International Journal of Theoretical Physics, Vol. 16, No. 5 (1977), pp. 321-331

Information-Theoretical Aspects of Quantum Measurement[1]

E. PRUGOVEČKI

Department of Mathematics, University of Toronto, Toronto, Canada M5S 1A1

Received: 21 *January* 1977

Abstract

We present criteria for comparing measurements on a given system from the point of view of the information they provide. These criteria lead to a concept of *informational* completeness of a set of observables, which generalizes the conventional concept of completeness. The entropy of a state with respect to an arbitrary sample space of potential measurement outcomes is defined, and then studied in the context of configuration space and fuzzy stochastic phase space.

1. *Introduction*

Obviously, one of the basic purposes of any measurement procedure is to gather information. Yet, little attempt has been made in setting general guidelines for ranking different measuring procedures (on a given physical system) according to the effectiveness with which this task is achieved.

In trying to provide such guidelines, we are faced from the start with the fundamental problem of presenting a precise definition of the term "information." At first sight, the solution might seem obvious since information theory is a well-developed discipline, and the notion of entropy of a random phenomenon

$$S = -\sum_{i=1}^{N} p_i \ln p_i \tag{1.1}$$

with N possible outcomes was long ago introduced by Shannon (1948) and others (cf. Aczel and Daroczy, 1975) as a suitable measure of information. Thus, as long as a measurement procedure can be regarded as supplying information in the form of assigning weights p_i to various possibilities from a finite sample space X, (1.1) can be taken as a measure of that information. However, many of the sample spaces in physics are infinite (having a cardinal

[1] Supported in part by the National Research Council of Canada.

number equal to or larger than that of the real line) and the probabilities are determined only within certain margins of error. This indicates that no absolute measure of information, but only an informational ordering might be feasible. Furthermore, this ordering might be only partial when measurement procedures over different sample spaces (e.g., position and angular momentum spaces) are compared. This turns out to be actually the case in quantum mechanics owing to the presence of incompatible observables, which prohibit the adoption of the most straightforward strategy of solving this last problem, namely, that of taking the Cartesian product of the two original spaces as a new sample space.

In Section 2 we introduce a partial ordering of *all* conceivable measurement procedures on a given system. This ordering is based upon set-theoretical inclusion of the states compatible with the data extractable from such measurements, and represents a natural extension of the type of reasoning that underlies some definitions (Haag and Kastler, 1964; Prugovečki, 1966) of *physical* equivalence of quantum theories. It also crystallizes part of the motivation behind the recent introduction (Prugovečki, 1976a) of the concept of fuzzy simultaneous measurement of noncommuting observables in quantum mechanics, since it clearly displays the informational gain resulting from such measurements.

In Section 3 we show that for preparatory measurements this partial ordering can be made into a linear ordering by introducing the concept of entropy of such measurements with respect to a given sample space—a concept based upon a corresponding concept of entropy for states.

2. *A Partial Ordering of Measurements According to their State-Resolution Capability*

One can view any measurement on an ensemble of replicas of the same system purely from an information-theoretical point of view, namely, as an informational input into the ensemble (and into the theoretical model used to compute the future stochastic behavior of that ensemble) if the measurement is preparatory, or as an informational output from the ensemble (which is then to be compared with the theoretical prediction) in case the measurement is determinative (Prugovečki, 1973). After a reduction of the data resulting from the measurement procedure is performed, this informational transaction can be used to single out a range $R(\mathscr{M})$ of states of the system, namely, those states that are compatible with the data obtained.

As a typical example, consider a determinative simultaneous measurement of the commuting observables $A_1 \ldots, A_n$ of a quantum-mechanical system, i.e., a measurement that, figuratively speaking, determines the values $\lambda_1, \ldots, \lambda_n$ of $A_1, \ldots, A_n$ which each element in the ensemble "would have had" if the measurement apparatus had caused absolutely no disturbance in its behavior (cf. Prugovečki, 1967, 1971, 1973 for details in terminology and notation). A reduction of the obtained data leads to a specification within error margins $\pm\epsilon(\Delta_i)$ of the acceptable a priori probabilities $P^{\hat{A}}(\Delta_i)$ for having obtained $(\lambda_1, \ldots, \lambda_n) \in \Delta_i$, where $\Delta_1, \Delta_2, \ldots$ are suitably selected (Borel) sets in $\mathbb{R}^n$. This information can then be used in deciding which density operator ρ would have represented the state of the ensemble at the instant of measurement if there had been no disturbance caused by the measuring procedure itself. Since for given ρ the theoretically predicted probability for measuring a result within Δ_i is $\mathrm{Tr}[\rho E^{\hat{A}}(\Delta_i)]$, where $E^{\hat{A}}(\Delta)$ is the spectral measure of the set $\hat{A} = \{A_1, \ldots, A_n\}$, we must have

$$|P^{\hat{A}}(\Delta_i) - \mathrm{Tr}[\rho E^{\hat{A}}(\Delta_i)]| \leqslant \epsilon(\Delta_i) \tag{2.1}$$

We note that in the absence of additional information about the system, any statistical operator ρ that satisfies (2.1) for all the preselected Δ_i's is a valid choice for the state of the ensemble. Thus, the set $R(\mathscr{M})$ of all ρ's satisfying (2.1) can be said to be compatible with the data obtained.

This observation suggests a natural partial ordering of all conceivable measurements in a given ensemble; we shall say that a measurement $\mathscr{M}$ *provides more information* than another measurement $\mathscr{M}_2$ if $R(\mathscr{M}_1)$ is a proper subset of $R(\mathscr{M}_2)$; then we write Inf $\mathscr{M}_1 >$ Inf $\mathscr{M}_2$. In case we only know that $R(\mathscr{M}_1)$ is a subset of $R(\mathscr{M}_2)$, without knowing whether it is a proper subset, we shall say that $\mathscr{M}_1$ provides no less information then $\mathscr{M}_2$, and write Inf $\mathscr{M}_1 \geqslant$ Inf $\mathscr{M}_2$. Thus, loosely speaking, the fewer states are compatible with the data provided by a certain measurement procedure $\mathscr{M}$, the more information $\mathscr{M}$ provides. If, in general, the ranges $R(\mathscr{M})$ specified by each $\mathscr{M}$ were finite sets, we could adopt the ordering according to the decrease in the number of their elements as the basis for transforming the partial ordering introduced above into a linear ordering. However, generally speaking, the opposite is the case, as illustrated by measurements $\mathscr{M}$ on commuting observables $A_1, \ldots, A_n$ with continuous spectra whose $R(\mathscr{M})$ are specified in (2.1): If $\epsilon(\Delta_i) > 0$, then the ranges $R(\mathscr{M})$ are uncountably infinite sets.

Determinative measurements of $\hat{A} = \{A_1, \ldots, A_n\}$ can be considered to provide a specification of probabilities ranges $P^{\hat{A}}(\Delta_i) \pm \epsilon(\Delta_i)$, namely the probabilities computable from the frequencies with which the determined data fall within chosen Borel sets in $\mathbb{R}^n$. The more precise the apparatus used in the measurement and the larger the sample of extracted data, the more information is obtained in the sense of narrowing down the range $R(\mathscr{M})$ of candidate states compatible with those data [e.g., the range of states that satisfy (2.1)]. Obviously, the informationally optional determinative procedure is the one that narrows $R(\mathscr{M})$ down to a single element—the "true" state of the system. Hence, in this context, given a theory and a set $\hat{A}$ of observables, it becomes important to understand when this can happen—if not in actuality at least in an idealized limit of indefinitely increasing number of raw data obtained with instruments of ever-increasing accuracy. Thus, we are interested in whether there exists for every state ρ a specification of probabilities $P^{\hat{A}}(\Delta_i)$ that singles out that state exclusively.

To state it more formally, we say a set $\hat{A}$ of observables is *informationally* *complete with respect to the state* ρ if there is a specification of probabilities $P^{\hat{A}}(\Delta)$ such that ρ is the only density operator satisfying

$$\mathrm{Tr}[E^{\hat{A}}(\Delta)\rho] = P^{\hat{A}}(\Delta) \tag{2.2}$$

for all the specified probabilities $P^{\hat{A}}(\Delta)$; if $\hat{A}$ is informationally complete with respect to all states of the system, then we shall say that $\hat{A}$ is *informationally complete globally*.

We note that a set $\hat{A}$ of commuting observables that is complete in the conventional sense is not informationally complete globally. In fact, such a set is informationally complete only with respect to pure states represented by simultaneous eigenvectors of all observables in the set. Moreover, if one of the observables has a purely continuous spectrum, then $\hat{A}$ is not informationally complete with respect to any states whatsoever.

A simple illustration of this point is provided by the position observables $\mathbf{Q}$ of a one-particle spinless nonrelativistic system: For every given probability density $\omega(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^3$, over configuration space $\mathbb{R}^3$, when there is one density matrix $\rho(\mathbf{x}, \mathbf{x}')$ that satisfies

$$\omega(\mathbf{x}) \equiv \rho(\mathbf{x}, \mathbf{x}), \tag{2.3}$$

there is an infinity of other solutions ρ of (2.3). Moreover, even if a simultaneous momentum probability density is specified, a unique state ρ is still not singled out.

This last fact can be most easily demonstrated on pure states. Consider any wave-packet $\psi \in L^2(\mathbb{R}^3)$,

$$\psi(\mathbf{x}) = |\psi(\mathbf{x})| \exp[i\theta(\mathbf{x})], \qquad 0 \leqslant \theta(\mathbf{x}) < 2\pi \tag{2.4}$$

for which (almost everywhere in $\mathbb{R}^3$) $|\psi(\mathbf{x})| = |\psi(-\mathbf{x})|$ but

$$\theta(\mathbf{x}) + \theta(-\mathbf{x}) \neq \text{const.} \pmod{2\pi} \tag{2.5}$$

Then the state represented by

$$\psi_1(\mathbf{x}) = |\psi(\mathbf{x})| \exp[-i\theta(-\mathbf{x})] = \psi^*(-\mathbf{x}) \tag{2.6}$$

is distinct from that represented by ψ. Yet

$$|\psi(\mathbf{x})|^2 = |\psi_1(\mathbf{x})|^2, \qquad |\tilde{\psi}(\mathbf{k})|^2 = |\tilde{\psi}_1(\mathbf{k})|^2 \tag{2.7}$$

for almost all values of $\mathbf{x}, \mathbf{k} \in \mathbb{R}^3$, since the first equality is true by definition, while the second follows from the observation that

$$\begin{aligned}\tilde{\psi}_1(\mathbf{k}) &= (2\pi)^{-3/2} \int_{\mathbb{R}^3} \exp(-i\mathbf{k}\mathbf{x})\psi^*(-\mathbf{x})\, d\mathbf{x} \\ &= (2\pi)^{-3/2} \int_{\mathbb{R}^3} \exp(i\mathbf{k}\mathbf{x}')\psi^*(\mathbf{x}')\, d\mathbf{x}' = \tilde{\psi}^*(\mathbf{k})\end{aligned} \tag{2.8}$$

On the other hand, the specification of the probabilities

$$P^{\mathbf{Q},\mathbf{P}}(\Delta) = \mathrm{Tr}[\rho E^{\mathbf{Q},\mathbf{P}}(\Delta)] \tag{2.9}$$

for all fuzzy (Borel) subsets Δ of a given open fuzzy set Δ_0 in the phase space Γ_s (resulting from simultaneous fuzzy measurements of position $\mathbf{Q}$ and momentum $\mathbf{P}$) do single out a unique ρ (Prugovečki, 1976b). Thus the set $\{\mathbf{Q}, \mathbf{P}\}$ is informationally complete globally. This feature of simultaneous (fuzzy) measurements of $\mathbf{Q}$ and $\mathbf{P}$ is all the more significant when it is realized that probability specifications over the entire configuration space are informationally equivalent in the sharp and fuzzy case, and the same holds true for momentum space (Ali and Doebner, 1976). Thus, it is not the "fuzziness" of measurement, but the intrinsic properties of $\mathbf{Q}$ and $\mathbf{P}$ that make these observables informationally complete.

It is interesting to note that the informational equivalence of sharp and fuzzy measurements can be displayed also quite effectively in classical statistical mechanics. In fact, let $\Gamma_{0,0}$ denote the conventional phase space of a system having n degrees of freedom, and let (for fixed $s, r > 0$)

$$\Gamma_{s,r} = \{(q, \chi_q^{(s)}) \times (p, \chi_p^{(r)}) \mid q, p \in R^n\} \tag{2.10}$$

$$\chi_q^{(s)}(x) = (\pi s^2)^{-n/2} \exp[-s^{-2}(x-q)^2] \tag{2.11}$$

be the fuzzy phase space corresponding to measurements of q and p with imperfectly precise instruments whose accuracy calibrations (Prugovečki, 1976a) at $(q, p) \in \mathbb{R}^{2n}$ are represented by the Gaussian confidence functions $\chi_q^{(s)}$ and $\chi_q^{(r)}$, respectively. [$\Gamma_{s,r}$ is the classical counterpart the quantum-mechanical fuzzy phase space Γ_s in (3.6), where, in the absence of the uncertainty relations, there is no lower bound on $s \cdot r$]. If $\rho(q, p)$ is the probability density in $\Gamma_{0,0}$ for a classical ensemble, then

$$\rho^{(s,r)}(q, p) = \int \chi_q^{(s)}(x)\chi_p^{(r)}(k)\rho(x, k)d^n x d^n k \tag{2.12}$$

is the corresponding probability density in $\Gamma_{s,r}$. It is evident that $\rho_1(q, p) \equiv \rho_2(q, p)$ implies

$$\rho_1^{(s,r)}(q, p) - \rho_2^{(s,r)}(q, p) \equiv 0 \tag{2.13}$$

Moreover, the converse is also true, as can be easily seen by taking the Fourier transform of both sides in (2.13). This yields

$$\exp\left\{-\frac{s^2}{4}x^2 - \frac{r^2}{4}k^2\right\} [\tilde{\rho}_1(x, k) - \tilde{\rho}_2(x, k)] \equiv 0 \tag{2.14}$$

and we conclude that $\tilde{\rho}_1(x, k) \equiv \tilde{\rho}_2(x, k)$, and therefore $\rho_1(q, p) \equiv \rho_2(q, p)$.

This simple argument shows that the set of all position and momentum observables of a classical system is informationally complete regardless of whether the measurements of these observables are sharp or fuzzy. Thus the situation is completely analogous to that of a quantum-mechanical system, except that in the latter case the role of sharp measurements of position and momentum is played by optimal fuzzy measurements corresponding to instruments whose accuracy calibrations yield Gaussians having minimal variances with respect to the Heisenberg uncertainty relations.

On first sight it might seem paradoxical that global probability specifications on both sharp and fuzzy phase space are informationally equivalent, although, intuitively speaking, any single measurement of a sharp point $(q, p) \in \Gamma_{0,0}$ supplies more information then the measurement of a corresponding fuzzy point $(q, p, \chi_{q,p}^{(s,r)}) \in \Gamma_{s,r}$. The paradox, however, disappears as soon as one realizes that any such specification requires an (uncountable) infinity of data since it involves absolutely accurate specifications of probability densities at all points in the respective Γ-space. On the other hand, any actual measurement supplies only a finite amount of data, and therefore this informational equivalence does not extend to two actual measurements $\mathscr{M}_1$ and $\mathscr{M}_2$ which supply exactly the same set of points in phase space, if in the case of, say, $\mathscr{M}_1$ one employs instruments that are more accurate (i.e., instruments whose confidence functions at each point on their reading scales have narrower spreads) than in the case of $\mathscr{M}_2$. On the contrary, it can be expected that in such a case we shall have $\mathrm{Inf}\, \mathscr{M}_1 > \mathrm{Inf}\, \mathscr{M}_2$.

3. *The Entropy of Preparatory Measurements*

As pointed out in the preceding section, the informationally optimal determinative measurements $\mathscr{M}$ are the ones that narrow down the corresponding ranges $R(\mathscr{M})$ to one-point sets, namely, the true states ρ in which the system was prior to measurement. Thus it would be meaningless to further differentiate between the informational efficiency of such optimal measurements since no improvements in the measuring procedure could supply any additional information: Since ρ is known, quantum theory provides the probability distributions for all other observables of the system.

This statement does not apply, however, to preparatory measurement, whose purpose is to ensure that immediately after the measurement certain relevant observables $A_1, \ldots, A_n$ have values within specified ranges. Thus, let us say that as a result of the preparatory measurements $\mathscr{M}_j, j = 1, 2$, we know that the system "has" values for $\{A_1, \ldots, A_n\}$ within the sets Δ_j, $j = 1, 2$, and that $R(\mathscr{M}_j) = \{\psi_j\}$, i.e., that both $\mathscr{M}_1$ and $\mathscr{M}_2$ have prepared pure states represented by the respective state vectors ψ_1 and ψ_2. Suppose now that Δ_1 is a proper subset of Δ_2. Whereas the partial ordering of Section 2 does not differentiate between $\mathscr{M}_1$ and $\mathscr{M}_2$, we can nevertheless claim that $\mathscr{M}_1$ has prepared a state with more information content than the state prepared by $\mathscr{M}_2$.

We shall impart to this observation a more universal meaning by introducing a linear ordering among preparatory measurements.

We begin by defining for each state ρ the concept of its *entropy* S_ρ *with respect to a sample space* X of (sharp or fuzzy) simultaneous values of the set $\hat{A}$ of observables

$$S_\rho(X) = -\int p_\rho(\xi; X) \log p_\rho(\xi; X)\, d\xi \tag{3.1}$$

Here the integrand $p_\rho \log p_\rho$ is taken to be zero at the points ξ where

$p_\rho(\xi, X) = 0$, and the integration symbol is to be interpreted as summation when the "integration" is carried out over the point spectra of the observables in $\hat{A}$. More specifically, if we assume that the observables $A_1, \ldots, A_n$ in $\hat{A}$ have absolutely continuous spectra, then the spectral measure $E^{\hat{A}}$ over the Borel sets of the (sharp or fuzzy) sample space X is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^n$ and the Radon-Nykodim derivative

$$p_\rho(\xi, X) = \lim_{\Delta \to \{\xi\}} |\Delta|^{-1} \operatorname{Tr}[\rho E^{\hat{A}}(\Delta)] \tag{3.2}$$

exists, so that (3.1) is a Lebesgue integral over $\mathbb{R}^n$. Since this last case is the one that embodies all the mathematically nontrivial problems that (3.1) gives rise to, we shall limit our subsequent comments to it.

Operationally, a sample space X consists of sharp (ξ, δ_ξ) or fuzzy (ξ, χ_ξ) sample points whose confidence functions χ_ξ are the result of accuracy calibrations of instruments used in determinative measurements of $\hat{A}$ (Prugovečki, 1976a, b). Typical examples of such spaces for a one-particle system are the sharp stochastic configuration space

$$\mathbb{R}_0{}^3 = \{(\mathbf{q}, \delta_\mathbf{q}) \mid \mathbf{q} \in R^3\} \tag{3.3}$$

the fuzzy stochastic configuration spaces

$$\mathbb{R}_s{}^3 = \{(\mathbf{q}, \chi_\mathbf{q}^{(s)}) \mid \mathbf{q} \in \mathbb{R}^3\} \tag{3.4}$$

$$\chi_\mathbf{q}^{(s)}(\mathbf{x}) = (\pi s^2)^{-3/2} \exp[-s^{-2}(\mathbf{x} - \mathbf{q})^2] \tag{3.5}$$

for $0 < s < \infty$, and the fuzzy stochastic phase spaces

$$\Gamma_s = \{(\mathbf{q}, \chi_\mathbf{q}^{(s)}) \times (\mathbf{p}, \chi_\mathbf{p}^{(s^{-1})}) \mid (\mathbf{q}, \mathbf{p}) \in \mathbb{R}^6\} \tag{3.6}$$

The corresponding probability densities, in terms of the density matrix $\rho(\mathbf{x}, \mathbf{x}')$, are

$$p_\rho(\mathbf{q}; \mathbb{R}_0{}^3) = \rho(\mathbf{q}, \mathbf{q}) \tag{3.7}$$

$$p_\rho(\mathbf{q}; \mathbb{R}_s{}^3) = \int_\mathbb{R} \chi_\mathbf{q}^{(s)}(\mathbf{x}) \rho(\mathbf{x}, \mathbf{x})\, d\mathbf{x} \tag{3.8}$$

$$p_\rho(\mathbf{q}, \mathbf{p}; \Gamma_s) = \int \chi_\mathbf{q}^{(s)}(\mathbf{x}) \chi_\mathbf{p}^{(s^{-1})}(\mathbf{k}) w_\rho(\mathbf{x}, \mathbf{k})\, d\mathbf{x}\, d\mathbf{k} \tag{3.9}$$

where w_ρ is the Wigner transform of ρ.

In the case when we are dealing with commuting observables $A_1, \ldots, A_n$ having finite point spectra, (3.1) coincides with (1.1) and is finite and nonnegative. This is not so, however, for observables with continuous spectra. Then, in general, $S_\rho(X)$ can assume any real value, including also the extreme values $\pm\infty$. For example, such is the case for the (sharp) configuration space $\mathbb{R}_0{}^3$, when

$$-\infty \leqslant S_\rho(\mathbb{R}_0{}^3) \leqslant +\infty \tag{3.10}$$

This fact can be easily ascertained by noting that for any probability density $\omega(\mathbf{q})$ in $\mathbb{R}^3$ there are wave-packets $\psi(\mathbf{x})$ for which $|\psi(\mathbf{q})|^2 = \omega(\mathbf{q})$, and, at the same time, there are many examples (Ash, 1965) of probability densities $\omega(\mathbf{x})$ for which

$$\int \omega(\mathbf{x}) \log \omega(\mathbf{x})\, d\mathbf{x} = \pm\infty \tag{3.11}$$

In view of (3.10), it is interesting to note that

$$0 < S_\rho(\Gamma_s) \leqslant +\infty \tag{3.12}$$

Indeed, for any density operator ρ we have

$$\rho = \Sigma |\psi_j\rangle \lambda_j \langle\psi_j|, \qquad \Sigma \lambda_j = 1 \tag{3.13}$$

where $\lambda_j \geqslant 0$ and $\langle\psi_i|\psi_j\rangle = \delta_{ij}$. Since (Prugovečki, 1976b)

$$|\psi_j(\mathbf{q}, \mathbf{p}; s)| \leqslant \|\psi_j\| = 1 \tag{3.14}$$

in units with $\hbar = 1$, we immediately get

$$p_\rho(\mathbf{q}, \mathbf{p}; \Gamma_s) = \Sigma \lambda_j |\psi_j(\mathbf{q}, \mathbf{p}; s)|^2 \leqslant 1 \tag{3.15}$$

Consequently

$$-p_\rho(\mathbf{q}, \mathbf{p}; \Gamma_s) \log p_\rho(\mathbf{q}, \mathbf{p}; \Gamma_s) \geqslant 0 \tag{3.16}$$

and since p_ρ is continuous and has only isolated zeros (Prugovečki, 1976b), we conclude that $S_\rho(\Gamma_s) > 0$.

In general, although the two extremes $S_\rho(X) = \pm\infty$ can occur in the continuous case, they are the exception rather than the rule. First of all, we evidently have $S_\rho(X) > -\infty$ whenever $p_\rho(\xi; X)$ is almost everywhere bounded. Secondly, when that is so, it can be shown by an easy generalization of the proof of Theorem 8.3.3 by Ash (1965) that $S_\rho(X) < +\infty$ if and only if the probability measure $p_\rho(\xi; X)$ has finite standard deviations σ_j,

$$\sigma_j{}^2 = \int_{\mathbb{R}^n} (\xi_j - \bar{\xi}_j)^2 p_\rho(\xi_1, \ldots, \xi_n; X)\, d\xi_1, \ldots d\xi_n \tag{3.17}$$

Furthermore, in that case

$$S_\rho(X) \leqslant \tfrac{1}{2} \log\, [(2\pi e)^n \sigma_1{}^2 \cdots \sigma_n{}^2] \tag{3.18}$$

This observation can be used to establish that there are states ρ with $S_\rho(\Gamma_s) = +\infty$. Indeed, according to the aforesaid we only have to exhibit states ρ having infinite standard deviations in Γ_s. Let us consider therefore a pure state $\rho = |\psi\rangle\langle\psi|$, which in the $L^2(\Gamma_s)$ space is represented by the function

$$\psi(\mathbf{q}, \mathbf{p}; s) = \exp[-\tfrac{1}{2}|\boldsymbol{\zeta}|^2] f_\psi(\boldsymbol{\zeta}) \tag{3.19}$$

where f_ψ is an entire function of $\boldsymbol{\zeta} = 2^{-1/2}(s^{-1}\mathbf{q} - is\mathbf{p})$:

$$f_\psi(\boldsymbol{\zeta}) = \sum_{n_1, n_2, n_3 = l}^{\infty} \alpha_{n_1, n_2, n_3} \zeta_1^{n_1} \zeta_2^{n_2} \zeta_3^{n_3} \tag{3.20}$$

It turns out that (Bargmann, 1961)

$$\|\psi\|^2 = \Sigma n_1!n_2!n_3!|\alpha_{n_1,n_2,n_3}|^2 \tag{3.21}$$

and that each set of coefficients $\{\alpha_{n_1,n_2,n_3}\}$ for which (3.21) is finite determines a $\psi \in L^2(\Gamma_s)$. For simplicity, consider only $\psi(\mathbf{q}, \mathbf{p}; s) \in L^2(\Gamma_s)$ whose expectation values in $\mathbf{q}$ and $\mathbf{p}$ are zero. Then obviously $|\psi(\mathbf{q}, \mathbf{p}; s)|^2$ will have finite variances in each component of $\mathbf{q}$ and $\mathbf{p}$ if and only if

$$\int |\zeta\psi(\mathbf{q}, \mathbf{p}; s)|^2 d\mathbf{q}\, d\mathbf{p} < \infty \tag{3.22}$$

i.e., if and only if for $j = 1, 2, 3$ the functions

$$\zeta_j f_\psi(\zeta) = \Sigma\, \alpha_{n_1,n_2,n_3} \zeta_1^{n_1+\delta_{1j}} \zeta_2^{n_2+\delta_{2j}} \zeta_3^{n_3+\delta_{3j}} \tag{3.23}$$

give rise to elements of $L^2(\Gamma_s)$ when they are substituted in place of f_ψ in (3.19). But in accordance with (3.21), a necessary and sufficient condition for these functions to belong to $L^2(\Gamma_s)$ is that

$$\Sigma(n_1 + \delta_{1j})!(n_2 + \delta_{2j})!(n_3 + \delta_{3j})!|\alpha_{n_1,n_2,n_3}|^2 < \infty \tag{3.24}$$

Since there obviously exist sets $\{\alpha_{n_1,n_2,n_3}\}$ for which (3.21) converges while (3.24) diverges, we have established the existence of states $\rho = |\psi\rangle\langle\psi|$ for which $S_\rho(\Gamma_s) = +\infty$.

In conclusion let us point out that the possibility of $S_\rho(X)$ being negative or infinite prevents its interpretation as an absolute measure of uncertainty (i.e., of lack of information). But $S_\rho(X)$ can be taken as a *relative* measure of information, i.e., used as a means of linearly ordering all states of a quantum-mechanical system in relation to the information they carry with respect to a sample space X: We shall say that ρ_1 *carries no less information* on the values of X than ρ_2 does (and write $\mathrm{Inf}_X\, \rho_1 \geqslant \mathrm{Inf}_X\, \rho_2$) if and only if $S_{\rho_1}(X_1) \leqslant S_{\rho_2}(X_2)$.

This ordering of states obviously implies right away a corresponding ordering of those preparatory measurements $\mathscr{M}$ that prepare a single state ρ [i.e., for which $R(\mathscr{M}) = \{\rho\}$] with regard to the information they provide about values in X. However, generally speaking, $R(\mathscr{M})$ will contain more than one element and the linear ordering $\mathrm{Inf}_X\, \mathscr{M}$ we shall introduce in that case has to be consistent with the partial ordering introduced in the preceding section, i.e., $\mathrm{Inf}\, \mathscr{M}_1 \geqslant \mathrm{Inf}\, \mathscr{M}_2$ should imply $\mathrm{Inf}_X\, \mathscr{M}_1 \geqslant \mathrm{Inf}_X\, \mathscr{M}_2$ (but not necessarily the converse!). Hence we define the entropy of any given preparatory measurement $\mathscr{M}$ with respect to X as a supremum,

$$S_{\mathscr{M}}(X) = \sup_{\rho \in R(\mathscr{M})} S_\rho(X) \tag{3.25}$$

We shall say that $\mathscr{M}_1$ *has prepared no less information on X than* $\mathscr{M}_2$ if and only if $S_{\mathscr{M}_1}(X) \leqslant S_{\mathscr{M}_2}(X)$; we write then $\mathrm{Inf}_X\, \mathscr{M}_1 \geqslant \mathrm{Inf}_X\, \mathscr{M}_2$. Clearly, this linear ordering is consistent with the partial ordering of Section 2.

We note that the entropy of a state as defined in the context of quantum statistical mechanics is a special case of (3.1): It represents the entropy of ρ with respect to the sample space X of sharp measurements of the Hamiltonian

H (i.e., total energy) of the system. Now, in statistical mechanics the state of a system that is isolated and has reached a state of equilibrium is taken to correspond to a narrow band $[E, E + \epsilon]$ in the values of H. Among the range R of all states consistent with this statement the one of maximum entropy is considered to be the most likely. This is the state ρ at which the supremum in (3.25) is reached, i.e., the state representing a microcanonical ensemble. On the other hand, if the system interacts with an infinite reservoir having a given temperature, among all states consistent with this fact the one for which S_ρ assumes its maximum value is again taken to be the most likely, only this time that is the state representing a canonical ensemble. Thus, if we regard the procedure of letting a system achieve a state of equilibrium as preparatory with respect to its internal energy, we see that (3.25) conforms to the postulates of quantum statistical mechanics.

4. *Conclusion*

The two central concepts that emerge from the preceding considerations are the informational completeness of a set of observables and the entropy of a state with respect to a given sample space X of potential measurement outcomes.

We have seen that a complete set of observables in the conventional sense is not informationally complete globally (i.e., not capable of uniquely pinpointing any given state of the system) except in trivial cases. On the other hand, position and momentum observables do provide (for spinless particles) sets that are informationally complete globally in quantum mechanics—thus paralleling their analogous role in classical mechanics.

Informational completeness can be also defined for families of vectors: A parametrized family $\phi(\lambda_1, \ldots, \lambda_i)$, $(\lambda_1, \ldots, \lambda_i) \in \Omega$, is called informationally complete if

$$\int_\Omega |\phi(\lambda_1, \ldots, \lambda_i)\rangle \, d\mu(\lambda_1, \ldots, \lambda_i) \langle \phi(\lambda_1, \ldots, \lambda_i)| = 1 \tag{4.1}$$

for some measures μ on the (Borel) subset Ω of $\mathbb{R}^n$ and if

$$|\langle \phi(\lambda_1, \ldots, \lambda_i) | \phi_1 \rangle|^2 = |\langle \phi(\lambda_1, \ldots, \lambda_i) | \phi_2 \rangle|^2 \tag{4.2}$$

for (μ-almost) all $(\lambda_1, \ldots, \lambda_i) \in \Omega$ implies $\phi_1 = c\phi_2$, $|c| = 1$. We note that in a separable Hilbert space, any orthonormal basis $\{\phi_{m_1, \ldots, m_n}\}$ of vectors labeled by the non-negative integral indices $m_1, \ldots, m_n$ satisfies (4.1), but is not informationally complete since it does not satisfy the condition derived from (4.2). On the other hand, the family of coherent states

$$\phi(z_1, \ldots, z_n) = \pi^{-n/2} \exp[-\tfrac{1}{2}(|z_1|^2 + \cdots + |z_n|^2)] \sum_{m_1, \ldots, m_n = 0}^{\infty} \times \frac{z^{m_1} \cdots z^{m_n}}{(m_1! \ldots, m_n!)^{1/2}} \phi_{m_1, \ldots, m_n} \tag{4.3}$$

is informationally complete in the $2n$ real variables $\mathrm{Re}\, z_j$, $\mathrm{Im}\, z_j$, $j = 1, \ldots, n$, since (Bargmann, 1961)

$$h^{-n} \int_{\mathbb{R}^{2n}} | \phi(z_1, \ldots, z_n) \rangle \, d^2 z_1, \ldots, d^2 z_n \langle \phi(z_1, \ldots, z_n) | = \mathbb{1} \qquad (4.4)$$

Thus, the nontrivial problem is not to find informationally complete families of vectors (there is an uncountable infinity of such) but rather to investigate which sets of observables admit an association with some such families $\{\phi(\lambda_1, \ldots, \lambda_{2n})\}$ in a manner that allows the interpretation of

$$\int_\Delta |\langle \phi(\lambda_1, \ldots, \lambda_{2n}) | \phi \rangle|^2 \, d\mu(\lambda_1, \ldots, \lambda_{2n}) \qquad (4.5)$$

as being the probability of determining (sharp or fuzzy) values $(\lambda_1, \ldots, \lambda_{2n})$ belonging to Δ in a simultaneous measurement of $A_1, \ldots, A_{2n}$ carried in the system in the state ϕ.

Once this stage of the program has carried through for a given set $\{A_1, \ldots, A_{2n}\}$ (position and momentum observables providing one such instance) one can talk about the entropy (4.1) of each state with respect to the resulting sample space X of, in general, stochastic points. As discussed in Section 3, this leads then to a method of distinguishing between different preparatory procedures from the point of view of their effectiveness in providing information about values in X.

References

Aczel, J., and Doroczy, Z. (1975). *On Measures of Information and their Characterizations*, Academic Press, New York and London.

Ali, S. T., and Doebner, H. D. (1976). *Journal of Mathematical Physics*, 17, 1105.

Ash, R. (1965). *Information Theory,* Sec. 8.3, Interscience Publishers, New York 1965.

Bargmann, V. (1961). *Communications in Pure and Applied Mathematics*, **14**, 187.

Haag, R., and Kastler, D. (1964). *Journal of Mathematical Physics*, **5**, 848.

Prugovečki, E. (1966). *Journal of Mathematical Physics*, **7**, 1070.

Prugovečki, E. (1967). *Canadian Journal of Physics*, **45**, 2173.

Prugovečki, E. (1971). *Quantum Mechanics in Hilbert Space*, Chap. IV, Sect. 1, Academic Press, New York and London.

Prugovečki, E. (1973). *Foundations of Physics*, 3, 3.

Prugovečki, E. (1976a). *Journal of Mathematical Physics*, 17, 517.

Prugovečki, E. (1976b). *Journal of Mathematical Physics*, 17, 1673.

Shannon, C. E., (1948). *Bell System Technical Journal*, 27, 379.