Title	Chapter 2 : Measurements (Axiom 1)
Sub Title	
Author	石川, 史郎(Ishikawa, Shiro)
Publisher	Keio University Press Inc.
Publication year	2006
Jtitle	Mathematical Foundations of Measurement Theory (測定理論の数学的基礎). (2006.) ,p.15- 54
JaLC DOI	
Abstract	
Notes	
Genre	Book
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO52003001-00000000-0015

Chapter 2

Measurements (Axiom 1)

Measurement theory (MT) is classified two subjects, i.e., "(pure) measurement theory (PMT)" and "statistical measurement theory (SMT)". That is,

$$\text{MT (=“measurement theory”)} \begin{cases} \text{PMT (=“(pure) measurement theory”) in Chapters } 2 \sim 7 \\ \\ \text{SMT (=“statistical measurement theory” in Chapters } 8 \sim) \end{cases} \tag{2.1}$$

PMT is essential, and it should be noted that *there is no SMT without PMT* (*cf.* Chapter 8). In Chapters 2 ∼ 7, we devote ourselves to PMT, which is formulated as follows:

$$\text{PMT} = \underset{\text{[Axiom 1 (2.37)]}}{\text{measurement}} + \underset{\text{[Axiom 2 (3.26)]}}{\text{the relation among systems}} \quad \text{in } C^*\text{-algebra} . \tag{2.2}$$
(=(1.4))

In this chapter we intend to explain "measurement (= Axiom 1)". (In Chapter 3 we will devote ourselves to Axiom 2 (i.e., "the relation among systems").)

2.1 Mathematical preparations

The theory of operator algebras (i.e, C^*-algebra and W^*-algebra) is a convenient mathematical tool to describe both classical and quantum mechanics (*cf.* [76]). Thus our theory is described in terms of C^*-algebras. Since our concern in this book is mainly concentrated on classical systems and not quantum systems, it may suffice to deal with only commutative C^*-algebras. In fact, most of our main results are related to classical systems. However, recall (1.4), that is:

$$\boxed{\text{PMT}} = \text{“Apply Axioms 1 and 2 to every phenomenon by an analogy of quantum mechanics”} \tag{2.3}$$
(=(1.4))

Thus we think that the essence of measurements can not be appreciated deeply without the knowledge of quantum measurements. In fact, the concept of measurements was first discovered and formulated by M. Born [13][1] as the most fundamental concept in quantum mechanics. Thus, we begin with general C^*-algebras, in which both classical and quantum systems are formulated.[2]

Let $\mathcal{A}$ be a linear associative algebra over the complex field $\mathbf{C}$. The algebra $\mathcal{A}$ is called a *Banach algebra* if it is associated to each element T a real number $\|T\|$, called the *norm* of T, with the properties:

$$\text{(i) } \|T\| \geq 0, \quad \text{(ii) } \|T\| = 0 \text{ if and only if } T = 0, \text{(i.e., the 0-element in } \mathcal{A}),$$
$$\text{(iii) } \|T+S\| \leq \|T\| + \|S\|, \quad \text{(iv) } \|\lambda T\| = |\lambda| \cdot \|T\|, \ \lambda \in \mathbf{C},$$
$$\text{(v) } \|TS\| \leq \|T\| \cdot \|S\|, \quad \text{(vi) } \ \mathcal{A} \text{ is complete with respect to the norm } \|\cdot\|.$$

A mapping $T \mapsto T^*$ of $\mathcal{A}$ into itself is called an *involution* (and T^* is called the *adjoint element* of T) if it satisfies the following conditions:

$$\text{(i) } (T^*)^* = T, \qquad \text{(ii) } (T+S)^* = T^* + S^*, \qquad \text{(iii) } (TS)^* = S^* T^*,$$

$$\text{(iv) } (\lambda T)^* = \bar{\lambda} T^*, \ \lambda \in \mathbf{C}.$$

A Banach algebra with an involution * is called a *Banach *-algebra.*

Definition 2.1. [C^*-algebra, identity, commutative C^*-algebra]. *A Banach *-algebra $\mathcal{A}$ (with the norm $\|\cdot\|_{\mathcal{A}}$) is called a C^*-algebra if it satisfies the C^*-condition, i.e., $\|T^*T\| = \|T\|^2$ for any $T \in \mathcal{A}$. A C^*-algebra $\mathcal{A}$ does not always have the identity element $I_{\mathcal{A}}$ (i.e., $I_{\mathcal{A}}T = TI_{\mathcal{A}} = T$ for all $T \in \mathcal{A}$), though in this book we usually suppose that a C^*-algebra $\mathcal{A}$ has the identity element $I_{\mathcal{A}}$. A C^*-algebra $\mathcal{A}$ is called* unital, *if it has the identity element $I_{\mathcal{A}}$. Also, a C^*-algebra $\mathcal{A}$ is called* commutative, *if it holds that $T_1T_2 = T_2T_1$ $(\forall T_1, T_2 \in \mathcal{A})$.*

■

An element F in $\mathcal{A}$ is called *self-adjoint* if it holds that $F = F^*$. A self-adjoint element F in $\mathcal{A}$ is called *positive* (and denoted by $F \geq 0$) if there exists an element F_0 in $\mathcal{A}$ such

[1] He proposed his theory in 1926, and he won the Nobel prize of physics in 1954.

[2] I am afraid that the mathematical preparation (in this section) discourages readers to want to read this book. Thus, it may be recommended to skip to Example 2.16 firstly. In order to read this book, it suffices to understand Example 2.16.

that $F = F_0^* F_0$ where F_0^* is the adjoint element of F_0. Also, a positive element F is called a *projection* if $F = F^2$ holds. Let $\mathcal{A}^*$ be the dual Banach space of $\mathcal{A}$. That is,

$$\mathcal{A}^* = \{\rho \mid \rho : \mathcal{A} \to \mathbf{C} \text{ is continuous linear } \}$$

with the norm $\|\rho\|_{\mathcal{A}^*}$ ($\equiv \sup\{|\rho(F)| \;:\; \|F\|_{\mathcal{A}} \leq 1\}$). $\Big($The linear functional $\rho(F)$ is sometimes denoted by ${}_{\mathcal{A}^*}\Big\langle \rho, F \Big\rangle_{\mathcal{A}}$.$\Big)$ Define the *mixed state space* $\mathfrak{S}^m(\mathcal{A}^*)$ such that:

$$\mathfrak{S}^m(\mathcal{A}^*) = \{\rho \in \mathcal{A}^* \mid \|\rho\|_{\mathcal{A}^*} = 1 \text{ and } \rho(F) \geq 0 \text{ for all } F \geq 0\}. \tag{2.4}$$

A mixed state ρ ($\in \mathfrak{S}^m(\mathcal{A}^*)$) is called a *pure state* if it satisfies that "$\rho = \theta\rho_1 + (1-\theta)\rho_2$ for some $\rho_1, \rho_2 \in \mathfrak{S}^m(\mathcal{A}^*)$ and $0 < \theta < 1$" implies "$\rho = \rho_1 = \rho_2$". Define

$$\mathfrak{S}^p(\mathcal{A}^*) \equiv \{\rho^p \in \mathfrak{S}^m(\mathcal{A}^*) \mid \rho^p \text{ is a pure state}\}, \tag{2.5}$$

which is called a *state space* (or *pure state space, phase space*). Note that $\mathfrak{S}^m(\mathcal{A}^*)$ is convex and compact in the weak* topology $\sigma(\mathcal{A}^*; \mathcal{A})$. Also, $\mathfrak{S}^p(\mathcal{A}^*)$ is characterized as the set of the extreme points of $\mathfrak{S}^m(\mathcal{A}^*)$. (Cf. [92, 76]). Since $\mathfrak{S}^p(\mathcal{A}^*)$ is the closed set of $\mathfrak{S}^m(\mathcal{A}^*)$, the $\mathfrak{S}^p(\mathcal{A}^*)$ is also compact in the weak* topology.

The following Examples 2.2 and 2.3 will promote the understanding of Definition 2.1.

Example 2.2. [Commutative C^*-algebras; $C(\Omega)$, or generally, $C_0(\Omega)$]. When $\mathcal{A}$ is a commutative C^*-algebra, that is, $T_1 \cdot T_2 = T_2 \cdot T_1$ holds for all T_1, $T_2 \in \mathcal{A}$, by Gelfand theorem (*cf.* [74, 76]), we can put $\mathcal{A} = C(\Omega)$, the algebra composed of all continuous complex-valued functions on a compact Hausdorff space Ω. $\Big($If the commutative C^*-algebra $\mathcal{A}$ is not necessarily assumed to be unital, we can put $\mathcal{A} = C_0(\Omega)$, the algebra composed of all continuous complex-valued functions *vanishing at infinity* on a *locally* compact Hausdorff space Ω.$\Big)$ The norm $\|f\|_{C(\Omega)}$ is, of course, defined by $\|f\|_{C(\Omega)} = \max\{|f(\omega)| \;:\; \omega \in \Omega\}$ $(\forall f \in C(\Omega))$. Also, we can easily see that it satisfies the C^*-condition, i.e., $\|f \cdot f^*\|_{C(\Omega)} = \|f\|^2_{C(\Omega)}$ where $f^*(\omega)$ ($\equiv \overline{f(\omega)}$) is defined by the conjugate "$\mathfrak{Re}[f(\omega)] - \mathfrak{Im}[f(\omega)]i$" $(\forall \omega \in \Omega)$ (where $\mathfrak{Re}$ is "real part", $\mathfrak{Im}$ is "imaginary part"). It is well known (i.e., Riese Theorem) that $C(\Omega)^* = \mathcal{M}(\Omega)$, i.e., the Banach space composed of all regular complex-valued measures on Ω. And therefore,

$$\mathfrak{S}^m(\mathcal{M}(\Omega)) = \{\rho \in \mathcal{M}(\Omega) \mid \rho \geq 0, \|\rho\|_{\mathcal{M}(\Omega)} = 1\}, \tag{2.6}$$

which is denoted by $\mathcal{M}_{+1}^m(\Omega)$. Also, it is clear that

$$\mathfrak{S}^p(\mathcal{M}(\Omega)) = \Big\{ \delta_\omega \in \mathcal{M}(\Omega) \mid \delta_\omega \text{ is a point measure at } \omega \in \Omega \Big\} \tag{2.7}$$

$\Big($i.e., ${}_{\mathcal{M}(\Omega)}\Big\langle \delta_\omega, f \Big\rangle_{C(\Omega)} = f(\omega)$ $(\forall f \in C(\Omega), \forall \omega \in \Omega)\Big)$, which is denoted by $\mathcal{M}_{+1}^p(\Omega)$, and called a *state space.* And therefore, we have the identification: $\Omega \approx \mathcal{M}_{+1}^p(\Omega)$ in the sense of

$$\Omega \ni \omega \longleftrightarrow \delta_\omega \in \mathcal{M}_{+1}^p(\Omega). \tag{2.8}$$

Thus the compact Hausdorff space Ω may be also called a *state space.*

■

Example 2.3. [Non-commutative C^*-algebras; $B(V)$ and $\mathcal{C}(V)$]. Let V be a (separable) Hilbert space with the inner product $\langle \cdot, \cdot \rangle_V$. Here we always assume that $\langle v_1, \alpha v_2 \rangle_V = \alpha \langle v_1, v_2 \rangle_V$ $(\forall v_1, v_2 \in V, \alpha \in \mathbf{C})$. (Cf. [4, 71].) Put

$$B(V) = \{ T \;:\; T \text{ is a bounded linear operator from a Hilbert space } V \text{ into itself } \}.^3$$

Define $\|T\|_{B(V)} = \sup\{\|Tv\|_V \;:\; \|v\|_V = 1\}$, and $(T_1T_2)(v) = T_1(T_2 v)$ $(\forall v \in V)$. And T^* is defined by the adjoint operator of T. Note that it holds that $\|T^*T\|_{B(V)} = \|T\|^2$ $(\forall T \in B(V))$. Thus, we can easily see that the $B(V)$ is a non-commutative C^*-algebra. Also note that

$$\mathcal{C}(V) \equiv \{ T \in B(V) : T \text{ is a compact operator } \} \tag{2.9}$$

is a C^*-subalgebra of $B(V)$. If the dimension of V is infinite, this C^*-algebra $\mathcal{C}(V)$ has no identity I. We see that

$$\mathcal{C}(V)^* = Tr(V)\Big(\equiv \{ T \in B(V) \;:\; \|T\|_{tr} < \infty \}\Big). \tag{2.10}$$

Here $Tr(V)$ is the class of trace operators with the trace norm $\|\cdot\|_{tr}$ such that:

$$\|\rho\|_{tr} = \sum_{n=1}^{\infty} \langle e_n, \sqrt{\rho^* \rho}\, e_n \rangle_V$$

where $\{e_n\}_{n=1}^\infty$ is the complete orthonormal system in V. It is well known that the value $\|\rho\|_{tr}$ is independent of the choice of a complete orthonormal basis $\{e_\lambda | \lambda \in \Lambda\}$ in V. And we see

$$\mathfrak{S}^m(\mathcal{C}(V)^*) = Tr_{+1}^m(V) \equiv \{ \rho \in Tr(V) : \rho \geq 0, \|\rho\|_{Tr(V)} = 1 \}. \tag{2.11}$$

[3] "bounded linear operator" = "continuous linear operator" (*cf.* [92])

And further,

$$(Tr(V))^* = B(V).$$

Also, it is well known that

$$\text{“}\rho \in \mathfrak{S}^p(\mathcal{C}(V)^*)\text{”} \Leftrightarrow \text{“there exists } \psi \in V\ (\|\psi\|_V = 1) \text{ such that } \rho = |\psi\rangle\langle\psi|\text{”}, \tag{2.12}$$

where the Dirac notation "$|\psi_1\rangle\langle\psi_2|$" $\Big(\in B(V)\Big)$, $\psi_1, \psi_2 \in V$, is defined by

$$(|\psi_1\rangle\langle\psi_2|)\phi = \langle\psi_2, \phi\rangle_V \psi_1 \qquad \text{for all } \phi \in V.$$

The state space $\mathfrak{S}^p(\mathcal{C}(V)^*)$ is denoted by $Tr^p_{+1}(V)$, that is,

$$\mathfrak{S}^p(\mathcal{C}(V)^*) \equiv Tr^p_{+1}(V).$$

Also, it is well-known that "$\rho \in \mathfrak{S}^m(\mathcal{C}(V)^*)$" $\Leftrightarrow$ "there exists an orthonormal system $\{\psi_n\}_{n=1}^\infty$ in V and non-negative real numbers $\{\lambda_n\}_{n=1}^\infty$ (where $\sum_{n=1}^\infty \lambda_n = 1$) such that $\rho = \sum_{n=1}^\infty \lambda_n |\psi_n\rangle\langle\psi_n|$".

■

The following theorem is one of the most important theorems in the theory of operator algebras.

Theorem 2.4. [GNS-construction, Gelfand, Naimark, Siegel, *cf.* [50, 76]]. *Let $\mathcal{A}$ be a C^*-algebra. Then there exists a $B(V)$ such that:*

$$\mathcal{A} \subseteq B(V). \tag{2.13}$$

That is, $\mathcal{A}$ can be identified with the norm-closed C^*-subalgebra of a certain $B(V)$.

■

Example 2.5. [Commutative C^*-algebra $\mathrm{Mat}^{\mathrm{D}}(n; \mathbf{C})$ as the subalgebra of $B(\mathbf{C}^n)$]. Let $\mathbf{C}^n$ be an n-dimensional Hilbert space with the inner product $\langle\cdot,\cdot\rangle_{\mathbf{C}^n}$ (that is, $\|z\|_{\mathbf{C}^n} = \sqrt{\sum_{k=1}^n |z_k|^2}$ $(\forall z = (z_1, z_2, ..., z_n) \in \mathbf{C}^n)$). Consider the non-commutative C^*-algebra

$$B(\mathbf{C}^n) \equiv \{T \ :\ T \text{ is a (bounded) linear operator from a Hilbert space } \mathbf{C}^n \text{ into itself }\},$$

which is clearly equal to

$$\mathrm{Mat}(n; \mathbf{C}) \equiv \{T \ :\ T \text{ is a complex } (n \times n)\text{-matrix }\}. \tag{2.14}$$

That is,

$$B(\mathbf{C}^n) = \mathrm{Mat}(n; \mathbf{C}).$$

Put

$$\mathrm{Mat}^{\mathrm{D}}(n; \mathbf{C}) = \{T \ : \ T \text{ is a complex } (n \times n)\text{-diagonal matrix } \}, \tag{2.15}$$

which is clearly a commutative C^*-subalgebra of $B(\mathbf{C}^n)$. Also, it is obvious that the $\mathrm{Mat}^{\mathrm{D}}(n; \mathbf{C})$ is equivalent to $C(\Omega)$, where Ω is the finite state space ($\{1, 2, ..., n\}$) with the discrete topology.[4] That is, we see the following identifications:

$$\mathrm{Mat}^{\mathrm{D}}(n; \mathbf{C}) \approx C(\{\omega_1, \omega_2, ..., \omega_n\}) \approx \mathbf{C}^n$$

where $\mathbf{C}^n$ is assumed to have the max-norm $\|z\|_{\mathbf{C}^n}^{\max} \Big(= \max_{k=1,2,...,n} |z_k| \ (\forall z = (z_1, z_2, ..., z_n) \in \mathbf{C}^n) \Big)$. Also, the multiplication $(z_1^1, z_2^1, ..., z_n^1) \cdot (z_1^2, z_2^2, ..., z_n^2)$ is defined by $(z_1^1 z_1^2, z_2^1 z_2^2, ..., z_n^1 z_n^2)$.

■

Remark 2.6. [(i): The identity]. Let $\mathcal{A}_0$ be a non-unital C^*-algebra. Theorem 2.4 (The GNS-construction) says that there exists a $B(V)$ such that $\mathcal{A}_0 \subseteq B(V)$. That is, $\mathcal{A}_0$ can be identified with the norm-closed subalgebra of $B(V)$. Thus we can define the C^*-algebra $\mathcal{A}_I$ such that it is the smallest algebra that includes $\{I\} \cup \mathcal{A}_0$ ($\subseteq B(V)$). Therefore, we can always add the identity I to $\mathcal{A}_0$, and construct the new unital C^*-algebra $\mathcal{A}_I$. This argument implies that the "unital condition" is not so strict. Thus, throughout this book, we usually deal with a unital C^*-algebra, though the $C_0(\Omega)$ is sometimes used.
[(ii): Minimal tensor C^*-algebras]. Here consider the minimal tensor C^*-algebra as follows: Let $\widehat{\mathcal{A}} \ \big(= \bigotimes_{k=1}^n \mathcal{A}_k \ \big)$ be the tensor product C^*-algebra of $\{\mathcal{A}_k \ : k = 1, 2, ..., n\}$. This can be easily constructed as follows: Since we can see, by Theorem 2.4 (GNS-construction), that

$$\mathcal{A}_k \subseteq B(V_k) \qquad (k = 1, 2, ..., n), \tag{2.16}$$

we can define $\bigotimes_{k=1}^n \mathcal{A}_k$ such that the smallest norm-closed sub-algebra (of $B(\bigotimes_{k=1}^n V_k)$) that contains

$$\Big\{ \bigotimes_{k=1}^n F_k \Big(\in B(\bigotimes_{k=1}^n V_k) \Big) \ \Big| \ F_k \in \mathcal{A}_k, k = 1, 2, ..., n \Big\} \tag{2.17}$$

[4] Throughout this book, we assume that a finite state space Ω ($\equiv \{\omega_1, \omega_2, ..., \omega_n\}$) has the discrete metric d_D (i.e., $d_D(\omega_1, \omega_2) = 1$ ($\omega_1 \neq \omega_2$), $= 0$ ($\omega_1 = \omega_2$)).

where $\bigotimes_{k=1}^n V_k$ is the tensor Hilbert space of $\{V_k \mid k=1,2,...,n\}$. Though the general theory of tensor product C^*-algebras $\bigotimes_{k=1}^n \mathcal{A}_k$ is not easy, we only use the following properties (i)∼(iii) of the tensor C^*-algebras:

(i) $T_1 \otimes T_2 \otimes \cdots \otimes T_n \in \widehat{\mathcal{A}}$ for any $T_k \in \mathcal{A}_k$, $k=1,2,...,n$,

(ii) $\rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_n \in \mathfrak{S}^p(\widehat{\mathcal{A}}^*)$ for any $\rho_k \in \mathfrak{S}^p(\mathcal{A}_k^*)$, $k=1,2,...,n$,

(iii) $(\rho_1 \otimes \rho_2 \otimes \cdots \otimes \rho_n)\ (T_1 \otimes T_2 \otimes \cdots \otimes T_n) = \prod_{k=1}^n \rho_k(T_k)$ for any $\rho_k \in \mathcal{A}_k^*$ and any $T_k \in \mathcal{A}_k$, $k=1,2,...,n$.

If we focus on only commutative cases, it is sufficient to know the fact that

$$\bigotimes_{k=1}^n C(\Omega_k) = C(\mathop{\times}_{k=1}^n \Omega_k) \qquad \text{and} \qquad \bigotimes_{k=1}^n \mathcal{M}(\Omega_k) = \mathcal{M}(\mathop{\times}_{k=1}^n \Omega_k), \tag{2.18}$$

where $\times_{k=1}^n \Omega_k$ is the product topological space of Ω_1,...,Ω_n. Therefore, for example, the above (iii) implies the elementary property of product measure (Fubini's theorem), i.e.,

$$\int_{\Omega_1 \times \Omega_2} f_1(\omega_1) \cdot f_2(\omega_2)(\rho_1 \otimes \rho_2)(d\omega_1 d\omega_2) = \int_{\Omega_1} f_1(\omega_1)\rho_1(d\omega_1) \cdot \int_{\Omega_2} f_2(\omega_2)\rho_2(d\omega_2)$$
$$(\forall f_1 \in C(\Omega_1), \quad \forall f_2 \in C(\Omega_2)). \tag{2.19}$$

For the deep studies of "tensor C^*-algebra", see [50].

■

2.2 Observables

Let X be a set. Let 2^X (or, $\mathcal{P}(X)$) be the power set of X. i.e., $2^X = \{\Xi \mid \Xi \subseteq X\}$. A set $\mathcal{F}(\subseteq 2^X)$ is called a *field* if the $\mathcal{F}$ is closed under the intersection (i.e., $\cap$) and the compliment (i.e., $[\ \cdot\ ]^c$), that is, if "Ξ_1, $\Xi_2 \in \mathcal{F}$" implies "$\Xi_1 \cap \Xi_2 \in \mathcal{F}$" and "$\Xi_1^c \in \mathcal{F}$", where $\Xi_1^c = X \setminus \Xi_1 = \{x \mid x \in X \wedge x \notin \Xi_1\}$. Note that $\Xi_1 \cup \Xi_2 = (\Xi_1^c \cap \Xi_2^c)^c$, $\Xi_1 \setminus \Xi_2 = \Xi_1 \cap \Xi_2^c$ and $\Xi_1 \triangle \Xi_2 = (\Xi_1 \cup \Xi_2) \setminus (\Xi_1 \cap \Xi_2)$. Thus the field $\mathcal{F}$ is also closed under the operations $\cup$, $\setminus$ and $\triangle$.

Also, a set $\mathcal{R}(\subseteq 2^X)$ is called a *ring* if the $\mathcal{R}$ is closed under the intersection (i.e., $\cap$) and the symmetric difference (i.e., $\triangle$), that is, if "Ξ_1, $\Xi_2 \in \mathcal{R}$" implies "$\Xi_1 \cap \Xi_2 \in \mathcal{R}$" and

"$\Xi_1 \triangle \Xi_2 \in \mathcal{R}$". Note that $\Xi_1 \cup \Xi_2 = (\Xi_1 \triangle \Xi_2) \triangle (\Xi_1 \cap \Xi_2)$, $\Xi_1 \setminus \Xi_2 = \Xi_1 \triangle (\Xi_1 \cap \Xi_2)$. Thus the ring $\mathcal{R}$ is also closed under the operations $\cup$ and $\setminus$ (*cf.* [29]).

Motivated by the Davies' idea (in quantum mechanics, *cf.* [17]), we propose the following definition.

Definition 2.7. [C^*-observables in a unital $\mathcal{A}$]. *A C^*-observable (or in short, observable, fuzzy observable)* $\mathbf{O} \equiv (X, \mathcal{F}, F)$ *in a unital C^*-algebra $\mathcal{A}$ is defined such that it satisfies that*

(i) [field]. *X is a set (called a "measured value set" or "label set"), and $\mathcal{F}$ is the subfield of the power set $\mathcal{P}(X)$ ($\equiv \{\Xi : \Xi \subseteq X\}$),*

(ii) *for every $\Xi \in \mathcal{F}$, $F(\Xi)$ is a positive element in $\mathcal{A}$ such that $F(\emptyset) = 0$ and $F(X) = I_A$ (where 0 is the 0-element in $\mathcal{A}$),*

(iii) *for any countable decomposition $\{\Xi_1, \Xi_2, ..., \Xi_n, ...\}$ of Ξ, $\Big($i.e., $\Xi, \Xi_n \in \mathcal{F}, \cup_{n=1}^{\infty} \Xi_n = \Xi$, $\Xi_n \cap \Xi_m = \emptyset$(if $n \neq m$)$\Big)$, it holds that $\rho\Big(F(\Xi)\Big) = \lim_{N \to \infty} \rho\Big(\sum_{n=1}^{N} F(\Xi_n)\Big)$ $(\forall \rho \in \mathfrak{S}^m(\mathcal{A}^*))$.*

Also, if $F(\Xi)$ is a projection for every Ξ ($\in \mathcal{F}$), a C^*-observable $(X, \mathcal{F}, F)$ is called a crisp C^*-observable (or, *a crisp observable, an idea*).

■

Remark 2.8. [(1): The case that X is finite]. In chapters 2∼8, we will usually deal with the case that X is finite. When we want to stress that X is finite, the $(X, \mathcal{F}, F)$ is often denoted by $(X, 2^X, F)$ or $(X, \mathcal{P}(X), F)$. Thus, in this case, the (iii) in Definition 2.7 means

$$F(\Xi_1 \cup \Xi_2) = F(\Xi_1) + F(\Xi_2) \quad (\forall \Xi_1, \forall \Xi_2 (\in 2^X) \text{ such that } \Xi_1 \cap \Xi_2 = \emptyset)).$$

[(2): C^*-observables in general C^*-algebras]. Although we are usually concerned with unital C^*-algebras, we add the generalization of Definition 2.7 as follows: *Let $\mathcal{A}$ be a C^*-algebra, which does not necessarily have the identity I. A C^*-observable (or in short, observable, fuzzy observable)* $\mathbf{O} \equiv (X, \mathcal{R}, F)$ in a C^*-algebra $\mathcal{A}$ *is defined such that it satisfies that*

(i) *X is a set, and $\mathcal{R}$ is the subring of the power set $\mathcal{P}(X)$ ($\equiv \{\Xi : \Xi \subseteq X\}$), that is, "$\Xi_1, \Xi_2 \in \mathcal{R}$" implies "$\Xi_1 \cap \Xi_2 \in \mathcal{R}$" and "$\Xi_1 \triangle \Xi_2 \in \mathcal{R}$",*

(ii) *for every* $\Xi \in \mathcal{R}$, $F(\Xi)$ *is a positive element in* $\mathcal{A}$ *such that* $F(\emptyset) = 0$ *(where* 0 *is the* 0*-element in* $\mathcal{A}$*),*

(iii) *for any countable decomposition* $\{\Xi_1, \Xi_2, ..., \Xi_n, ...\}$ *of* Ξ, $(\Xi, \Xi_n \in \mathcal{R})$, *it holds that* $\rho\Big(F(\Xi)\Big) = \lim_{N\to\infty} \rho\Big(\sum_{n=1}^N F(\Xi_n)\Big) \quad (\forall \rho \in \mathfrak{S}^m(\mathcal{A}^*))$,

(iv) *there exists a sequence* $\{\Xi_n^0\}_{n=1}^\infty$ *in* $\mathcal{R}$ *such that* $\Xi_1^0 \subseteq \Xi_2^0 \subseteq \cdots$ *and* $X = \cup_{n=1}^\infty \Xi_n^0$ *and* $\lim_{n\to\infty} \rho\Big(F(\Xi_n^0)\Big) = 1 \quad (\forall \rho \in \mathfrak{S}^m(\mathcal{A}^*))$.

Also, if $F(\Xi)$ *is a projection for every* Ξ $(\in \mathcal{R})$, *a* C^**-observable* $(X, \mathcal{R}, F)$ *is called a crisp* C^**-observable.*

■

Definition 2.9. [Image observable]. *Let* $\mathbf{O} \equiv (X, \mathcal{F}, F)$ *be an observable in a* C^**-algebra* $\mathcal{A}$. *Let* $\mathcal{G}$ *be a subfield of* 2^Y. *Let* $h : X \to Y$ *be a measurable map, i.e.,* $h^{-1}(\Gamma) \in \mathcal{F}$ $(\forall \Gamma \in \mathcal{G})$. *Then, we can define the observable* $\mathbf{O}_{[h]}$ $(\equiv (Y, \mathcal{G}, F \circ h^{-1}))$ *in* $\mathcal{A}$ *such that:*

$$(F \circ h^{-1})(\Gamma) = F(h^{-1}(\Gamma)) \qquad (\Gamma \in \mathcal{G}). \tag{2.20}$$

The $\mathbf{O}_{[h]} \equiv (X, \mathcal{F}, G \circ h^{-1})$ *is called the image observable of* $\mathbf{O} \equiv (Y, \mathcal{G}, G)$ *(in a* C^**-algebra* $\mathcal{A}$*) concerning the map* $h : X \to Y$. *The image observable* $\mathbf{O}_{[h]}$ *is also denoted by* $h(\mathbf{O})$.

■

Definition 2.10. [Quasi-product observable]. *For each* $k = 1, 2, ..., n$, *consider an observable* $\mathbf{O}_k \equiv (X_k, \mathcal{F}_k, F_k)$ *in a* C^**-algebra* $\mathcal{A}$. *Define the field* $\bigotimes_{k=1}^n \mathcal{F}_k$ $(\subseteq 2^{\times_{k=1}^n X_k})$ *such as the smallest field (on* $\times_{k=1}^n X_k$*) that contains* $\times_{k=1}^n \Xi_k$, $\Xi_k \in \mathcal{F}_k$. *The product field* $\bigotimes_{k=1}^n \mathcal{F}_k$ *is usually denoted by* $\times_{k=1}^n \mathcal{F}_k$. $\Big($*Throughout this book, the notation* $\times_{k=1}^n \mathcal{F}_k$ *does not mean the set* $\{\times_{k=1}^n \Xi_k : \Xi_k \in \mathcal{F}_k\}$.$\Big)$ *An observable* $\widehat{\mathbf{O}} \equiv (\times_{k=1}^n X_k, \times_{k=1}^n \mathcal{F}_k, \widehat{F})$ *in* $\mathcal{A}$ *is called the quasi-product observable of* $\{\mathbf{O}_k : k = 1, 2, ..., n\}$ $\Big($*or, quasi-product observable with marginal observables* $\{\mathbf{O}_k : k = 1, 2, ..., n\}\Big)$ *if it holds that*

$$\widehat{F}(X_1 \times \cdots \times X_{k-1} \times \Xi_k \times X_{k+1} \times \cdots \times X_n) = F_k(\Xi_k) \quad (\forall \Xi_k \in \mathcal{F}_k, \forall k = 1, ..., n). \tag{2.21}$$

The quasi-product observable $\widehat{\mathbf{O}}$ *(of* $\{\mathbf{O}_k\}_{k=1}^n$*) is denoted by*

$$\mathop{\mathsf{X}}_{k=1,2,...,n}^{\mathrm{qp}} \mathbf{O}_k, \ or, \ \Big(\mathop{\times}_{k=1}^{n} X_k, \mathop{\times}_{k=1}^{n} \mathcal{F}_k, \mathop{\mathsf{X}}_{k=1,2,...,n}^{\mathrm{qp}} F_k\Big), \ or \ \Big(\mathop{\times}_{k=1}^{n} X_k, \bigotimes_{k=1}^{n} \mathcal{F}_k, \mathop{\mathsf{X}}_{k=1,2,...,n}^{\mathrm{qp}} F_k\Big), \tag{2.22}$$

i.e., $\widehat{\mathbf{O}} = \overset{\text{qp}}{\times}_{k=1,2,...,n}\mathbf{O}_k$, $\widehat{F} = \overset{\text{qp}}{\times}_{k=1,2,...,n}F_k$. *Also,* $\overset{\text{qp}}{\times}_{k=1,2,...,n}F_k$ *is sometimes written by* $\overset{\widehat{\mathbf{O}}}{\times}_{k=1,2,...,n}F_k$.

■

Note that the existence and the uniqueness of the quasi-product observable of $\{\mathbf{O}_k : k=1,2,...,n\}$ are not guaranteed in general. However, when $\mathbf{O}_k$, $k=1,2,...,n$, commute, i.e.,

$$F_k(\Xi_k)F_{k'}(\Xi_{k'}) = F_{k'}(\Xi_{k'})F_k(\Xi_k) \quad \text{for all } \Xi_k \in \mathcal{F}_k,\ \Xi_{k'} \in \mathcal{F}_{k'} \text{ such that } k \neq k' \ , \tag{2.23}$$

we can construct the quasi-product observable $(\times_{k=1}^n X_k, \times_{k=1}^n \mathcal{F}_k, \widetilde{F})$ in $\mathcal{A}$ such that:

$$\widetilde{F}(\Xi_1 \times \Xi_2 \times \cdots \times \Xi_n) = F_1(\Xi_1)F_2(\Xi_2)\cdots F_n(\Xi_n). \tag{2.24}$$

This kind of quasi-product observable is called a *product observable* and denoted by

$$\mathop{\times}_{k=1}^{n} \mathbf{O}_k \quad \Big(= \mathop{\times}_{k=1,2,...,n} \mathbf{O}_k, \text{ or, } (\mathop{\times}_{k=1}^{n} X_k, \mathop{\times}_{k=1}^{n} \mathcal{F}_k, \mathop{\times}_{k=1}^{n} F_k), \text{ or, } (\mathop{\times}_{k=1}^{n} X_k, \bigotimes_{k=1}^{n} \mathcal{F}_k, \mathop{\times}_{k=1}^{n} F_k), \Big). \tag{2.25}$$

$\times_{k=1}^n$ is sometimes written by $\prod_{k=1}^n$, and thus, we write: $\times_{k=1}^n \mathbf{O}_k = \prod_{k=1}^n \mathbf{O}_k$, $\times_{k=1}^n X_k = \prod_{k=1}^n X_k$, etc. Also, note that the product observable $\times_{k=1}^n \mathbf{O}_k$ always exists for any $\mathbf{O}_k$ in a commutative C^*-algebra $C(\Omega)$.

Summing up the above arguments, we can state the following theorem.

Theorem 2.11. *For each* $k \in K \equiv \{1,2,...,|K|\}$, *consider an observable* $\mathbf{O}_k \equiv (X_k, \mathcal{F}_k, F_k)$ *in a* C^**-algebra* $\mathcal{A}$. *If the commutativity condition:*

$$F_{k_1}(\Xi_{k_1})F_{k_2}(\Xi_{k_2}) = F_{k_2}(\Xi_{k_2})F_{k_1}(\Xi_{k_1}) \quad (\forall \Xi_{k_1} \in \mathcal{F}_{k_1},\ \forall \Xi_{k_2} \in \mathcal{F}_{k_2},\ k_1 \neq k_2) \tag{2.26}$$

holds, then we can construct a product observable $\widehat{\mathbf{O}} \equiv (\times_{k\in K} X_k, \times_{k\in K} \mathcal{F}_k,\ \widetilde{F} \equiv \times_{k\in K} F_k)$ *such that:*

$$\widetilde{F}(\Xi_1 \times \Xi_2 \times \cdots \times \Xi_{|K|}) = F_1(\Xi_1)F_2(\Xi_2)\cdots F_{|K|}(\Xi_{|K|}). \tag{2.27}$$

Note that the uniqueness (of quasi-product observables) is not guaranteed even under the above commutativity condition. Also, note that the product observable $\times_{k=1}^n \mathbf{O}_k$ *always exists for any* $\mathbf{O}_k$ *in a commutative* C^**-algebra* $C(\Omega)$.

■

Theorem 2.12. *Let* $\mathbf{O} \equiv (X, \mathcal{R}, F)$ *be a* C^**-observable in a general* C^**-algebra* $\mathcal{A}$ *(i.e., it does not necessarily have the identity). Let* $\mathcal{A}_1$ *be a* C^**-algebra with the identity* I *(generated by the* $\mathcal{A}$ *such as in Remark 2.6(i)). Then, there uniquely exists the observable* $(X, \mathcal{F}, \widetilde{F})$ *be a* C^**-observable in* $\mathcal{A}_1$ *such that:*

$$\text{(i)} \qquad \mathcal{F} = \mathcal{R} \cup \{X \setminus \Gamma \mid \Gamma \in \mathcal{R}\}$$

$$\text{(ii)} \qquad \widetilde{F}(\Xi) = \begin{cases} F(\Xi) & (\Xi \in \mathcal{R}) \\ I - F(\Xi^c) & (\Xi^c = (X \setminus \Xi) \in \mathcal{R}). \end{cases}$$

Proof. It suffices to show that $\mathcal{F}$ is the field. Let $\Xi_1 \in \mathcal{R}$ and $\Xi_2 \in \{X \setminus \Gamma \mid \Gamma \in \mathcal{R}\}$. Thus $\Xi_2 = X \setminus \Gamma$ (for some $\Gamma \in \mathcal{R}$). Then, we see $\Xi_1 \cap \Xi_2 = \Xi_1 \cap (X \setminus \Gamma) = \Xi_1 \cap (\Xi_1 \setminus \Gamma)$ $\in \mathcal{F}$. Also, $\Xi_1 \cup \Xi_2 = (\Xi_1^c \cap \Xi_2^c)^c = (\Xi_1^c \cap \Gamma)^c = (\Gamma \setminus \Xi_1)^c \in \mathcal{F}$. Also, it is clear that "$\Xi \in \mathcal{F}$" $\Longrightarrow$ "$\Xi^c \in \mathcal{F}$". Thus, we see that $\mathcal{F}$ is the field. □

The following theorem (and Theorem 9.8) will be often used throughout this book.

Theorem 2.13. [*cf.* [42]]. *Let* $\mathcal{A}$ *be a* C^**-algebra. Let* $\mathbf{O}_1 \equiv (X_1, \mathcal{F}_1, F_1)$ *and* $\mathbf{O}_2 \equiv (X_2, \mathcal{F}_2, F_2)$ *be* C^**-observables in* $\mathcal{A}$ *such that at least one of them is crisp.* $\Big($*So, without loss of generality, we assume that* $\mathbf{O}_2$ *is crisp*$\Big)$. *Then, the following statements are equivalent:*

(i) *There exists a quasi-product observable* $\mathbf{O}_{12} \equiv (X_1 \times X_2, \mathcal{F}_1 \times \mathcal{F}_2, F_1 \overset{\text{qp}}{\times} F_2)$ *with marginal observables* $\mathbf{O}_1$ *and* $\mathbf{O}_2$.

(ii) $\mathbf{O}_1$ *and* $\mathbf{O}_2$ *commute, that is,* $F_1(\Xi_1)F_2(\Xi_2) = F_2(\Xi_2)F_1(\Xi_1)$ $(\forall \Xi_1 \in \mathcal{F}_1, \forall \Xi_2 \in \mathcal{F}_2)$.

Furthermore, if the above statements (i) and (ii) hold, the uniqueness of the quasi-product observable $\mathbf{O}_{12}$ *of* $\mathbf{O}_1$ *and* $\mathbf{O}_2$ *is guaranteed.*

Proof. It suffices to prove it in the case that $\mathcal{A}$ has the identity. When $\mathbf{O}_1 \equiv (X_1, \mathcal{F}_1, F_1)$ and $\mathbf{O}_2 \equiv (X_2, \mathcal{F}_2, F_2)$ are both crisp observables, it is proved in [17]. By the same way, we can prove this theorem. It is clear that (ii) $\Longrightarrow$ (i) since we can construct a C^*-observable $(X_1 \times X_2, \mathcal{F}_1 \times \mathcal{F}_2, H)$ such that:

$$H(\Xi_1 \times \Xi_2) = F_1(\Xi_1)F_2(\Xi_2) \quad (\forall \Xi_1 \in \mathcal{F}_1, \forall \Xi_2 \in \mathcal{F}_2).$$

Thus, it suffices to prove that (i) $\Longrightarrow$ (ii). Assume that (i) holds. Let Ξ_1 and Ξ_2 be any element in $\mathcal{F}_1$ and $\mathcal{F}_2$ respectively. Put $\Xi_1^1 = \Xi_1$, $\Xi_1^2 = X_1 \backslash \Xi_1$, $\Xi_2^1 = \Xi_2$ and $\Xi_2^2 = X_2 \backslash \Xi_2$. Put $H = F_1 \overset{\text{qp}}{\times} F_2$. Note that:

$$0 \le H(\Xi_1^i \times \Xi_2^j) \le H(X_1 \times \Xi_2^j) \equiv F_2(\Xi_2^j) \ \ (= \text{"projection"}). \tag{2.28}$$

This implies that $H(\Xi_1^i \times \Xi_2^j)$ and $F_2(\Xi_2^j)$ commute, and so, $H(\Xi_1^i \times \Xi_2^j)$ and $I - F_2(\Xi_2^j)$ commute. Hence, $F_1(\Xi_1)$ $(= H(\Xi_1^1 \times \Xi_2^1) + H(\Xi_1^1 \times \Xi_2^2))$ and $F_2(\Xi_2)$ $(= F_2(\Xi_2^1))$ commute. Therefore, we get that (i) $\Longrightarrow$ (ii).

Next we prove the uniqueness of H under the assumption (i) (and so (ii)). Note that $0 \le H(\Xi_1^i \times \Xi_2^j) \le H(\Xi_1^i \times X_2) \equiv F_1(\Xi_1^i)$. This implies, by the commutativity condition (ii) and (2.28), that

$$0 \le H(\Xi_1^i \times \Xi_2^j) \le F_2(\Xi_2^j) F_1(\Xi_1^i) F_2(\Xi_2^j) = F_1(\Xi_1^i) F_2(\Xi_2^j). \tag{2.29}$$

Therefore we see that $I = \sum_{i,j=1,2} H(\Xi_1^i \times \Xi_2^j) \le \sum_{i,j=1,2} F_1(\Xi_1^i) F_2(\Xi_2^j) = I$. Then, we obtain that $H(\Xi_1 \times \Xi_2) = F_1(\Xi_1) F_2(\Xi_2)$, that is, H is unique. Therefore, we finish the proof. □

2.3 The meanings of observables and crisp observables

In the conventional classical [resp. quantum] mechanics, the term "observable" usually means a *real valued continuous function on a state space* Ω [resp. *a self-adjoint operator in* $B(V)$]. Thus, the "observable" (defined in Definition 2.7) should be a kind of generalization of the above conventional "observable". In what follows we will see it.

Now we shall consider the several aspects (and properties) of the observable $\mathbf{O} \equiv (X, \mathcal{F}, F)$ in a C^*-algebra $\mathcal{A}$. Examining Definition 2.7, we can easily see

(A_1) An observable $\mathbf{O}$ $\big(\equiv (X, \mathcal{F}, F) \big)$ in $\mathcal{A}$ can be regarded as the $\mathcal{A}$-valued probability space[5], i.e., the additive set-function:

$$\mathcal{F} \ni \Xi \mapsto F(\Xi) \in \mathcal{A}.$$

Also, we may find the similarity between an observable $\mathbf{O}$ and *the resolution of the identity* I in what follows. Assume, for simplicity, that X is countable (i.e., $X \equiv \{x_1, x_2, ...\}$). Then, it is clear that

[5]In this book, the term "probability space" is used as "a positive measure space whose total measure is equal to 1". That is, the term "probability space" is used as the pure mathematical concept, and thus, it is not always assured to be related to the concept of "probability".

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

(ii) $\sum_{k=1}^{\infty} F(\{x_k\}) = I_{\mathcal{A}}$ in the sense of weak topology of $\mathcal{A}$,

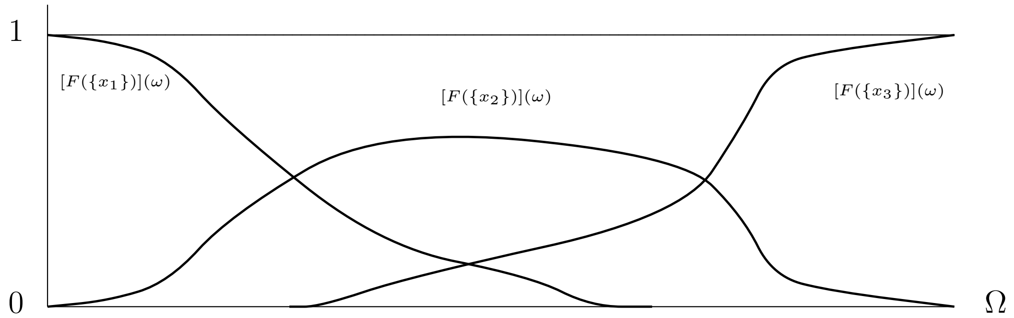
which imply that the $[F(\{x_k\}) \ : \ k = 1, 2, ..., n]$ can be regarded as *the resolution of the identity element* $I_{\mathcal{A}}$. Thus we say that

(A_2) An observable $\mathbf{O}$ $\big(\equiv (X, \mathcal{F}, F) \big)$ in $\mathcal{A}$ can be regarded as

$$\text{“\textit{the fuzzy decomposition}”} \tag{2.30}$$

that is, *the resolution of the identity* $I_{\mathcal{A}}$, i.e., $[F(\{x_k\}) \ : \ k = 1, 2, ..., n]$.

“The figure of $\mathbf{O} \equiv (\{x_1, x_2, x_3\},\ 2^{\{x_1,x_2,x_3\}},\ F)$ in $C(\Omega)$”



Also, we note that

(A_3) An observable $\mathbf{O}$ $\big(\equiv (X, \mathcal{F}, F) \big)$ in $\mathcal{A}$ can be characterized as a kind of generalization of a self-adjoint element in $\mathcal{A}$.

This is shown as follows: For simplicity, assume that $\mathcal{A} = B(\mathbf{C}^N)$. And put

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \cdots, \quad e_N = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \tag{2.31}$$

Thus we see that

$$|e_1\rangle\langle e_1| = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}, \quad |e_2\rangle\langle e_2| = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}, \cdots, |e_N\rangle\langle e_N| = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}.$$

The spectral theorem says that a self-adjoint matrix $\widehat{F}$ ($\in B(\mathbf{C}^N)$) can be represented by

$$\widehat{F} = U \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_N \end{bmatrix} U^* \\ = U\Big(\lambda_1|e_1\rangle\langle e_1| + \lambda_2|e_2\rangle\langle e_2| + \cdots + \lambda_N|e_N\rangle\langle e_N|\Big)U^* \\ = \sum_{n=1}^{N} \lambda_n |Ue_n\rangle\langle Ue_n| \tag{2.32}$$

where $\lambda_n \in \mathbf{R}$ ($\forall n = 1,2,...,N$) and U is a unitary matrix in $B(\mathbf{C}^N)$. For any Ξ ($\in \mathcal{B}_{\mathbf{R}} =$ "Borel field")[6], put

$$F(\Xi) = \sum_{\lambda_n \in \Xi} |Ue_n\rangle\langle Ue_n|. \tag{2.33}$$

Here it should be noted that $F(\Xi)$ is a projection for all Ξ ($\in \mathcal{B}_{\mathbf{R}}$). This implies the the following identification:

$$\underset{\text{(self-adjoint operator)}}{\widehat{F}} \longleftrightarrow \underset{\text{(crisp observable)}}{(\mathbf{R}, \mathcal{B}_{\mathbf{R}}, F) \quad \text{in } B(\mathbf{C}^N)}. \tag{2.34}$$

That is because $\widehat{F}$ is represented by (2.32), i.e.,

$$\widehat{F} = \int_{\mathbf{R}} \lambda F(d\lambda).$$

Next assume that $\mathcal{A} = C(\Omega)$, where Ω is, for simplicity, assumed to be the finite set $\{\omega_1, \omega_2, \omega_3, ..., \omega_N\}$ with the discrete topology. Consider a real valued continuous function $\widehat{F} : \Omega \to \mathbf{R}$. Define the observable $(\mathbf{R}, \mathcal{B}_{\mathbf{R}}, F)$ in $C(\Omega)$ such that:

$$[F(\Xi)](\omega) = \begin{cases} 1 & \text{if } \omega \in \widehat{F}^{-1}(\Xi) \\ \\ 0 & \text{if } \omega \notin \widehat{F}^{-1}(\Xi) \end{cases} \qquad (\forall \omega \in \Omega, \quad \forall \Xi \in \mathcal{B}_{\mathbf{R}}). \tag{2.35}$$

Note that

$$\widehat{F}(\omega) = \sum_{n=1}^{N} \widehat{F}(\omega_n)\Big([F(\{\omega_n\})](\omega)\Big) = \sum_{\lambda \in \mathbf{R}} \lambda [F(\{\lambda\})](\omega) \ \Big(= [\int_{\mathbf{R}} \lambda F(d\lambda)](\omega)\Big) \qquad (\forall \omega \in \Omega).$$

This implies the the following identification:

$$\underset{\text{(real valued function on } \Omega)}{\widehat{F}} \longleftrightarrow \underset{\text{(crisp observable)}}{(\mathbf{R}, \mathcal{B}_{\mathbf{R}}, F) \quad \text{in } C(\Omega)}. \tag{2.36}$$

Therefore, we say, by (2.34) and (2.36), that

[6] "Borel field" = "the smallest σ-field that contains all open sets"

(A_4) "crisp observable $(\mathbf{R}, \mathcal{B}_{\mathbf{R}}, F)$" in $\mathcal{A}$ $\underset{\text{identification}}{\longleftrightarrow}$ "self-adjoint element" in $\mathcal{A}$.

$\Big($where $\mathcal{A} = B(\mathbf{C}^n)$ or $\mathcal{A} = C(\{\omega_1, \omega_2, ..., \omega_N\})\Big)$. Here, the "self-adjoint element" in $\mathcal{A}$ (i.e., "crisp observable $(\mathbf{R}, \mathcal{B}_{\mathbf{R}}, F)$" in $\mathcal{A}\Big)$ is sometimes called a *"quantity* (or, *system theoretical quantity")* *in* $\mathcal{A}$.

Remark 2.14. [OR (= operation research) and game theory]. In OR [resp. game theory [85]], we are mainly concerned with the problem: "Study the maximal point [resp. the saddle point] of $\widehat{F}$!"

■

2.4 Measurement (Axiom 1)

Under the mathematical preparations in the previous sections, now we can describe the fundamental concepts of measurement theory (2.2) (=(1.4)).

With any *system* S, a C^*-algebra $\mathcal{A}$ can be associated in which measurement theory of that system can be formulated. A *state* of the system S is represented by a pure state ρ^p ($\in \mathfrak{S}^p(\mathcal{A}^*)$, i.e., a *state space*). Also, an *observable* is represented by a C^*-observable $\mathbf{O}$ $\equiv (X, \mathcal{F}, F)$ in the C^*-algebra $\mathcal{A}$.[7] The *measurement of an observable* $\mathbf{O}$ *for the system* S *with (or, in) the state* ρ^p *is represented by* $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[\rho^p]})$ *in the* C^**-algebra* $\mathcal{A}$*. Also, we can obtain a measured value* x ($\in X$) *by the measurement* $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[\rho^p]})$.

The axiom presented below is analogous to (or, a kind of generalizations of) Born's probabilistic interpretation of quantum mechanics [13]. We of course assert that the axiom is a principle for all measurements, i.e., classical and quantum measurements. Cf. [41, 42].

AXIOM 1. [Measurement axiom]. *Consider a measurement* $\mathbf{M}_{\mathcal{A}}(\mathbf{O} \equiv (X, \mathcal{F}, F), S_{[\rho^p]})$ *formulated in a* C^**-algebra* $\mathcal{A}$*. Assume that the measured value* x ($\in X$) *is obtained by the measurement* $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[\rho^p]})$*. Then, the probability that the* x ($\in X$) *belongs to a set* Ξ ($\in \mathcal{F}$) *is given by* $\rho^p(F(\Xi))$ $\Big(\equiv {}_{\mathcal{A}^*}\Big\langle \rho^p, F(\Xi) \Big\rangle_{\mathcal{A}} \Big)$. (2.37)

---

[7] I like to image the following correspondence (measurement theory and philosophy):

"state" ↔ "matter" "observable" ↔ "idea" (= "form")

We introduce the following classification in measurement theory:

$$
\text{measurement theory} \begin{cases} \text{classical measurement theory (for classical systems)} \\ \\ \text{quantum measurement theory (for quantum systems)} \end{cases} \tag{2.38}
$$

where a C^*-algebra $\mathcal{A}$ is commutative or non-commutative.

Recall the (1.3), that is, quantum mechanics (*cf.* [71]) is formulated by

$$
\boxed{\text{"quantum mechanics"}} = \underset{\text{("Born's quantum measurements")}}{\text{measurement}} + \underset{\text{("Schrödinger equation")}}{\text{the rule of time evolution}} \tag{1.3}
$$

Of course, Axiom 1 corresponds to "Born's quantum measurements". Note that quantum measurement theory is well authorized as a principle of quantum mechanics (*cf.* [17, 34, 84]). Our interest in this book is mainly concentrated on classical systems. Therefore, in most cases, it suffices to assume that $\mathcal{A} = C(\Omega)$.

2.5 Remarks

In this section we add some remarks concerning Axiom 1.

[**(I): Probability**]. It should be noted that the term "probability" appears in Axiom 1. Following the common knowledge of quantum mechanics (*cf.* [71, 84]), we believe that any scientific statement including the term "probability" is meaningless without the concept of "measurement". That is, we say that

($\sharp$) *"There is no probability without measurements".*

Throughout this book, the above spirit ($\sharp$) is quite important.

[**(II): It is prohibited to take measurements twice**]. The quasi-product observable (or, the product observable) is used to represent *"the measurement of* (*more than one*) *observables"* as follows: For example, consider *"the measurement of* $\mathbf{O}_1$ *and* $\mathbf{O}_2$ *for the system with the state* ρ^p $(\in \mathfrak{S}^p(\mathcal{A}^*))$*."* If the quasi-product observable $\mathbf{O}_1 \overset{\text{qp}}{\times} \mathbf{O}_2$ of $\mathbf{O}_1$ and $\mathbf{O}_2$ exists, the measurement is represented by $\mathbf{M}_{\mathcal{A}}(\mathbf{O}_1 \overset{\text{qp}}{\times} \mathbf{O}_2, S_{[\rho^p]})$ $\Big($and not "$\mathbf{M}_{\mathcal{A}}(\mathbf{O}_1, S_{[\rho^p]})$ $+ \mathbf{M}_{\mathcal{A}}(\mathbf{O}_2, S_{[\rho^p]})$"$\Big)$. If the quasi-product observable $\mathbf{O}_1 \overset{\text{qp}}{\times} \mathbf{O}_2$ does not exist, the measurement does not also exist. That is, the symbol "$\mathbf{M}_{\mathcal{A}}(\mathbf{O}_1, S_{[\rho^p]}) + \mathbf{M}_{\mathcal{A}}(\mathbf{O}_2, S_{[\rho^p]})$" is nonsense. Thus we can say that

(♯) *only one measurement is permitted to be conducted even in the classical measurement theory.*

which is the well-known fact in quantum mechanics. The measurement $\mathbf{M}_{\mathcal{A}}\big(\mathbf{O}_1 \overset{\text{qp}}{\times} \mathbf{O}_2, S_{[\rho^p]}\big)$ is sometimes called a *simultaneous measurement (or iterated measurement)* of two observables $\mathbf{O}_1$ and $\mathbf{O}_2$. That is, it is prohibited to take measurements twice in measurement theory. For example, the following statement:

- "Take two measurements $\mathbf{M}_{\mathcal{A}}\big(\mathbf{O}_1, S_{[\rho^p]}\big)$ and $\mathbf{M}_{\mathcal{A}}\big(\mathbf{O}_2, S_{[\rho^p]}\big)$."

is prohibited.

[**(III): Sample space**]. Let ρ^m be a mixed state, i.e., $\rho^m \in \mathfrak{S}^m(\mathcal{A}^*)$. Applying Hopf extension theorem (*cf.* [92]), we can get the measure space $(X, \overline{\mathcal{F}}, \overline{\rho^m(F(\,\cdot\,))}\ )$ such that $\overline{\rho^m(F(\Xi))} = \rho^m(F(\Xi))$ for all $\Xi \in \mathcal{F}$ where $\overline{\mathcal{F}}$ is the smallest σ-field that contains $\mathcal{F}$. For simplicity, the $\overline{\rho^m(F(\,\cdot\,))}$ is also denoted by $\rho^m(F(\,\cdot\,))$ or ${}_{\mathcal{A}^*}\Big\langle \rho^m, F(\,\cdot\,)\Big\rangle_{\mathcal{A}}$. Axiom 1 makes us call the measure space $(X, \overline{\mathcal{F}},\ \overline{\rho^p(F(\,\cdot\,))}\ )$ $\Big($or in short, $(X, \mathcal{F},\ \rho^p(F(\,\cdot\,))\ )\Big)$ a *sample space* concerning a measurement $\mathbf{M}_{\mathcal{A}}(\mathbf{O} \equiv (X, \mathcal{F}, F), S_{[\rho^p]}\ )$.

[**(IV): Conditional probability**]. Let $\mathbf{O} \equiv (X, \mathcal{F}, F)$ and $\mathbf{O}' \equiv (Y, \mathcal{G}, G)$ be observables in $\mathcal{A}$. Let $\widehat{\mathbf{O}}$ be a quasi-product observable of $\mathbf{O}$ and $\mathbf{O}'$, that is, $\widehat{\mathbf{O}} \equiv \mathbf{O} \overset{\text{qp}}{\times} \mathbf{O}' = (X \times Y, \mathcal{F} \times \mathcal{G}, F \overset{\text{qp}}{\times} G)$. Assume that we know that the measured value (x, y) $(\in X \times Y)$ obtained by a measurement $\mathbf{M}_{\mathcal{A}}(\widehat{\mathbf{O}}, S_{[\rho^p]})$ belongs to $\Xi \times Y$ $(\in \mathcal{F} \times \mathcal{G})$. Then, it is clear that the unknown measured value y $(\in Y)$ is distributed under the conditional probability $P_\Xi(\cdot)$, where

$$P_\Xi(\Gamma) = \frac{{}_{\mathcal{A}^*}\langle \rho^p, F(\Xi) \overset{\text{qp}}{\times} G(\Gamma)\rangle_{\mathcal{A}}}{{}_{\mathcal{A}^*}\langle \rho^p, F(\Xi)\rangle_{\mathcal{A}}} \left(= \frac{\rho^p(F(\Xi) \overset{\text{qp}}{\times} G(\Gamma))}{\rho^p(F(\Xi))} \right) \quad (\forall \Gamma \in \mathcal{G}).$$

[**(V): Commutativity and simultaneous measurability**]. Let ρ^p be a pure state, i.e., $\rho^p \in \mathfrak{S}^p(\mathcal{A}^*)$. Let $\mathbf{O} \equiv (X, \mathcal{F}, F)$ and $\mathbf{O}' \equiv (Y, \mathcal{G}, G)$ be crisp observables in $\mathcal{A}$. Now we have the following problem:

- What is the simultaneous measurability condition of $\mathbf{O}$ and $\mathbf{O}'$ for the fixed ρ^p?

This is answered in [39] as follows:

- ρ^p-commutativity, i.e., $F(\Xi)G(\Gamma)\rho^p = G(\Gamma)F(\Xi)\rho^p \qquad$ for all $\Xi \in \mathcal{F}, \Gamma \in \mathcal{G}$.

However, in this book we are not concerned with such arguments.

[(VI): Schrödinger's cat paradox]. Note that Schrödinger's cat does not appear in the world of MT. Let us explain it as follows: In 1935 (*cf.* [77]) Schrödinger published an essay describing the conceptual problems in quantum mechanics. A brief paragraph in this essay described the cat paradox.

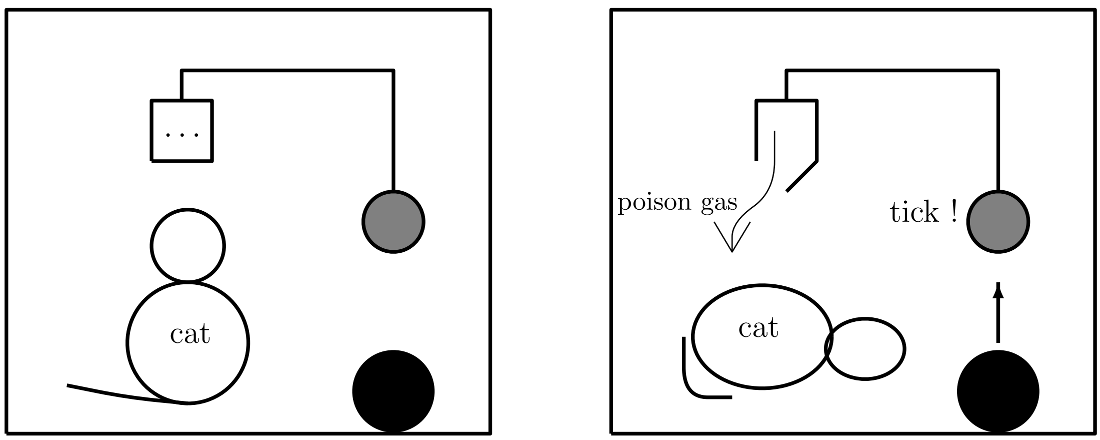
- Suppose we put a cat in a cage with a radioactive atom, a Geiger counter, and a poison gas bottle; further suppose that the atom in the cage has a half-life of one hour, a fifty-fifty chance of decaying within the hour. If the atom decays, the Geiger counter will tick; the triggering of the counter 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. Now the question:

  (Q) We then ask: What is the state of the cat after the hour?

  The answer according to quantum mechanics is that

  (A) the cat is in a state which can be thought of as half-alive and half-dead, that is, the state such as $\frac{\text{“Fig.(a)”} + \text{“Fig.(b)”}}{2}$

Fig.(a) Fig.(b)



Of course, this answer (A) is curious. This is the so-called Schrödinger's cat paradox. This paradox is due to the fact that micro mechanics and macro mechanics are mixed in the above situation. On the other hand, as seen in (2.38), micro mechanics (= quantum measurement theory) and macro mechanics (= classical measurement theory) are always separated in MT. Therefore, Schrödinger's cat does not appear in the world of MT, though this may be a surface solution of Schrödinger's cat paradox.

2.6 Examples

Again recall the (1.4), i.e.,

"measurement theory (or in short, PMT)"

$$=\underset{\text{“Axiom 1 (2.37)”}}{[\text{measurements}]} + \underset{\text{[Axiom 2 (3.26)]}}{[\text{the relation among systems}]} \quad \text{in } C^*\text{-algebra } \mathcal{A} \tag{2.39a (=(1.4a))}$$

or more precisely,

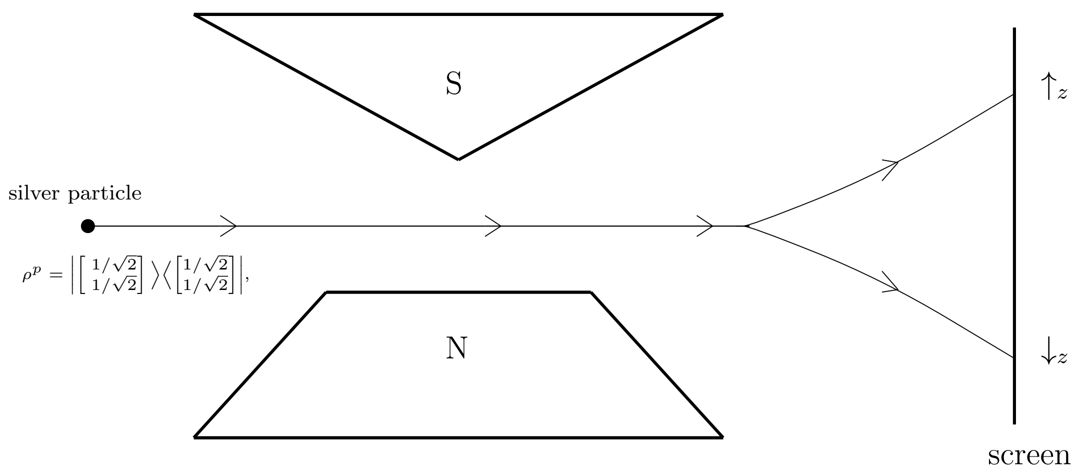
$$= \text{“Apply (2.39a) to every phenomenon by an analogy of quantum mechanics”} \tag{2.39b (=(1.4b))}$$

Thus, in order to understand PMT, we need a little knowledge of quantum mechanics.

The following example is enough tested ;[8] and thus, it is the most firm in PMT

Example 2.15. [(i): The spin observable concerning the z-axis, Stern and Gerlach's experiment]. Assume that we examine the beam (of silver particles) after passing through the magnetic field. Then, as seen in the following figure, we see that all particles are deflected either equally upwards or equally downwards in a 50:50 ratio.

"Stern and Gerlach's experiment (1922)"



Consider the two dimensional Hilbert space $V = \mathbf{C}^2$, And therefore, we get the non-commutative C^*-algebra $\mathcal{A} = B(V)$, that is, the algebra composed of all 2×2 matrices.

[8] A lot of tests of quantum mechanics have been conducted. Especially Aspect's experiment [8] is well authorized. (Cf. §2.9 Bell's inequality) Recall that "quantum system theory" $\subset$ "PMT". Thus, quantum mechanics must be enough tested though the experimental test of PMT is generally meaningless. (Cf. Remark 1.1(e).)

Note that $\mathcal{A} = B(V) = \mathcal{C}(V) = \mathcal{C}_I(V)$ (*cf.* Example 2.3 and Remark 2.6 (i)) since the dimension of V is finite. Define $\mathbf{O}^z \equiv (Z, 2^Z, F^z)$, the spin observable concerning the z-axis, such that, $Z = \{\uparrow_z, \downarrow_z\}$ and

$$F^z(\{\uparrow_z\}) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad F^z(\{\downarrow_z\}) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \tag{2.40}$$

$$F^z(\emptyset) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad F^z(\{\uparrow_z, \downarrow_z\}) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

For example, consider the measurement $\mathbf{M}_{B(\mathbf{C}^2)}\big(\mathbf{O}^z \equiv (Z = \{\uparrow_z, \downarrow_z\}, 2^Z, F^z), S_{[\rho^p]}\big)$, where $\rho^p = \Big|\begin{bmatrix} \alpha \\ \beta \end{bmatrix}\Big\rangle\Big\langle\begin{bmatrix} \alpha \\ \beta \end{bmatrix}\Big|$, $|\alpha|^2 + |\beta|^2 = 1$. That is, consider

$$\text{the measurement } \mathbf{M}_{B(\mathbf{C}^2)}\big(\mathbf{O}^z \equiv (Z = \{\uparrow_z, \downarrow_z\}, 2^Z, F^z), S_{[\rho^p]}\big)$$
$$\Big(= \text{“the measurement of the observable } \mathbf{O}^z \text{ for a particle with the state } \rho^p\text{”}\Big).$$

Then, the probability that the measured value "$\uparrow_z$" [resp. "$\downarrow_z$"] is obtained by the measurement $\mathbf{M}_{B(\mathbf{C}^2)}\big(\mathbf{O}^z, S_{[\rho^p]}\big)$ is given by $\rho^p(F^z(\{\uparrow_z\})) = |\alpha|^2$ [resp. $\rho^p(F^z(\{\downarrow_z\})) = |\beta|^2$]. Thus, if $\rho^p = \Big|\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}\Big\rangle\Big\langle\begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}\Big|$, we see that $\rho^p(F^z(\{\uparrow_z\})) = 1/2$ [resp. $\rho^p(F^z(\{\downarrow_z\}))$ $= 1/2$]. For the further argument, see §2.9 (Bell's thought experiment).

[(ii): The other spin observables]. Also, we can define $\mathbf{O}^x \equiv (X, 2^X, F^x)$, the spin observable concerning the x-axis, such that, $X = \{\uparrow_x, \downarrow_x\}$ and

$$F^x(\{\uparrow_x\}) = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}, \quad F^x(\{\downarrow_x\}) = \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix}. \tag{2.41}$$

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

$$F^y(\{\uparrow_y\}) = \begin{bmatrix} 1/2 & i/2 \\ -i/2 & 1/2 \end{bmatrix}, \quad F^y(\{\downarrow_y\}) = \begin{bmatrix} 1/2 & -i/2 \\ i/2 & 1/2 \end{bmatrix}, \tag{2.42}$$

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

■

The following example (= "urn problem") is the most important in the classical PMT, though it is somewhat artificial. That is, we believe that it is not too much to say that

- the probability in Axiom 1 for classical systems is essentially the same as the probability in the following urn problem. (2.43)

However, it should be noted that no serious test for the urn problem has been conducted.[9] It is generally considered to be self-evident without serious experiments. Recall that theoretical informatics does not require serious experiments (*cf.* §1.4).

Example 2.16. [The urn problem (i)]. There are three urns U_1, U_2 and U_3. The urn U_1 [resp. U_2, U_3] contains 8 white and 2 black balls [resp. 4 white and 6 black balls, 1 white and 9 black balls]. That is,

	white balls	black balls
urn U_1	8	2
urn U_2	4	6
urn U_3	1	9

(2.44)

Here, consider the following measurement M_2^c:

M_2^c := "Pick out one ball from the urn U_2, and recognize the color of the ball"

In measurement theory, the "measurement" M_2^c is formulated as follows: Define the state space Ω by $\Omega = \{\omega_1, \omega_2, \omega_3\}$. Here,

$$\omega_1 = [8:2], \quad \omega_2 = [4:6], \quad \omega_3 = [1:9].$$

Thus, we see that

U_1 $\cdots$ "the urn with the state ω_1"

U_2 $\cdots$ "the urn with the state ω_2"

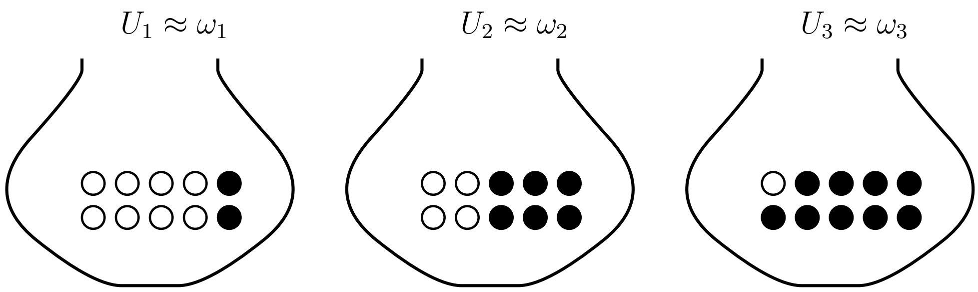
U_3 $\cdots$ "the urn with the state ω_3"

In this sense, we have the identification;

$$U_1 \approx \omega_1, \quad U_2 \approx \omega_2, \quad U_3 \approx \omega_3.$$

That is,

[9] [Fuzzy statement and precise statement]. Such a test (i.e., the experimental test of an urn problem) is usually considered to be no more than the good theme of a child's homework. However, the question *"Why is a serious test (concerning the urn problem) not required?"* may be profound. The reason can be understood if we think that the urn problem is a model within theoretical informatics. Cf. §1.4. That is, any model, represented by a precise statement, must be tested in theoretical physics. On the other hand, a model in theoretical informatics is not required to be tested, that is, it suffices to be useful. Cf. (I_{14}) in §1.3. We can say that the urn problem is as true as the statement "A cat is stronger than a mouse". It should be noted that the statement "A cat is stronger than a mouse" is "almost experimentally true" (*cf.* (I_9)) in §1.2, though it is ambiguous, fuzzy, vague, etc.



And further, define the observable $\mathbf{O} = (\{w, b\}, 2^{\{w,b\}}, F)$ in $C(\Omega)$ such that

$$\begin{aligned} &F(\{w\})(\omega_1) = 0.8, \qquad F(\{b\})(\omega_1) = 0.2, \\ &F(\{w\})(\omega_2) = 0.4, \qquad F(\{b\})(\omega_2) = 0.6, \\ &F(\{w\})(\omega_3) = 0.1, \qquad F(\{b\})(\omega_3) = 0.9, \end{aligned} \tag{2.45}$$

where 'w' and 'b' mean white and black respectively. Then, we see that

$$M_2^c = \mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[\delta_{\omega_2}]}). \tag{2.46}$$

Of course, the probability that a measured value w [resp. b] is obtained is, by Axiom 1, given by

$$F(\{w\})(\omega_2) = 0.4 \qquad [\text{ resp. } F(\{b\})(\omega_2) = 0.6] \tag{2.47}$$

[The urn problem (ii)] Further, assume that the (white or black) balls in the urns U_1, U_2 and U_3 are also made of "stone" or "metal". For example, assume that the urn U_1 [resp. U_2, U_2] contains 4 stone and 6 metal balls [resp. 5 stone and 5 metal balls, 1 stone and 9 metal balls]. That is,

	stone balls	metal balls
urn U_1	4	6
urn U_2	5	5
urn U_3	7	3

(2.48)

Here, consider the following measurement M_2^m:

$M_2^m :=$ "Pick out one ball from the urn U_2, and recognize the materials of the ball"

The measurement M_2^m is formulated as follows: Define the state space Ω by $\Omega = \{\omega_1, \omega_2, \omega_3\}$. Here,

$$\omega_1 = [4 : 6], \quad \omega_2 = [5 : 5], \quad \omega_3 = [7 : 3].$$

Thus, we see that

$$U_1 \quad \cdots \quad \text{“the urn with the state } \omega_1\text{”}$$
$$U_2 \quad \cdots \quad \text{“the urn with the state } \omega_2\text{”}$$
$$U_3 \quad \cdots \quad \text{“the urn with the state } \omega_3\text{”}$$

In this sense, we have the identification;

$$U_1 \approx \omega_1, \quad U_2 \approx \omega_2, \quad U_3 \approx \omega_3.$$

And further, define the observable $\mathbf{O}' = (\{s, m\}, 2^{\{s,m\}}, G)$ in $C(\Omega)$ such that

$$\begin{aligned} G(\{s\})(\omega_1) &= 0.4, & G(\{m\})(\omega_1) &= 0.6, \\ G(\{s\})(\omega_2) &= 0.5, & G(\{m\})(\omega_2) &= 0.5, \\ G(\{s\})(\omega_3) &= 0.7, & G(\{m\})(\omega_3) &= 0.3. \end{aligned} \tag{2.49}$$

Thus, we see:

$$M_2 = \mathbf{M}_{C(\Omega)}(\mathbf{O}', S_{[\delta_{\omega_2}]}). \tag{2.50}$$

For example, the probability that a measured value s [resp. m] is obtained is, by Axiom 1, given by

$$G(\{s\})(\omega_2) = 0.5 \qquad [\text{ resp. } G(\{m\})(\omega_2) = 0.5]. \tag{2.51}$$

[The urn problem (iii)] However, it should noted that some information is not represented in the tables (2.44) and (2.48). That is, the situation is, for example, stated precisely as follows:

(1) the urn U_1 contains 10 balls such as

	stone balls	metal balls
white balls	4	4
black balls	0	2

(2.52)

(2) the urn U_2 contains 10 balls such as

	stone balls	metal balls
white balls	4	0
black balls	1	5

(2.53)

(3) the urn U_3 contains 10 balls such as

	stone balls	metal balls
white balls	1	0
black balls	6	3

(2.54)

Here, consider the following measurement M_2^{cm}:

M_2^{cm} := "Pick out one ball from the urn U_2, and recognize the color and materials of the ball".

The measurement M_{12} is formulated as follows: Put $\Omega = \{\omega_1, \omega_2, \omega_2\}$. Define the state space Ω by $\Omega = \{\omega_1, \omega_2, \omega_3\}$. Here,

$$\omega_1 = \begin{bmatrix} 4 & 4 \\ 0 & 2 \end{bmatrix}, \quad \omega_2 = \begin{bmatrix} 4 & 0 \\ 1 & 5 \end{bmatrix}, \quad \omega_3 = \begin{bmatrix} 1 & 0 \\ 6 & 3 \end{bmatrix}.$$

Thus, we see that

U_1 $\cdots$ "the urn with the state ω_1"

U_2 $\cdots$ "the urn with the state ω_2"

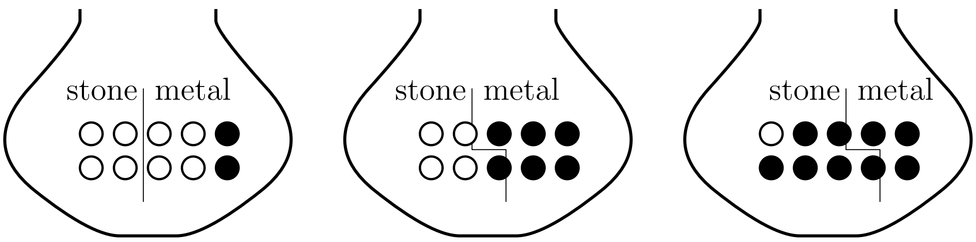
U_3 $\cdots$ "the urn with the state ω_3"

In this sense, we have the identification;

$$U_1 \approx \omega_1, \quad U_2 \approx \omega_2, \quad U_3 \approx \omega_3.$$

That is,

$U_1 \approx \omega_1$ $\qquad$ $U_2 \approx \omega_1$ $\qquad$ $U_3 \approx \omega_1$



And further, define the observable $\widehat{\mathbf{O}} = (\{w, b\} \times \{s, m\}, 2^{\{w,b\} \times \{s,m\}}, H(\equiv F \overset{\text{qp}}{\times} G))$ in $C(\Omega)$ such that

$$H(\{(w,s)\})(\omega_1) = 0.4,\ H(\{(w,m)\})(\omega_1) = 0.4,\ H(\{(b,s)\})(\omega_1) = 0.0,\ H(\{(b,m)\})(\omega_1) = 0.2,$$
$$H(\{(w,s)\})(\omega_2) = 0.4,\ H(\{(w,m)\})(\omega_2) = 0.0,\ H(\{(b,s)\})(\omega_2) = 0.1,\ H(\{(b,m)\})(\omega_2) = 0.5,$$
$$H(\{(w,s)\})(\omega_3) = 0.1,\ H(\{(w,m)\})(\omega_3) = 0.0,\ H(\{(b,s)\})(\omega_3) = 0.6,\ H(\{(b,m)\})(\omega_3) = 0.3,$$
(2.55)

which is, of course, constructed by (2.52) + (2.53) + (2.54) . Then, we see that

$$M_{12} = \mathbf{M}_{C(\Omega)}(\widehat{\mathbf{O}}, S_{[\delta_{\omega_2}]}). \tag{2.56}$$

Of course, the probability that a measured value (w, s) [resp. (w, m), (b, s), (b, m)] is obtained is, by Axiom 1, given by

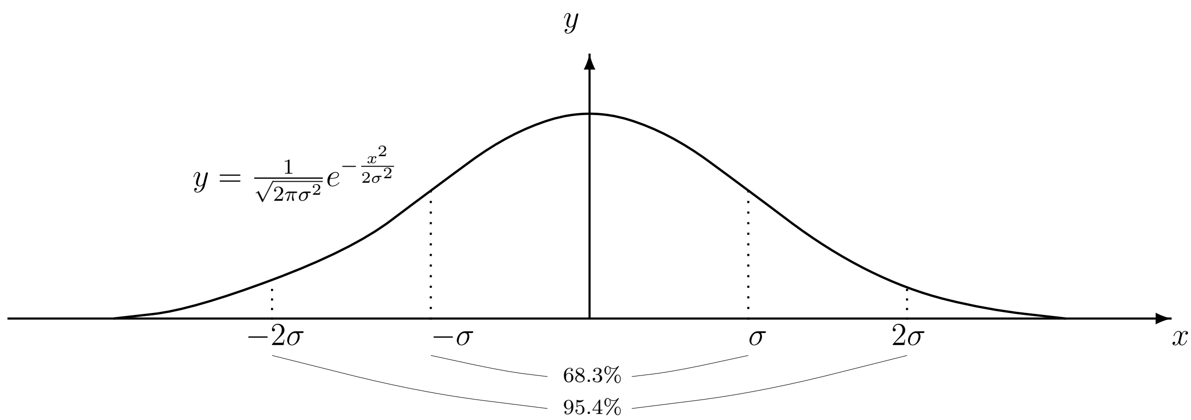
$$\begin{aligned} & F(\{(w,s)\})(\omega_2) = 0.4 \\ & [\text{ resp. } F(\{(w,m)\})(\omega_2) = 0.0, F(\{(b,s)\})(\omega_2) = 0.1, F(\{(b,m)\})(\omega_2) = 0.5]. \end{aligned} \tag{2.57}$$

■

Example 2.17. [Gaussian observable[10]]. [(i): Gaussian observable in $C(\Omega)$]. Put $\Omega = [a, b]$ ($\subseteq \mathbf{R}$, the real line), i.e., the closed interval And let σ be a fixed positive real. Define the *normal observable* (or *Gaussian observable*) $\mathbf{O}_{G^\sigma} \equiv (\mathbf{R}, \mathcal{B}_{\mathbf{R}}, G^\sigma)$ in $C(\Omega)$ such that:

$$[G^\sigma(\Xi)](\omega) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_\Xi e^{-\frac{(x-\omega)^2}{2\sigma^2}} dx \quad (\forall \Xi \in \mathcal{B}_{\mathbf{R}}, \forall \omega \in \Omega \equiv [a, b]),$$

which will be often used in this book.



Here, $\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\sigma}^{\sigma} e^{-\frac{x^2}{2\sigma^2}} dx = 0.683...$ and $\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-2\sigma}^{2\sigma} e^{-\frac{x^2}{2\sigma^2}} dx = 0.954...$ Also, note that

$$\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-1.96\sigma}^{1.96\sigma} e^{-\frac{x^2}{2\sigma^2}} dx \approx 0.95, \quad \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{1.65\sigma} e^{-\frac{x^2}{2\sigma^2}} dx \approx 0.95 \tag{2.58}$$

[10]Why is the Gaussian observable fundamental? We should not be too serious with the question. That is because we do not necessarily need a complete reason in theoretical informatics (*cf.* Chapter 1), though the differential geometrical reason must be indispensable for theoretical physics. In informatics, what is important is "useful or not". And we know that the Gaussian observable is quite useful. Also recall that every equation (e.g., Boltzmann's kinetic equation, Navier-Stokes equation, etc.) in theoretical informatics is somewhat empirical. As mentioned in (I_9) in §1.2, we think that "useful" $\Longrightarrow$ "almost experimentally true".

[(ii).Gaussian observable in $C_0(\mathbf{R}^d)$]. Consider a commutative C^*-algebra $C_0(\mathbf{R}^d)$ and the Borel ring $(\mathbf{R}^d, \mathcal{B}^{\rm bd}_{\mathbf{R}^d})$, where $\mathcal{B}^{\rm bd}_{\mathbf{R}^d} = \{\Xi \in \mathcal{B}_{\mathbf{R}^d} : \Xi$ is a bounded Borel set in $\mathbf{R}^d$ $\}$. And define the *d-dimensional Gaussian observable* $\mathbf{O}_\Sigma \equiv (\mathbf{R}^d, \mathcal{B}^{\rm bd}_{\mathbf{R}^d}, F^\Sigma)$ in $C_0(\mathbf{R}^d)$ such that:

$$
[F^\Sigma(\Xi)](\vec{\omega}) = \frac{1}{\sqrt{2\pi}^d |\Sigma|^{1/2}} \int_\Xi \exp[-\frac{1}{2}(\vec{x}-\vec{\omega})^t \Sigma^{-1}(\vec{x}-\vec{\omega})] d\vec{x} \qquad (\forall \Xi \in \mathcal{B}^{\rm bd}_{\mathbf{R}^d}, \quad \forall \vec{\omega} \in \mathbf{R}^d), \tag{2.59}
$$

where the Σ is a covariance $(d \times d)$-matrix, i.e., a positive definite $(d \times d)$-matrix. Of course, the probability that a measured value obtained by the measurement $\mathbf{M}_{C_0(\mathbf{R}^d)}(\mathbf{O}_\Sigma, S_{[\delta_{\vec{\omega}_0}]})$ belongs to Ξ $(\in \mathcal{B}^{\rm bd}_{\mathbf{R}^d})$ is given by $[F^\Sigma(\Xi)](\vec{\omega}_0)$.

■

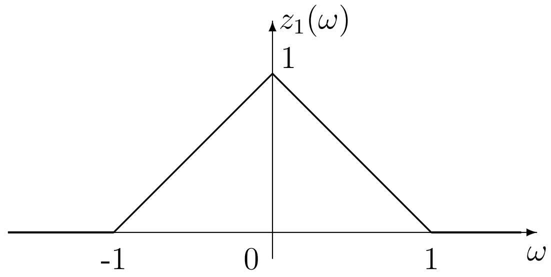
Example 2.18. [Discrete Gaussian observable]. Put $\Omega \equiv [a, b]$ ($\subseteq \mathbf{R}$, the real line), the closed interval. Let $\sigma > 0$. And let N be a sufficiently large fixed integer. Put $X_N \equiv \{\frac{k}{N} \mid k = 0, \pm 1, \pm 2, ..., \pm N^2\}$. And define the *discrete Gaussian observable* $\mathbf{O}_{\sigma^2, N}$ $\equiv (X_N, 2^{X_N}, F_{\sigma,N})$ in the commutative C^*-algebra $C([a, b])$ such that:

$$
\begin{aligned}
&[F_{\sigma,N}(\{k/N\})](\omega) \\
&= \begin{cases}
\frac{1}{\sqrt{2\pi\sigma^2}} \int_{N-\frac{1}{2N}}^{\infty} \exp[-\frac{(x-\omega)^2}{2\sigma^2}] dx & (k = N^2, \forall \omega \in [a, b]), \\
\frac{1}{\sqrt{2\pi\sigma^2}} \int_{\frac{k}{N}-\frac{1}{2N}}^{\frac{k}{N}+\frac{1}{2N}} \exp[-\frac{(x-\omega)^2}{2\sigma^2}] dx & (\forall k = 0, \pm 1, \pm 2, ..., \pm (N^2-1), \quad \forall \omega \in [a, b]), \\
\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{-N+\frac{1}{2N}} \exp[-\frac{(x-\omega)^2}{2\sigma^2}] dx & (k = -N^2, \forall \omega \in [a, b]).
\end{cases}
\end{aligned} \tag{2.60}
$$

And thus, for any Ξ ($\subseteq X_N$), we define $[F_{\sigma,N}(\Xi)](\omega) = \sum_{\frac{k}{N} \in \Xi} [F_{\sigma,N}(\{k/N\})]$ (ω). This $\mathbf{O}_{\sigma^2,N}$, as well as the d-dimensional Gaussian observable $\mathbf{O}_\Sigma$ (in Example 2.17), is the most important observable in classical measurements. ■

Example 2.19. [Fuzzy numbers observable (= triangle observable = round error observable)]. Let Δ be any positive number. Define the membership function (i.e., triangle fuzzy number) $\mathcal{Z}_\triangle$ $\Big(\in C_0(\mathbf{R})$, where $\mathbf{R}$ is the real line with the usual topology $\Big)$ such that:

$$
\mathcal{Z}_\triangle(\omega) = \begin{cases}
1 - \frac{\omega}{\Delta} & 0 \le \omega \le \Delta \\
\frac{\omega}{\Delta} + 1 & -\Delta \le \omega \le 0 \\
0 & \text{otherwise} .
\end{cases}
$$



Put $\mathbb{Z}_{\Delta} \equiv \{\Delta k \; : \; k \in \mathbb{Z} \equiv \{0, \pm 1, \pm 2, ...\}\}$. Define the C^*-observable $\mathbf{O}_{\mathcal{Z}_{\Delta}} \equiv (\mathbb{Z}_{\Delta}, \mathcal{P}_0(\mathbb{Z}_{\Delta}), \zeta^{\Delta}_{(\cdot)})$ in the commutative C^*-algebra $C_0(\mathbf{R})$ such that $\zeta^{\Delta}_{\Xi}(\omega) = \sum_{x \in \Xi} \mathcal{Z}_{\Delta}(\omega - x)$ $(\forall \Xi \in \mathcal{P}_0(\mathbb{Z}_{\Delta}),\ \forall \omega \in \mathbf{R})$. This C^*-observable is called a *fuzzy numbers observable* in $C_0(\mathbf{R})$. Putting $\Delta = 1$, we frequently use the fuzzy numbers observable $\mathbf{O}_{\mathcal{Z}} \equiv (\mathbb{Z}, \mathcal{P}_0(\mathbb{Z}), \zeta_{(\cdot)})$ in this book.

■

Example 2.20. [(i): Exact observable]. Let $\mathbb{Z}$ be the set of all integers, i.e., $\mathbb{Z} = \{0, \pm 1, \pm 2, ...\}$. And put $\mathcal{P}_0(\mathbb{Z}) = \{A(\subseteq \mathbb{Z}) \mid A \text{ is finite }\}$. Consider a commutative C^*-algebra $C_0(\mathbb{Z})$. And define the exact observable $\mathbf{O}_{\text{EXA}} \equiv (\mathbb{Z}, \mathcal{P}_0(\mathbb{Z}), E_{(\cdot)})$ in $C_0(\mathbb{Z})$ such that:

$$E_{\Xi}(n) = \begin{cases} 1 & n \in \Xi(\in \mathcal{P}_0(\mathbb{Z})) \\ \\ 0 & n \notin \Xi(\in \mathcal{P}_0(\mathbb{Z})) \end{cases} \tag{2.61}$$

which is called *the exact observable* (or, *fundamental observable*) in $C_0(\mathbb{Z})$. Of course we want to define the exact observable in $C_0(\mathbf{R})$ (or, $C([a, b])$). However, it is impossible in the C^*-algebraic formulation. For this, we must prepare the W^*-algebraic formulation (*cf.* Chapter 9).

[(ii): Approximate exact observable]. Though the exact observable in $C([0, 1])$ can not be defined, we have the approximate exact observable $\mathbf{O}^A_{\text{EXA}}$ in $C([0, 1])$ as follows: Let N be a sufficiently large integer. Put $X_N = \{\frac{1}{N}, \frac{2}{N}, \frac{3}{N}, ..., \frac{N}{N}(\equiv 1)\}$. Define the approximate exact observable $\mathbf{O}^A_{\text{EXA}} \equiv (X_N, \mathcal{P}(X_N), F)$ in $C([0, 1])$ such that:

$$[F(\{\frac{1}{N}\})](\omega) = \begin{cases} 1 & (0 \le \omega \le \frac{1}{N} - \frac{1}{N^2}) \\ -\frac{N^2}{2}(\omega - \frac{1}{N}) + \frac{1}{2} & (\frac{1}{N} - \frac{1}{N^2} \le \omega \le \frac{1}{N} + \frac{1}{N^2}) \\ 0 & (\frac{1}{N} + \frac{1}{N^2} \le \omega \le 1) \end{cases}$$

$$[F(\{\frac{N}{N}\})](\omega) = \begin{cases} 0 & (0 \le \omega \le \frac{N-1}{N} - \frac{1}{N^2}) \\ \frac{N^2}{2}(\omega - \frac{N-1}{N}) + \frac{1}{2} & (\frac{N-1}{N} - \frac{1}{N^2} \le \omega \le \frac{N-1}{N} + \frac{1}{N^2}) \\ 1 & (\frac{N-1}{N} + \frac{1}{N^2} \le \omega \le \frac{N}{N} - \frac{1}{N^2}) \end{cases}$$

For $n = 2, 3, ..., N-1$,

$$[F(\{\frac{n}{N}\})](\omega) = \begin{cases} 0 & (0 \le \omega \le \frac{n-1}{N} - \frac{1}{N^2}) \\ \frac{N^2}{2}(\omega - \frac{n-1}{N}) + \frac{1}{2} & (\frac{n-1}{N} - \frac{1}{N^2} \le \omega \le \frac{n-1}{N} + \frac{1}{N^2}) \\ 1 & (\frac{n-1}{N} + \frac{1}{N^2} \le \omega \le \frac{n}{N} - \frac{1}{N^2}) \\ -\frac{N^2}{2}(\omega - \frac{n}{N}) + \frac{1}{2} & (\frac{n}{N} - \frac{1}{N^2} \le \omega \le \frac{n}{N} + \frac{1}{N^2}) \\ 0 & (\frac{n}{N} + \frac{1}{N^2} \le \omega \le 1) \end{cases}$$

Note that the observable (i.e., fuzzy numbers observable) in Example 2.19 is also regarded as "approximate exact observable", if Δ is sufficiently small.

■

Example 2.21. [Null observable]. Define the observable $\mathbf{O}^{(\mathrm{nl})} \equiv (\{0,1\}, 2^{\{0,1\}}, F^{(\mathrm{nl})})$ in $\mathcal{A}$ such that:

$$F^{(\mathrm{nl})}(\emptyset) \equiv 0, \quad F^{(\mathrm{nl})}(\{0\}) \equiv 0, \quad F^{(\mathrm{nl})}(\{1\}) \equiv 1_{\mathcal{A}}, \quad F^{(\mathrm{nl})}(\{0,1\}) \equiv 1_{\mathcal{A}} \quad \text{in } \mathcal{A}, \tag{2.62}$$

which may be called the *null observable* (or, *existence observable*). Then, we have the measurement $\mathbf{M}_{\mathcal{A}}(\mathbf{O}^{(\mathrm{nl})} \equiv (\{0,1\}, 2^{\{0,1\}}, F^{(\mathrm{nl})}), S_{[\rho^p]})$. Note that:

($\sharp$) the probability that measured value (by $\mathbf{M}_{\mathcal{A}}(\mathbf{O}^{(\mathrm{nl})}, S_{[\rho^p]})$) is equal to 1 ($\in \{0,1\}$) is given by 1. That is, the measured value is always equal to 1 ($\in \{0,1\}$).

Thus, we think that "to take the measurement $\mathbf{M}_{\mathcal{A}}(\mathbf{O}^{(\mathrm{nl})}, S_{[\rho^p]})$" is the same as "to assure the existence of the system".

■

2.7 Operations of observables

Recall the identification (2.36), that is, we have the following identification:

$$\underset{\text{(real valued function on }\Omega)}{\widehat{F}_k} \longleftrightarrow \underset{\text{(crisp observable)}}{\mathbf{O}_k = (\mathbf{R}, \mathcal{B}_{\mathbf{R}}, F_k)} \text{ in } C(\Omega) \qquad (k = 1, 2, ..., n). \tag{2.63}$$

Note that $\widehat{F}_1 + \widehat{F}_2$, $\widehat{F}_1 \cdot \widehat{F}_2$, etc. are meaningful in the ordinary sense since $\widehat{F}_1$ and $\widehat{F}_2$ are real-valued functions. This makes us ask the following question.

- For each $k = 1, 2, ..., n$, consider an observable $\mathbf{O}_k \equiv (X_k, \mathcal{F}_k, F_k)$ in a C^*-algebra $\mathcal{A}$. Are $\mathbf{O}_1 + \mathbf{O}_2$, $\mathbf{O}_1 \cdot \mathbf{O}_2$, etc. meaningful in general? Or, how the operations of observables are defined?

This will be answered in what follows.

For each $k = 1, 2, ..., n$, consider an observable $\mathbf{O}_k \equiv (X_k, \mathcal{F}_k, F_k)$ in a C^*-algebra $\mathcal{A}$. Put $\mathbf{O} = \overset{\text{qp}}{\times}_{k=1,2,...,n} \mathbf{O}_k$. Let $g : \times_{k=1}^{n} \to Y$ be a measurable map, where Y has the subfield $\mathcal{G}$ of 2^Y. Then we can define the observable $(Y, \mathcal{G}, G)$, which is symbolically represented by $g(\mathbf{O}_1, \mathbf{O}_2, ..., \mathbf{O}_n)$, as follows:

- the $(Y, \mathcal{G}, G)$ is the image observable of the quasi-product observable $\mathbf{O} \equiv (\times_{k=1}^{n} X_k, \times_{k=1}^{n} \mathcal{F}_k, \widehat{F})$ concerning g (if it exists). That is,

$$(Y, \mathcal{G}, G) = g(\mathbf{O}) \tag{2.64}$$

i.e.,

$$G(\Gamma) = \widehat{F}(g^{-1}(\Gamma)) \qquad (\forall \Gamma \in \mathcal{G}). \tag{2.65}$$

Example 2.22. [The addition of triangle observables]. Let $\mathbf{O}_{\mathcal{Z}} \equiv (\mathbb{Z}, \mathcal{P}_0(\mathbb{Z}), \zeta_{(\cdot)})$ be the fuzzy numbers observable in $C_0(\mathbf{R})$ (*cf.* Example 2.19). Now let us calculate $\mathbf{O}_{\mathcal{Z}} + \mathbf{O}_{\mathcal{Z}}$ as follows: Note that the product observable $\mathbf{O}_{\mathcal{Z}} \times \mathbf{O}_{\mathcal{Z}} \equiv (\mathbb{Z}^2, \mathcal{P}_0(\mathbb{Z}^2), \zeta_{(\cdot)} \times \zeta_{(\cdot)})$ is represented by

(i) $|m - n| \geq 2$

$$[\zeta_{\{m\}} \times \zeta_{\{n\}}](\omega) = 0 \tag{2.66}$$

(ii) $|m - n| = 1$

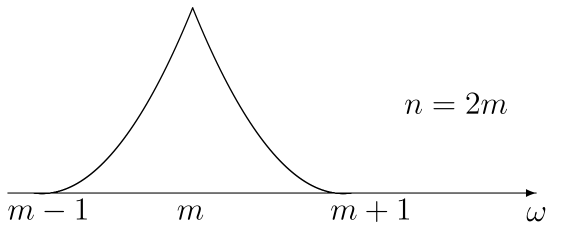
$$[\zeta_{\{m\}} \times \zeta_{\{n\}}](\omega) = \begin{cases} 0 & \omega \leq \min\{m, n\} \\ \frac{(x-m)(x-n)}{2} & \min\{m, n\} \leq \omega \leq \max\{m, n\} \\ 0 & \omega \leq \min\{m, n\} \end{cases}$$

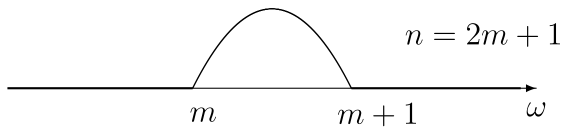
(iii) $m = n$

$$[\zeta_{\{m\}} \times \zeta_{\{m\}}](\omega) = \begin{cases} 0 & \omega \leq m - 1 \\ (x - (m-1))^2 & m - 1 \leq \omega \leq m \\ (x - (m+1))^2 & m \leq \omega \leq m + 1 \\ 0 & m + 1 \leq \omega \end{cases} \tag{2.67}$$

Thus we see

$$
\begin{aligned}
&(\zeta+\zeta)_{\{n\}}(\omega) \\
&= \begin{cases}
\underset{\text{(when } n=2m)}{(\zeta+\zeta)_{\{2m\}}(\omega)} = \begin{cases} 0 & \omega \le m-1 \\ (\omega-(m-1))^2 & m-1 \le \omega \le m \\ (\omega-(m+1))^2 & m \le \omega \le m+1 \\ 0 & m+1 \le \omega \end{cases} \\
\underset{\text{(when } n=2m+1)}{(\zeta+\zeta)_{\{2m+1\}}(\omega)} = \begin{cases} 0 & \omega \le m \\ -(\omega-(2m+1)/2)^2+1/2 & m \le \omega \le m+1 \\ 0 & m+1 \le \omega \end{cases}
\end{cases}
\end{aligned}
\tag{2.68}
$$





Therefore we get the $\mathbf{O}_{\mathbb{Z}} + \mathbf{O}_{\mathbb{Z}} \equiv (\mathbb{Z}, \mathcal{P}_0(\mathbb{Z}), (\zeta+\zeta)_{(\cdot)})$ in $C_0(\mathbf{R})$, where

$$
(\zeta+\zeta)_{\Xi}(\omega) = \sum_{n \in \Xi} (\zeta+\zeta)_{\{n\}}(\omega) \qquad (\Xi \in \mathcal{P}_0(\mathbb{Z}), \omega \in \Omega).
$$

■

Example 2.23. (χ^2-observable). Consider the (1-dimensional) Gaussian observable $\mathbf{O}_{\sigma^2} \equiv (\mathbf{R}, \mathcal{B}_{\mathbf{R}}^{\mathrm{bd}}, G^\sigma)$ in $\mathcal{A} \equiv C_0(\mathbf{R})$ such that:

$$
[G^\sigma(\Xi)](\mu) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\Xi} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \qquad (\forall \mu \in \mathbf{R}\ \forall \Xi \in \mathcal{B}_{\mathbf{R}}^{\mathrm{bd}}),
$$

(where σ^2 is a variance). And further, for each ϕ $(= 0, 1, 2, ...)$, define the product observable $(\mathbf{O}_{\sigma^2})^{\phi+1}$ such that

$$
(\mathbf{O}_{\sigma^2})^{\phi+1} = (\mathbf{R}^{\phi+1}, \mathcal{B}_{\mathbf{R}^{\phi+1}}^{\mathrm{bd}}, (G^\sigma)^{\phi+1}) \qquad (\text{ in } \mathcal{A} \equiv C_0(\mathbf{R})
$$

where

$$
(G^\sigma)^{\phi+1}(\Xi_1 \times \Xi_2 \times \cdots \times \Xi_{\phi+1}) = G^\sigma(\Xi_1) \times G^\sigma(\Xi_2) \times \cdots \times G^\sigma(\Xi_{\phi+1}).
$$

Define the map $g : \mathbf{R}^{\phi+1} \to \mathbf{R}$ such that

$$\mathbf{R}^{\phi+1} \ni (x_1, x_2, x_3, ..., x_{\phi+1}) \mapsto \sum_{k=1}^{\phi+1} \frac{(x_k - \frac{\sum_{j=1}^{\phi+1} x_j}{\phi+1})^2}{\sigma^2} \in \mathbf{R}.$$

The image observable $g((\mathbf{O}_{\sigma^2})^{\phi+1})$ is called the χ^2*-observable with* ϕ*, the degree of freedom.*

■

2.8 Frequency probabilities

The meaning of "probability" in Axiom 1 seems to be a matter of common knowledge in quantum mechanics. However, we, in this section, study the relation between "the probability in Axiom 1" and "frequency probability".

For each $k = 1, 2, ..., n$, consider a measurement $\mathbf{M}_{\mathcal{A}_k}\big(\mathbf{O}_k \equiv (X, \mathcal{P}(X), F_k), S_{[\rho_k^p]}\big)$ in a C^*-algebra $\mathcal{A}_k$, where we assume, for simplicity, that X is finite. Put $\widehat{\mathcal{A}} = \bigotimes_{k=1}^n \mathcal{A}_k$, i.e., the tensor product C^*-algebra of $\{\mathcal{A}_k \; : k = 1, 2, ..., n\}$. Here, consider the tensor-product C^*-observable $\bigotimes_{k=1}^n \mathbf{O}_k \equiv (X^n, \mathcal{P}(X^n), \widehat{F} \equiv \bigotimes_{k=1}^n F_k \;)$ in $\widehat{\mathcal{A}}$ $(\equiv \bigotimes_{k=1}^n \mathcal{A}_k \;)$ such that:

$$\widehat{F}(\Xi_1 \times \Xi_2 \times \cdots \times \Xi_n) = F_1(\Xi_1) \otimes F_2(\Xi_2) \otimes \cdots \otimes F_n(\Xi_n) \quad (\forall \Xi_k \in \mathcal{P}(X),\; k = 1, 2, ..., n). \tag{2.69}$$

Therefore, we get the measurement $\mathbf{M}_{\otimes \mathcal{A}_k}(\bigotimes_{k=1}^n \mathbf{O}_k, S_{[\bigotimes_{k=1}^n \rho_k^p]})$ in $\bigotimes_{k=1}^n \mathcal{A}_k$, which is also denoted by $\bigotimes_{k=1}^n \mathbf{M}_{\mathcal{A}_k}(\mathbf{O}_k, S_{[\rho_k^p]})$ and called the *repeated measurement* (or, *"parallel measurement"*) of $\mathbf{M}_{\mathcal{A}_k}(\mathbf{O}_k,\, S_{[\rho_k^p]})$'s. Put $\mathcal{M}_{+1}^m(X) = \{\nu : \nu$ is a positive measure on X such that $\nu(X) = 1 \}$ and define the map $g : X^n \to \mathcal{M}_{+1}^m(X)$ such that:

$$[g(x_1, x_2, ..., x_n)](\Xi) = \frac{\sharp[\{k : x_k \in \Xi\}]}{n} \quad (\forall \Xi \in \mathcal{P}(X)), \tag{2.70}$$

where $\sharp[B] =$ "the number of the elements of a set B".

Then we have the following proposition.

Proposition 2.24. [The weak law of large numbers, cf [56]]. *Suppose the above notations. For any* $\epsilon > 0$ *and any* Ξ $(\in \mathcal{P}(X))$*, define* $\widehat{D}_{\Xi,\epsilon}$ $(\in \mathcal{P}(X^n)\;)$ *by*

$$\widehat{D}_{\Xi,\epsilon} = \Big\{\widehat{x} = (x_1, x_2, ..., x_n) \in X^n \;:\; \Big|[g(\widehat{x})](\Xi) - \frac{1}{n}\sum_{k=1}^n \rho_k^p(F_k(\Xi))\Big| < \epsilon\Big\}. \tag{2.71}$$

Then we see that

$$1-\frac{1}{4\epsilon^2 n} \;\leq\; (\otimes_{k=1}^n \rho_k^p)\Big(\widehat{F}(\widehat{D}_{\Xi,\epsilon})\Big) \leq\; 1, \quad (\forall \Xi \in \mathcal{P}(X), \forall \epsilon > 0, \forall n). \tag{2.72}$$

Proof. We easily see that $[g(\widehat{x})](\Xi) = \frac{1}{n}\sum_{k=1}^n \chi_{\Xi}(\pi_k(\widehat{x}))$ $(\forall \widehat{x} = (x_1, x_2, ..., x_n) \in X^n)$, where $\pi_k : X^n \to X$ is defined by $\pi_k(\widehat{x}) \equiv \pi_k(x_1, x_2, ..., x_k, ..., x_n) = x_k$ and $\chi_{\Xi} : X \to \mathbf{R}$ is the characteristic function of Ξ (i.e., $\chi_{\Xi}(x) = 1$ $(x \in \Xi)$, $= 0$ $(x \notin \Xi)$). Using the terms in Kolmogorov's probability theory, we can say that $\chi_{\Xi}(\pi_k(\,\cdot\,))$, $k = 1, 2, ..., n$, are independent variables on a probability space $\Big(X^n, \mathcal{P}(X^n), \widehat{P}(\cdot) \equiv (\otimes_{k=1}^n \rho_k^p)(\widehat{F}(\cdot))\Big)$. Also it is clear that $\int_{X^n} \chi_{\Xi}(\pi_k(\widehat{x}))\widehat{P}(d\widehat{x}) = \int_{X^n}[\chi_{\Xi}(\pi_k(\widehat{x}))]^2 \widehat{P}(d\widehat{x}) = \rho_k^p(F_k(\Xi))$ $(k = 1, 2, ..., n)$. Therefore, by Čebyšev inequality, we see

$$\begin{aligned}
\widehat{P}\Big(X^n \setminus \widehat{D}_{\Xi,\epsilon}\Big) &= \widehat{P}\Big(\Big\{\widehat{x} \in X^n \;:\; \Big|\frac{\sum_{k=1}^n \chi_{\Xi}(\pi_k(\widehat{x})))}{n} - \frac{\sum_{k=1}^n \rho_k^p(F_k(\Xi))}{n}\Big| \geq \epsilon\Big\}\Big) \\
&\leq \frac{1}{\epsilon^2 n^2}\int_{X^n} \Big|\sum_{k=1}^n \Big(\chi_{\Xi}(\pi_k(\widehat{x})) - \rho_k^p(F_k(\Xi))\Big)\Big|^2 \widehat{P}(d\widehat{x}) \\
&= \frac{1}{\epsilon^2 n^2}\sum_{k=1}^n \int_{X^n} |\chi_{\Xi}(\pi_k(\widehat{x})) - \rho_k^p(F_k(\Xi))|^2 \widehat{P}(d\widehat{x}) \\
&\leq \frac{1}{\epsilon^2 n}\max_{1\leq k\leq n}\big[\rho_k^p(F_k(\Xi))(1-\rho_k^p(F_k(\Xi)))\big] \leq \frac{1}{4\epsilon^2 n},
\end{aligned} \tag{2.73}$$

which implies (2.72). This completes the proof. □

Now we can show the following theorem as an immediate consequence of Proposition 2.24. It clarifies the "probability" in Axiom 1 from the statistical point of view.

Theorem 2.25. [Frequency probability, *cf.* [42]]. *Put* $\mathcal{A}_k = \mathcal{A}$, $\rho_k^p = \rho^p$ *and* $\mathbf{O}_k = \mathbf{O}$ $\equiv (X, \mathcal{P}(X), F)$, $k = 1, 2, ..., n$, *in Proposition 2.24. Consider the repeated measurement* $\mathbf{M}_{\otimes\mathcal{A}}(\bigotimes_{k=1}^n \mathbf{O}, S_{[\otimes_{k=1}^n \rho^p]})$ *in* $\bigotimes_{k=1}^n \mathcal{A}$. *Then, we see that*

$$1-\frac{1}{4\epsilon^2 n} \;\leq\; (\otimes_{k=1}^n \rho^p)\Big((\bigotimes_{k=1}^n F)\big(\{\widehat{x} \in X^n : \big|\rho^p(F(\Xi)) - \frac{\sharp[\{k : x_k \in \Xi\}]}{n}\big| < \epsilon\}\big)\Big) \leq\; 1,$$

$$(\forall \Xi \in \mathcal{P}(X), \forall \epsilon > 0, \forall n).$$

Here note, by Axiom 1, that $(\otimes_{k=1}^n \rho^p)\Big((\bigotimes_{k=1}^n F)(\widehat{\Xi})\Big)$ *is the probability that a measured value by* $\mathbf{M}_{\otimes\mathcal{A}}(\bigotimes_{k=1}^n \mathbf{O}, S_{[\otimes_{k=1}^n \rho^p]})$ *belongs to* $\widehat{\Xi}$. *Therefore, if* n *is sufficiently large, for a measured value* $\widehat{x}$ $(= (x_1, x_2, ..., x_n) \in X^n)$ *by* $\mathbf{M}_{\otimes\mathcal{A}}(\bigotimes_{k=1}^n \mathbf{O}, S_{[\otimes_{k=1}^n \rho^p]})$, *we can consider* *(in the sense of (2.72)) that*

$$\rho^p(F(\Xi)) \approx \frac{\sharp[\{k : x_k \in \Xi\}]}{n}. \tag{2.74}$$

■

The (2.74) says that

- "probability in Axiom 1" = "frequency probability".

Thus, there is a reason that the probability space $(X, \mathcal{F}, \rho^p(F(\cdot)))$ is called *a sample space obtained by a measurement* $\mathbf{M}_{\mathcal{A}}\big(\mathbf{O}, S_{[\rho^p]}\big)$.

Remark 2.26. ["repeated measurement = iterated measurement" for $S_{[\delta_{\omega_0}]}$]. As seen in this section, we think that

"take a measurement M_{ω_0} N times" $\Leftrightarrow$ "take a measurement $\mathbf{M}_{\otimes_{n=1}^N C(\Omega)}(\otimes_{n=1}^N \mathbf{O}, S_{[\otimes_{n=1}^N \delta_{\omega_0}]})$ "

Thus, in classical measurements, we have the following identification:

"take a measurement $\mathbf{M}_{\otimes_{n=1}^N C(\Omega)}(\otimes_{n=1}^N \mathbf{O}, S_{[\otimes_{n=1}^N \delta_{\omega_0}]})$ " $\Leftrightarrow$ "take a measurement $\mathbf{M}_{C(\Omega)}(\mathbf{O}^N, S_{[\delta_{\omega_0}]})$ "

That is because it holds that

$${}_{\otimes_{n=1}^N \mathcal{M}(\Omega)}\Big\langle \otimes_{n=1}^N \delta_{\omega_0}, \otimes_{n=1}^N F(\Xi_n) \Big\rangle_{\otimes_{n=1}^N C(\Omega)} = {}_{\mathcal{M}(\Omega)}\Big\langle \delta_{\omega_0}, \times_{n=1}^N F(\Xi_n) \Big\rangle_{C(\Omega)}.$$

However, it should be noted that it does not always hold that "repeated measurement = iterated measurement" in statistical measurement theory (mentioned in Chapter 8) and quantum measurement theory.

■

Definition 2.27. [Semi-distance, moment method (inference for a pure state in repeated measurement)].
[(i): Semi-distance]. *Let Y be a set. If the map $\Delta : Y \times Y \to \mathbf{R}$ satisfies the following (a)∼(d):*

(a): $\Delta(x, y) \geq 0$ $(\forall x, y \in Y)$, *(b):* "$x = y$" $\Rightarrow \Delta(x, y) = 0$,
(c): $\Delta(x, y) = \Delta(y, x)$ $(\forall x, y \in Y)$, *(d):* $\Delta(x, y) \leq \Delta(x.z) + \Delta(z, y)$ $(\forall x, y, z \in Y)$,

then, the Δ is called a semi-distance on Y. In addition, if "(b'): $x = y \Leftrightarrow \Delta(x, y) = 0$" is assumed, then the Δ is called a distance (or metric) on Y.

[(ii): Moment method]. *Assume the* ρ_0^p *(in* $\mathbf{M}_{\mathcal{A}}\big(\mathbf{O} \equiv (X, \mathcal{F}, F),\ S_{[\rho_0^p]}\big)$*) is unknown. And further, we get the sample space* $(X, \mathcal{F}, \nu_0)$ *from the measured value* $\widehat{x}$ $(= (x_1, x_2,\ ...,\ x_n) \in X^n)$ *obtained by* $\mathbf{M}_{\otimes\mathcal{A}}(\bigotimes_{k=1}^n \mathbf{O}, S_{[\otimes_{k=1}^n \rho_0^p]})$*. That is,* $\nu_0(\Xi) \approx \frac{\sharp[\{k : x_k \in \Xi\}]}{n}$*. Note, by (2.74), that* $\rho^p(F(\Xi)) \approx \nu_0(\Xi)$ $(\forall \Xi \in \mathcal{F})$*. Let* Δ *be a semi-distance on* $\mathcal{M}_{+1}^m(X)$.[11] *Then, there is a very reason to infer the unknown* ρ_0^p $(\in \mathfrak{S}^p(\mathcal{A}^*))$ *such that*

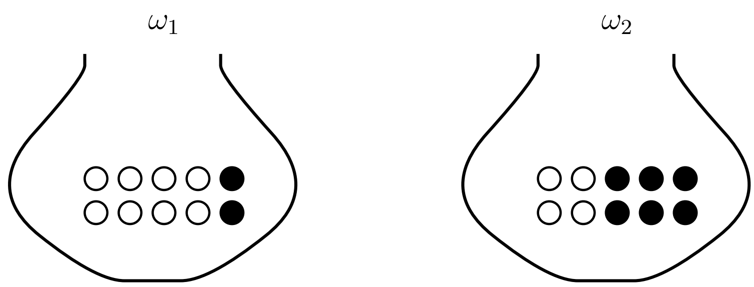
$$\Delta(\nu_0, \rho_0^p(F(\,\cdot\,))\,) = \min_{\rho^p \in \mathfrak{S}^p(\mathcal{A}^*)} \Delta(\nu_0, \rho^p(F(\,\cdot\,))\,).$$

This method is called "generalized moment method" or "moment method". Cf. §9.4. Note that the "semi-distance Δ *on* $\mathcal{M}_{+1}^m(X)$*" is not always unique. In this sense, the moment method is somewhat artificial.*

■

Example 2.28. [The urn problem by the moment method]. There are two urns ω_1 and ω_2. The urn ω_1 [resp. ω_2] contains 8 white and 2 black balls [resp. 4 white and 6 black balls]. Assume that they can not be distinguished in appearance. Choose one urn from the two. Assume that you do not know whether the chosen urn is ω_1 or ω_2. Now you sample, randomly, with replacement after each ball. In 7 samples, you get (w, b, b, w, b, w, b) in sequence where "w" = "white", "b" = "black".

(*Q*) *Which is the chosen urn,* ω_1 *or* ω_2*?*



[*Answer*]. We regard Ω $\big(\equiv \{\omega_1, \omega_2\}\big)$ as the state space. And consider the observable $\mathbf{O}\big(\equiv (X \equiv \{w, b\}, 2^{\{w,b\}}, F)\big)$ in $C(\Omega)$ where

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

Note that we have the real sample space $(X \equiv \{w, b\}, 2^{\{w,b\}}, \nu_0)$ such that:

$$\nu_0(\emptyset) = 0, \quad \nu_0(\{w\}) = 3/7, \quad \nu_0(\{b\}) = 4/7, \quad \nu_0(\{w, b\}) = 1.$$

[11] The definition of the semi-distance Δ may be too strong for the generalized moment method. However, in this book we focus on the above definition.

Also, note that the measurement

$$\mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[\delta_{\omega_1}]}) \quad [\text{resp. } \mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[\delta_{\omega_2}]})]$$

has the sample space

$$(X \equiv \{w, b\}, 2^{\{w,b\}}, [F(\cdot)](\omega_1)) \quad [\text{resp. } (X \equiv \{w, b\}, 2^{\{w,b\}}, [F(\cdot)](\omega_2))].$$

Thus, it suffices to compare

$$\Delta(\nu_0, [F(\cdot)](\omega_1)) \quad \text{and} \quad \Delta(\nu_0, [F(\cdot)](\omega_2)),$$

where Δ is a certain distance on $\mathcal{M}^m_{+1}(\{w, b\})$. For example define the distance Δ such that:

$$\Delta(\nu_1, \nu_2) = |\nu_1(\{w\}) - \nu_2(\{w\})| + |\nu_1(\{b\}) - \nu_2(\{b\})| \quad (\forall \nu_1, \nu_2 \in \mathcal{M}^m_{+1}(\{w, b\})).$$

Then, we see

$$\Delta(\nu_0, [F(\cdot)](\omega_1)) = |3/7 - 8/10| + |4/7 - 2/10| = 52/70$$

and

$$\Delta(\nu_0, [F(\cdot)](\omega_2)) = |3/7 - 4/10| + |4/7 - 6/10| = 10/70.$$

Thus, we can, by the moment method, infer that the unknown urn is ω_2.

■

2.9 Appendix (Bell's thought experiment)

(Continued from Example 2.15. Also see the footnote below[12])

2.9.1 EPR thought experiment

Although the original "EPR experiment (*cf.* [22])" was proposed in the framework of classical mechanics (*cf.* Chapter 12), the following argument is the quantum form of the "EPR experiment".[13]

---

[12] All appendixes in this book can be skipped.

[13] The argument in §2.9.1 is essentially the same as EPR-experiment (i.e., EPR-paradox, *cf.* [22]), which will be again discussed in §12.7.

Now consider the quantum system composed of two particles with the singlet state ρ_s (concerning z-axis) formulated in $B(\mathbf{C}^2 \bigotimes \mathbf{C}^2)$, where $\mathbf{C}^2 \bigotimes \mathbf{C}^2$ is the tensor Hilbert space of $\mathbf{C}^2$ and $\mathbf{C}^2$. The singlet state ρ_s is represented by $\rho_s = |\psi_s\rangle\langle\psi_s| \ \big(\in \mathfrak{S}^p(B(\mathbf{C}^2 \bigotimes \mathbf{C}^2)^*)\big)$, where

$$\psi_s = \frac{1}{\sqrt{2}}\big(\vec{e}_1 \otimes \vec{e}_2 - \vec{e}_2 \otimes \vec{e}_1\big) \quad (\in \mathbf{C}^2 \otimes \mathbf{C}^2). \quad \vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \in \mathbf{C}^2, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \in \mathbf{C}^2. \tag{2.75}$$

And consider the measurement $\mathbf{M}_{B(\mathbf{C}^2)\otimes B(\mathbf{C}^2)}\big(\mathbf{O}^z \otimes \mathbf{O}^z \equiv (Z^2 = \{\uparrow_z, \downarrow_z\}^2, 2^{Z^2}, F^z \otimes F^z), S_{[\rho_s]}\big)$, where

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

$$F^z(\emptyset) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad F^z(\{\uparrow_z, \downarrow_z\}) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Taking the measurement $\mathbf{M}_{B(\mathbf{C}^2)\otimes B(\mathbf{C}^2)}\big(\mathbf{O}^z \otimes \mathbf{O}^z \equiv (Z^2 = \{\uparrow_z, \downarrow_z\}^2, 2^{Z^2}, F^z \otimes F^z),\ S_{[\rho_s]}\big)$, we see that

(a) the probability that a measured value $(\uparrow_z, \uparrow_z)$ is obtained is equal to

$$=\rho_s\Big(F^z(\{\uparrow_z\}) \otimes F^z(\{\uparrow_z\})\Big)$$
$$={}_{\mathbf{C}^2\otimes\mathbf{C}^2}\Big\langle \frac{1}{\sqrt{2}}\big(\vec{e}_1 \otimes \vec{e}_2 - \vec{e}_2 \otimes \vec{e}_1\big), [F^z(\{\uparrow_z\}) \otimes F^z(\{\uparrow_z\})]\frac{1}{\sqrt{2}}\big(\vec{e}_1 \otimes \vec{e}_2 - \vec{e}_2 \otimes \vec{e}_1\big) \Big\rangle_{\mathbf{C}^2\otimes\mathbf{C}^2}$$
$$=0$$

(b) the probability that a measured value $(\uparrow_z, \downarrow_z)$ is obtained is equal to

$$=\rho_s\Big(F^z(\{\uparrow_z\}) \otimes F^z(\{\downarrow_z\})\Big)$$
$$={}_{\mathbf{C}^2\otimes\mathbf{C}^2}\Big\langle \frac{1}{\sqrt{2}}\big(\vec{e}_1 \otimes \vec{e}_2 - \vec{e}_2 \otimes \vec{e}_1\big), [F^z(\{\uparrow_z\}) \otimes F^z(\{\downarrow_z\})]\frac{1}{\sqrt{2}}\big(\vec{e}_1 \otimes \vec{e}_2 - \vec{e}_2 \otimes \vec{e}_1\big) \Big\rangle_{\mathbf{C}^2\otimes\mathbf{C}^2}$$
$$=1/2$$

(c) the probability that a measured value $(\downarrow_z, \uparrow_z)$ is obtained is equal to

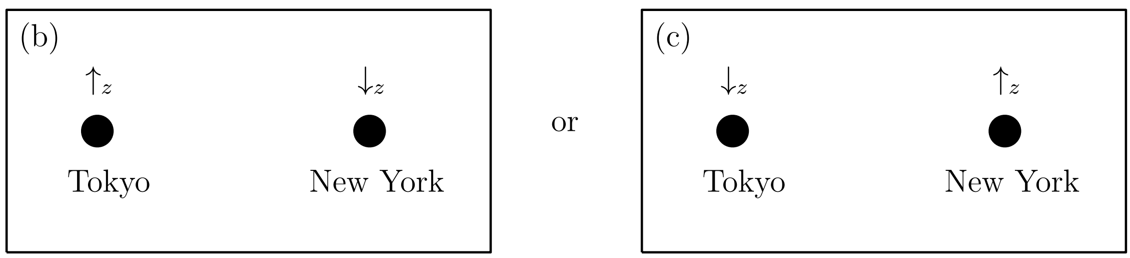
$$=\rho_s\Big(F^z(\{\downarrow_z\}) \otimes F^z(\{\uparrow_z\})\Big)$$
$$={}_{\mathbf{C}^2\otimes\mathbf{C}^2}\Big\langle \frac{1}{\sqrt{2}}\big(\vec{e}_1 \otimes \vec{e}_2 - \vec{e}_2 \otimes \vec{e}_1\big), [F^z(\{\downarrow_z\}) \otimes F^z(\{\uparrow_z\})]\frac{1}{\sqrt{2}}\big(\vec{e}_1 \otimes \vec{e}_2 - \vec{e}_2 \otimes \vec{e}_1\big) \Big\rangle_{\mathbf{C}^2\otimes\mathbf{C}^2}$$
$$=1/2$$

(d) the probability that a measured value $(\downarrow_z, \downarrow_z)$ is obtained is equal to

$$
\begin{aligned}
&= \rho_s\Big(F^z(\{\downarrow_z\}) \otimes F^z(\{\downarrow_z\})\Big) \\
&= {}_{\mathbf{C}^2\otimes\mathbf{C}^2}\Big\langle \frac{1}{\sqrt{2}}\big(\vec{e}_1 \otimes \vec{e}_2 - \vec{e}_2 \otimes \vec{e}_1\big), [F^z(\{\downarrow_z\}) \otimes F^z(\{\downarrow_z\})] \frac{1}{\sqrt{2}}\big(\vec{e}_1 \otimes \vec{e}_2 - \vec{e}_2 \otimes \vec{e}_1\big) \Big\rangle_{\mathbf{C}^2\otimes\mathbf{C}^2} \\
&= 0.
\end{aligned}
$$

Here, it should be noted that we can assume that the x_1 and the x_2 (in $(x_1, x_2) \in \{ (\uparrow_z, \uparrow_z), (\uparrow_z, \downarrow_z), (\downarrow_z, \uparrow_z), (\downarrow_z, \downarrow_z)\}$) are respectively obtained in Tokyo and in New York (or, in the earth and in the polar star).

(probability$\frac{1}{2}$) (probability$\frac{1}{2}$)



This fact is, figuratively speaking, explained as follows:

- Immediately after the particle in Tokyo is measured and the measured value $\uparrow_z$ [resp. $\downarrow_z$] is observed, the particle in Tokyo informs the particle in New York "Your measured value has to be $\downarrow_z$ [resp. $\uparrow_z$]".

Therefore, the above fact implies that quantum mechanics says that *there is something faster than light.* This is essentially the same as *the de Broglie paradox* (*cf.* [20]. Also see §9.3.3). That is,

- if we admit quantum mechanics, we must also admit the fact that there is something faster than light. (*cf.* [18, 78]). (2.76)

Of course we admit PMT, and therefore, we believe that there is something faster than light.

2.9.2 Bell's thought experiment

In this section, we review Bell's thought experiment in (quantum) measurement theory. (Cf. [9, 18, 78].) All the idea is, of course, owed to J.S. Bell [9]. Thus, we do not intend to assert our originality in this section. The argument is divided into two steps (i.e., [Step: I] and [Step: II]). [Step: I] is essentially the same as the previous section (i.e., §2.9.1).

[Step: I]. Let $a = (\alpha_1, \alpha_2)$ be any element in $\mathbf{R}^2$ such that $\|a\|_{\mathbf{R}^2} \equiv (|\alpha_1|^2 + |\alpha_2|^2)^{1/2} = 1$. Put

$$\sigma_a = \begin{bmatrix} 0 & \alpha_1 - \alpha_2\sqrt{-1} \\ \alpha_1 + \alpha_2\sqrt{-1} & 0 \end{bmatrix} \in B(\mathbf{C}^2), \quad \vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \in \mathbf{C}^2, \quad \vec{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \in \mathbf{C}^2.$$

It is easy to see that the self-adjoint matrix $\sigma_a : \mathbf{C}^2 \to \mathbf{C}^2$ has a unique spectral representation : $\sigma_a = F_a^{(1)} - F_a^{(-1)}$, where $F_a^{(1)}$ and $F_a^{(-1)}$ are orthogonal projections on $\mathbf{C}^2$ such that

$$F_a^{(1)} = \frac{1}{2}\begin{bmatrix} 1 & \alpha_1 - \alpha_2\sqrt{-1} \\ \alpha_1 + \alpha_2\sqrt{-1} & 1 \end{bmatrix}, \quad F_a^{(-1)} = \frac{1}{2}\begin{bmatrix} 1 & -\alpha_1 + \alpha_2\sqrt{-1} \\ -\alpha_1 - \alpha_2\sqrt{-1} & 1 \end{bmatrix}.$$

Define the observable $\mathbf{O}_a \equiv \big(X \equiv \{1, -1\}, \mathcal{P}(X), F_a\big)$ in $B(\mathbf{C}^2)$ such that $F_a(\{1\}) = F_a^{(1)}$ and $F_a(\{-1\}) = F_a^{(-1)}$

Now consider the quantum system composed of two particles with the singlet state ρ_s (concerning z-axis) formulated in $B(\mathbf{C}^2 \bigotimes \mathbf{C}^2)$, where $\mathbf{C}^2 \bigotimes \mathbf{C}^2$ is the tensor Hilbert space of $\mathbf{C}^2$ and $\mathbf{C}^2$. The singlet state ρ_s is represented by $\rho_s = |\psi_s\rangle\langle\psi_s| \ \big(\in \mathfrak{S}^p(B(\mathbf{C}^2 \bigotimes \mathbf{C}^2)^*)\big)$, where

$$\psi_s = \frac{1}{\sqrt{2}}\big(\vec{e}_1 \otimes \vec{e}_2 - \vec{e}_2 \otimes \vec{e}_1\big) \quad (\in \mathbf{C}^2 \otimes \mathbf{C}^2).$$

Put $a = (\alpha_1, \alpha_2)$, $b = (\beta_1, \beta_2) \in \mathbf{R}^2$ where $\|a\|_{\mathbf{R}^2} = \|b\|_{\mathbf{R}^2} = 1$. And define the tensor product observable $\mathbf{O}_{ab}$ ($\equiv \mathbf{O}_a \otimes \mathbf{O}_b) = (X^2, \mathcal{P}(X^2), F_a \bigotimes F_b)$ in $B(\mathbf{C}^2 \otimes \mathbf{C}^2)$ such that

$$(F_a \bigotimes F_b)(\{(x_1, x_2)\}) = F_a(\{x_1\}) \bigotimes F_b(\{x_2\}) \qquad (\forall (x_1, x_2) \in X^2 \equiv \{-1, 1\}^2).$$

Thus we get a measurement $\mathbf{M}_{B(\mathbf{C}^2 \otimes \mathbf{C}^2)}(\mathbf{O}_{ab}, S_{[\rho_s]})$ in $B(\mathbf{C}^2 \otimes \mathbf{C}^2)$. Axiom 1 says that the probability that a measured value x ($= (x_1, x_2)) \in X^2$ ($\equiv \{1, -1\}^2$) obtained by the measurement $\mathbf{M}_{B(\mathbf{C}^2 \otimes \mathbf{C}^2)}(\mathbf{O}_{ab}, S_{[\rho_s]})$ belongs to a set B ($\subseteq X^2$) is given by $\nu_{\text{EPR}}(B)$, where $\nu_{\text{EPR}}(B) = \sum_{x \equiv (x_1, x_2) \in B} \rho_s\big((F_a \otimes F_b)(\{(x_1, x_2)\})\big)$. Therefore, we see, for example, that

(♮) if we know that $x_1 = 1$, quantum mechanics says that the probability that $x_2 = 1$ [resp. $x_2 = -1$] is given by

$$\frac{\nu_{\text{EPR}}(\{1\} \times \{1\})}{\nu_{\text{EPR}}(\{1\} \times \{1, -1\})} \quad \Big[\text{resp.} \frac{\nu_{\text{EPR}}(\{1\} \times \{-1\})}{\nu_{\text{EPR}}(\{1\} \times \{1, -1\})}\Big]$$

and further, if we know that $x_1 = -1$, the probability that $x_2 = 1$ [resp. $x_2 = -1$] is given by

$$\frac{\nu_{\text{EPR}}(\{-1\} \times \{1\})}{\nu_{\text{EPR}}(\{-1\} \times \{1, -1\})} \quad \Big[\text{resp.} \frac{\nu_{\text{EPR}}(\{-1\} \times \{-1\})}{\nu_{\text{EPR}}(\{-1\} \times \{1, -1\})}\Big].$$

[Step: II]. Let $a^1(=(\alpha_1^1,\alpha_2^1))$, $a^2(=(\alpha_1^2,\alpha_2^2))$, $b^1(=(\beta_1^1,\beta_2^1))$ and $b^2(=(\beta_1^2,\beta_2^2))$ be elements in $\mathbf{R}^2$ such that $\|a^1\|_{\mathbf{R}^2}=\|a^2\|_{\mathbf{R}^2}=\|b^1\|_{\mathbf{R}^2}=\|b^2\|_{\mathbf{R}^2}=1$. Further, consider the parallel measurement $\bigotimes_{i,j=1,2}\mathbf{M}_{B(\mathbf{C}^2\otimes\mathbf{C}^2)}(\mathbf{O}_{a^ib^j},S_{[\rho_s]})$ in $\otimes_{i,j=1,2}B(\mathbf{C}^2\otimes\mathbf{C}^2)$ ($\equiv B(\otimes_{i,j=1,2}(\mathbf{C}^2\otimes\mathbf{C}^2)))$, that is,

$$\begin{aligned}&\bigotimes_{i,j=1,2}\mathbf{M}_{B(\mathbf{C}^2\otimes\mathbf{C}^2)}(\mathbf{O}_{a^ib^j},S_{[\rho_s]})\\ &=\mathbf{M}_{B(\otimes_{i,j=1,2}(\mathbf{C}^2\otimes\mathbf{C}^2))}\big((\mathop{\times}_{i,j=1,2}X^2,\mathcal{P}(\mathop{\times}_{i,j=1,2}X^2),\bigotimes_{i,j=1,2}(F_{a^i}\otimes F_{b^j})\big),S_{[\otimes_{i,j=1,2}\rho_s]}\big).\end{aligned}$$

Here note that $\otimes_{i,j=1,2}\rho_s=\rho_s\otimes\rho_s\otimes\rho_s\otimes\rho_s=|\psi_s\otimes\psi_s\otimes\psi_s\otimes\psi_s\rangle\langle\psi_s\otimes\psi_s\otimes\psi_s\otimes\psi_s|$ and

$$\mathop{\times}_{i,j=1,2}X^2\ni\big((x_1^{11},x_2^{11}),(x_1^{12},x_2^{12}),(x_1^{21},x_2^{21}),(x_1^{22},x_2^{22})\big)=x\in X^8\equiv\{-1,1\}^8.$$

Axiom 1 (2.37) says that the probability that a measured value $x\in X^8$ ($\equiv\{1,-1\}^8$) obtained by the parallel measurement $\bigotimes_{i,j=1,2}\mathbf{M}_{B(\mathbf{C}^2\otimes\mathbf{C}^2)}(\mathbf{O}_{a^ib^j},S_{[\rho_s]})$ belongs to a set B ($\subseteq X^8$) is given by $\nu_{\rm BTE}(B)$, where $\nu_{\rm BTE}(B)=\sum_{x\in B}\prod_{i,j=1,2}\rho_s\big((F_{a^i}\otimes F_{b^j})(\{(x_1^{ij},x_2^{ij})\})\big)$. That is, we have the sample space $(X^8,\mathcal{P}(X^8),\nu_{\rm BTE})$, which is induced by the parallel measurement $\bigotimes_{i,j=1,2}\mathbf{M}_{B(\mathbf{C}^2\otimes\mathbf{C}^2)}$ $(\mathbf{O}_{a^ib^j},\ S_{[\rho_s]})$.

Define the $\{-1,1\}$-valued functions g_k^{ij} on X^8, $(i,j,k=1,2)$, such that

$$g_k^{ij}((x_1^{11},x_2^{11}),(x_1^{12},x_2^{12}),(x_1^{21},x_2^{21}),(x_1^{22},x_2^{22}))=x_k^{ij}\qquad(\forall i,\forall j,\forall k\in\{1,2\}).\tag{2.77}$$

Note that it holds that

$$\begin{aligned}&\nu_{\rm BTE}\big((g_1^{11})^{-1}(\{1\})\big)=\nu_{\rm BTE}\big((g_1^{12})^{-1}(\{1\})\big),\quad \nu_{\rm BTE}\big((g_1^{21})^{-1}(\{1\})\big)=\nu_{\rm BTE}\big((g_1^{22})^{-1}(\{1\})\big),\\ &\nu_{\rm BTE}\big((g_2^{11})^{-1}(\{1\})\big)=\nu_{\rm BTE}\big((g_2^{12})^{-1}(\{1\})\big),\quad \nu_{\rm BTE}\big((g_2^{21})^{-1}(\{1\})\big)=\nu_{\rm BTE}\big((g_2^{22})^{-1}(\{1\})\big).\end{aligned}$$

Here note that (*cf.* (3.42) in §3.7 later)

$$g_1^{11}\neq g_1^{12},\quad g_1^{21}\neq g_1^{22},\quad g_2^{11}\neq g_2^{21},\quad g_2^{12}\neq g_2^{22}.\tag{2.78}$$

Moreover, define the correlation functions $P(g_1^{ij},g_2^{ij})$ $(i,j=1,2)$ by

$$P(g_1^{ij},g_2^{ij})\equiv\int_{X^8}g_1^{ij}(x)\cdot g_2^{ij}(x)\nu_{\rm BTE}(dx),\tag{2.79}$$

which may be also denoted by $P(a^i,b^j)$. A simple calculation shows that $P(a^i,b^j)=-(\alpha_1^i\beta_1^j+\alpha_2^i\beta_2^j)$. Thus, putting

$$a^1=(0,1),\quad b^1=(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}),\quad a^2=(1,0)\quad\text{and}\quad b^2=(\frac{1}{\sqrt{2}},-\frac{1}{\sqrt{2}}),$$

we see that

$$|P(a^1, b^1) - P(a^1, b^2)| \ + \ |P(a^2, b^1) + P(a^2, b^2)| = 2\sqrt{2}. \tag{2.80}$$

This is precisely Bell's calculation concerning Bell's though experiment.

The (2.80) can be tested by the repeated measurement $\bigotimes_{k=1}^{K} \big(\bigotimes_{i,j=1,2} \mathbf{M}_{B(\mathbf{C}^2 \otimes \mathbf{C}^2)} (\mathbf{O}_{a^i b^j}, S_{[\rho_s]}) \big)$. Let $\widehat{x} = \{ \big((x_{1,k}^{11}, x_{2,k}^{11}), (x_{1,k}^{12}, x_{2,k}^{12}), (x_{1,k}^{21}, x_{2,k}^{21}), (x_{1,k}^{22}, x_{2,k}^{22}) \big) \}_{k=1}^{K}$ be a measured value of the repeated measurement. Then, we see that

$$P(a^i, b^j) \approx \frac{1}{K} \sum_{k=1}^{K} x_{1,k}^{ij} x_{2,k}^{ij}$$

for sufficiently large K. Thus, the experimental test: "$2\sqrt{2}$ or not?" is possible. In fact, Aspect's experiment [8] is generally believed to guarantee the (2.80). It is, of course, important since quantum mechanics must be always tested.

(Continued in §3.7 (Appendix(Bell's inequality)))