Research Report

KSTS/RR-02/013
Dec. 27, 2002

Fundamentals of measurements

by

Shiro Ishikawa

Shiro Ishikawa
Department of Mathematics
Keio University

Department of Mathematics
Faculty of Science and Technology
Keio University

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

KSTS/RR-02/013
December 27, 2002

Fundamentals of Measurements*

Shiro Ishikawa
Department of Mathematics, Faculty of Science and Technology, Keio University
3-14-1 Hiyoshi, Kohoku-ku, Yokohama, 23-8522, Japan
ishikawa@math.keio.ac.jp

§1. Introduction

The phase: *"There is no science without measurements"* is an old saying, which emphasizes the importance of the "measurement". On the other hand, we do not have the authorized "measurement theory" yet. This fact seems to be strange if the concept of "measurement" is quite important in science. Thus, the propose of this lecture is to propose "measurement theory", which is expected to be the most fundamental theory in science.

Firstly, we begin with the conventional system theory (or shortly, "conventional ST", "CST",). It usually starts from the following equations:

$$\text{“CST”} \;\; = \begin{cases} \frac{dx(t)}{dt} = f(x(t), u_1(t), t), \quad x(0) = x_0 & \cdots \quad (\text{ state equation }) \;, \\ \qquad y(t) = g(x(t), u_2(t), t) & \cdots \quad (\text{ measurement equation}) \end{cases} \tag{1}$$

It should be noted that the term "measurement" appears in the beginning of the conventional system theory.

Also, quantum mechanics has the following form:

$$\text{“quantum mechanics”} \;\; = \;\; \text{“Schrödinger equation”} + \;\; \text{“Born's quantum measurements”} \tag{2}$$

*The annual conference of the Japan Society for Industrial and Applied Mathematics, (2002, September 19), Plenary Lecture,
English version (page 1-18), Japanese translation (page 19-35)

Here, it should be also noted that the term "measurement" appears in the beginning of quantum mechanics. That is, the conventional system theory and quantum theory are built in the same form —- "the rule of the time evolutioni"and "measurement". Of course, we believe that the above two measurements are closely related to each other, that is, these two have the same origin.

In [5∼13], we proposed "general system theory (in short, GST)" that unifies the above two (1) and (2). Our proposal is as follows. Born's quantum measurement theory is formulated in terms of Hilbert spaces. Thus it is easy to generalize and describe Born's quantum measurement theory in terms of C^*-algebra. Therefore, we can propose as follows.

$$\begin{aligned} & \text{"general system theory (or in short, GST)"} \\ = \ & \text{"the rule of time evolution"} + \text{"measurements"} \end{aligned} \tag{3}$$

This C^*-algebraic formulation of quantum mechanics is expected to include the above two (1) and (2), (see Remark 4.2 later). Of course, the (3) is merely our staring point. What we should do is to derive many interesting results from the (3).

Note that the conventional system theory (1) and quantum theory (2) are due to Newtonian mechanics. The discovery of quantum mechanics (2) is, of course, applauded sufficiently. On the other hand, the proposal of the conventional system theory (1) may be underestimated. However, we think that it is difficult to discover the concept of "measurement" in classical systems rather than in quantum mechanics. In fact, we have an opinion that even Newton missed (or underestimated) the importance of "measurement". Therefore, we assert that the (1), as well as the (2), should be applauded.

Here, we want to state our opinion for the question *"What is the general system theory (3)?"*. The general system theory (3) is constructed from mechanics (Newtonian mechanics and quantum mechanics). Thus we consider that

$$\begin{aligned} & \text{"the general system theory (3)"} \\ = \ & \text{"the mathematical representation of the mechanical world view"} \end{aligned} \tag{4}$$

That is, we consider that the general system theory (3) is a kind of epistemology called "the mechanical world view", namely, an epistemology to understand and analyze every

KSTS/RR-02/013
December 27, 2002

phenomenon (that appears in our usual life) — economics, psychology, engineering, OR and so on) — by an analogy of mechanics (or precisely, in the mathematical structure of mechanics). Although there may be several opinions for our assertion, everyone may agree that it is the most powerful method in all scientifical methods (cf. §6).

Summarizing the above argument, we say

(i) The mechanical world view is the most powerful scientific epistemology.

(ii) The general system theory (3) is the mathematical representation of the mechanical world view.

If this is true, we can, for the first time, understand the true meaning of the phase *"There is no science without measurements".*

[Remark 1.1] As the mathematical tools to analyze "every phenomenon", we may choose

$$\text{"the theory of differential equations"} + \text{"probability theory"} . \tag{5}$$

Thus some may consider that the (5) should be the starting point of system theory. However, the pioners, who constructed the conventional system theory (1), did not choose the (5) but the (1). We believe that their selection is quite excellent. *Why is the (5) improper?* Everyone should reconfirm the reason through this lecture.

§2. Several results derived from the GST (3)

Before we proceed to the main section 3, we mention several results derived from the GST (3). For this, we must first emphasize the wideness of the GST (3). That is,

($\sharp_1$) The three concepts: "measurement", "inference" and "control" are equivalent. Therefore, statistics and control theory are unified in terms of measurements.

For the further argument, see Remark 3.5 later.

Next we want to answer the question *"Why is the GST (3) essential?"*. Of course, since (1)⊂(3) (cf. Remark 4.2 later), the GST (3) is more applicable than the convention ST (1). In applications, this is one of advantages of the (3). However, we believe that the true reason that the GST (3) is needed is as follows. For example, we now consider the following problems:

KSTS/RR-02/013
December 27, 2002

(i). Justify the syllogism. (i.e., Show that the syllogism is true.)

(ii). Justify Fisher's maximal likelihood function method.

(iii). Justify Kalman filter.

(iv). Justify Zadeh's fuzzy sets theory.

(v). Show that the method to use differential equations is true.

Note that these problems have the same form such that "Justify ... ". Therefore we must first answer the question: *What is "justification"?*. For example, if some want to assert "Zadeh's fuzzy sets theory is true" (or "It is not true"), what must they do? The following opinion is completely wrong:

"To justify it" = "To describe it in terms of mathematics"

This is not true. That is because it is not difficult that any doubtful theory can be presented in terms of mathematics. After all, some may reach a vague conclusion such as "To justify it" is "To describe it in mathematics and moreover to show that it is useful". In fact, we have an opinion that the conventional statistics is generally considered to be justified under such vague definition. (For our justification of statistics, see Section 5 later.) However we believe that "true"≠"mathematical + useful". Our opinion is as follows.

$$\text{“To justify it”} = \text{“To regard it as one of aspect of the GST (3)”} \tag{6}$$

In general, we believe that "To justify it" is "To regard it as one of aspect of a certain world view". Therefore, it is quite important to investigate the question "Does another world view exist?". If we have another world view, we may have another definition of "justification". Though we do not know another world view yet, it must be always encouraged to try to propose another world view (cf. §6). The above (6) is, from the scientific point of view, our answer for the question *"Why is the GST (3) is needed?"*. In other words we consider that the (4) is equivalent to

$$\text{“GST (3)”} = \text{The Constitution of Science (i.e. engineering, economics, OR, etc.)} \tag{7}$$

KSTS/RR-02/013
December 27, 2002

if another world view is not proposed yet. Thus, we assert that "true"="constitutional under the GST (3)". Therefore, every theory (or, method) should be tested whether it is constitutional or unconstitutional.

Now we itemize our results derived from the GST (3) in what follows. The number corresponds to the reference number. These should be read under the spirit of (6) (or, (7)).

[Ref.2]. Heisenberg's uncertainty is formulated in terms of approximately simultaneous measurements. It clarifies the relation between Heisenberg's uncertainty and EPR-paradox.

[Refs. 3,4]. The concept of quantum particle's trajectories is formulated in terms of repeated measurements. It promotes the understanding of the Willson Chamber. And moreover, the numerical result of "two slit problem" is presented. Also, the concept of "trajectory" defined by the time series of measured values is also essential to the classical system (cf. [9]).

[Ref. 5]. The syllogism (i.e., "$A \to B, B \to C$" $\Longrightarrow$ "$A \to C$") does not hold in quantum systems. But it is always true in classical systems. These are proved in the GST (3). Also, we show that the strange logic (i.e., "$A \to B, B \to C$" $\Longrightarrow$ "$C \to A$") is true in some sense. This may be related to fuzzy logic. It is essential to prove "logic" in the GST (3). That is because it guarantees that the "logic" is valid in our usual life (i.e., in the world that is expected to be dominated by the mechanical world view)..

[Ref. 6]. In this paper, the assertion (4) is proposed. Many ideas (e.g. the entropy of measurement) are presented. If some consider that there is a gap between our formulation of fuzzy theory and Zadeh's fuzzy theory, they must conclude that Zadeh's fuzzy theory is out of the mechanical world view. Therefore, if they want to assert that Zadeh's fuzzy theory is true, they must propose another world view in order to justify Zadeh's fuzzy theory. We consider that this challenge is quite difficult (cf. §6).

[Ref. 7]. The "fuzzy modus ponens" is justified in the GST (3). We emphasize that every "logic" that is expected to be applied to our usual life must be proved in the GST (3).

[Ref. 8]. Factor analysis and principal component analysis in statistics are justified in the GST (3). We consider that, in most cases, the region that a statistical method is applied to is not clear in the conventional statistics. This is due the fact that the conventional

KSTS/RR-02/013
December 27, 2002

statistics is not formulated under a certain world view. In this sense, our approach has an advantage.

[Ref. 9]. Newtonian mechanics is reconstructed in the GST (3). Through the reconstruction, the relation between "the time evolution of the state" and "the trajectory of a particle" is clarified. And moreover, the measurement theoretical definition of Brown motion is proposed. In order to do it, Kolmogorov's extension theorem (that guarantees the existence of probability space) is no sufficient. Thus, the W^*-algebraic extension theorem is proposed. This theorem is most fundamental in measurement theory since it guarantees the existence of measurement.

[Ref. 10]. Statistics is formulated in the GST (3). However, the argument about Bayesian statistic is somewhat incomplete. This is completed in [13].

[Ref. 11]. We are convinced that the difference between statistical mechanics and Newtonian mechanics is due to the difference between rough measurements and precise measurements. Therefore equilibrium statistical mechanics can not be justified in Newtonian mechanics but in the classical GST (cf. Remark 4.2 later). This spirit is surely true. Therefore we now consider that the true foundation of statistical mechanics should not be due to the principle of a prior probability (cf. Remark 5.1 later). In this sense, our argument in [11] must be reconsidered.

[Ref. 12]. Kalman filter is formulated in the GST (3). This is shown under the quite weak conditions. Also, it should be noted that In the light of [13], we can see that Kalman filter is not related to probability density functions but weights. Also, we show that Bayes' Kalman filter is the dual concept of Fisher's Kalman filter.

[Ref. 13]. This completes the improper part of [10]. It will be mentioned in §5 later.

§**3. Measurememts**

In this and the nect section, we shall propose the GST (3) in terms of mathematics.

The GST is formulated in the framework of C^*-algebras. Let $\mathcal{A}$ be a C^*-algebra (e.g., [15]), that is, a Banach algebra with the involution "$*$" and the norm $||\cdot||$ satisfying the C^*-condition: $||F^*F|| = ||F||^2$ $(\forall F \in \mathcal{A})$. For simplicity, in this paper we always assume that $\mathcal{A}$ has the identity I. The linear functional $\rho(F)$ on $\mathcal{A}$ is denoted by ${}_{\mathcal{A}^*}\langle \rho, F \rangle_{\mathcal{A}}$, where $\rho \in \mathcal{A}^*$ (the dual Banach space) and $F \in \mathcal{A}$. An element F $(\in \mathcal{A})$ is said to be

KSTS/RR-02/013
December 27, 2002

self-adjoint if it holds that $F = F^*$. Also, a self-adjoint element F is called a *positive element* if $F = F_0^* F_0$ holds for some F_0 $(\in \mathcal{A})$. Define the *mixed state class* $\mathfrak{S}^m(\mathcal{A}^*)$ by $\{\rho \in \mathcal{A}^* \ : \ ||\rho||_{\mathcal{A}^*} = 1$ and $\rho(F) \geq 0$ for all $F \geq 0$ $\}$. And define the *pure state space* (or in short, *state space*) $\mathfrak{S}^p(\mathcal{A}^*)$ by "the set of all extreme points of $\mathfrak{S}^m(\mathcal{A}^*)$". Note that $\mathfrak{S}^p(\mathcal{A}^*)$ and $\mathfrak{S}^m(\mathcal{A}^*)$ are compact Hausdorff spaces in the sense of the weak*-topology $\tau(\mathcal{A}^*, \mathcal{A})$.

[**Example 3.1**] A C^*-algebra $\mathcal{A}$ is said to be *commutative* if it holds that $F_1 F_2 = F_2 F_1$ for all F_1, $F_2 \in \mathcal{A}$. Gelfand theorem says that any commutative C^*-algebra $\mathcal{A}$ can be identified with some $C(\Omega)$, the algebra composed of all complex valued continuous functions f on a compact Hausdorff space Ω. Here, the norm is defined by $||f|| = \sup_{\omega \in \Omega} |f(\omega)|$. Riesz representation theorem (e.g, [16]) says that $C(\Omega)^* = \mathcal{M}(\Omega)$, i.e., the Banach space composed of all regular complex-valued measures on Ω. We see and denote that $\mathfrak{S}^m(C(\Omega)^*)$ $= \{\rho^m \in \mathcal{M}(\Omega) | \ \rho^m \geq 0, ||\rho^m||_{\mathcal{M}(\Omega)} = 1\} \equiv \mathcal{M}_{+1}^m(\Omega)$, and $\mathfrak{S}^p(C(\Omega)^*) = \{\delta_\omega \in \mathcal{M}(\Omega) |$ δ_ω is a point measure at $\omega \in \Omega$, i.e., ${}_{\mathcal{M}(\Omega)}\langle \delta_\omega, f \rangle_{C(\Omega)} = f(\omega)$ ($\forall f \in C(\Omega)$, $\forall \omega \in \Omega$) $\} \equiv$ $\mathcal{M}_{+1}^p(\Omega)$. Under the identification: $\Omega \ni \omega \longleftrightarrow \delta_\omega \in \mathcal{M}_{+1}^p(\Omega)$, the Ω is also called a *state space*. The state space Ω and the mixed state ρ^m $(\in \mathcal{M}_{+1}^m(\Omega))$ is respectively called a *parameter space* and a *weight* in the conventional statistics.

[**Example 3.2**] Let $\mathbf{C}^n$ be the n-dimensional unitary space, i.e. the n-dimensional Hilbert space. Put $B(\mathbf{C}^n) = \{T : \mathbf{C}^n \to \mathbf{C}^n | \ T$ is a bounded linear operator $\}$, i.e., the set of all $n \times n$ matrices. Then, $B(\mathbf{C}^n)$ is a noncommutative C^*-algebra. We see that $\mathfrak{S}^p(B(\mathbf{C}^n)^*) = \{|x\rangle\langle x|($ Dirac notation $) \ : x \in \mathbf{C}^n, ||x|| = 1\}$.

As a natural generalization of Davies' idea ([1]) in quantum mechanics, a C^*-*observable* $\mathbf{O} \equiv (X, 2^X, F)$ in a C^*-algebra $\mathcal{A}$ is defined such that it satisfies that (i) X is a finite set, and 2^X $(\equiv \{\Xi \mid \Xi \subseteq X\})$, (in this paper, we focus on the case that X is finite) (ii) F is a map from 2^X into $\mathcal{A}$ such that $F(\Xi) \geq 0$ $(\forall \Xi \in 2^X)$, $F(\emptyset) = 0$ and $F(X) = I$, (iii) if $x_1 \neq x_2$, it holds that $F(\{x_1, x_2\}) = F(\{x_1\}) + F(\{x_2\})$. If $\mathbf{O}_1 \equiv (X, 2^X, F)$ and $\mathbf{O}_2 \equiv (Y, 2^Y, G)$ commute (i.e., $F(\{x\})G(\{y\}) = G(\{y\})F(\{x\})$ $(\forall (x,y) \in X \times Y)$), the *product observable* $\mathbf{O}_1 \times \mathbf{O}_2$ is defined by $(X \times Y,\ 2^{X \times Y},\ F \times G)$ where $(F \times G)(\{(x,y)\}) = F(\{x\})G(\{y\})$ $(\forall (x,y) \in X \times Y)$.

[**Remark 3.3**] Here we focus on the C^*-algebraic formulation. The W^*-algebraic formulation is also convenient (cf. particularly [9]). This formulation is handy to deal with "limit" and "convergence". For the general X (i.e., infinite set), see [6,9].

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 $(\in \mathfrak{S}^p(\mathcal{A}^*))$, an *observable* is represented by a C^*-observable $\mathbf{O} \equiv (X, 2^X, F)$ in the C^*-algebra $\mathcal{A}$. Also, a *measurement of the observable* $\mathbf{O}$ *for the system* S *with the state* ρ^p is denoted by $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[\rho^p]})$. We can obtain a *measured value* x $(\in X)$ by the measurement $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[\rho^p]})$. The measurement $\mathbf{M}_{\mathcal{A}}(\mathbf{O}_1 \times \mathbf{O}_2, S_{[\rho^p]})$ is sometimes called an iterated (or, simultaneous) measurement of two observables $\mathbf{O}_1$ and $\mathbf{O}_2$. Also, a *quantity* is represented by the self-adjoint element F $(\in \mathcal{A}$).

The axiom presented below is analogous to (or, a kind of generalization of) Born's probabilistic interpretation of quantum mechanics.

AXIOM 1. [Measurements]. *Consider a measurement* $\mathbf{M}_{\mathcal{A}}$ $(\mathbf{O} \equiv (X, 2^X,\ F),\ S_{[\rho^p]})$ *formulated in a* C^**-algebra* $\mathcal{A}$. *Then, the probability that a measured value* x $(\in X)$ *is obtained by the measurement* $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[\rho^p]})$ *is given by* ${}_{\mathcal{A}^*}\langle \rho^p, F(\{x\}) \rangle_{\mathcal{A}}$.

We have the classification of the GST as follows

$$\text{"GST(3)"} = \begin{cases} \text{"classical GST"} \\ \text{"quantum GST"} \end{cases} \tag{8}$$

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

[Remark 3.4] The above "probability" is individualistic. By the repeated $\bigotimes_{j=1}^N \mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[\rho^p]})$, we get the sample space $(X, \mathcal{F},\ \rho^p(F(\cdot))$) (Theorem 2.2 in [6]).

[Remark 3.5] When we take a measurement $\mathbf{M}_{\mathcal{A}}(\mathbf{O},\ S_{[\rho^p]})$, we usually know no information about the state ρ^p. Therefore, the measurement $\mathbf{M}_{\mathcal{A}}(\mathbf{O},\ S_{[\rho^p]})$ is often denoted by $\mathbf{M}_{\mathcal{A}}(\mathbf{O},\ S_{[\ \cdot\ ]})$. As mentioned in Section 2, statistics (S) and control theory (C) is respectively regarded as one of aspect of measurements as follows.

(S) We get a measured value $x (\in X)$ by a measurement $\mathbf{M}_{\mathcal{A}}(\mathbf{O},\ S_{[\ \cdot\ ]})$. Then, infer the state $[\ \cdot\ ]$.

(C) We want to get a measured value $x (\in X)$ by a measurement $\mathbf{M}_{\mathcal{A}}(\mathbf{O},\ S_{[\ \cdot\ ]})$. Then, settle the state $[\ \cdot\ ]$.

Of course, Fisher's maximal likelihood method is one of the answers of the above problems (S) and (C). However, it should be noted that Fisher's maximal likelihood method is due to Axiom 1. Cf. Section 5 later.

KSTS/RR-02/013
December 27, 2002

§4 The rule of time evolution

Let $\mathcal{A}$ be a C^*-algebra. A continuous linear operator $\Psi : \mathcal{A} \to \mathcal{A}$ is called a Markov operator, if it satisfies that (i) $\Psi(F) \geq 0$ for any positive element F in $\mathcal{A}$, (ii) $F(I_{\mathcal{A}}) = I_{\mathcal{A}}$, where $I_{\mathcal{A}}$ is the identity in $\mathcal{A}$. Here note that, for any observable $(X, \mathcal{F}, F)$ in $\mathcal{A}$, the $(X, \mathcal{F}, \Psi F)$ is an observable in $\mathcal{A}$, which is denoted by $\Psi\mathbf{O}$. Also, a Markov operator Ψ : $\mathcal{A} \to \mathcal{A}$ is called a homomorphism, if it satisfies that (i) $\Psi(F)\Psi(G) = \Psi(FG)$ for any F and G in $\mathcal{A}$, (ii) $(\Psi(F))^* = \Psi(F^*)$ for any F in $\mathcal{A}$. Let $\Psi^* : \mathcal{A}^* \to \mathcal{A}^*$ be the dual operator of a Markov operator $\Psi : \mathcal{A} \to \mathcal{A}$. Then the following mathematical results are well known. (a) $\Psi^*(\mathfrak{S}^m(\mathcal{A}^*)) \subseteq \mathfrak{S}^m(\mathcal{A}^*)$, (b) $\Psi^*(\mathfrak{S}^p(\mathcal{A}^*)) \subseteq \mathfrak{S}^p(\mathcal{A}^*)$ if $\Psi : \mathcal{A} \to \mathcal{A}$ is homomorphic. Cf. [16].

Let $\mathbf{R}$ be the real line, which is assumed to represent the time axis. Assume that a system S has the state ρ^p ($\in \mathfrak{S}^p(\mathcal{A}^*)$ at time $t = 0$, that is, the initial state is equal to ρ^p. Put $\mathbf{R}^2_{\leq} = \{(t_i, t_2) \in \mathbf{R}^2 : t_1 \leq t_2\}$. The pair $\mathbf{S}_{[\rho^p]} \equiv [S_{[\rho^p]}, \{\Psi_{t_1,t_2} : \mathcal{A} \to \mathcal{A}\}_{(t_1,t_2) \in \mathbf{R}^2_{\leq}}]$ is call a C^*-dynamical system with an intial system $S_{[\rho^p]}$ if it satisfies the following conditions (i)$\sim$(ii):

(i) $\rho^p \in \mathfrak{S}^p(\mathcal{A}^*)$,

(ii) for every $(t_1, t_2) \in \mathbf{R}^2_{\leq}$, $\Psi_{t_1,t_2} : \mathcal{A} \to \mathcal{A}$ is a Markov operator, and moreover, it holds that $\Psi_{t_1,t_2}\Psi_{t_2,t_3} = \Psi_{t_1,t_3}$ for all $(t_1, t_2), (t_2, t_3) \in \mathbf{R}^2_{\leq}$.

Now we can state the following axiom, which corresponds to "the rule of time evolution" in (3).

AXIOM 2. [The time evolution of a system]. *Let $S_{[\rho^p]}$ be a system with the (initial) state $\rho^p \in \mathfrak{S}^p(\mathcal{A}^*)$. The time evolution of the system S is represented by a C^*-dynamical system $\mathbf{S}_{[\rho^p]} \equiv [S_{[\rho^p]}, \{\Psi_{t_1,t_2} : \mathcal{A} \to \mathcal{A}\}_{(t_1,t_2) \in \mathbf{R}^2_{\leq}}]$. Also, the Heisenberg picture connects Axiom 1 with the C^*-dynamical system $\mathbf{S}_{[\rho^p]}$. (Cf. Remark 4.1 presented below).*

Though there is a reason to consider that Axiom 2 for non-deterministic relations should be called "Method 2", in this paper we do not dare to do so. Also, the term "deterministic" may not be proper in quantum mechanics, since "deterministic quantum general system" usually has a "stochastic" aspect. However, in this paper we dare to call so. We believe that this axiom dominates all systems, i.e., classical and quantum systems.

[Remark 4.1] (The Heisenberg picture). We must add some remark on the relation between Axiom 1 and Axiom 2. Assume $0 \le t_1 \le t_2 \le \cdots \le t_n$. And consider an observable $\mathbf{O}_{t_k} \equiv (X_{t_k}, \mathcal{F}_{t_k}, F_{t_k})$ for each time $_{t_k}$. The measurement of the observables $\{\mathbf{O}_{t_k}\}_{k=1}^n$ for the dynamical system $\mathbf{S}_{[\rho^p]} \equiv [S_{[\rho^p]}, \ \{\Psi_{t_1,t_2} : \mathcal{A} \to \mathcal{A}\}_{(t_1,t_2) \in \mathbf{R}^2_{\le}}]$, which is denoted by $\mathbf{M}(\{\mathbf{O}_{t_k}\}_{k=1}^n, \mathbf{S}_{[\rho^p]})$, is represented in what follows. Put $\widetilde{\mathbf{O}}_{t_n} = \mathbf{O}_{t_n}$. Consider the observable $\Psi_{t_{n-1},t_n}\widetilde{\mathbf{O}}_{t_n} = (X_{t_n}, \mathcal{F}_{t_n}, \Psi_{t_{n-1},t_n}F_{t_n})$, which is regarded as an observable at time t_{n-1}. By the Heisenberg picture, the $\Psi_{t_{n-1},t_n}\widetilde{\mathbf{O}}_{t_n}$ is identified with $\mathbf{O}_{t_n}$. Thus, we can define the quasi-product observable $\widetilde{\mathbf{O}}_{t_{n-1}} \equiv (X_{t_{n-1}} \times X_{t_n}, \ \mathcal{F}_{t_{n-1}} \times \mathcal{F}_{t_n}, \ F_{t_{n-1}} \times^{\mathbf{O}_{t_{n-1}}} \Psi_{t_{n-1},t_n}F_{t_n})$ of $\mathbf{O}_{t_{n-1}}$ and $\Psi_{t_{n-1},t_n}\widetilde{\mathbf{O}}_{t_n}$. Here note that the existence and the uniqueness of $\widetilde{\mathbf{O}}_{t_{n-1}}$ are not guaranteed in general. Similarly, we can define the observable $\widetilde{\mathbf{O}}_{t_{n-1}}$ of $\mathbf{O}_{t_{n-2}}$ and $\Psi_{t_{n-2},t_{n-1}}\widetilde{\mathbf{O}}_{t_{n-1}}$. And finally, we get the observable $\widetilde{\mathbf{O}}_{t_1}$ of $\mathbf{O}_{t_1}$ and $\Psi_{t_1,t_2}\widetilde{\mathbf{O}}_{t_2}$. Putting $\widetilde{\mathbf{O}}_0 = \Psi_{0,t_1}\mathbf{O}_{t_1}$, we have the measurement $\mathbf{M}_{\mathcal{A}}(\widetilde{\mathbf{O}}_0, S_{[\rho^p]})$, which is the representation of $\mathbf{M}(\{\mathbf{O}_{t_k}\}_{k=1}^n, \mathbf{S}_{[\rho^p]})$. However, note again that the existence and the uniqueness of $\widetilde{\mathbf{O}}_0$ are not guaranteed in general.

Now we can summarize the arguments in Sections 3 and 4 as follows.

$$
\begin{array}{ccccc}
\text{general system theory} & = & \text{the rule of time evolution} & + & \text{measurements} \\
 & & \text{(Axiom 2)} & & \text{(Axiom 1)}
\end{array}
\tag{3}
$$

Here Axioms 1 and 2 instruct us how to use the terms: "state", "observable", "measurement", "measured value", "probability", "time evolution", etc. To describe every phenomenon by these several words is the spirit of the mechanical world view.

[Remark 4.2] The following diagram will promote the better understanding of our argument.

$$
\begin{array}{ccc}
\underset{(\mathcal{A}=B(H))}{\text{Quantum Mechanics (2)}} & \underset{\text{generalization (a)}}{\longrightarrow} & \underset{(\text{general } \mathcal{A})}{\text{GST (3)}} \\
 & & \downarrow \textit{specialization(b)} \\
 & & \underset{(\text{commutative } \mathcal{A})}{\text{classical GST(8)}} \\
 & & \downarrow \textit{specialization(c)} \\
\underset{\text{(Newtonian eq.)}}{\text{Newtonian Mechanics}} & \underset{\text{generalization (d)}}{\longrightarrow} & \underset{\text{(state and measurement eqs.)}}{\text{conventional ST (1)}}
\end{array}
$$

As mentioned in Section 1, we believe that the generalization (d) is admirable. For the

KSTS/RR-02/013
December 27, 2002

specialization (c), see [7]. Thus we see that (1) ∪ (2) ⊂ (3). However, it should be noted that the main purpose of our papers [5∼13] is to assert that the classical GST (8) is, beyond comparison. richer than the conventional ST (1). That is, the conventional ST (1) is too poor to to assert the (4) (or, (7)), i.e., system theory is the constitution of science. Therefore, we hope that the GST (or, the classical GST) is generally accepted as the standard frame of system theory.

[Remark 4.3] Note that Axiom 2 guarantees that the method to use differential equations is true. However Axiom 2 is not sufficient in order to characterize Kalman filter (or particularly, regression analysis) in general. Thus Axiom 2 should be generalized as follows: "linear order $\mathbf{R} \longrightarrow$ partial ordered set T with a tree-like structure" and "$t_1 \neq t_2 (t_1, t_2 \in T) \Rightarrow \mathcal{A}_1 \neq \mathcal{A}_2$ in general". Therefore, "the rule of time evolution" should be replaced by "the relation among systems". Cf. [12]

[Remark 4.4] Some may be interested in the relation between GST (3) and Kolmogorov's probability theory. Note that Kolmogorov's probability theory is mathematics. Since the GST (3) is described in terms of mathematics, we can use many mathematical theories in the analysis of the GST (3). For example, we may use the theory of Hilbert spaces or finite mathematics. In the same sense, Kolmogorov's probability theory may be used. Therefore it is nonsense to compare the GST (3) with Kolmogorov's probability theory. Of course, it is certain that Kolmogorov's probability theory and the theory of differential equations are quite useful in the analysis of the GST (3). Again recall the difference between the (1) and the (5).

[Remark 4.5] We are not necessarily interested in the quantum part of the GST (3). Our interest is rather the classical GST. In fact, our most results in [5∼13] are mainly related to the classical GST. However, it is sure that the knowledge of the quantum GST promotes the better understanding of the classical GST. For example, *"There is no probability without measurements"* is a common sense in quantum theory. This common sense is also quite important in the consideration of the classical GST. It should be recalled that the classical GST is due to quantum mechanics (cf. Remark 4.2). Therefore, we consider that it is not advisable to return to quantum mechanics.

§5 An application of the GST (3) ——on the relation between Fisher's and Bayes' methods

As mentioned in Section 2, we have an opinion that statistics (even Fisher's method) is not justified yet. Therefore in this section we clarify the relation between between Fisher's and Bayes' methods in statistics as one of applications of the classical GST (cf. [13]). We consider that the term "subjectivity" in Bayesian statistics is too extraordinary in comparison with other science. Thus we are not concerned with the "subjectivity" in our approach.

From here and onward, we focus on the classical GST (i.e., $\mathcal{A} = C(\Omega)$), in which Fisher's method and Bayes' method will be clarified. That is, we show that, in Fisher's method [resp. in Bayes' method], the information (obtained by a measurement) is described in terms of quantities (i.e., real valued continuous functions on Ω) [resp. in terms of weights ($\in \mathcal{M}_{+1}^m(\Omega)$)].

A continuous linear operator $\Phi : C(\Omega) \to C(\Omega)$ is said to be a Markov operator if it satisfies that $\Phi(f) \geq 0$ $(\forall f \geq 0)$ and $\Phi(I) = I$. A subclass $\mathcal{S}_\Omega$ of the set of all Markov operators is called a *shuffle,* if it holds that

$$\underset{\text{(The shuffle } \mathcal{S}_\Omega \text{ has no effect)}}{\Phi(f) = f \quad (\forall \Phi \in \mathcal{S}_\Omega)} \Leftrightarrow \underset{\text{(no information about the card)}}{f(\omega) = kI \ \text{(i.e., constant)}} \tag{9}$$

The notes under the formula (1) should be read as the interpretation of the (1) in card games. Thus, there is a reason to consider that the likelihood quantity kI represents the situation that we have no information. (A non-negative quantity is called a *likelihood quantity* in Fisher's method.)

Assume a shuffle $\mathcal{S}_\Omega$. Under the hypothesis that *we know that the likelihood quantity of the system $S_{[\cdot]}$ is equal to a positive constant function kI* (in other words, *we have no information about the state δ_ω concerning the shuffle $\mathcal{S}_\Omega$*), the measurement $\mathbf{M}_{C(\Omega)}$ $(\mathbf{O} \equiv (X,\, 2^X, F),\, S_{[\delta_\omega]})$ is written by $\mathbf{M}_{C(\Omega)}(\mathbf{O},\, S_{[\cdot]}(kI)_{lq}^{\mathcal{S}_\Omega})$. (This is the precise form in Remark 3.5.) The conventional Fisher's likelihood function method says that

($\mathrm{F_C}$) From the fact that the measured value x $(\in X)$ is obtained by the $\mathbf{M}_{C(\Omega)}(\mathbf{O},\, S_{[\cdot]}(kI)_{lq}^{\mathcal{S}_\Omega})$, we know that the likelihood quantity of the system $S_{[\cdot]}$ is equal to $k[F(\{x\})](\omega)$. Thus, there is a reason to regard the unknown state $[\cdot]$ as the state $\omega_0 (\in \Omega)$ such that $k[F(\{x\})](\omega_0) = \max_{\omega \in \Omega} k[F(\{x\})](\omega)$.

For completeness, the reason of the ($\mathrm{F_C}$) is added as follows. Assume that $[F(\{x\})](\omega_1) < [F(\{x\})](\omega_2)$. Then, Axiom 1 says that the fact that the measured value x $(\in X)$ is obtained by the $\mathbf{M}_{C(\Omega)}(\mathbf{O},\, S_{[\delta_{\omega_1}]})$ happens more rarely than the fact that the measured value x $(\in X)$ is obtained by the $\mathbf{M}_{C(\Omega)}(\mathbf{O},\, S_{[\delta_{\omega_2}]})$ happens. Thus, there is a reason to regard the unknown state $[\cdot]$ as the state ω_0.

It is usual to assume that we have a little bit of information before a measurement. Thus we must proceed in what follows. Under the hypothesis that *we know that the likelihood quantity of the system $S_{[\cdot]}$ is equal to G_0,* the $\mathbf{M}_{C(\Omega)}(\mathbf{O},\, S_{[\delta_\omega]})$ is written by $\mathbf{M}_{C(\Omega)}(\mathbf{O},\, S_{[\cdot]}(G_0)_{lq}^{\mathcal{S}_\Omega})$.

Here we have the following problem:

($\mathrm{P_G}$) How to infer the new likelihood quantity of the system $S_{[\cdot]}$ from the fact that the measured value x $(\in X)$ is obtained by the $\mathbf{M}_{C(\Omega)}(\mathbf{O},\, S_{[\cdot]}(G_0)_{lq}^{\mathcal{S}_\Omega})$.

This is equivalent to the following problem:

($\mathrm{P'_G}$) How to infer the likelihood quantity of the system $S_{[\cdot]}$ from the fact that the measured value (y_0, x) $(\in \{y_0, y_1\} \times X)$ is obtained by the iterated measurement $\mathbf{M}_{C(\Omega)}(\mathbf{O}_0 \times \mathbf{O},\, S_{[\cdot]}(kI)_{lq}^{\mathcal{S}_\Omega})$, where $\mathbf{O}_0 = (\{y_0, y_1\},\, 2^{\{y_0, y_1\}}, G)$ and $G(\{y_0\}) = G_0$, $G(\{y_1\}) = I - G_0$.

Thus, from ($\mathrm{F_C}$) and "($\mathrm{P_G}$)$\leftrightarrow$($\mathrm{P'_G}$)", the ($\mathrm{P_G}$) is solved as follows:

[F] (The answer of the (P_G)): We know that the new likelihood quantity G_{new} of the system $S_{[\cdot]}$ is equal to $R_{\mathbf{M}[\mathbf{O};\{x\}]}(G_0)$. Here, the reduction map $R_{\mathbf{M}[\mathbf{O};\{x\}]} : C(\Omega) \to C(\Omega)$ is defined by $R_{\mathbf{M}[\mathbf{O};\{x\}]}\ (G) = F(\{x\})G\ (\forall G \in C(\Omega))$.

Now we understand Fisher's method (i.e., "no information kI" + [F]). For completeness, note that the information in Fisher's method is described in terms of likelihood quantities G_0.

Here we have a problem: *"What is Bayes' method ?".* Though there may be several opinions, in what follows we assert that Bayes' statistics is the statistics in which the information is described in terms of weights, and further, two methods are different representations of the same thing.

A shuffle $\mathcal{S}_\Omega^B$ (or, $\mathcal{S}_\Omega^B(\nu)$) is called a *mild shuffle* (or, *Bayes shuffle*), if there exists a unique weight ν ($\in \mathcal{M}_{+1}^m(\Omega)$) such that $\Phi^*(\nu) = \nu$ ($\forall \Phi \in \mathcal{S}_\Omega^B$), and further, if the ν satisfies that $\nu(U) > 0$ for any open set U ($\subseteq \Omega$). (Here the $\Phi^* : \mathcal{M}(\Omega) \to \mathcal{M}(\Omega)$ is the dual operator of the Φ.) The invariant weight ν is not destroyed by the Bayes shuffle $\mathcal{S}_\Omega^B$, or, it is invariant before and after the Bayes shuffle $\mathcal{S}_\Omega^B$. Thus, we consider that the invariant weight ν is also the representation of the situation that we have no information. Therefore, if we start from the Bayes shuffle $\mathcal{S}_\Omega^B(\nu)$, the concept of "no information" has two representations, i.e., kI and ν. If Ω is finite and if the bijection shuffle (cf. Example 5.2 later) is assumed, the invariant weight is always represented by the normalized counting measure $\nu_{_{|\Omega|}}(D) = |D|/|\Omega|$ ($\forall D \in 2^\Omega$). Also, if Ω is the compact domain in the n-dimensional Euclidean space $\mathbf{R}^n$ and if the $\mathcal{S}_\Omega^B(\nu)$ is the "diffusion shuffle" (i.e., the semi-group (e.g., [16]) generated by the heat equation on Ω with reflecting barrier $\partial\Omega$), then the invariant measure ν is represented by the normalized Lebesgue measure.

Assume a Bayes shuffle $\mathcal{S}_\Omega^B(\nu)$. Under the hypothesis that *we have no information* (*concerning the* $\mathcal{S}_\Omega^B(\nu)$) *about the state* δ_ω (in other words, *we know that the weight of the system* $S_{[\cdot]}$ *is equal to* ν), the $\mathbf{M}_{C(\Omega)}(\mathbf{O} \equiv (X, 2^X, F),\ S_{[\delta_\omega]})$ is written by $\mathbf{M}_{C(\Omega)}(\mathbf{O},\ S_{[\cdot]}(\nu)_w^{\mathcal{S}_\Omega^B})$. In general the symbol $\mathbf{M}_{C(\Omega)}(\mathbf{O},\ S_{[\cdot]}(\rho_0^m)_w^{\mathcal{S}_\Omega^B}))$, ($\rho_0^m \in \mathcal{M}_{+1}^m(\Omega)$), is assumed to be represents *the measurement* $\mathbf{M}_{C(\Omega)}$ ($\mathbf{O} \equiv (X, 2^X, F)$, $S_{[\delta_\omega]}$) *under the assumption that we know that the weight of the system* $S_{[\cdot]}$ *is* ρ_0^m. Thus, as the dual form of the [F], we get the following [B].

[B] Assume that we know that the weight of the system $S_{[\cdot]}$ is equal to ρ_0^m. After we get the measured value x by the measurement $\mathbf{M}_{C(\Omega)}$ ($\mathbf{O}$, $S_{[\cdot]}(\rho_0^m)_w^{\mathcal{S}_\Omega^B}$), we know that the new weight of the system $S_{[\cdot]}$ is equal to ρ_{new}^m ($\in \mathcal{M}_{+1}^m(\Omega)$) such that $\rho_{\mathrm{new}}^m\ (B) = \frac{\int_B [F(\{x\})](\omega)\rho_0^m(d\omega)}{\int_\Omega [F(\{x\})](\omega)\rho_0^m(d\omega)}$ ($\forall B \in \mathcal{B}_\Omega$, Borel field).

Define the map $\overline{R}_{\mathbf{M}[\mathbf{O};\{x\}]}^* : \mathcal{M}_{+1}^m(\Omega) \to \mathcal{M}_{+1}^m(\Omega)$ such that $\overline{R}_{\mathbf{M}[\mathbf{O};\{x\}]}^*(\rho^m) = \frac{R_{\mathbf{M}[\mathbf{O};\{x\}]}^*(\rho^m)}{\|R_{\mathbf{M}[\mathbf{O};\{x\}]}^*(\rho^m)\|_{\mathcal{M}(\Omega)}}$ ($\forall \rho^m \in \mathcal{M}_{+1}^m(\Omega)$), where $R_{\mathbf{M}[\mathbf{O};\{x\}]}^* : \mathcal{M}(\Omega) \to \mathcal{M}(\Omega)$ is the dual

KSTS/RR-02/013
December 27, 2002

operator of $R_{\mathbf{M}[\mathbf{O};\{x\}]}$. Then we see that $\rho_{\text{new}}^m = \overline{R}^*_{\mathbf{M}[\mathbf{O};\{x\}]}(\rho_0^m)$. Thus we consider that the [B] is the dual representation of the [F]. In this sense, the [B] is independent of the Bayes shuffle. However, as shown below, the Bayes shuffle $\mathcal{S}_\Omega^B$ is essential to the interpretation of [B].

Now we shall show that the concept of "no information" connects Fisher's method and Bayes' method. Assume a Bayes shuffle $\mathcal{S}_\Omega^B(\nu)$, which is, of course, also a shuffle. Since both kI and ν represent the concept of "no information" (or, they are not destroyed by the Bayes shuffle $\mathcal{S}_\Omega^B(\nu)$), we can compare $\mathbf{M}_{C(\Omega)}$ ($\mathbf{O}$, $S_{[\cdot]}(\rho_0^m)_w^{\mathcal{S}_\Omega^B}$) and $\mathbf{M}_{C(\Omega)}$ ($\mathbf{O}$, $S_{[\cdot]}(G_0)_{lq}^{\mathcal{S}_\Omega^B}$) as follows. Since $\nu(U) > 0$ for any open set U ($\subseteq \Omega$), we can define the bijective correspondence: [likelihood quantity G_0] $\underset{T_\nu}{\hookrightarrow}$ [weight ρ_0^m] such that $[T_\nu(G_0)](D) \equiv \rho_0^m(D) = \frac{\int_D G_0(\omega)\nu(d\omega)}{\int_\Omega G_0(\omega)\nu(d\omega)}$ ($\forall D \in \mathcal{B}_\Omega$), or $T_\nu^{-1}(\rho_0^m) = \frac{d\rho_0^m}{d\nu} = \frac{G_0}{\int_\Omega G_0(\omega)\nu(d\omega)}$, i.e., Radon-Nikodym derivative. Clearly, it holds that $T_\nu(kI) = \nu$, or more generally, $T_\nu(R_{\mathbf{M}[\mathbf{O};\{x\}]}(G_0)) = \overline{R}^*_{\mathbf{M}[\mathbf{O};\{x\}]}(\rho_0^m)$ if $T_\nu G_0 = \rho_0^m$. Therefore, the correspondence (or, translation) T_ν is compatible with [F] and [B]. (Cf. Example 5.2 later). Thus we can conclude that two methods are different representation of the same thing. That is, the Bayes shuffle (or, the concept of "no information") enables Fisher's method and Bayes' method to be translated from each other.

Of course, Bayes' method can say something more than Fisher's method since the class of shuffles is restricted. For example, the entropy of the measurement $\mathbf{M}_{C(\Omega)}$ ($\mathbf{O}$, $S_{[\cdot]}(\nu)_w^{\mathcal{S}_\Omega^B}$) (i.e., $\sum_{x\in X}\int_\Omega [F(\{x\})](\omega)\log[F(\{x\})](\omega)\nu(d\omega) - \sum_{x\in X} P(x)\log P(x)$, where $P(x) = \int_\Omega [F(\{x\})](\omega)\nu(d\omega)$, cf. [5]) is defined in only Bayes' method.

The interpretation of the invariant weight ν is not anything more than the definition says. It is usual that a certain *natural* measure μ (such as "counting measure", "volume", etc.) on Ω is ready prior to statistics. The natural measure μ may be also defined by an invariant measure of a certain class of transformations on Ω. Thus, we may often see that $\nu = \mu$. However, it does not necessarily mean that two measures ν and μ have the same interpretation. For completeness, we add the following remark.

[**Remark 5.1**] (The probabilistic interpretation of the Bayes shuffle $\mathcal{S}_\Omega^B(\nu)$). If $P[x;$ $\mathbf{M}_{C(\Omega)}(\mathbf{O},\ S_{[\cdot]}(\rho_0^m)_w^{\mathcal{S}_\Omega^B})]$, *i.e., the probability that a measured value* x ($\in X$) *is obtained by the measurement* $\mathbf{M}_{C(\Omega)}(\mathbf{O},\ S_{[\cdot]}(\rho_0^m)_w^{\mathcal{S}_\Omega^B})$*), is given by* ${}_{\mathcal{M}(\Omega)}\langle \rho_0^m, F(\{x\})\rangle_{C(\Omega)}$, we can easily see that for any $\mathbf{O}_1 \equiv (Y, 2^Y, G)$ and $y(\in Y)$, $P[y; \mathbf{M}_{C(\Omega)}(\mathbf{O}_1,\ S_{[\cdot]}(\rho_{\text{new}}^m)_w^{\mathcal{S}_\Omega^B})] = \frac{P[(x,y);\mathbf{M}_{C(\Omega)}(\mathbf{O}\times\mathbf{O}_1, S_{[\cdot]}(\rho_0^m)_w^{\mathcal{S}_\Omega^B})]}{P[x;\mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[\cdot]}(\rho_0^m)_w^{\mathcal{S}_\Omega^B})]} = {}_{\mathcal{M}(\Omega)}\langle \rho_{\text{new}}^m, G(\{y\})\rangle_{C(\Omega)}$. Therefore, if $P[x;$ $\mathbf{M}_{C(\Omega)}$ ($\mathbf{O}$, $S_{[\cdot]}(\nu)_w^{\mathcal{S}_\Omega^B})] = {}_{\mathcal{M}(\Omega)}\langle \nu, F(\{x\})\rangle_{C(\Omega)}$ holds (i.e., if "the principle of equal probability" can be assumed), the [B] has the probabilistic interpretation (cf. Example 5.2 later). However, it is not a direct consequence of our theory (="GST").

The following example will promote the understanding of our proposal.

[Example 5.2] (The urn problem). There are two urns ω_r and ω_b, whose prices is respectively 14 dollars and 21 dollars. The urn ω_r [resp. ω_b] contains 8 red and 2 blue balls [resp. 4 red and 6 blue balls]. Assume that they can not be distinguished in appearance (i.e., they are shuffled). Choose one urn from the two. Now you sample, randomly, with replacement after each ball. In 2 samples, you get 1 red and 1 blue in sequence, i.e., ("red", "blue"). *Which is the chosen urn, ω_r or ω_b ?* Further, assume that you continuously get "red". *How about the case ?* Also, answer the question *"How much is the average price of the two urns (at this moment)?".* Lastly, study the case that the urn is chosen by a fair coin tossing.

In what follows this problem is studied in Fisher's method [resp. Bayes' method]. Put $\Omega = \{\omega_r, \omega_b\}$. $\mathbf{O} = (\{r,b\}, 2^{\{r,b\}}, F)$ where $[F(\{r\})](\omega_r) = 0.8$, $[F(\{b\})](\omega_r) = 0.2$, $[F(\{r\})](\omega_b) = 0.4$, $[F(\{b\})](\omega_b) = 0.6$. Assume the Bayes shuffle $\mathcal{S}_\Omega^B$ such that $\mathcal{S}_\Omega^B = \{\Phi_\phi | \ \phi : \Omega \to \Omega$ is a bijection $\}$, where $\Phi_\phi : C(\Omega) \to C(\Omega)$ is defined by $[\Phi_\phi(f)](\omega) = f(\phi(\omega))$ $(\forall f \in C(\Omega), \forall \omega \in \Omega)$. (Note that Bayes' method can not be applicable, if the shuffle $\mathcal{S}_\Omega \equiv \{\Phi_\phi | \ \phi : \Omega \to \Omega$ is a map $\}$ is first assumed.) The situation of no information in Fisher's method [resp. Bayes' method] is represented by kI $(k > 0)$ [resp. $\nu_{|2|}$ (i.e., $\nu_{|2|}(\{\omega_r\}) = \nu_{|2|}(\{\omega_b\}) = 1/2$)]. Thus, it suffices to consider the measurement $\mathbf{M}_{C(\Omega)}(\mathbf{O} \times \mathbf{O}, \ S_{[\cdot]}(kI)_{lq}^{\mathcal{S}_\Omega^B})$ [resp. $\mathbf{M}_{C(\Omega)}(\mathbf{O} \times \mathbf{O}, \ S_{[\cdot]}(\nu_{|2|})_w^{\mathcal{S}_\Omega^B})$]. Since the measured value (r, b) was obtained, the new likelihood quantity $G_{\rm new}$ and the new weight $\rho^m_{\rm new}$ is respectively given as follows.

$$G_{\rm new}(\omega_r)(= kI \cdot [F(\{r\})](\omega_r) \cdot [F(\{b\})](\omega_r)) = 0.16k, \quad G_{\rm new}(\omega_b) = 0.24k,$$

(by the direct calculation or translation: $T_{\nu_{|2|}}(G_{\rm new}) = \rho^m_{\rm new}$),

$$\rho^m_{\rm new}(\{\omega_r\})(= \frac{\int_{\{\omega_r\}}[F(\{r\})](\omega)[F(\{b\})](\omega)\nu_{|2|}(d\omega)}{\int_\Omega [F(\{r\})](\omega)[F(\{b\})](\omega)\nu_{|2|}(d\omega)}) = \frac{16}{40} = \frac{2}{5}, \ \rho^m_{\rm new}(\{\omega_b\}) = \frac{3}{5}.$$

Thus there is a reason to infer that $[\cdot] = \omega_b$ (cf. Fisher's likelihood function method $(\mathrm{F_C})$). For the further case, it suffices to consider the measurement $\mathbf{M}_{C(\Omega)}(\mathbf{O}, \ S_{[\cdot]}(G_{\rm new})_{lq}^{\mathcal{S}_\Omega^B})$ [resp. $\mathbf{M}_{C(\Omega)}(\mathbf{O}, \ S_{[\cdot]}(\rho^m_{\rm new})_w^{\mathcal{S}_\Omega^B})$]. Thus we similarly calculate that $G_{{\rm new}^2}(\omega_r)$ ($= [G_{\rm new}](\omega_r) \cdot [F(\{r\})](\omega_r)$) $= 0.128k$, $G_{{\rm new}^2}(\omega_b) = 0.096k$, [resp. $\rho^m_{{\rm new}^2}(\{\omega_r\}) = \frac{4}{7}$, $\rho^m_{{\rm new}^2}(\{\omega_b\}) = \frac{3}{7}$]. (Here again note that $G_{{\rm new}^2} \underset{T_{\nu_{|2|}}}{\longleftrightarrow} \rho^m_{{\rm new}^2}$) Putting $p(\omega_r) = 14$ and $p(\omega_b) = 21$, we see that their (weighted) average price at this moment is equal to $\int_\Omega p(\omega)\rho^m_{{\rm new}^2}(d\omega) = 17$ (dollars). (This is the result of Bayesian statistics.) Here, it should be noted that, in the above arguments, any statement is not described in terms of probabilities, though the probability in Axiom 1 is essential to the justification of $(\mathrm{F_C})$. However, if the urn is chosen by a fair coin tossing, Bayes' method acquires the probabilistic interpretation. Therefore, we can, for example, calculate that $P[r; \mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[\cdot]}(\rho^m_{\rm new})_w^{\mathcal{S}_\Omega^B})] = {}_{\mathcal{M}(\Omega)}\langle \rho^m_{\rm new}, F(\{r\})\rangle_{C(\Omega)} = 0.56$. For completeness, note that this probability is due to "Method 1 in Remark 5.1" and not Axiom 1. The two probabilities in Axiom 1 and "Method 1" must not be confused.

KSTS/RR-02/013
December 27, 2002

§6. Conclusions

In this lecture we introduce the general system theory (or shortly, GST)

$$\text{“GST”} \;=\; \underset{\text{(Axiom 2, cf. Remark 4.3)}}{\text{“the relation between systems”}} \;+\; \underset{\text{(Axiom 1)}}{\text{“measurement theory”}} \tag{3}$$

as the mathematical representation of the mechanical world view. Here Axioms 1 and 2 instruct us how to use the terms: “state”, “observable”, “measurement”, “measured value”, “probability”, “time evolution”, etc. To describe every phenomenon by these several words is the spirit of the mechanical world view. We believe that the GST is powerful enought to say that

$$\text{“To justify a theory”} \;=\; \text{“To regard a theory as one of aspects of the GST”} \tag{6}$$

Therefore, we may say that

$$\text{“GST”} \;=\; \text{“The mathematical representation of the mechanical world view”} \tag{4}$$

$$=\; \text{“such a thing as the constitution of science”} \tag{7}$$

Therefore, we believe that the saying *“There is no science without measurements”* is realized in the GST (3).

Also, in Section 5 we justify statistics as follows.

$$\underset{\text{(Axioms 1,2)}}{\text{GST}} \quad \underset{\text{(statistics)}}{\Longrightarrow} \quad \left\{ \begin{array}{ll} \underset{\text{likelihood quantity (} \mathcal{S}_\Omega \text{)}}{\Longrightarrow} & \underset{\text{(“no information (} = kI\text{)” + [F])}}{\text{Fisher's method}} \\ & \quad \updownarrow \text{ translation } T_\nu (\text{if } \mathcal{S}_\Omega = \mathcal{S}_\Omega^B) \\ \underset{\text{weight (} \mathcal{S}_\Omega^B \text{)}}{\Longrightarrow} & \underset{\text{(“no information (} = \nu\text{)” + [B])}}{\text{Bayes' method}} \end{array} \right.$$

Here the “translation T_ν” between Fisher's and Bayes' methods is essential. Therefore, we conclude that Fisher's and Bayes' methods are true, or they are constitutional under the constitution: “classical GST”.

Lastly we shall mention our present interests as follows.

(A) Every method (particularly, in OR, psychology, etc.) should be checked under the GST. If a *useful* and *unconstitutional* method will be discovered, it is very exciting. That is because it may imply that we have a good chance to propose another world view (cf. the (B) below).

This must be executed as fast as possible. And

(B) Propose another world view.

As mentioned in Section 2, we believe that "To justify it" is "To regard it as one of aspect of a certain world view". If we have another world view, we may have another definition of "justification". Therefore, the above (B) is quite important. Though we do not know another world view yet, it must be always encouraged to try to propose another world view. Although there is an opinion that Frieden's recent work [17] is hopeful as another world view, now we can not decide whether it is hopeful or not. We consider that it is desirable that many world views will be proposed. However we are convinced that the GST plays a central role in the argument about "world view" since the GST is quite orthodox (or, it is due to mechanics). We hope that the GST will be generally accepted as the standard frame of system theory, or as the mathematical representation of the mechanical world view..

References

[1] E. B. Davies, *Quantum theory of open systems*, Academic Press 1976

[2] S. Ishikawa, *Uncertainty relation in simultaneous measurements for arbitrary observables* , *Rep. Math. Phys.* **29**, 257-273 (1991)

[3] S. Ishikawa, *Uncertainties an an interpretation of nonrelativistic quantum theory Internat. J. Theoret. Phys.* , **30**, 401-417 (1991)

[4] S. Ishikawa, T. Arai, T. Kawai *Numerical Analysis of trajectories of a quantum particle in two-slit experiment. Internat. J. Theoret. Phys.* , **33**, 1265-1274 (1994)

[5] S. Ishikawa, *Fuzzy inferences by algebraic method*, *Fuzzy Sets and Systems* **87**, 181-200 (1997)

[6] S. Ishikawa, *A quantum mechanical approach to a fuzzy theory*, *Fuzzy Sets and Systems* **90**, 277-306 (1997)

[7] S. Ishikawa, *Fuzzy logic in measurements*, *Fuzzy Sets and Systems* **100**, 291-300 (1998)

[8] S. Ishikawa, A. Iida, *A system theoretical characterization of factor analysis*, EUFIT'98 , Vol. 2. 1257-1261 (1998), (September 7-10, Aachen)

[9] S. Ishikawa, T. Arai, T. Takamura, A dynamical system theoretical approach to Newtonian mechanics, *Far east journal of dynamical systems* **1**, 1-34 (1999)

[10] S. Ishikawa, *Statistics in measurements*, *Fuzzy Sets and Systems* **116**, 141-154 (2000)

[11] S. Ishikawa A dynamical system theoretical approach to equilibrium statistical mechanics *Far east journal of dynamical systems* **3**, 9-28 (2001)

[12] S. Ishikawa The existence theorem of Kalman filter in measurement theory *Far east journal of dynamical systems* **3**, 175-204 (2001)

KSTS/RR-02/013
December 27, 2002

[13] S. Ishikawa, A system theoretical characterization of the relation between Fisher's and Bayes' methods , *Far east journal of theoretical statistics* **7**, 19-33 (2002)

[14] A. Kolmogorov, *Foundations of probability* (*translation*), Chelsea Publishing Co. 1950

[15] S. Sakai, *C^*-algebras and W^*-algebras*, Ergebnisse der Mathematik und ihrer Grenzgebiete (Band 60), Springer-Verlag, Berlin, Heidelberg, New York 1971

[16] K. Yosida, *Functional analysis*, Springer-Verlag (Sixth Edition) 1980

[17] B. R. Frieden *Fisher information in physics* , Cambridge University Press (1998)

[Notice]
For the details of Ishikawa's papers, see http://www.math.keio.ac.jp/˜ ishikawa.

KSTS/RR-02/013
December 27, 2002

測定理論の数理[1]

石川 史郎
(慶應義塾大学理工学部数理科学科)

ishikawa@math.keio.ac.jp

§1. 序.

「測定なくして科学なし」とは昔から言い古されている言葉である．この言葉の意味は様々な解釈が可能かもしれないが，いずれにしても「測定」の重要さを的確に表している言葉であると思う．ところが,「測定理論」と称されるべき科学理論が一般に認知されているかというと疑問である．最も重要な科学的概念にしかるべき理論がないという事実は不可解である．我々はこの不可解な事実をチャンスと見て，ここでは,「測定理論」のあるべき姿としての一つの提案をしたい.

まず，システム理論から始める．システム理論は次の式からスタートするのが普通である.

$$\begin{cases} \frac{dx(t)}{dt} = f(x(t), u_1(t), t), \quad x(0) = x_0 & \cdots \ (\text{ 状態方程式 }), \\ \quad y(t) = g(x(t), u_2(t), t) & \cdots \ (\text{ 測定方程式 }) \end{cases} \tag{1}$$

「測定」という言葉がシステム理論の最初に述べられていることに注目していただきたい.

また，量子力学は次の形で書ける.

$$\text{「量子力学」} = \text{「 Schrödinger 方程式」} + \text{「ボルンの測定理論」} \tag{2}$$

ここでも「測定」という言葉が最初に宣言されている．すなわち，システム理論と量子力学は同じ形—— 「時間発展のルール」と「測定」—— で定式化されている. このことは偶然で，システム理論と量子力学の二つの「測定」はまったく関係ないという意見もあるかもしれないが，我々はそうは考えない.

ここで我々は上の二つの「測定」を統合し，一つの抽象的な「測定理論」を提案する. 我々の提案は簡単で次のようなものである.「ボルンの測定理論」は通常ヒルベルト空間の言葉で書かれている．従って，これを，C^*-algebra の言葉で記述することは容易にできる (cf. §3)．この C^*-algebra 版の「ボルンの測定理論」において，特に可換 C^*-algebra の場合を「古典測定理論」と考える. 従って，いかなる（古典および量子）システムも，ある C^*-algebra $\mathcal{A}$ で定式化できて,

$$\text{「(古典および量子)システム理論」} = \text{「時間発展のルール」} + \text{「測定理論」} \tag{3}$$

[1] 1 英語版 page 1-18，日本語版 p19-35

KSTS/RR-02/013
December 27, 2002

の形で表現できると考える．すなわち,「(1) $\cup$ (2) $\subset$ (3)」である．もちろん, 重要なことは, 一般性ではなく,「システム理論（3)」からどれほどの知見が得られるかであり，これが我々の関心事である．

システム理論 (1) と量子力学 (2) はともにニュートン力学が進化したものと考えるとき，量子力学の発見は最大の賛辞をもって評価されている．これに比べて，システム理論 (1) の提案の評価は十分であるとは言い難い．しかしながら，ニュートン力学に「測定」という概念を導入してシステム理論 (1) を提案した先人達は量子力学の発見者達と同等の名誉に浴するべきと思う．このことは決して大げさに言っているつもりはない．古典測定の方が量子測定より見えにくいのである．事実，ニュートンすら見逃した（または， 軽視した ） ことなのだから．

さて，ここで「システム理論 ((1) またはより一般に (3)) は何か?」について我々の意見を述べておこう．システム理論（3）は力学（ニュートン力学と量子力学）を抽象化して作られたものであるから，我々は

$$\text{「システム理論 (3)」} = \text{「力学的世界観の数学的表現」} \tag{4}$$

と考える．すなわち，この世の（日常生活レベルの）すべての現象— 経済学，心理学，工学等—を (ある程度の無理を承知で多少強引に) 力学のアナロジーとして（ より正確にいうならば，力学の持つ数学構造のなかに押し込めて）理解しようとする認識の方法がシステム理論であると考える．これに対してはいろいろと意見があるかもしれないが，これほど強力な認識の方法が他にないことは誰しも異論ないことと思う (cf. §6).

以上をまとめて

(i) システム理論 (3) は「力学的世界観」の数学表現である．

(ii) 「力学的世界観」は唯一と言っていい程一般的かつ強力な科学的認識論である．

と考える．したがって，冒頭で述べた「測定なくして科学なし」もあながち大げさとも言えない.

[**注 1**] この世の現象を理解するするための数学的道具として

$$\text{微分方程式} \ + \text{確率} \tag{5}$$

と考え，これをシステム理論の出発点とするのも一理あると考えるかもしれない．しかし，システム理論を構築した先人達は (5) ではなく (1) を主張した．ここに彼らの英知が集約されていると思う．彼らの英知— なぜ（5:確率）ではなく（1: 測定）なのか? —を以下の節を通して再確認していただきたい.

§2. システム理論 (3) から得られる様々な結果

KSTS/RR-02/013
December 27, 2002

本論に入る前に，システム理論 (3) から得られる様々な結果について述べておく．次の主張はシステム理論 (3) の大枠をつかむのには重要である．

($\sharp_1$) 「測定」,「推論」,「制御」の 3 つは同値な概念である．したがって，統計学と制御理論は「測定理論」はの中でを統合されるべきである．

ストレートに言うならば,「測定」の言葉が便利であるが，同じことをすこしひねっていうときは「推論」,「制御」の言葉が使われる．言い換えると，工学（ または応用分野）では「逆問題」を扱うことが多いので,「推論」,「制御」という言葉がよく使われるが，これは「測定の逆問題」の中で現れる概念にすぎない (cf. 注 3.5).「計測と制御」という名前のジャーナルがあるくらいだから，この ($\sharp_1$) は常識かもしれないが，システム理論 (3) の中ではより明確になる．もちろん，問題の出どころが違えばその性質も異なるから，それを解析するテクニックも異なるのが普通で, 研究者のグループも異なっているという現状は当然のことである．上の ($\sharp_1$) は,「純理論的に言えば」という条件下の主張である.「システム理論（3)」さえ知っていれば，統計学, 制御理論, 古典力学，量子力学等は勉強する必要がないというのは原理的な話で，世の中はそんなに甘くはない．

ここで,「なぜシステム理論（3）が必要なのか?」について確認しておく．もちろん，(1)⊂(3) なのだから，システム理論（3）はシステム理論（1）より適用範囲が広い．このことは工学的に有利なことは当然である．しかし，最も基本的な「システム理論（3）の必要な理由」は以下のことであると我々は考えている．たとえば，次の 5 つの問題を考えよう．

(i). 三段論法を justification せよ．(＝ 三段論法が正しい方法であることを示せ）

(ii). Fisher 統計を justification せよ．

(iii). Kalman filter を justification せよ．

(iv). Zadeh の fuzzy sets 理論を justification せよ．

(v). 微分方程式を使って解析する方法は正しいことを示せ．

いずれも 「justification せよ」という形の問いかけであることに注意してもらいたい．このような「問いかけ」は日常的かもしれないが，はたしてそれ自体意味を持つのだろうか? この 5 つの問題は（肯定的結論でも否定的結論でもどちらでもいいが）結論が出ているのであろうか? 等々が我々の興味の中心である (cf. Refs[2∼13])．なんとなく勘違いしそうなこととして,

「justification すること 」 ＝「数学の言葉で書くこと」

というのは正しくない．どんないかがわしい理論も高級な数学を付加して, それを提示することはそんなに難しいことではないのだから．結局,「正しい方法」とは工学的な言い方で,「（適当な数学表現がされていて）役にたつ方法」ということなのだろうか? これは

KSTS/RR-02/013
December 27, 2002

かなり有力で，実際「統計学が正しい」と一般に言われるときは，このような基準で言われていると我々は思っている (我々の justification は §5 で述べる）．しかし，今まで「役にたたない」と思われていたことでも，周りの環境が変われば「役にたつ」ようになるかもしれない．このように考えると「justification とは何か?」(= 「正しい方法とは何か?」) に答えることは単純ではない．ここで，結論をいうと，

$$\text{「justification すること 」} = \text{「システム理論（3）の一つの側面と見なすこと」} \tag{6}$$

が我々の主張である．一般に，ある理論（または，方法）を「justification する 」とは，その理論（または，方法）を「最も大きな理論」の一つの側面（または，方法）と見なすことであると思う．システム理論（3）を大げさに「世界観」(または,「認識論」）と呼んだのはこのような理由からである．ともかく,「justification」とか「正しい」とかいう概念は，一つの「世界観」を固定してそのもとで意味をもつ概念であると信じている．したがって，もし「別の世界観」があるのならば当然「別の正しい」という概念が考えられる (cf. §6). システム理論（3）は，システム理論（1）の一般化という側面を持つだけではなく，認識論—- 正しい方法かそうでないかを分別する能力を有する理論—-としての側面を持つ．これがシステム理論（3）の最も基本的な存在価値であると考える．すなわち，我々は

$$\text{システム理論（3）} = \text{「科学の憲法のようなもの」} \tag{7}$$

と考えている．

システム理論（3）の一般的かつ原理的な側面（（4）, (6), (7)）をバックアップするためには必要なことは，哲学をすることではなく,「システム理論（3）からどのような知見がえられるか?」という問いかけに具体的にひとつづつ答えていくことである．以下に今までに得られた結果を列挙しておく．番号は参考文献の番号に対応している．どれも (6) (または (7)) の観点で読んでほしい．

[2]. ハイゼンベルグの不確定性原理を同時測定を用いて定式化し，これを証明した．この結果から EPR-paradox とハイゼンベルグの不確定性原理との関係がクリアーになった．

[3,4]. trajectory の概念を逐次測定の測定値の時系列と考え，量子粒子の trajectory を解析した．時に，ウィルソンの霧箱の意味を明確にした．trajectory を粒子が通過した跡と考えずに，測定値の時系列と考えることは古典システムでは特に重要となる (cf. [9]).

[5]. 量子システムにおいては，三段論法 ($A \Rightarrow B, B \Rightarrow C$ ならば $A \Rightarrow C$) は一般には成立しないが，古典システムにおいては三段論法は正しいことを測定理論の定理として証明した．また， $A \Rightarrow B, B \Rightarrow C$ ならば，ある意味で， $C \Rightarrow A$ というような奇妙なことも測定理論の定理として言える．このことは， fuzzy 推論と関連する．種々のロジックを測定理論の中の定理とみなすという考えは, ロジックを現実に適用する際の正当性を保証するという意味で本質的である．

[6]. この論文で,「システム理論 (3)」が提案された．測定のエントロピー等いろいろなアイディアが提案されていて意欲的であったが, 勇み足の部分もあった．以後の論文で徐々

KSTS/RR-02/013
December 27, 2002

に改善されていった．
[7]. 上の [5] の続編で fuzzy logic を測定理論の一定理と見なす立場で論じた．数学的にはいろいろな logic を提案されているが，その中でシステム理論 (3) の定理と見なせるものだけが,（力学的世界観の下で）現実的であるという立場を強く主張した．事実，我々は日常的に三段論法をよく使うが，これは三段論法がシステム理論 (3) の一定理と見なせるからである．
[8]. 因子分析，主成分分析を測定理論の立場から定式化した．測定理論の色づけがされているので，意味がクリアになっている．統計学一般に言えることであるが，ある手法を使うときに，それが適用範囲なのかそうでないのかが明確でない場合が多い．測定理論の言葉で定式化すればそれが明確になる．
[9]. システム理論 (3) はニュートン力学が進化したものと見て，システム理論 (3) の立場からニュートン力学を再構築した．ニュートン力学に測定理論を付け加えることで，状態の時間発展と trajectory とが全く異なる概念であることが明らかになった．またブラウン運動の測定理論的定式化を提案した．これらのためには，Kolmogorov の拡張定理 (確率空間 (=測定データの空間）の存在定理）の W^*-algebra 版の一般化定理（測定の存在定理）が基本になる．
[10]. 統計学の測定理論的定式化を提案した．ただ完全な形は［13］で明らかになった．
[11]. 統計力学を測定理論的立場から定式化した．すなわち，どのような雑な測定を仮定すれば，ニュートン力学から統計力学へ移行できるかを論じた．これは未完で続きを準備中である．
[12]. Kalman フィルタの測定理論的定式化を提案した．非常に一般的な条件の下でその存在が保証される．また［13］と合わせて読めば，Kalman フィルタの表現するものは確率分布ではなく，weight であることがわかる．また，フィッシャー流の Kalman フィルタとベイズ流の Kalman フィルタが双対の関係になることを強調した．
[13]. 統計学の測定理論的定式化 [10] の完全版である．これについては §5 で論じる．

以上であるが，システム理論（3）の中で大抵のことは定式化できて，それと同時に新しい知見が得られる. 測定理論の立場からの既存の方法（理論）の定式化の作業は継続しなければならない．この作業を継続する中で，そこで得られた知見から新しい分野が開花することを期待している．大抵のことが定式化できるののならば,「しっかりした手法と信じられているものの中で，定式化できないものは何か?」という問いかけは重要である．このような例を我々は知らないが，これについては徹底的に考える価値があると思う (cf. §6).

§3. 測定理論

この節と次節でシステム理論（3）の概略を述べる．$\mathcal{A}$ を C^*-algebra とする．ここでは，簡単のため単位元 I をもつとする．$\mathcal{A}^*$ をその dual Banach space とする．線形汎関数は $\rho(F)$

KSTS/RR-02/013
December 27, 2002

または ${}_{\mathcal{A}^*}\Big\langle \rho, F \Big\rangle_{\mathcal{A}}$ と書く ($F \in \mathcal{A}, \rho \in \mathcal{A}^*$). また，$\mathfrak{S}^m(\mathcal{A}^*) = \{\rho \in \mathcal{A}^* : \rho \geq 0, ||\rho||_{\mathcal{A}^*} = 1\}$ および $\mathfrak{S}^p(\mathcal{A}^*) =$ "$\mathfrak{S}^m(\mathcal{A}^*)$ の extreme points 全体" として，それぞれを混合状態と純粋状態の空間と呼ぶ．単に「状態」というときは純粋状態を意味するとする．

ここでは次の例 3.1 と 3.2 だけを確認すれば十分である.（最小限というならば例 3.1 だけでよい．）

[例 3.1] コンパクト集合 Ω 上の連続複素数値関数全体からなる集合を $C(\Omega)$ とすれば，これは可換 C^*-algebra となる．$C(\Omega)^* = \mathcal{M}(\Omega)$, すなわち，$\Omega$ 上の測度全体の集合．$\mathfrak{S}^m(C(\Omega)^*)$ $= \mathcal{M}^m_{+1}(\Omega)$, すなわち，$\Omega$ 上の正測度で全測度が 1 であるもの全体. $\mathfrak{S}^p(C(\Omega)^*) = \mathcal{M}^p(\Omega)$, すなわち，$\Omega$ 上のポイント測度 δ_ω ($\omega \in \Omega$) 全体. したっがて，$\mathfrak{S}^p(C(\Omega)^*) = \Omega$ と同一視できる．よって,「$\delta_\omega = \omega$」のように書く場合もある．

[例 3.2] n-次元ユニタリ空間 $\mathbf{C}^n$ 上の線形作用素全体（すなわち，$n \times n$ 複素数値行列全体）からなる集合を $B(\mathbf{C}^n)$ とすれば，これは非可換 C^*-algebra となる．$\mathfrak{S}^p(B(\mathbf{C}^n)^*)$ $= \{|x\rangle\langle x|$(Dirac の記号) $: x \in \mathbf{C}^n, ||x|| = 1\}$. したっがて，$\mathfrak{S}^p(B(\mathbf{C}^n)^*) = \{x \ : x \in \mathbf{C}^n, ||x|| = 1\}$ と同一視できる．

3 つ組 $\mathbf{O} \equiv (X, \mathcal{F}, F)$ が C^*-algebra $\mathcal{A}$ 内の observable（観測量，物理量とも言う）であるとは，つぎの (1)~(iii) を満たすことと定義する.（これは，Davies [1] の observable の定義の C^*-algebra 版の拡張である.）

(i). X は有限集合，$\mathcal{F}$ はベキ集合 2^X

(ii). $F : \mathcal{F} \to \mathcal{A}$ で, $F(\Xi) \geq 0$ ($\forall \Xi \in \mathcal{F}$). また $F(\emptyset) = 0$ と $F(X) = I$

(iii). 任意の $x_1, x_2 \in X$ ($x_1 \neq x_2$) に対して，$F(\{x_1, x_2\}) = F(\{x_1\}) + F(\{x_2\})$ が成立する．

2 つの observable $\mathbf{O}_1 \equiv (X_1, \mathcal{F}_1, F_1)$ と $\mathbf{O}_2 \equiv (X_2, \mathcal{F}_2, F_2)$ を考える．observable $\mathbf{O}_1 \times^q \mathbf{O}_2$ $\equiv (X_1 \times X_2, \mathcal{F}_1 \times \mathcal{F}_2, F_1 \times^q F_2)$ が $\mathbf{O}_1$ と $\mathbf{O}_2$ の quasi-product observable であるとは $(F_1 \times^q F_2)(\{x\} \times Y) = F_1(\{x\})$ ($\forall x \in X$) かつ $(F_1 \times^q F_2)(X \times \{y\}) = F_2(\{y\})$ ($\forall y \in Y$) を満たすときを言う．また，observable $\mathbf{O}_1 \times \mathbf{O}_2 \equiv (X_1 \times X_2, \mathcal{F}_1 \times \mathcal{F}_2, F_1 \times F_2)$ が $\mathbf{O}_1$ と $\mathbf{O}_2$ の product observable であるとは $(F_1 \times F_2)(\{x\} \times \{y\}) = F_1(\{x\})F_2(\{y\})$ ($\forall (x, y) \in X \times Y$) を満たすときを言う．

[注 3.3]. ここでは，後に必要な最小限のことしか述べていない．極限操作が C^*-algebra では不自由なので W^*-algebra を採用する場合もある.（ cf. [6,9]）.

いかなるシステム S に対しても，ある C^*-algebra $\mathcal{A}$ を適当に選べば，その中でそのシステム S は完全に記述できると考える．システム S の状態は，純粋状態 ρ^p ($\in \mathfrak{S}^p(\mathcal{A}^*)$) で表現される．また，observable（観測量，物理量とも言う）は observable $\mathbf{O} \equiv (X, \mathcal{F}, F)$ で表現される．状態 ρ^p ($\in \mathfrak{S}^p(\mathcal{A}^*)$) をもつシステム S に対する observable $\mathbf{O}$ の測定を $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[\rho^p]})$ と記す．もちろん，この測定により，測定値 x ($\in X$) を得る．また，quasi-product observable (or, product observable) の測定を同時測定または逐次測定と呼ぶ.

次の公理は,「ボルンの量子測定公理」の C^*-algebra 版にすぎない．Cf. [5,6].

KSTS/RR-02/013
December 27, 2002

AXIOM 1. 6 [測定公理]. 測定 $\mathbf{M}_{\mathcal{A}}(\mathbf{O} \equiv (X, \mathcal{F}, F), S_{[\rho^p]})$ を考える．このとき，測定値 $x\ (\in X)$ が $\Xi\ (\in \overline{\mathcal{F}})$ に属す確率は $\rho^p(F(\Xi))$ で与えられる．

C^*-algebra $\mathcal{A}$ が可換か非可換により，測定理論は

$$
\text{"システム理論（3）"} = \begin{cases} \text{"古典システム理論"} \\ \text{"量子 システム理論"} \end{cases} \tag{8}
$$

と分類される．

[**注 3.4**] 上の確率は,「individualistic」であるが，各 $\rho^p(\in \mathfrak{S}^p(\mathcal{A}^*))$ に対して，十分大きな N の繰り返し測定 $\bigotimes_{j=1}^N \mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[\rho^p]})$ を行うことにより，その頻度確率としてサンプル空間 $(X, \mathcal{F}, \rho^p(F(\cdot)))$ を得ることができる．(Theorem 2.2 in [6]).

[**注 3.5**]. 実際の測定 $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[\rho^p]})$ において，システムの状態 ρ^p は未知であることが普通である．すなわち,「状態 ρ^p を知りたいので，測定を行なう」と考えるのが普通の設定である.「未知である」ことを強調する意味で，$\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[\rho^p]})$ を $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[\,\cdot\,]})$ と書くこともある．統計学 (S) と制御理論 (C) は「測定」の逆問題で次のように定式化できる．

(S) 測定 $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[\,\cdot\,]})$ により測定値 $x(\in X)$ を得た．このとき状態 $[\,\cdot\,]$ を推定せよ．

(C) 測定 $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[\,\cdot\,]})$ により測定値 $x(\in X)$ を得ることができるように状態 $[\,\cdot\,]$ を設定せよ．

この問題 (S) と (C) の 1 つの解答は Fisher の最尤法で与えられる．しかし，最尤法は Axiom 1 で基礎づけされていることに注意されたい．(Cf. 5 節).

§4 時間発展のルール

$\mathcal{A}$ を C^*-algebra とする．連続線形作用素 $\Psi : \mathcal{A} \to \mathcal{A}$ は次の (i), (ii) を満たすときマルコフ作用素と呼ばれる．(i) $\Psi(F) \geq 0$ ($\forall\ F(\geq 0) \in \mathcal{A}$, (ii) $F(I_{\mathcal{A}}) = I_{\mathcal{A}}$, ここに $I_{\mathcal{A}}$ は $\mathcal{A}$ の単位元. $\mathbf{O} \equiv (X, \mathcal{F}, F)$ を $\mathcal{A}$ 内の observable とすると，$(X, \mathcal{F}, \Psi F)$ も observable となる．これを $\Psi\mathbf{O}$ と書く．また，マルコフ作用素 $\Psi : \mathcal{A} \to \mathcal{A}$ は次の (i), (ii) を満たすとき準同型作用素と呼ばれる．(i) $\Psi(F)\Psi(G) = \Psi(FG)$ ($\forall F, G \in \mathcal{A}$), (ii) $(\Psi(F))^* = \Psi(F^*)$ ($\forall F \in \mathcal{A}$). また，$\Psi^* : \mathcal{A}^* \to \mathcal{A}^*$ をマルコフ作用素 $\Psi : \mathcal{A} \to \mathcal{A}$ の共役作用素とする．このとき次が成り立つ．

(a) $\Psi^*(\mathfrak{S}^m(\mathcal{A}^*)) \subseteq \mathfrak{S}^m(\mathcal{A}^*)$,

(b) もし $\Psi : \mathcal{A} \to \mathcal{A}$ が準同型ならば，$\Psi^*(\mathfrak{S}^p(\mathcal{A}^*)) \subseteq \mathfrak{S}^p(\mathcal{A}^*)$ Cf. [15].

実数全体 $\mathbf{R}$ を時間軸と見る．時刻 $t = 0$ でのシステム S の状態（初期状態）を ρ^p $(\in \mathfrak{S}^p(\mathcal{A}^*))$ とする．集合 $\mathbf{R}^2_{\leq}$ を $\{(t_i, t_2) \in \mathbf{R}^2 : t_1 \leq t_2\}$ で定義する．組 $\mathbf{S}_{[\rho^p]} \equiv [S_{[\rho^p]}, \{\Psi_{t_1,t_2} : \mathcal{A} \to \mathcal{A}\}_{(t_1,t_2)\in\mathbf{R}^2_{\leq}}]$ は次の (i), (ii) を満たすとき初期システム $S_{[\rho^p]}$ をもつ C^*-力学システムと呼ばれる．

(i) $\rho^p \in \mathfrak{S}^p(\mathcal{A}^*)$,

(ii) すべての $(t_1, t_2) \in \mathbf{R}^2_{\leq}$ に対して，$\Psi_{t_1,t_2} : \mathcal{A} \to \mathcal{A}$ はマルコフ作用素，しかも $\Psi_{t_1,t_2}\Psi_{t_2,t_3} = \Psi_{t_1,t_3}$ $(\forall\, (t_1,t_2),(t_2,t_3) \in \mathbf{R}^2_{\leq})$ が成立する．

さて，ここでシステム理論（3）の時間発展のルールを次のように述べることができる．
AXIOM 2. [時間発展のルール (cf. [6,9])]．$S_{[\rho^p]}$ を初期状態 $\rho^p \in \mathfrak{S}^p(\mathcal{A}^*)$ をもつシステムとする．システム S の時間発展は C^*-力学システム $\mathbf{S}_{[\rho^p]} \equiv [S_{[\rho^p]}, \{\Psi_{t_1,t_2} : \mathcal{A} \to \mathcal{A}\}_{(t_1,t_2)\in\mathbf{R}^2_{\leq}}]$ で表現できる．また，ハイゼンベルグピクチャーにより，*Axiom 1* と C^*-力学システム $\mathbf{S}_{[\rho^p]}$ との関連をつけることができる．(*Cf.* 注 *4.1*).

古典力学および量子力学は可逆な決定過程であるから同型作用素からなる C^*-力学システムで記述できる．しかし，上ではより一般にマルコフ作用素からなる C^*-力学システムを考えた．

[**注 4.1**] Axiom 2 の中に書かれている「ハイゼンベルグピクチャー」について補足する．$0 \leq t_1 \leq t_2 \leq \cdots \leq t_n$ とする．各 t_k に対して，observable $\mathbf{O}_{t_k} \equiv (X_{t_k}, \mathcal{F}_{t_k}, F_{t_k})$ を考える．力学システム $\mathbf{S}_{[\rho^p]} \equiv [S_{[\rho^p]}, \{\Psi_{t_1,t_2} : \mathcal{A} \to \mathcal{A}\}_{(t_1,t_2)\in\mathbf{R}^2_{\leq}}]$ に対しての observable $\{\mathbf{O}_{t_k}\}_{k=1}^n$ の測定—— $\mathbf{M}_{\mathcal{A}}(\{\mathbf{O}_{t_k}\}_{k=1}^n, \mathbf{S}_{[\rho^p]})$ —- を考える．$\widetilde{\mathbf{O}}_{t_n} = \mathbf{O}_{t_n}$ と置く．observable $\Psi_{t_{n-1},t_n}\widetilde{\mathbf{O}}_{t_n} = (X_{t_n}, \mathcal{F}_{t_n}, \Psi_{t_{n-1},t_n}F_{t_n})$ は時刻 t_{n-1} の observable と見なすことができる．ハイゼンベルグピクチャー, により，$\Psi_{t_{n-1},t_n}\widetilde{\mathbf{O}}_{t_n}$ は $\mathbf{O}_{t_n}$ と同一視できる．ここで,（$\mathbf{O}_{t_{n-1}}$ と $\Psi_{t_{n-1},t_n}\widetilde{\mathbf{O}}_{t_n}$ との）quasi-product observable $\widetilde{\mathbf{O}}_{t_{n-1}} \equiv (X_{t_{n-1}} \times X_{t_n}, \mathcal{F}_{t_{n-1}} \times \mathcal{F}_{t_n}, F_{t_{n-1}} \overset{q}{\times} \Psi_{t_{n-1},t_n}F_{t_n})$ を考える．もちろん，$\widetilde{\mathbf{O}}_{t_{n-1}}$ の存在は量子システムでは一般には言えない．また古典システムにおいても一意性は保証されていない．同様に，$\mathbf{O}_{t_{n-2}}$ と $\Psi_{t_{n-2},t_{n-1}}\widetilde{\mathbf{O}}_{t_{n-1}}$ との quasi-observable $\widetilde{\mathbf{O}}_{t_{n-1}}$ を定義する．これを繰り返して最終的には，$\mathbf{O}_{t_1}$ と $\Psi_{t_1,t_2}\widetilde{\mathbf{O}}_{t_2}$ との quasi-observable $\widetilde{\mathbf{O}}_{t_1}$ を得る．$\widetilde{\mathbf{O}}_0 = \Psi_{0,t_1}\mathbf{O}_{t_1}$ と置いて，測定 $\mathbf{M}_{\mathcal{A}}(\widetilde{\mathbf{O}}_0, S_{[\rho^p]})$ を得る．これが $\mathbf{M}_{\mathcal{A}}(\{\mathbf{O}_{t_k}\}_{k=1}^n, \mathbf{S}_{[\rho^p]})$ の表現である．ただし，前にも述べたように，$\widetilde{\mathbf{O}}_0$ の存在と一意性がいつも保証されているわけではない．

ここまででシステム理論 (3) の概略を述べた．まとめると，以下のようになる．

$$
\text{「（古典および量子）システム理論」} = \underset{\text{(Axiom 2)}}{\text{「時間発展のルール」}} + \underset{\text{(Axiom 1)}}{\text{「測定理論」}} \tag{3}
$$

ここで，Axioms 1,2 は,「状態」,「observable」,「測定」,「測定値」,「確率」,「時間発展（システム間の関係，注 4.2 ）」という言葉の使い方を規定していると思えばよい．そして，これらの言葉だけを用いて,「世の中の現象すべて」を表現しようというのが，我々が主張する「力学的世界観」の精神である．したがって，もし他の言葉を用いるときは，その言葉はこれらの基本的な言葉を使って定義されていなければならない．

[**注 4.2**] 我々の議論の流れは次の図が示す．

KSTS/RR-02/013
December 27, 2002

$$
\begin{array}{ccc}
\text{量子力学 (2)} & \xrightarrow[\text{一般化 (a)}]{} & \text{システム理論 (3)} \\
(\mathcal{A}=B(H)) & & (\text{general } \mathcal{A}) \\
& & \downarrow \text{特殊化 } (b) \\
& & \text{古典システム理論 (8)} \\
& & (\text{commutative } \mathcal{A}) \\
& & \downarrow \text{特殊化 } (c) \\
\text{Newtonian 力学} & \xrightarrow[\text{一般化 (d)}]{} & \text{（旧来の）システム理論 (1)} \\
(\text{Newtonian eq.}) & & (\text{状態および測定方程式})
\end{array}
$$

第 1 節で強調したように，一般化 (d) は称賛されべきと考える．特殊化 (c) は [7] を見よ．したがって，(1) $\cup$ (2) $\subset$ (3) と言える．我々の仕事 [2∼13] の大部分は，古典 システム理論 (8) が（旧来の）システム理論 (1) よりも比較にならない程 rich であることを示すことであった.（旧来の）システム理論 (1) では到底（7）のような主張はできない．従って，システム理論 (3) (or, 古典 システム理論 (8)) がシステム理論の標準的な枠組として一般に認められることを願っている．

[**注 4.3**] 「微分方程式で解析する方法は正しい」ことは Axiom 2 で保証されていること注意されたい.(もちろん，どんな「正しい方法」でも使い方を間違えれば話しにならないことは当然である.）しかし，Axiom 2 はもうすこし一般形にしておいた方がよい．すなわち，時間軸は木構造をもつ半順序集合に，また各 t に依存した C^*-algebra $\mathcal{A}_t$ を考えるように一般化する (cf. [12])．したがって,「時間発展のルール」は「システム間の関係」となる．このようにしておけば回帰分析等も定式化できる．これ以上の一般化は不要と思っている．

[**注 4.4**] システム理論 (3) (特に「測定理論」）とコルモゴロフの確率論 [14] の関係ががが気になる向きもあるかもしれない．コルモゴロフの確率論は数学の一分野なのは当然のことである．システム理論 (3) は数学の言葉で記述されているので，それを解析するときいろいろな数学が使える．ヒルベルト空間や順列組合せの数学をつかう場合もあるだろうし，これと同じ意味でコルモゴロフの確率論を使う場合もある．測定理論とコルモゴロフの確率論の関係はこのようなもので, 決して対比するようなものでない．もちろん，システム理論 (3) の解析において，コルモゴロフの確率論は (微分方程式論と並んで) 最も役に立つ数学の 1 つであることは確かである．

[**注 4.5**] システム理論 (3) の量子システムの部分を強調する気持ちは毛頭ない．我々の興味はむしろ古典システムである．事実, §2 での結果の大部分は古典システムに関する結果である．しかし，古典システムを考えるときに量子システムのことがヒントになることがしばしばある．たとえば「測定なくして確率なし」は量子力学の常識であるが，これは古

典システムを考えるときにも重要なことである．従って，我々はたいていの場合，$C(\Omega)$ でなく，一般の C^*-algebra $\mathcal{A}$ で考えているが，出てくる結果は古典システムに関する結果という場合が多い．そもそもシステム理論 (3) は量子力学から動機づけられたのだから，また量子力学に回帰するのは賢明でないことが多い．古典システムが狙い目と思う．

§5 システム理論（3）の 1 つの応用——フィッシャー推定とベイズ推定の関係

ここではシステム理論（3）の 1 つの応用として，フィッシャー推定とベイズ推定の関係について述べる．§2 で述べた緒結果のどれを取り上げてもよかったが，誰もが一度は学習済みであるテーマ—- 統計学の基礎づけ—- であるということでこれ [13] を選んだ．もちろん，いつまでも（200 年間以上も）「ベイズ=主観」ではおかしいので，これをなんとかしようという気概もある．古典システムについて考えるのでこの節では一貫して $\mathcal{A} = C(\Omega)$ とする．また，ここでは Axiom 1 だけで Axiom 2 には関与しない．

$C(\Omega)$ 上のマルコフ作用素からなる集合 $\mathcal{S}_\Omega$ が次の (9) を満たすとき $\mathcal{S}_\Omega$ をシャフルと呼ぶ．

$$\underset{(\text{ シャフル } \mathcal{S}_\Omega \text{ が有効でない })}{\Phi(f) = f \quad (\forall \Phi \in \mathcal{S}_\Omega)} \Leftrightarrow \underset{(\text{ カードについての情報がない})}{f(\omega) = kI \ (\text{ i.e., constant })} \tag{7}$$

式（6）の下部分はトランプカードを例にした式（6）の解釈として読んでもらいたい．したがって正定数関数 kI $(\in C(\Omega))$ は情報 0 の状況を表していると考える．また，いくらかの情報を持っているという状況は非負関数 G $(\in C(\Omega))$ —- likelihood quantity —- で表されるとする．

シャフル $\mathcal{S}_\Omega$ を 1 つ固定する．システム $S_{[\,\cdot\,]}$ の likelihood quantity が正定数関数のとき（ すなわち，$S_{[\,\cdot\,]}$ の情報がまったくないとき）測定 $\mathbf{M}_{C(\Omega)}$ $(\mathbf{O} \equiv (X, 2^X, F),\, S_{[\delta_\omega]})$ を $\mathbf{M}_{C(\Omega)}(\mathbf{O},\, S_{[\,\cdot\,]}(kI)^{\mathcal{S}_\Omega}_{lq})$ と書く (cf. 注 3.5）．フィッシャーの likelihood function 方法によれば，

($\mathrm{F_C}$) 測定 $\mathbf{M}_{C(\Omega)}(\mathbf{O},\, S_{[\,\cdot\,]}(kI)^{\mathcal{S}_\Omega}_{lq})$ により測定値 x $(\in X)$ を得たとき，システム $S_{[\,\cdot\,]}$ の likelihood quantity は $k[F(\{x\})](\omega)$ となる．したがって，未知状態 $[\,\cdot\,]$ は $k[F(\{x\})](\omega_0)$ $= \max_{\omega \in \Omega}\, k[F(\{x\})](\omega)$ をみたす状態 $\omega_0(\in \Omega)$ とみなすことは妥当である．

念のため ($\mathrm{F_C}$) の理由づけをしておく．$[F(\{x\})](\omega_1) \ll [F(\{x\})](\omega_2)$ と仮定しよう．Axiom 1 によれば「測定 $\mathbf{M}_{C(\Omega)}(\mathbf{O},\, S_{[\delta_{\omega_1}]})$ により測定値 x $(\in X)$ を得る確率」$= [F(\{x\})](\omega_1)$ $\ll [F(\{x\})](\omega_2) =$ 「$\mathbf{M}_{C(\Omega)}(\mathbf{O},\, S_{[\delta_{\omega_2}]})$ により測定値 x $(\in X)$ を得る確率」である．したがって, 測定値 x を得たという事実からは，$[\,\cdot\,] = \omega_2$ と推定するほうが, $[\,\cdot\,] = \omega_1$ と推定するより妥当性がある．もし，$[\,\cdot\,] = \omega_1$ ならば非常に稀なことが起こったことになってしまう．

測定の前にシステムの状態についてまったく情報がないという場合（すなわち，上の（F_C）のような場合）はむしろ稀で，たいていの場合は多少の情報を測定者が持っていると仮定した方が自然である．以下このような場合を考える．

システム $S_{[\cdot]}$ の likelihood quantity が G_0 ($0 \leq G_0 \in C(\Omega)$) であることを知っているという前提のもとで，測定 $\mathbf{M}_{C(\Omega)}(\mathbf{O},\, S_{[\delta_\omega]})$ を $\mathbf{M}_{C(\Omega)}(\mathbf{O},\, S_{[\cdot]}(G_0)_{lq}^{\mathcal{S}_\Omega})$ と書くことにする．ここで次の問題を考える．

($\mathrm{P_G}$) 測定 $\mathbf{M}_{C(\Omega)}(\mathbf{O},\, S_{[\cdot]}(G_0)_{lq}^{\mathcal{S}_\Omega})$ により測定値 x ($\in X$) を得たという事実からどのような新しい likelihood quantity を得るか?

この問題は次と同値である．

($\mathrm{P'_G}$) 逐次測定 $\mathbf{M}_{C(\Omega)}(\mathbf{O}_0{\times}\mathbf{O}, S_{[\cdot]}(kI)_{lq}^{\mathcal{S}_\Omega})$ (ここに $\mathbf{O}_0 = (\{y_0, y_1\}, 2^{\{y_0,y_1\}}, G)$ は $G(\{y_0\}) = G_0$, $G(\{y_1\}) = I - G_0$ で定義する) により測定値 (y_0, x) ($\in \{y_0, y_1\} \times X$) が得られたという事実からいかに新しい likelihood quantity を得るか?

さて，($\mathrm{F_C}$) と同値関係 "($\mathrm{P_G}$) $\leftrightarrow$ ($\mathrm{P'_G}$)" から ($\mathrm{P'_G}$ は $\mathrm{F_C}$ で解けるので) 上の ($\mathrm{P_G}$) は次のように解ける．

[F] ((P_G) の解答): $\mathrm{P_G}$ の条件下で，likelihood quantity $G_0F(\{x\})$ を得る．変形写像 $R_{\mathbf{M}[\mathbf{O};\{x\}]} : C(\Omega) \to C(\Omega)$ を $R_{\mathbf{M}[\mathbf{O};\{x\}]}\ (G) = F(\{x\})G$ ($\forall G \in C(\Omega)$) で定義するならば，$R_{\mathbf{M}[\mathbf{O};\{x\}]}(G_0)$ を得るともいえる．(変形写像の一般形は [12] では本質的な役割をする．)

従って，我々はフィッシャーの方法とは「 "情報 0 ($= kI$)" + [F] 」と考える．すなわち，フィッシャーの方法とは,（測定で得られた）情報を likelihood quantity ($\in C(\Omega)$ かつ 非負 ）で記述する方法である．

次にベイズの方法に進む．結論を先に言うならば，ベイズの方法とは,（測定で得られた）情報を weight ($\in \mathcal{M}_{+1}^m(\Omega)$) で記述する方法である．このためには以下で述べるようにシャフルの性質を多少制限する必要があるが，結局は，二つの方法 (フィッシャーの方法とベイズの方法) は，同じものの別の表現と言える．したがって，二つの方法の間の翻訳法が重要になる．

シャフル $\mathcal{S}_\Omega^B$ (または $\mathcal{S}_\Omega^B(\nu)$) は次の（BS）を満たすとき Bayes シャフルと呼ぶ.

(BS) $\Phi^*(\nu) = \nu$ ($\forall \Phi \in \mathcal{S}_\Omega^B$) を満たす weight ν ($\in \mathcal{M}_{+1}^m(\Omega)$) が一意に存在して，しかも任意の開集合 U ($\subseteq \Omega$) に対して $\nu(U) > 0$ が成立する．

(ここに $\Phi^* : \mathcal{M}(\Omega) \to \mathcal{M}(\Omega)$ は Φ の共役作用素). 不変 weight ν は Bayes シャフル $\mathcal{S}_\Omega^B$ により壊されることはない．言い換えれば，それは Bayes シャフル $\mathcal{S}_\Omega^B$ の前後で不変である．したがって，不変 weight ν もまた情報 0 の状況を表現すると考える．したがって,

我々が Bayes シャフル $\mathcal{S}_\Omega^B(\nu)$ からスタートするならば，情報 0 は二つの表現—— likelihood quantity kI (フィッシャーの方法) と weight ν (ベイズの方法) —— をもつことになる．

不変 weight ν は Bayes シャフル $\mathcal{S}_\Omega^B$ に依存することは当然である．もし Ω が有限集合（離散距離を考える）で，全単射からなる Bayes シャフル $\mathcal{S}_\Omega^B$ (cf. 例 5.2 ）を考えるならば, 不変 weight は正規個数測度 $\nu_{_{|\Omega|}}(D) = |D|/|\Omega|$ $(\forall D \in 2^\Omega)$ となる．また，Ω が n-次元ユークリッド空間 $\mathbf{R}^n$ 内のコンパクト領域で，しかも "拡散シャフル" $\mathcal{S}_\Omega^B(\nu)$ （すなわち反射壁をもつ熱方程式の解を表す半群 [16] からなるシャフル) ならば，不変 weight ν は ルベーグ測度を正規化したものとなる．

一つの Bayes シャフル $\mathcal{S}_\Omega^B(\nu)$ を固定しておく．記号 $\mathbf{M}_{C(\Omega)}(\mathbf{O},\, S_{[\,\cdot\,]}(\nu)_w^{\mathcal{S}_\Omega^B})$ で，$\mathcal{S}_\Omega^B(\nu)$ に関して情報 0 という状況での測定 $\mathbf{M}_{C(\Omega)}(\mathbf{O} \equiv (X, 2^X, F),\, S_{[\delta_\omega]})$ を表す．また，より一般に，記号 $\mathbf{M}_{C(\Omega)}(\mathbf{O},\, S_{[\,\cdot\,]}(\rho_0^m)_w^{\mathcal{S}_\Omega^B}))$, ($\rho_0^m \in \mathcal{M}_{+1}^m(\Omega)$), で，情報 ρ_0^m という状況での測定 $\mathbf{M}_{C(\Omega)}$ $(\mathbf{O} \equiv (X, 2^X, F),\, S_{[\delta_\omega]})$ を表す.

さて，[F] の共役形として次を得る.

[B] 測定 $\mathbf{M}_{C(\Omega)}$ $(\mathbf{O},\, S_{[\,\cdot\,]}(\rho_0^m)_w^{\mathcal{S}_\Omega^B})$ により測定値 x を得たのちに，システム $S_{[\,\cdot\,]}$ の新しい weight は $\rho_{\rm new}^m$ $(\in \mathcal{M}_{+1}^m(\Omega))$ となる．ただしここに $\rho_{\rm new}^m$ $(B) = \frac{\int_B [F(\{x\})](\omega)\rho_0^m(d\omega)}{\int_\Omega [F(\{x\})](\omega)\rho_0^m(d\omega)}$ $(\forall B \in \mathcal{B}_\Omega$, Borel field).

ここで，我々はベイズの方法とは「 "情報 0 $(= \nu$)" + [B] 」と考える．

写像 $\overline{R}_{\mathbf{M}[\mathbf{O};\{x\}]}^{\,*} : \mathcal{M}_{+1}^m(\Omega) \to \mathcal{M}_{+1}^m(\Omega)$ を $\overline{R}_{\mathbf{M}[\mathbf{O};\{x\}]}^{\,*}(\rho^m) = \frac{R_{\mathbf{M}[\mathbf{O};\{x\}]}^{*}(\rho^m)}{\|R_{\mathbf{M}[\mathbf{O};\{x\}]}^{*}(\rho^m)\|_{\mathcal{M}(\Omega)}}$ ($\forall \rho^m \in \mathcal{M}_{+1}^m(\Omega)$) で定義する．(ここに $R_{\mathbf{M}[\mathbf{O};\{x\}]}^{*} : \mathcal{M}(\Omega) \to \mathcal{M}(\Omega)$ は $R_{\mathbf{M}[\mathbf{O};\{x\}]}$ の共役作用素, cf. [F]). このとき $\rho_{\rm new}^m = \overline{R}_{\mathbf{M}[\mathbf{O};\{x\}]}^{\,*}(\rho_0^m)$ となる．したがって，[B] は [F] の共役表現と見ることができる．しかし重要なのは以下に示す (フィッシャーの方法とベイズの方法の間の) 翻訳である

さて，"情報 0 " の概念がフィッシャーの方法とベイズの方法の間の翻訳を可能にすることを示そう．Bayes シャフル $\mathcal{S}_\Omega^B(\nu)$ はまたシャフルでもあるから，"情報 0 " は二つの表現—— kI と ν ——- をもつ．よって，以下のように測定 $\mathbf{M}_{C(\Omega)}$ $(\mathbf{O},\, S_{[\,\cdot\,]}(\rho_0^m)_w^{\mathcal{S}_\Omega^B})$ と測定 $\mathbf{M}_{C(\Omega)}$ $(\mathbf{O},\, S_{[\,\cdot\,]}(G_0)_{lq}^{\mathcal{S}_\Omega^B})$ を比較することができる．

$\nu(U) > 0$ ($\forall$ 開集合 U $(\subseteq \Omega)$) だから，全単射: [likelihood quantity G_0] $\underset{T_\nu}{\longleftrightarrow}$ [weight ρ_0^m] を $[T_\nu(G_0)](D) \equiv \rho_0^m(D) = \frac{\int_D G_0(\omega)\nu(d\omega)}{\int_\Omega G_0(\omega)\nu(d\omega)}$ $(\forall D \in \mathcal{B}_\Omega)$, または $T_\nu^{-1}(\rho_0^m) = \frac{d\rho_0^m}{d\nu} = \frac{G_0}{\int_\Omega G_0(\omega)\nu(d\omega)}$, (すなわち Radon-Nikodym 微分 ）で定めることができる．明らかに，$T_\nu(kI) = \nu$ が成り立つ，またより一般に，$T_\nu(R_{\mathbf{M}[\mathbf{O};\{x\}]}(G_0)) = \overline{R}_{\mathbf{M}[\mathbf{O};\{x\}]}^{\,*}(\rho_0^m)$ （if $T_\nu G_0 = \rho_0^m$) も明らかである．したがって，対応 (= 翻訳) T_ν は [F] と [B] の間で矛盾を生ずることなく両立する (Cf. 例 5.2). すなわち，Bayes シャフル (または "情報 0 " の概念) はフィッシャーの方法と ベイズの方法との間の翻訳 T_ν を可能にする．

もちろん，Bayes シャフルには条件 (BS) が加えられている分だけ，ベイズの方法の方がフィッシャーの方法より多くのことを主張できることは当然である．たとえば，ベイズの方法

KSTS/RR-02/013
December 27, 2002

では，測定 $\mathbf{M}_{C(\Omega)}\,(\mathbf{O}, S_{[\,\cdot\,]}(\nu)_w^{\mathcal{S}_\Omega^B})$ のエントロピーは $\sum_{x \in X} \int_\Omega [F(\{x\})](\omega) \log [F(\{x\})](\omega) \nu(d\omega)$ $- \sum_{x \in X} P(x) \log P(x)$ （ ここに $P(x) = \int_\Omega [F(\{x\})](\omega)\nu(d\omega)$）と定めることができる (cf. [3])．エントロピーとは情報のある性質を数量化したものが，これはフィッシャーの方法ではできない．またこの例はエントロピーの概念に「確率」が不可欠ではないということも示唆している．フィッシャーの方法は広く，ベイズの方法は深いわけだから，すべてが翻訳可能というわけではない．

不変 weight ν が自動的に確率解釈をもつと考えてはならない．念のため，次の注意をする．

[**注 5.1** (weight の確率解釈)]．もし $P[x;\, \mathbf{M}_{C(\Omega)}(\mathbf{O},\, S_{[\,\cdot\,]}(\rho_0^m)_w^{\mathcal{S}_\Omega^B})]$ —- 測定 $\mathbf{M}_{C(\Omega)}(\mathbf{O},$ $S_{[\,\cdot\,]}(\rho_0^m)_w^{\mathcal{S}_\Omega^B}))$ により測定値 x $(\in X)$ が得られる確率—- が ${}_{\mathcal{M}(\Omega)}\langle \rho_0^m, F(\{x\}) \rangle_{C(\Omega)}$ で与えられならば，各 $\mathbf{O}_1 \equiv (Y, 2^Y, G)$ と各 $y (\in Y)$ に対して，$P[y; \mathbf{M}_{C(\Omega)}(\mathbf{O}_1,\, S_{[\,\cdot\,]}(\rho_{\mathrm{new}}^m)_w^{\mathcal{S}_\Omega^B})]$ $= \frac{P[(x,y); \mathbf{M}_{C(\Omega)}\big(\mathbf{O} \times \mathbf{O}_1, S_{[\,\cdot\,]}(\rho_0^m)_w^{\mathcal{S}_\Omega^B})]}{P[x; \mathbf{M}_{C(\Omega)}\big(\mathbf{O}, S_{[\,\cdot\,]}(\rho_0^m)_w^{\mathcal{S}_\Omega^B})]} = {}_{\mathcal{M}(\Omega)}\langle \rho_{\mathrm{new}}^m, G(\{y\}) \rangle_{C(\Omega)}$ が成り立つ． したがって，もし $P[x;\, \mathbf{M}_{C(\Omega)}\,(\mathbf{O},\, S_{[\,\cdot\,]}(\nu)_w^{\mathcal{S}_\Omega^B})] = {}_{\mathcal{M}(\Omega)}\langle \nu, F(\{x\}) \rangle_{C(\Omega)}$ を仮定すれば（ すなわち，等重率を仮定すれば）[B] は確率解釈をもつ．(cf. 例 5.2)．しかしながら，等重率は “システム理論 (3)” の範囲外であることに注意されたい．これは我々の理論の力強さを表していると考える．適用範囲か否かを言明できることは理論として最も重要なことであると思うからである．

次の例は我々の主張の理解を助けると思う．

[**例 5.2.**] (壷の問題)．外見からは区別がつかない二つの壷 ω_r と ω_b がある．(すなわち，十分シャフルされているとする.）それぞれの値段を 1400 円 と 2100 円とする．壷 ω_r [resp. ω_b] には 8 個の赤玉 と 2 個の青玉 [resp. 4 個の赤玉 と 6 個の青玉] が入っているとする．二つの壷 ω_r と ω_b のうちどちらか一つを選ぶ．その壷の中から一つの玉をランダムにとりだしその色を確認してまたもとの壷に戻す．これを繰り返すとする．順序を考慮して，その 2 回の結果は (“赤”, “青’) であったとする．ここで以下の問題を考える．

[問 1] 二つの壷 ω_r と ω_b のうちどちらか選ばれたと推定せよ?

さらに続けてその壷から 玉と取りだしたとして，それが赤玉だったとする．

[問 2] この場合はどうか?

また

[問 3] この時点でのこの壷の値段を推定せよ．

最後に

[問 4] もし始めに公正なコイン投げで壷が選ばれた場合についても考察せよ．

KSTS/RR-02/013
December 27, 2002

[解答] 上の問題をフィッシャーの方法 [resp. ベイズの方法] で解析する．$\Omega = \{\omega_r, \omega_b\}$. $\mathbf{O} = (\{r,b\}, 2^{\{r,b\}}, F)$ と置く．ここに，$[F(\{r\})](\omega_r) = 0.8$, $[F(\{b\})](\omega_r) = 0.2$, $[F(\{r\})](\omega_b) = 0.4$, $[F(\{b\})](\omega_b) = 0.6$. である．また，Bayes シャフル $\mathcal{S}_\Omega^B$ を $\mathcal{S}_\Omega^B = \{\Phi_\phi | \ \phi : \Omega \to \Omega$ は 全単射 $\}$ とする．ここに $\Phi_\phi : C(\Omega) \to C(\Omega)$ は $[\Phi_\phi(f)](\omega) = f(\phi(\omega))$ $(\forall f \in C(\Omega), \forall \omega \in \Omega)$ で定める．(もしより強力なシャフル $\mathcal{S}_\Omega \equiv \{\Phi_\phi | \ \phi : \Omega \to \Omega$ は写像 $\}$ を仮定したならば，ベイズの方法は適用できないことに注意せよ．) 情報 0 の状況はフィッシャーの方法 [resp. ベイズの方法] では likelihood quantity kI $(k > 0)$ [resp. weight $\nu_{|2|}$ (i.e., $\nu_{|2|}(\{\omega_r\}) = \nu_{|2|}(\{\omega_b\}) = 1/2$)] で表せる．したがって，測定 $\mathbf{M}_{C(\Omega)}(\mathbf{O} \times \mathbf{O}, S_{[\,\cdot\,]}(kI)_{lq}^{\mathcal{S}_\Omega^B})$ [resp. $\mathbf{M}_{C(\Omega)}(\mathbf{O} \times \mathbf{O}, S_{[\,\cdot\,]}(\nu_{|2|})_w^{\mathcal{S}_\Omega^B})$] を考えればよい．測定値 (r,b) が得られたのであるから新しい likelihood quantity $G_{\rm new}$ と新しい weight $\rho_{\rm new}^m$ は次のようになる．

$$G_{\rm new}(\omega_r)(= kI \cdot [F(\{r\})](\omega_r) \cdot [F(\{b\})](\omega_r))= 0.16k, \quad G_{\rm new}(\omega_b) = 0.24k,$$

(直接の計算または翻訳: $T_{\nu_{|2|}}(G_{\rm new}) = \rho_{\rm new}^m$ から),

$$\rho_{\rm new}^m(\{\omega_r\})(= \frac{\int_{\{\omega_r\}}[F(\{r\})](\omega)[F(\{b\})](\omega)\nu_{|2|}(d\omega)}{\int_\Omega [F(\{r\})](\omega)[F(\{b\})](\omega)\nu_{|2|}(d\omega)})= \frac{16}{40} = \frac{2}{5}, \ \rho_{\rm new}^m(\{\omega_b\}) = \frac{3}{5}.$$

したがって問 1 の解答として，$[\,\cdot\,] = \omega_b$ (cf. フィッシャーの likelihood function 方法 $(\mathrm{F_C})$) と推定することは妥当である．

問 2 の場合，測定 $\mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[\,\cdot\,]}(G_{\rm new})_{lq}^{\mathcal{S}_\Omega^B})$ [resp. $\mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[\,\cdot\,]}(\rho_{\rm new}^m)_w^{\mathcal{S}_\Omega^B})$] を考えればよい．上と同様の計算から，$G_{\rm new^2}(\omega_r)$ ($= [G_{\rm new}](\omega_r) \cdot [F(\{r\})](\omega_r)$) $= 0.128k$, $G_{\rm new^2}(\omega_b) = 0.096k$, [resp. $\rho_{\rm new^2}^m(\{\omega_r\}) = \frac{4}{7}$, $\rho_{\rm new^2}^m(\{\omega_b\}) = \frac{3}{7}$] を得る．($G_{\rm new^2} \underset{T_{\nu_{|2|}}}{\longleftrightarrow} \rho_{\rm new^2}^m$ に注意せよ．)

さて，$p(\omega_r) = 1400$ と $p(\omega_b) = 2100$ と置いて，この時点でのその壷の平均的値段は $\int_\Omega p(\omega)\rho_{\rm new^2}^m(d\omega) = 1700$ (円) と計算できる (問 3)．(これはベイズの方法の結果であり，フィッシャーの方法では言えない)．ここまでの議論では,「確率」がインプリシットに扱われていることに注意されたい．もちろん，これが統計学の標準的な設定である．

しかし，問 4 の設定の場合には事態は一変して，ベイズの方法は「確率解釈」を獲得する．(cf. 注 5.1). したがって，上の平均的値段 1700 円は通常の期待値意味をもつ．また，たとえば (注 5.1 の記号を使って)，$P[r; \mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[\,\cdot\,]}(\rho_{\rm new}^m)_w^{\mathcal{S}_\Omega^B})] = {}_{\mathcal{M}(\Omega)}\langle \rho_{\rm new}^m, F(\{r\}) \rangle_{C(\Omega)} = 0.56$ と計算できる．念の為，この確率は等重率 (cf. 注 5.1) の結果であり，Axiom 1 の結果でないことに注意せよ．

§6. 結言

「力学的世界観」の数学表現として,

$$\text{「（古典および量子）システム理論」} = \underset{\text{(Axiom 2, cf. 注 4．3)}}{\text{「システム間の関係」}} + \underset{\text{(Axiom 1)}}{\text{「測定理論」}} \tag{3}$$

KSTS/RR-02/013
December 27, 2002

を提案した．すなわち，Axioms 1 と 2 は,「状態」,「observable」,「測定」,「測定値」,「確率」,「時間発展（システム間の関係，注 4.2)」という言葉の使い方を規定していて,「世の中の現象すべて」をこれらの言葉だけを用いて表現しようというのが，我々が主張する「力学的世界観」の精神である．この枠組は非常に一般的でかつ強力であると確信している．すなわち

$$\text{「ある理論を justification すること 」} = \text{「システム理論（3）の一つの側面と見なすこと」} \tag{6}$$

という意味で「システム理論（3)」は一つの世界観を表現していると考える．すなわち

$$\text{"システム理論（3）"} = \text{"力学的世界観の数学表現"} \tag{4}$$

$$= \text{"科学の憲法のようなもの"} \tag{7}$$

と考える．従って，(3) の中の「測定理論」は冒頭で述べた「測定なくして科学なし」の名言に恥じないと確信している．

また，システム理論 (3) の一つの側面として統計学の基礎づけを以下のように行なった．

$$\underset{\text{(Axioms 1,2)}}{\text{システム理論（3）}} \quad \underset{(\text{統計学})}{\Longrightarrow} \quad \begin{cases} \underset{\text{likelihood quantity } (\mathcal{S}_\Omega)}{\Longrightarrow} \quad \underset{(\text{“情報 0 } (= kI)\text{” + [F]})}{\text{フィッシャーの方法}} \\ \qquad\qquad \updownarrow \text{翻訳 } T_\nu (\text{if } \mathcal{S}_\Omega = \mathcal{S}_\Omega^B) \\ \underset{\text{weight } (\mathcal{S}_\Omega^B)}{\Longrightarrow} \quad \underset{(\text{“情報 0 } (= \nu)\text{” + [B]})}{\text{ベイズの方法}} \end{cases}$$

ここでフィッシャーの方法と ベイズの方法との間の翻訳 T_ν が特に重要である. (7）のような言い方をするならば，統計学は合憲であるといえる．

ここで我々の現時点での関心事について述べておく．もちろん，

- §2 で述べたことの継続として，システム理論（3）からどのような知見が引き出せるか一つづつ進めていく．実際の手続きとしては，いろいろな手法（特に，心理学や OR の手法は有望）が合憲であるか違憲であるかを一つづつ check することである．

はやらなければならない不可欠なことである．また，

- システム理論（3）—- すなわち「力学的世界観」—— 以外の科学的認識の方法があるのか?

にも関心がある. §2 の「システム理論の理学的側面 」で述べたように，いろいろな理論を「justification 」するとは，その理論を大きな理論（世界観）の一つの側面と見なすことであると，我々は考える．したがって，もし「ある理論」—— これはシステム理論（3）の一つの側面と見なせないと仮定しよう—— の正当性を誰かが主張したいならば，彼は別

の「大きな理論（世界観）」を提案して，この「大きな理論（世界観）」の有用性を十分検証して，その後にその「ある理論」が彼が提案した「大きな理論（世界観）」の一つの側面であることを示さなければならない．したがって，別の大きな理論（世界観）があるかどうかは非常に重要であると考える．

別の世界観の候補としては最近の Frieden の仕事 [17] は注目しておいたほうがいいかもしれない．システム理論 (3) は極めてオーソドックスなのでこれからも中心的な位置に居続けると思うが、このような理論がいろいろと出てきて競合相手になることは賑やかしくて好ましいことと思う。

References

[1] E. B. Davies, *Quantum theory of open systems*, Academic Press 1976

[2] S. Ishikawa, *Uncertainty relation in simultaneous measurements for arbitrary observables* , *Rep. Math. Phys.* **29**, 257-273 (1991)

[3] S. Ishikawa, *Uncertainties an an interpretation of nonrelativistic quantum theory* *Internat. J. Theoret. Phys.* , **30**, 401-417 (1991)

[4] S. Ishikawa, T. Arai, T. Kawai *Numerical Analysis of trajectories of a quantum particle in two-slit experiment.* *Internat. J. Theoret. Phys.* , **33**, 1265-1274 (1994)

[5] S. Ishikawa, *Fuzzy inferences by algebraic method*, *Fuzzy Sets and Systems* **87**, 181-200 (1997)

[6] S. Ishikawa, *A quantum mechanical approach to a fuzzy theory*, *Fuzzy Sets and Systems* **90**, 277-306 (1997)

[7] S. Ishikawa, *Fuzzy logic in measurements*, *Fuzzy Sets and Systems* **100**, 291-300 (1998)

[8] S. Ishikawa, A. Iida, *A system theoretical characterization of factor analysis*, EUFIT'98 , Vol. 2. 1257-1261 (1998), (September 7-10, Aachen)

[9] S. Ishikawa, T. Arai, T. Takamura, A dynamical system theoretical approach to Newtonian mechanics, *Far east journal of dynamical systems* **1**, 1-34 (1999)

[10] S. Ishikawa, *Statistics in measurements*, *Fuzzy Sets and Systems* **116**, 141-154 (2000)

[11] S. Ishikawa A dynamical system theoretical approach to equilibrium statistical mechanics *Far east journal of dynamical systems* **3**, 9-28 (2001)

[12] S. Ishikawa The existence theorem of Kalman filter in measurement theory *Far east journal of dynamical systems* **3**, 175-204 (2001)

[13] S. Ishikawa, A system theoretical characterization of the relation between Fisher's and Bayes' methods , *Far east journal of theoretical statistics* **7**, 19-33 (2002)

KSTS/RR-02/013
December 27, 2002

[14] A. Kolmogorov, *Foundations of probability* (*translation*), Chelsea Publishing Co. 1950

[15] S. Sakai, C^**-algebras and* W^**-algebras*, Ergebnisse der Mathematik und ihrer Grenzgebiete (Band 60), Springer-Verlag, Berlin, Heidelberg, New York 1971

[16] K. Yosida, *Functional analysis*, Springer-Verlag (Sixth Edition) 1980

[17] B. R. Frieden *Fisher information in physics* , Cambridge University Press (1998)

注. S. Ishikawa の論文の詳細は http://www.math.keio.ac.jp/˜ ishikawa を見よ.