Progress of Theoretical Physics, Vol. 64, No. 3, September 1980

# A Remark on Machida-Namiki Theory of Measurement

Huzihiro ARAKI

*Research Institute for Mathematical Sciences*
*Kyoto University, Kyoto 606*

(Received April 11, 1980)

An essential feature of Machida-Namiki theory of measurement is mathematically formulated in terms of continuous superselection rules. A general theory of probabilistic description of states is presented as its background.

## § 1. Introduction

Recently Machida and Namiki[1] proposed a theory of measurement in quantum mechanics, in which a continuous superselection rule plays an important role for the reduction of wave packets. The purpose of this paper is to provide an abstract background for such a theory. We do not mean by this to develop the most general framework. Rather we try to abstract an essential feature of Machida-Namiki argument and at the same time provide a brief explanation about some aspects of continuous superselection rules for physicists who might not be familiar with some of mathematical background.

For physicists, a pure state in quantum mechanics usually means a (unit) vector in a Hilbert space and a mixture state means a density matrix, i.e., a (trace-class) positive operator of trace 1. However, this corresponds to the simplest case in the mathematical sense where all bounded linear operators on the Hilbert space under consideration are somehow related to observables. A more general situation arises if we restrict observables to operators belonging to a smaller algebra $\mathfrak{A}$ (the simplest case being $\mathfrak{A}$ = the set of all bounded linear operators).

In this paper, we consider still a rather limited case, the only generalization being the presence of the center of the algebra of observables, which represents something like classical observables (corresponding to superselection rules). Such a framework will be explained in § 2. A simple mechanism for the reduction of wave packets will be explained in § 3 in terms of abstract examples. We believe that these examples represent mathematically an essential feature of Machida-Namiki argument.

The superselection rules in the above model could be considered corresponding to classical observables of the measurement apparatus rather than something of fundamental nature. In the Appendix we shall briefly describe a general framework which sheds some light on the notion of classical observables.

## § 2. Continuous superselection rules

Let $L$ be a direct integral of Hilbert space $L_\lambda$ with respect to an absolutely continuous measure $d\mu(\lambda)$, namely a vector $f$ in $L$ is a function of a real variable $\lambda$ with its value $f(\lambda)$ belonging to $L_\lambda$ and the inner product of two vectors $f_1$ and $f_2$ is given by

$$(f_1, f_2) = \int (f_1(\lambda), f_2(\lambda))_\lambda d\mu(\lambda), \tag{2·1}$$

where $(\ ,\ )_\lambda$ denotes the inner product in $L_\lambda$. Such a vector will be denoted as $f = \int f(\lambda)\, d\mu(\lambda)$.

The purpose of introducing such a structure is to restrict operators $A$ which are to represent observables to those having the following form:

$$(Af)(\lambda) = A(\lambda) f(\lambda), \tag{2·2}$$

where $A(\lambda)$ for each $\lambda$ is a linear operator on $L_\lambda$. Such $A$ is denoted by

$$A = \int A(\lambda)\, d\mu(\lambda). \tag{2·3}$$

In mathematics, a state is formulated as an (expectation) functional giving to each observable operator $A$ a number $\rho(A)$ called its expectation value in such a way that this functional is linear and positive for any positive operator (i.e., $\rho(A^*A) \geqq 0$). If we restrict $A$ as above to operator of the form (2·3), then

$$\rho(A) = \int \mathrm{tr}(\rho(\lambda) A(\lambda))\, d\mu(\lambda) \tag{2·4}$$

is a bona-fide state if $\rho(\lambda)$ for each $\lambda$ is a density matrix in the usual sense, up to a multiplicative factor and $\int \mathrm{tr}\, \rho(\lambda)\, d\mu(\lambda) = 1$. It should be noted that such a state cannot necessarily be described by a density matrix $\rho$ on the total space $L$ in the usual form $\rho(A) = \mathrm{tr}(\rho A)$ if we allow all bounded operators $A$ of the form (2·3) and if $d\mu(\lambda)$ is an absolutely continuous measure. (This is implicitly implied by the example of the next section, which produces a phenomena forbidden for states given by a density matrix due to Wigner's argument[2),3)].)

## § 3. Reduction of wave packets

*Example 1* The Hilbert space $K$ for the measured system is assumed to be of $n$ dimension ($n < \infty$) with an orthonormal vectors $u_j$ ($j = 1, \cdots, n$) which we want to separate by a measurement. Let $L$ described in the preceding section be the Hilbert space for the measuring apparatus. An arbitrary initial pure state of the measured system is represented by

$$E(\psi) = |\psi\rangle\langle\psi|, \quad \psi = \sum_j c_j u_j. \tag{3·1}$$

Let an initial state of the measuring apparatus be $\rho$ given by (2·4). The combined system is represented in the tensor product $K\otimes L$. Let the Hamiltonian $H$ of the combined system be such that

$$H\left(\psi\otimes\int f(\lambda)\,d\mu(\lambda)\right)=(H_1\psi)\otimes\int\lambda f(\lambda)\,d\mu(\lambda) \tag{3·2}$$

(namely, operator $H_1$ on the space $K$ and the multiplication of $\lambda$ on the space $L$) with $H_1u_j=\lambda_ju_j$ and all $\lambda_j$ distinct. The state $\alpha_t$ of the combined system at time $t$ is then

$$\alpha_t(A)=\int\mathrm{tr}\{(E(\psi)\otimes\rho(\lambda))\,e^{i\lambda Ht}A(\lambda)\,e^{-i\lambda Ht}\}\,d\mu(\lambda), \tag{3·3}$$

where $A(\lambda)$ is now an operator on $K\otimes L_\lambda$. By (3·2), we have

$$\begin{aligned}\alpha_t(A)&=\sum_{j,k}c_jc_k^*\int e^{i\lambda(\lambda_k-\lambda_j)t}\,\mathrm{tr}\{(|u_j\rangle\langle u_k|\otimes\rho(\lambda))\,A(\lambda)\}\,d\mu(\lambda)\\&\to\sum_j|c_j|^2\int\mathrm{tr}\{(E(u_j)\otimes\rho(\lambda))\,A(\lambda)\}\,d\mu(\lambda)\end{aligned} \tag{3·4}$$

as $t\to\infty$. The last assertion about the limit is due to the Riemann-Lebesgue Lemma, which says

$$\int e^{i\lambda t}F(\lambda)\,d\mu(\lambda)\to0\quad\text{as}\quad t\to\infty \tag{3·5}$$

for any $d\mu$-integrable function $F$ and absolutely continuous $d\mu(\lambda)$.

This shows that a wave packet $\psi$ of the measured system tends to a mixture $\sum_j|c_j|^2E(u_j)$ as $t\to\infty$ through a contact with the measuring apparatus.

*Example 2* In the above example, the measuring apparatus is not affected by the interaction. (Namely it works as a separation apparatus but not as a measuring apparatus.) By a slight modification, we can obtain a change of the measuring apparatus, which reflects the state of the measured system.

Instead of $L$, we consider $L\otimes L'$ as the Hilbert space for the measuring apparatus, where $L$ is as before and $L'$ is of finite dimension for simplicity. To the previous $H$ we add

$$H'=\sum_jE(u_j)\otimes1\otimes H_j\,, \tag{3·6}$$

where $H_j$ is some selfadjoint operator on $L'$ depending on $j$. If the initial state of the combined system is

$$\alpha_0(A)=\int\mathrm{tr}((E(\psi)\otimes\rho(\lambda)\otimes\sigma)\,A(\lambda))\,d\mu(\lambda) \tag{3·7}$$

for some density matrix $\sigma$ on $H'$, then

$$\alpha_t(A) = \alpha_0(e^{i(H+H')t} A\, e^{-i(H+H')t})$$
$$= \sum_{j,k} c_j c_k^* \int e^{i\lambda(\lambda_k-\lambda_j)t} \operatorname{tr}\{(|u_j\rangle\langle u_k| \otimes \rho(\lambda) \otimes e^{-iH_jt}\sigma e^{iH_kt})\, A(\lambda)\} d\mu(\lambda)$$
$$\sim \sum_j |c_j|^2 \operatorname{tr}\{(E(u_j) \otimes \rho(\lambda) \otimes \sigma_j(t))\, A(\lambda)\} d\mu(\lambda) \qquad (3\cdot 8)$$

as $t \to \infty$ where $\sim$ denotes vanishing of the difference of two sides in the limit and

$$\sigma_j(t) = e^{-iH_jt}\sigma e^{iH_jt} \qquad (3\cdot 9)$$

reflects the state of the measured system. The proof is the same as before if we rewrite the trace using complete sets of eigenvectors of $H_j$ and $H_k$ as follows:

$$\sum_{l,m} \operatorname{tr}\{(|u_j\rangle\langle u_k| \otimes \rho(\lambda)\, |v_l\rangle\langle w_m|)\, A(\lambda)\}\, e^{it(h_l - h_m')} (v_l, \sigma w_m), \qquad (3\cdot 10)$$

where $H_j v_l = h_l v_l$ and $H_k w_m = h_m' w_m$.

## Appendix

### ——*Classical Observables*——

In the situation described in § 2, an operator $A$ for which $A(\lambda)$ is a multiple of the identity operator is classical observables in the sense described below in a more general setting of probabilistic description not necessarily restricted to quantum mechanics.

**A.1.**

In any measurement, there are the measuring apparatus by which the measurement is done and the measured object on which the measurement is done. These two will interact with each other during a certain time interval and then the apparatus will show the result of the measurement. In order to distinguish different results of measurement, we usually label them by real numbers, called the measured values.

The measured object is initially prepared by a certain definite procedure and if the preparing procedure is different, the measured object so prepared will in general be in a different condition. We use the letters $\alpha_1$, $\alpha_2$, etc. to distinguish various conditions of the measured object. On the other hand, different measuring apparatus will in general measure different property of the measured object and the letters $Q_1$, $Q_2$, etc. will be used to distinguish various properties measured. The set of all possible $\alpha$ will be denoted by $\mathcal{\Sigma}$ and the set of all possible $Q$ will be denoted by $\mathcal{O}$. From an operational point of view, $\alpha$ in $\mathcal{\Sigma}$ is specified by the exact instruction of the procedure by which the measured object is prepared and $Q$ in $\mathcal{O}$ is specified by the exact instruction as to the design and operation of the measuring apparatus.

From experiments on microscopic objects, one learns that the result of a measurement, when repeated, fluctuates between various values in general, even if $\alpha$ and $Q$ are fixed, and that a simple physical law can be formulated only in the statistical sense. Thus a meaningful quantity which can be obtained from experiments is the probability $w_\alpha^Q(q)$ of obtaining the measured value $q$ when the measurement of $Q$ is done on $\alpha$. To obtain $w_\alpha^Q(q)$, one repeats the measurement of $Q$ on many samples in the same condition $\alpha$ (namely, samples prepared by a fixed procedure) and, if the measured value $q$ is obtained $n$ times out of $N$ runs of such a measurement, $n/N$ is the experimental value for $w_\alpha^Q(q)$. The physical law is then some rule which correlates $w_\alpha^Q(q)$ for various $\alpha$, $Q$ and $q$.

Because of the statistical nature just described, the terminology *ensemble* will be used for each $\alpha$. It is also called a state. Each $Q$ will be called an *observable*. Thus $w_\alpha^Q(q)$ is the probability of obtaining the measured value $q$ when the observable $Q$ is measured on the ensemble $\alpha$. In order that $w_\alpha^Q(q)$ is a probability, we have the requirement

$$w_\alpha^Q(q) \geqq 0, \qquad \sum_q w_\alpha(q) = 1 . \tag{A·1}$$

*Example* Consider a measurement of a proton by a Geiger counter. The measured object is the proton and the measuring apparatus is the Geiger counter. If the proton comes from an accelerator, the exact structure of the accelerator as well as its position and the time of operation will specify $\alpha$. The structure of the Geiger counter as well as its position and the time of operation will specify $Q$. In a simple case, the Geiger counter will measure whether the proton passes a certain space region $B$ where the Geiger counter is placed, during the time interval $\tau = (t_1, t_2)$ when it is being operated. The result of the measurement may be distinguished by one of the two real numbers 1 and 0 according as the counter clicks or not. Then $w_\alpha^Q(1)$ will be the probability that the proton is in the region $B$ during the time interval $\tau$. On the other hand, $w_\alpha^Q(0)$ will be the probability that the proton is not in the region $B$ during the time interval $\tau$ and we have $w_\alpha^Q(0) + w_\alpha^Q(1) = 1$. The $w_\alpha^Q(q)$ for $q \neq 0$ or 1, is taken to be 0.

We may measure $w_\alpha^Q(1)$ for various $Q$ corresponding to various regions and time intervals. The totality of such information will give a good picture of the proton under consideration. (Needless to say, all our examples of experiments will be vastly simplified and idealized compared to actual ones.)

**A.2.**

At this point, we might have the following question: $w_\alpha^Q(q)$ is a probability which can be measured only if the measurement of the same observable $Q$ is repeated on many samples of a fixed ensemble $\alpha$. However the specification of the preparing apparatus of the ensemble includes its position and time of operation and hence there is only one sample available for each ensemble. How can one measure the probability $w_\alpha^Q(q)$ with only one sample?

The well-known remedy for this is to use a large number of equivalent

measurements. Namely one prepares a proton by operating exactly the same accelerator at many different times $t$, $t+\Delta_1$, $t+\Delta_2$, etc. and makes the measurement by operating exactly the same Geiger counter at correspondingly displaced time intervals $\tau=(t_1, t_2)$, $\tau+\Delta_1=(t_1+\Delta_1, t_2+\Delta_1)$, $\tau+\Delta_2=(t_1+\Delta_2, t_2+\Delta_2)$, etc. (In practice, the position and direction of the whole experimental set-up is also displaced due to the movement of the earth.) One assumes that these experiments at different times (and at different places) are all equivalent and one uses them to obtain the measured value of $w_\alpha^Q(q)$. (This is the assumption of the invariance of the theory under the time and space displacement.)

It should be noted that the Geiger counter operated at $\tau+\Delta$ defines an observable $Q'$ which is different from the observable $Q$ corresponding to the same Geiger counter operated at $\tau$. Similarly the ensemble $\alpha'$ prepared by the accelerator operated at $t+\Delta$ is different from the original $\alpha$ corresponding to the accelerator operated at the time $t$. The invariance assumption may be formulated as

$$w_\alpha^Q(q)=w_{\alpha'}^{Q'}(q) \qquad \text{for all } q. \tag{A·2}$$

Hence, in practice, we measure this probability by performing one run of experiment for each pair $(\alpha, Q)$, $(\alpha', Q')$, etc.

If one wants to measure two different observables $Q_1$ and $Q_2$ on the same ensemble $\alpha$, one may first perform the measurement of $Q_1$ using a set of equivalent meaeurements for pairs $(\alpha, Q_1)$, $(\alpha', Q_1')$, etc., and then perform the measurement of $Q_2$ using a new set of equivalent measurements for pairs $(\alpha'', Q_2'')$, $(\alpha''', Q_2''')$, etc. This, then, enables one to obtain the probabilities $w_\alpha^{Q_1}(q)$ and $w_\alpha^{Q_2}(q)$ for the same $\alpha$ even if the measurements of $Q_1$ and $Q_2$ on $\alpha$ interfere with each other.

**A.3.**

In some ideal case, it may happen that two seemingly different measuring apparatuses measure exactly the same property of the measured object. In such a case we will find two distinct measuring apparatuses $Q_1$ and $Q_2$ satisfying

$$w_\alpha^{Q_1}(q)=w_\alpha^{Q_2}(q) \tag{A·3}$$

for all ensembles $\alpha$ and for all values $q$. If (A·3) holds, we identify $Q_1$ and $Q_2$:

$$Q_1 \equiv Q_2. \tag{A·4}$$

Mathematically, we will consider the equivalence class in $\mathcal{O}$ defined by (A·4).

In a similar manner, if

$$w_{\alpha_1}^Q(q)=w_{\alpha_2}^Q(q) \tag{A·5}$$

is satisfied for all $Q$ and $q$, we define

$$\alpha_1 \equiv \alpha_2. \tag{A·6}$$

**A.4.**

The set of all real numbers $q$ such that $w_\alpha^Q(q) \neq 0$ for some ensemble $\alpha$ is called the *spectrum* of $Q$ and denoted by $\mathrm{sp}\,Q$. In other words, the spectrum of $Q$ is the set of all possible measured values of $Q$.

If the $\mathrm{sp}\,Q$ has a bound (i.e., if any $q \in \mathrm{sp}\,Q$ satisfy $|q| < A$ for some $A$), the observable $Q$ is called *bounded*. (We are restricting ourselves to observables with *discrete spectrum*.)

**A.5.**

In the previous example, we have assigned the measured value 1 or 0 according to whether the counter clicks or not. The description of the measuring instrument together with this specific assignment of real numbers to each possible result of the measurement defines the observable $Q$. We may assign some other values, say 6 and 5, instead of 1 and 0. In that case, we have another observable $Q'$, although the measuring instrument for $Q'$ is the same as that for $Q$. The probability functions for $Q$ and $Q'$ are related to each other by $w_\alpha^{Q'}(q+5) = w_\alpha^Q(q)$ and we say that $Q' = Q+5$. In general, if $f(x)$ is a function of the real variable $x$ ($f(x)$ should be defined for all $x$ in $\mathrm{sp}\,Q$ but need not be defined for other $x$), then the function $Q' = f(Q)$ of the observable $Q$ is defined as an observable $Q'$ satisfying

$$w_\alpha^{Q'}(q') = \sum_{q \in f^{-1}(q')} w_\alpha^Q(q). \tag{A·7}$$

Here $f^{-1}(q')$ is the set of all $q$ such that $q' = f(q)$. There may be many such $q$ and the summation is taken over all such $q$. If there is no such $q$, $w_\alpha^{Q'}(q')$ is defined to be 0.

If $q' = f(q)$ has at most one solution (namely, if $f^{-1}$ is singlevalued on the spectrum of $Q$), $Q$ may be considered as a function of $f(Q)$ and both $Q$ and $f(Q)$ are equally good for the description of the results of the measurement under consideration. In other words, with a given measuring apparatus, one may associate an infinite number of observables which are functions of each other.

In many cases, some specific structure of the theory (or the physical law of the system) enables one to single out a specific $Q$ as a canonical choice. For example, it is most convenient to consider the electric charge $Q$ of a particle instead of, say, its cubic power $Q^3$, because $Q$ satisfies an "additive conservation law", while $Q^3$ does not.

**A.6.**

An expectation value (or an average) of an observable $Q$ in an ensemble $\alpha$ is defined, as in any probability theory, by

$$\mathrm{Exp}(Q/\alpha) = \sum_q q w_\alpha^Q(q). \tag{A·8}$$

If $Q$ is bounded, the sum always converges. If we are allowed to use functions, then we can compute back the probability $w$ from the expectation values. Namely,

using the characteristic function

$$\chi_q(x)=\begin{cases}1 & \text{if } x=q,\\ 0 & \text{if } x\neq q,\end{cases} \tag{A·9}$$

we have

$$w_\alpha^Q(q)=\mathrm{Exp}\,(\chi_q(Q)/\alpha). \tag{A·10}$$

An observable, whose spectrum consists of at most two points 0 and 1 like $\chi_q(Q)$, is called a question. In particular $\chi_q(Q)$ is called a spectral question of $Q$.

**A.7.**

We now want to introduce the notion of a mixture and a pure state. Suppose one makes various measurements on a fixed ensemble of a positively charged particle coming from a fixed accelerator. Suppose that the charged particle in question happens to be a proton 7 times out of 10 and a $\pi^+$-meson 3 times out of 10, on the average. If we may use a detector which tells us whether the particle is a proton or a $\pi^+$-meson without disturbing the various measurements in question, we can select either only protons or only $\pi^\pm$-mesons. Let $\alpha_P$ and $\alpha_\pi$ denote the ensembles of a charged particle so prepared with the combination of an accelerator and a selector. We then expect the following equality for all $Q$ and $q$:

$$w_\alpha^Q(q)=0.7w_{\alpha_P}^Q(q)+0.3w_{\alpha_\pi}^Q(q). \tag{A·11}$$

In general, if ensembles $\alpha$, $\alpha_1$ and $\alpha_2$ satisfy

$$w_\alpha^Q(q)=\lambda w_{\alpha_1}^Q(q)+(1-\lambda)\,w_{\alpha_2}^Q(q) \tag{A·12}$$

for all $Q\in\mathcal{O}$ and $q$ and for a fixed real number $\lambda$ between 0 and 1, we call $\alpha$ a *mixture of* $\alpha_1$ *and* $\alpha_2$ with the weights $\lambda$ and $1-\lambda$ and denote

$$\alpha=\lambda\alpha_1\oplus(1-\lambda)\,\alpha_2. \tag{A·13}$$

For this definition, we do not care whether the preparing processes for three ensembles $\alpha$, $\alpha_1$ and $\alpha_2$ are somehow related, as in the above example, or are totally unrelated to each other. If $\lambda\neq 0$, $\alpha_1$ is called a *subensemble* of $\alpha$.

For given two ensembles $\alpha_1$ and $\alpha_2$, the mixture ensemble $\lambda\alpha_1\oplus(1-\lambda)\,\alpha_2$ represents their statistical average and can always be prepared (rather artificially) by some mechanism (maybe the experimentalist himself) which chooses one out of the two preparing apparatuses for $\alpha_1$ and $\alpha_2$, in a random fashion with the weights $\lambda$ and $(1-\lambda)$. Therefore we assume that *if* $\alpha_1$ *and* $\alpha_2$ *are in* $\Sigma$, *then* $\lambda\alpha_1\oplus(1-\lambda)\,\alpha_2$ *is also in* $\Sigma$ *for any* $\lambda$ *between* 0 *and* 1.

On the other hand, the converse procedure (called the purifying procedure) is not so trivial. An ensemble $\alpha$ is called a *mixture* if there exists *some* $\alpha_1$ and $\alpha_2$ in $\Sigma$, $\neq\alpha$ and a $\lambda$ between 0 and 1 satisfying (1·12). Otherwise, $\alpha$ is called a *pure state*. The set of all pure states will be denoted by $\Sigma_p$.

The statistical description in terms of the probabilities $w_\alpha^Q(q)$ is not restricted to quantum physics but can be applied to classical physics. The distinctive feature of the classical physics is that the measurement of any observable $Q$ on a pure state $\alpha$ always yields a definite answer and hence the probability function $w_\alpha^Q(q)$ is 0 for all but one values of $q$, if $\alpha$ is a pure state. Namely, any statistical feature in classical physics is represented by a mixture and if one restricts one's attention to pure states, one has a deterministic physical law for classical physics. Thus the notion of a mixture and a pure state is important for a distinction of the statistical feature already existent in classical physics and the statistical feature intrinsic in non-classical physics.

**A.8.**

It sometimes happens that some observable $Q$ may behave classically even in quantum theory, namely $w_\alpha^Q(q)$ will be 0 for all but one values of $q$ if $\alpha$ is a pure state. Such an observable $Q$ will be called a *classical observable.*

For a given observable $Q$, there may exist an ensemble $\alpha$ for which $w_\alpha^Q(q)$ is 0 for all $q$ except for $q=q_0$. In other words, the measurement of an observable $Q$ on a particular ensemble $\alpha$ may always give a definite measured value $q_0$. Such $\alpha$ is called an *eigenstate* (or eigenensemble) of $Q$ belonging to the *eigenvalue* $q_0$.

If $Q$ is a classical observable, every pure state $\alpha$ is an eigenstate of $Q$.

**A.9.**

The above definition of classical observables is meaningful only if there exist sufficiently many pure states. The following argument justifies such an assumption. (The argument in the C*-algebra context is due to Haag and Kastler.[4])

In actual measurement, we can obtain only a finite number of data. This means that (i) we can distinguish only a finite number of distinctive results of measurements (and hence we can measure only those observables with a discrete spectrum such as questions), (ii) we can obtain information about only a finite number of observables and (iii) the measured value has a finite error (statistical or other errors). Thus information about a state $\alpha$ given by an actual measurement is expressed by

$$|w_\alpha^{Q_j}(q_j) - w_j| < \varepsilon_j, \qquad (j=1, \cdots, N) \tag{A·14}$$

or equivalently (see § A. 6) by

$$|\mathrm{Exp}(Q_j'/\alpha) - \varphi_j| < \varepsilon_j, \qquad (j=1, \cdots, N) \tag{A·15}$$

where $w_j$ is the measured probability and $\varphi_j$ is the measured expectation value. In principle, we are supposed to be able to choose $(Q_j, q_j)$ and make $\varepsilon_j$ as small as possible (at least in the following argument).

We may view the set of numbers $w_\alpha^Q(q)$ with $Q$ and $q$ varying over all possibilities as coordinates of the state $\alpha$. Then a state can be represented by a point in the (direct) product of the compact interval $[0, 1]$ indexed by $(Q, q)$.

The mixture of states is then a convex combination of points. The information about the state $\alpha$ given by (A·14) is that a finite number of its coordinates are in open intervals $(w_j-\varepsilon_j, w_j+\varepsilon_j)$. Such a "cylinder" set generates the direct product topology (in some context, called weak topology) and the direct product of $[0,1]$ is compact with respect to such a topology [Tihonov theorem].

A theory (or a model) is supposed to express states and observables by mathematical objects, for which rules for computing $w_\alpha^Q(q)$ (or $\mathrm{Exp}(Q/\alpha)$) are given. More specifically, at the present level of generality, a convex subset of the direct product $\Pi$ of $[0,1]$ (indexed by $(Q,q)$) is a model for the set of states of the system under consideration. For a description of observed states (and its time development, for example), it is enough to find at least one point in the given model corresponding to the information about the state $\alpha$ given by (A·14), i.e., a point whose $(Q_j, q_j)$-coordinate is in the open interval $(w_j-\varepsilon_j, w_j+\varepsilon_j)$ for $j=1, \cdots, N$. On the other hand, we might want to exclude all those points which are irrelevant for the description of states of a given system.

From this point of view, two convex subsets of $\Pi$ with the same closure have the equivalent capacity for description of states because if an open set given by (A·14) has at least one point $x$ in one convex set, then it has at least one point in common with another (because a sufficiently small neighbourhood of $x$ is completely in the open set given by (A·14) and must have a non-empty intersection with another convex set). On the other hand, if a point $y$ is outside the closure of a convex set $S$, then there exist a finite number of $(Q_j, q_j)$ $(j=1, \cdots, N)$ such that $(Q_j, q_j)$-coordinates separate $y$ from $S$ by a non-zero distance $d$. By choosing $\varepsilon_j$ sufficiently small, we should be able to distinguish in principle (up to a restriction of a finiteness of the number of states which can be tested) two cases (i) $y$ is not relevant in the description of any of the states under consideration or (ii) $S$ is insufficient for the description of the totality of states under consideration. Therefore two convex subsets of $\Pi$ with different closures describe different sets of states in principle. In this sense, we call two subsets of $\Pi$ *physically equivalent* if and only if they have the same closure. The choice between two physically equivalent models for the description of a given system can be made purely from mathematical convenience.

In particular, for any convex subset of $\Pi$, its closure in $\Pi$ is a compact convex set due to the compactness of $\Pi$ and is physically equivalent to the original convex subset. Therefore we may always take a convex compact subset of $\Pi$ as a model of the set of states of a given system. By Krein-Milman theorem, a compact convex set (in a Hausdorff locally convex topological vector space $\Pi$) has pure states (mathematically called extremal points) whose mixtures are dense in itself. Therefore our discussion on classical observables is meaningful for such a choice of model.

**A.10.**

We shall now discuss classical observables in the context of operator algebra

approach.

Let $\Omega(=R^N)$ be a phase space of classical mechanics. We may take as observables the set $\mathfrak{A}(=C_0(\Omega))$ of all continuous functions on $\Omega$ vanishing at $\infty$. It is a (C*-) algebra with respect to pointwise sum and multiplication. An arbitrary state $\varphi$ of $\mathfrak{A}$ (mathematically defined as a positive linear functional on $\mathfrak{A}$ with norm 1) can be described in terms of a probability measure $\mu_\varphi$ on $\Omega$ by

$$\varphi(f) = \int_\Omega f(x)\, d\mu_\varphi(x). \tag{A·16}$$

(In our general framework, $\varphi(f)$ is the expectation value of $f$ in the state $\varphi$ and a function $F(f)$ of $f$ is the function $F(f(x))$ in the usual sense.) An arbitrary pure state $\chi$ is given by a Dirac measure $\mu=\delta_a$ at some point $a$ of $\Omega$:

$$\chi(f) = f(a). \tag{A·17}$$

The time translation is usually described in terms of an orbit $x(t;a)$ for any given initial condition $x(0;a)=a$. This induces the time translation automorphisms of observables $f\in\mathfrak{A}$:

$$f_t(y) \equiv f(x(t;y)). \tag{A·18}$$

The time translation of a state $\varphi$ is described by

$$\varphi_t(f) = \varphi(f_t). \tag{A·19}$$

For example, the pure state $\chi$ given by (A·17) (concentrated at a point $a$) moves to $\chi_t$ concentrated at $x(t;a)$. Therefore the time translation automorphisms $f\to f_t$ of observables contains the information about the motion $a\to x(t;a)$ of a point in the phase space.

The algebra $\mathfrak{A}$ considered above is an abelian C*-algebra. Actually any abelian C*-algebra $\mathfrak{A}$ has the same type of description. Namely the set $\Omega$ of all pure states of $\mathfrak{A}$ is a locally compact Hausdorff space (with respect to weak-*topology), $\mathfrak{A}$ can be identified with the algebra $C_0(\Omega)$ of all continuous functions on $\Omega$ (Gelfand isomorphism) vanishing at $\infty$ and states on $\mathfrak{A}$ can be identified with probability measures on $\Omega$ (so-called Radon measures). The time translation can be described either by a one-parameter group of automorphisms of $\mathfrak{A}$ or continuous orbits on $\Omega$, just as in the above example of $\Omega=R^N$.

In the above situation, all observables $f\in\mathfrak{A}$ takes a sharp value $f(a)$ on any pure state $a\in\Omega$, so that all observables are classical observables, i.e., the commutative C*-algebras describes classical physics. (If we insist on observables with discrete spectrum, then we have to use the elements of the von Neumann algebra $\mathfrak{A}$**.)

In the more general situation where $\mathfrak{A}$ is not commutative, then the central elements of $\mathfrak{A}$ are classical observables. In the case of algebras of operators of the form (2·3), the center of algebra consists of the operators $A$ of the form (2·3)

where each $A(\lambda)$ is a multiple of the identity operator. Thus the direct integral with $\lambda$-variables in § 2 corresponds to the diagonalization of classical observables of the system. In this example, pure states are not given by vectors or density matrices of the Hilbert space $L$. If we restrict observables to those $A$ of the form (2·3) where $A(\lambda)$ is continuous in $\lambda$, vectors of $L_\lambda$ give pure states, which can be approximated by states obtained by vectors of $L$ in the weak sense described earlier. Our definition of classical observables applies to this example if we include all these pure states into our consideration. By using the continuous classical observables, we then attain the reduction of wave packets as described in § 3.

**Acknowledgements**

The author is indebted to Professor S. Machida, Professor M. Namiki and Professor M. Yanase for drawing the author's attention to the problems and for kind discussions. A part of the Appendix (A. 1∼A. 8) is taken from the beginning part of the author's lecture at ETH, Zürich in 1961∼2 and the rest of the Appendix is probably common knowledge of operator algebra theorists. He owes the inclusion of this appendix to the encouragement and insistence of Professor M. Yanase.

**References**

1) S. Machida and M. Namiki, Prog. Theor. Phys. **63** (1980), 1833.
2) E. P. Wigner, Am. J. Phys. **31** (1963), 6.
3) A. Fine, Phys. Rev. **D2** (1970), 2783.
4) R. Haag and D. Kastler, J. Math. Phys. **5** (1964), 848.