle of equal weight – The most famous unsolved problem

Let us reconsider Monty Hall problem (Problem 8.14, Problem8.15) in what follows. We think that the following is one of the most reasonable answers (also, see Problem 15.5).

Problem 8.17. [Monty Hall problem (The principle of equal weight)]

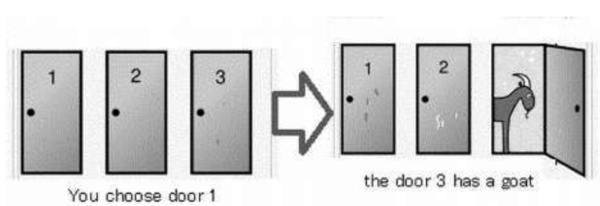
Suppose you are on a game show, and you are given a choice of three doors (i.e., "number 1," "number 2," "number 3"). Behind one door is a car, behind the others, goats.

($\sharp_2$) *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 ?



For further information see my homepage

Answer: By the same way of Problem8.15 and Problem8.16 (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 (8.11). With 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 :

(a) By the dice-throwing, you get $\begin{bmatrix} 1,2 \\ 3,4 \\ 5,6 \end{bmatrix}$, then, take $\begin{bmatrix} \mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[\omega_k]}) \\ \mathsf{M}_{L^\infty(\Omega)}(\Phi\mathsf{O}, S_{[\omega_k]}) \\ \mathsf{M}_{L^\infty(\Omega)}(\Phi^2\mathsf{O}, S_{[\omega_k]}) \end{bmatrix}$

as a measurement. We, by the argument in Chapter 11 (*cf.* the formula (10.7))[2], see the following identifications:

$$\begin{aligned} \mathsf{M}_{L^\infty(\Omega)}(\Phi\mathsf{O}, S_{[\omega_k]}) &= \mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[\phi(\omega_k)]}), \mathsf{M}_{L^\infty(\Omega)}(\Phi^2\mathsf{O}, S_{[\omega_k]}) \\ &= \mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[\phi^2(\omega_k)]}). \end{aligned}$$

Thus, the above (a) is equal to

(b) By the dice-throwing, you get $\begin{bmatrix} 1,2 \\ 3,4 \\ 5,6 \end{bmatrix}$ then, take $\begin{bmatrix} \mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[\omega_k]}) \\ \mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[\phi(\omega_k)]}) \\ \mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[\phi^2(\omega_k)]}) \end{bmatrix}$

as a measurement. 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 8.17 is the same as Problem 8.16. Hence, you should choose the door 2. □

♠**Note 8.4.** 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 case as more probable than others, 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 8.18. [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 $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, \overline{S}_{[*]}(w_e))$ $\Big($ or, $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}, S_{[*]}(\nu_e))$ $\Big)$, where

$$w_e(\omega_k) = 1/n \ \ (\forall k = 1, 2, ..., n) \qquad \text{or} \qquad \nu_e = \frac{1}{n}\sum_{k=1}^{n} \delta_{\omega_k}.$$

[2]Thus, from the pure theoretical point of view, this problem should be discussed after Chapter 11.

For further information see my homepage

Proof. The proof is a easy consequence of the above Monty Hall problem (or, see [35, 43]). □

♠**Note 8.5.** 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 8.18

($\sharp_3$) the principle of equal weight in Proclaim 15.4

8.8 Averaging information (Entropy)

As one of applications (of Bayes' theorem), we now study the "entropy (*cf.* ref. [108])" of the measurement. This section is due to the following references.

($\sharp$) Ref. [30]: 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

($\sharp$) Ref. [35]: S. Ishikawa, "Mathematical Foundations of Measurement Theory," Keio University Press Inc. 2006.

Let us begin with the following definition.

Definition 8.19. [**Entropy**(*cf.* refs. [30, 35])] Assume

$$\text{Classical basic structure } [C_0(\Omega) \subseteq L^\infty(\Omega,\nu) \subseteq B(L^2(\Omega,\nu))] \ .$$

Consider a mixed measurement $\mathsf{M}_{L^\infty(\Omega,\nu)}$ $(\mathsf{O}=(X,2^X,F), \overline{S}_{[*]}(w_0))$ 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 $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O},\overline{S}_{[*]}(w_0))$ is given by

$$P(\{x_n\}) = \int_\Omega [F(\{x_n\})](\omega) w_0(\omega)\nu(d\omega). \tag{8.25}$$

Furthermore, when a measured value x_n is obtained, the information $I(\{x_n\})$, from Bayes' theorem 8.11, 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 $H\big(\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O},\overline{S}_{[*]}(w_0))\big)$ of the mixed measurement $\mathsf{M}_{L^\infty(\Omega)}$ $(\mathsf{O}, \overline{S}_{[*]}(w_0))$ is naturally defined by

$$H\big(\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O},\overline{S}_{[*]}(w_0))\big) = \sum_{n=1}^\infty P(\{x_n\}) \cdot I(\{x_n\}). \tag{8.26}$$

Also, the following is clear:

$$H\big(\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O},\overline{S}_{[*]}(w_0))\big)=\sum_{n=1}^{\infty}\int_\Omega [F(\{x_n\})](\omega)\log[F(\{x_n\})](\omega)w_0(\omega)\nu(d\omega)$$
$$-\sum_{n=1}^{\infty}P(\{x_n\})\log P(\{x_n\}). \tag{8.27}$$

Example 8.20. [The offender is male 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

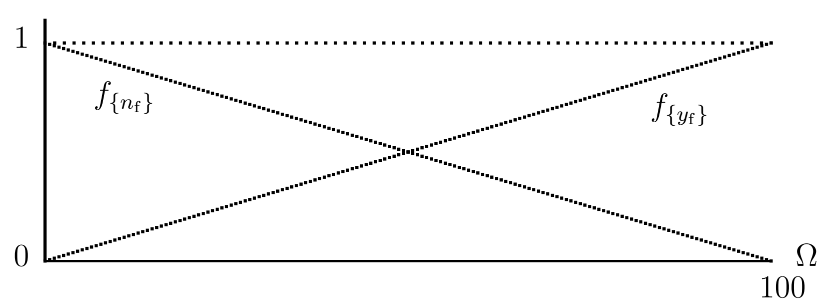
$$[M(\{y_{\rm m}\})](\omega_n)=m_{y_{\rm m}}(\omega_n)=\begin{cases}0 & (n \text{ is odd})\\ 1 & (n \text{ is even})\end{cases}$$
$$[M(\{n_{\rm m}\})](\omega_n)=m_{n_{\rm m}}(\omega_n)=1-[M(\{y_{\rm m}\})](\omega_n)$$

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 8.18), 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 $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{\rm m},\overline{S}_{[*]}(w_e))$ and $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{\rm f},\overline{S}_{[*]}(w_e))$. Then, we see:

$$H\big(\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{\rm m},\overline{S}_{[*]}(w_e))\big) = \int_\Omega m_{y_{\rm m}}(\omega)w_e(\omega)\nu(d\omega)\cdot\log\int_\Omega m_{y_{\rm m}}(\omega)w_e(\omega)\nu(d\omega)$$

$$
\begin{aligned}
& -\int_\Omega m_{\{n_{\rm m}\}}(\omega)w_e(\omega)\nu(d\omega)\cdot\log\int_\Omega m_{n_{\rm m}}(\omega)w_e(\omega)\nu(d\omega) \\
= \;& -\frac{1}{2}\log\frac{1}{2}-\frac{1}{2}\log\frac{1}{2}=\log_2 2=1 \;\; \text{(bit)}^{3}.
\end{aligned}
$$

Also,

$$
\begin{aligned}
H\big(\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{\rm f},\overline{S}_{[*]}(w_e))\big) \;=\;& \int_\Omega f_{y_{\rm f}}(\omega)\log f_{y_{\rm f}}(\omega)w_e(\omega)\nu(d\omega) \\
&+\int_\Omega f_{n_{\rm f}}(\omega)\log f_{n_{\rm f}}(\omega)w_e(\omega)\nu(d\omega)-\int_\Omega f_{y_{\rm f}}(\omega)w_e(\omega)\nu(d\omega)\cdot\log\int_\Omega f_{y_{\rm f}}(\omega)w_e(\omega)\nu(d\omega) \\
&-\int_\Omega f_{n_{\rm f}}(\omega)w_e(d\omega)\cdot\log\int_\Omega f_{n_{\rm f}}(\omega)w_e(\omega)\nu(d\omega) \\
&\doteqdot 2\int_0^1\lambda\log_2\lambda d\lambda+1=-\frac{1}{2\log_e 2}+1=0.278\cdots\text{(bit)}
\end{aligned}
$$

Therefore, as eyewitness information, “male or female” has more valuable than “fast or slow”.

8.9 Fisher statistics: Monty Hall problem [three prisoners problem]

This section is extracted from the following:

> Ref. [57]: S. Ishikawa; The Final Solutions of Monty Hall Problem and Three Prisoners Problem (arXiv:1408.0963v1 [stat.OT] 2014)

In Sections 8.9 $\sim$ 8.11, I will discuss Monty Hall Problem and Three Prisoners Problem in parallel. As Monty Hall Problem has already been discussed, only Three Prisoners Problem may be read. However, reading the two in parallel has the following advantages.
It is usually said that

Monty Hall problem and three prisoners problem are the so-called isomorphism problem.

But, we think that the meaning of “isomorphism problem” is ambiguous, or, it is not able to be clarified without measurement or the dualism. Therefore, in order to understand “isomorphism”, we simultaneously discuss the two

- $\begin{cases}\text{Monty Hall problem}\\ \text{three prisoners problem}\end{cases}$

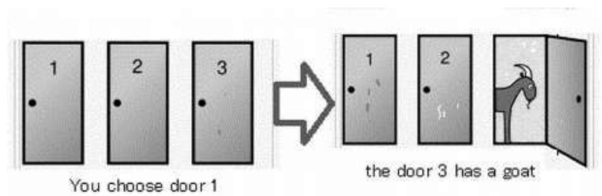
8.9.1 Fisher statistics: Monty Hall problem and three prisoners problem

Problem 8.21. (=Problem 8.15: [Monty Hall problem]).

Suppose you are on a game show, and you are given 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

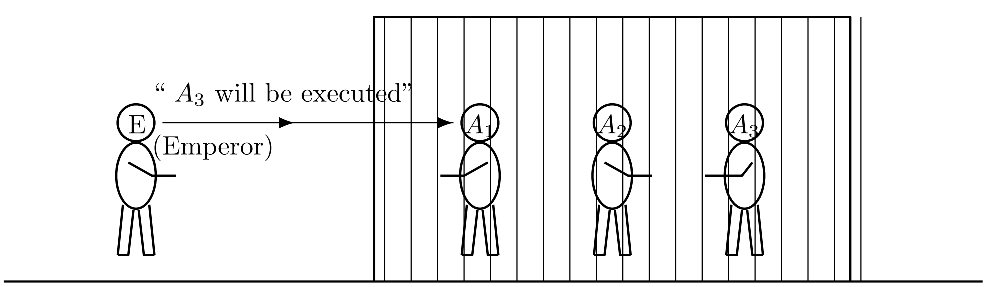
However, you pick a door, say "Door A_1", and the host, who knows what's behind the doors, opens another door, say "Door A_3," which has a goat.

He says to you, "Do you want to pick Door A_2 ?" Is it to your advantage to switch your choice of doors ?



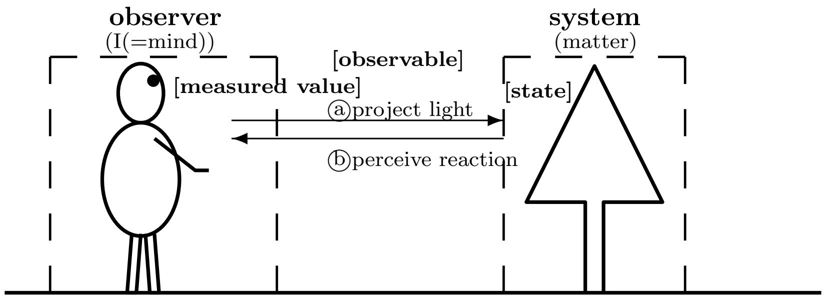
Problem 8.22. [three prisoners problem].

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 of 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 $\frac{1}{3(=\text{Num}\{A_1,A_2,A_3\}])}$ to $\frac{1}{2(=\text{Num}\{A_1,A_2\}])}$. This prisoner A_1's happiness may or may not be reasonable ?



8.9.2 The answer in Fisher statistics: Monty Hall problem and three prisoners problem

Let rewrite the spirit of dualism (Descartes figure) as follows.



Descartes Figure 8.7: The image of "measurement(=①+②)" in dualism

In the dualism, we have the confrontation

$$\text{"observer} \longleftrightarrow \text{system"}$$

as follows.

Table 8.1: 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, the first answer to $\left[\begin{array}{l}\text{Problem 8.21 (Monty-Hall problem)}\\ \text{Problem 8.22 (Three prisoners problem)}\end{array}\right]$ is given in classical pure measurement theory. The two will be simultaneously solved as follows. The spirit of dualism (in Figure 8.7) urges us to declare that

$$\text{(A)} \quad \left[\begin{array}{l}\text{"observer} \approx \text{you" and "system} \approx \text{three doors" in Problem 8.21}\\ \text{"observer} \approx \text{prisoner } A_1\text{" and "system} \approx \text{emperor's mind" in Problem 8.22}\end{array}\right]$$

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

$$\left[\begin{array}{l}\delta_{\omega_m} \Leftrightarrow \text{ the state that the car is behind the door } A_m\\ \delta_{\omega_m} \Leftrightarrow \text{ the state that the prisoner } A_m \text{ is will be executed}\end{array}\right]$$
$$(m = 1, 2, 3) \tag{8.28}$$

Define the observable $\mathsf{O}_1 \equiv (\{1,2,3\}, 2^{\{1,2,3\}}, F_1)$ in $L^\infty(\Omega)$ such that

$$\begin{array}{lll}[F_1(\{1\})](\omega_1) = 0.0, & [F_1(\{2\})](\omega_1) = 0.5, & [F_1(\{3\})](\omega_1) = 0.5,\\ [F_1(\{1\})](\omega_2) = 0.0, & [F_1(\{2\})](\omega_2) = 0.0, & [F_1(\{3\})](\omega_2) = 1.0,\\ [F_1(\{1\})](\omega_3) = 0.0, & [F_1(\{2\})](\omega_3) = 1.0, & [F_1(\{3\})](\omega_3) = 0.0,\end{array} \tag{8.29}$$

For further information see my homepage

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 a 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 $\left[\begin{array}{l} \text{the host said "Door 3 has a goat"} \\ \text{the emperor said "Prisoner } A_3 \text{ will be executed"} \end{array}\right]$.

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

$$
\begin{aligned}
&[F_1(\{3\})](\omega_2) = 1.0 = \max\{0.5, \ \ 1.0, \ \ 0.0\} \\
&= \max\{[F_1(\{3\})](\omega_1), [F_1(\{3\})](\omega_2), [F_1(\{3\})](\omega_3)\}.
\end{aligned} \tag{8.30}
$$

Therefore, Theorem 5.6 (Fisher's maximum likelihood method) says :

(B_1) In Problem 8.21 (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 8.22 (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 8.22 is not within Fisher's maximum likelihood method.

8.10 Bayesian statistics: Monty Hall problem and three prisoners problem

This section is extracted from the following:

Ref. [57]: S. Ishikawa; The Final Solutions of Monty Hall Problem and Three Prisoners Problem (arXiv:1408.0963v1 [stat.OT] 2014)

8.10.1 Bayesian statistics: Monty Hall problem and three prisoners problem

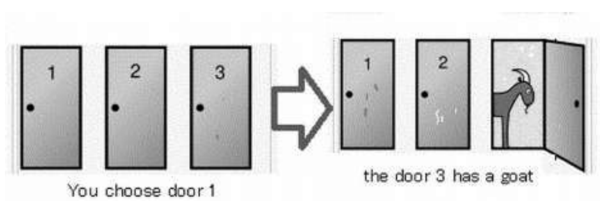
Problem 8.23. [(=Problem8.16) Monty Hall problem (the case that the host throws the dice)].

Suppose you are on a game show, and you are given a 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 is decided by a cast of a 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 ?

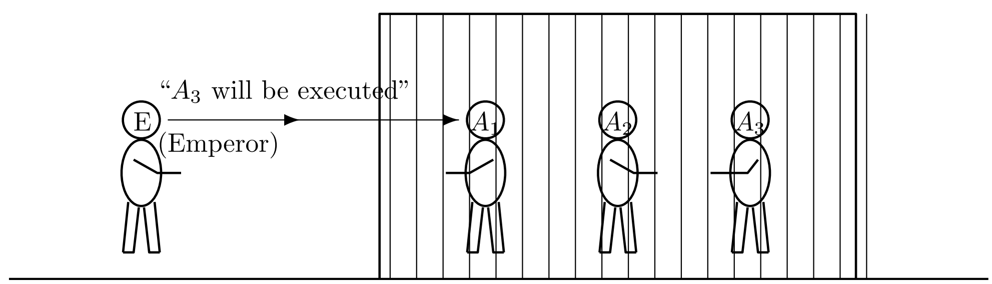


Problem 8.24. [three prisoners problem].

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 of 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 ?



8.10.2 The answer in Bayesian statistics: Monty Hall problem and three prisoners problem

In the dualism, we have the confrontation

$$\text{“observer}\longleftrightarrow\text{system”}$$

as follows.

Table 8.2: 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

Let $\Omega = \{\omega_1, \omega_2, \omega_3\}$ be a state space with the discrete metric. Each pure state $\delta_{\omega_m} (\in \mathfrak{S}^p(C(\Omega)^*))$ means as follows.

$$\begin{aligned} &\delta_{\omega_m} \Leftrightarrow \text{The state such that a car is behind the door } A_m \\ &[\text{resp.} \quad \delta_{\omega_m} \Leftrightarrow \text{ the state such that a prisoner } A_m \text{ is pardoned }] \\ &(m = 1, 2, 3) \end{aligned} \tag{8.31}$$

The observable $\mathsf{O}_1 \equiv (\{1, 2, 3\}, 2^{\{1,2,3\}}, F_1)$ is defined by

$$\begin{aligned} &[F_1(\{1\})](\omega_1) = 0.0, \quad &[F_1(\{2\})](\omega_1) = 0.5, \quad &[F_1(\{3\})](\omega_1) = 0.5, \\ &[F_1(\{1\})](\omega_2) = 0.0, \quad &[F_1(\{2\})](\omega_2) = 0.0, \quad &[F_1(\{3\})](\omega_2) = 1.0, \\ &[F_1(\{1\})](\omega_3) = 0.0, \quad &[F_1(\{2\})](\omega_3) = 1.0, \quad &[F_1(\{3\})](\omega_3) = 0.0, \end{aligned} \tag{8.32}$$

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 mixed measurement $\mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_1, S_{[*]}(\nu_0))$ (where, $\nu_0 = p_1\delta_{\omega_1} + p_2\delta_{\omega_2} + p_3\delta_{\omega_3}$), 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 8.11 says that the posterior state ν_{post} ($\in \mathcal{M}_{+1}^m(\Omega)$) is given by

$$\nu_{\text{post}} = \frac{F_1(\{3\}) \times \nu_0}{\langle \nu_0, F_1(\{3\}) \rangle}. \tag{8.33}$$

That is,

$$\nu_{\text{post}}(\{\omega_1\}) = \frac{\frac{p_1}{2}}{\frac{p_1}{2} + p_2}, \quad \nu_{\text{post}}(\{\omega_2\}) = \frac{p_2}{\frac{p_1}{2} + p_2}, \quad \nu_{\text{post}}(\{\omega_3\}) = 0. \tag{8.34}$$

Then,

(I1) In Problem 8.23,

$$\begin{cases} \text{if } \nu_{\text{post}}(\{\omega_1\}) < \nu_{\text{post}}(\{\omega_2\}) \ (\ p_1 < 2p_2), \text{ you should pick Door } A_2 \\ \text{if } \nu_{\text{post}}(\{\omega_1\}) = \nu_{\text{post}}(\{\omega_2\}) \ (p_1 = 2p_2), \text{ you may pick Doors } A_1 \text{ or } A_2 \\ \text{if } \nu_{\text{post}}(\{\omega_1\}) > \nu_{\text{post}}(\{\omega_2\}) \ (\ p_1 > 2p_2), \text{ you should not pick Door } A_1 \end{cases}$$

(I2) In Problem 8.24,

$$\begin{cases} \text{if } \nu_0(\{\omega_1\}) < \nu_{\text{post}}(\{\omega_1\}) \ (p_1 < 1 - 2p_2), \text{ the prisoner } A_1\text{'s happiness increases} \\ \text{if } \nu_0(\{\omega_1\}) = \nu_{\text{post}}(\{\omega_1\}) \ (\ p_1 = 1 - 2p_2), \text{ the prisoner } A_1\text{'s happiness is invariant} \\ \text{if } \nu_0(\{\omega_1\}) > \nu_{\text{post}}(\{\omega_1\}) \ (\ p_1 > 1 - 2p_2), \text{ the prisoner } A_1\text{'s happiness decreases} \end{cases}$$

For further information see my homepage

8.11 Equal probability: Monty Hall problem and three prisoners problem

This section is extracted from the following:

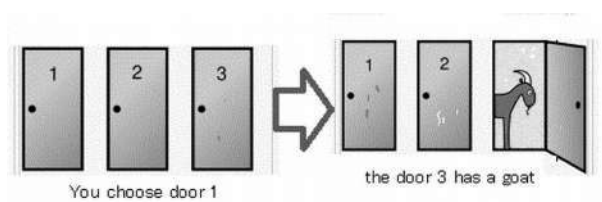
> ref. [57]: S. Ishikawa; The Final Solutions of Monty Hall Problem and Three Prisoners Problem (arXiv:1408.0963v1 [stat.OT] 2014)

Problem 8.25. [(=Problem8.16)Monty Hall problem (the case that you throws the dice)].

Suppose you are on a game show, and you are given a 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 ?

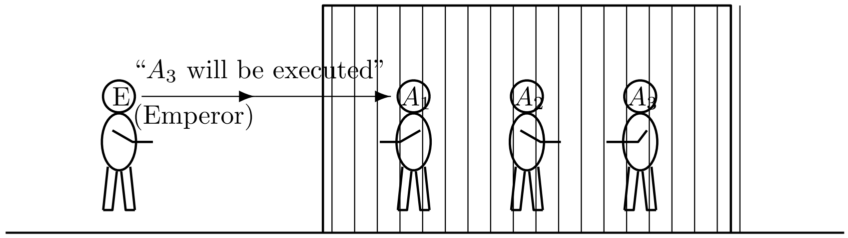


Problem 8.26. [three prisoners problem(the case that the prisoner throws the dice)].

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 of 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 $\frac{1}{3(=\text{Num}[\{A_1,A_2,A_3\}])}$ to $\frac{1}{2(=\text{Num}[\{A_1,A_2\}])}$. This prisoner A_1's happiness may or may not be reasonable ?



Answer : By Theorem 8.18 (The principle of equal weight), the above Problems 8.25 and 8.26 is respectively the same as Problems 8.23 and 8.24 in the case that $p_1 = p_2 = p_3 = 1/3$. Therefore,

(B_1) Problem8.25 [Monty Hall problem (the case that you throw a fair dice)]

$$\nu_{\text{post}}(\{\omega_1\}) < \nu_{\text{post}}(\{\omega_2\}) \text{ (i.e., } p_1 = 1/3 < 2/3 = 2p_2\text{),}$$
thus, you should choose the door A_2

(B_2) Problem8.26 [three prisoners problem (the case that the questioner was decided by the fair dice throw)],

$$\nu_0(\{\omega_1\}) = \nu_{\text{post}}(\{\omega_1\}) \text{ (i.e., } p_1 = 1/3 = 1 - 2p_2\text{),}$$
Thus, the happiness of the prisoner A_1 is invariant

♠**Note 8.6.** These problems (i.e., Monty Hall problem and the three prisoners problem) continued attracting the philosopher's interest. This is not due to the fact that these are easy to make a mistake for high school students, but

these problems include the essence of "dualism".

8.12 Bertrand's paradox

Theorem8.18 (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.

8.12.1 Bertrand's paradox("randomness" depends on how you look at)

Here, let us review the argument about the Bertrand paradox (*cf.* refs. [25, 35, 55]). Consider the following problem:

Problem 8.27. (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 $\sqrt{3}$?

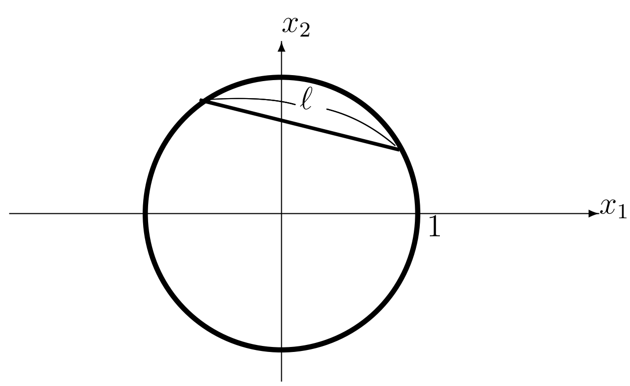


Figure 8.8: Bertrand' paradox

Define the rotation map $T^{\theta}_{\rm rot} : \mathbb{R}^2 \to \mathbb{R}^2$ $(0 \le \theta < 2\pi)$ and the reverse map $T_{\rm rev} : \mathbb{R}^2 \to \mathbb{R}^2$ such that

$$T^{\theta}_{\rm rot} x = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad T_{\rm rev} x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

For further information see my homepage

Problem 8.28. (Bertrand paradox and its answer) Given a circle with the radius 1.

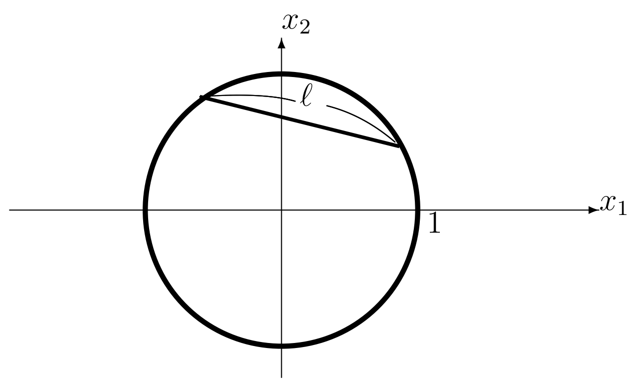


Figure 8.9: Bertrand' paradox

Put $\Omega = \{l \mid l \text{ is a chord}\}$, that is, the set of all chords.

(A) Can we uniquely define an invariant probability measure on Ω ?

Here, "invariant" means "invariant concerning the rotation map $T^{\theta}_{\mathrm{rot}}$ and reverse map T_{rev}".

In what follows, we show that the above invariant measure exists but it is not determined uniquely.

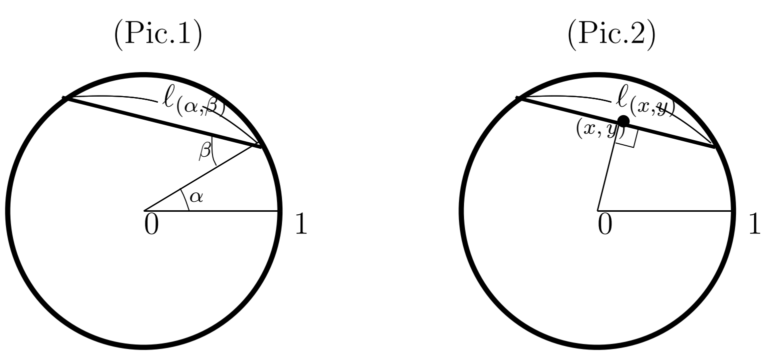


Figure 8.10: Two cases in Bertrand's paradox

cm

[The first answer (Pic.1(in Figure 8.10))]. In Pic.1, we see that the chord ℓ is represented by a point (α, β) in the rectangle $\Omega_1 \equiv \{(\alpha, \beta) \mid 0 < \alpha \leq 2\pi,\ 0 < \beta \leq \pi/2(\text{radian})\}$. That is, we have the following identification:

$$\Omega(= \text{the set of all chords}) \ni \omega = \ell_{(\alpha,\beta)} \underset{\text{identification}}{\longleftrightarrow} (\alpha, \beta) \in \Omega_1(\subset \mathbb{R}^2).$$

Note that we have the natural probability measure ν_1 on Ω_1 such that $\nu_1(A) = \frac{\mathrm{Meas}[A]}{\mathrm{Meas}[\Omega_1]} = \frac{\mathrm{Meas}[A]}{\pi^2}$ $(\forall A \in \mathcal{B}_{\Omega_1})$, where " Meas" = " Lebesgue measure". Transferring the probability measure ν_1 on Ω_1

For further information see my homepage

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 $T^\theta_{\rm rot}$ and reverse map $T_{\rm rev}$.

Therefore, we have a natural measurement $\mathsf{M}_{L^\infty(\Omega,\rho_1)}(\mathsf{O}_E \equiv (\Omega, \mathcal{B}_\Omega, F_E), S_{[*]}(1))$, where O_E is the exact observable, i.e., $[F_E((\Xi)](\omega) = \chi_{\Xi}(\omega)$ $(\forall \Xi \in \mathcal{B}_\Omega, \forall \omega \in \Omega)$. Consider the identification:

$$\Omega \supseteq \Xi_{\sqrt{3}} \underset{\text{identification}}{\longleftrightarrow} \{(\alpha,\beta) \in \Omega_1 \;:\; \text{"the length of } \ell_{(\alpha,\beta)}\text{"} < \sqrt{3}\} \subseteq \Omega_1$$

Then, Axiom$^{\rm (m)}$ 1 says that the probability that a measured value belongs to $\Xi_{\sqrt{3}}$ is given by

$$\begin{aligned}
&\int_\Omega [F_E(\Xi_{\sqrt{3}})](\omega)\ \rho_1(d\omega) = \int_{\Xi_{\sqrt{3}}} 1\ \ \rho_1(d\omega)\\
=&m_1(\{\ell_{(\alpha,\beta)} \approx (\alpha,\beta) \in \Omega_1 \mid \ \text{"the length of } \ell_{(\alpha,\beta)}\text{"} \le \sqrt{3}\})\\
=&\frac{\text{Meas}[\{(\alpha,\beta) \mid 0 \le \alpha \le 2\pi,\ \pi/6 \le \beta \le \pi/2\}]}{\text{Meas}[\{(\alpha,\beta) \mid 0 \le \alpha \le 2\pi,\ 0 \le \beta \le \pi/2\}]}\\
=&\frac{2\pi \times (\pi/3)}{\pi^2} = \frac{2}{3}.
\end{aligned}$$

[The second answer (Pic.2**(in Figure 8.10))].** 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:

$$\Omega(= \text{the set of all chords}) \ni \omega = \ell_{(x,y)} \underset{\text{identification}}{\longleftrightarrow} (x,y) \in \Omega_2 (\subset \mathbb{R}^2).$$

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,

$$\mathcal{M}_{+1}(\Omega) \ni \rho_2 \underset{\text{identification}}{\longleftrightarrow} \nu_2 \in \mathcal{M}_{+1}(\Omega_2)$$

($\sharp$) It is clear that the measure ρ_2 is invariant concerning the rotation map $T^\theta_{\rm rot}$ and reverse map $T_{\rm rev}$.

Therefore, we have a natural measurement $\mathsf{M}_{L^\infty(\Omega,\rho_2)}(\mathsf{O}_E \equiv (\Omega, \mathcal{B}_\Omega, F_E), S_{[*]}(1))$. Consider the identification:

$$\Omega \supseteq \Xi_{\sqrt{3}} \underset{\text{identification}}{\longleftrightarrow} \{(x,y) \in \Omega_2 \;:\; \text{"the length of } \ell_{(\alpha,\beta)}\text{"} < \sqrt{3}\} \subseteq \Omega_1$$

Then, Axiom$^{\rm (m)}$ 1 says that the probability that a measured value belongs to $\Xi_{\sqrt{3}}$ is given by

$$\int_\Omega [F_E(\Xi_{\sqrt{3}})](\omega)\ \rho_2(d\omega) = \int_{\Xi_{\sqrt{3}}} 1\ \ \rho_2(d\omega)$$

$$
\begin{aligned}
&= \nu_2(\{\ell_{(x,y)} \approx (x,y) \in \Omega_2 \mid \text{ "the length of } \ell_{(x,y)}\text{"} \le \sqrt{3}\}) \\
&= \frac{\text{Meas}[\{(x,y) \mid 1/4 \le x^2 + y^2 \le 1\}]}{\pi} = \frac{3}{4}.
\end{aligned}
$$

Conclusion 8.29. Thus, even if there is a custom to regard a natural probability measure (i.e., an invariant measure concerning natural map) as "random", the first answer and the second answer say that

(♯) the uniqueness in (A) of Problem 8.28 is denied.

That is, the invariant measure concerning a natural map does not always mean 'probability'.

For further information see my homepage

Chapter 9

Axiom 2 – causality

Measurement theory (= QL) has the following classification:

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

This is formulated as follows.

(B):

(B_1):

$$
\underset{\text{(=quantum language)}}{\boxed{\text{pure measurement theory}}} := \overbrace{\underset{(cf.\ §2.7)}{\overset{\text{[(pure)Axiom 1]}}{\boxed{\text{pure measurement}}}} + \underset{(cf.\ §9.3)}{\overset{\text{[Axiom 2]}}{\boxed{\text{Causality}}}}}^{\text{a kind of spells (a priori judgment)}} + \overbrace{\underset{(cf.\ §3.1)}{\overset{\text{[quantum linguistic Copenhagen interpretation]}}{\boxed{\text{Linguistic Copenhagen interpretation}}}}}^{\text{manual to use spells}}
$$

(B_2):

$$
\underset{\text{(=quantum language)}}{\boxed{\text{mixed measurement theory}}} := \overbrace{\underset{(cf.\ §8.1)}{\overset{\text{[(mixed)Axiom}^{(m)}\text{ 1]}}{\boxed{\text{mixed measurement}}}} + \underset{(cf.\ §9.3)}{\overset{\text{[Axiom 2]}}{\boxed{\text{Causality}}}}}^{\text{a kind of spells (a priori judgment)}} + \overbrace{\underset{(cf.\ §3.1)}{\overset{\text{[quantum linguistic Copenhagen interpretation]}}{\boxed{\text{Linguistic Copenhagen interpretation}}}}}^{\text{manual to use spells}}
$$

In this chapter, we devote ourselves to the last theme $\underset{(cf.\ §9.3)}{\overset{\text{[Axiom 2]}}{\boxed{\text{Causality}}}}$, which is common to both (B_1) and (B_2).

9.1 The most important unsolved problem – what is causality ?

This section is extracted from ref.[49]. 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 Copenhagen interpretation in a wider sense.

9.1.1 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

Smoke is not located on the place which does not have fire.

However the situation is not so simple as you think. Consider, for example,

This morning I feel good. Is it because that I slept sound yesterday ? or

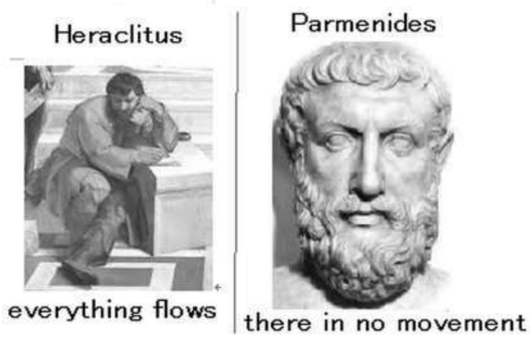
Is it because I go to favorite golf from now on ?

You will find the difficulty in using the word "causality". In daily conversation, it is used in many contexts, mixing up "a cause (past)", "a reason (connotation)", and "the purpose and a motive (future)".

Pioneering research on movement and change may be found in

- Heraclitus(BC.540 -BC.480): "Everything flows."
- Parmenides (born around BC. 515): "There is no movement." (Zeno's teacher)

Although their assertions are not clear, they recognized that "movement and change" were the primarily important keywords in "world description".



[The beginning of World description]

$$=[\textit{The discovery of movement and change}\,] = \begin{cases} \text{Heraclitus} \\ \text{Parmenides} \end{cases}$$

Aristotle (BC384–BC322) further investigated the essence of movement and change, and concluded that

all the movements had the "purpose".

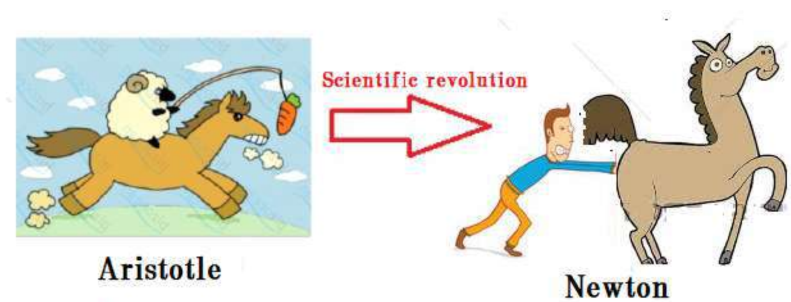
For example, a stone falls because it has the purpose to go downward, and smoke rises because it has the purpose to go upward. Under the influence of Aristotle, "*Purpose*" had remained as a mainstream idea of "Movement" for a long period of 1500 years or more.

We were freed from the spell of "Purpose", only after Galileo, Bacon, Descartes, and Newton et al. discovered the essence of movement and change lies in "Causality".

Revolution from "Purpose" to "Causality"

is the greatest paradigm shift in the history of science. It is not an exaggeration even if we call the shift "*birth of modern science*".

the birth of world description		the birth of modern science
Movement	"purpose" ⟶	Causality
(Heraclitus, Parmenides, Zeno)	Aristotle :(About 1500 years)	(Galileo, Bacon, Descartes, Newton)

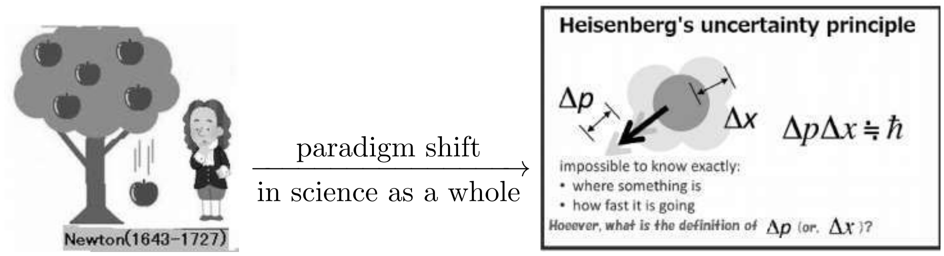


♠**Note 9.1.** I cannot emphasize too much the importance of the discovery of the term: "causality". That is,

(♯) 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.

Also, as mentioned in Preface, note that my purpose of this book is to the propaganda of the follows:



That is, the ‘quantum mechanical paradigm’ that began 100 years ago is no longer finished and will spread to all sciences, not just physics.

9.1.2 Four answers to “what is causality ?” (*cf.* Sec. 10.2.1 in ref. [76])

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 9.1. Problem:

”What is causality?”

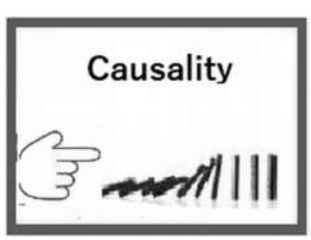
is the most important outstanding problems in modern science.

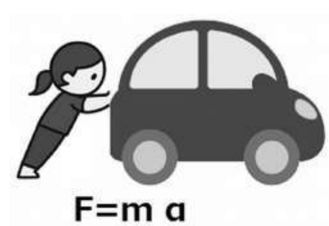
Answer this problem!

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

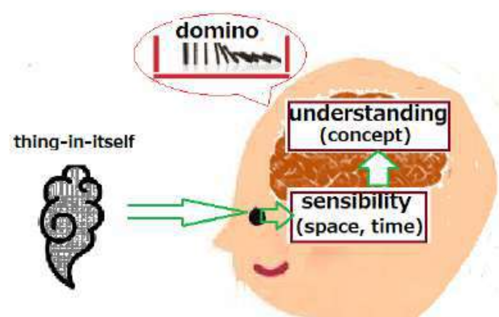
"Newtonian mechanics⟶ Electricity and magnetism⟶ Theory of relativity⟶ $\cdots$ "

is the mainstream of science.

However, there are also other ideas, i.e. three "non-realistic causalities" as follows.

(B) **[Cognitive causality]** David Hume, Immanuel Kant, etc. thought as follows. :

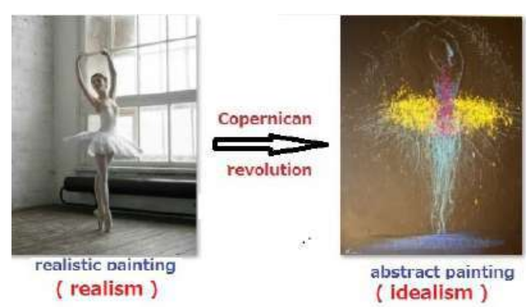
> We can not say that "Causality" actually exists in the world, or that it does not exist in the world. And when we think that "something" in the world is "causality", we should just believe that it has "causality".



Most readers may regard this as "a kind of rhetoric", however, some readers may believe it. It may look like that, because you are looking through the prejudice of "causality." This is Kant's famous "Copernican revolution" (i.e., "Kant was awakened from his dogmatic slumber by Hume's idea and came up with the Copernican revolution"), that is,

"cognition constitutes the world."

which is considered that the cognition circuit of causality is installed in the brain, and when it is stimulated by "something" and reacts, "there is causal relationship."

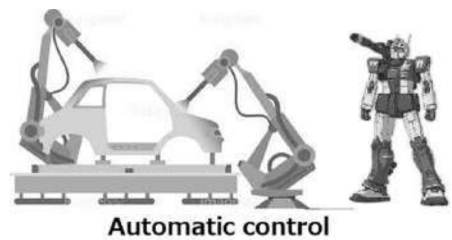


♠**Note 9.2.** About his discovery of "the Copernican revolution", Kant says in his book "Prolegomena" (1783):

(♯) *I freely admit that it was the remembrance of David Hume which, many years ago, first interrupted my dogmatic slumber and gave my investigations in the field of speculative philosophy a completely different direction.*

Readers may ask, "Why did Kant, an honest and humble man, make such an exaggerated statement?" It is a matter of course that Kant had great confidence such that it was the greatest discovery in the history of philosophy. I agree to his opinion. □

(C) **[Causality in applied mathematics (Dynamical system theory)]**



Since dynamical system theory (= statistics) 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

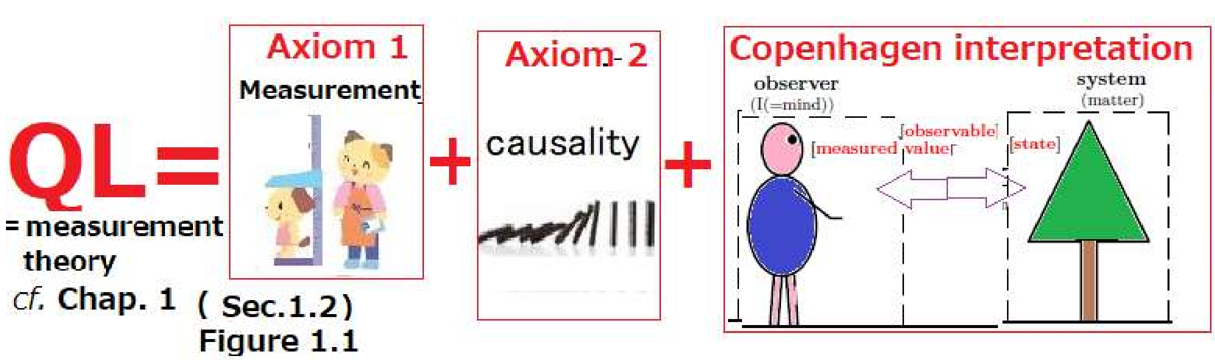
$$\begin{cases} \frac{d\omega_1}{dt}(t) = v_1(\omega_1(t), \omega_2(t), \ldots, \omega_n(t), t) \\ \frac{d\omega_2}{dt}(t) = v_2(\omega_1(t), \omega_2(t), \ldots, \omega_n(t), t) \\ \cdots\cdots \\ \frac{d\omega_n}{dt}(t) = v_n(\omega_1(t), \omega_2(t), \ldots, \omega_n(t), t) \end{cases} \tag{9.1}$$

and, we think that

(♯) 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 worldview, it is quite useful. In this sense, I think that (C) should be evaluated more.

(D) **[Linguistic causal relationship (Measurement Theory)]**



The causal relationship of measurement theory is decided by the Axiom 2 (causality; Sec. 1.1) of Chap. 1. If I say in detail,:

For further information see my homepage

- Although measurement theory (= quantum language) consists of the two Axioms 1 and 2, it is the Axiom 2 that is concerned with causal relationship. When describing something in quantum language and using Axiom 2 (causality; Sec. 1.1) , we think that it has causality.

Summary 9.2. 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,

In science, causality is just as mentioned in the above (D).

This is my answer to "What is causality?".

9.2 Causality in QL – Mathematical preparation

9.2.1 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

$$
\begin{array}{ccccccccc}
\underset{C^*\text{-pure state}}{\mathfrak{S}^p(\mathcal{A}^*)} & \subset & \underset{C^*\text{-mixed state}}{\mathfrak{S}^m(\mathcal{A}^*)} & \subset & \mathcal{A}^* & & & & \\
 & & & & \uparrow \text{dual} & & & & \\
 & & & & \boxed{\mathcal{A}} & \underset{\substack{\text{subalgebra}\\ \text{weak-closure}}}{\xrightarrow{\subseteq}} & \boxed{\overline{\mathcal{A}}} & \underset{\text{subalgebra}}{\xrightarrow{\subseteq}} & \boxed{B(H)} \\
 & & & & & & \downarrow \text{pre-dual} & & \\
 & & & & & & \underset{W^*\text{-mixed state}}{\overline{\mathfrak{S}}^m(\overline{\mathcal{A}}_*)} \subset \overline{\mathcal{A}}_* & &
\end{array}
\tag{9.2}
$$

Remark 9.3. $[\overline{\mathcal{A}}_* \subseteq \mathcal{A}^*]$: 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

$$\left|_{\overline{\mathcal{A}}_*}\Big(\rho, F\Big)_{\overline{\mathcal{A}}}\right| \leq C\|F\|_{B(H)} = C\|F\|_{\mathcal{A}} \tag{9.3}$$

Thus, we can consider that $\rho \in \mathcal{A}^*$. That is, in the sense of (9.3), 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,

$${}_{\overline{\mathcal{A}}_*}\Big(\rho, F\Big)_{\overline{\mathcal{A}}} = {}_{\mathcal{A}^*}\Big(\widehat{\rho}, F\Big)_{\mathcal{A}} \qquad (\forall F \in \mathcal{A}(\subseteq \overline{\mathcal{A}})). \tag{9.4}$$

Definition 9.4. [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

$$\text{(a)} \quad {}_{(\overline{\mathcal{A}}_1)_*}\Big(\rho_1, \Phi_{1,2}F_2\Big)_{\overline{\mathcal{A}}_1} = {}_{(\overline{\mathcal{A}}_2)_*}\Big((\Phi_{1,2})_*\rho_1, F_2\Big)_{\overline{\mathcal{A}}_2} \qquad (\forall \rho_1 \in (\overline{\mathcal{A}}_1)_*, \forall F_2 \in \overline{\mathcal{A}}_2) \tag{9.5}$$

$$\text{(b)} \quad (\Phi_{1,2})_*(\overline{\mathfrak{S}}^m((\overline{\mathcal{A}}_1)_*)) \subseteq \overline{\mathfrak{S}}^m((\overline{\mathcal{A}}_2)_*) \tag{9.6}$$

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

$$\text{(a)} \quad {}_{(\overline{\mathcal{A}}_1)_*}\Big(\rho_1, \Phi_{1,2}F_2\Big)_{\overline{\mathcal{A}}_1} = {}_{\mathcal{A}^*_2}\Big(\Phi^*_{1,2}\widehat{\rho}_1, F_2\Big)_{\mathcal{A}_2} \qquad (\forall \rho_1 = \widehat{\rho}_1 \in (\overline{\mathcal{A}}_1)_*(\subseteq \mathcal{A}^*_1), \forall F_2 \in \mathcal{A}_2) \tag{9.7}$$

$$\text{(b)} \quad (\Phi_{1,2})^*(\mathfrak{S}^p(\mathcal{A}^*_1)) \subseteq \mathfrak{S}^m(\mathcal{A}^*_2) \tag{9.8}$$

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

$$(\Phi_{1,2})^*(\mathfrak{S}^p(\mathcal{A}^*_1)). \subseteq \mathfrak{S}^p(\mathcal{A}^*_2). \tag{9.9}$$

For further information see my homepage

♠**Note 9.3.** [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) – (iv):

(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

$$_{L^1(\Omega_1)}\Big(\rho_1, \Phi_{1,2} F_2\Big)_{L^\infty(\Omega_1)} = {}_{\mathcal{M}(\Omega_2)}\Big(\Phi^*_{1,2} \widehat{\rho}_1, F_2\Big)_{C_0(\Omega_2)}$$
$$(\forall \rho_1 = \widehat{\rho}_1 \in \mathcal{M}(\Omega_1), \forall F_2 \in C_0(\Omega_2))$$

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

$$[\Phi_{1,2} f_2](\omega_1) = f_2(\phi_{1,2}(\omega_1)) \quad (\forall f_2 \in C(\Omega_2), \forall \omega_1 \in \Omega_1). \tag{9.10}$$

This $\phi_{1,2} : \Omega_1 \to \Omega_2$ is called a *deterministic causal map*. Here, it is clear that

$$\Omega_1 \approx \mathfrak{S}^p(C_0(\Omega_1)^*) \ni \delta_{\omega_1} \underset{\Phi^*_{12}}{\longrightarrow} \delta_{\phi_{12}(\omega_1)} \in \mathfrak{S}^p(C_0(\Omega_2)^*) \approx \Omega_2.$$

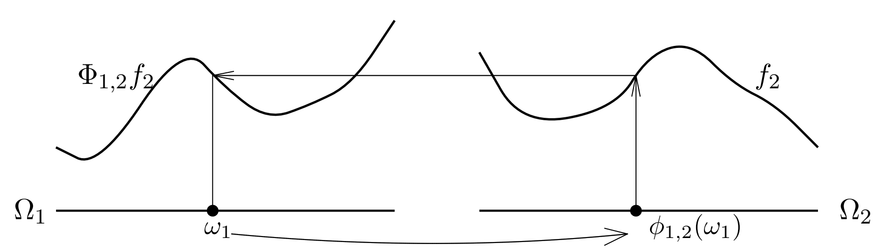


Figure 9.1: Deterministic causal map $\phi_{1,2}$ and deterministic causal operator $\Phi_{1,2}$

For further information see my homepage

Theorem 9.5. [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 (9.10) 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.* ref. [114]) $[\Phi_{1,2}]_*(\overline{\rho}_1) = d\mu_2/d\nu_2$. That is because

$$D_2 \in \mathcal{B}_{\Omega_2}, \ \ \nu_2(D_2) = 0 \quad \Longrightarrow \quad \nu_1(\phi_{1,2}^{-1}(D_2)) = 0 \quad \Longrightarrow \quad \mu_2(D_2) = 0. \tag{9.11}$$

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 9.6. 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. □

9.2.2 Simple example – Finite causal operator is represented by matrix

Example 9.7. [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 (9.10), 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 9.8. [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

$$\rho_1 = a_1 \delta_{\omega_1^1} + a_2 \delta_{\omega_1^2} + a_3 \delta_{\omega_1^3} \quad (0 \leqq a_1, a_2, a_3 \leqq 1, \quad a_1 + a_2 + a_3 = 1).$$

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

$$\Phi^*_{1,2}(\rho_1) = \beta_1 \delta_{\omega_2^1} + \beta_2 \delta_{\omega_2^1} = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix},$$

$$\rho_1 = \alpha_1 \delta_{\omega_1^1} + \alpha_2 \delta_{\omega_1^2} + \alpha_3 \delta_{\omega_1^3} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix},$$

we 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,

$$C_0(\Omega_1) \ni f_1 \underset{\text{(identification)}}{\longleftrightarrow} \begin{bmatrix} f_1(\omega_1^1) \\ f_1(\omega_1^2) \\ f_1(\omega_1^3) \end{bmatrix} \in \mathbb{C}^3,$$

$$C_0(\Omega_2) \ni f_2 \underset{\text{(identification)}}{\longleftrightarrow} \begin{bmatrix} f_2(\omega_2^1) \\ f_2(\omega_2^2) \end{bmatrix} \in \mathbb{C}^2.$$

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 9.9. [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

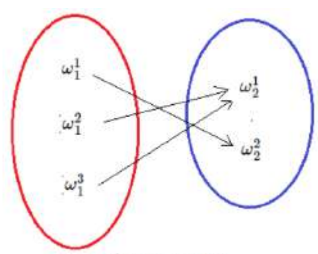
$$\Phi^*_{1,2}(\rho_1) = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}.$$

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

$$\phi_{1,2}(\omega_1^1) = \omega_2^2, \quad \phi_{1,2}(\omega_1^2) = \omega_2^1, \quad \phi_{1,2}(\omega_1^3) = \omega_2^1.$$



$\phi_{1,2} : \Omega_1 \to \Omega_2$

9.2.3 Sequential causal operator – A chain of causalities

Let $(T, \leq)$ be a finite tree, 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: $\pi(t_3) = \pi(t_4) = t_2$, $\pi(t_2) = \pi(t_5) = t_1$, $\pi(t_1) = \pi(t_6) = \pi(t_7) = t_0$

 For further information see my homepage

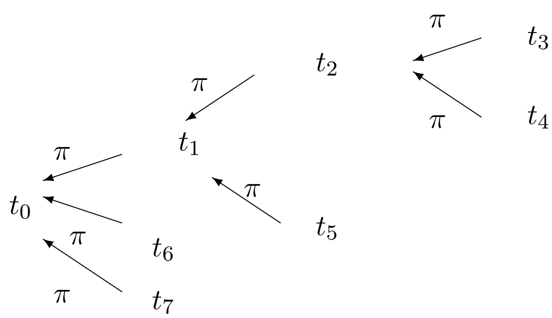


Figure 9.2: Tree: $(T = \{t_0, t_1, ..., t_7\}, \pi : T \setminus \{t_0\} \to T)$

Definition 9.10. [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.

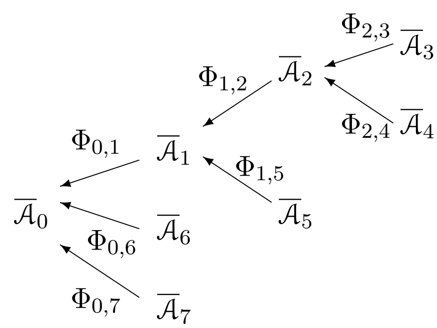
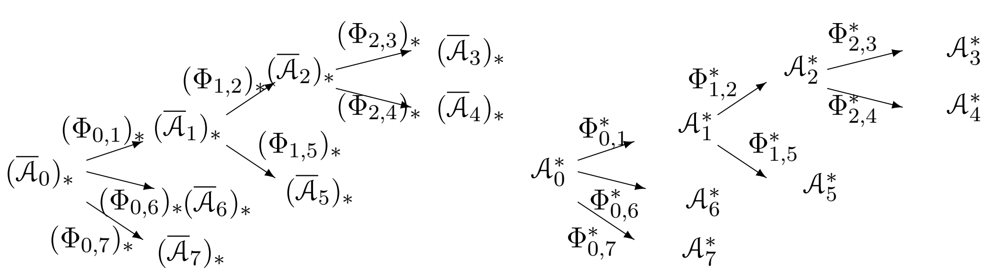


Figure 9.3: Heisenberg picture(sequential causal operator)

Definition 9.11. (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

For further information see my homepage

Figure 9.4: Schrödinger picture (dual sequential causal operator)

Remark 9.12. [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 Copenhagen interpretation says "state does not move", and thus, we consider that

- $\begin{cases} \text{the Heisenberg picture is formal} \\ \text{the Schrödinger picture is makeshift} \end{cases}$

For further information see my homepage

9.3 Axiom 2 – Smoke is not located on the place which does not have fire

In this section, propose Axiom 2 (Causality), and thus all of QL are presented as follows.

(♯)

(♯₁): [pure measurement theory] (=quantum language)

:= [pure measurement] ([(pure)Axiom 1]; (*cf.* §2.7)) + [(deterministic) Causality] ([Axiom 2]; (*cf.* §9.3)) + [Linguistic Copenhagen interpretation] ([quantum linguistic Copenhagen interpretation]; (*cf.* §3.1))

— pure measurement + Causality: a kind of spells (a priori judgment); Linguistic Copenhagen interpretation: manual to use spells

(♯₂): [mixed measurement theory] (=quantum language)

:= [mixed measurement] ([(mixed)Axiom$^{(m)}$ 1]; (*cf.* §8.1)) + [Causality] ([Axiom 2]; (*cf.* §9.3)) + [Linguistic Copenhagen interpretation] ([quantum linguistic Copenhagen interpretation]; (*cf.* §3.1))

— mixed measurement + Causality: a kind of spells (a priori judgment); Linguistic Copenhagen interpretation: manual to use spells

9.3.1 Axiom 2 (A chain of causal relations)

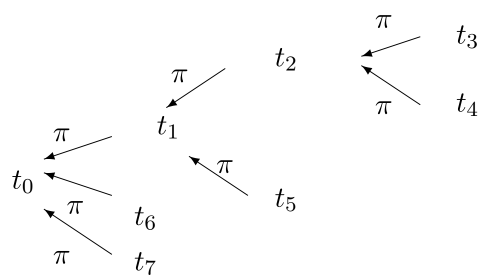
Now we can propose Axiom 2 of causality, which is a measurement theoretical representation of the maxim “Smoke is not located on the place which does not have fire”:

(C): Axiom 2 (A chain of causalities)

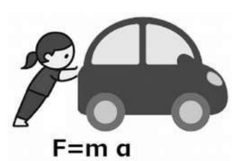
For each $t(\in T=$"tree")), consider the basic structure:

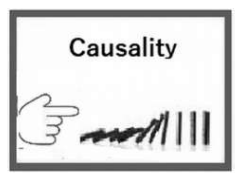
$$[\mathcal{A}_t \subseteq \overline{\mathcal{A}}_t \subseteq B(H_t)].$$

Then, the *chain of causalities* is represented by a *sequential causal operator* $\{\Phi_{t_1,t_2} : \overline{\mathcal{A}}_{t_2} \to \overline{\mathcal{A}}_{t_1}\}_{(t_1,t_2)\in T^2_{\leqq}}$. Also, when Φ_{t_1,t_2} is always deterministic, it is called a sequential deterministic causal operator.



Later Figure 9.2: Tree: $(T = \{t_0, t_1, ..., t_7\}, \pi : T \setminus \{t_0\} \to T)$



♠**Note 9.4.** Note that there is no mention of 'time' yet. Time is discussed in sections 9.7 and 9.8. Therefore, QL (i.e., the linguistic Copenhagen interpretation) asserts "Causality precedes time". Therefore, if the concept of time is extended to include parallel time and tree-typed time, these are also considered to be a type of time.

♠**Note 9.5.** 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 (($\sharp_1$)–($\sharp_4$)) moved the world.

9.3.2 Sequential causal operator – State equation, etc.

In what follows, we shall exercise the chain of causality in quantum language.

Example 9.13. [State equation] Let $T = \mathbb{R}$ be a tree which represents the time axis. 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

$$\begin{cases} \frac{d\omega_1}{dt}(t) = v_1(\omega_1(t), \omega_2(t), \ldots, \omega_n(t), t) \\ \frac{d\omega_2}{dt}(t) = v_2(\omega_1(t), \omega_2(t), \ldots, \omega_n(t), t) \\ \cdots\cdots \\ \frac{d\omega_n}{dt}(t) = v_n(\omega_1(t), \omega_2(t), \ldots, \omega_n(t), t). \end{cases} \tag{9.12}$$

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 (9.12). 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 9.14. [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 $\pi(t) = t - 1$ $(\forall t = 1, 2, \ldots)$. 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 9.5, 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 9.6.** In order to analyze multiple Markov processes and time-lag processes, such ideas in Example 9.14 are needed.

9.4 Kinetic equation in classical and quantum mechanics

9.4.1 Hamiltonian (Time-invariant system)

In this section, we consider the simplest kinetic equation in classical and quantum systems. Consider the state space Ω such that $\Omega = \mathbb{R}^2$, that is,

$$\mathbb{R}^2 = \mathbb{R}_q \times \mathbb{R}_p = \{(q,p) = (\text{position} , \text{momentum}) \mid q,p \in \mathbb{R}\} \tag{9.13}$$

Hamiltonian $\mathcal{H}(q,p)$ defined by the total energy takes the form of

$$\begin{aligned}&[\text{Hamiltonian } (= \mathcal{H}(q,p))] \\ =&[\text{kinetic energy}(= \frac{p^2}{2m})] + [\text{potential energy}(= V(q))]\end{aligned} \tag{9.14}$$

for a typical case of one particle with mass $=m$.

9.4.2 Newtonian equation (=Hamilton's canonical equation)

Concerning Hamiltonian $\mathcal{H}(q,p)$, *Hamilton's canonical equation* is defined by

$$\text{Hamilton's canonical equation} = \begin{cases} \frac{dp}{dt} = -\frac{\mathcal{H}(q,p)}{\partial q} \\ \\ \frac{dq}{dt} = \frac{\mathcal{H}(q,p)}{\partial p} \end{cases} \tag{9.15}$$

And thus, in the case of (9.14), we get

$$\text{Hamilton's canonical equation} = \begin{cases} \frac{dp}{dt} = -\frac{\mathcal{H}(q,p)}{\partial q} = -\frac{\partial V(q,p)}{\partial q} \\ \\ \frac{dq}{dt} = \frac{\partial \mathcal{H}(q,p)}{\partial p} = \frac{p}{m} \end{cases} \tag{9.16}$$

which is the same as Newtonian equation. That is,

$$m\frac{d^2q}{dt^2} = [\text{Mass}] \times [\text{Acceleration}] = -\frac{\partial V(q,p)}{\partial q}(= \text{Force})$$

Now, let us describe the above (9.16) in quantum language. For each $t \in T = \mathbb{R}$, define the state space Ω_t by

$$\Omega_t = \Omega = \mathbb{R}^2 = \mathbb{R}_q \times \mathbb{R}_p = \{(q,p) = (\text{position} , \text{momentum}) \mid q,p \in \mathbb{R}\} \tag{9.17}$$

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 (9.16) is defined by

$$\Omega_{t_1} \ni \omega_{t_1} \mapsto \phi_{t_1,t_2}(\omega_{t_1}) = \omega_{t_2} \in \Omega_{t_2}. \tag{9.18}$$

Since (9.18) 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

$$[\Phi_{t_1,t_2}(f_{t_2})](\omega_{t_1}) = f_{t_2}(\phi_{t_1,t_2}(\omega_{t_1})) \qquad (\forall f_{t_2} \in L^\infty(\Omega_2), \forall \omega_{t_1} \in \Omega_{t_1}, t_1 \leq t_2). \tag{9.19}$$

9.4.3 Schrödinger equation (quantized Hamiltonian)

The quantization is the following procedure:

$$\textbf{quantization}^1 \begin{cases} \text{total energy} E & \xrightarrow[\text{quantization}]{} \frac{\hbar\sqrt{-1}\partial}{\partial t} \\ \\ \text{momentum } p & \xrightarrow[\text{quantization}]{} \frac{\hbar\partial}{\sqrt{-1}\partial q} \\ \\ \text{position } q & \xrightarrow[\text{quantization}]{} q \end{cases} \tag{9.20}$$

Substituting the quantization (9.20) to the classical Hamiltonian:

$$E = \mathcal{H}(q,p) = \frac{p^2}{2m} + V(q)$$

we get

$$\hbar\sqrt{-1}\frac{\partial}{\partial t} = \mathcal{H}(q, \frac{\hbar}{\sqrt{-1}}\frac{\partial}{\partial q}) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial q^2} + V(q) \tag{9.21}$$

And therefore, we get the *Schrödinger equation*:

$$\hbar\sqrt{-1}\frac{\partial u(t,q)}{\partial t} = \mathcal{H}(q, \frac{\hbar}{\sqrt{-1}}\frac{\partial}{\partial q})u(t,q) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial q^2}u(t,q) + V(q)u(t,q). \tag{9.22}$$

Putting $u(t,\cdot) = u_t \in L^2(\mathbb{R})$ $(\forall t \in T = \mathcal{R})$, we denote the Schrödinger equation (9.22) by

$$u_t = \frac{1}{\hbar\sqrt{-1}}\mathcal{H}u_t.$$

Solving this formally, we see

$$u_t = e^{\frac{\mathcal{H}}{\hbar\sqrt{-1}}t}u_0 \quad (\text{Thus, the state representation is } |u_t\rangle\langle u_t| = |e^{\frac{\mathcal{H}}{\hbar\sqrt{-1}}t}u_0\rangle\langle e^{\frac{\mathcal{H}}{\hbar\sqrt{-1}}t}u_0| \;) \tag{9.23}$$

[1]Learning the (9.20) by rote, we can derive Schrödinger equation (9.22). However, the meaning of "quantization" is not clear.

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 $\{\Phi^*_{t_1,t_2} : \mathcal{T}r(H_{t_1}) \to \mathcal{T}r(H_{t_2})\}_{(t_1,t_2)\in T^2_{\leq}}$ is defined by

$$\Phi^*_{t_1,t_2}(\rho) = e^{\frac{\mathcal{H}}{\hbar\sqrt{-1}}(t_2-t_1)}\rho e^{\frac{-\mathcal{H}}{\hbar\sqrt{-1}}(t_2-t_1)} \quad (\forall \rho \in \mathcal{T}r(H_{t_1}) = (B(H_{t_1}))_* = \mathcal{C}(H_{t_1})^*). \tag{9.24}$$

And therefore, the sequential causal operator $\{\Phi_{t_1,t_2} : B(H_{t_2}) \to B(H_{t_1})\}_{(t_1,t_2)\in T^2_{\leq}}$ is defined by

$$\Phi_{t_1,t_2}(A) = e^{\frac{-\mathcal{H}}{\hbar\sqrt{-1}}(t_2-t_1)} A e^{\frac{\mathcal{H}}{\hbar\sqrt{-1}}(t_2-t_1)} \quad (\forall A \in B(H_{t_2})). \tag{9.25}$$

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 $\{\Phi_{t_1,t_2} : B(H_{t_2}) \to B(H_{t_1})\}_{(t_1,t_2)\in T^2_{\leq}}$ is deterministic. Since we deal with the time-invariant system, putting $t = t_2 - t_1$, we see that (9.25) is equal to

$$A_t = \Phi_t(A_0) = e^{\frac{-\mathcal{H}}{\hbar\sqrt{-1}}t} A_0 e^{\frac{\mathcal{H}}{\hbar\sqrt{-1}}t}. \tag{9.26}$$

And thus, we get the differential equation:

$$\begin{aligned} \frac{dA_t}{dt} &= \frac{-\mathcal{H}}{\hbar\sqrt{-1}} e^{\frac{-\mathcal{H}}{\hbar\sqrt{-1}}t} A_0 e^{\frac{\mathcal{H}}{\hbar\sqrt{-1}}t} + \frac{-\mathcal{H}}{\hbar\sqrt{-1}} e^{\frac{-\mathcal{H}}{\hbar\sqrt{-1}}t} A_0 e^{\frac{\mathcal{H}}{\hbar\sqrt{-1}}t} \frac{\mathcal{H}}{\hbar\sqrt{-1}} \\ &= \frac{-\mathcal{H}}{\hbar\sqrt{-1}} A_t + A_t \frac{\mathcal{H}}{\hbar\sqrt{-1}} = \frac{1}{\hbar\sqrt{-1}}\Big(A_t\mathcal{H} - \mathcal{H}A_t\Big) \end{aligned} \tag{9.27}$$

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.

Schrödinger
(1887-1961)

Heisenberg
(1901-1976)

9.5 Exercise: Solving 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

$$V_0(q) = \begin{cases} 0 & (0 \le q \le 2) \\ \infty & (\text{ otherwise }) \end{cases}$$

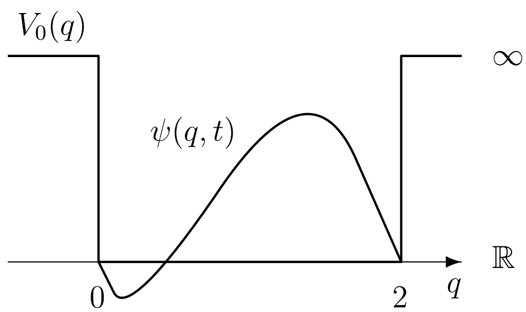


Figure 9.5: Particle in a box

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

$$\phi(q,t) = T(t)X(q) = C_3 \exp(\frac{in^2\pi^2\hbar t}{8m}) \sin(n\pi q/2) \qquad (n = 1, 2, ...).$$

For further information see my homepage

Assume the initial condition:

$$\psi(q,0) = c_1 \sin(\pi q/2) + c_2 \sin(2\pi q/2) + c_3 \sin(3\pi q/2) + \cdots .$$

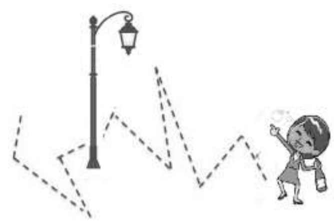
where $\int_{\mathbb{R}} |\psi(q,0)|^2 dq = 1$. Then we see

$$\psi(q,t) = c_1 \exp(\frac{i\pi^2 \hbar t}{8m}) \sin(\pi q/2) + c_2 \exp(\frac{i4\pi^2 \hbar t}{8m}) \sin(2\pi q/2) + c_3 \exp(\frac{i9\pi^2 \hbar t}{8m}) \sin(3\pi q/2) + \cdots .$$

And thus, we have the time evolution of the state by

$$\rho_t = |\psi(\cdot,t)\rangle\langle\psi(\cdot,t)| \qquad (\in \mathfrak{S}^p(Tr(H)) \subseteq B(H)) \qquad (\forall t \geq 0)$$

9.6 Random walk and quantum decoherence



9.6.1 Diffusion process

Example 9.15. [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 $\pi(t) = t - 1$ $(t = 1, 2, ...)$. For each $t(\in T)$, a state space Ω_t is define by $\Omega_t = \mathbb{Z}$. Then, we have the sequential causal operator $\{\Phi_{\pi(t),t}(= \Phi) : L^\infty(\Omega_t) \to L^\infty(\Omega_{\pi(t)})\}_{t \in T \setminus \{0\}}$.

For further information see my homepage

9.6.2 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,

$$P_n \text{ is a projection (i.e., } P_n = (P_n)^2\text{), and } \quad \sum_{n=1}^{\infty} P_n = I.$$

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.
Hence

$$(\Psi_{\mathbb{P}})_*(\mathcal{T}r_{+1}^p(H)) \subseteq \mathcal{T}r_{+1}(H).$$

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 9.16. [Quantum decoherence] For the relation between quantum decoherence and quantum Zeno effect, see § 10.4. Also, for the relation between quantum decoherence and Schrödinger's cat, see § 10.5. In this note, we assume that the non-deterministic causal operator belongs to the mixed measurement theory. Thus, we consider quantum language (= measurement theory) is classified as follows.

(A) measurement theory (=quantum language)
- pure type (A_1)
  - classical system : Fisher statistics
  - quantum system : usual quantum mechanics
- mixed type (A_2)
  - classical system : including Bayesian statistics and Kalman filter
  - quantum system : quantum decoherence

For further information see my homepage

9.7 Leibniz-Clarke Correspondence: What is space-time?

This section is published in the following:

- ref. [70]: S. Ishikawa; *Leibniz-Clarke correspondence, Brain in a vat, Five-minute hypothesis, McTaggart's paradox, etc. are clarified in quantum language*
Open Journal of philosophy, Vol. 8, No.5 , 466-480, 2018,
(https://www.scirp.org/Journal/PaperInformation.aspx?PaperID=87862)
- ref. [71]; S. Ishikawa; *Leibniz-Clarke correspondence, Brain in a vat, Five-minute hypothesis, McTaggart's paradox, etc. are clarified in quantum language*; [Revised version] ; Keio Research report; 2018; KSTS/RR-18/001, 1-15 (https://philpapers.org/rec/ISHLCB)
(http://www.math.keio.ac.jp/academic/research_pdf/report/2018/18001.pdf)

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 the quantum linguistic answer to these problems. As seen later, our answer is similar to Leibniz's relationalism concerning space-time. In this sense, we consider that Leibniz is one of the discoverers of the linguistic Copenhagen interpretation

9.7.1 "What is space?" and "What is time?")

Note that

"space" and "time" are not written in Axioms 1 and 2 (in QL);

We must therefore, like God, make "space" and "time" as follows.

9.7.1.1 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

$$
\text{(B)} \quad \begin{cases} (\text{B}_1)\text{: a classical particle in the one dimensional space } \mathbb{R}_q \\ \\ (\text{B}_2)\text{: a quantum particle in the one dimensional space } \mathbb{R}_q \end{cases}
$$

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:

$$
(\text{C}_1) \qquad [C_0(\mathbb{R}_q \times \mathbb{R}_p) \subseteq L^\infty(\mathbb{R}_q \times \mathbb{R}_p) \subseteq B(L^2(\mathbb{R}_q \times \mathbb{R}_p))]
$$

Also, concerning quantum system, we have the quantum basic structure:

(C$_2$) $[\mathcal{C}(L^2(\mathbb{R}_q) \subseteq B(L^2(\mathbb{R}_q) \subseteq B(L^2(\mathbb{R}_q)]$

Summing up, we have the basic structure

$$\text{(C)}\;\; [\mathcal{A} \subseteq \overline{\mathcal{A}} \subseteq B(H)] \begin{cases} \text{(C}_1\text{): classical } [C_0(\mathbb{R}_q \times \mathbb{R}_p) \subseteq L^\infty(\mathbb{R}_q \times \mathbb{R}_p) \subseteq B(L^2(\mathbb{R}_q \times \mathbb{R}_p)] \\ \\ \text{(C}_2\text{): quantum } [\mathcal{C}(L^2(\mathbb{R}_q) \subseteq B(L^2(\mathbb{R}_q) \subseteq B(L^2(\mathbb{R}_q)] \end{cases}$$

Since we always start from a basic structure in quantum language, we consider that

$$\begin{aligned} &\text{How to describe “space” in quantum language} \\ \Leftrightarrow\; &\text{How to describe [(A):space] by [(C):basic structure]} \end{aligned} \tag{9.28}$$

This is done in the following steps.

Assertion 9.17. [The linguistic Copenhagen interpretation concerning ”space”]
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)

For further information see my homepage

9.7.1.2 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.

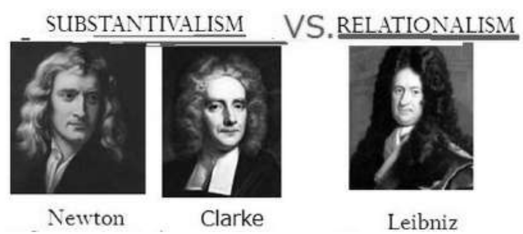
Assertion 9.18. [The linguistic Copenhagen interpretation concerning ”time”]
How to describe “time” in quantum language

(F_1) Let T be a tree. 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”.

9.7.2 Leibniz-Clarke Correspondence



The above argument urges us to recall Leibniz-Clarke Correspondence (1715–1716: *cf.* ref. [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”, i.e., a point in space is regarded as a parameter ($\approx$ state).

(H_2) Time is an order of occurring in succession which changes one after another.

Therefore, if “ thing ” does not exits, the space-time does not exist.

Therefore, I regard this correspondence as

$$\underset{\text{(realistic view)}}{\boxed{\text{Newton } (\approx \text{Clarke})}} \quad \underset{\text{vs.}}{\longleftrightarrow} \quad \underset{\text{(linguistic view)}}{\boxed{\text{Leibniz}}}$$

which should be compared to

Einstein (realistic view)	$\longleftrightarrow$ vs.	Bohr (linguistic view)

Again, we emphasize that Leibniz's relationalism in Leibniz-Clarke correspondence is clarified in quantum language, and it should be regarded as one of the most important parts of the linguistic Copenhagen interpretation of quantum theory.

♠**Note 9.7.** 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 0.1 in Preface):

①: Newton, Clarke (realistic world view) ··· (space-time in physics) **realistic space-time** "What is space-time?" (successors: Einstein, etc.)

②: Leibniz (linguistic world view) ··· (space-time in measurement theory) **linguistic space-time** "How should space-time be represented?" (i.e., spectrum, tree)

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

($\sharp_1$) Newton proposed the scientific language called Newtonian mechanics,
on the other hand,
Leibniz could not propose a scientific language

After all, we conclude that

($\sharp_2$) **the philosopher's failure is that they did not propose a language**.

Talking cynically, we say that

($\sharp_3$) Philosophers continued investigating "linguistic Copenhagen interpretation" (="how to use Axioms 1 and 2") without language (i.e., Axiom 1(measurement:§2.7) and Axiom 2(causality:§9.3)).

♠**Note 9.8.** I want to believe that "realistic" vs. "linguistic" is always hidden behind the great disputes in the history of the world view (*cf.* ref. [76]). That is,

realistic world view	$\longleftrightarrow$ vs.	linguistic world view (idealistic)

For example,

For further information see my homepage

Table 10.1: Philosophical controversy that has been ongoing for 2,500 years [(monistic) realistic worldview] vs. [(dualistic) linguistic/idealistic world view]

dispute ＼ [R] vs. [L]	Realistic worldview (monism, realism, no measurement)	Idealistic worldview (dualism, idealism, measurement)
ⓐ: motion	Hērakleitos	Parmenides
ⓑ:Ancient Greece	Aristotle	Plato
ⓒ: Problem of universals	Ockham	Anselmus
ⓓ: space-time	Newton (Clarke)	Leibniz
ⓔ: quantum theory	Einstein	Bohr
ⓕ:philosophy of science	Carnap	Quine
ⓖ: fuzzy sets	Kalman	Zadeh

For a detailed discussion, see ref. [76].

9.8 Zeno's paradox and Motion function method (in classical system)

Zeno's paradox is humanity's oldest unsolved scientific problem. Thus, numerous challenges have therefore been made to solve Zeno's paradox. For example,

(i) solving it with Newtonian mechanics.

(ii) Solving it in the framework of relativity.

(iii) solving it in the framework of quantum mechanics, etc.

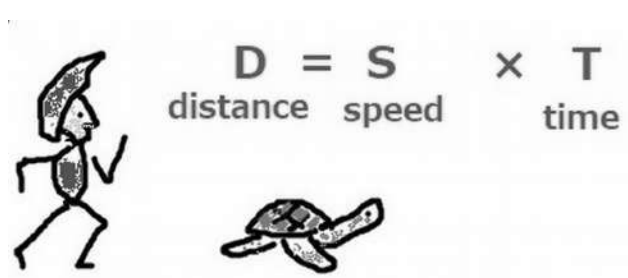
Why were these challenges not generally approved?

The reason, I think, is that Newtonian mechanics, relativity, quantum mechanics are not a theory of everyday science. And thus, I would like to consider that

(♯) to solve Zeno's paradox ⇔ to discover a theory of everyday science (i.e., classical QL), and clarify Zeno's paradox in classical QL

Thus, let us prove Zeno's paradox in classical QL as follows.

9.8.1 Zeno's paradox (e.g., flying arrow)



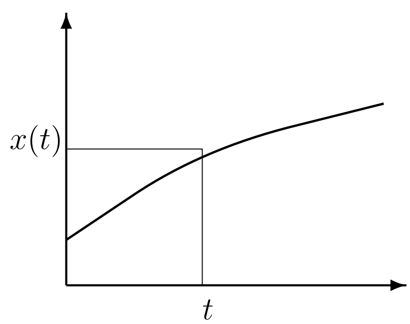
If we obey the motion function method, we can easily solve Zeno's paradoxes (e.g., Flying arrow) as follows.

Answer 9.19. (=Answer 2.11 in ref.[76]) Under the motion function method, we discuss "Flying arrow" as follows.

- Consider the motion function $x(t)$, that is, for each time t, the position $x(t)$ of the arrow is corresponded. It is obvious that

  ($\sharp$) "for each time t, the position $x(t)$ of the arrow is corresponded" does not imply that the motion function $x(t)$ is a constant function.

  Therefore, the arrow is not necessarily at rest.



□

9.8.2 The Schrödinger picture and the Heisenberg picture are equivalent in the classical system

(The general case (the Schrödinger picture and the Heisenberg picture are equivalent) will be discussed in section 10.1.)

According to Leibniz, "time" is just a "parameter" that can be conveniently created. Let's introduce "parallel time" and "Series time. Here, parallel time represents the time lapse of a dice throw or the law of large numbers, etc.(*cf.* ref. [77]). Let $\Omega(\subseteq \mathbb{R}^N)$ (where N is assumed to be sufficiently large natural number) be a compact space, and let $\mathcal{B}(\in \mathcal{P}(\Omega))$ be the Borel field of Ω. $(\Omega, \mathcal{B}(\Omega), \nu)$ be measure space such that $\nu(\Omega) = 1$. Assume that $\nu(D) > 0$ for all open set $D(\subseteq \Omega)$ such that $D \neq \emptyset$. Thus, we consider the W^*-algebraic basic structure $[C(\Omega) \subseteq L^\infty(\Omega, \nu) \subseteq B(L^\infty(\Omega, \nu))]$. Consider a classical dynamical system (Ω, ϕ_{t_1,t_2}). Assume that $t_1, t_2 \in T = [0, 1]$ such that $0 \leq t_1 \leq t_2 \leq 1$, a map $\phi_{t_1,t_2}(\cdot) : \Omega \to \Omega$ is bi-continuous and satisfies the following condition:

($\sharp_1$) $\lim_{t_2 \to t_1} \phi_{t_1,t_2}(\omega) = \omega \quad (\omega \in \Omega)$

($\sharp_2$) $[\phi_{t_2,t_3} \circ \phi_{t_1,t_2}](\omega) = \phi_{t_2,t_3}(\phi_{t_1,t_2}(\omega)) = \phi_{t_1,t_3}(\omega) \quad (\omega \in \Omega, 0 \leq t_1 \leq t_2 \leq t_3 \leq 1)$

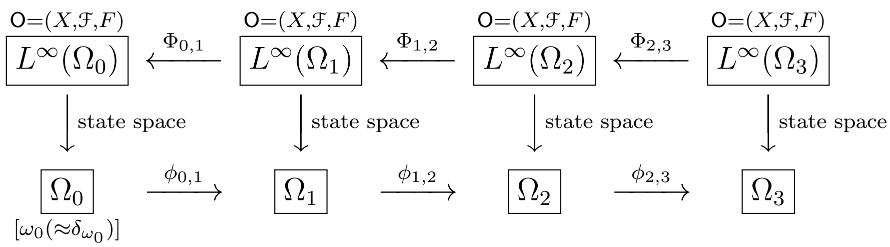
As mentioned before

For further information see my homepage

(K) there exists a homomorphism $\Phi_{t_1,t_2} : L^\infty(\Omega) \to L^\infty(\Omega)$ such that

$$[\Phi_{t_1,t_2}(g_{t_2})](\omega_{t_1}) = g_{t_2}(\phi_{t_1,t_2}(\omega_{t_1})) \qquad (\forall \omega_{t_1} \in \Omega, \forall g_{t_2} \in L^\infty(\Omega)),$$

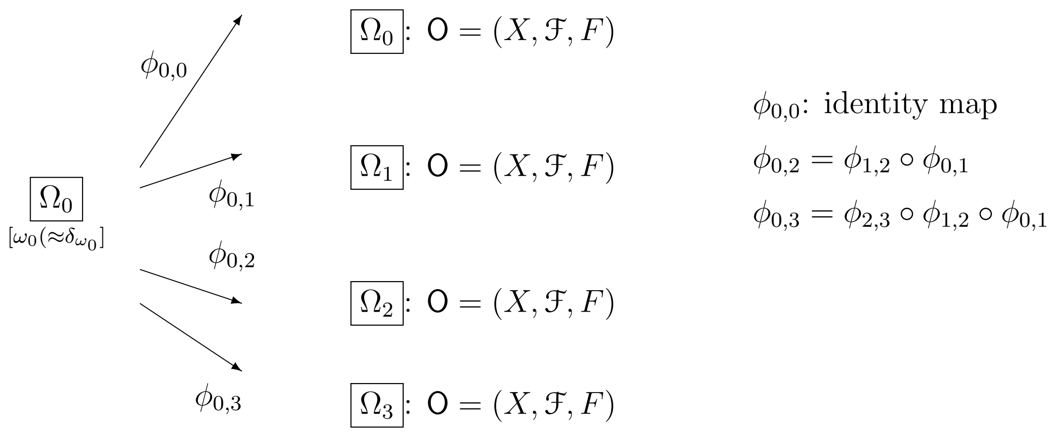
Consider the following time series (i.e., the case that $N = 3$, $\Omega_i = \Omega$, $i = 0, 1, 2, 3$)
(L)



where $\mathsf{O} = (X, \mathcal{F}, F)$ is arbitrary observable in $L^\infty(\Omega)$.

[(i) Schrödinger pictures (a state moves) :Parallel time)] of (L):

Figure (the case that $N = 3$; $\Omega = \Omega_i$, $i = 0, 1, 2, 3$)



Assume that the state $\omega_0 (\in \Omega)$ at time $t_0 (= 0)$ evolves in time to become $\phi_{0,t_k}(\omega_0)$ $(k = 0, 1, ..., N)$ as follows:

($\sharp_3$) state $\phi_{0,t_0}(\omega_0) = \omega_0$ at time $t_0 = 0/n (= 0)$
state $\phi_{0,t_1}(\omega_0)$ at time $t_1 = 1/n$
state $\phi_{0,t_2}(\omega_0)$ at time $t_2 = 2/n$
$\cdots$
state $\phi_{0,t_k}(\omega_0)$ at time $t_k = k/n$
$\cdots$
state $\phi_{0,t_n}(\omega_0)$ at time $t_n = n/n = 1$

And assume:

(M) At each time $t_0 (= 0), t_1 (= 1/n), ..., t_k (= k/n), ..., t_n (= 1)$, measurement $\mathsf{M}_{C(\Omega)}(\mathsf{O} = (X, \mathcal{F}, F), S_{[\phi_{0,t_k}(\omega_0)]})$ is taken.

That is, putting $T_n = \{t_0(=0), t_1(=1/n), \cdots, t_n(=1)\}$, we take the tensor product exact measurement:

$$
\begin{aligned}
&\bigotimes_{t_k \in T_n} \mathsf{M}_{C(\Omega)}(\mathsf{O} = (X, \mathcal{F}, F), S_{[(\phi_{0,t_k}(\omega_0))]}) \\
=&\mathsf{M}_{C(\Omega^{T_n})}(\bigotimes_{t_k \in T_n} \mathsf{O}_{E_\Omega} = (\Omega^{T_n}, \mathcal{B}(\Omega^{T_n}), \bigotimes_{t_k \in T_n} F), S_{[(\phi_{0,t_k}(\omega_0))_{t_k \in T_n}]})
\end{aligned}
$$

Then, we see that, for any $\Xi_k \subseteq X$ $(k = 1, 2, ..., n)$,

(N) the probability that the measured value belongs to $\times_{i=0}^{k} \Xi_k$ is given by

$$
\mathop{\times}_{k=0}^{n} [F(\Xi_k)](\phi_{0,t_k}(\omega_0))
$$

[(ii) Heisenberg picture (observable moves: (Series time)]

Figure (the case that $N = 3$; $\Omega = \Omega_i$, $\;i = 0, 1, 2, 3$)

$$
\underset{[\omega_0]}{\overset{\mathsf{O}=(X,\mathcal{F},F)}{\boxed{L^\infty(\Omega_0)}}} \xleftarrow{\Phi_{0,1}} \overset{\mathsf{O}=(X,\mathcal{F},F)}{\boxed{L^\infty(\Omega_1)}} \xleftarrow{\Phi_{1,2}} \overset{\mathsf{O}=(X,\mathcal{F},F)}{\boxed{L^\infty(\Omega_2)}} \xleftarrow{\Phi_{2,3}} \overset{\mathsf{O}=(X,\mathcal{F},F)}{\boxed{L^\infty(\Omega_3)}}
$$

As mentioned in the above, assume that the state $\omega_0 (\in \Omega)$ at time $t_0(=0)$, and $T_n = \{t_0(=0), t_1(=1/n), \cdots, t_{n-1}(=(n-1)/n), t_n(=1)\}$. Assume that, at each $t_0(=0), t_1(=1/n), \cdots, t_{n-1}(=(n-1)/n), t_n(=1)$, an observable $\mathsf{O} = (X, \mathcal{P}(X), F)$ is set.

($\flat_4$) the observable $\mathsf{O}(= (X, \mathcal{F}, F))$ at time $t_n(=1)$ is identified with the observable $\Phi_{t_{n-1},t_n}\mathsf{O}(= (X, \mathcal{F}, \Phi_{t_{n-1},t_n} F))$ at time t_{n-1}. At time t_{n-1}, we originally have an observable O, and the product of this O and $\Phi_{t_{n-1},t_n}\mathsf{O}$ gives the observable at time t_{n-1}:

$$
\mathsf{O} \times (\Phi_{t_{n-1},t_n}\mathsf{O}) \quad \big(= (X^2, \boxtimes_{k=1}^{2} \mathcal{F}, \widehat{F}_{n-1}) \big)
$$

Similarly, the observable it time t_{n-2} is represented by

$$
\mathsf{O} \times (\Phi_{t_{n-2},t_{n-1}}(\mathsf{O} \times (\Phi_{t_{n-1},t_n}\mathsf{O}))) \quad \big(= (X^3, \boxtimes_{k=1}^{3} \mathcal{F}, \widehat{F}_{n-2}) \big)
$$

Further, the observable at time t_{n-3} is represented by,

$$
\mathsf{O} \times (\Phi_{t_{n-3},t_{n-2}}(\mathsf{O} \times (\Phi_{t_{n-2},t_{n-1}}(\mathsf{O} \times (\Phi_{t_{n-1},t_n}\mathsf{O})))) \quad \big(= (X^4, \boxtimes_{k=1}^{4} \mathcal{F}, \widehat{F}_{n-3}) \big)
$$

Iteratively, after all, the observable at time t_0 is represented by,

$$
\begin{aligned}
\widehat{\mathsf{O}}_{t_0} &= \mathsf{O} \times (\Phi_{t_0,t_1}(\cdots(\mathsf{O} \times (\Phi_{t_{n-3},t_{n-2}}(\mathsf{O} \times (\Phi_{t_{n-2},t_{n-1}}(\mathsf{O} \times (\Phi_{t_{n-1},t_n}\mathsf{O})))))\cdots)) \\
&= (X^{n+1}, \boxtimes_{k=1}^{n+1} \mathcal{F}, \widehat{F}_0)
\end{aligned}
$$

Thus, we get the measurement $\mathsf{M}_{L^\infty(\Omega)}(\widehat{\mathsf{O}}_{t_0}, S_{[\omega_0]})$ at time $t = 0$. Therefore, putting $\Xi_k \subseteq X$ $(k = 1, 2, ..., n)$, we see that

For further information see my homepage

(O) the probability that its measured value belongs to $\times_{i=1}^{k}\Xi_k$ is given by $[\widehat{F}_0(\Xi_0 \times \Xi_1 \times \cdots \times \Xi_n)](\omega_0)$

Here, we see

$$
\begin{aligned}
&[\widehat{F}_0(\Xi_0 \times \Xi_1 \times \cdots \times \Xi_n)](\omega_0) \\
=&[F(\Xi_0)](\omega_0) \times \Phi_{t_{n-1},t_{n-2}}[\widehat{F}_1(\Xi_1 \times \cdots \times \Xi_n)](\omega_0) \\
=&[F(\Xi_0)](\omega_0) \times [\widehat{F}_1(\Xi_1 \times \cdots \times \Xi_n)](\omega_1) \\
\cdots& \\
=&\mathop{\times}_{k=0}^{n}[F(\Xi_k)](\phi_{0,t_k}(\omega_0))
\end{aligned}
$$

Here, note that (N)=(O) holds. Thus, we can conclude that

(P) Schrödinger and Heisenberg pictures are equivalent in the classical system

9.8.3 Derivation of the motion function method from (classical) quantum language

In the above, we see that the Schrödinger picture (N) and the Heisenberg picture (O) are equivalent in classical system. From here, consider the case of exact observables, i.e.,

$$\mathsf{O} = (X, \mathcal{F}, F) = (\Omega, \mathcal{B}(\Omega), E_\Omega) = \mathsf{O}_{E_\Omega}$$

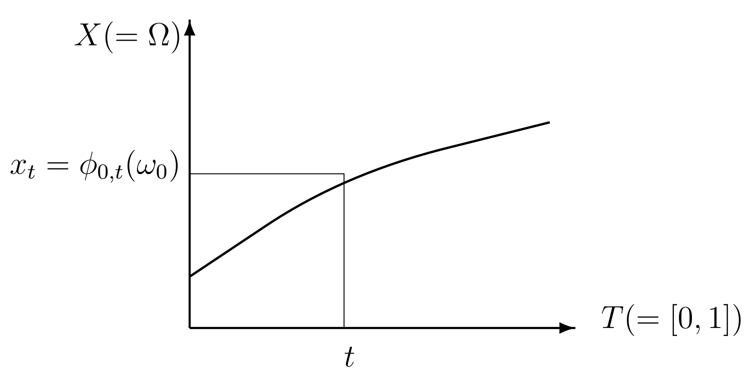
where $\mathcal{B}(\Omega)$ is the Borel field, $[E_\Omega(\Xi)](\omega) = 1 (\omega \in \Xi), = 0 (\omega \notin \Xi)$.

Put $T = [0, 1]$. And further, consider the infinite tensor product exact measurement

$$
\begin{aligned}
&\bigotimes_{t\in T} \mathsf{M}_{L^\infty(\Omega)}(\mathsf{O}_{E_\Omega} = (\Omega, \mathcal{B}(\Omega), E_\Omega), S_{[\phi_{0,t}(\omega_0)]}) \\
=&\mathsf{M}_{L^\infty(\Omega^T)}(\bigotimes_{t\in T} \mathsf{O}_{E_\Omega} = (\Omega^T, \mathcal{B}(\Omega^T), \bigotimes_{t\in T} E_\Omega), S_{[(\phi_{0,t}(\omega_0))_{t\in T}]})
\end{aligned}
$$

Thus, we see

(Q) When the tensor product exact measurement $\mathsf{M}_{L^\infty(\Omega^T)}(\bigotimes_{t\in T} \mathsf{O}_{E_\Omega} = (\Omega^T, \mathcal{B}(\Omega^T), E_{\Omega^T}), S_{[(\phi_{0,t}(\omega_{\omega_0})_{t\in T}]})$ is taken, the probability that the measured value $(x_t)_{t\in T} (\in \Omega^T)$ belongs to any open set which includes $(\omega_t)_{t\in T} (\in \Omega^T)$ is 1. In the same sense, the measured value $(x_t)_{t\in T} (\in \Omega^T)$ is surely equal to $(\phi_{0,t}(\omega_0))_{t\in T}$



For further information see my homepage

In general, define the position map $P' : \Omega(= X) \to X'$ such that

$$\Omega(= X) \ni [\text{state}] \overset{P'}{\mapsto} [\text{position}](= X')$$

Then, the motion function $m : T \to X'$ can be written as follows.

$$m(t) = P'(\phi_{0,t}(\omega_0)) \qquad (\forall t \in T)$$

♠**Note 9.9.** Readers may ask the following questions.

- Why is the author concerned with the Zeno's paradox (like schoolchildren's problem)?

The reason is as follows.

Seeing Figure 0.1, I think that

($\sharp$) the aim of Western philosophy from a scientific perspective is to propose a theory of everyday science, under which the Zeno's paradox is solved.

Also, recall that the main purpose of this book is to propose 'classical QL' as the theory of everyday science.

This is why I was concerned with Zeno's paradox. Also, if so, this would explain why western philosophy has been obsessed with Zeno's paradox for some 2,500 years. That is, I think that the followings are equivalent

($\flat_1$) To solve Zeno's paradox

($\flat_2$) to propose the theory of everyday science

♠**Note 9.10.** Many scientists may not understand the meaning of the 'philosophical theory of time'. In fact, I do not know what 'Bergson's theory of time' means either. However, as discussed in this chapter, the theory of time within QL would be understandable. Note that the tree structure should be linear, i.e., $T = \{t_0, t_1, ...t_n\}$, if we consider time series. That is,

$$T: \quad t_0 \longleftarrow t_1 \longleftarrow t_2 \longleftarrow ... \longleftarrow t_n$$

or, more generally,

$$T: \quad [t_0, \infty)$$

This implies that

($\sharp$) the beginning of time always exists. However, there is not always an end to time.

For further information see my homepage

Chapter 10

Simple measurement and causality

By chapter 10, we have learned all of quantum language, that is,

$(\sharp_1)$: pure measurement theory (=quantum language)

:= [(pure)Axiom 1] pure measurement (*cf.* §2.7) + [Axiom 2] (deterministic) Causality (*cf.* §9.3) [a kind of spells (a priori judgment)] + [quantum linguistic Copenhagen interpretation] Linguistic Copenhagen interpretation (*cf.* §3.1) [manual to use spells]

$(\sharp_2)$: mixed measurement theory (=quantum language)

:= [(mixed)Axiom$^{(m)}$ 1] mixed measurement (*cf.* §8.1) + [Axiom 2] Causality (*cf.* §9.3) [a kind of spells(a priori judgment)] + [quantum linguistic Copenhagen interpretation] Linguistic Copenhagen interpretation (*cf.* §3.1) [manual to use spells]

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.*

10.1 The Heisenberg picture and the Schrödinger picture

In Sec. 9.8.2 I discussed the Schrödinger picture and the Heisenberg picture are equivalent in the classical system, In this section I discuss the Schrödinger picture and the Heisenberg picture in quantum systems.

10.1.1 State does not move – the Heisenberg picture

We consider that

$$\text{"only one measurement"} \Longrightarrow \text{"state does not move"}$$

That is because

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

We are tempted to think that this is associated with Parmenides' words:

$$\textit{There is no movement,} \tag{10.1}$$

which is related to the Heisenberg picture. This will be explained in what follows.

Theorem 10.1. [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 Ξ $(\in \mathcal{F})$. And consider the countable decomposition $\{\Xi_1, \Xi_2, \ldots, \Xi_n, \ldots\}$ of Ξ $\Big($ i.e., $\Xi = \bigcup_{n=1}^{\infty} \Xi_n$, $\Xi_n \in \mathcal{F}, (n=1,2,\ldots)$, $\Xi_m \cap \Xi_n = \emptyset$ $(m \neq n)$ $\Big)$. 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 with 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,

$$\overline{\mathcal{A}}_0 \xleftarrow{\Phi_{0,1}} \overline{\mathcal{A}}_1. \tag{10.2}$$

Therefore, we have the pre-dual operator $(\Phi_{0,1})_*$ and the dual operator $\Phi^*_{0,1}$:

$$(\overline{\mathcal{A}}_0)_* \xrightarrow[(\Phi_{0,1})_*]{} (\overline{\mathcal{A}}_1)_* \qquad\qquad \mathcal{A}_0^* \xrightarrow[\Phi^*_{0,1}]{} \mathcal{A}_1^*. \tag{10.3}$$

If $\Phi_{0,1} : \overline{\mathcal{A}}_1 \to \overline{\mathcal{A}}_0$ is deterministic, we see that

$$\mathcal{A}_0^* \supset \mathfrak{S}^p(\mathcal{A}_0^*) \ni \rho \xrightarrow[\Phi^*_{0,1}]{} \Phi^*_{0,1}\rho \in \mathfrak{S}^p(\mathcal{A}_1^*) \subset \mathcal{A}_1^*. \tag{10.4}$$

For further information see my homepage

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 $\overline{\mathcal{A}}_1$.

Then, we see:

Explanation 10.2. [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

$$_{\mathcal{A}_0^*}\Big(\rho_0, \Phi_{0,1}(F_1(\Xi))\Big)_{\overline{\mathcal{A}}_0}. \tag{10.5}$$

Explanation 10.3. [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

$$\mathsf{M}_{\overline{\mathcal{A}}_1}(\mathsf{O}_1, S_{[\Phi_{0,1}^*\rho_0]}).$$

For further information see my homepage

Thus,

(b3) the probability that a measured value belongs to $\Xi (\in \mathcal{F})$ is given by

$$ {}_{\mathcal{A}_1^*}\Big(\Phi_{0,1}^* \rho_0, F_1(\Xi)\Big)_{\overline{\mathcal{A}}_1}, \tag{10.6}$$

which is equal to

$$ {}_{\mathcal{A}_0^*}\Big(\rho_0, \Phi_{0,1}(F_1(\Xi))\Big)_{\overline{\mathcal{A}}_0}. \tag{10.7}$$

In the above sense (i.e., (10.6) and (10.7)), we conclude that, under the condition (A_1),

the Heisenberg picture and the Schrödinger picture are equivalent.

That is,

$$\underset{\text{(Heisenberg picture)}}{\boxed{\mathsf{M}_{\overline{\mathcal{A}}_0}(\Phi_{0,1}\mathsf{O}_1, S_{[\rho_0]})}} \quad \underset{\text{(identification)}}{\longleftrightarrow} \quad \underset{\text{(Schrödenger picture)}}{\boxed{\mathsf{M}_{\overline{\mathcal{A}}_1}(\mathsf{O}_1, S_{[\Phi_{0,1}^*\rho_0]})}} \tag{10.8}$$

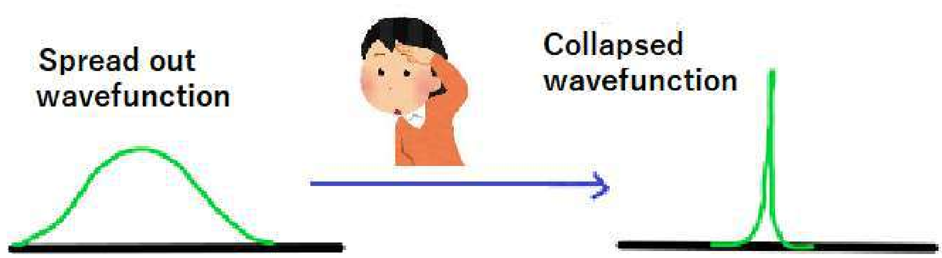
Remark 10.4. In the above, the conditions (A_1) is indispensable, that is,

(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. On the other hand, the Heisenberg picture is always formulated. Hence we consider that

$$\bullet \begin{cases} \text{the Heisenberg picture is formal} \\ \text{the Schrödinger picture is makeshift} \end{cases}$$

10.2 The wave function collapse (i.e., the projection postulate)



The linguistic 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:

- Ref. [59] *Linguistic Copenhagen interpretation of quantum mechanics; Projection Postulate,* JQIS, Vol. 5(4) , 150-155, 2015

10.2.1 Problem: How should the von Neumann-Lüders projection postulate be understood?

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

$$P_\lambda = P(\{\lambda\}) \qquad (\forall \lambda \in \Lambda) \tag{10.9}$$

Axiom 1 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

$$\mathrm{Tr}_H(\rho P_{\lambda_0})(= \langle u, P_{\lambda_0} u\rangle = \|P_{\lambda_0} u\|^2), \quad (\text{ where } \rho = |u\rangle\langle u|) \tag{10.10}$$

Also, the von Neumann-Lüders projection postulate (in so called Copenhagen interpretation, *cf.* refs. [110, 88]) 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 ρ_{post} is given by

$$\rho_{\text{post}} = \frac{P_{\lambda_0}|u\rangle\langle u|P_{\lambda_0}}{\|P_{\lambda_0} u\|^2}$$

And therefore, when a next measurement $\mathsf{M}_{B(H)}(\mathsf{O}_F :=(X, \mathcal{F}, F)$, $S_{[\rho_{\text{post}}]})$ 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

$$\mathrm{Tr}_H(\rho_{\text{post}} F(\Xi))\Big(= \langle \frac{P_{\lambda_0} u}{\|P_{\lambda_0} u\|}, F(\Xi)\frac{P_{\lambda_0} u}{\|P_{\lambda_0} u\|}\rangle\Big) \tag{10.11}$$

Problem 10.5. In the linguistic Copenhagen interpretation, the phrase:"post-measure- ment state" in the (A$_2$) is meaningless. Also, the above (=(A$_1$)+(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 Copenhagen interpretation.

10.2.2 The derivation of von Neumann-Lüders projection postulate in the linguistic Copenhagen 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 10.2.1, 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 $u \in H$,

$$\Psi_*(|u\rangle\langle u|) = |\sum_{\lambda \in \Lambda}(P_\lambda u \otimes e_\lambda)\rangle\langle\sum_{\lambda \in \Lambda}(P_\lambda u \otimes e_\lambda)| \tag{10.12}$$

For further information see my homepage

or

$$\Psi_*(|u\rangle\langle u|) = \sum_{\lambda\in\Lambda} |P_\lambda u \otimes e_\lambda\rangle\langle P_\lambda u \otimes e_\lambda| \tag{10.13}$$

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

$$\begin{aligned}
&\mathrm{Tr}_H[(|u\rangle\langle u|)\Psi(F(\Xi) \otimes G(\{\lambda_0\}))] = {}_{Tr(H)}\big(|u\rangle\langle u|, \Psi(F(\Xi) \otimes G(\{\lambda_0\}))\big)_{B(H)} \\
=&{}_{Tr(H\otimes K)}\big(\Psi_*(|u\rangle\langle u|), F(\Xi) \otimes G(\{\lambda_0\})\big)_{B(H\otimes K)} = \mathrm{Tr}_{H\otimes K}[(\Psi_*(|u\rangle\langle u|))(F(\Xi) \otimes G(\{\lambda_0\}))] \\
=&\mathrm{Tr}_{H\otimes K}[(|\sum_{\lambda\in\Lambda}(P_\lambda u \otimes e_\lambda)\rangle\langle\sum_{\lambda\in\Lambda}(P_\lambda u \otimes e_\lambda)|)(F(\Xi) \otimes |e_{\lambda_0}\rangle\langle e_{\lambda_0}|)] \\
=&\langle P_{\lambda_0} u, F(\Xi) P_{\lambda_0} u\rangle \quad (\forall \Xi \in \mathcal{F})
\end{aligned}$$

(In a similar way, the same result is easily obtained in the case of (10.13)).
Thus, we see the following.

(D_1) if $\Xi = X$, then

$$\mathrm{Tr}_H[(|u\rangle\langle u|)\Psi(F(X) \otimes G(\{\lambda_0\}))] = \langle P_{\lambda_0} u, P_{\lambda_0} u\rangle = \|P_{\lambda_0} u\|^2 \tag{10.14}$$

(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

$$\frac{\langle P_{\lambda_0} u, F(\Xi) P_{\lambda_0} u\rangle}{\|P_{\lambda_0} u\|^2} \Big(= \langle \frac{P_{\lambda_0} u}{\|P_{\lambda_0} u\|}, F(\Xi) \frac{P_{\lambda_0} u}{\|P_{\lambda_0} u\|} \rangle \Big) \quad (\forall \Xi \in \mathcal{F}) \tag{10.15}$$

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,

$$(10.10) \Leftrightarrow (10.14), \quad (10.11) \Leftrightarrow (10.15)$$

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 10.5 (i.e., $\Psi(\mathsf{O}_F \otimes \mathsf{O}_G)$).

Remark 10.6. So called Copenhagen interpretation may admit the post-measurement state (*cf.* ref. [24]). 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 approved. That is because, if the post-measurement state is admitted, a series of problems occur, that is, "When is a measurement taken?", "When does the wave function collapse happen?", or "How fast is the wave function collapse?", which is beyond Axioms 1 and 2. Hence, the projection postulate is usually regarded as "postulate". On the other hand, in the linguistic Copenhagen 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"*.

Postulate 10.7. [Projection postulate, *cf.* ref. [59]] As mentioned in the above, the statement (A_2) (= von Neumann-Lüders projection postulate) is wrong. However, 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

$$\rho_{\text{post}} = \frac{P_{\lambda_0}|u\rangle\langle u|P_{\lambda_0}}{\|P_{\lambda_0}u\|^2} \tag{10.16}$$

10.3 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

$$\nabla^2 = \frac{\partial^2}{\partial q_1^2} + \frac{\partial^2}{\partial q_2^2} + \frac{\partial^2}{\partial q_3^2},$$

we consider Schrödinger equation (concerning one particle):

$$i\hbar\frac{\partial}{\partial t}\psi(\mathbf{q},t) = \Big[\frac{-\hbar^2}{2m}\nabla^2 + V(\mathbf{q},t)\Big]\psi(\mathbf{q},t) \tag{10.17}$$

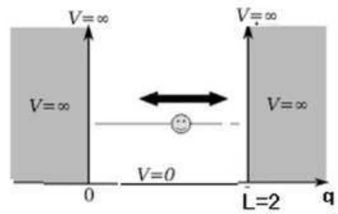
where m is the mass of the particle, V is a potential energy.

For simplicity, we discuss one dimensional case $\mathbb{R}$, and 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 10.8. [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:

$$\begin{cases} \text{initial state:} \psi(\cdot, t_0) = \psi_{t_0} \\ i\hbar \frac{\partial}{\partial t}\psi(q,t) = \left[\frac{-\hbar^2}{2m}\frac{\partial^2}{\partial q^2} + V(q,t)\right]\psi(q,t) \end{cases} \tag{10.18}$$



Consider the same situation in §10.5, i.e., a particle with the mass m in the box of closed interval $[0, 2]$ in one dimensional space $\mathbb{R}$.

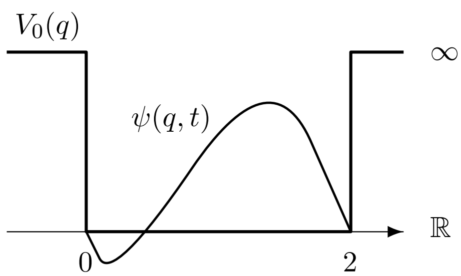


Figure 10.1(1)(time t_0)

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

$$V_1(q) = \begin{cases} 0 & (0 \le q < 1) \\ \infty & (q = 1) \\ 0 & (1 < q \le 2) \\ \infty & (\text{ otherwise }) \end{cases} \tag{10.19}$$

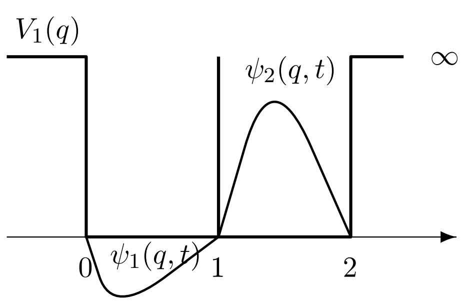


Figure 10.1(2)(partition)

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)$\big]$.

For further information see my homepage

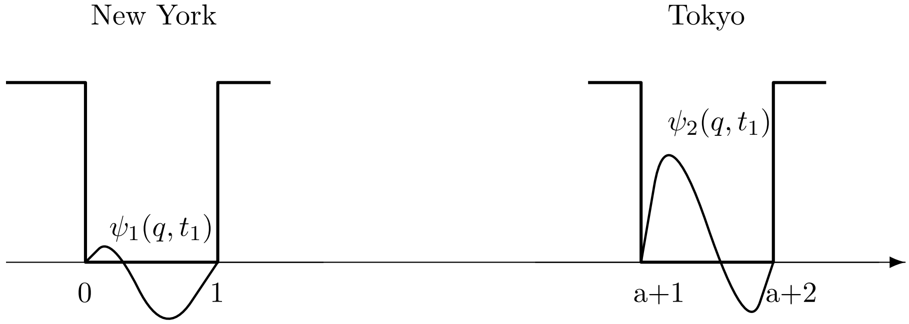


Figure 10.1(3)(time t_1)

Here, $1 \ll a$. Solving the Schrödinger equation (10.18), 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

$$\Phi_{t_0,t_1}(A)=U^*_{t_0,t_1}AU_{t_0,t_1}\qquad (\forall A\in B(L^2(\mathbb{R}_{t_2})))$$

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 10.9.
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

$$\begin{bmatrix}\langle u_{t_0},\Phi_{t_0.t_1}F(\{N\})u_{t_0}\rangle=\int_0^1|\psi_1(q,t_1)|^2dq\\ \langle u_{t_0},\Phi_{t_0.t_1}F(\{T\})u_{t_0}\rangle=\int_{a+1}^{a+2}|\psi_2(q,t_1)|^2dq\\ \langle u_{t_0},\Phi_{t_0.t_1}F(\{E\})u_{t_0}\rangle=0\end{bmatrix}.$$

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,

$$\mathbf{(A}_1\mathbf{)=(A}_2\mathbf{)}$$

Remark 10.10. 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 Projection postulate 10.7, we say that at the moment the wave function ψ_2 vanishes. That is,

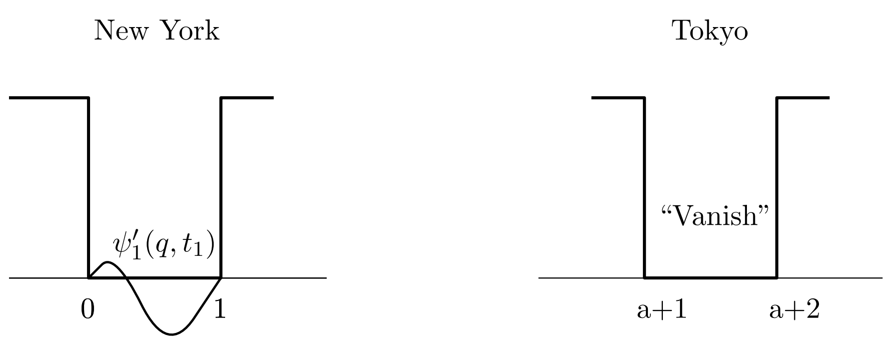


Figure 10.1(4) (The wave function after measurement)

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

$$\psi_1(\cdot,t_1)+\psi_2(\cdot,t_1) \xrightarrow[\text{the collapse of wave function}]{} \psi'_1(\cdot,t_1) \tag{10.20}$$

Also, note that New York $\big[$resp. Tokyo$\big]$ may be the earth $\big[$resp. the polar star$\big]$. 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.* refs. [13, 107]). This is a true paradox, which is not clarified even in quantum language.

10.4 Quantum Zeno effect

This section is extracted from

- Ref. [50]: S. Ishikawa; Heisenberg uncertainty principle and quantum Zeno effects in the linguistic Copenhagen interpretation of quantum mechanics
(arXiv:1308.5469 [quant-ph] 2014)

10.4.1 Quantum decoherence: non-deterministic sequential causal operator

Let us start from a review of Section 9.6.2 (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

$$\begin{aligned} &\mathrm{Tr}((\Psi_{\mathbb{P}})_*(|u\rangle\langle u|)) \\ =&\mathrm{Tr}(\sum_{n=1}^{\infty} |P_n u\rangle\langle P_n u|) = \sum_{n=1}^{\infty}\sum_{k=1}^{\infty} |\langle e_k, P_n u\rangle|^2 = \sum_{n=1}^{\infty} \|P_n u\|^2 = \|u\|^2 \qquad (\forall u \in H) \end{aligned}$$

Hence

$$(\Psi_{\mathbb{P}})_*(\mathcal{T}r_{+1}^{p}(H)) \subseteq \mathcal{T}r_{+1}(H).$$

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 10.11. [Quantum decoherence in quantum Zeno effect *cf.* ref. [47]]. 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 $\mathcal{H}$ 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 (9.21). 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

$$\underset{\rho=|\psi\rangle\langle\psi|}{\boxed{B(H_0)}} \xleftarrow[\Phi_{0,1}^{(N)}]{} \underset{\mathsf{O}_1 := (X,\mathcal{F},F)}{\boxed{B(H_1)}}$$

Thus, we have a measurement:

$$\mathsf{M}_{B(H_0)}(\Phi_{0,1}^{(N)} \mathsf{O}_1, S_{[\rho]})$$

(or more precisely, $\mathsf{M}_{B(H_0)}(\Phi_{0,1}^{(N)} \mathsf{O} := (X, \mathcal{F}, \Phi_{0,1}^{(N)} F),\ S_{[|\psi\rangle\langle\psi|]})$). 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

$$\mathrm{Tr}(|\psi\rangle\langle\psi| \cdot \Phi_{0,1}^{(N)} F(\Xi)). \tag{10.21}$$

Now we shall explain "quantum Zeno effect" in the following example.

Example 10.12. [Quantum Zeno effect]

Hot soup is hard to cool down when you see it.

. . . .

Let $\psi \in H$ such that $\|\psi\| = 1$. Define the spectrum decomposition

$$\mathbb{P} = [P_1(= |\psi\rangle\langle\psi|), P_2(= I - P_1)]. \tag{10.22}$$

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

$$F(\{x_1\}) = |\psi\rangle\langle\psi|(= P_1), \quad F(\{x_2\}) = I - |\psi\rangle\langle\psi|(= P_2).$$

Now we can calculate (10.21)(i.e., the probability that a measured value x_1 is obtained) as follows.

$$\begin{aligned}(10.21) &= \langle\psi, ((\Psi_S^{1/N})_*(\Psi_{\mathbb{P}})_*)^N(|\psi\rangle\langle\psi|)\psi\rangle \\ &\geq |\langle\psi, e^{-\frac{i\mathcal{H}}{\hbar N}}\psi\rangle\langle\psi, e^{\frac{i\mathcal{H}}{\hbar N}}\psi\rangle|^N \\ &\approx \Big(1 - \frac{1}{N^2}\big(||(\frac{\mathcal{H}}{\hbar})\psi||^2 - |\langle\psi, (\frac{\mathcal{H}}{\hbar})\psi\rangle|^2\big)\Big)^N \to 1 \qquad (N \to \infty)\end{aligned} \tag{10.23}$$

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 roughly say in Schrödinger picture that

the state $|\psi\rangle\langle\psi|$ does not move.

Remark 10.13. The above argument is motivated by B. Misra and E.C.G. Sudarshan (ref. [93]). However, the title of their paper: "The Zeno's paradox in quantum theory" is not appropriate. That is because

(B) the spectrum decomposition $\mathbb{P}$ should not be regarded as an observable (or moreover, measurement).

The effect in Example 10.12 should be called "brake effect" and not "watched pot effect".

10.5 Schrödinger's cat, Wigner's friend and Laplace's demon

10.5.1 Schrödinger's cat and Wigner's friend

Let us explain Schrödinger's cat paradox in the Schrödinger picture.

Problem 10.14. [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.

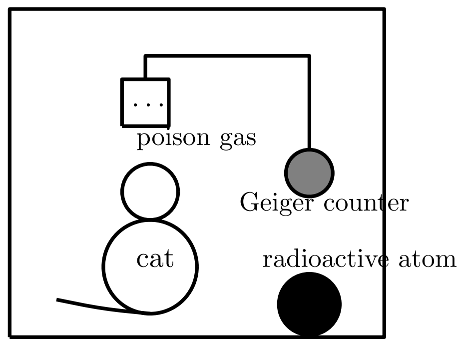
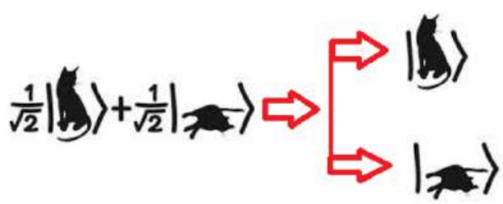


Figure 10.2: Schrödinger's cat

Here, we have the following question:

(b) Assume that, after one hour, you look at the inside of the box. Then, do you know whether the cat is dead or alive after one hour ?
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" !



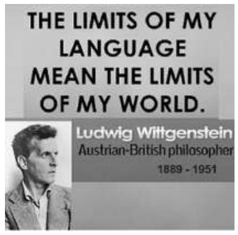
♠**Note 10.1.** [Wigner's friend]: Instead of the above (b), we consider as follows.

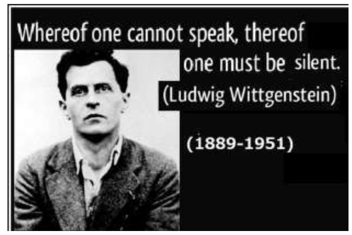
(b′) after one hour, Wigner's friend look at the inside of the box, and thus, he knows whether the cat is dead or alive after one hour. And further, after two hours, Wigner's friend informs you of the fact. How is the cat ?

This problem is not difficult. That is because the linguistic Copenhagen 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.





10.5.2 The usual answer

Answer 10.15. [The first answer to Problem 10.14 (i.e., The pure state, Projection postulate 10.7)].

Put $\mathbf{q} = (q_{11}, q_{12}, q_{13}, q_{21}, q_{22}, q_{23}, \dots, q_{n1}, q_{n2}, q_{n3}) \in \mathbb{R}^{3n}$. 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):

$$\begin{cases} i\hbar \frac{\partial}{\partial t}\psi(\mathbf{q},t) = \Big[\sum_{i=1}^n \frac{-\hbar^2}{2m_i} \nabla_i^2 + V(\mathbf{q},t) \Big] \psi(\mathbf{q},t) \\ \\ \psi_0(\mathbf{q}) = \psi(\mathbf{q},0) : \text{initial condition} \end{cases} \tag{10.24}$$

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 (10.24). 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} $(\in \mathcal{T}r^p_{+1}(H))$ after 60^2 seconds is represented by

$$\rho_{60^2} = |\psi_{60^2}\rangle\langle\psi_{60^2}| \tag{10.25}$$

(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

$$\psi_{60^2}=\frac{1}{\sqrt{2}}(\psi_{\text{life}}+\psi_{\text{death}}). \tag{10.26}$$

$$(\text{where } \psi_{\text{life}}\in V_{\text{life}}, \|\psi_{\text{life}}\|=1 \quad \psi_{\text{death}}\in V_{\text{death}}, \|\psi_{\text{death}}\|=1)$$

Hence. we can conclude that

(A$_4$) the state (or, wave function) of the cat (after one hour) is represented by (10.26), that is,

$$\frac{\text{“Fig.}(\sharp_1)\text{”}+\text{“Fig.}(\sharp_2)\text{”}}{\sqrt{2}}$$

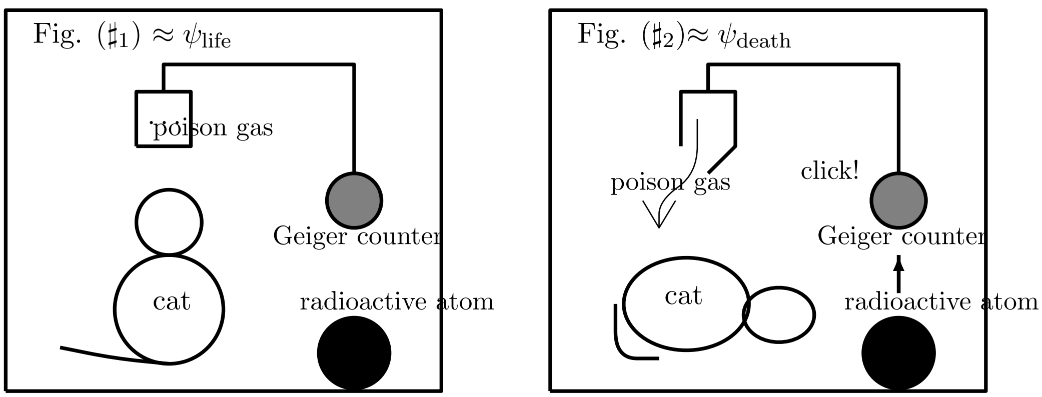


Figure 10.3: Schrödinger's cat(half and half)

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 projection postulate 10.7 (precisely speaking, by the misunderstanding of projection postulate 10.7),

$$\xrightarrow[\text{the collapse of wave function}]{\text{to the moment of opening a window}}\begin{cases}\text{“alive”}(=|\psi_{\text{life}}\rangle\langle\psi_{\text{life}}|)\\ \\ \text{“dead”}(=|\psi_{\text{death}}\rangle\langle\psi_{\text{death}}|)\end{cases}$$

□

10.5.3 The answer using decoherence

Answer 10.16. [The second answer to Problem 10.14)].

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

$$\rho'_{60^2} = \frac{1}{2}\Big(|\psi_{\rm life}\rangle\langle\psi_{\rm life}| + |\psi_{\rm death}\rangle\langle\psi_{\rm death}|\Big)$$

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

$$\begin{bmatrix} \mathcal{T}r(H)\Big(\rho_0, \Phi_{0,60^2}F(\{\text{life}\})\Big)_{B(H)} = \langle\psi'_{60^2}, F(\{\text{life}\})\psi_{60^2}\rangle = 0.5 \\ \mathcal{T}r(H)\Big(\rho_0, \Phi_{0,60^2}F(\{\text{death}\})\Big)_{B(H)} = \langle\psi'_{60^2}, F(\{\text{death}\})\psi_{60^2}\rangle = 0.5 \end{bmatrix}.$$

Also, "the moment of measuring" and "the collapse of wave function" are prohibited in the linguistic Copenhagen interpretation, but the statement (B$_2$) holds in quantum language. □

10.5.4 Summary (Laplace's demon)

Summary 10.17. [Schrödinger's cat in quantum language]

Here, let us examine

Answer10.15 :(A$_5$) vs. **Answer10.16 :(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$_3$) **Answer 10.16:(B$_2$)** is superior to **Answer 10.15:(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 micro-world, chaos appears. This kind of supremacy of physics is not natural, and thus, we consider that these are beyond "the limit of Newtonian mechanics"

And,

- when we want to apply Newton mechanics to phenomena beyond "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.

$$\underset{\text{physics}}{\boxed{\text{Newtonian mechanics}}} \xrightarrow[\text{linguistic turn}]{\text{beyond the limits}} \underset{\text{metaphysics}}{\boxed{\text{dynamical system theory; statistics}}} \tag{10.27}$$

In the same sense, we consider that quantum mechanics has "the limit". That is,

- Schrödinger's cat is beyond quantum mechanics.

And thus,

- When we want to apply quantum mechanics to phenomena beyond "the limit of quantum mechanics", we often use the quantum decoherence. Although this approach is not physics but metaphysics, it is quite powerful.

$$\underset{\text{physics}}{\boxed{\text{quantum mechanics}}} \xrightarrow[\text{linguistic turn}]{\text{beyond the limits}} \underset{\text{metaphysics}}{\boxed{\text{quantum language}}}$$

♠**Note 10.2.** 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 super realistic-view (i.e., the realistic-view over which the degree passed). Thus, we consider the following correspondence:

$$\underset{\text{physics}}{\boxed{\text{Newtonian mechanics}}} \xrightarrow[\text{super realistic-view}]{\text{beyond the limits}} \underset{\text{physics ?}}{\boxed{\text{Laplace's Demon}}} \tag{10.28}$$

This should be compared with the formula (10.27).

10.6 Wheeler's Delayed choice experiment: " Particle or wave ?" is a foolish question

This section is extracted from

(♯) [56] S. Ishikawa, *The double-slit quantum eraser experiments and Hardy's paradox in the quantum linguistic Copenhagen interpretation,* arxiv:1407.5143[quantum-ph], (2014)

10.6.1 "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)

World-views \ P or W	Particle(=symbol)	Wave(= math. represent)
Aristotle	hyle	eidos
Newton mechanics	point mass	state (=(position, momentum))
Statistics	population	parameter
Quantum mechanics	particle	state ($\approx$ wave function)
Quantum language	system (=measuring object)	state

In table 11.1, Newtonian mechanics (i.e., mass point $\leftrightarrow$ state) may be easiest to understand. In view of this table, we understand "particle" and "wave" are not contradictory concepts, so that it is possible to think

(A_1) "Particle or wave" is a foolish question.

On the other hand

(A_2) we have Wheeler's delayed choice experiment on "particle or wave".

So let me answer the interesting question:

(A_3) How is Wheeler's delayed choice experiment described in quantum mechanics ?

10.6.2 Preparation

Let us start from a 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

$$F(\{1\}) = |f_1\rangle\langle f_1|, \quad F(\{2\}) = |f_2\rangle\langle f_2| \quad \text{and} \quad G(\{1\}) = |g_1\rangle\langle g_1|, \quad G(\{2\}) = |g_2\rangle\langle g_2|$$

where

$$g_1 = \frac{f_1 - f_2}{\sqrt{2}}, \qquad g_2 = \frac{f_1 + f_2}{\sqrt{2}}.$$

10.6.3 de Broglie's paradox in $B(\mathbb{C}^2)$ (No interference)

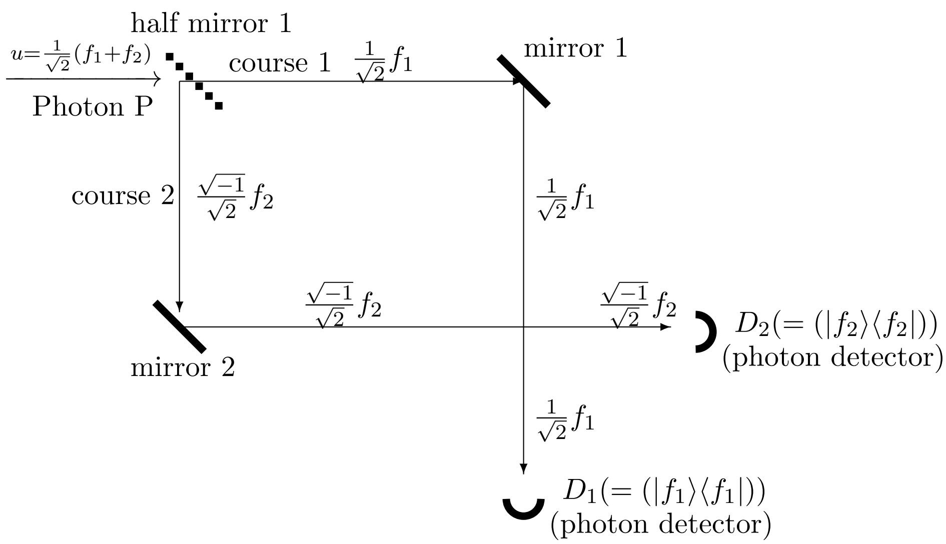


Figure 10.4(1). $[D_1 + D_2]$=Observable O_f

Now we shall explain, in the Schrödinger picture, Figure 10.4(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 at 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 at the mirror 2, and goes to the photon detector D_2.

This is, in the Heisenberg picture, represented by the following measurement:

$$\mathsf{M}_{B(\mathbb{C}^2)}(\Phi \mathsf{O}_f, S_{[\rho]}) \tag{10.29}$$

Then, we see:

For further information see my homepage

(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

$$\begin{bmatrix} \langle Uu, F(\{1\})Uu \rangle \\ \langle Uu, F(\{2\})Uu \rangle \end{bmatrix} = \begin{bmatrix} |\langle Uu, f_1 \rangle|^2 \\ |\langle Uu, f_2 \rangle|^2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} \tag{10.30}$$

Remark 10.18. [Projection postulate] By the analogy of Section 10.2 (The projection postulate), Figure 10.4(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,

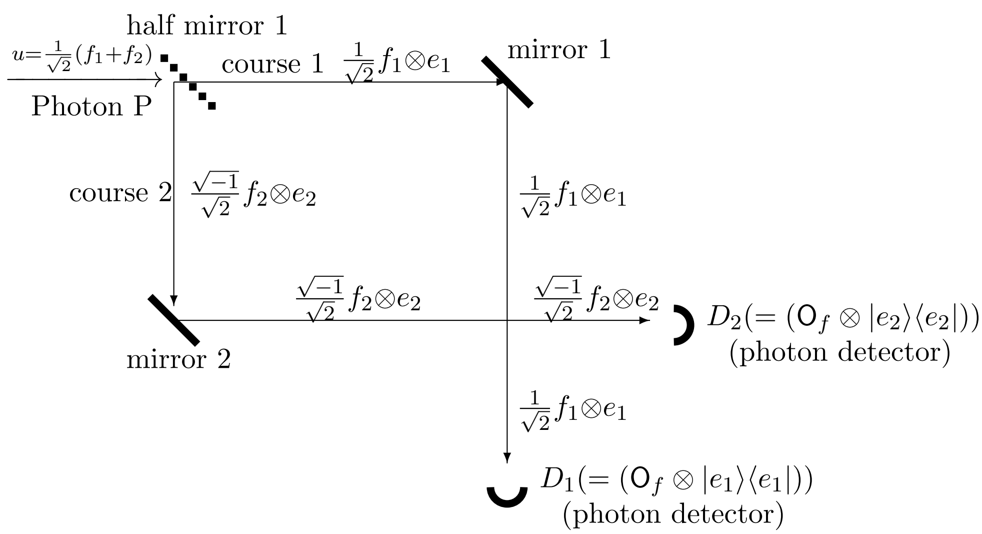


Figure 10.4(1′). $[D_1 + D_2]$=$\mathsf{O}_f \otimes \mathsf{O}_E$

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.

10.6.4 Mach-Zehnder interferometer (Interference)

Next, consider the following figure:

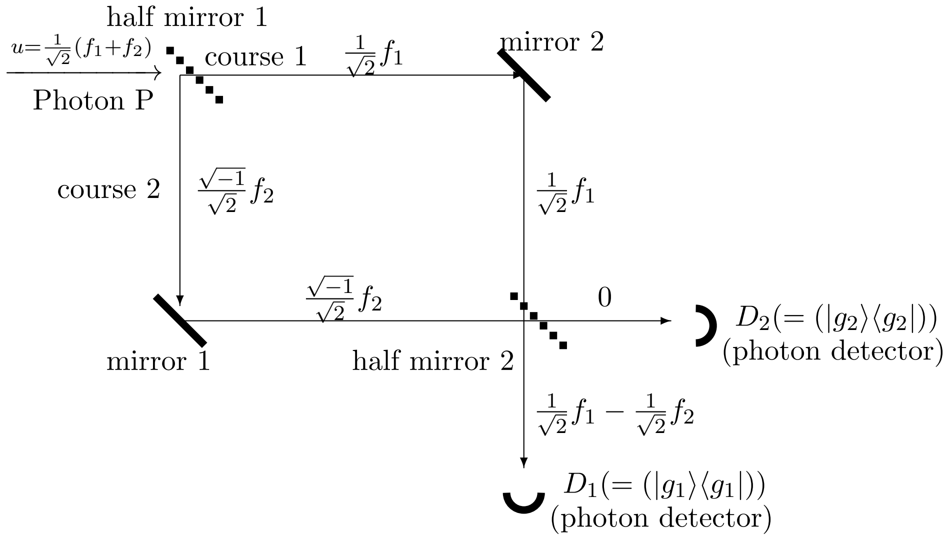


Figure 10.4(2). $[D_1 + D_2]$=ObservableO_g

Now we shall explain, by the Schrödinger picture, Figure 10.4(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 at 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 at 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:
$\mathsf{M}_{B(\mathbb{C}^2)}(\Phi^2\mathsf{O}_g, S_{[\rho]})$. 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

$$\begin{bmatrix}\langle u, \Phi^2 G(\{1\})u\rangle\\ \langle u, \Phi^2 G(\{2\})u\rangle\end{bmatrix} = \begin{bmatrix}|\langle u, UUg_1\rangle|^2\\ |\langle u, UUg_2\rangle|^2\end{bmatrix} = \begin{bmatrix}1\\0\end{bmatrix}. \tag{10.31}$$

10.6.5 Another case

Consider the following Figure 10.4(3).

For further information see my homepage

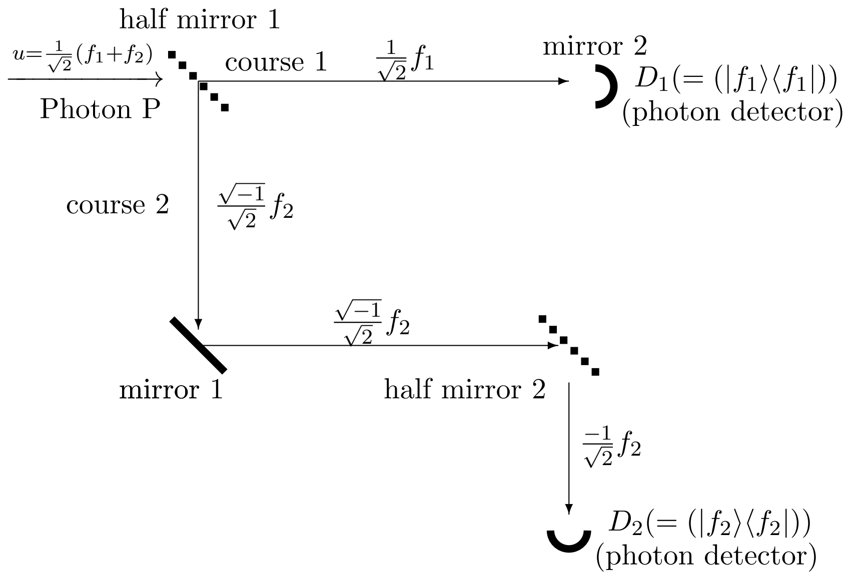


Figure 10.4(3). $[D_2 + D_1]$ =ObservableO_f

Now we shall explain, by the Schrödinger picture, Figure 10.4(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 at the mirror 1, and further reflected in the half-mirror 2, and goes to the photon detector D_2.

This is, in the Heisenberg picture, represented by the following measurement:

$$\mathsf{M}_{B(\mathbb{C}^2)}(\Phi^2\mathsf{O}_f, S_{[\rho]}). \tag{10.32}$$

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.

10.6.6 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 10.4(1))– §11.5 (Figure 10.4(3)) lies in the observables (= measuring instrument). That is,

For further information see my homepage

$$\begin{cases} \S 11.5.3 \text{ (Figure 10.4(1))} & \xrightarrow[\text{Heisenberg picture}]{} \Phi \mathsf{O}_f \\ \S 11.5.4 \text{ (Figure 10.4(2))} & \xrightarrow[\text{Heisenberg picture}]{} \Phi^2 \mathsf{O}_g \\ \S 11.5.5 \text{ (Figure 10.4(3))} & \xrightarrow[\text{Heisenberg picture}]{} \Phi^2 \mathsf{O}_f \end{cases}$$

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 10.4(1), Figure 10.4(2) and Figure 10.4(3) is chosen" — can not be described paradoxically in quantum language.

Hence, Wheeler's delayed choice experiment is not a paradox in quantum language, or in the sense of Wittgenstein's words (i.e., the spirit of the linguistic world view):

What we cannot speak about we must pass over in silence.

However, it should be noted that the non-locality paradox (i.e., "there is something faster than light") is not solved even in quantum language.

♠**Note 10.3.** What we want to assert in this book may be the following:

(♯) everything (except "there is something faster than light") can not be described paradoxically in quantum language

10.7 Hardy's paradox: total probability is less than 1

In this section, we shall introduce the Hardy's paradox (*cf.* ref.[18]) 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

$$P(\alpha_{11} f_1 \otimes f_1 + \alpha_{12} f_1 \otimes f_2 + \alpha_{21} f_2 \otimes f_1 + \alpha_{22} f_2 \otimes f_2)$$

---

[1]This section is extracted from

(♯) [56] S. Ishikawa, *The double-slit quantum eraser experiments and Hardy's paradox in the quantum linguistic Copenhagen interpretation,* arxiv:1407.5143[quantum-ph],(2014)

$$= -\alpha_{12} f_1 \otimes f_2 - \alpha_{21} f_2 \otimes f_1 + \alpha_{22} f_2 \otimes f_2$$

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.

10.7.1 Observable $\mathsf{O}_g \otimes \mathsf{O}_g$

Consider the following figure

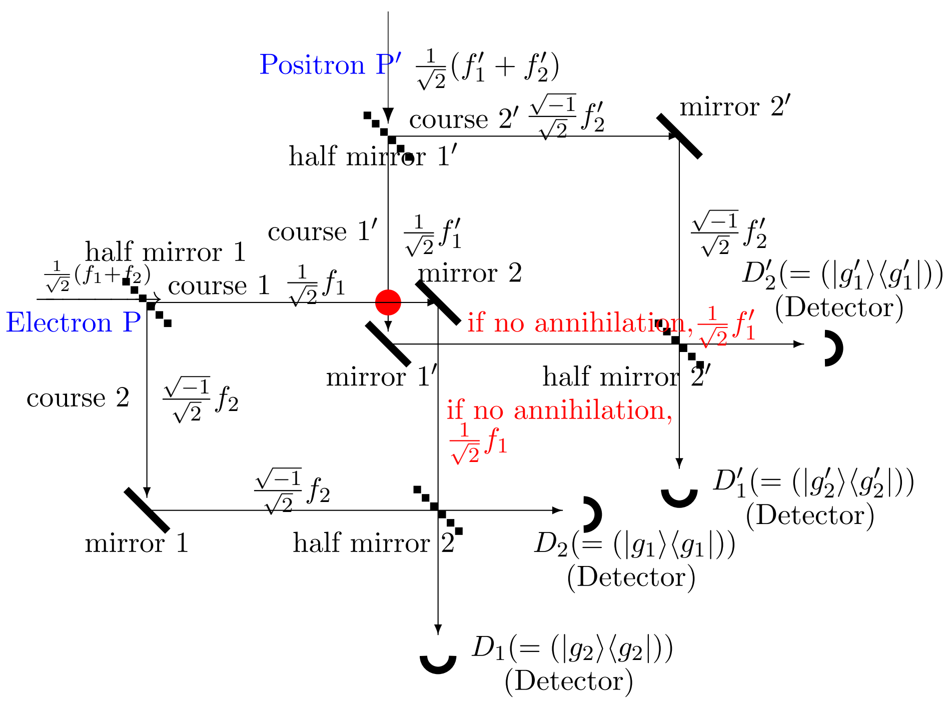


Figure 10.5(1). Electron P and Positron P′ are annihilated at ●

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:

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} (= f_1 = f'_1) \xrightarrow[\text{pass through half-mirror}]{} \begin{bmatrix} 1 \\ 0 \end{bmatrix} (= f_1 = f'_1)$$

For further information see my homepage

$$\begin{bmatrix}0\\1\end{bmatrix}(=f_2=f_2') \xrightarrow[\text{be reflected in half-mirror, and } \times\sqrt{-1}]{} \sqrt{-1}\begin{bmatrix}0\\1\end{bmatrix}(=f_2=f_2')$$

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

$$\begin{aligned}
&(\beta_1 f_1 + \beta_2 f_2) \otimes (\beta_1' f_1' + \beta_2' f_2') = \beta_1\beta_1' f_1 \otimes f_1' + \beta_1\beta_2' f_1 \otimes f_2' + \beta_2\beta_1' f_2 \otimes f_1' + \beta_2\beta_2' f_2 \otimes f_2'\\
&\xrightarrow[\text{(half-mirror)}]{}\\
&\beta_1\beta_1' f_1 \otimes f_1' + \sqrt{-1}\beta_1\beta_2' f_1 \otimes f_2' + \sqrt{-1}\beta_2\beta_1' f_2 \otimes f_1' - \beta_2\beta_2' f_2 \otimes f_2'\\
&\xrightarrow[\text{(annihilation(i.e., } f_1 \otimes f_1' = 0))]{}\\
&\sqrt{-1}\beta_1\beta_2' f_1 \otimes f_2' + \sqrt{-1}\beta_2\beta_1' f_2 \otimes f_1' - \beta_2\beta_2' f_2 \otimes f_2'\\
&\xrightarrow[\text{(second half-mirror)}]{}\\
&-\beta_1\beta_2' f_1 \otimes f_2' - \beta_2\beta_1' f_2 \otimes f_1' + \beta_2\beta_2' f_2 \otimes f_2'
\end{aligned}$$

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,

$$\begin{aligned}
&\widehat{H}_{gg}(\{(1,1)\}) = |g_1 \otimes g_1\rangle\langle g_1 \otimes g_1|, \quad \widehat{H}_{gg}(\{(1,2)\}) = |g_1 \otimes g_2\rangle\langle g_1 \otimes g_2|,\\
&\widehat{H}_{gg}(\{(2,1)\}) = |g_2 \otimes g_1\rangle\langle g_2 \otimes g_1|, \quad \widehat{H}_{gg}(\{(2,2)\}) = |g_2 \otimes g_2\rangle\langle g_2 \otimes g_2|
\end{aligned}$$

Consider the measurement:

$$\mathsf{M}_{B(\mathbb{C}^2\otimes\mathbb{C}^2)}(\widehat{\Psi}\widehat{\mathsf{O}}_{gg}, S_{[\widehat{\rho}]}) \tag{10.33}$$

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

$$\begin{aligned}
&\langle u \otimes u, P\widehat{H}_{gg}(\{(1,2)\})P(u \otimes u)\rangle\\
=&\frac{|\langle (f_1 + f_2) \otimes (f_1 - f_2), f_1 \otimes f_2 + f_2 \otimes f_1 + f_2 \otimes f_2\rangle|^2}{16}
\end{aligned}$$

For further information see my homepage

$$= \frac{|\langle f_1 \otimes f_1 - f_1 \otimes f_2 + f_2 \otimes f_1 - f_2 \otimes f_2, \;\; f_1 \otimes f_2 + f_2 \otimes f_1 + f_2 \otimes f_2 \rangle|^2}{16} = \frac{1}{16}$$

Similarly,

$$\langle u \otimes u, P\widehat{H}_{gg}(\{(2,1)\})P(u \otimes u)\rangle = \frac{1}{16}$$

Remark 10.19. 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 $\frac{1}{4}$.

10.7.2 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:

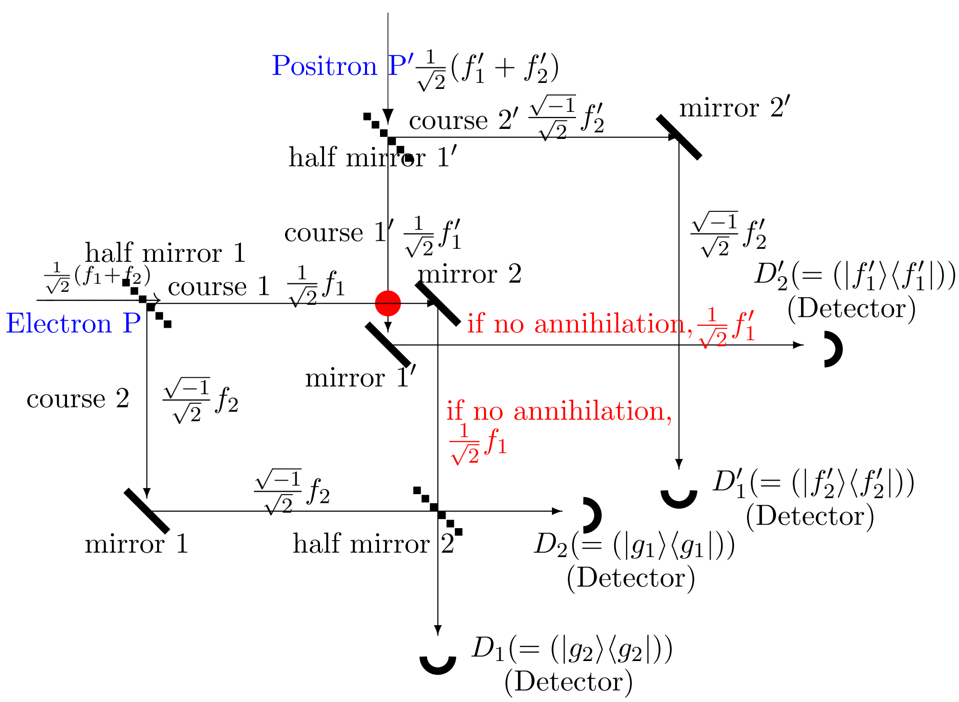


Figure 10.5(2). Electron P and Positron P′ are annihilated at ●

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:

$$\mathsf{M}_{B(\mathbb{C}^2 \otimes \mathbb{C}^2)}(\widehat{\Psi}\widehat{\mathsf{O}}_{gf}, S_{[\widehat{\rho}]}) \tag{10.34}$$

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 10.20. 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.

10.8 quantum eraser experiment

Let us explain quantum eraser experiment(*cf.* [111]). This section is extracted from

(♯) [56] S. Ishikawa, *The double-slit quantum eraser experiments and Hardy's paradox in the quantum linguistic Copenhagen interpretation,* arxiv:1407.5143[quantum-ph],(2014)

10.8.1 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 $|\alpha_1|^2 + |\alpha_2|^2 = 1$.

10.8.2 Interference

Consider the measurement:

$$\mathsf{M}_{B(\mathbb{C}^2 \otimes H)}(\mathsf{O}_x \otimes \mathsf{O}, S_{[|\psi\rangle\langle\psi|]}) \tag{10.35}$$

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[Real part]$(\overline{\alpha}_1\alpha_2\langle u_1, F(\Xi)u_2\rangle)$), we get the following graph.

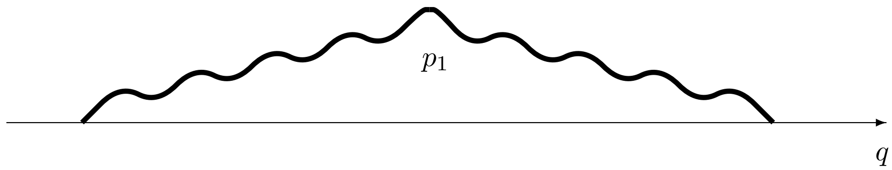


Figure 10.6(1): The graph of p_1

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

$$
\begin{aligned}
&\langle\psi,(F_x(\{-1\})\otimes F(\Xi))\psi\rangle\\
=&\langle\alpha_1e_1\otimes u_1+\alpha_2e_2\otimes u_2,(F_x(\{-1\}\otimes F(\Xi)))(\alpha_1e_1\otimes u_1+\alpha_2e_2\otimes u_2)\rangle\\
=&\frac{1}{2}\langle\alpha_1e_1\otimes u_1+\alpha_2e_2\otimes u_2,\alpha_1(e_1-e_2)\otimes F(\Xi)u_1+\alpha_2(-e_1+e_2)\otimes F(\Xi)u_2\rangle\\
=&\frac{1}{2}\Big(|\alpha_1|^2\langle u_1,F(\Xi)u_1\rangle+|\alpha_2|^2\langle u_2,F(\Xi)u_2\rangle-\overline{\alpha}_1\alpha_2\langle u_1,F(\Xi)u_2\rangle-\alpha_1\overline{\alpha}_2\langle u_2,F(\Xi)u_1\rangle\Big)\\
=&\frac{1}{2}\Big(|\alpha_1|^2\langle u_1,F(\Xi)u_1\rangle+|\alpha_2|^2\langle u_2,F(\Xi)u_2\rangle-2[\text{Real part}](\overline{\alpha}_1\alpha_2\langle u_1,F(\Xi)u_2\rangle)\Big)
\end{aligned}
$$

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[Real part]$(\overline{\alpha}_1\alpha_2\langle u_1, F(\Xi)u_2\rangle)$), we get the following graph.

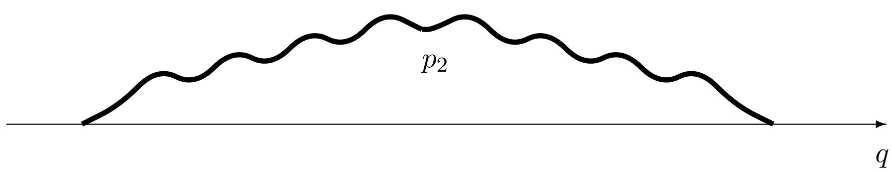


Figure 10.6(2): The graph of p_2

10.8.3 No interference

Consider the measurement:

$$
\mathsf{M}_{B(\mathbb{C}^2\otimes H)}(\mathsf{O}_x\otimes\mathsf{O},S_{[|\psi\rangle\langle\psi|]})\tag{10.36}
$$

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

$$
\begin{aligned}
&\langle\psi,(I\otimes F(\Xi))\psi\rangle\\
=&\langle\alpha_1e_1\otimes u_1+\alpha_2e_2\otimes u_2,(I\otimes F(\Xi))(\alpha_1e_1\otimes u_1+\alpha_2e_2\otimes u_2)\rangle\\
=&\langle\alpha_1e_1\otimes u_1+\alpha_2e_2\otimes u_2,\alpha_1e_1\otimes F(\Xi)u_1+\alpha_2e_2\otimes F(\Xi)u_2\rangle\\
=&|\alpha_1|^2\langle u_1,F(\Xi)u_1\rangle+|\alpha_2|^2\langle u_2,F(\Xi)u_2\rangle
\end{aligned}
$$

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.

$$p_3 = p_1 + p_2$$

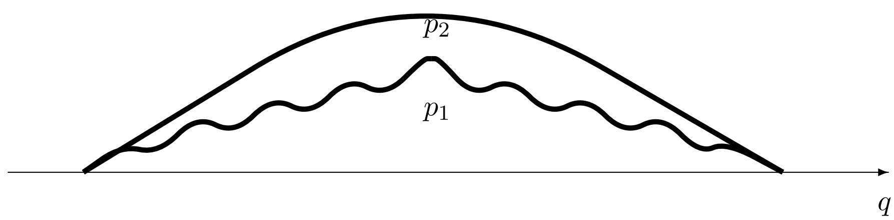


Figure 10.6(3): The graph of $p_3 = p_1 + p_2$

Remark 10.21. 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 [□□].

Chapter 11

Realized causal observable in general theory

What we have studied is

$$
\left\{
\begin{array}{l}
(\sharp_1): \underset{\text{(=quantum language)}}{\boxed{\text{pure measurement theory}}} \\
\quad := \underbrace{\underset{(cf. \ \S 2.7)}{\overset{[\text{(pure)Axiom 1}]}{\boxed{\text{pure measurement}}}} + \underset{(cf. \ \S 9.3)}{\overset{[\text{Axiom 2}]}{\boxed{\text{Causality}}}}}_{\text{a kind of spells (a priori judgment)}} + \underbrace{\underset{(cf. \ \S 3.1)}{\overset{[\text{quantum linguistic Copenhagen interpretation}]}{\boxed{\text{Linguistic Copenhagen interpretation}}}}}_{\text{manual to use spells}} \\
\\
(\sharp_2): \underset{\text{(=quantum language)}}{\boxed{\text{mixed measurement theory}}} \\
\quad := \underbrace{\underset{(cf. \ \S 8.1)}{\overset{[\text{(mixed)Axiom}^{(m)} \text{ 1}]}{\boxed{\text{mixed measurement}}}} + \underset{(cf. \ \S 9.3)}{\overset{[\text{Axiom 2}]}{\boxed{\text{Causality}}}}}_{\text{a kind of spells (a priori judgment)}} + \underbrace{\underset{(cf. \ \S 3.1)}{\overset{[\text{quantum linguistic Copenhagen interpretation}]}{\boxed{\text{Linguistic Copenhagen interpretation}}}}}_{\text{manual to use spells}}
\end{array}
\right.
$$

As mentioned in the previous chapter, what is important is

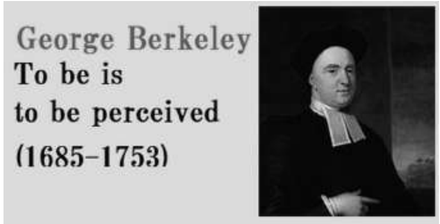
- *to exercise the relationship of measurement and causality.*

In this chapter, we discuss the relationship more systematically.

11.1 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 11.1.** 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. The readers should understand:

- "realized causal observable" is a direct consequence of the linguistic Copenhagen interpretation, that is,

Only one measurement is permitted.

Now we shall review the following theorem:

Theorem 11.1. [=Theorem 10.1: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 10.1 □

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 11.2. [(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 9.10) 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.

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 $[\mathsf{O}_T] = [\{\mathsf{O}_t\}_{t\in T}$, $\{\overline{\mathcal{A}}_t \xrightarrow{\Phi_{\pi(t),t}} \overline{\mathcal{A}}_{\pi(t)}\}_{t\in T\setminus\{t_0\}})]$.

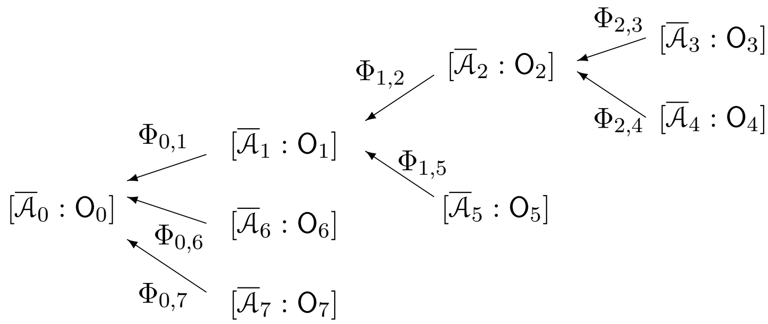


Figure 11.1 : Simple example of sequential causal observable

Now we can show our present problem.

Problem 11.3. 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 ?

Let us solve this problem as follows. Note that the linguistic Copenhagen interpretation says :

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 11.4. [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

$$\widehat{\mathsf{O}}_s = \begin{cases} \mathsf{O}_s & (\text{ if } s \in T \setminus \pi(T) \text{ }) \\ \\ \mathsf{O}_s \times (\times_{t\in\pi^{-1}(\{s\})} \Phi_{\pi(t),t}\widehat{\mathsf{O}}_t) & (\text{ if } s \in \pi(T) \text{ }) \end{cases} \tag{11.1}$$

(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$[\mathsf{O}_{T(t_0)}] = [\{\mathsf{O}_t\}_{t\in T}, \{\Phi_{\pi(t),t} : \overline{\mathcal{A}}_t \to \overline{\mathcal{A}}_{\pi(t)}\}_{t\in T\setminus\{t_0\}}]$.

Note that

(♯) 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 11.2.** In the above (11.1), 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 11.5. [A simple classical example] Suppose that a tree $(T \equiv \{0, 1, ..., 6, 7\}, \pi)$ has an ordered structure such that $\pi(1) = \pi(6) = \pi(7) = 0$, $\pi(2) = \pi(5) = 1$, $\pi(3) = \pi(4) = 2$.

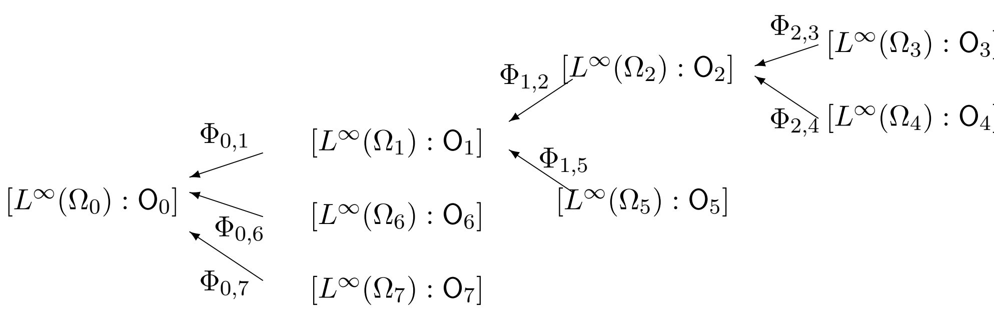


Figure 11.2 : Simple classical example of sequential causal observables

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:

$$\begin{array}{ccccc}
L^\infty(\Omega_0) & \overset{\Phi_{0,1}}{\longleftarrow} & L^\infty(\Omega_1)P & \overset{\Phi_{1,2}}{\longleftarrow} & L^\infty(\Omega_2) \\
F_0 \times \Phi_{0,6}\widehat{F}_6 \times \Phi_{0,7}\widehat{F}_7 & & F_1 \times \Phi_{1,5}\widehat{F}_5 & & \\
\downarrow & & \downarrow & & \\
\widehat{F}_0 & \overset{\Phi_{0,1}}{\longleftarrow} & \widehat{F}_1 & \overset{\Phi_{1,2}}{\longleftarrow} & \widehat{F}_2 \\
{\scriptstyle (F_0 \times \Phi_{0,6}\widehat{F}_6 \times \Phi_{0,7}\widehat{F}_7 \times \Phi_{0,1}\widehat{F}_1)} & & {\scriptstyle (F_1 \times \Phi_{1,5}\widehat{F}_5 \times \Phi_{1,2}\widehat{F}_2)} & & {\scriptstyle (F_2 \times \Phi_{2,3}\widehat{F}_3 \times \Phi_{2,4}\widehat{F}_4)}
\end{array}$$

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

$$\widehat{\mathsf{O}}_0 \equiv (\times_{t=0}^{7} X_t, \boxtimes_{t=0}^{7} \mathcal{F}_t, \widehat{F}_0 (= F_0 \times (\times_{t=1,6,7} \Phi_{0,t}\widehat{F}_t))$$

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 a 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

$$\begin{aligned}
&\widehat{F}_0(\Xi_0 \times \Xi_1 \times \Xi_2 \times \Xi_3 \times \Xi_4 \times \Xi_5 \times \Xi_6 \times \Xi_7)] \\
=&F_0(\Xi_0) \times \Phi_{0,1}\Big(F_1(\Xi_1) \times \Phi_{1,5}F_5(\Xi_5) \times \Phi_{1,2}\Big(F_2(\Xi_2) \times \Phi_{2,3}F_3(\Xi_3) \times \Phi_{2,4}F_4(\Xi_4)\Big)\Big) \\
&\times \Phi_{0,6}(F_6(\Xi_6)) \times \Phi_{0,7}(F_7(\Xi_7))
\end{aligned} \tag{11.2}$$

(In quantum case, the existence of $\widehat{\mathsf{O}}_0$ in not guaranteed). □

Remark 11.6. 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 $\mathsf{O}_t^{(\text{exi})}{=}(X_t, \{\emptyset, X_t\}, F_t^{(\text{exi})})$. Then, we see that

$$\begin{aligned}
&\widehat{F}_0(\Xi_0 \times \Xi_1 \times X_2 \times \Xi_3 \times \Xi_4 \times \Xi_5 \times X_6 \times X_7) \\
=&F_0(\Xi_0) \times \Phi_{0,1}\Big(F_1(\Xi_1) \times \Phi_{1,5}F_5(\Xi_5) \times \Phi_{1,2}\Big(\Phi_{2,3}F_3(\Xi_3) \times \Phi_{2,4}F_4(\Xi_4)\Big)\Big).
\end{aligned} \tag{11.3}$$

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

$$\begin{aligned}
&\widehat{F}'_0(\Xi_0 \times \Xi_1 \times \Xi_3 \times \Xi_4 \times \Xi_5) \\
=&F_0(\Xi_0) \times \Phi_{0,1}\Big(F_1(\Xi_1) \times \Phi_{1,5}F_5(\Xi_5) \times \Phi_{1,4}F_4(\Xi_4) \times \Phi_{1,3}F_3(\Xi_3) \times \Phi_{1,4}F_4(\Xi_4)\Big)
\end{aligned} \tag{11.4}$$

which is different from the fact (11.2). We may sometimes omit "existence observable". However, we have to do it with careful cautions.

Thus, we can answer Problem 11.3 as follows.

Problem 11.7. [=Problem 11.3] (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

For further information see my homepage

(B) 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

$$
{}_{\mathcal{A}_0^*}\Big(\rho_{t_0}, \widehat{F}_{t_0}(\widehat{\Xi})\Big)_{\overline{\mathcal{A}}_{t_0}} \tag{11.5}
$$

The following theorem, which holds in classical systems, is frequently used.

Theorem 11.8. [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

$$
[\widehat{F}_{t_0}(\mathop{\times}_{t\in T} \Xi_t)](\omega_{t_0}) = \mathop{\times}_{t\in T}[\Phi_{t_0,t}F_t(\Xi_t)](\omega_{t_0}) = \mathop{\times}_{t\in T}[F_t(\Xi_t)](\phi_{t_0,t}\omega_{t_0}).
$$

$$
(\forall \omega_{t_0} \in \Omega_{t_0}, \forall \Xi_t \in \mathcal{F}_t)
$$

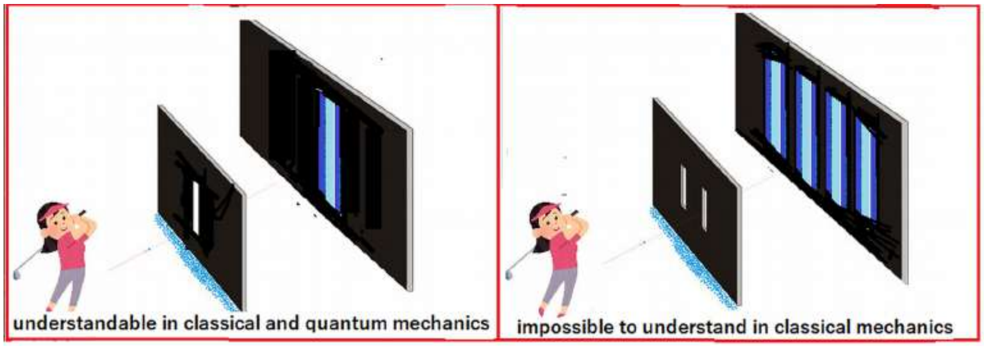
Proof. It suffices to prove the simple classical case of Example 11.5. Using Theorem 9.6 repeatedly, we see that

$$
\begin{aligned}
&\widehat{F}_0 = F_0 \times (\mathop{\times}_{t=1,6,7} \Phi_{0,t}\widehat{F}_t) \\
=&F_0 \times (\Phi_{0,1}\widehat{F}_1 \times \Phi_{0,6}\widehat{F}_6 \times \Phi_{0,7}\widehat{F}_7) = F_0 \times (\Phi_{0,1}\widehat{F}_1 \times \Phi_{0,6}F_6 \times \Phi_{0,7}F_7) \\
=&\Big(\mathop{\times}_{t=0,6,7} \Phi_{0,t}F_t\Big) \times (\Phi_{0,1}\widehat{F}_1) = \Big(\mathop{\times}_{t=0,6,7} \Phi_{0,t}F_t\Big) \times \Phi_{0,1}(F_1 \times (\mathop{\times}_{t=2,5} \Phi_{1,t}\widehat{F}_t)) \\
=&\Big(\mathop{\times}_{t=0,1,6,7} \Phi_{0,t}F_t\Big) \times \Phi_{0,1}(\mathop{\times}_{t=2,5} \Phi_{1,t}\widehat{F}_t) = \Big(\mathop{\times}_{t=0,1,6,7} \Phi_{0,t}F_t\Big) \times \Phi_{0,1}(\Phi_{1,2}\widehat{F}_2 \times \Phi_{1,5}\widehat{F}_5) \\
=&\Big(\mathop{\times}_{t=0,1,5,6,7} \Phi_{0,t}F_t\Big) \times \Phi_{0,1}(\Phi_{1,2}\widehat{F}_2) \\
=&\Big(\mathop{\times}_{t=0,1,5,6,7} \Phi_{0,t}F_t\Big) \times \Phi_{0,1}(\Phi_{1,2}(F_2 \times (\mathop{\times}_{t=3,4} \Phi_{2,t}\widehat{F}_t))) \\
=&\mathop{\times}_{t=0}^{7} \Phi_{0,t}F_t
\end{aligned}
$$

This completes the proof. □

For further information see my homepage

11.2 Double-slit experiment



11.2.1 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 11.9. $\mathsf{M}_{B(H_0)}(\Phi_{0,t_2}\mathsf{O}_2 = (\mathbb{R}, \mathcal{B}_{\mathbb{R}}, \Phi_{0,t_2}F_2), S_{[|u_0\rangle\langle u_0|]})$

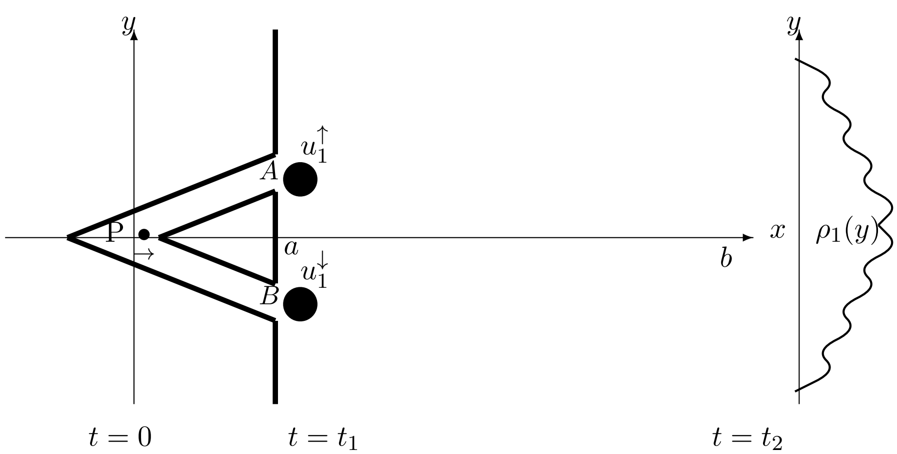


Figure 11.3(1) Potential $V(x, y) = \infty$ on the thick line, $= 0$ (elsewhere)

For further information see my homepage

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

$$\Phi_{s,t}A = e^{\frac{\mathcal{H}(t-s)}{i\hbar}}Ae^{-\frac{\mathcal{H}(t-s)}{i\hbar}} \qquad (\forall A \in B(H_t) = B(L^2(\mathbb{R}^2)))$$

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 11.3.** Precisely speaking, we say as follows. Let Δ, ϵ be small positive real numbers. For each $k \in \mathbb{Z} = \{k \mid k = 0, \pm1, \pm2, \pm3, , , , , \}$, 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 10.7.

11.2.2 Which-way path experiment

Picture 11.10. 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 $\{\uparrow\}\times(-\infty, y]$

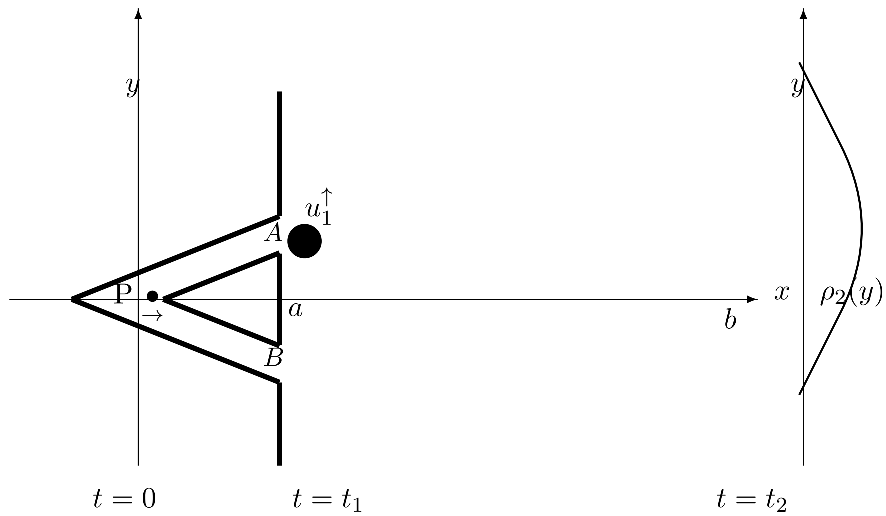


Figure 11.3(2) Potential $V(x,y)=\infty$ on the thick line, $=0$ (elsewhere)

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

$$[F_1(\{\uparrow\})](x,y)=\left\{\begin{array}{ll}1 & y\geq 0\\ 0 & y<0\end{array}\right.$$
$$[F_1(\{\downarrow\})](x,y)=1-[F_1(\{\uparrow\})](x,y)$$

According to Section 10.2 (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 11.4.** Precisely speaking, in the above case, it suffices to consider the following procedure (1) and (ii):

For further information see my homepage

(i) for time t_1, the projection observable O_1 is measured in the sense of Postulate 10.7

(ii) 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 Postulate 10.7.

11.3 Wilson cloud chamber in double slit experiment

In this section, we shall analyze a discrete trajectory of a quantum particle, which is assumed to be 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. [27, 28, (1991, 1994, S. Ishikawa, *et al.*)].

11.3.1 Trajectory of a particle is nonsense

We shall consider a particle P in the one-dimensional real line $\mathbb{R}$, whose initial wave 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 $u(x)$ (or precisely, $|u\rangle\langle u|$) as follows:

$$u(x) = \begin{cases} l/\sqrt{2}, x \in (-3/2, -1/2) \cup (1/2, 3/2) \\ \\ 0, \ \text{otherwise} \end{cases} \tag{11.6}$$

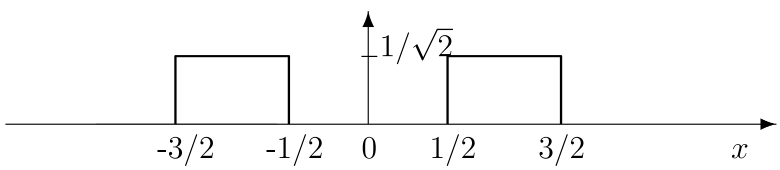


Figure 11.4 The initial wave function $u(x)$

Let A_0 be a position observable in H, that is,

$$(A_0 v)(x) = xv(x) \qquad (\forall x \in \mathbb{R}, \quad \text{for } v \in H = L^2(\mathbb{R}))$$

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,

$$-i\hbar \frac{dA_t}{dt} = \mathcal{H}A_t - A_t\mathcal{H}, \quad -\infty < t < \infty, \text{ where } A_0 = A. \tag{11.7}$$

The one-parameter unitary group U_t is defined by $\exp(-itA)$. An easy calculation shows that

$$A_t = U_t^* A U_t = U_t^* x U_t = x + \frac{\hbar t}{im}\frac{d}{dx}. \tag{11.8}$$

For further information see my homepage

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

$$\underset{\text{initial wave function:}u_0}{\overset{\text{position observable: } A_0}{\boxed{B(H_0)}}} \xleftarrow[\Phi_{0,1/4}]{\qquad} \overset{\text{position observable: } A_0}{\boxed{B(H_{1/4})}}$$

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 nonsense.

11.3.2 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 4.14, that is,

Definition 11.11. (=Definition 4.14). 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,

$$\langle u \otimes s, \widehat{A}(u \otimes s)\rangle = \langle u, Au\rangle, \qquad \langle u \otimes s, \widehat{B}(u \otimes s)\rangle = \langle u, Bu\rangle \qquad (11.9)$$
$$(\forall u \in H, \|u\|_H = 1)$$

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$_2$) $\Delta_{\widehat{N}_1}^{\widehat{\rho}_{us}}$ $(= \|(\widehat{A} - A \otimes I)(u \otimes s)\|)$ and $\Delta_{\widehat{N}_2}^{\widehat{\rho}_{us}}$ $(= \|(\widehat{B} - B \otimes I)(u \otimes s)\|)$ are called errors of the approximate simultaneous measurement $\mathsf{M}_{B(H \otimes K)}(\mathsf{O}_{\widehat{A}} \times \mathsf{O}_{\widehat{B}}, S_{[\widehat{\rho}_{us}]})$.

For further information see my homepage

Now, let us constitute approximate observables $(K, s, \widehat{A}, \widehat{B})$ 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

$$\langle s, As \rangle = \langle s, Bs \rangle = 0. \tag{11.10}$$

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 (11.10),

$$\langle u \otimes s, \widehat{A}(u \otimes s) \rangle = \langle u \otimes s, (A \otimes I + 2I \otimes A)(u \otimes s) \rangle = \langle u, Au \rangle, \tag{11.11}$$

$$\langle u \otimes s, \widehat{A}(u \otimes s) \rangle = \langle u \otimes s, (B \otimes I - 2I \otimes A)(u \otimes s) \rangle = \langle u, Bu \rangle. \tag{11.12}$$

$$(\forall u \in H, i = 1, 2)$$

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:

$$\delta_0 = \Delta_{\widehat{N}_1}^{\widehat{\rho}_{us}} = \|(\widehat{A} - A \otimes I)(u \otimes s)\| = \|2(I \otimes A)(u \otimes s)\| = 2\|As\| \tag{11.13}$$

$$\delta_{1/4} = \Delta_{\widehat{N}_2}^{\widehat{\rho}_{us}} = \|(\widehat{B} - B \otimes I)(u \otimes s)\| = (1/2)\|(I \otimes B)(u \otimes s)\| = (1/2)\|Bs\| \tag{11.14}$$

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.

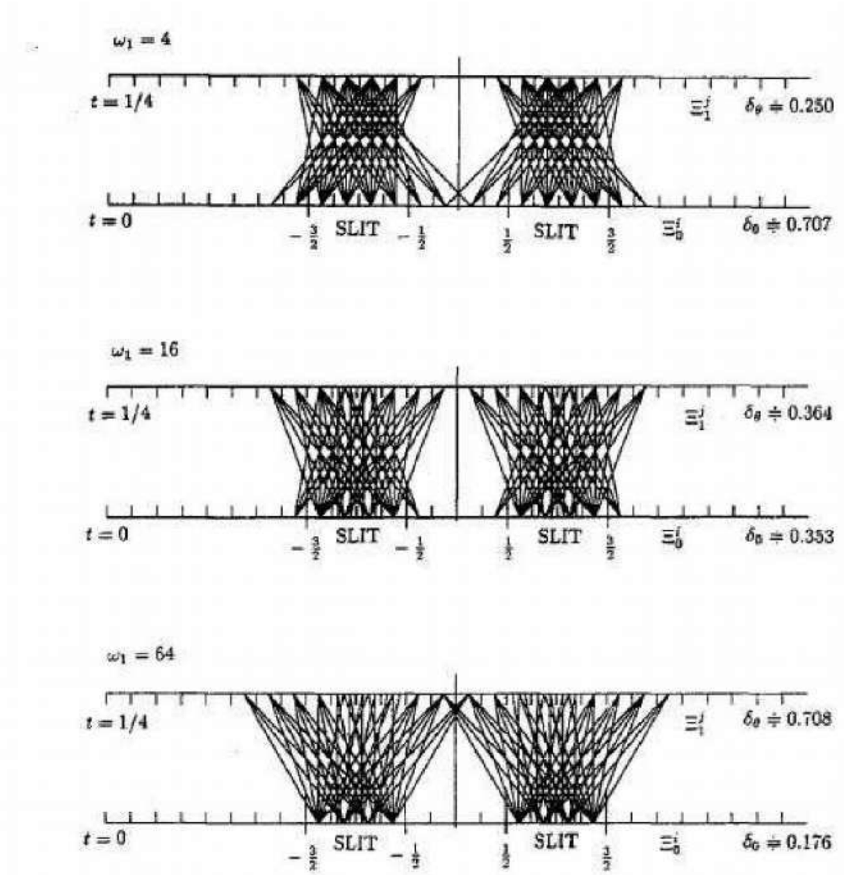


Figure 11.5: The lines connecting two points (i.e., x_k and x'_k)$(k = 1, 2, ...)$

Here, note that $\delta_\theta(= \delta_{1/4})$ and δ_0 are depend on ω_1.

♠**Note 11.5.** For further arguments, see the following references.

($\sharp_1$) [27]: S. Ishikawa, *Uncertainties and an interpretation of non-relativistic quantum theory,* International Journal of Theoretical Physics 30, 401–417 (1991)
doi: 10.1007/BF00670793

($\sharp_2$) [28]: 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

For further information see my homepage

Chapter 12

Why does statistics work? : Fisher statistics (II)

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

$$
\underset{\text{(=quantum language)}}{\boxed{\text{measurement theory}}} := \underbrace{\overset{\text{[Axiom 1]}}{\underset{(cf.\ \S 2.7)}{\boxed{\text{Measurement}}}} + \overset{\text{[Axiom 2]}}{\underset{(cf.\ \S 9.3)}{\boxed{\text{Causality}}}}}_{\text{a kind of spell(a priori judgment)}} + \underbrace{\overset{\text{[quantum linguistic Copenhagen interpretation]}}{\underset{(cf.\ \S 3.1)}{\boxed{\text{Linguistic Copenhagen interpretation}}}}}_{\text{manual how to use spells}}
$$

In Chapter 5 (Why does Fisher statistics work in our world? (I)), we discussed "inference" in relation to "measurement". In this chapter, we discuss "inference" in the relation to both "measurement" and "causality".

12.1 "Inference = Control" in quantum language

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

$$
\text{"inference"} = \text{"control"}
$$

In this sense, we conclude that statistics and dynamical system theory are essentially the same.

12.1.1 Inference problem (statistics)

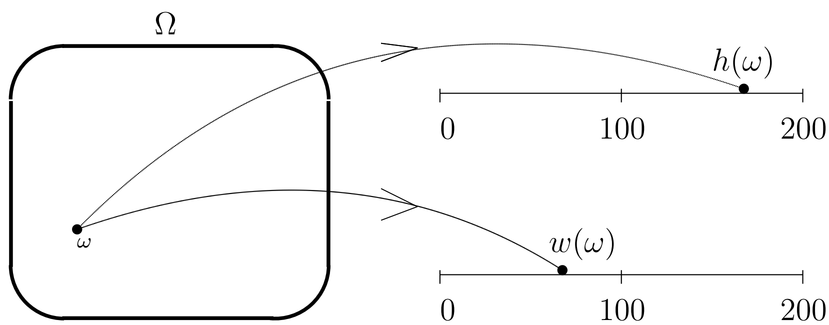
Problem 12.1. [Who is the high school student who saved the drowning girl?] Let $\Omega \equiv \{\omega_1, \omega_2, ..., \omega_{100}\}$ be a set of all students of a certain high school. Define $h : \Omega \to [0, 200]$ and $w : \Omega \to [0, 200]$ such that

$$\begin{aligned} h(\omega_n) = & \text{ "the height of a student } \omega_n\text{"} \quad (n = 1, 2, ..., 100) \\ w(\omega_n) = & \text{ "the weight of a student } \omega_n\text{"} \quad (n = 1, 2, ..., 100) \end{aligned} \tag{12.1}$$

For simplicity, put, $N = 5$. For example, see the following.

Table 12.1: 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 of 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 has is as follows:

(i) he is a student of the principal's high school.

(ii) his height [resp. weight] is about 165 cm [resp. about 65 kg].

(iii) Assume that the height and weight of high school students follow independent normal distributions $N(\mu_1, \sigma_1)$ and $N(\mu_2, \sigma_2)$, and further, assume that $\sigma_2/\sigma_1 = \sqrt{2}$ though it may not be natural.

Now we have the following question:

(b) Under the above assumption (a_1) and (a_2), how does the principal infer who he is.

This will be answered in Answer 12.3. ///

To answer this problem, we must prepare the following Theorem.

Theorem 12.2. 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 $[\{\mathsf{O}_t(= (X_t, \mathcal{F}_t, F_t))\}_{t\in T}, \{\Phi_{\pi(t),t} : L^\infty(\Omega_t) \to L^\infty(\Omega_{\pi(t)})\}_{t\in T\setminus\{t_0\}}]$. Thus, we have a measurement

$$\mathsf{M}_{L^\infty(\Omega_{t_0})}(\widehat{\mathsf{O}}_T{=}(\times_{t\in T} X_t, \boxtimes_{t\in T}\mathcal{F}_t, \widehat{F}_{t_0}), S_{[*]}).$$

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

$$[\widehat{F}_{t_0}(\widehat{\Xi})](\omega_{t_0}) = \max_{\omega\in\Omega_{t_0}} [\widehat{F}_{t_0}(\widehat{\Xi})](\omega).$$

(Fisher's maximum likelihood method).

///

The proof is a direct consequence of Axiom 2 (causality; §9.3) and Fisher maximum likelihood method (Theorem 5.6). Thus, we omit it.

Answer 12.3. [(Continued from Problem 12.1 (Inference problem))] Let $(T{=}\ \{0, 1, 2\}, \pi : T \setminus \{0\} \to T)$ be the parent map representation of a tree, where it is assumed that

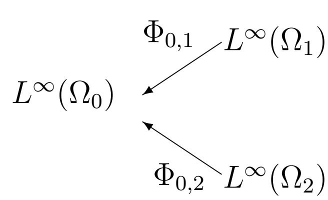
$$\pi(1) = \pi(2) = 0.$$

Put $\Omega_0 = \{\omega_1, \omega_2, \ldots, \omega_5\}$, $\Omega_1 =$ interval $[100, 200]$, $\Omega_2 =$ interval $[30, 110]$. Here, we consider that

$\Omega_0 \ni \omega_n \cdots\cdots$ a state such that "the girl is helped by a student ω_n" $\quad (n = 1, 2, ..., 5)$

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

$$[\Phi_{0,t} f_t](\omega) = f_t(\phi_{0,t}(\omega)) \qquad (\forall\omega \in \Omega_0, \forall f_t \in L^\infty(\Omega_t)).$$



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

$$\Xi_1 = \left[165 - \frac{1}{N}, 165 + \frac{1}{N}\right], \qquad \Xi_2 = \left[65 - \frac{1}{N}, 65 + \frac{1}{N}\right].$$

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, the measured value belongs to $\Xi_1 \times \Xi_2$. Using Theorem 12.2, we say:

($\sharp$) 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,

$$
\begin{aligned}
(\sharp) &\Longrightarrow \max_{\omega \in \Omega_0} \frac{1}{\sqrt{(2\pi)^2 \sigma_1^2 \sigma_2^2}} \iint_{\Xi_1 \times \Xi_2} \exp\,[-\frac{(x_1 - h(\omega))^2}{2\sigma_1^2} - \frac{(x_2 - w(\omega))^2}{2\sigma_2^2}] dx_1 dx_2 \\
&\Longrightarrow \max_{\omega \in \Omega_0} \exp\,[-\frac{(165 - h(\omega))^2}{2\sigma_1^2} - \frac{(65 - w(\omega))^2}{2\sigma_2^2}] \\
&\Longrightarrow \min_{\omega \in \Omega_0} [\frac{(165 - h(\omega))^2}{2\sigma_1^2} + \frac{(65 - w(\omega))^2}{4\sigma_1^2}] \quad \big(\ (a_2\text{:iii}) \text{ says that } 2\sigma_1^2 = \sigma_2^2\ \big) \\
&\Longrightarrow \text{When } \omega = \omega_4, \text{ minimum } 2(165 - 170)^2 + (65 - 60)^2 \text{ is attained} \\
&\Longrightarrow \text{The student is } \omega_4.
\end{aligned}
$$

Therefore, we can infer that the student who helps the girl is ω_4. □

12.1.2 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 (12.2). That is,

For further information see my homepage

$$\boxed{\text{dynamical system theory}} = \begin{cases} \text{(i)} : \frac{d\omega(t)}{dt} = v(\omega(t), t, e_1(t), \beta) & \cdots (\text{ state equation}) \\ \qquad \text{(initial } \omega(0)=\alpha) & \\ \text{(ii)} : x(t) = g(\omega(t), t, e_2(t)) & \cdots (\text{ measurement}) \end{cases} \tag{12.2}$$

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 12.4. [Control problem] We have a rectangular water tank filled with water.

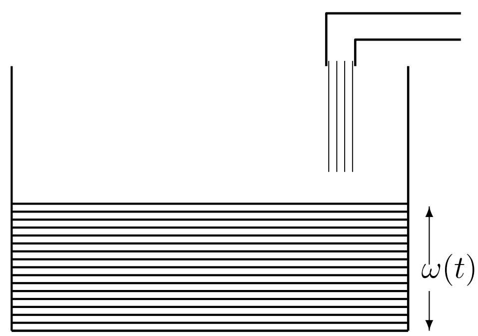


Figure 12.1: Water tank

Assume that the height of water at time t is given by the following function $\omega(t)$:

$$\frac{d\omega}{dt} = \beta_0, \text{ then } \omega(t) = \omega_0 + \theta t, \tag{12.3}$$

where ω_0 and θ are unknown fixed parameters such that ω_0 is the height of water filling the tank at the beginning and θ is the increasing height of water per unit time. The measured height $x(t)$ of water at time t is assumed to be represented by

$$x(t) = \omega_0 + \theta t + e(t),$$

where $e(t)$ represents a noise (or more precisely, a measurement error) with some suitable conditions. And assume that as follows:

$$x(1) = 1.9, \quad x(2) = 3.0, \quad x(3) = 4.7. \tag{12.4}$$

Under this setting, we consider the following problem:

(c_1) **[Control]:** Settle the state (ω_0, θ) such that measured data (12.4) will be obtained.

or, equivalently,

(c_2) **[Inference]:** when measured data (12.4) is obtained, infer the unknown state (ω_0, θ).

For further information see my homepage

This will be answered in Answer 12.8.

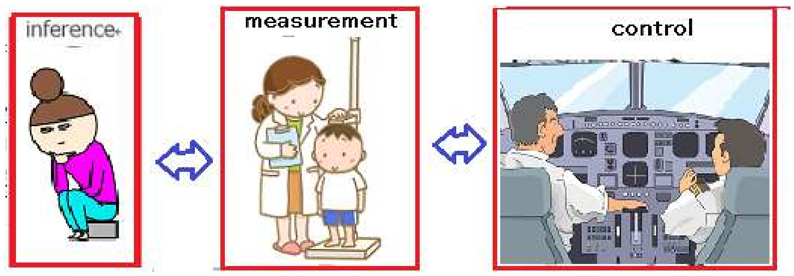
///

Note that

$$(c_1)=(c_2)$$

from a mathematical point of view. Thus, we consider :

(d) *Inference problem and control problem are the same problem. And these are characterized as the reverse problem of measurements. Thus, the three are essentially the same.*



Thus, statistics, measurement theory, dynamical system theory, control theory are essentially the same.

Remark 12.5. [Remark on dynamical system theory (*cf.* [35])] Again recall the formulation (12.2) of dynamical system theory, in which

(♯) 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, confusions between noises and measurement errors frequently occur. This weakness is clarified in quantum language, as shown in Answer 12.8.

12.2 [Parameter≈State] in QL

The following theorem is a slight extension of Theorem 12.2

Theorem 12.6. [Parameter≈State] in QL Let $(T=\{t_0, t_1, \ldots, t_N\}, \pi : T\setminus\{t_0\} \to T)$ be a tree. Let Θ be a (locally) compact set (i.e., parameter space), which is regarded as a kind of state space. For each $\theta(\in \Theta)$, consider a sequential causal observable $[\{\mathsf{O}_t\}_{t\in T}, \{\Phi^\theta_{\pi(t),t} : L^\infty(\Omega_t) \to L^\infty(\Omega_{\pi(t)})\}_{t\in T\setminus\{t_0\}}$]. Let $\widehat{\mathsf{O}}^\theta_T$ $=(\times_{t\in T} X_t,\ \boxtimes_{t\in T}\mathcal{F}_t,\ \widehat{F}^\theta_{t_0})$ be the realized causal observable of a sequential causal observable $[\{\mathsf{O}_t\}_{t\in T}, \{\Phi^\theta_{\pi(t),t} : L^\infty(\Omega_t) \to L^\infty(\Omega_{\pi(t)})\}_{t\in T\setminus\{t_0\}}$]. Consider a measurement

$$\mathsf{M}_{L^\infty(\Omega_{t_0})}(\widehat{\mathsf{O}}^\theta_T=(\mathop{\times}_{t\in T} X_t,\ \boxtimes_{t\in T}\mathcal{F}_t,\ \widehat{F}^\theta_{t_0}), S_{[*]}) \qquad (\theta\in\Theta)$$

which can be identified with the following.

$$\mathsf{M}_{L^\infty(\Omega_{t_0}\times\Theta)}(\widehat{\mathsf{O}}_T^\theta{=}(\mathop{\times}_{t\in T} X_t,\ \boxtimes_{t\in T}\mathcal{F}_t,\ \widehat{F}_{t_0}^\theta), S_{[(*_\Omega,*_\Theta)]})$$

. . . .

And Fisher's maximum likelihood method

Assume that a measured value obtained by the measurement belongs to $\widehat{\Xi}$ $(\in\ \boxtimes_{t\in T}\mathcal{F}_t)$. Then, Fisher's maximum likelihood method (Theorem 5.6) says that there is a reason to infer that

$$[\,*\,](=[*_{\Omega_0},*_\Theta])=(\omega_{t_0},\theta_0),$$

where (ω_{t_0},θ_0) $(\in \Omega_{t_0}\times\Theta)$ is defined by

$$[\widehat{F}_{t_0}(\widehat{\Xi})](\omega_{t_0},\theta_0)=\max_{(\omega,\theta)\in\Omega_{t_0}\times\Theta}[\widehat{F}_{t_0}(\widehat{\Xi})](\omega,\theta).$$

///

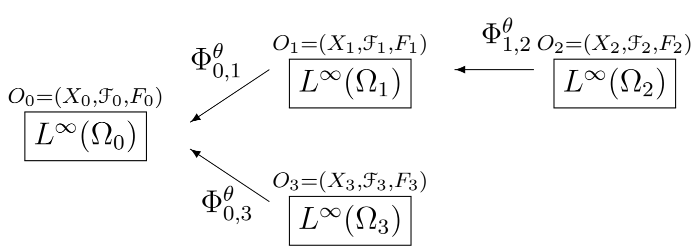
The proof is a direct consequence of Axiom 2 (causality; §9.3) and Fisher's maximum likelihood method (Theorem 5.6). Thus, we omit it.

♠**Note 12.1.** Perhaps the above should have been called a 'method' (or, 'one of the Copenhagen interpretation') rather than a 'theorem'. Even up to this point, we should recall that what is called a 'parameter' in statistics is called a 'state' in QL.

The above is too general, so consider the simple case as follows.

Corollary 12.7. [The simple form of Theorem 12.6]

Put $T=\{0,1,2,3\}$,



Thus, we get the realized causal observable:

$$\widehat{\mathsf{O}}_T^\theta = (\times_{t\in T} X_t,\ \boxtimes_{t\in T}\mathcal{F}_t, \widehat{F}_{t_0}^\theta) \quad \text{in } L^\infty(\Omega_0)$$

where

$$\widehat{F}_{t_0}^\theta = F_0(\Xi_0)\big[(\Phi_{0,3}^\theta F_3(\Xi_3))\Big(\Phi_{0,1}^\theta\big((F(\Xi_1)(\Phi_{1,2}^\theta F_2(\Xi_2))\big)\Big)\big]$$

Consider a measurement

$$\mathsf{M}_{L^\infty(\Omega_{t_0})}(\widehat{\mathsf{O}}_T^\theta = (\times_{t\in T} X_t,\ \boxtimes_{t\in T}\mathcal{F}_t, \widehat{F}_{t_0}^\theta), S_{[*]}) \qquad (\theta\in\Theta)$$

which can be identified with the following.

$$\mathsf{M}_{L^\infty(\Omega_{t_0}\times\Theta)}(\widehat{\mathsf{O}}_T^\theta = (\times_{t\in T} X_t,\ \boxtimes_{t\in T}\mathcal{F}_t, \widehat{F}_{t_0}^\theta), S_{[(*_\Omega, *_\Theta)]})$$

Assume that a measured value obtained by the measurement belongs to $\widehat{\Xi}$ $(\in \boxtimes_{t\in T}\mathcal{F}_t)$.

Then, Fisher's maximum likelihood method (Theorem 5.6) says that there is a reason to infer that

$$[\,*\,](=[*_{\Omega_0}, *_\Theta]) = (\omega_{t_0}, \theta_0),$$

where (ω_{t_0}, θ_0) $(\in \Omega_{t_0}\times\Theta)$ is defined by

$$[\widehat{F}_{t_0}(\widehat{\Xi})](\omega_{t_0}, \theta_0) = \max_{(\omega,\theta)\in\Omega_{t_0}\times\Theta} [\widehat{F}_{t_0}(\widehat{\Xi})](\omega, \theta).$$

///

♠**Note 12.2.** It should be noted that there is a consistent spirit of the linguistic Copenhagen interpretation of 'measurement only once' in Theorem 12.6.

Answer 12.8. [Continued from Problem 12.4 (Control problem); Theorem 12.6] Put $\Omega_0 = \Omega_1 = \Omega_2 = \Omega_3 = \mathbb{R}$. and put

$$\Omega_0 \ni \omega_0 \xmapsto{\phi_{01}} \omega_0 + \theta = \omega_1 \in \Omega_1$$

$$\Omega_1 \ni \omega_1 \xmapsto{\phi_{12}} \omega_1 + \theta = \omega_2 \in \Omega_2$$

$$\Omega_2 \ni \omega_2 \xmapsto{\phi_{23}} \omega_2 + \theta = \omega_3 \in \Omega_3$$

For further information see my homepage

Thus, we see:

$$
\underset{\boxed{L^\infty(\Omega_0)}}{O_0{=}(X_0,\mathcal{F}_0,F_0)} \xleftarrow{\Phi^\theta_{0,1}} \underset{\boxed{L^\infty(\Omega_1)}}{O_1{=}(X_1,\mathcal{F}_1,F_1)} \xleftarrow{\Phi^\theta_{1,2}} \underset{\boxed{L^\infty(\Omega_2)}}{O_2{=}(X_2,\mathcal{F}_2,F_2)} \xleftarrow{\Phi^\theta_{2,3}} \underset{\boxed{L^\infty(\Omega_3)}}{O_3{=}(X_3,\mathcal{F}_3,F_3)}
$$

where $\mathsf{O}_0 = (X_0, \mathcal{F}_0, F_0)$ is the existence observable (*cf.*Definition 2.20), so, it can be neglected. Also, $\mathsf{O}_0 = \mathsf{O}_1 = \mathsf{O}_2 = \mathsf{O}_3$ is the normal observable O_{G_σ} with a standard deviation σ, i.e., $\mathsf{O}_{G_\sigma}{=}(\mathbb{R}, \mathcal{B}_\mathbb{R}, G_\sigma)$ where

$$
[G_\sigma(\Xi)](\omega) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_\Xi e^{-\frac{(x-\omega)^2}{2\sigma^2}} dx \quad (\forall \Xi \in \mathcal{B}_\mathbb{R}, \forall \omega \in \Omega_t).
$$

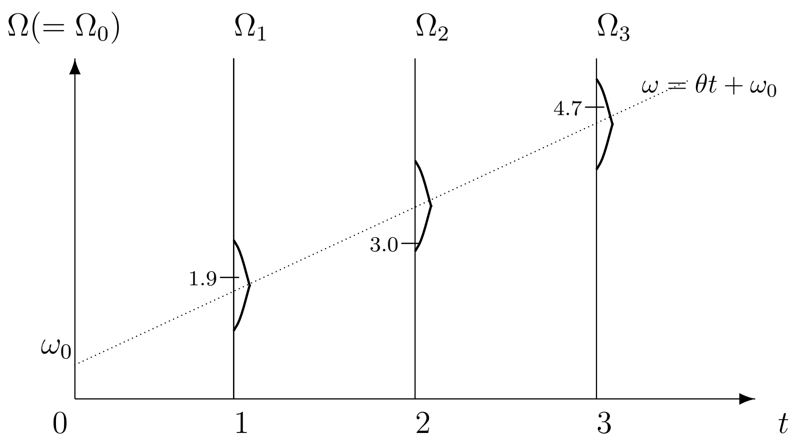


Figure 12.2 Problem: Find the equation $\omega = \theta t + \omega_0$ of the dashed line

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 11.8)

$$
\begin{aligned}
&[\widehat{F}_0(\Xi_1 \times \Xi_2 \times \Xi_3)](\omega_0) = \big[\Phi_{0,1}\big(G_\sigma(\Xi_1)\Phi_{1,2}(G_\sigma(\Xi_2)\Phi_{2,3}(G_\sigma(\Xi_3)))\big)\big](\omega_0) \\
=&[\Phi_{0,1}G_\sigma(\Xi_1)](\omega_0) \cdot [\Phi_{0,2}G_\sigma(\Xi_2)](\omega_0) \cdot [\Phi_{0,3}G_\sigma(\Xi_3)](\omega_0) \\
=&[G_\sigma(\Xi_1)](\phi_{0,1}(\omega_0)) \cdot [G_\sigma(\Xi_2)](\phi_{0,2}(\omega_0)) \cdot [G_\sigma(\Xi_3)](\phi_{0,3}(\omega_0)) \\
=&[G_\sigma(\Xi_1)](\omega_0 + \theta)) \cdot [G_\sigma(\Xi_2)](\omega_0 + 2\theta)) \cdot [G_\sigma(\Xi_3)](\omega_0 + 3\theta)) \\
&\qquad (\forall \Xi_1, \Xi_2, \Xi_3 \in \mathcal{B}_\mathbb{R},\ \forall \omega_0, \theta \in \Omega_0 \times \Theta)
\end{aligned}
$$

Our problem (i.e., Problem 12.4) is as follows,

($\sharp_1$) Find the parameter (θ, ω_0) (i.e.,$\mathsf{M}_{L^\infty(\Omega_0)}(\ \widehat{\mathsf{O}}^\theta_T,\ S_{[\omega_0]})\ )$ that is most likely to yield the measured value $(1.9, 3.0, 4.7)$.

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 (ω_0, θ) $(\in \Omega_0 \times \Theta)$ such that

$$[\widehat{F}_0(\Xi_1 \times \Xi_2 \times \Xi_3)](\omega_0, \theta) = \max_{(\omega_0,\theta)} [\widehat{F}_0(\Xi_1 \times \Xi_2 \times \Xi_3)].$$

Since N is assumed to be sufficiently large, we see

$$\begin{aligned}
(\sharp_2) \Longrightarrow & \max_{(\omega_0,\theta)\in\Omega_0} [\widehat{F}_0(\Xi_1 \times \Xi_2 \times \Xi_3)](\omega_0, \theta) \\
\Longrightarrow & \max_{(\omega_0,\theta)\in\Omega_0} \frac{1}{\sqrt{2\pi\sigma^2}^3} \iiint_{\Xi_1\times\Xi_2\times\Xi_3} e^{[-\frac{(x_1-(\omega_0+\theta))^2+(x_2-(\omega_0+2\theta))^2+(x_3-(\omega_0+3\theta))^2}{2\sigma^2}]} \\
& \qquad \times dx_1 dx_2 dx_3 \\
\Longrightarrow & \max_{(\omega_0,\theta)\in\Omega_0} \exp(-J/(2\sigma^2)) \\
\Longrightarrow & \min_{(\omega_0,\theta)\in\Omega_0} J
\end{aligned}$$

where

$$J = (1.9 - (\omega_0 + \theta))^2 + (3.0 - (\omega_0 + 2\theta))^2 + (4.7 - (\omega_0 + 3\theta))^2.$$

($\frac{\partial}{\partial\omega_0}\{\cdots\} = 0, \frac{\partial}{\partial\theta}\{\cdots\} = 0$)

$$\begin{aligned}
\Longrightarrow & \begin{cases} (1.9 - (\omega_0 + \theta)) + (3.0 - (\omega_0 + 2\theta)) + (4.7 - (\omega_0 + 3\theta)) = 0 \\ (1.9 - (\omega_0 + \theta)) + 2(3.0 - (\omega_0 + 2\theta)) + 3(4.7 - (\omega_0 + 3\theta)) = 0 \end{cases} \\
\Longrightarrow & \quad (\omega_0, \theta) = (0.4, 1.4)
\end{aligned}$$

Therefore, in order to obtain a measured value (1.9, 3.0, 4.7), it suffices to put

$$(\omega_0, \theta) = (0.4,\ 1.4).$$

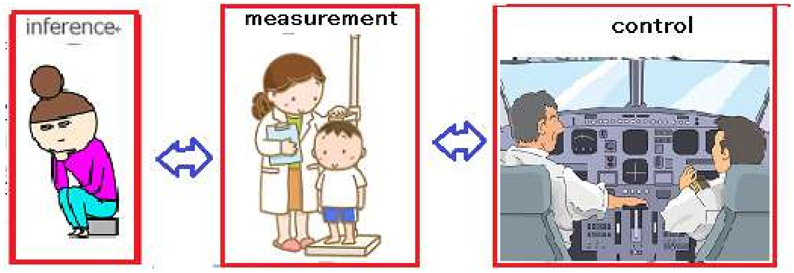
□

For completeness, note that,

- From a theoretical point of view,

$$\text{“\textit{inference}”} = \text{“\textit{control}”}$$

Thus, we conclude that statistics and dynamical system theory are essentially the same.



♠**Note 12.3.** Comparing Sec. 6.4 (Regression) and Answer 12.8, we may say the Theorem 12.6 is a kind of the generalization of regression analysis. I have previously overemphasized this. This emphasis caused confusion among readers, so I will not emphasize it in this publication.

Chapter 13

Least-squares method and Regression analysis

Although regression analysis has a history of great achievements, it seems to have been wrongly understood in essence. For example, the fundamental terms in regression analysis (e.g., "regression", "least-squares method", "explanatory variable", "response variable", etc.) are historical conventions, and do not express their roles adequately in the regression analysis. In this chapter, we show that the least squares method acquires a right position in quantum language as follows.

$$
\underset{\text{(Section 13.1)}}{\boxed{\text{The least squares method}}} \xrightarrow[\text{quantum language}]{\text{describe by}} \underset{\text{(Section 13.2)}}{\boxed{\text{Regression analysis}}} \xrightarrow[\text{generalization}]{\text{natural}} \underset{\text{(Section 13.3)}}{\boxed{\text{Generalized linear model}}} \qquad (\sharp)
$$

In this story, the terms "explanatory variable" and "response variable" are clarified in the framework of quantum language. To develop a general theory of regression analysis, it suffices to work with Theorem 12.6. However, from a practical point of view, we need the above scheme ($\sharp$). This chapter is extracted from

> Ref. [54]: S. Ishikawa; Regression analysis in quantum language
> arxiv:1403.0060[math.ST], (2014)

13.1 The least squares method

Let us start from a 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

$$
\mathbb{R} \ni a \mapsto x = \phi^{(\beta_1,\beta_2)}(a) = \beta_1 a + \beta_0 \in \mathbb{R}. \tag{13.1}
$$

where the pair $(\beta_1, \beta_2)(\in \mathbb{R}^2)$ is assumed to be unknown. Define the error σ by

$$
\sigma^2(\beta_1, \beta_2) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \phi^{(\beta_1,\beta_2)}(a_i))^2 \Big(= \frac{1}{n}\sum_{i=1}^{n}(x_i - (\beta_1 a_i + \beta_0))^2 \Big). \tag{13.2}
$$

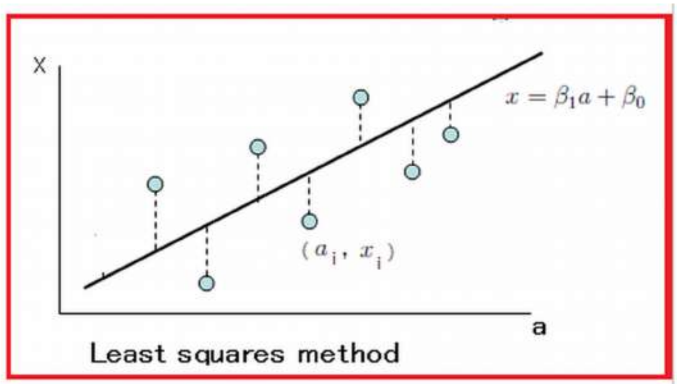
Then, we have the following minimization problem:

Problem 13.1. [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

$$\sigma^2(\hat{\beta}_0, \hat{\beta}_1) = \min_{(\beta_0,\beta_1)\in\mathbb{R}^2} \sigma^2(\beta_0, \beta_1)\Big(= \min_{(\beta_0,\beta_1)\in\mathbb{R}^2} \frac{1}{n}\sum_{i=1}^n (x_i - (\beta_1 a_i + \beta_0))^2\Big), \tag{13.3}$$

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"),

$$\frac{\partial \sigma^2(\beta_0, \beta_1)}{\partial \beta_0} = \sum_{i=1}^n (x_i - \beta_0 - \beta_1 a_i) = 0, \quad (i = 1, ..., n), \tag{13.4}$$

$$\frac{\partial \sigma^2(\beta_0, \beta_1)}{\partial \beta_1} = \sum_{i=1}^n (x_i - \beta_0 - \beta_1 a_i) a_i = 0, \quad (i = 1, ..., n). \tag{13.5}$$

Solving it, we get that

$$\hat{\beta}_1 = \frac{s_{ax}}{s_{aa}}, \quad \hat{\beta}_0 = \overline{x} - \frac{s_{ax}}{s_{aa}}\overline{a}, \quad \hat{\sigma}^2\Big(= \frac{1}{n}\sum_{i=1}^n (x_i - (\hat{\beta}_1 a_i + \hat{\beta}_0))^2\Big) = s_{xx} - \frac{s_{ax}^2}{s_{aa}}, \tag{13.6}$$

where

$$\bar{a} = \frac{a_1 + \cdots + a_n}{n}, \qquad \bar{x} = \frac{x_1 + \cdots + x_n}{n}, \tag{13.7}$$

$$s_{aa} = \frac{(a_1 - \bar{a})^2 + \cdots + (a_n - \bar{a})^2}{n}, \quad s_{xx} = \frac{(x_1 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}{n}, \tag{13.8}$$

$$s_{ax} = \frac{(a_1 - \bar{a})(x_1 - \bar{x}) + \cdots + (a_n - \bar{a})(x_n - \bar{x})}{n}. \tag{13.9}$$

For further information see my homepage

♠**Note 13.1.** [Applied mathematics]. Note that the above result is in (applied) mathematics, that is,

- the above is neither in statistics nor in quantum language.

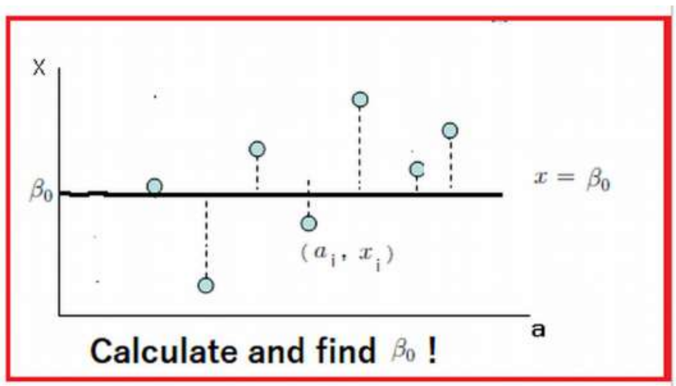
The purpose of this chapter is to add a quantum linguistic story to Problem 13.1 (i.e., the least-squares method).

13.2 Regression analysis

13.2.1 The simplest problem

Let us start from the simplest problem.

Problem 13.2. [The simplest problem].
[**(I): Applied math**]
Let $\{(a_i, x_i)\}_{i=1}^n$ be a sequence in the two dimensional real space $\mathbb{R}^2$.



Find the $\hat{\beta}_0$ $(\in \mathbb{R})$ such that

$$\sigma^2(\hat{\beta}_0) = \min_{(\beta_0) \in \mathbb{R}} \sigma^2(\beta_0) \Big(= \min_{(\beta_0) \in \mathbb{R}} \frac{1}{n} \sum_{i=1}^n (x_i - \beta_0)^2 \Big),$$

Of course, it is easy. That is,

$$\hat{\beta}_0 = \frac{x_1 + x_2 + \ldots + x_n}{n} \tag{$*$}$$

[**(II): The argument in QL**]
It should be noted that this problem is similar to the inference problem of the simultaneous normal measurement (in Example 5.10): $\mathsf{M}_{L^\infty(\mathbb{R}\times\mathbb{R}_+)}(\mathsf{O}^n = (\mathbb{R}^n, \mathcal{B}_{\mathbb{R}}^n, G^n), S_{[*]})$, where

$$
\begin{aligned}
&[G^m(\Xi_1 \times \Xi_2 \times \cdots \times \Xi_n)](\omega) \\
=&[(\mathop{\times}_{k=1}^n G)(\Xi_1 \times \Xi_2 \times \cdots \times \Xi_n)](\omega) = \mathop{\times}_{k=1}^n [G(\Xi_k)](\omega)
\end{aligned}
$$

$$= \bigtimes_{k=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} \int_{\Xi_k} \exp\Big[-\frac{1}{2\sigma^2}(x_k-\mu)^2 \Big] dx_k$$

$$(\forall \Xi_k \in \mathcal{B}_X (= \mathcal{B}_{\mathbb{R}}),\ \forall \omega = (\mu, \sigma) \in \Omega (= \mathbb{R} \times \mathbb{R}_+))$$

Recall that Fisher's maximum likelihood method (Theorem5.6) says that the unknown state $[*] = (\mu, \sigma)$ $(\in \mathbb{R} \times \mathbb{R}_+)$ is inferred as follows.

$$\mu = \overline{\mu}(x) = \frac{x_1 + x_2 + \ldots + x_n}{n}, \qquad (**)$$

$$\sigma = \overline{\sigma}(x) = \sqrt{\frac{\sum_{k=1}^{n}(x_k - \overline{\mu}(x))^2}{n}}.$$

[(III): The purpose of this chapter]
The above (i.e., $(*)$=$(**)$) is easy. However, our purpose of this chapter is to investigate a quantum linguistic understanding of Problem 13.1 just like the above [(I) and [(II)].

13.2.2 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

$$\tau(i) = 0 \qquad (\forall i = 1, 2, \cdots, n). \tag{13.10}$$

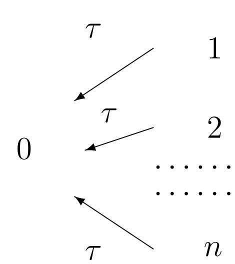


Figure 13.1: Parallel structure

♠**Note 13.2.** In regression analysis, we usually deal with "classical deterministic causal relation". Thus, Theorem 11.8 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

$$\Omega_0 = \mathbb{R}^2 = \Big\{ \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} \ :\ \beta_0, \beta_1 \in \mathbb{R} \Big\}, \tag{13.11}$$

$$\Omega_i = \mathbb{R} = \Big\{ \mu_i \ :\ \mu_i \in \mathbb{R} \Big\} \quad (i = 1, 2, \cdots, n) \tag{13.12}$$

where the Lebesgue measures m_i are assumed.

For further information see my homepage

Assume that

$$a_i \in \mathbb{R} \qquad (i = 1, 2, \cdots, n), \tag{13.13}$$

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

$$\Omega_0 = \mathbb{R}^2 \ni \beta = (\beta_0, \beta_1) \mapsto \psi_{a_i}(\beta_0, \beta_1) = \beta_0 + \beta_1 a_i = \mu_i \in \Omega_i = \mathbb{R} \tag{13.14}$$

which is equivalent to the deterministic causal operator $\Psi_{a_i} : L^\infty(\Omega_i) \to L^\infty(\Omega_0)$ such that

$$[\Psi_{a_i}(f_i)](\omega_0) = f_i(\psi_{a_i}(\omega_0)) \quad (\forall f_i \in L^\infty(\Omega_i), \ \ \forall \omega_0 \in \Omega_0, \forall i \in 1, 2, \cdots, n). \tag{13.15}$$

Thus, under the identification: $a_i \Leftrightarrow \psi_{a_i} \Leftrightarrow \Psi_{a_i}$, the term *"explanatory variable"* means a kind of causal relation Ψ_{a_i}.

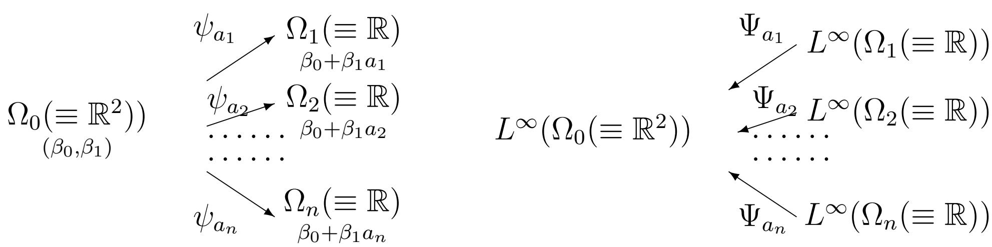


Figure 13.2: Parallel structure (Causal map ψ_{a_i}, Causal operator Ψ_{a_i})

For each $i = 1, 2, \cdots, n$, define *normal observables* $\mathsf{O}_i{\equiv}(\mathbb{R}, \mathcal{B}_{\mathbb{R}}, G_\sigma)$ in $L^\infty(\Omega_i(\equiv \mathbb{R}))$ such that

$$[G_\sigma(\Xi)](\mu) = \frac{1}{(\sqrt{2\pi\sigma^2})} \int_\Xi \exp\Big[-\frac{(x-\mu)^2}{2\sigma^2}\Big] dx \qquad (\forall \Xi \in \mathcal{B}_{\mathbb{R}}, \forall \mu \in \Omega_i(\equiv \mathbb{R})) \tag{13.16}$$

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

$$[\Psi_{a_i}(G_\sigma(\Xi))](\beta) = [(G_\sigma(\Xi))](\psi_{a_i}(\beta)) = \frac{1}{(\sqrt{2\pi\sigma^2})} \int_\Xi \exp\Big[-\frac{(x-(\beta_0+a_i\beta_1))^2}{2\sigma^2}\Big] dx \tag{13.17}$$

$$(\forall \Xi \in \mathcal{B}_{\mathbb{R}}, \forall \beta = (\beta_0, \beta_1) \in \Omega_0(\equiv \mathbb{R}^2)$$

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

$$\begin{aligned}
&[(\mathop{\times}_{i=1}^n \Psi_{a_i} G_\sigma)(\mathop{\times}_{i=1}^n \Xi_i)](\beta) = \mathop{\times}_{i=1}^n \Big([\Psi_{a_i} G_\sigma)(\Xi_i)](\beta)\Big) \\
=&\frac{1}{(\sqrt{2\pi\sigma^2})^n} \int \cdots \int_{\times_{i=1}^n \Xi_i} \exp\Big[-\frac{\sum_{i=1}^n (x_i - (\beta_0 + a_i\beta_1))^2}{2\sigma^2}\Big] dx_1 \cdots dx_n \\
=&\int \cdots \int_{\times_{i=1}^n \Xi_i} p_{(\beta_0,\beta_1,\sigma)}(x_1, x_2, \cdots, x_n) dx_1 \cdots dx_n.
\end{aligned} \tag{13.18}$$

For further information see my homepage

$$(\forall \bigtimes_{i=1}^{n} \Xi_i \in \mathcal{B}_{\mathbb{R}^n}, \forall \beta = (\beta_0, \beta_1) \in \Omega_0(\equiv \mathbb{R}^2))$$

Assuming that σ is a 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

$$[F(\bigtimes_{i=1}^{n} \Xi_i)](\beta, \sigma) = [(\bigtimes_{i=1}^{n} \Psi_{a_i} G_\sigma)(\bigtimes_{i=1}^{n} \Xi_i)](\beta) \quad (\forall \Xi_i \in \mathcal{B}_{\mathbb{R}}, \forall (\beta, \sigma) \in \mathbb{R}^2(\equiv \Omega_0) \times \mathbb{R}_+). \tag{13.19}$$

Problem 13.3. [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,

- Infer the β_0, β_1, σ from the measured value $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$.

That is, represent $(\beta_0, \beta_1, \sigma)$ by $(\hat{\beta}_0(x), \hat{\beta}_1(x), \hat{\sigma}(x))$ as 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

$$\frac{\partial L}{\partial \beta_0} = 0 \quad \Longrightarrow \quad \sum_{i=1}^{n} (x_i - (\beta_0 + a_i \beta_1)) = 0 \tag{13.20}$$

$$\frac{\partial L}{\partial \beta_1} = 0 \quad \Longrightarrow \quad \sum_{i=1}^{n} a_i (x_i - (\beta_0 + a_i \beta_1)) = 0 \tag{13.21}$$

$$\frac{\partial L}{\partial \sigma^2} = 0 \quad \Longrightarrow -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^{n} (x_i - \beta_0 - \beta_1 a_i)^2 = 0 \tag{13.22}$$

Therefore, using the notations (13.7)-(13.9), we obtain that

$$\hat{\beta}_0(x) = \overline{x} - \hat{\beta}_1(x) \overline{a} = \overline{x} - \frac{s_{ax}}{s_{aa}} \overline{a}, \quad \hat{\beta}_1(x) = \frac{s_{ax}}{s_{aa}} \tag{13.23}$$

and

$$(\hat{\sigma}(x))^2 = \frac{\sum_{i=1}^{n} \Big(x_i - (\hat{\beta}_0(x) + a_i \hat{\beta}_1(x))\Big)^2}{n}$$

$$
\begin{aligned}
&= \frac{\sum_{i=1}^n \left(x_i - (\overline{x} - \frac{s_{ax}}{s_{aa}}\overline{a}) - a_i \frac{s_{ax}}{s_{aa}} \right)^2}{n} = \frac{\sum_{i=1}^n \left((x_i - \overline{x}) + (\overline{a} - a_i)\frac{s_{ax}}{s_{aa}} \right)^2}{n} \\
&= s_{xx} - 2 s_{ax} \frac{s_{ax}}{s_{aa}} + s_{aa} (\frac{s_{ax}}{s_{aa}})^2 = s_{xx} - \frac{s_{ax}^2}{s_{aa}}.
\end{aligned}
\tag{13.24}
$$

Note that the above (13.23) and (13.24) are the same as (13.6). Therefore, Problem 13.3 (i.e., regression analysis in quantum language) is a quantum linguistic story of the least squares method (Problem 13.1).

Remark 13.4. Again, note that

(A) the least squares method (13.6) and the regression analysis (13.23) and (13.24) are the same.

Therefore, a small mathematical technique (the least squares method) can be understood in a grand story of regression analysis in quantum language. The readers may think that

(B) *Why do we choose "complicated (Problem 13.3)" rather than "simple (Problem 13.1)" approaches ?*

Of course, such a reason is unnecessary for quantum language ! That is because

(C) *the spirit of quantum language says*

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. The discussion of 'confidence interval and hypothesis testing' is omitted in this book, see refs. [62, 64].

13.3 Generalized linear model

Put $T = \{0, 1, 2, \cdots, i, \cdots, n\}$, which is the same as the tree (13.10), that is,

$$
\tau(i) = 0 \qquad (\forall i = 1, 2, \cdots, n). \tag{13.25}
$$

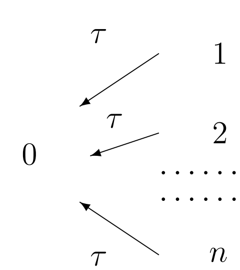


Figure 13.3: Parallel structure

For each $i \in T$, define a locally compact space Ω_i such that

$$\Omega_0 = \mathbb{R}^{m+1} = \left\{ \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_m \end{bmatrix} \;:\; \beta_0, \beta_1, \cdots, \beta_m \in \mathbb{R} \right\} \tag{13.26}$$

$$\Omega_i = \mathbb{R} = \left\{ \mu_i \;:\; \mu_i \in \mathbb{R} \right\} \quad (i = 1, 2, \cdots, n). \tag{13.27}$$

Assume that

$$a_{ij} \in \mathbb{R} \qquad (i = 1, 2, \cdots, n, \;\; j = 1, 2, \cdots, m, (m+1 \le n)). \tag{13.28}$$

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

$$\Omega_0 = \mathbb{R}^{m+1} \ni \beta = (\beta_0, \beta_1, \cdots, \beta_m) \mapsto$$
$$\psi_{a_{i\bullet}}(\beta_0, \beta_1, \cdots, \beta_m) = \beta_0 + \sum_{j=1}^{m} \beta_j a_{ij} = \mu_i \in \Omega_i = \mathbb{R} \tag{13.29}$$
$$(i = 1, 2, \cdots, n)$$

Summing up, we see

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix} \mapsto \begin{bmatrix} \psi_{a_{1\bullet}}(\beta_0, \beta_1, \cdots, \beta_m) \\ \psi_{a_{2\bullet}}(\beta_0, \beta_1, \cdots, \beta_m) \\ \psi_{a_{3\bullet}}(\beta_0, \beta_1, \cdots, \beta_m) \\ \vdots \\ \psi_{a_{n\bullet}}(\beta_0, \beta_1, \cdots, \beta_m) \end{bmatrix} = \begin{bmatrix} 1 & a_{11} & a_{12} & \cdots & a_{1m} \\ 1 & a_{21} & a_{22} & \cdots & a_{2m} \\ 1 & a_{31} & a_{32} & \cdots & a_{3m} \\ 1 & a_{41} & a_{42} & \cdots & a_{4m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} \cdot \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{bmatrix} \tag{13.30}$$

which is equivalent to the deterministic Markov operator $\Psi_{a_{i\bullet}} : L^\infty(\Omega_i) \to L^\infty(\Omega_0)$ such that

$$[\Psi_{a_{i\bullet}}(f_i)](\omega_0) = f_i(\psi_{a_{i\bullet}}(\omega_0)) \quad (\forall f_i \in L^\infty(\Omega_i), \;\; \forall \omega_0 \in \Omega_0, \forall i \in 1, 2, \cdots, n). \tag{13.31}$$

Thus, under the identification: $\{a_{ij}\}_{j=1,\cdots,m} \Leftrightarrow \Psi_{a_{i\bullet}}$, the term "explanatory variable" means a kind of causality.

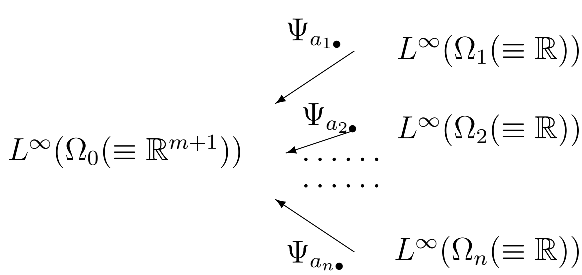


Figure 1.4: Parallel structure(Causal relation $\Psi_{a_{i\bullet}}$)

Therefore, we have an 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

$$[\Psi_{a_{i\bullet}}(G_\sigma(\Xi))](\beta) = [(G_\sigma(\Xi))](\psi_{a_{i\bullet}}(\beta))$$

$$= \frac{1}{(\sqrt{2\pi\sigma^2})} \int_{\Xi} \exp\Big[- \frac{(x - (\beta_0 + \sum_{j=1}^m a_{ij}\beta_j))^2}{2\sigma^2} \Big] dx. \tag{13.32}$$

$$(\forall \Xi \in \mathcal{B}_{\mathbb{R}}, \forall \beta = (\beta_0, \beta_1, \cdots, \beta_m) \in \Omega_0 (\equiv \mathbb{R}^{m+1}))$$

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

$$[(\mathop{\times}_{i=1}^n \Psi_{a_{i\bullet}} G_\sigma)(\mathop{\times}_{i=1}^n \Xi_i)](\beta) = \mathop{\times}_{i=1}^n \Big([\Psi_{a_{i\bullet}} G_\sigma)(\Xi_i)](\beta) \Big)$$

$$= \frac{1}{(\sqrt{2\pi\sigma^2})^n} \int \cdots \int_{\times_{i=1}^n \Xi_i} \exp\Big[- \frac{\sum_{i=1}^n (x_i - (\beta_0 + \sum_{j=1}^m a_{ij}\beta_j))^2}{2\sigma^2} \Big] dx_1 \cdots dx_n. \tag{13.33}$$

$$(\forall \mathop{\times}_{i=1}^n \Xi_i \in \mathcal{B}_{\mathbb{R}^n}, \forall \beta = (\beta_0, \beta_1, \cdots, \beta_m) \in \Omega_0 (\equiv \mathbb{R}^{m+1}))$$

Assuming that σ is a variable, we have an 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

$$[F(\mathop{\times}_{i=1}^n \Xi_i)](\beta, \sigma) = [(\mathop{\times}_{i=1}^n \Psi_{a_{i\bullet}} G_\sigma)(\mathop{\times}_{i=1}^n \Xi_i)](\beta)$$

$$(\forall \mathop{\times}_{i=1}^n \Xi_i \in \mathcal{B}_{\mathbb{R}^n}, \forall (\beta, \sigma) \in \mathbb{R}^{m+1} (\equiv \Omega_0) \times \mathbb{R}_+). \tag{13.34}$$

Thus, we have the following problem.

Problem 13.5. [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,

Infer $\beta_0, \beta_1, \cdots, \beta_m, \sigma$ from the measured value $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$

or

Represent $(\beta_0, \beta_1, \cdots, \beta_m, \sigma)$ by $(\hat{\beta}_0(x), \hat{\beta}_1(x), \cdots, \beta_m(x), \hat{\sigma}(x))$ as functions of x.

The answer is easy, since it is a slight generalization of Problem 13.3. Also, it suffices to follow ref. [8]. However, note that the purpose of this chapter is to propose Problem 13.5 (i.e., the quantum linguistic formulation of the generalized linear model) and not to give the answer to Problem 13.5.

For further information see my homepage

Remark 13.6. As a generalization of regression analysis, we also see measurement error model (*cf.* §5.5 (117 page) in ref. [35]), That is, we have two different generalizations such as

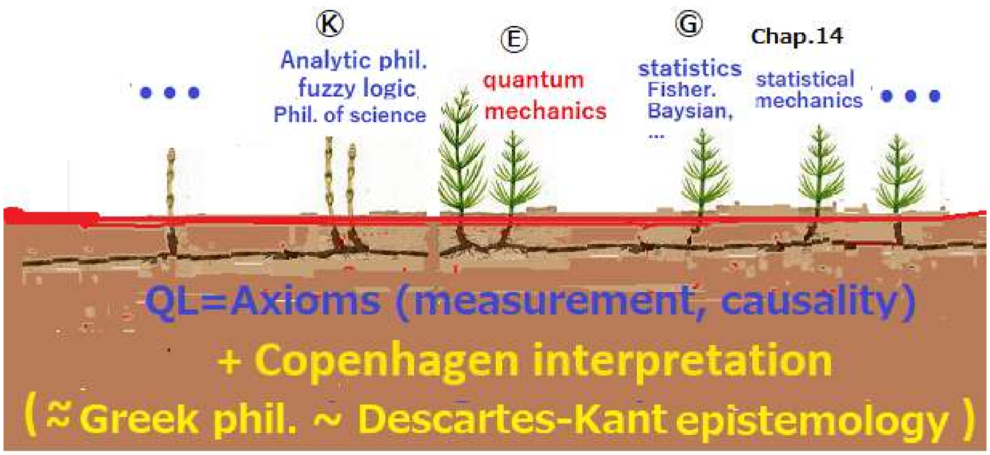
$$\boxed{\text{Regression analysis}} \xrightarrow[\text{generalization}]{} \begin{cases} ① : \boxed{\text{generalized linear model}} \\ ② : \boxed{\text{measurement error model}} \end{cases} \tag{13.35}$$

However, we think that ① is natural as the generalization of regression analysis (*cf.* ref. [35]).

Chapter 14

Equilibrium statistical mechanics

This chapter propose the quantum linguistic formulation of statistical mechanics as follows.



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

$$\begin{cases} \text{the kinetic aspect (i.e., causality)} & \cdots \text{ in Section 14.1} \\ \text{the probabilistic aspect (i.e., measurement)} & \cdots \text{ in Section 14.2} \end{cases}$$

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:

[45] S. Ishikawa, "Ergodic Hypothesis and Equilibrium Statistical Mechanics in the Quantum Mechanical World View," WJM, Vol. 2, No. 2, 2012, pp. 125-130. doi: 10.4236/wim.2012.22014, or ref. [38], ref. [62](Ver.5; Chap.17).

14.1 Equilibrium statistical mechanical phenomena concerning Axiom 2 (causality)

14.1.1 Equilibrium statistical mechanical phenomena

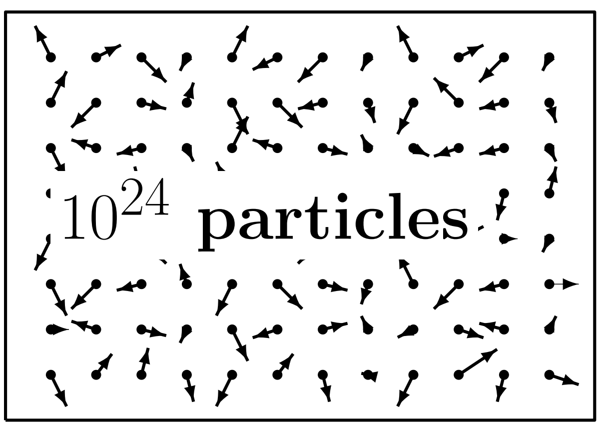
Hypothesis 14.1. [Equilibrium statistical mechanical hypothesis]. Assume that about $N(\approx 10^{24} \approx 6.02 \times 10^{23} \approx$ "the Avogadro constant") particles (for example, hydrogen molecules) move in a box with about 20 liters. It is natural to assume the following phenomena ① − ④:

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.



(14.1)

♠**Note 14.1.** Let me illustrate the above ② – ④ with a simple 'metaphor'. Suppose that 100 kindergarten children play on swings, slides and sand in a kindergarten yard during a one-hour lunch break. Assume, however, that there are enough swings, slides and sandboxes for all of them and that there is no queueing time. The, the above ② – ④ can be illustrated by the following 'metaphor'.

② All the kindergartners are bored and change their play one after the other. For example, one of the preschoolers played as follows.

(♯) [Swing] (5min) → [Slide] (3min) → [Sand] (6min) → [Slide] (7min) → [Swing] (9min) → [Sand] (8min) → [Swing] (9min) → [Sand] (6min) → [Swing] (7min)

For example, no children play only on the swings during the lunch break.

③ All the children have the same preferences. Therefore, the total duration of each of the three play activities is the same for all children. For example, every child are as follows.

Total time spent playing on the swings	30min
Total time spent playing on the slides	18min
Total time spent playing in the sandpit	12min

④ All children play with a spirit of "independence and self-respect". In other words, they are rarely influenced by the play of other children. For example, they do not act in groups, such as playing on the swings, then the slide, with other close friends.

You can read the following by imagining this②–④. .

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).

14.1.2 About ① in Hypothesis 14.1

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

$$\mathcal{H}\big((q_{1n}, q_{2n}, q_{3n}, p_{1n}, p_{2n}, p_{3n})_{n=1}^N\big) = \text{momentum energy} + \text{potential energy}$$

$$=[\sum_{n=1}^{N} \sum_{k=1,2,3} \frac{(p_{kn})^2}{2 \times \text{particle's mass}}]+U((q_{1n}, q_{2n}, q_{3n})_{n=1}^{N}). \tag{14.2}$$

Fix a positive $E > 0$. And define the measure ν_E on the energy surface Ω_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 Ω_E. Let $\{\psi_t^E\}_{-\infty<t<\infty}$ be the flow on the energy surface Ω_E induced by the Newton equation with the Hamiltonian $\mathcal{H}$, or equivalently, Hamilton's canonical equation:

$$\frac{dq_{kn}}{dt} = \frac{\partial \mathcal{H}}{\partial p_{kn}}, \quad \frac{dp_{kn}}{dt} = -\frac{\partial \mathcal{H}}{\partial q_{kn}}, \qquad (k=1,2,3, \quad n=1,2,\ldots,N). \tag{14.3}$$

Liouville's theorem (*cf.*[86]) says that the measure ν_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 Ω_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 (14.3), 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 9.4).

14.1.3 About ② in Hypothesis 14.1

Now let us begin with the well-known ergodic theorem (*cf.* [86]). 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 $S_{P_1} \subsetneq \Omega_E$. 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.* ref. [86]) says that the above ②′ is equivalent to the following equality:

$$\underbrace{\int_{\Omega_E} f(\omega)\overline{\nu}_E(d\omega)}_{\text{((state) space average)}} = \lim_{T\to\infty} \frac{1}{T} \underbrace{\int_\alpha^{\alpha+T} f(\psi_t^E(\omega_0))dt}_{\text{(time average)}} \tag{14.4}$$

$$(\forall \alpha \in \mathbb{R}, \forall f \in C(\Omega_{E}), \quad \forall \omega_0 \in \Omega_{E})$$

After all, the ergodic property ②′ ($\Leftrightarrow$ (14.4)) says that if T is sufficiently large, it holds that

$$\int_{\Omega_E} f(\omega)\overline{\nu}_{E}(d\omega) \approx \frac{1}{T}\int_{\alpha}^{\alpha+T} f(\psi_t^{E}(\omega_0))dt. \tag{14.5}$$

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 Copenhagen interpretation (§3.1) :*there is no probability without measurement*).

14.1.4 About ③ and ④ in Hypothesis 14.1

Put $K_N = \{1, 2, \dots, N(\approx 10^{24})\}$. For each k ($\in K_N$), define the coordinate map $\pi_k : \Omega_{E}(\subset \mathbb{R}^{6N}) \to \mathbb{R}^6$ such that

$$\begin{aligned}\pi_k(\omega) = \pi_k(q,p) =& \pi_k((q_{1n}, q_{2n}, q_{3n}, p_{1n}, p_{2n}, p_{3n})_{n=1}^N) \\ =&(q_{1k}, q_{2k}, q_{3k}, p_{1k}, p_{2k}, p_{3k})\end{aligned} \tag{14.6}$$

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 ($\subseteq K_N = \{1, 2, \dots, N\ (\approx 10^{24})\}$), 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

$$X_n^{\omega_0}(t) = \pi_n(\psi_t^{E}(\omega_0)) \qquad (\forall t \in [0,T]). \tag{14.7}$$

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:

$$\begin{aligned}&\overline{m}_{T}(\{t \in [0,T]\ :\ X_n^{\omega_0}(t) \in \Xi\}) \approx \rho_{E}(\Xi) \\ &(\forall \Xi \in \mathcal{B}_{\mathbb{R}^6}, n = 1, 2, \dots, N)\end{aligned} \tag{14.8}$$

④′ $\{X_n^{\omega_0}\}_{n=1}^N$ is *approximately independent*, in the sense that, for any $K_0 \subset \{1, 2, \dots, N(\approx 10^{24})\}$ such that $1 \leqq \sharp[K_0] \ll N$ (that is, $\frac{\sharp[K_0]}{N} \approx 0$), it holds that

$$\begin{aligned}&\overline{m}_{T}(\{t \in [0,T] : X_k^{\omega_0}(t) \in \Xi_k (\in \mathcal{B}_{\mathbb{R}^6}), k \in K_0\}) \\ \approx& \mathop{\times}_{k \in K_0} \overline{m}_{T}(\{t \in [0,T] : X_k^{\omega_0}(t) \in \Xi_k (\in \mathcal{B}_{\mathbb{R}^6})\}).\end{aligned}$$

For further information see my homepage

Here, we can assert the advantage of our method in comparison with Ruelle's method (*cf*.ref. [104]) as follows.

Remark 14.2. [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 (14.7) and (14.5), that, for $K_0(\subseteq K_N)$ such that $1\le \sharp[K_0]\ll N$,

$$
\begin{aligned}
&\overline{m}_T(\{t\in[0,T]\ :\ X_k^{\omega_0}(t)\in \Xi_k(\in \mathcal{B}_{\mathbb{R}^6}),k\in K_0\})\\
=&\overline{m}_T(\{t\in[0,T]:\pi_k(\psi_t^E(\omega_0)\in \Xi_k(\in \mathcal{B}_{\mathbb{R}^6}),k\in K_0\})\\
=&\overline{m}_T(\{t\in[0,T]:\psi_t^E(\omega_0)\in((\pi_k)_{k\in K_0})^{-1}(\mathop{\times}_{k\in K_0}\Xi_k)\})\\
\approx&\,\overline{\nu}_E\big(((\pi_k)_{k\in K_0})^{-1}(\mathop{\times}_{k\in K_0}\Xi_k)\big)\\
\equiv&\big(\overline{\nu}_E\circ((\pi_k)_{k\in K_0})^{-1}\big)(\mathop{\times}_{k\in K_0}\Xi_k).
\end{aligned}
\tag{14.9}
$$

Particularly, putting $K_0=\{k\}$, we see:

$$
\overline{m}_T(\{t\in[0,T]\ :\ X_k^{\omega_0}(t)\in\Xi\})\approx(\overline{\nu}_E\circ\pi_k^{-1})(\Xi)\quad(\forall\Xi\in\mathcal{B}_{\mathbb{R}^6}).
\tag{14.10}
$$

Hence, we can describe the ③ and ④ in terms of $\{\pi_k\}$ in what follows.

Hypothesis 14.3. [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 (14.9) 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

$$
\bigotimes_{k\in K_0}\rho_E(=\text{"product measure"})\approx\overline{\nu}_E\circ((\pi_k)_{k\in K_0})^{-1}.
\tag{14.11}
$$

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 $D_{K_N}^{(q,p)}\approx\rho_E$.

14.1.5 Ergodic Hypothesis

Now, we have the following theorem (*cf*.ref. [45]):

Theorem 14.4. [Ergodic hypothesis]. Assume Hypothesis 14.3 (or equivalently, ③ and ④). Then, for any $\omega_0 = (q(0), p(0)) \in \Omega_E$, it holds that

$$[D_{K_N}^{(q(t),p(t))}](\Xi) \approx \overline{m}_T(\{t \in [0,T] \; : \; X_k^{\omega_0}(t) \in \Xi\}) \\ (\forall \Xi \in \mathcal{B}_{\mathbb{R}^6}, k = 1, 2, \ldots, N(\approx 10^{24})) \tag{14.12}$$

for almost all t. That is, $0 \leqq \overline{m}_T(\{t \in [0,T] : (14.12) \text{ does not hold}\}) \ll 1$.

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.* ref. [85]) says that

$$D_{K_0}^{(q(t),p(t))} \approx \overline{\nu}_E \circ \pi_k^{-1} \; (\approx \rho_E \;) \tag{14.13}$$

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 (14.13), it holds that, for each k $(= 1, 2, \ldots, N \; (\approx 10^{24}))$,

$$\begin{aligned} D_{K_N}^{(q(t),p(t))} &= \frac{1}{N} \sum_{l=1}^{L} [\sharp[K_{(l)}] \times D_{K_{(l)}}^{(q(t),p(t))}] \\ &\approx \frac{1}{N} \sum_{l=1}^{L} [\sharp[K_{(l)}] \times \rho_E] \approx \overline{\nu}_E \circ \pi_k^{-1} \; (\approx \rho_E \;), \end{aligned} \tag{14.14}$$

for almost all time t. Thus, by (14.10), we get (14.12). Hence, the proof is completed.

We believe that Theorem 14.4 is just what should be represented by the *"ergodic hypothesis"* such that

"population average of N particles at each t"
="time average of one particle".

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 (14.5)) 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 14.5. [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 Ω_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.

For further information see my homepage

14.2 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 14.1. Also, Recall the linguistic Copenhagen interpretation (§3.1) : *there is no probability without measurement.* Note that the (14.12) 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

$$[F_0(\Xi)](q,p) = [D^{(q,p)}_{K_N}](\Xi)\Big(\equiv \frac{\sharp[\{k \mid \pi_k(q,p) \in \Xi\}]}{\sharp[K_N]}\Big)$$
$$(\forall \Xi \in \mathcal{B}_{\mathbb{R}^6}, \forall (q,p) \in \Omega_{_E}(\subset \mathbb{R}^{6N})). \tag{14.15}$$

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 $\mathsf{M}_{L^\infty(\Omega_E)}(\mathsf{O}_0 := (\mathbb{R}^6, \mathcal{B}_{\mathbb{R}^6}, F_0), S_{[\delta_{\psi_t(}$ belongs to $\Xi(\in \mathcal{B}_{\mathbb{R}^6})$ is given by $\rho_{_E}(\Xi)$. That is because Theorem 14.4 says that $[F_0(\Xi)](\psi_t(q_{_0}, p_{_0}))$ $\approx \rho_{_E}(\Xi)$ (almost every time t).

Also, let $\Psi^E_t : L^\infty(\Omega_{_E}) \to L^\infty(\Omega_{_E})$ be a deterministic Markov operator determined by the continuous map $\psi^E_t : \Omega_{_E} \to \Omega_{_E}$ (*cf.* Section 14.1.2). Then, it clearly holds $\Psi^E_t \mathsf{O}_0 = \mathsf{O}_0$. 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 Copenhagen 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 14.6. [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 $\Omega_{_E}$". 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 [104] in terms of measurement theory, we must use mixed measurement theory (*cf.* Chapter 8). However, this trial will end in failure.

14.3 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 [104] (which has been authorized for a long time)"

And, we assert the superiority of our method to Ruele's method in Remarks 14.2, 14.5, 14.6.

For further information see my homepage

Chapter 15

How to describe "belief"

Recall the spirit of quantum language (i.e., the spirit of the quantum mechanical world view), that is,

(♯) 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 (Sec. 8.6) in

- Ref. [35]: S. Ishikawa, "Mathematical Foundations of Measurement Theory," Keio University Press Inc. 2006.

15.1 Belief, probability and odds

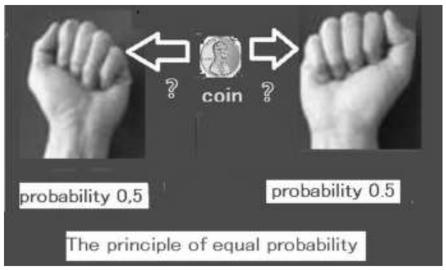
For instance, we want to formulate the following "probability":

(A) the "probability" that Japan will win the victory in the next FIFA World Cup.

This is possible (*cf.* [35]), 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 8, we studied the mixed measurement: that is,

$$
\underset{\text{(=quantum language)}}{\boxed{\text{mixed measurement theory}}} := \underbrace{\overset{[\text{(mixed)Axiom}^{(m)}\ 1]}{\underset{(cf.\ \S 8.1)}{\boxed{\text{mixed measurement}}}} + \overset{[\text{Axiom 2}]}{\underset{(cf.\ \S 9.3)}{\boxed{\text{Causality}}}}}_{\text{a kind of spells (a priori judgment)}}
$$
$$
+ \underbrace{\overset{[\text{quantum linguistic Copenhagen interpretation}]}{\underset{(cf.\ \S 3.1)}{\boxed{\text{Linguistic Copenhagen interpretation}}}}}_{\text{manual to use spells}} \tag{15.1}
$$

The purpose of this chapter is to characterize "belief" as a kind of mixed measurement.

15.1.1 A simple example; how to describe "belief" in quantum language

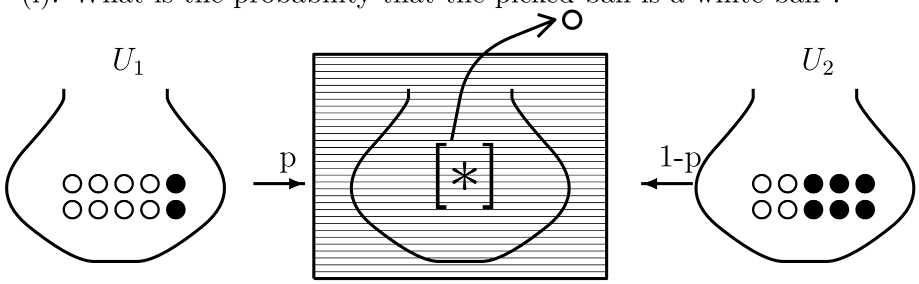
We begin with a simplest example (*cf.* Problem 8.5) as follows.

Problem 15.1. [= Problem 8.5; 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 15.1:(Mixed measurement)

Answer 15.2. (=Answer 8.13)
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 $\rho^{(p)}_{\text{prior}} (\in \mathcal{M}_{+1}(\Omega))$ such that

$$\rho^{(p)}_{\text{prior}} = p\delta_{\omega_1} + (1-p)\delta_{\omega_2},$$

where δ_ω is the point measure at ω. Thus, we have the mixed measurement:

$$\mathsf{M}_{C(\Omega)}(\mathsf{O} \times \mathsf{O}_U := (\{W,B\} \times \{U_1, U_2\}, 2^{\{W,B\}\times\{U_1,U_2\}}, F \times G_U), S_{[*]}(\rho^{(p)}_{\text{prior}})). \tag{15.2}$$

Axiom$^{(m)}$ 1 gives the answer to the (i) in Problem 15.1 as follows.

(D) the probability that a measured value (x, y) obtained by the mixed measurement $\mathsf{M}_{C(\Omega)}(\mathsf{O} \times \mathsf{O}_U, S_{[*]}(\rho^{(p)}_{\text{prior}}))$ belongs to $\{W\} \times \{U_1, U_2\}$ is given by

$$_{\mathcal{M}(\Omega)}(\rho^{(p)}_{\text{prior}}, F(\{W\}))_{C(\Omega)} = 0.8p + 0.4(1-p).$$

Since a white ball is obtained, Answer 8.13 (=Bayes' theorem) says that a new mixed state $\rho^{(p)}_{\text{post}} (\in \mathcal{M}_{+1}(\Omega))$ is given by

$$\rho^{(p)}_{\text{post}} = \frac{F(\{W\})\rho^{(p)}_{\text{prior}}}{\int_\Omega [F(\{W\})](\omega)\rho^{(p)}_{\text{prior}}(d\omega)} = \frac{0.8p}{0.8p + 0.4(1-p)}\delta_{\omega_1} + \frac{0.4(1-p)}{0.8p + 0.4(1-p)}\delta_{\omega_2} \tag{15.3}$$

Hence, the answer of the (ii) is given by

$$\mathcal{M}(\Omega)(\rho_{\text{post}}^{(p)}, G_U(\{U_1\}))_{C(\Omega)} = \frac{0.8p}{0.8p + 0.4(1-p)}.$$

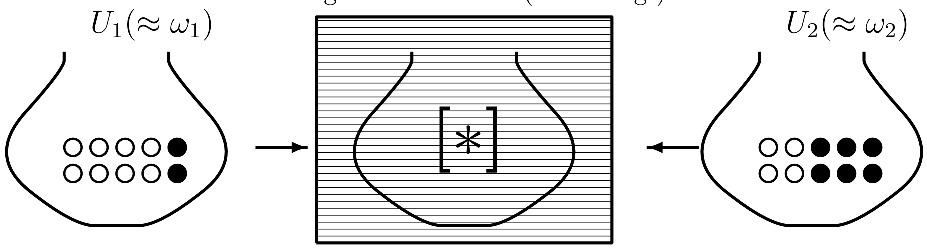
By an analogy of the above Problem 15.1 (for simplicity, we put: $p = 1/4$), we consider as follows.

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 15.1:

Figure 15.2: Belief (or voting)



25 people believe that $[*] = U_1$, 75 people believe that $[*] = U_2$.

Now, we have the following problem:

Problem 15.3. Consider Situation (E) and Situation (C) ($p = 1/4$, $1 - p = 3/4$). Then,

(F$_1$) Can Situation (E) be understood like Situation (C) ?

or, in the same sense,

(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 ?

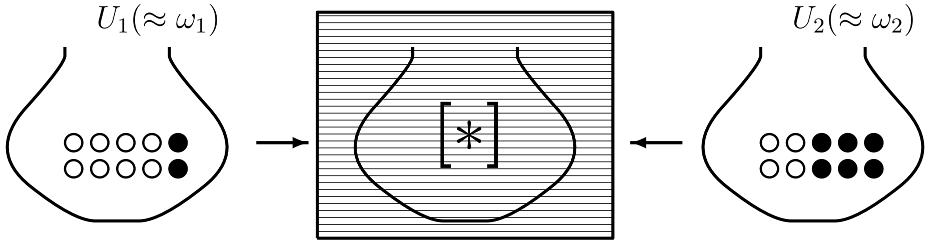
15.1.2 The affirmative answer to Problem 15.3

Since 100 people know the situation of the urn (i.e., Figure 15.2, the assumption (E)) implies (G)(=Figure 15.3), that is,

$$\textbf{(G)} \begin{cases} \text{25 people (in 100 people) believe that } [*] = U_1 \\ \quad \Longrightarrow \begin{cases} \text{(G}_1\text{): 20 people guess (or bet) that a white ball will be picked} \\ \text{(G}_2\text{): 5 people guess (or bet) that a black ball will be picked} \end{cases} \\ \text{75 people (in 100 people) believe that } [*] = U_2 \\ \quad \Longrightarrow \begin{cases} \text{(G}_3\text{): 30 people guess (or bet) that a white ball will be picked} \\ \text{(G}_4\text{): 45 people guess (or bet) that a black ball will be picked} \end{cases} \end{cases}$$

For further information see my homepage

Figure 15.3: The odds in bookmaker



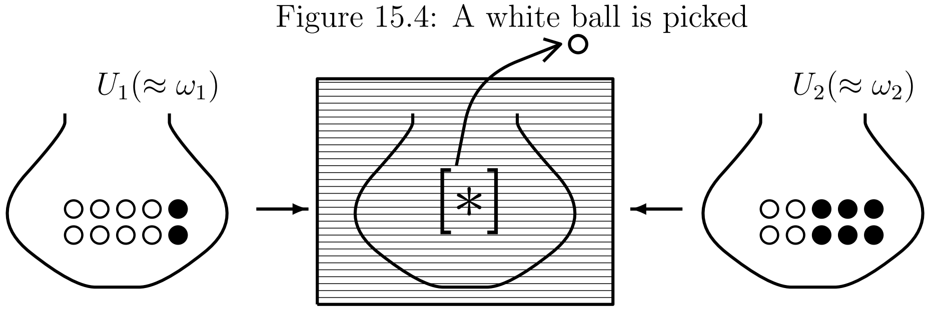
25 people believe that $[*] = U_1$. 75 people believe that $[*] = U_2$.

(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 15.4: A white ball is picked



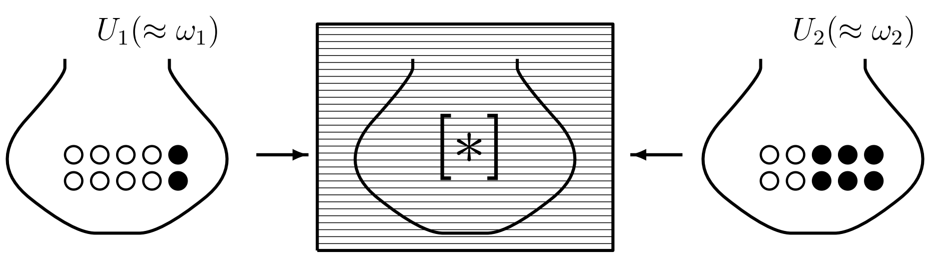
25 people believe that $[*] = U_1$. 75 people believe that $[*] = U_2$.

(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.~~

After all, we get the following figure:

Figure 15.5: After all, we get the new odds



40 % people believe that $[*] = U_1$, 60 % people believe that $[*] = U_2$.

Thus, we see that

$$\underset{\frac{1}{4}\delta_{\omega_1}+\frac{3}{4}\delta_{\omega_2}}{\overset{\text{(prior state)}}{\boxed{\text{Fig. 15.3}}}} \longrightarrow \overset{\text{(a white ball is picked)}}{\boxed{\text{Fig. 15.4}}} \longrightarrow \underset{\frac{2}{5}\delta_{\omega_1}+\frac{3}{5}\delta_{\omega_2}}{\overset{\text{(post state)}}{\boxed{\text{Fig. 15.5}}}} \tag{15.4}$$

Considering the mixed measurement (i.e., the (15.2) in the case that $p = 1/4$):

$$\mathsf{M}_{C(\Omega)}(\mathsf{O} \times \mathsf{O}_U = (\{W,B\} \times \{U_1, U_2\}, 2^{\{W,B\}\times\{U_1,U_2\}}, F \times G_U), S_{[*]}(\rho^{(1/4)}_{\text{prior}})) \tag{15.5}$$

we see that the above (15.4) is the same as the Bayesian result (15.3).

Note that the measurement (15.5) 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

$$[\widehat{F}(\{W\})](\theta_k) = 4/5, \quad [\widehat{F}(\{B\})](\theta_k) = 1/5, \quad (k = 1, 2, ..., 25)$$
$$[\widehat{F}(\{W\})](\theta_k) = 2/5, \quad [\widehat{F}(\{B\})](\theta_k) = 3/5, \quad (k = 26, 27, ..., 100) \tag{15.6}$$
$$[\widehat{G}_U(\{U_1\})](\theta_k) = 1, \quad [\widehat{G}_U(\{U_2\})](\theta_k) = 0, \quad (k = 1, 2, ..., 25)$$
$$[\widehat{G}_U(\{U_1\})](\theta_k) = 0, \quad [\widehat{G}_U(\{U_2\})](\theta_k) = 1, \quad (k = 26, 27, ..., 100) \tag{15.7}$$

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

$$\mathsf{M}_{C(\Theta)}(\widehat{\mathsf{O}} \times \widehat{\mathsf{O}}_U = (\{W,B\} \times \{U_1, U_2\}, 2^{\{W,B\}\times\{U_1,U_2\}}, \widehat{F} \times \widehat{G}_U), S_{[*]}(\nu_0)) \tag{15.8}$$

which is identified with the measurement (15.5) 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,

$$\boxed{\text{(H)=(15.8): the Heisenberg picture}} \underset{\text{identification}}{\overset{\Phi}{\longleftarrow}} \boxed{\text{(15.5): the Schrödinger picture}}$$

Thus, as a particular case of the above arguments, we can answer Problem 15.3 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.

15.2 The principle of equal odds weight

From the above arguments, we see that

Proclaim 15.4. [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

$$\rho_{\text{prior}} = \frac{1}{n}\sum_{k=1}^{n} \delta_{\omega_k}. \tag{15.9}$$

Explanation. In betting, it is certain that everybody wants to choose an unpopular ω_k. Thus, I believe that everybody agrees with Proclaim 15.4. 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 15.1.** 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 8.18

($\sharp_3$) the principle of equal weight in Proclaim 15.4

which are essentially the same.

In order to promote the readers' understanding of the difference between Theorem 8.18 and Proclaim 15.4, we show the following example, which should be compared with Problem 5.14 and Problem 8.17

Problem 15.5. [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 ?*

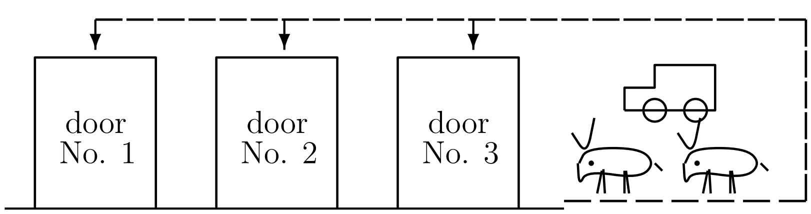
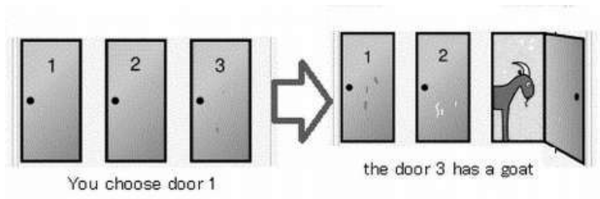


Figure 15.6: Monty Hall problem



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 15.4 and Bayes theorem (B_2). That is, the proof is similar to Problem 8.16 . □

Chapter 16

Postscript: Everyday science

This research report examines [LCI area] (i.e., the area within the green line) in the diagram below.

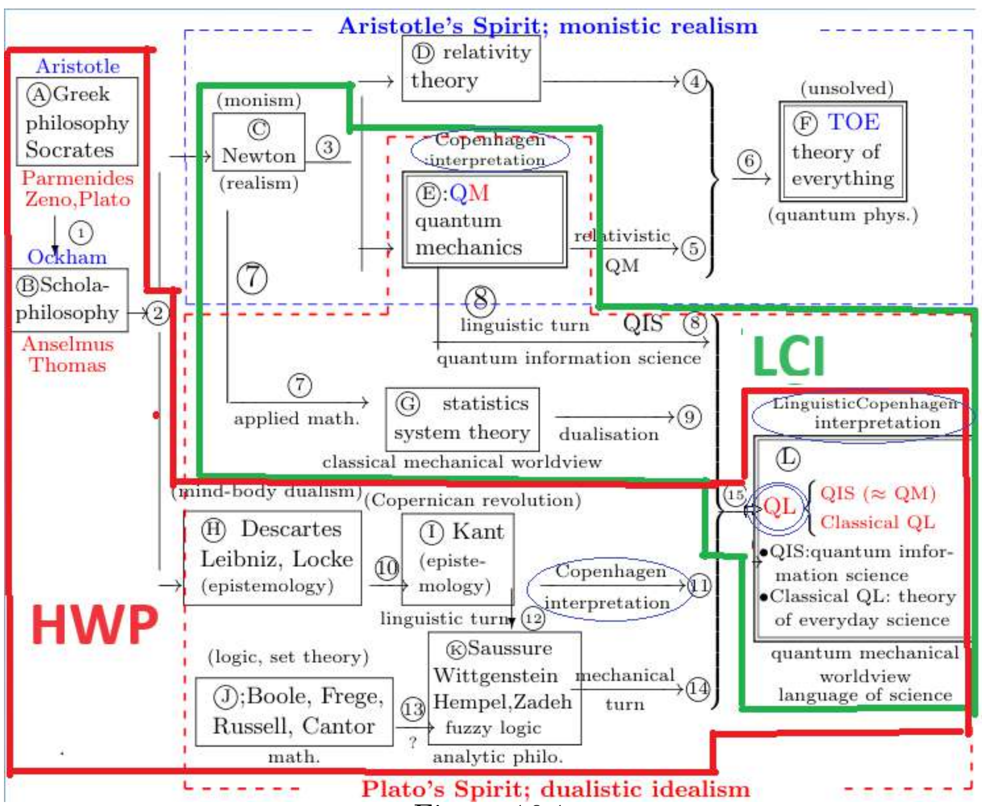


Figure 16.1

Also, recall that [HWP] and [LCI] is respectively discussed in refs. [63, 76] and refs. [62, 65]. Figure 16.1 implies that

(A) [My vision of the future on the interpretation of quantum mechanics]:
The diagram above alludes to the shape of the end of the century-long interpretative problem of quantum mechanics (the Copenhagen Interpretation versus the Other Interpretations (e.g., many world, Bohmian mechanics, etc.)). Interpretations other than the Copenhagen interpretation will develop and dissolve beyond the framework of quantum mechanics as we move in the direction of ⑤ and ⑥. The Copenhagen Interpretation will go in the direction of ⑧. In other words, the Copenhagen Interpretation is not a rule governing the small scientific theory

of quantum mechanics, but a rule governing the vast scientific theory (i.e. quantum language), including statistics. It would be a waste to confine the Copenhagen interpretation only to quantum mechanics Ⓔ. Note "Copenhagen interpretation" in Ⓘ and Ⓛ! The above is "my vision of the future of the problem of interpretation of quantum mechanics", or rather, anyone who knows quantum language would think this way. Thus, the Copenhagen Interpretation is eternal.

16.1 My favorite results (Best 10)

Under QL (= QM(=QIS) + classical QL), I showed a lot of important results in this book and my recent book [76]. There are many different opinions on their importance, in my personal opinion, the following are my top ten best jobs

....

(i) Solving of Zeno's paradox (*cf.* Sec. 9.8, or [76])

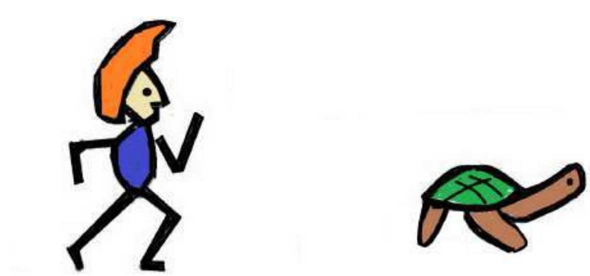
For quite some time now (*cf.* ref. [37]), I have believed that Zeno's paradox is due to the fact that "everyday science" is not established as a scientific theory. The proofs were gradually simplified (*cf.* refs. [62, 63]) and the current recommended proof is given in Sec. 9.8 of this book. That is, I believe the following equivalence:

(a) Solving Zeno's paradox

(b) Completing the philosophy of science

(c) the discovery of "the theory of everyday science"

Zeno's paradox (a) may be a symbol of philosophical puzzle, and thus it a 'problem that should not be solved'. Also, there may be various opinions about "what is the philosophy of science?". Therefore some philosophers may disagree with (a) and (b), but if they do, that's fine by me. That is because my true assertion is (c). Also, see Note 9.9.



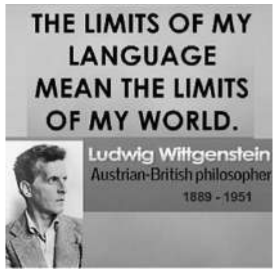
....

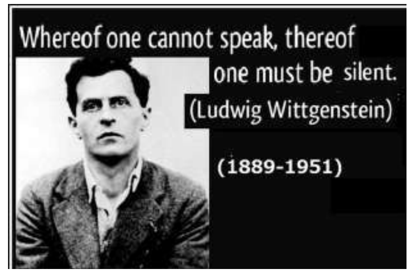
(ii) Quantum linguistic understanding of analytic philosophy (i.e., Wittgenstein's TLP (Tractatus Logico-philosophicus, *cf.* ref. [113, 76])).

The starting point of analytic philosophy is Wittgenstein's TLP: in TLP, Wittgenstein asked "Why does logic work in our world?" and examined the question:

(♯) Why does logic work in our world?

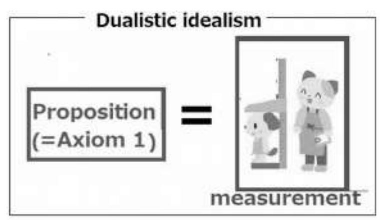
As mentioned in (E_7) of Sec. 3.1, he wrote down in TLP gems expressing the basic spirit of the linguistic Copenhagen interpretation.



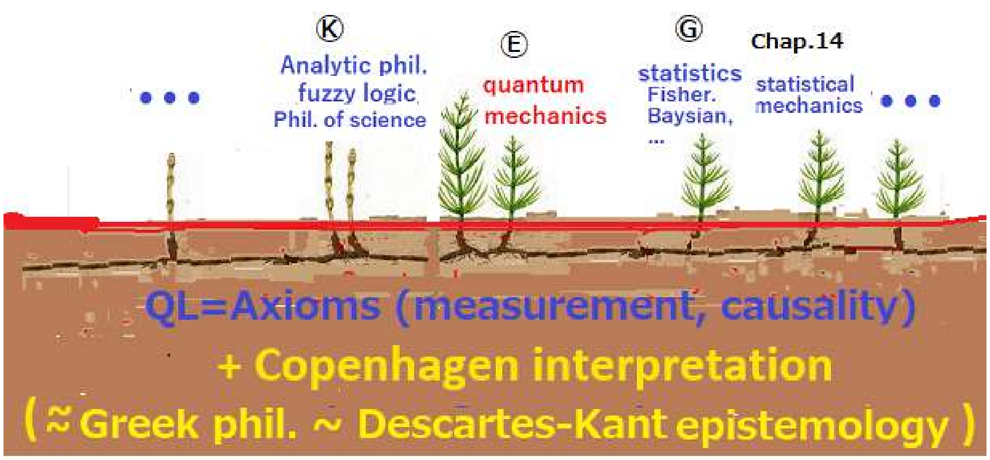


However, his TLP is a kind of poetry collection, and theoretically no one could read it. I assert that this question (♯) can be anssered in QL (*cf.* Chap. 12 ref. [76]). And thus, I believe that Wittgenstein's dream has come true in QL (*cf.* Chap. 12 ref. [76]) as 'fuzzy logic'. Further, fuzzy logic can solve analytic-synthetic distinction problem (Carnap-Quine controversy). In fuzzy logic, this controversy was won by Quine. Since fuzzy logic asserts that

$$\text{fuzzy proposition} = \text{measurement} \quad (= \text{Axiom 1})$$



Look below.



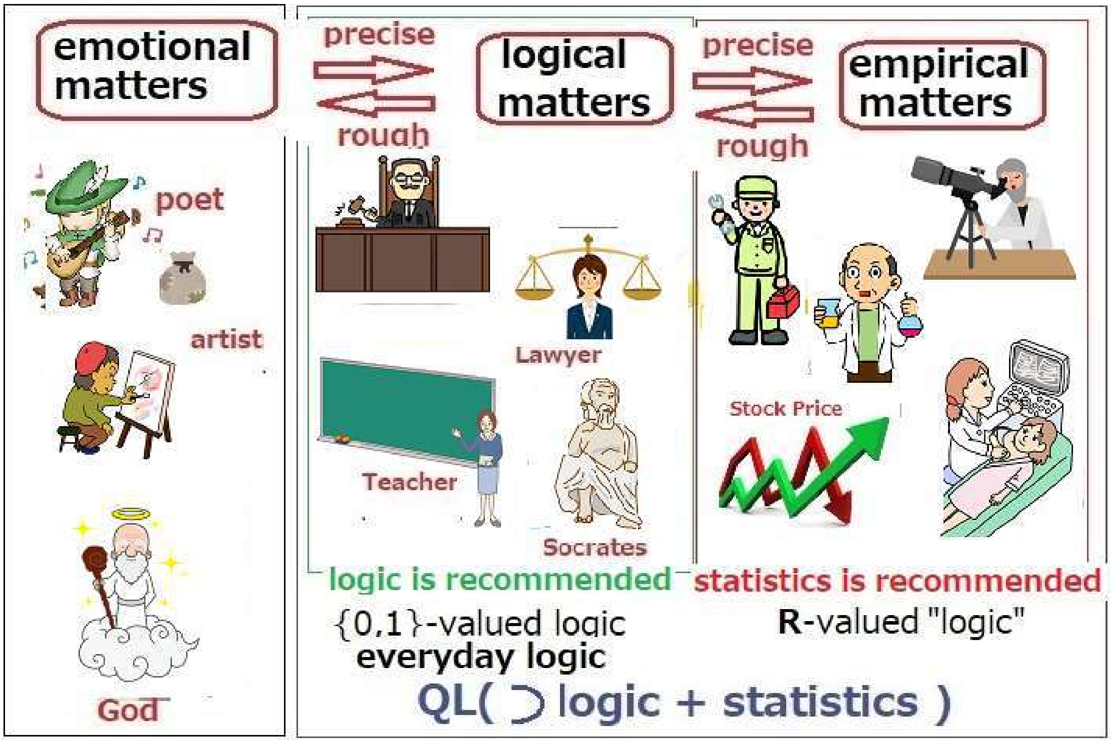
My opinion on analytic philosophy (Wittgenstein) is written below.

(a) It was a mistake to dismiss Descartes-Kant's epistemology as metaphysics.

(b) It is true that analytical philosophy was born out of the study of mathematical logic. However, there is no direct relationship between mathematical logic and analytical philosophy. I think it is unfortunate that 100 years have passed without us realizing this.

(c) The argument of analytic philosophy is, "Be logical!". But the argument of quantum language is, "Be scientific!" (="Speak quantum language!"). Analytic philosophy could not find the similarity between logic and statistics.

For further information see my homepage



(d) Wittgenstein was not wrong when he chose "logic" as the theme of his philosophy. About 20 years ago, I believed that logic had nothing to do with philosophy. Wittgenstein sensibly knew the difference between "logic" and "mathematical logic," and I am sure that he was one of the greatest philosophers.

. . . .

(iii) The clearance of 'Hempel's ravens paradox' in philosophy of science. (See ⑭ above, and [76]).

Hempel's raven paradox is a central problem in the philosophy of science, and its resolution means the completion of the scientific part of the philosophy of science. As mentioned in this book, we consider QL to be a foundational theory of everyday science. And it should be noted that this paradox arises from the consideration of set theory as the fundamental theory of the philosophy of science. That is, after learning about Hempel's ravens paradox, I became convinced that

- the purpose of philosophy of science
  = to discover the language of 'everyday science' (=classical QL).

I have to admire Hempel for approaching the heart of the philosophy of science without knowledge of quantum language

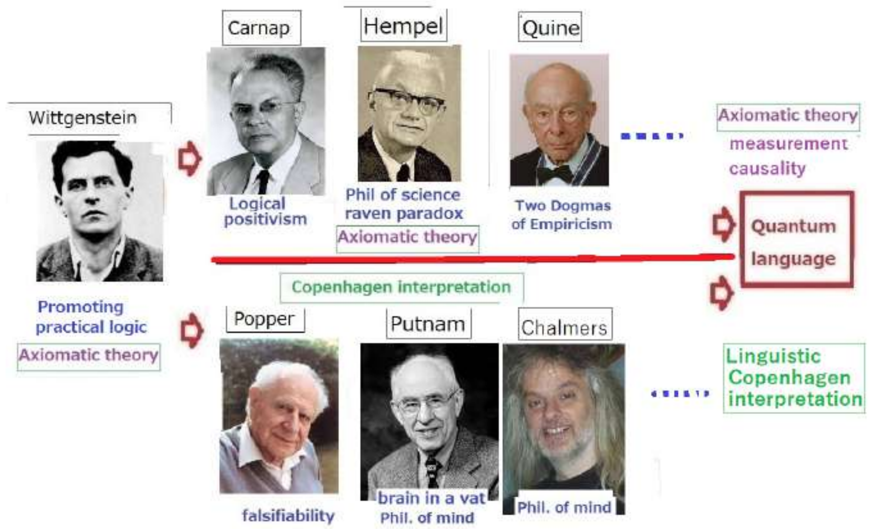
For further information see my homepage

The next question is the frequently raised question of analytic philosophy.

- Has there ever been even a single research result that could be called ‘analytic philosophy’, which is distinct from philosophy?

I can say that, if the perfection of analytic philosophy is QL, we can assert that almost all (dualistic idealist) philosophical problems can be solved within analytic philosophy.

With regard to post-Wittgenstein analytic philosophy, I consider the following (*cf.* ref. [76]).



Just to be clear, I say it again:

- **“Philosophy of Science = QL”**

. . . .

(iv) Solving the problem of universals (*cf.* ref. [76])

The problem of universals is the greatest controversy (Aristotle’s monism vs. Plato’s dualism) in Scholastic philosophy. The controversy was muddled because Plato’s dualism was not clearly understood; if Plato’s dualism is replaced by The dualism of QL, the meaning of the controversy becomes clearer. However, Anselmus’ and Thomas Aquinas’ understanding of dualism was groundbreaking at the time and laid the foundation for Descartes’ discovery of mind-body dualism.

Their (Anselmus’ Thomas Aquinas’ and Descartes’) arguments are so outstanding that it is hard to believe that they are considerations made without knowledge of quantum language (the scientific end point of dualistic idealism).

As mentioned in [76], I consider that what Descartes did was rewrite Thomas's philosophy for the general public (using the magic phrase [I think therefore I am]).

. . . .

(v) Solving of Hume's problem of induction and the grue paradox (*cf.* ref. [76]).

This paradox is due to the fact that 'everyday science' is not established as a scientific theory. Therefore, it is automatically solved if classical QL is accepted as the theory of everyday science. That is, the law of large numbers is one of the most important theorems in the theory of everyday science. Also, I think that the concept of 'parallel time' is needed in classical QL (*cf.* sections 9.7 and 9.8).[1]

Glue's emerald paradox is well-known but not written in an understandable way. However, Glue's emerald paradox can be solved immediately if Hume's problem of induction is seen as a law of large numbers within a quantum language.

. . . .

[1] As mentioned in Sec. 9.7 (Leibniz-Clarke correspondence), we think that

we, like God, can create 'space' and 'time' at our convenience,

since "space" and "time" do not exist in Axioms 1 and 2 (in QL).

(vi) The proposal of the linguistic understanding of von Neumann-Lüders projection postulate (i.e., Postulate 10.7 in Sec. 10.2).

And we use this postulate to clarify the paradox of Schrödinger cat, though what we have done is 'clarification', not 'resolution'.

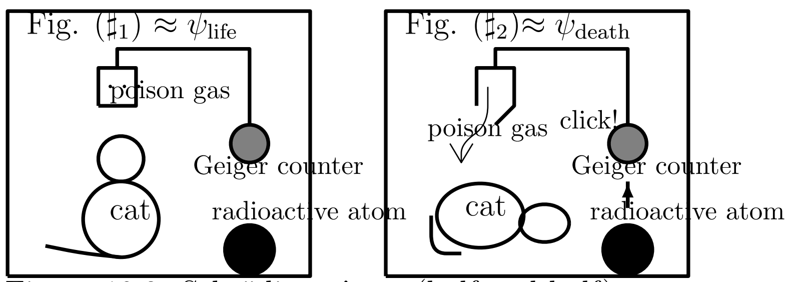


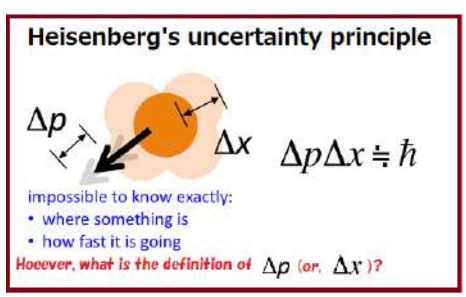
Figure 10.3: Schrödinger's cat(half and half)

As mentioned in Chap. 10, note that, without Postulate 10.7, we can not mention several famous quantum paradoxes (such as the paradox of Schrödinger cat).

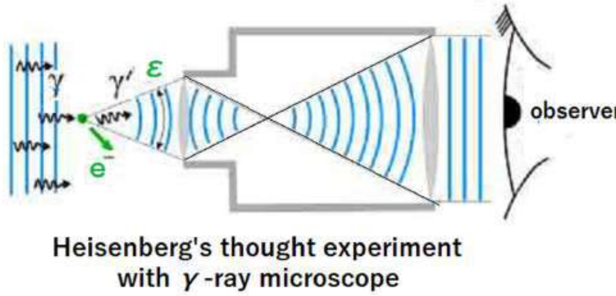
. . . .

(vii) The discovery of Theorem 4.16 (= the true Heisenberg's uncertainty principle $\Delta_p \cdot \Delta_x \geq \hbar/2$)

That is, the definitions of Δ_p and Δ_x are proposed

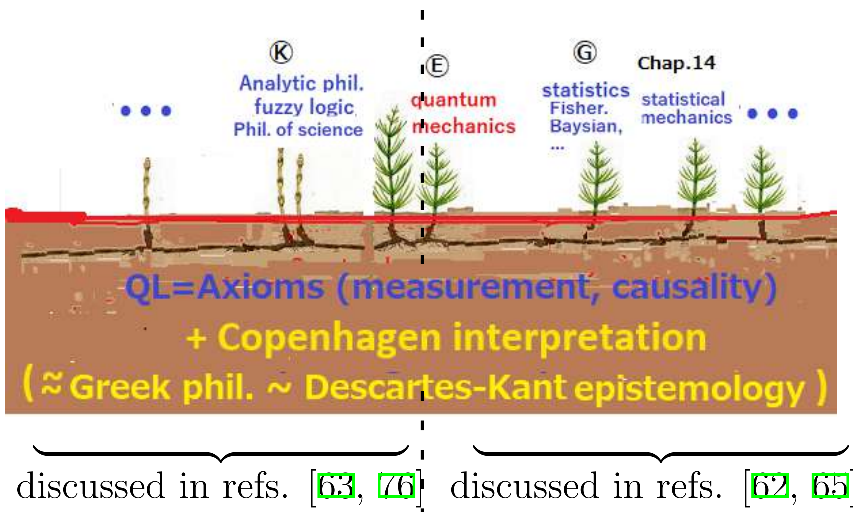


I am of the opinion that just as Descartes' cogito proposition was used as a catchphrase for Descartes-Kant's epistemology, Heisenberg's uncertainty principle (Proposition 4.11) from the γ-ray microscope thought experiment was used as propaganda for quantum mechanics.



. . . .

(viii) I assert that 'theoretical statistics' should be constructed in the frame of QL (i.e., [QL is the language of (everyday) science], or [science is 'speaking in QL'])[2]. I believe that this assertion is the biggest in science. This is the main theme in this book. See Ⓖ below.



(See Chap. 5 Statistics (I), Chap. 8 Bayes statistics and Chap. 12 Statistics (II))

I am aware that there are negative opinions about the scientific part of the history of Western philosophy (ancient Greek philosophy, Scholastic philosophy, Descartes-Kant's epistemology, analytic philosophy), but I am rather positive about dualistic idealism, since these lead to QL as we saw in Figure 0.1 in Preface (or, ref. [76]).

. . . .

(ix) Solving the Monty Hall Problem (and the proof of the principle of equal weight)[3]

($\sharp_1$) Monty Hall Problem in Fisher statistics · · · Problem 5.14

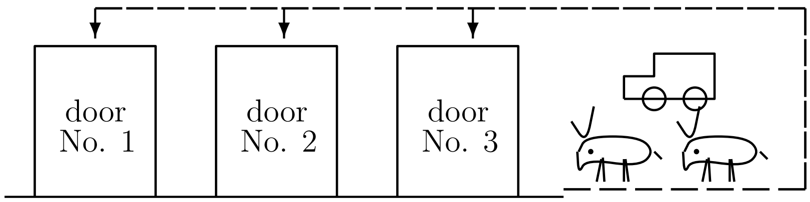
($\sharp_2$) Monty Hall Problem by the moment method · · · Remark 5.15

($\sharp_3$) Monty Hall Problem by Bayes' method · · · Problem 8.16

($\sharp_4$) Monty Hall Problem by the principle of equal weight · · · Problem 15.5

[2] I do not agree with the claims of analytic philosophy (i.e., mathematics is the language of science).

[3] As far as I checked (*e.g.*, ref. [103]), the solution ($\sharp_1$) is new.

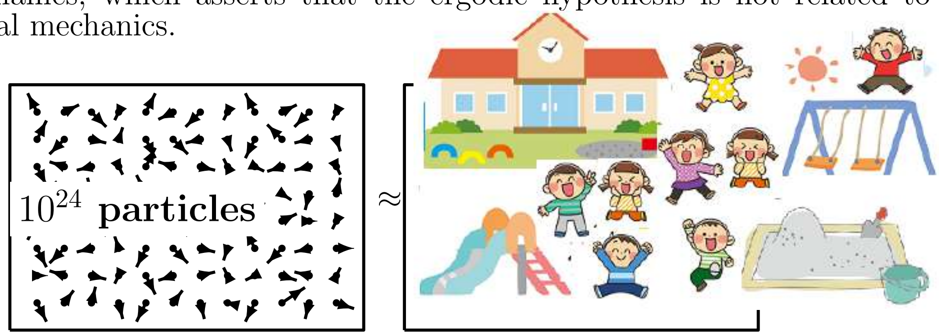


The solution ($\sharp_3$) in Bayesian statistics is well known. However, I think that my solution ($\sharp_1$) (*cf.* [44] (using Fisher's maximum likelihood method)) should be standard, since Fisher statistics is more fundamental than Bayesian statistics (*cf.* ($\sharp$) in Sec. 8.4).

I like these puzzles such as 'two envelopes problem', 'three prisoners problem', 'Bertrand's paradox', etc.

. . . .

(x) In Chap. 14, I propose the quantum linguistic characterization of equilibrium statistical mechanics, which asserts that the ergodic hypothesis is not related to equilibrium statistical mechanics.



Clarifying the relationship between statistics and statistical mechanics is an open question, which could be fully answered. That is, I think that

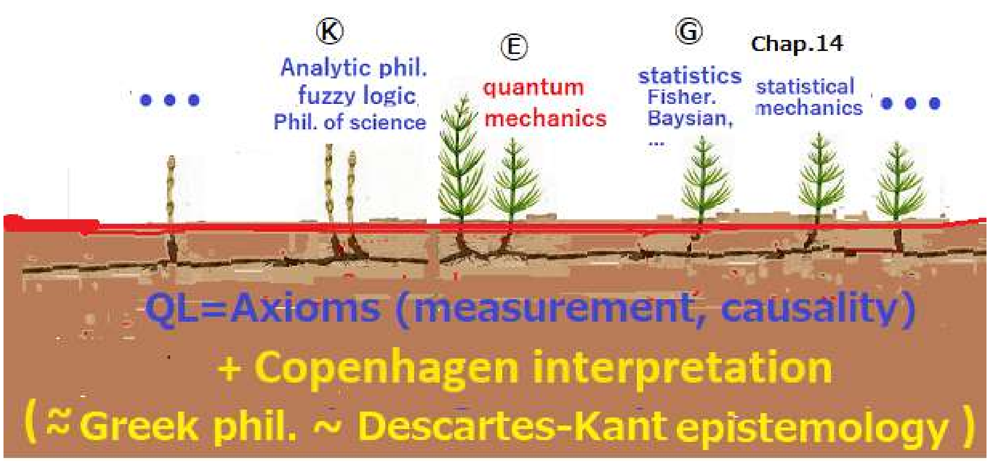


Figure 16.1

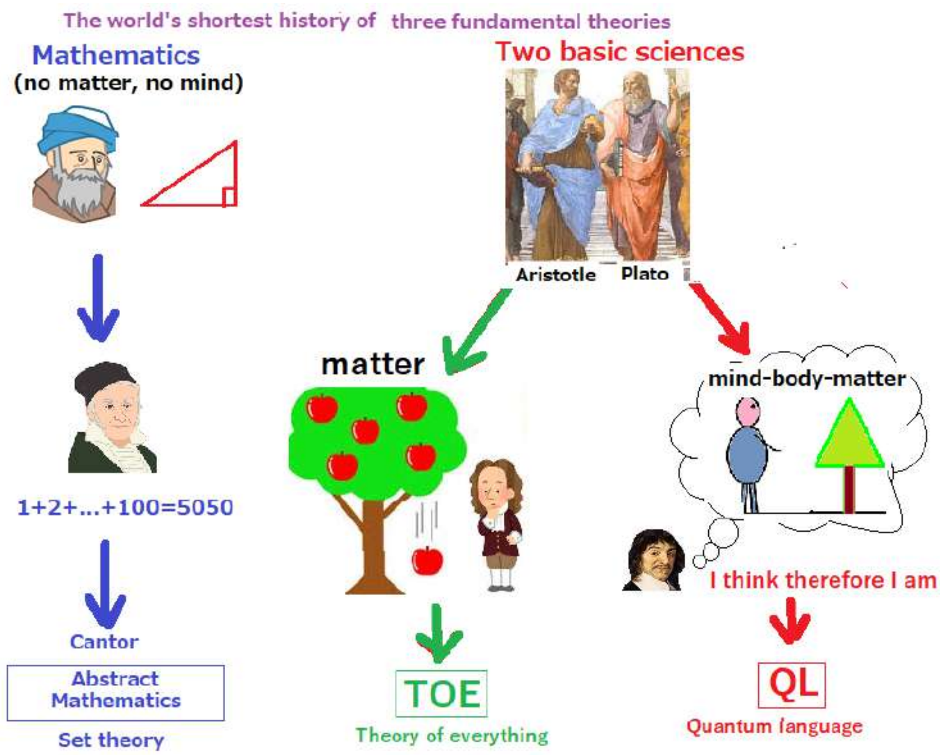
This is a matter of course, since we assert that classical QL is proposed as a theory of everyday science.

. . . .

For further information see my homepage

16.2 At the end.

(i):In the first year of university, regardless of whether you are a liberal arts or science student, there are lectures on mathematics, physics, and statistics. This means that these three are the "most important basic theories" and correspond to the following three.



Mathematics and QL have reached their destination, and now we are in the details. However, the TOE of physics is still unknown. I think that human wisdom must solve these three problems. Physicists need to hurry, otherwise AI will do it first!

(ii): Another question I've been wondering about is, "Is there a fourth fundamental theory?" I believe there isn't, but I'm not confident.

(iii): I add Chap. 17 [Appendix: Socrates' absolutism was perfected by QL(See. Sec.17.1)]. I highly recommend you read it.

Shiro Ishikawa[4]
20 July in 2024

[4]For the further information concerning quantum language (notices on improvements to the results of this publication, etc.), see home page: **https://ishikawa.math.keio.ac.jp/indexe.html**

References

[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-Podolsky-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] Clauser,J.F., Horne M.A., Shimony,A, Holt,R.A., Proposed experiment to test local hidden variable theories, *Phys,Rev,Lett,* 23(15), 880-884 (1969)

[11] F. Click, *The Astonishing Hypothesis: The Scientific Search For The Soul,* New York: Charles Scribner's Sons., 1994.

[12] Davies, E.B. *Quantum theory of open systems,* Academic Press 1976

[13] de Broglie, L. *L'interpretation de la mecanique ondulatoire,* Journ. Phys. Rad. 20, 963 (1959)

[14] Einstein, A., Podolsky, B. and Rosen, N. *Can quantum-mechanical description of reality be considered completely?* Physical Review Ser 2(47) 777–780 (1935)

[15] R. P. Feynman *The Feynman lectures on Physics; Quantum mechanics* Addison-Wesley Publishing Company, 1965

[16] G.A. Ferguson, Y. Takane, *Statistical analysis in psychology and education* (Sixth edition). New York: McGraw-Hill. (1989)

[17] Goodman, Nelson (1983) Fact, fiction, and forecast, Harvard University Press

[18] L. Hardy, *Quantum mechanics, local realistic theories, and Lorentz-invariant realistic theories,* Physical Review Letters 68 (20): 2981-2984 1992

[19] Hawking, Stephen *A brief History of Time,* Bantam Dell Publishing Group 1988

[20] Heisenberg, W. *Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik,* Z. Phys. 43, 172–198 (1927)

[21] Hempel, C.G. (1945) Studies in the Logic of Confirmation (I), Mind, 54 1-26

[22] Hempel, C.G. (1965) Aspects of Scientific Explanations and other Essays in the Philosophy of Science, Free Press

[23] Holevo, A.S. *Probabilistic and statistical aspects of quantum theory,* North-Holland publishing company (1982)

[24] D.Howard, Who invented the "Copenhagen Interpretation"? A study in mythology, Philosophy of Science, 71 2004, 669-682

[25] Isaac, R. *The pleasures of probability,* Springer-Verlag (Undergraduate texts in mathematics) 1995

[26] Ishikawa,S. *Uncertainty relation in simultaneous measurements for arbitrary observables,* Rep. Math. Phys., **9**, 257-273, 1991
doi: 10.1016/0034-4877(91)90046-P

[27] Ishikawa, S. *Uncertainties and an interpretation of nonrelativistic quantum theory,* International Journal of Theoretical Physics 30 401–417 (1991) doi: 10.1007/BF00670793

[28] 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

[29] Ishikawa,S. *Fuzzy inferences by algebraic method,* Fuzzy Sets and Systems 87, 181–200 (1997) doi:10.1016/S0165-0114(96)00035-8

[30] Ishikawa,S. *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

[31] Ishikawa,S. *Fuzzy Logic in Measurements,* Fuzzy Sets and Systems 100, 291-300 (1998) (https://doi.org/10.1016/S0165-0114(97)00154-1)

[32] Ishikawa,S., et al. *A dynamical system theoretical approach to Newtonian mechanics*, Far east journal of dynamical systems Vol.1, Issue 1, 1-34 (1999)
(http://www.pphmj.com/abstract/191.htm)

[33] Ishikawa,S. *Statistics in measurements*, Fuzzy sets and systems, Vol. 116, No. 2, 141-154, 2000
doi:10.1016/S0165-0114(98)00280-2

[34] Ishikawa,S., *A system theoretical characterization of the relation between Fisher's and Bayes' methods* Far East Journal of Dynamical Systems, Volume 7, Issue 1, Pages 19 - 33 (May 2002) (http://www.pphmj.com/abstract/1914.htm)

[35] Ishikawa,S. *Mathematical Foundations of Measurement Theory,* Keio University Press Inc. 335pages, 2006, (http://www.keio-up.co.jp/kup/mfomt/)

[36] Ishikawa,S., *A new formulation of measurement theory,* Far East Journal of Dynamical Systems, Volume 10, Issue 1, Pages 107 - 117 (February 2008) (http://www.pphmj.com/abstract/3096.htm)

[37] Ishikawa,S., *Dynamical systems, measurements, quantitative languages and Zeno's paradox,* Far East Journal of Dynamical Systems, Volume 10, Issue 3, Pages 277 - 292 (October 2008) (http://www.pphmj.com/abstract/3595.htm)

[38] Ishikawa,S., *Ergodic problem in quantitative language,* Far East Journal of Dynamical Systems, Volume 11, Issue 1, Pages 33 - 48 (February 2009) (http://www.pphmj.com/abstract/3893.htm)

[39] Ishikawa,S., *A measurement theoretical characterization of dynamical system theory and statistics,* Far East Journal of Dynamical Systems, Volume 11, Issue 2, Pages 161 - 178 (June 2009) (http://www.pphmj.com/abstract/3781.htm)

[40] Ishikawa,S., *Fisher's method, Bayes' method, and Kalman filter in measurement theory* Far East Journal of Dynamical Systems, Volume 29, Issue 1, Pages 9 - 23 (September 2009) (http://www.pphmj.com/abstract/4316.htm)

[41] Ishikawa,S. *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)

[42] Ishikawa,S. *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)

[43] Ishikawa,S. *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&)

[44] Ishikawa,S. *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)

[45] Ishikawa,S. *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
(https://www.scirp.org/journal/PaperInformation.aspx?PaperID=18861)

[46] Ishikawa,S. *The linguistic Copenhagen interpretation of quantum mechanics,*
arXiv:1204.3892v1[physics.hist-ph],(2012) (http://arxiv.org/abs/1204.3892)

[47] Ishikawa,S. *Zeno's paradoxes in the Mechanical World View,* arXiv:1205.1290v1 [physics.hist-ph], (2012)

[48] Ishikawa,S. *What is Statistics?; The Answer by Quantum Language,* arXiv:1207.0407 [physics.data-an] 2012. (http://arxiv.org/abs/1207.0407)

[49] Ishikawa,S. *Measurement Theory in the Philosophy of Science,* arXiv:1209.3483 [physics.hist-ph] 2012. (http://arxiv.org/abs/1209.3483)

[50] Ishikawa,S. *Heisenberg uncertainty principle and quantum Zeno effects in the linguistic Copenhagen interpretation of quantum mechanics,* arxiv:1308.5469 [quant-ph],(2013)

[51] Ishikawa,S. K. Kikuchi *The Confidence Interval Methods in Quantum Language,* https://arxiv.org/abs/1312.6757 [math.ST] 2013

[52] Ishikawa,S. *A quantum linguistic characterization of the reverse relation between confidence interval and hypothesis testing,* arxiv:1401.2709[math.ST],(2014)

[53] Ishikawa,S. *ANOVA (analysis of variance) in the quantum linguistic formulation of statistics,* arxiv:1402.0606[math.ST],(2014)

[54] Ishikawa,S. *Regression analysis in quantum language,* arxiv:1403.0060[math.ST],(2014)

[55] Ishikawa,S. K. Kikuchi: *Kalman filter in quantum language,* arXiv:1404.2664 [math.ST] 2014. (http://arxiv.org/abs/1404.2664)

[56] Ishikawa,S. *The double-slit quantum eraser experiments and Hardy's paradox in the quantum linguistic Copenhagen interpretation,* arxiv:1407.5143[quantum-ph],(2014)

[57] Ishikawa,S. *The Final Solutions of Monty Hall Problem and Three Prisoners Problem,* arXiv:1408.0963 [stat.OT] 2014. (http://arxiv.org/abs/1408.0963)

[58] Ishikawa,S. *Two envelopes paradox in Bayesian and non-Bayesian statistics* arXiv:1408.4916v4 [stat.OT] 2014. (http://arxiv.org/abs/1408.4916)

[59] Ishikawa,S. *Linguistic Copenhagen 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)
Or see the following preprint;
(http://www.math.keio.ac.jp/academic/research_pdf/report/2015/15009.pdf)
(https://doi.org/10.48550/arXiv.1511.07777)

[60] Ishikawa,S., A Final solution to mind-body problem by quantum language, *Journal of quantum information science*, Vol. 7, No.2 , 48-56, 2017, DOI: 10.4236/jqis.2017.72005 (http://www.scirp.org/Journal/PaperInformation.aspx?PaperID=76391)

[61] Ishikawa,S., Bell's inequality should be reconsidered in quantum language , Journal of quantum information science, Vol. 7, No.4 , 140-154, 2017, DOI: 10.4236/jqis.2017.74011
(http://www.scirp.org/Journal/PaperInformation.aspx?PaperID=80813)

[62] Ishikawa,S., Linguistic Copenhagen interpretation of quantum mechanics; Quantum Language, Research Report, Dept. Math. Keio University, (http://www.math.keio.ac.jp/en/academic/research.html) Version 1 KSTS/RR-15/001 (2015); 416 pages, Version 2 KSTS/RR-16/001 (2016); 426 pages, Version 3 KSTS/RR-17/007 (2017); 434 pages, Version 4 KSTS/RR-18/002 (2018); 449 pages, Version 5 KSTS/RR-19/003 (2019); 473 pages, Version 6 https://philarchive.org/rec/ISHLCI-3; 352 pages
Ver. 1:(http://www.math.keio.ac.jp/academic/research_pdf/report/2015/15001.pdf)
Ver. 2:(http://www.math.keio.ac.jp/academic/research_pdf/report/2016/16001.pdf)
Ver. 3:(http://www.math.keio.ac.jp/academic/research_pdf/report/2017/17007.pdf)
Ver. 4:(http://www.math.keio.ac.jp/academic/research_pdf/report/2018/18002.pdf)
Ver. 5:(http://www.math.keio.ac.jp/academic/research_pdf/report/2019/19003.pdf)
Ver. 6:(https://philarchive.org/rec/ISHLCI-3) (2023)

[63] Ishikawa,S., *History of Western Philosophy from the quantum theoretical point of view*, Research Report (Department of mathematics, Keio university, Yokohama), Version 1 (KSTS-RR-16/005, 2016, 142 pages), Version 2 (KSTS-RR-17/004, 2017, 132 pages), Version 3 (KSTS-RR-20/001, 2020, 296 pages), Version 4 (KSTS-RR-21/001, 2021, 306 pages), Version 5 (KSTS-RR-22/001, 2022, 443 pages),
Ver. 1:(http://www.math.keio.ac.jp/academic/research_pdf/report/2016/16005.pdf)
Ver. 2:(http://www.math.keio.ac.jp/academic/research_pdf/report/2017/17004.pdf)
Ver. 3:(http://www.math.keio.ac.jp/academic/research_pdf/report/2020/20001.pdf)
Ver. 4:(http://www.math.keio.ac.jp/academic/research_pdf/report/2021/21001.pdf)
Ver. 5:(https://www.math.keio.ac.jp/wp-content/uploads/2022/12/22001.pdf)

[64] Ishikawa,S. Linguistic Interpretation of Quantum Mechanics –Towards World-Description in Quantum Language– Shiho-Shuppan Publisher, 1st edition(409 p. (2015)), 2nd edition (411 p. (2017)) (https://www.shiho-shuppan.com/index.php?Linguistic%20Interpretation%20of%20Quantum%20Mechanics%20S.ISHIKAWA) (Out of Print) For the revised edition, see [65].

[65] Ishikawa,S. Linguistic Copenhagen Interpretation of Quantum Theory –Why does statistics work in our world?– (2023)) Shiho-Shuppan press ISBN 978-4-907625-67-2
(https://www.shiho-shuppan.com/index.php?LCIQT)

[66] Ishikawa,S., Leibniz-Clarke correspondence, Brain in a vat, Five-minute hypothesis, McTaggart's paradox, etc. are clarified in quantum language , Open Journal of philosophy, Vol. 8, No.5 , 466-480, 2018, DOI: 10.4236/ojpp.2018.85032
(https://www.scirp.org/Journal/PaperInformation.aspx?PaperID=87862)

[67] Ishikawa,S., Leibniz-Clarke correspondence, Brain in a vat, Five-minute hypothesis, McTaggart's paradox, etc. are clarified in quantum language [Revised version] 2018, (https://philpapers.org/rec/ISHLCB)
(http://www.math.keio.ac.jp/academic/research_pdf/report/2018/18001.pdf)

[68] Ishikawa,S., Leibniz-Clarke correspondence, Brain in a vat, Five-minute hypothesis, McTaggart's paradox, etc. are clarified in quantum language , Open Journal of philosophy, Vol. 8, No.5 , 466-480, 2018, DOI: 10.4236/ojpp.2018.85032
(https://www.scirp.org/Journal/PaperInformation.aspx?PaperID=87862)

[69] Ishikawa,S., Leibniz-Clarke correspondence, Brain in a vat, Five-minute hypothesis, McTaggart's paradox, etc. are clarified in quantum language [Revised version] (2018), (https://philpapers.org/rec/ISHLCB)
(http://www.math.keio.ac.jp/academic/research_pdf/report/2018/18001.pdf)

[70] Ishikawa,S., Leibniz-Clarke correspondence, Brain in a vat, Five-minute hypothesis, McTaggart's paradox, etc. are clarified in quantum language , Open Journal of philosophy, Vol. 8, No.5 , 466-480, 2018, DOI: 10.4236/ojpp.2018.85032
(https://www.scirp.org/Journal/PaperInformation.aspx?PaperID=87862)

[71] Ishikawa,S., Leibniz-Clarke correspondence, Brain in a vat, Five-minute hypothesis, McTaggart's paradox, etc. are clarified in quantum language [Revised version] (2018), (https://philpapers.org/rec/ISHLCB)
(http://www.math.keio.ac.jp/academic/research_pdf/report/2018/18001.pdf)

[72] Ishikawa, S., (2019) Philosophy of science for scientists; The probabilistic interpretation of science, Journal of quantum information science, Vol. 9, No.3 , 140-154, DOI: 10.4236/jqis.2019.93007
(https://www.scirp.org/Journal/paperinformation.aspx?paperid=95447)

[73] Ishikawa, S., (2020) Wittgenstein's picture theory in the quantum mechanical worldview, Journal of quantum information science, Vol. 10, No.4, 104-125, DOI:10.4236/jqis.2020.104007 (https://www.scirp.org/journal/paperabs.aspx?paperid=106233)

[74] Ishikawa, S., (2021) Wittgenstein's picture theory in the linguistic Copenhagen interpretation of dualistism, PhilPaper, (https://philpapers.org/rec/SHIWPT)

[75] Ishikawa, S., (2021) Fuzzy Logic in the Quantum Mechanical Worldview ; Related to Zadeh, Wittgenstein, Moore, Saussure, Quine, Lewis Carroll, etc. Journal of applied mathematics and physics, Vol. 9, No.7, 1583-1610, DOI:10.4236/jamp.2021.97108 (https://www.scirp.org/journal/paperinformation.aspx?paperid=110830)

[76] Ishikawa,S., *History of Western Philosophy from a perspective of quantum theory— Introduction to theory of everyday science–* Shiho-Shuppan Publisher, 425 p. (2023) (https://www.shiho-shuppan.com/index.php?HWPQT)

[77] Ishikawa,S., K. Kikuchi, (2021) *Quantum Fuzzy Logic and Time, Journal of Applied mathematics and physics,* Vo.9 No.11 2021, 2609-2622 (https://www.scirp.org/journal/paperinformation.aspx?paperid=112972)

[78] Ishikawa,S., (2022) *The problem of universals from the scientific point of view: Thomas Aquinas should be more appreciated* Open Journal of Philosophy, Vol. 12, No. 1, 86-104, (https://www.scirp.org/journal/paperinformation.aspx?paperid=115252)

[79] Ishikawa,S., (2022) *Empiricism, Rationalism, and Kantian synthesis in quantum linguistic point of view* Open Journal of Philosophy, Vol. 12, No. 2, 182-198, (https://www.scirp.org/journal/paperinformation.aspx?paperid=117114)

[80] Ishikawa,S., (2024) *History of Western philosophy and quantum language (including quantum mechanics, statistics, fuzzy logic, etc.)* Journal of Applied mathematics and physics, Vo.12 No.5 2024, 1769-1795 DOI: 10.4236/jamp.2024.125110 (https://www.scirp.org/journal/paperinformation?paperid=133335)

[81] Kikuchi, K. Ishikawa,S. *Psychological tests in Measurement Theory, Far east journal of theoretical statistics,* 32(1) 81-99, (2010) ISSN: 0972-0863

[82] Kikuchi,K. *Axiomatic approach to Fisher's maximum likelihood method,* Non-linear studies, 18(2) 255-262, (2011)

[83] Kalman, R. E. *A new approach to linear filtering and prediction problems,* Trans. ASME, J. Basic Eng. 82, 35 (1960)

[84] I. Kant, *Critique of Pure Reason (Edited by P. Guyer, A. W. Wood),* Cambridge University Press, 1999

[85] A. Kolmogorov, *Foundations of the Theory of Probability (Translation),* Chelsea Pub Co. Second Edition, New York, 1960,

[86] U. Krengel, "Ergodic Theorems," Walter de Gruyter. Berlin, New York, 1985.

[87] Lee, R. C. K. *Optimal Estimation, Identification, and Control,* M.I.T. Press 1964

[88] G. Lüders, Über die Zustandsänderung durch den Messprozess, Ann. Phys. (Leipzig) (6)8,322-328, 1951

[89] J. M. E. McTaggart, *The Unreality of Time,* Mind (A Quarterly Review of Psychology and Philosophy), Vol. 17, 457-474, 1908

[90] G. Martin, *Aha! Gotcha: Paradoxes to Puzzle and Delight* Freeman and Company, 1982

[91] Peirce. C.S. (1958) Collected Papers of Charles Sanders Pierce, Harvard Univ. Press.

[92] Popper, K.R. (1959) The logic of Scientific Discover, London Hutchinson

[93] B. Misra and E. C. G. Sudarshan, *The Zeno's paradox in quantum theory,* Journal of Mathematical Physics 18 (4): 756-763 (1977)

[94] N.D. Mermin, *Boojums all the way through, Communicating Science in a Prosaic Age,* Cambridge university press, 1994.

[95] 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)

[96] 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,

[97] Popper, Karl *The Logic of Scientific Discovery,* Routledge. ISBN 978-0-415-27844-7. (1959).

[98] Prugovečki, E. *Quantum mechanics in Hilbert space,* Academic Press, New York. (1981).

[99] Putnam, H. *Reason, Truth, and History,* Cambridge University Press. (1982).

[100] Quine, W.V. (1951) *Two Dogmas of Empiricism (in "From a Logical Point of View")* Harvard U. P., 1951

[101] Redhead, M. Incompleteness, nonlocality, and realism, Oxford University Press, Oxford (1987)

[102] Robertson, H.P. *The uncertainty principle,* Phys. Rev. 34, 163 (1929)

[103] Rosenhause, Jason *The Monty Hall problem: The remarkable story of Math's most contentious Brain Teaser,* Oxford university press. (2009)

[104] D. Ruelle, "Statistical Mechanics, Rigorous Results," World Scientific, Singapore, 1969.

[105] Russel. B. *The Analysis of Mind,* London: George Allen and Unwin. (1921).

[106] Sakai, S. *C^*-algebras and W^*-algebras,* Ergebnisse der Mathematik und ihrer Grenzgebiete (Band 60), Springer-Verlag, Berlin, Heidelberg, New York 1971

[107] Selleri, F. *Die Debatte um die Quantentheorie,* Friedr. Vieweg&Sohn Verlagsgesellscvhaft MBH, Braunschweig (1983)

[108] Shannon, C.E., Weaver. W *A mathematical theory of communication,* Bell Syst. Tech.J. 27 379–423, 623–656, (1948)

[109] Shao, Jun *Mathematical Statistics* (Springer Texts in Statistics) Springer; 2nd ed. (2003)

[110] von Neumann, J. *Mathematical foundations of quantum mechanics* Springer Verlag, Berlin (1932)

[111] S. P.Walborn, et al. *"Double-Slit Quantum Eraser,"* Phys.Rev.A 65, (3), 2002

[112] 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)

[113] Wittgenstein, L *Tractatus Logico-philosophicus,* Oxford: Routledge and Kegan Paul, 1921

[114] Yosida, K. *Functional analysis,* Springer-Verlag (Sixth Edition) 1980

[115] Zadeh, L.A. *Fuzzy sets, Information and Control,* Vol. 8, 338-353 1965

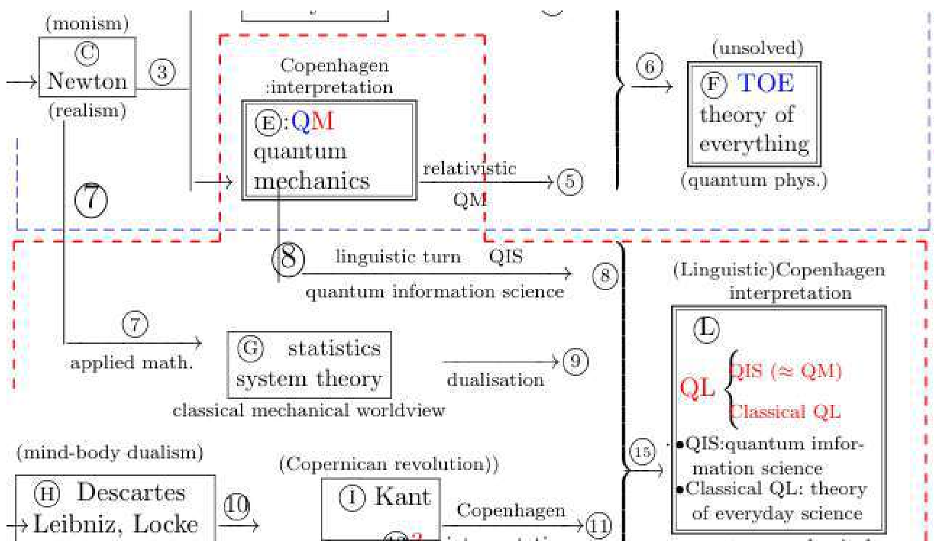
[116] Zadeh, L.A. (2004) *Fuzzy Logic Systems, origin, concepts and trends,*

Index

Notation

Chapter 17

Appendix: Socrates' absolutism was perfected by QL



The diagram above alludes to the shape of the end of the century-long interpretative problem of quantum mechanics (the Copenhagen Interpretation versus the Other Interpretations). Interpretations other than the Copenhagen interpretation will develop and dissolve beyond the framework of quantum mechanics as we move in the direction of ⑤ and ⑥. The Copenhagen Interpretation will go in the direction of ⑧. In other words, the Copenhagen Interpretation is not a rule governing the small scientific theory of quantum mechanics, but a rule governing the vast scientific theory (i.e. quantum language), including statistics. It would be a waste to confine the Copenhagen interpretation only to quantum mechanics Ⓔ. It doesn't matter if the Copenhagen interpretation is slightly inferior to other interpretations in quantum mechanics. Thus,

the Copenhagen Interpretation is eternal.

17.1 Socrates' absolutism was perfected by QL

The following is an excerpt from my website (https://ishikawa.math.keio.ac.jp/indexe.html).

Linguistic Copenhagen interpretation of Quantum language

Socrates' absolutism was perfected by QL

[I]:The whole picture of Quantum Language (=QL) is explained in detail in the following two:

[HWP]: History of Western Philosophy from the quantum theoretical point of view; [Ver.5] (2023) [https://philarchive.org/rec/ISHHOW-3]

The main point is the derivation of fuzzy logic from QL. It is the part of [analytical philosophyⓀ → QLⓁ] written in Figure 1 below.

[LCI]: Linguistic Copenhagen interpretation of quantum mechanics: Quantum Language [Ver. 6] (2023) [https://philarchive.org/rec/ISHLCI-3]

The main point is the derivation of statistics from QL. It is the part of [statisticsⒼ → QLⓁ] written in Figure 1 below.

The above two can be summarized as follows:

- It is common knowledge that "Newtonian mechanics" is not a branch of mathematics, but a theory of everyday cases of "the theoory of relativity (theory of monistic realism)." Now, in university, "logic" is emphasized in the humanities and "statistics" in the sciences, but "logic" and "statistics" are not branches of mathematics but these of "quantum language (theory of dualistic idealism)." In other words, logic and statistics is used for the rough matters and statistics for the fairly precise stuff respectively .

However, this homepage is written so that it can be read without knowing [HWP] and [LCI].
==

[II : Latest Results; The perfection of Socrates' absolutism.]:

This homepage will introduces the following latest paper:

[SOC]: History of Western Philosophy and QL (2024) [J. applied math. and Physics (Pdf download free)]

I chose this because I want the claim of [SOC] (i.e., the completion of Socrates' absolutism) to become common knowledge in general. Also, I think that every university student is able to read it regardless of their field of study.

Below is the Abstract:

What did they want to achieve? Were they just having fun with fashion?

- [Abstract]: The purpose of philosophy is diverse, but many philosophers acknowledge that the mainstream of Western philosophy (Socrates, Plato, Aristotle, Thomas Aquinas, Descartes, Kant, Wittgenstein) has progressed towards the completion of Socrates' absolutism. However, can absolutism still maintain its central position after analytical philosophy? There are pessimistic views on this issue, like that of R. Rorty, a leading figure of neo-pragmatism. Recently, I have proposed quantum language (including quantum mechanics, statistics, fuzzy sets, etc.). I believe this theory is not only one of the most fundamental scientific theories but also the scientific ultimate endpoint of Western philosophy. If so, Socrates' dream has come true. The purpose of this paper is to discuss the above and convey to the reader that quantum language has the power to cause a paradigm shift from a classical mechanical worldview to a quantum mechanical worldview.

Now, it is no exaggeration to say that the following [Figure 1: Diagram of two scientific world descriptions] is all I have to say. The argument for this diagram is as follows.

(A_0) There are two types of scientific thinking (monistic realism and dualistic idealism). I am not familiar with the former (e.g. TOE) because it is difficult to understand, but I consider as 'Newtonian mechanistic thinking' for the moment. But we are interested in the latter, everyday (non-physical) scientific thinking, which is a quantum language Ⓛ evolved from statistics Ⓖ or fuzzy logic Ⓚ. We therefore use 'scientific' here in the latter sense (sometimes also 'QL-like' to avoid confusion, e.g. 'QL-like' is used to mean 'scientific').

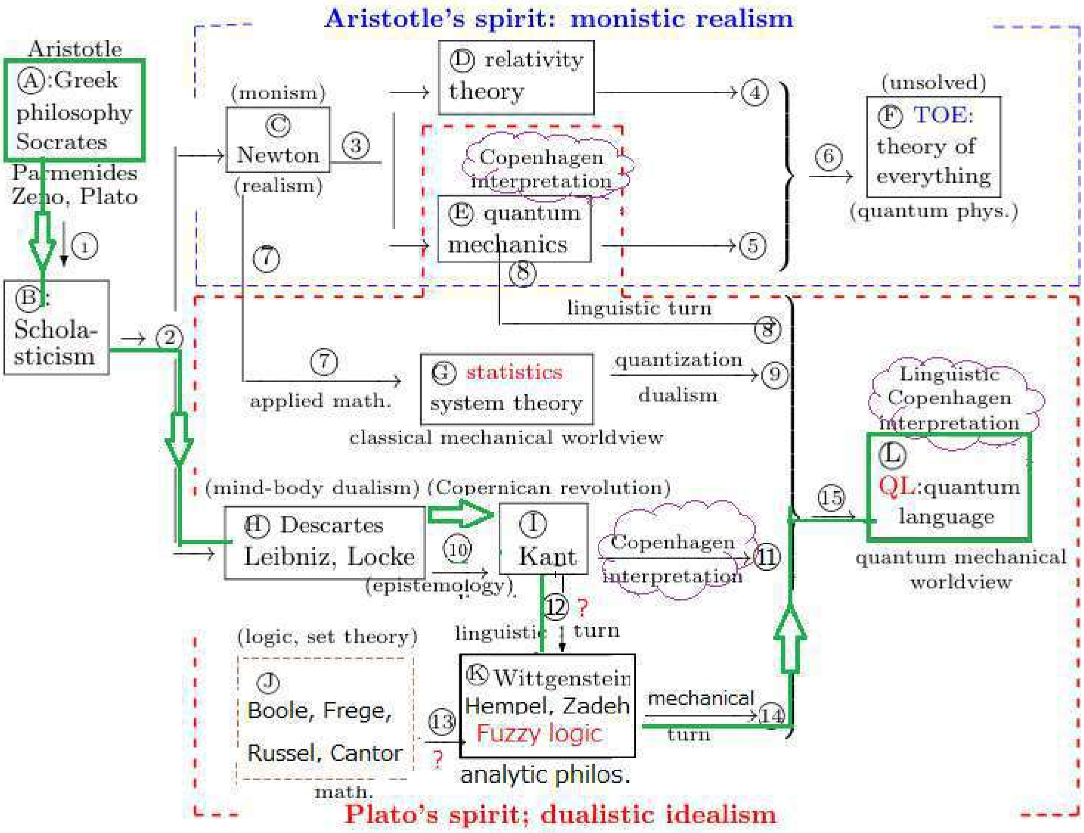


Figure 1: The locaction QL in the History of Western Philosophy

(the green road =the main stream of western philosophy history)

===
===

[III:Let's Get to the Main Topic; This is not the 'end of philosophy' but the 'completion of philosophy']:

We adopt the general convention of considering Socrates as the founder of philosophy. Therefore, we have:

Socrates (absolutism: pursuit of truth) vs. Protagoras (relativism: mastering rhetoric)

In ancient Athens, it was customary for citizens to gather in the agora, a public square, to freely debate. So how did one "win an argument"?

- Protagoras, the relativist, responded to this question by saying "improve your rhetoric skills"
- Socrates, the absolutist, said "speak the truth" (or, "Make the correct expression.")

. If I were in the agora, I would probably agree with Protagoras, but that's not where philosophy begins. So Socrates' disciples pursued the question, "What is absolute truth?" This pursuit has continued through

Plato, Aristotle, Augustine, Anselm, Thomas Aquinas, Descartes, Kant, and Wittgenstein,

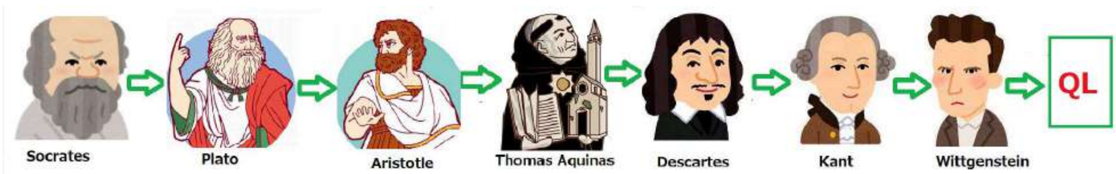
and has formed the mainstream of Western philosophical history. However, despite being pursued by the most brilliant geniuses of every era for the past 2,500 years, no clear answer has yet been found. That's why some people, like Rorty (the flag bearer of neo-pragmatism), say, "Let's give up on the pursuit of truth here." If Rorty says something like that, I would think that Rorty's opinion may be correct, but still, the stubborn pursuit of true

If so, I think everyone would agree with the following:

(A_1) The most important problem in Western philosophy is the completion of Socratic absolutism, i.e., the final settlement of the mainstream (Plato, Aristotle, Augustine, Anselmus, Thomas Aquinas, Descartes, Kant, Wittgenstein)

And the answer of this paper is as follows.

(A_2) Next for Wittgenstein is QL, which is the perfection of Socratic absolutism.



Some readers might be thinking:

- Why are there no names like Spinoza, Hegel, Nietzsche, Husserl, Heidegger, Sartre etc?

The reason is simple: their achievements are not scientific (=QL). That does not mean that they are philosophically inferior to QL philosophers. However, the spirit that permeates the mainstream of Western philosophy is "scientific."

==
==

[IV:The explanation of the answer (A_2)]:

Quantum language (=QL) is a mathematical extension of quantum mechanics (=QM), so it has the following form similar to quantum mechanics:

$$\boxed{\text{Quantum language}} = \overset{\text{Axiom 1}}{\boxed{\text{measurement}}} + \overset{\text{Axiom 2}}{\boxed{\text{causality}}} + \boxed{\text{Copenhagen interpretation}} \quad (1)$$

The reader may naturally request that "[Axioms 1 and 2] be clearly stated here," but since this is at the graduate school level, I will omit this.
Axiom 1 is an axiom about measurement, and Axiom 2 is an axiom about motion (equation of motion). However, since QL is a mathematical extension of quantum mechanics, it is no longer physics, and Axioms 1 and 2 become something like incomprehensible spells. Therefore, quantum mechanics is physics (realism), but QL is idealism. As shown in the diagram below, measurement is a concept consisting of three parts: the measurer (measured value), the measuring device (observable), and the measured thing (state). However, following convention, it is called dualism (rather than trinism). In summary, quantum language is dualistic idealism, and the mainstream of Western philosophy is also dualistic idealism. In other words, the two can be discussed within the following Cartesian diagram.

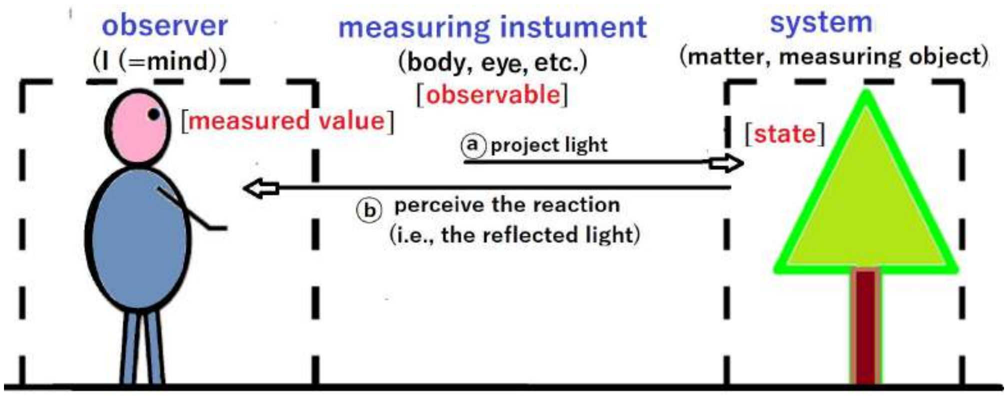


Figure 2: Descartes Figure [The three keywords : mind (measured value)-- body(observable)-- matter(state)]

[Axioms 1 and 2] are spells, and we don't know how to use them. Therefore, we need a manual on how to use [Axioms 1 and 2]. This manual is the Copenhagen interpretation. Therefore,

(B_1) The Copenhagen Interpretation is a manual on how to use [Axioms 1 and 2]

The word "manual" means that you can learn how to use [Axioms 1 and 2] by trial and error without looking at a manual, and Dr. Mermin's famous "Shut up and calculate!" can be considered a similar definition to (B_1). Or rather, I would like to consider the following definition of (B_2) to be the true one.
However, there is another way of thinking about this. This is when you do not know [Axioms 1 and 2]. In

this case, it is as follows.

(B_2) The Copenhagen Interpretation is a memo that records things that are obvious in the world of dualistic idealism.

(To be more specific, it is a memo that records things that are obvious in the world of dualistic idealism, but not obvious to our normal senses.)

The author's preference is definition (B_2) and does not wish to adopt definition (B_1) if possible. Our Copenhagen interpretation is specifically enumerated as follows [footnote[1]]

(c_1)Always think with the Cartesian Figure in mind.

(c_2)Measure only once.

(c_3)Measurers have no space-time.

(c_4)State does not change (There is no movement)

(c_5)... (Maybe endless)

etc. Also, I think as follows.

- The Copenhagen Interpretation is attached to quantum language, not to quantum mechanics. However, when quantum mechanics is viewed as [quantum mechanics $\subset$ quantum language], the Copenhagen interpretation appears to be attached to quantum mechanics. The many-worlds interpretation of quantum mechanics does not consider [quantum mechanics $\subset$ quantum language], so it is not associated with the Copenhagen interpretation.

Many readers may think, "I can understand definition (B_1), but definition (B_2) is impossible." However, what is surprising is that most of the Copenhagen Interpretation was discovered before the birth of quantum mechanics. In fact, things like 2, ..., 4 above were discovered by Parmenides (c. 520 BC - 450 BC), who was 50 years older than Socrates. However, Parmenides was too much of a genius, and it is probably the consensus among philosophy lovers that the position of the father of philosophy should belong to Socrates. Even so, Parmenides had a clear vision of the world of dualistic idealism and discovered the Copenhagen Interpretation. Moreover, he discovered it under definition (B2) (without assuming Axioms 1 and 2), which means he was nothing short of a genius. The field of Western philosophical history is full of geniuses, but Parmenides (as well as Thomas Aquinas and Descartes) is surprising. Thus we see:

(D) The mainstream of Western philosophy has pursued the question, "What is the world of dualistic idealism (the world in which "things" and "mind" are intertwined)?" In other words, the Copenhagen Interpretation has been sought.

That the search for the Copenhagen interpretation runs through the mainstream of Western philosophy [Plato, Augustine, Anselmus, Thomas Aquinas, Descartes, Kant, (excluding Aristotle and Wittgenstein)] is evident in the following table.

[Plato, Aristotle, Scholasticism, Descartes, Wittgenstein and QL]

Plato (Allegory of the Sun)	actual world	Idea(=sunlight)	Idea world
Aristotle (monism)	/	/	hule [eidos]
Scholasticism (Anselmus)		universal	individual
Scholasticism (Thomas Aquinas)	human intellect (universale post rem)	divine intellect (universale ante rem)	individual (universal in re)
Descartes, Locke, Kant (epistemology)	[A](= mind)	[B](=body (≈ sensory organ)) (Mediating of A and C)	[C](= matter)
Wittgenstein (analytic philosophy)		logic	
quantum language	**measurer** [measured value]	**measuring instrument** [observable]	**system** [state]

Figure 3: The mainstream of Western philosophical history [the history of clarifying the three keywords (mind, body, matter)] (Aristotle is a monistic realist)

But the story is not simple. The problem is the [logic] part of analytic philosophy (Wittgenstein). Anyone looking at [Figure 3] would think that

(E_1) Isn't "analytical philosophy" unique and out of touch with the mainstream of Western philosophy?

とか

(E_2) Descartes-Kant's epistemology is meaningless because it is metaphysics. True philosophy begins with analytic philosophy!

etc. It is natural that a variety of opinions erupt. Figure 1 is shown again:

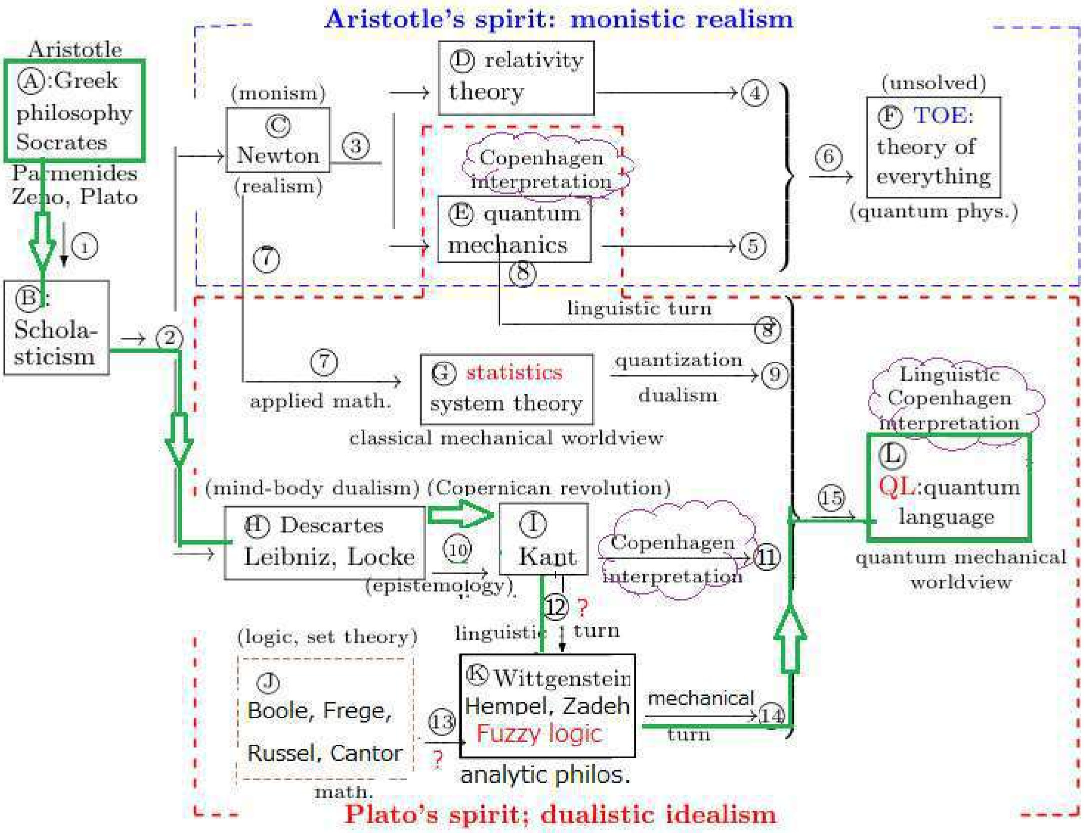


Figure 1: The locaction QL in the History of Western Philosophy

Now, as we saw above, everyone would think that

(E_3) "⑫: linguistic turn" is unreasonable (i.e., Ⓘ and Ⓚ not connected)

Wittgenstein discarded (Descartes-Kant) epistemology as metaphysics and established analytical philosophy.

Wittgenstein's brilliant eloquence led philosophy lovers to believe in the linguistic turn: Ⓘ → Ⓚ, and this has continued for 100 years.
If you think about it normally, even if "logic" is the foundation of mathematics and is important, philosophy does not go out of its way to emphasize the "importance of logic." If mathematicians or physicists emphasize the "importance of logic," it would be somewhat understandable, but they do not go out of their way to say such obvious things. To begin with, his TLP (= Tractatus Logico-philosophicus) is not written logically, and Wittgenstein is not a philosopher who thinks logically.

About 10 years ago, I thought the following:

(E_4) Analytical philosophy is a philosophy that the magician Wittgenstein created in the midst of the scientific revolution (abstract mathematics, quantum mechanics, and relativity) in the early 20th century in order to break the "Epistemological rut." He was blinded by the surprising fact that logic is the language of (abstract) mathematics and jumped on "logic," but it is doubtful that "logic" is the main theme of philosophy. It has nothing to do with Socratic absolutism.

==
==

[V: The most important issue in the history of Western philosophy is the relationship between Kant Ⓘ and Wittgenstein Ⓚ]:

About 10 years ago, my opinion was (E_4), but now I think as follows.

(F) **Wittgenstein was wrong to discard "epistemology". However, Wittgenstein's choice of "logical" is quite commendable. Maybe 50 points. But to get 100 points, he should have chosen "scientific".**

Let me show this below.
We must free ourselves from the spell of Wittgenstein and look squarely at the absolute truth of Socrates. In other words, we must look at it from the perspective of QL. Then, by expanding the essential part of Figure 1, we obtain the following:

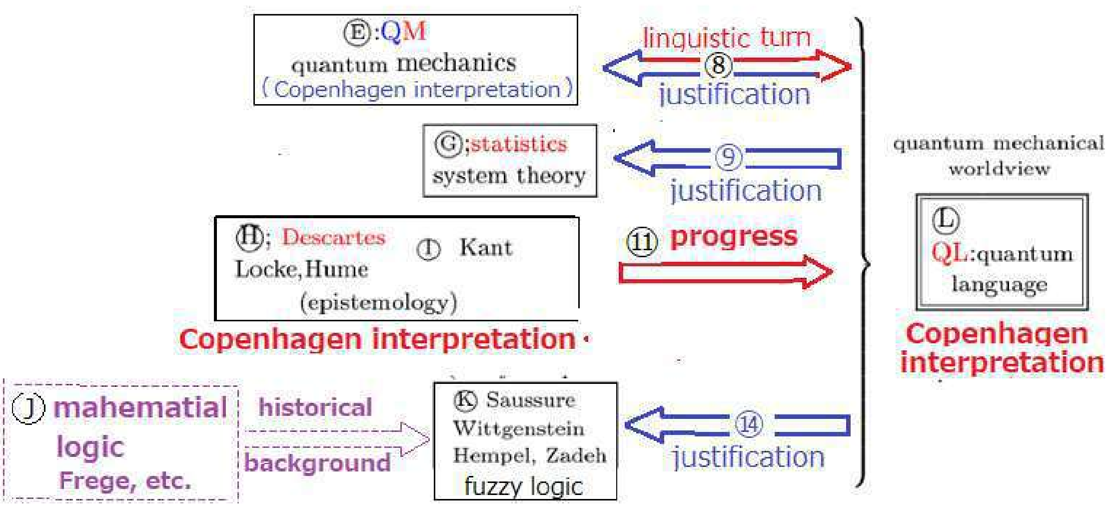


Figure 4: The view [⑧,⑨,⑪,⑭] from QL (=end point)

The following has the same meaning as Figure 4:

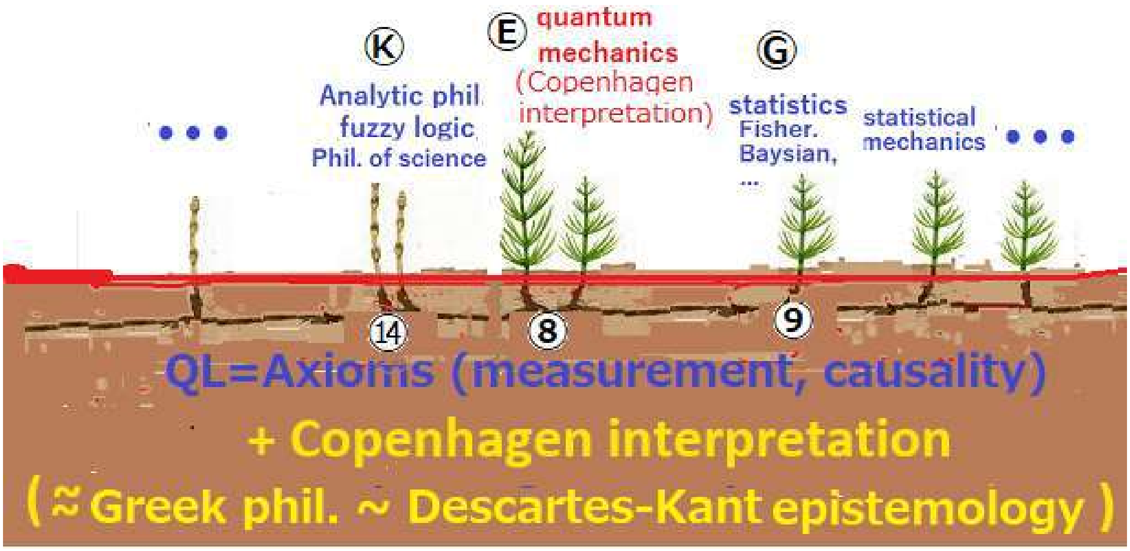


Figure 5:The strangeness of quantum mechanics arises from the strangeness of the Copenhagen Interpretation.
Statistics is a theory that has sealed off the Copenhagen Interpretation to prevent its strangeness from surfacing, and has chosen to be a type of applied mathematics.

Hence we see :

(G_1) QM(⑧),statistcs(⑨) and fuzzy logic(⑭) are derived from QL. Therefore, these are connected These are also connected to the underlying Copenhagen interpretation (≈ Descartes=Kant epistemology) [footnote[2]]

Therefore, quantum theory, statistics, and analytical philosophy (fuzzy logic) can all be deduced from LQ (= [AXioms 1 and 2] + Copenhagen interpretation).
Therefore,

(G_2) Wittgenstein was not wrong to emphasise the importance of being 'logical'. But he really should have emphasised the importance of 'scientific (i.e. QL)'. It was also a big mistake to abandon 'Cartesian-Kantian epistemology'.

That is,

(H) Like anything, you can't see the whole picture until you see it from the top of the mountain.

Just to be clear, analytical philosophy Ⓚ has no relation to mathematical logic Ⓙ, just as physics Ⓕ has no relation to mathematical logic Ⓙ. Research that considers the results of mathematical logic, such as Gödel's incompleteness theorem, from a philosophical perspective may be important, but this has nothing to do with what Wittgenstein aimed for.

==
==

[VI: Conclution] :

The above showed that.

(A_2) QL is the perfection of Socratic absolutism (i.e., dualistic idealism)

(I think this is the most important claim in philosophy. If it were not so, we would not be able to answer the question, "What was the mainstream of the history of Western philosophy?" In other words, we would not know what the great philosophers depicted below were trying to do.)

We were not doing philosophy as fashion. We were running a relay race aiming for the QL (= goal).

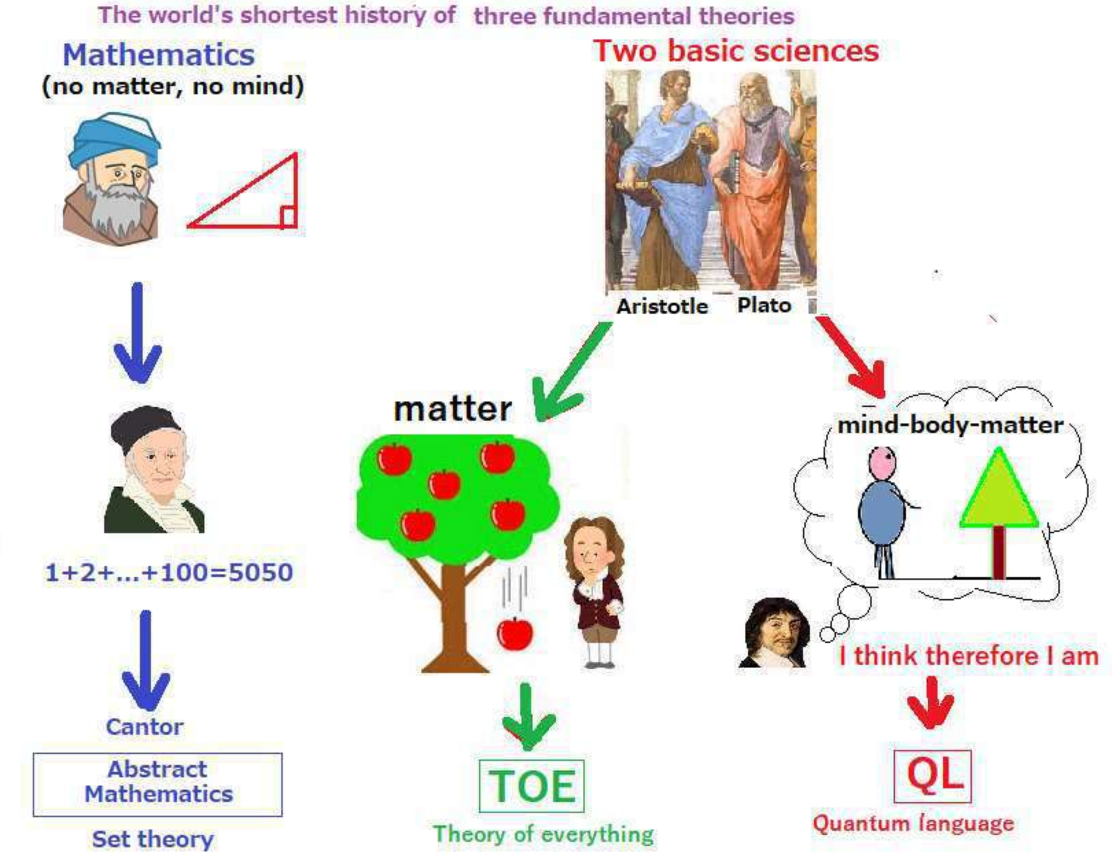
Now we can finally say "what Wittgenstein wanted to say".

(I) What we cannot speak about in QL, we must pass over in silence.

==

[VII:Supplement]:

(i):In the first year of university, regardless of whether you are a liberal arts or science student, there are lectures on mathematics, physics, and statistics. This means that these three are the "most important basic theories" and correspond to the following three.



(Abbreviation for Figure 1)

Mathematics and QL have reached their destination, and now we are in the details. However, the TOE of physics is still unknown. I think that human wisdom must solve these three problems. Physicists need to hurry, otherwise AI will do it first!

(ii): Another question I've been wondering about is, "Is there a fourth fundamental theory?" I believe there isn't, but I'm not confident.

==

[footnote]

1. The author's preference is (B2) and does not want to adopt (B1) if possible. Therefore, we do not agree that something like 'contraction of the wave function' is one of the Copenhagen interpretations; in QL, the 'projective canon' is a theorem (cf. [LCI]). In the same sense that there is 'no perfect manual', we believe there is no 'perfect Copenhagen Interpretation'. There may be no Copenhagen Interpretation that clears all the examples given by researchers in the 'philosophy of mind' (e.g. the 'aquarium brain'). However, we think that a Copenhagen Interpretation that covers the problems we are likely to encounter in practice is possible. Some researchers might want to debate whether to adopt the 'Heisenberg cut' as one of the Copenhagen interpretations, but the author's aesthetic sense would have to be reluctant to do so. However, this is not to say that the 'Heisenberg Cut' is totally rejected. We hope that many people will try to propose a 'user-friendly Copenhagen interpretation'. back
2. What Wittgenstein writes in his Tractatus that cannot be derived from QL is unscientific. back