Foundations of Physics, Vol. 6, No. 3, 1976

New Information-Theoretic Foundations for Quantum Statistics

William Band[1] and James L. Park[1,2]

Received May 19, 1975

When the state of a physical system is not fully determined by available data, it should be possible nevertheless to make a systematic guess concerning the unknown state by applying the principles of information theory. The resulting theoretical blend of informational and mechanical constructs should then constitute a modern structure for statistical physics. Such a program has been attempted by a number of authors, most notably Jaynes, with seeming success. However, we demonstrated in a recent publication that the standard list of so-called "mutually exclusive and exhaustive" quantum states that is commonly employed by these authors is in fact not exhaustive. It follows that the information-theoretic foundations of quantum statistics must be reformulated. The present paper discusses the fundamental problems involved and establishes a format for the correct application of information theory to quantum mechanical situations.

1. REFORMULATION OF THE CENTRAL PROBLEM OF QUANTUM STATISTICS

The orthodox information-theoretic treatment[(1–3)] of the foundations of quantum statistics starts from the premise that a given quantum system may be regarded as being in some unknown pure state $|\Psi\rangle$, which is one among a complete list of orthonormal alternatives $\{|\psi_n\rangle\}$. The members of that list are in one-to-one correspondence with quantal state propositions of this form: System $\mathscr{S}$ is in state $|\psi_n\rangle$. This collection of propositions is assumed to be exhaustive and mutually exclusive, so that, in accordance with normal

[1] Department of Physics, Washington State University, Pullman, Washington.
[2] Work supported by a grant from Research Corporation.

procedure in information theory, a subjective probability distribution $\{W_n\}$ may be defined over the propositions, which satisfies the constraint

$$\sum_n W_n = 1 \tag{1}$$

The information-theoretic entropy, or missing information function, may then be defined in the usual way as

$$I \equiv -\kappa \sum_n W_n \ln W_n \tag{2}$$

where κ is an arbitrary constant. By maximizing I subject to physical constraints on the $\{W_n\}$, it is possible to obtain the fundamental theorems of quantum statistics by well-known procedures.

The density operator formalism of quantum mechanics is customarily introduced into the analysis as a compact artifice for describing "states of ignorance" as to which among the $\{|\psi_n\rangle\}$ is the correct $|\Psi\rangle$. Thus the informational situation wherein each subjective probability in $\{W_n\}$ represents the likelihood that its associated state vector in $\{|\psi_n\rangle\}$ is the correct one is said to be characterized by the density operator

$$\rho = \sum_n W_n |\psi_n\rangle\langle\psi_n| \tag{3}$$

Due to the orthonormality of the $\{|\psi_n\rangle\}$, it is immediately possible to combine (2) and (3) to derive an expression for information-theoretic entropy which is often called the Gibbs–von Neumann form:

$$I = -\kappa \operatorname{Tr}(\rho \ln \rho) \tag{4}$$

It is true that this formula was introduced for thermodynamic entropy by von Neumann[4] in the course of his investigations concerning the density operator. It is also true that it bears a mathematical family resemblance to the classical construct Gibbs[5] called $\bar{\eta}$, the average index of probability of phase. However, we shall find shortly that the functional of ρ given by (4) is not the proper quantal analog to $\bar{\eta}$ and that Gibbs does not deserve the blame for von Neumann's formula.

Once the formula (4) has been adopted as the information measure in quantum statistics, that discipline acquires a formal elegance in its foundations which is easy to admire. Unfortunately, the entire framework has been erected, as it were, on pseudo-quantum-mechanical quicksand. In a recent publication[6] we scrutinized the traditional arguments that have been advanced in behalf of the basic list of quantal state propositions described above. We found that the list has been derived from a set of common misconceptions concerning the foundations of quantum theory. Moreover, the list does not possess the crucial property of exhaustiveness, which is necessary if Eq. (1) is to be valid. To be sure, the propositions in question are mutually exclusive, but not, as is usually stated, because the vectors $\{|\psi_n\rangle\}$ are mutually orthogonal.

We do not argue in the present paper that the informational approach to quantum statistics is intrinsically bad. On the contrary, we believe that abstract information theory offers perhaps the most cogent foundation for quantum statistics ever given; but we do assert that information theory has never been applied correctly to modern quantum mechanics. Indeed the program of discussing quantal situations involving incomplete information has never even been initiated properly. The first step in an information-theoretic analysis is to identify an exhaustive list of mutually exclusive propositions over which to define the subjective probabilities that enter into the missing information function. Since this first step has been taken improperly in the past, we must reformulate information-theoretic quantum statistics by starting anew from the very beginning.

Jaynes[7] has argued that Gibbs himself based classical statistical mechanics upon principles that were to be rediscovered and called information theory in a later generation. We are inclined to agree; indeed it seems to us eminently sensible to emulate Gibbs as far as possible in our attempt to establish a firm foundation for quantum statistical mechanics. However, in order to follow Gibbs' classical footsteps, we must be clear as to which quantal constructs are the theoretical analogs to the key ingredients of classical statistics. In the paper[6] cited previously in connection with the unacceptability of the standard list of state propositions, we also developed a chart (Table I) of correct analogies between classical and quantal constructs.

Among the analogies in Table I, neither A nor C is unusual. The

Table I

Construct	Classical representative	Quantal counterpart
A. System	Phase space	Hilbert space
B. State, or preparation, of system	Phase point (q, p)	Density operator ρ
C. Observable	Function of phase	Hermitian operator with complete orthonormal eigenvector set
D. Ignorance of true state	Gibbsian coefficient of probability of phase	Subjective probability distribution defined over density operators

significant departures from standard discussions of quantum statistics lie in B and D. Analogy B recognizes that the *density operator is an irreducible feature of basic quantum mechanics even when there is total information;* analogy D affirms the same point by noting that *ignorance as to the true quantum state must be represented by a subjective probability distribution defined over the density operators.* This format stands in marked contrast to the traditional view sketched earlier that an orthogonal set of pure states $\{|\psi_n\rangle\}$ surely contains the true state and that a density operator of the form (3) merely represents ignorance as to which $|\psi_n\rangle$ is correct.

We begin our new information-theoretic foundations of quantum statistics with the definition of a new *logical spectrum*, i.e., an exhaustive list of mutually exclusive state propositions. The typical proposition in that list has this form: System $\mathscr{S}$ is prepared in the manner characterized by quantum state ρ. The new logical spectrum comprises one such statement for every density operator ρ defined on the Hilbert space of the system of interest.

This list of propositions, unlike its predecessor, is exhaustive. In quantum mechanics, it is axiomatic that for every repeatable preparation, there is a density operator ρ. If enough data are gathered pertaining to a given preparation, ρ may be uniquely determined but will not necessarily have the pure form $|\Psi\rangle\langle\Psi|$. To determine ρ, what is required is an enormity of data sufficient to compute the mean values in the quantal ensemble of a set of observables, which we have called a *quorum.*[8] Whenever less than a full set of such quorum means is available, the density operator becomes unknown; and quantum statistics is born.

The new logical spectrum is not only exhaustive, but it is also mutually exclusive. States that have heretofore been regarded, on the basis of faulty quantal reasoning, as not being mutually exclusive are all included. A thorough analysis of this point may be found elsewhere.[6]

Having taken the first essential step, identification of the logical spectrum, it is now seemingly straightforward to apply information theory to quantum mechanical situations where the set of quorum means is incomplete. There is, however, an immediate difficulty—the new logical spectrum includes a *continuum* of alternatives. We return later (Section 3) to this rather serious complication, but for now let us assume that some kind of physical information has enabled us to eliminate all but a discrete list $\{\rho_n\}$ of possible density operators. The set $\{\rho_n\}$ is not necessarily of the form $\{|\psi_n\rangle\langle\psi_n|\}$, as some or all of the ρ_n may be mixed states. To apply information theory, we associate with each ρ_n a subjective probability W_n. Since the $\{\rho_n\}$ legitimately constitute an exhaustive set of mutually exclusive alternatives, Eq. (1) is valid; and Eq. (2) defines the information-theoretic entropy, which is to be maximized subject to whatever physical constraints restrict the $\{W_n\}$.

Even if the true ρ_n were known, all that quantum mechanics could predict would be the mean values of observables. Thus if A were measured on each element of an ensemble generated by repeated preparations in the manner characterized by ρ_n, the data gathered would have an arithmetic mean value $\bar{A}$ given by the well-known formula

$$(\bar{A})_n = \mathrm{Tr}(\rho_n A) \tag{5}$$

If ρ_n has only a likelihood W_n of being the true density operator, then the *expectation value* (of the quantal mean value) for an observable A is given by

$$\langle A \rangle = \sum_n W_n (\bar{A})_n \tag{6}$$

By combining (5) and (6), we obtain

$$\langle A \rangle = \mathrm{Tr}(\tilde{\rho} A) \tag{7}$$

where

$$\tilde{\rho} \equiv \sum_n W_n \rho_n \tag{8}$$

The operator $\tilde{\rho}$ has all the properties of a density operator, and in (7) it is seen to be useful for computing the quantum statistical expectation value by the same formula (5) that is used in ordinary quantum mechanics to compute mean values. Thus it is possible to say that assignment of a subjective probability distribution $\{W_n\}$ over a set of alternative density operators $\{\rho_n\}$ is equivalent in its practical consequences to believing that the unknown density operator is $\tilde{\rho}$ as defined in (8). In this sense, $\tilde{\rho}$ retrieves some of the spirit of the conventional ignorance interpretation of density operators described above for Eq. (3) but rejected later when correct quantal counterparts to Gibbsian quantities were considered. We discuss this point further in the next section.

It is extremely important to observe that the von Neumann formula (4) will play no role in our (general) reformulation of information-theoretic statistics, since

$$I \equiv -\kappa \sum_n W_n \ln W_n \neq -\kappa \,\mathrm{Tr}(\tilde{\rho} \ln \tilde{\rho}) \tag{9}$$

Finally, let $\{\hat{W}_n\}$ denote the subjective probability distribution for which I has its maximum value compatible with available information. The *best guess* for $\tilde{\rho}$ would then be given by

$$\hat{\rho} = \sum_n \hat{W}_n \rho_n \qquad (10$$

The calculation of $\hat{\rho}$ is the central problem of quantum statistics.

In a future publication treating *equilibrium* statistics from the perspective developed here, we shall find that the general inequality in (9) sometimes becomes an equality when $\hat{\rho}$ is substituted for $\tilde{\rho}$. In this sense the von Neumann entropy formula sometimes acquires a limited informational meaning, but our rejection of (4) as a general measure of quantal missing information still stands.

2. SUBJECTIVE AND OBJECTIVE PROBABILITIES

Philosophers concerned with the foundations of probability theory have long been divided into two famous camps. The subjectivist school promotes the notion that a probability is primarily a measure of likelihood or a degree of rational belief. According to this thesis, a probability can be meaningful even when it is inherently unmeasurable, for it is an abstract tool of thought which serves to regularize intellectual processes involving inductive inference. The antipodal view is that of the objectivist school, which holds that the probability concept acquires meaning and utility only after it has been equipped with an operational definition. That definition is of course the well-known identification of the probability of an event as the relative frequency of occurrence of that event in a statistical ensemble of identical situations. In spite of the philosophical gulf which separates these two schools of thought, there seem to be at least two bridges between them: Both employ the same mathematical calculus of probability, and the subjectivists regard the frequentist viewpoint as a special case valid whenever an ensemble is available.

Both versions of probability have been used in physics, often without careful distinction and hence without much controversy. However, with the advent of modern quantum mechanics, physicists encountered for the first time a *probabilistic* theory which claimed also to be a *fundamental* theory. Thus quantum mechanics is supposed to supplant classical mechanics as the irreducible theoretical framework with which we comprehend nature; yet quantum mechanics does not make exact predictions even in principle.

Of course it has always been the case that the exact predictions of the old mechanics were too precise for empirical verification and that a layer of probabilistic reasoning, the theory of errors, had to mediate between the exactitude of the theory and the comparative haziness of empirical experience. But this use of the probability concept in physics can be regarded as a means to cope with *ignorance* due to the crudity of laboratory mensuration, a technological difficulty.

In quantum mechanics, on the other hand, probability would remain in the fundamental predictions even if technological innovation enabled infinitely accurate measurement of every observable. Thus the probabilities in quantum mechanics have nothing to do with human ignorance; instead the quantal probabilities are themselves, so to speak, fundamental attributes of physical reality. It follows that for basic quantum theory, the appropriate version of probability is that of the objectivists.

We are of course aware that in the past a subjectivist interpretation of quantal probability has sometimes been promulgated even by eminent quantum theorists,[9] so that in effect the subjectivist–objectivist dichotomy in philosophy exists also in the world of theoretical quantum physics. However, it would be inappropriate here to elaborate on the mischief that has been visited upon the foundations of quantum theory by this intrusion of subjectivist thought; for that we refer the interested reader to the literature.[10]–[12]

In previous work[8,13] we have adopted the following statement of the axioms of quantum theory, which leaves no doubt as to the objective meaning of quantal probabilities:

I. With every physical *system*, there is associated a Hilbert space $\mathscr{H}$.

II. Each linear Hermitian operator A with a complete orthonormal eigenvector set corresponds to an *observable* of the system; any function of A corresponds to that same function of the observable represented by A.

III. With every repeatable empirical method of *preparation* of the system, there is associated a statistical operator ρ, the quantum state; the arithmetic mean $\bar{A}$ of a collective of A data gathered by measurements of A upon an *ensemble* of systems each prepared (identically) in the manner ρ is given by

$$\bar{A} = \mathrm{Tr}(\rho A) \tag{11}$$

With this axiomatization, the concept of probability can mean nothing more than relative frequency in an ensemble. A quantal probability is simply a mean value, in the sense of Axiom III, of an observable A whose eigenvalues are 0 (no) and 1 (yes). We hold that these postulates capture the essence of basic quantum theory, and that they are in fact the ones actually employed in the analysis of empirical findings.

Nevertheless, as is already evident in Section 1, we are not thoroughgoing objectivists concerning every conceivable application of the probability concept. Indeed there are interesting situations within and without physics where a paucity of knowledge should not be construed as an insurmountable obstacle to analysis. In just such cases, the subjectivist interpretation of probability coupled with information theory provides a systematic method for making educated guesses. Statistical physics is a splendid example of the potency of the subjectivist methodology. As we have already explained, when less than a quorum of mean values is known for a given preparation, the density operator ρ cannot be determined; and quantum statistics is born when information theory with its subjective probabilities is used to make the best guess.

This conception of quantum statistics as a blend of objective quantal probabilities and subjective information-theoretic probabilities is not qualitatively new. In the present work, however, the line of demarcation between the two types of probability is not in its orthodox place. Given a density operator displayed in a spectral expansion,

$$\rho = \sum_n w_n \mid \psi_n \rangle\langle \psi_n \mid \tag{12}$$

we do not make the traditional ignorance interpretation that the system is really in one of the states $\{\mid \psi_n\rangle\}$ and that the $\{w_n\}$ are subjective probabilities reflecting a lack of knowledge as to the true state. The irrationality of that interpretation as a general principle has been demonstrated elsewhere.[6] The $\{w_n\}$ do have a probabilistic interpretation, but they are relative frequencies rather than degrees of belief. Consider the projection operator $\mid \psi_m\rangle\langle\psi_m \mid$; its eigenvalues are 0 and 1 and it represents a "yes–no" question concerning a physical system. If that question is posed and answered through measurements performed upon the systems of a quantal ensemble generated by a repeatable preparation characterized by the density operator ρ in (12), then according to quantum mechanics, the quantity

$$\mathrm{Tr}(\mid \psi_m\rangle\langle\psi_m \mid \rho) = w_m \tag{13}$$

will be the relative frequency of "yes" answers. Thus, contrary to a popular though ill-founded custom, we regard all probabilities that may be extracted from the density operator ρ, including the eigenvalues $\{w_n\}$, as objective relative frequencies.

Now, as has been explained in Section 1, if there is a set of density operators $\{\rho_n\}$ but uncertainty as to which one is correct, we may describe this lack of knowledge, or state of ignorance, by defining a subjective probability distribution $\{W_n\}$ over the set $\{\rho_n\}$; and this is equivalent to guessing that the unknown density operator is given by

$$\tilde{\rho} = \sum_n W_n \rho_n \tag{14}$$

Thus in quantum statistics it is possible to use the density operator formalism in a manner which explicitly displays subjective probabilities $\{W_n\}$ arising from ignorance, and keeps them separate from the objective probabilities which inhere in the alternative quantum states $\{\rho_n\}$. The point we wish to emphasize here is that $\{\rho_n\}$ will not generally be a set of orthogonal projectors, and hence the W_n will not generally be the eigenvalues of $\tilde{\rho}$. This guessed density operator $\tilde{\rho}$ does of course have its own spectral expansion of the form (12):

$$\tilde{\rho} = \sum_n \tilde{w}_n \mid \tilde{\psi}_n\rangle\langle\tilde{\psi}_n \mid \tag{15}$$

This expansion, however, does not admit of an ignorance interpretation in which the eigenvalues $\{\tilde{w}_n\}$ would be probabilities for the eigenvector states $\{| \tilde{\psi}_n\rangle\}$, for such an interpretation would contradict the hypothesis from which $\tilde{\rho}$ was derived. The only exception would be the rare situation where $\{\rho_n\}$ happened to be a set of orthogonal projectors, in which case, by coincidence, the $\{\tilde{w}_n\}$ would match the $\{W_n\}$ and the spectral expansion of $\tilde{\rho}$ could be given the old ignorance interpretation.

In general, the only density operator decompositions that can safely be given an ignorance interpretation are those of the form (14) which have arisen in the context of an information-theoretic analysis and explicitly constructed to exhibit the subjective probabilities.

3. QUORUM PARAMETERS AND THE CONTINUOUS CHOICE PROBLEM

The abstract structure of quorum theory[8] may be explained as follows. Let $\mathscr{H}$ denote the Hilbert space associated with a system, and let $\mathscr{L}$ denote the space of linear Hermitian operators whose domain is $\mathscr{H}$. If $\mathscr{H}$ is N-dimensional, then $\mathscr{L}$ is N^2-dimensional. Among the operators of $\mathscr{L}$ are the density operators, defined by the requirements that a density operator must be nonnegative definite and of trace unity. The set of all density operators constitutes a convex region $\mathscr{D}$ within $\mathscr{L}$.

In the operator space $\mathscr{L}$ we may select N^2 linearly independent operators $\{Q_j\}$ to serve as a basis, and then expand other operators as linear combinations of the $\{Q_j\}$. A standard definition for the scalar product of two operators A, B in $\mathscr{L}$ is the trace formula

$$(A \mid B) \equiv \mathrm{Tr}(AB) \tag{16}$$

It is then convenient to say that two Hermitian operators are orthogonal whenever the trace of their product vanishes.

Now, since $\mathscr{D}$ is a region of $\mathscr{L}$, any arbitrary density operator ρ may be expanded in terms of the $\{Q_j\}$:

$$\rho\{q_j\} = \sum_{j=1}^{N^2} q_j Q_j \tag{17}$$

To get a physical interpretation of the real parameters $\{q_j\}$, suppose that measurements of the N^2 observables $\{Q_j\}$ are performed on systems prepared in the manner characterized by ρ. If a collective of data is obtained for each Q_j and the corresponding arithmetic mean value $\bar{Q}_j$ computed, we could then write N^2 relations of the form

$$\left\{\bar{Q}_k = (\rho \mid Q_k) = \sum_{j=1}^{N^2} q_j(Q_j \mid Q_k)\right\} \tag{18}$$

In principle, if ρ were unknown, Eq. (18) could be inverted to determine the $\{q_j\}$ and hence ρ as a function of the data $\{\bar{Q}_k\}$. Because the set of observables $\{Q_j\}$ is just sufficient to determine, through its quantal mean values, an unknown quantum state, such a set is the *quorum* of observables mentioned earlier. In previous publications,[8] we have considered the problem of physically identifying quorum observables; but for the present analysis that practical matter is unimportant.

It is convenient to think of the quorum parameters $\{q_j\}$ as the coordinates of a point in an auxiliary space $\mathscr{L}'$, which might be construed as a quantal "phase space" in that a point $\{q_j\}$ determines via (17) a unique quantum state ρ. There is, however, a flaw in this analogy, due to the requirement that ρ be nonnegative definite. This essential property of ρ imposes very complicated constraints upon the values of the $\{q_j\}$. The additional restriction that ρ have trace unity further constrains the $\{q_j\}$. As a result the quantal "phase space" is only a convex region $\mathscr{D}'$ rather than the whole auxiliary space $\mathscr{L}'$ defined by unrestricted quorum parameters.

Because of this complication, we have yet to discover a really tractable scheme for writing down the full list of elements of $\mathscr{D}$ or $\mathscr{D}'$ which index the propositions of our new logical spectrum for quantum statistics. However, in addition to that technical difficulty, we confront also a more serious dilemma—the domain $\mathscr{D}$ is a continuum.

The continuous choice problem has plagued probability theory for centuries, and our encounter with it here is in no way peculiar to quantum mechanics. Nevertheless it is instructive to look at the difficulty as it appears in the context of quorum theory. Suppose for some reason we have assigned equal subjective probability to every density operator in $\mathscr{D}$ and wish to compute the expection value of an observable A. Expanding A in terms of a quorum $\{Q_j\}$,

$$A = \sum_k a_k Q_k \tag{19}$$

and substituting into (11) and (17), we obtain

$$\bar{A}\{q_j\} = \sum_{jk} q_j a_k (Q_j \mid Q_k) \tag{20}$$

Equation (20) expresses the mean value of observable A as a function of the coordinates $\{q_j\}$ in "quantal phase space" $\mathscr{D}'$. To compute the desired expectation value, we next average (20) over $\mathscr{D}'$:

$$\langle A\rangle = \Big(\int_{\mathscr{D}'} \bar{A}\{q_j\} \prod_j dq_j\Big)\Big/\Big(\int_{\mathscr{D}'} \prod_j dq_j\Big) \tag{21}$$

Unfortunately this value of $\langle A\rangle$ is not really independent of the choice of quorum $\{Q_j\}$ and hence of quorum parameters $\{q_j\}$. In fact even with a fixed quorum $\{Q_j\}$, a convention different from (17) for defining quorum parameters would lead to a different value of $\langle A\rangle$. To see this, consider a general one-to-one coordinate transformation in $\mathscr{L}'$:

$$\{q_j = q_j(\{y_k\});\quad y_k = y_k(\{q_j\})\} \tag{22}$$

Still using the same hypothesis that every density operator in $\mathscr{D}$ is equally probable, we can now recompute $\langle A\rangle$ by averaging $\bar{A}\{q_j(y_k)\}$ over $\mathscr{D}'$ in terms of the new quorum parameters $\{y_k\}$. This procedure gives

$$\langle A\rangle = \Big(\int_{\mathscr{D}'} \bar{A}\{q_j(y_k)\} \prod_k dy_k\Big)\Big/\Big(\int_{\mathscr{D}'} \prod_k dy_k\Big) \tag{23}$$

$$= \Big[\int \bar{A}\{q_j\}\, J\Big(\frac{\{y_k\}}{\{q_j\}}\Big) \prod_j dq_j\Big]\Big/\Big[\int J\Big(\frac{\{y_k\}}{\{q_j\}}\Big) \prod_j dq_j\Big] \tag{24}$$

where J denotes the Jacobian of transformation (22).

Obviously, the values of $\langle A\rangle$ in Eqs. (21) and (24) will in general be unequal, and quantum statistical mechanics thus confronts the unsolved problem of the continuum of alternatives.

The question naturally arises here as to why Gibbs did not run aground at this point in classical statistical mechanics, for there, too, the choice of states is continuous. By reinterpreting the $\{q_j\}$ and $\{y_k\}$ as coordinates in classical phase space, and regarding $\bar{A}$ now as a function of phase representing some classical observable, the formulas (21) and (24) immediately become valid classical expectation values for the analogous equal probability case. Yet in classical statistics there is no ambiguity; the two values for $\langle A\rangle$ are equal. The reason may contain a clue to eventual resolution of the problem for quantum statistics. In classical statistics there is an agreement, sometimes tacit, sometimes explicitly postulated, that only canonical coordinates shall be used in formulating expectation values. Since the Jacobian of any canonical transformation is unity, the classical counterparts to expressions (21) and (24) are identical.

Perhaps in quantum physics there is some similar aesthetic criterion of elegance or symmetry or parsimony that would incline us to prefer one class of quorum parameters over another; at present we cannot imagine what it would be. It may be that the ultimate resolution will be similar to the invariance argument recently used by Jaynes[14] to discuss the classic continuous choice problem due to Bertrand.

4. LAPLACIAN INDIFFERENCE AND INFORMATION-THEORETIC ENTROPY

In normal applications of information theory to lists of discrete alternatives, the familiar information-theoretic entropy function of Eq. (2) is the starting point. However, it should always be borne in mind that the form (2) is derived from abstract premises which include the Laplacian doctrine of insufficient reason, or indifference. According to this principle, if there is no known reason to prefer one alternative to another, the best approach is to regard all alternatives as equally probable. That the principle of indifference is embedded in normal information theory is easily seen by maximizing (2) subject to no constraints other than (1), an exercise which immediately yields equal a priori probability for each alternative.

Thus in statistical mechanics there has always been an axiom asserting that the logical spectrum consists of "a priori equally probable" alternatives. Applied to a discrete list of orthogonal state vectors, this axiom is customarily invoked to justify the use of (2) in the traditional version of quantum statistics. In classical statistics, on the other hand, where the alternatives comprise a continuum, this axiom only makes sense when coupled to the convention noted in the last section which requires the use of canonical coordinates. Without that convention, classical statistics rests upon ambiguous foundations.

In the present development of quantum statistics, we confront a situation where the notion of "a priori equal probability" is ill-defined due to the nondenumerability of the points in $\mathscr{D}'$, and where there is no natural, or "canonical," set of quorum parameters which would eliminate the troublesome Jacobian in (24). We must therefore abandon the principle of Laplacian indifference, and adopt a definition of information-theoretic entropy which does not presuppose that principle. The generalization of (2) needed in continuous choice problems is available in the literature[15–17]; in terms of our quorum parameters it takes the form

$$I \equiv -\kappa \int_{\mathscr{D}'} w\{q_j\} \ln[w\{q_j\}/p\{q_j\}] \prod_j dq_j \tag{25}$$

Here $w\{q_j\}$ and $p\{q_j\}$ are subjective probability densities defined over $\mathscr{D}'$. The function $p\{q_j\}$ is the prior probability density; when I is maximized subject only to the constraint of normalization of $w\{q_j\}$, the result is

$$\hat{w}\{q_j\} = p\{q_j\} \tag{26}$$

The assignment of a prior distribution $p\{q_j\}$ is the analog, in the case of a continuous logical spectrum, to the familiar Laplacian rule of equal a priori probabilities in the discrete alternatives case. Under parameter transformations (22), both $p\{q_j\}$ and $w\{q_j\}$ transform as scalar densities:

$$\begin{aligned} p\{q_j\} &\rightarrow p\{q_j(y_k)\}\, J(\{y_k\}/\{q_j\}) \\ w\{q_j\} &\rightarrow w\{q_j(y_k)\}\, J(\{y_k\}/\{q_j\}) \end{aligned} \tag{27}$$

The expression (25) for I is therefore invariant under a change of quorum parameters.

The ambiguity displayed by (21) and (24) would therefore be resolved if $p\{q_j\}$ were given. Thus the theoretical impasse arising from the lack of a criterion for choosing quorum parameters is now reduced to the problem of finding some rational metaphysical principle for the selection of a prior probability distribution $p\{q_j\}$ appropriate for quantum statistics. We shall explore this vexing question further in a future paper devoted to the case of thermal equilibrium.

Our present objective of establishing a format for the correct application of information theory to quantum mechanical situations has now been achieved: the missing information function I in the form (25) is to be maximized subject to whatever physical constraints define the empirical situation, and the best-guess quantum state $\hat{\rho}$ is then given by

$$\hat{\rho} = \int_{\mathscr{D}'} \hat{w}\{q_j\}\, \rho\{q_j\} \prod_j dq_j \tag{28}$$

where $\hat{w}\{q_j\}$ maximizes (25) and $\rho\{q_j\}$ is defined by (17). In the aforementioned sequel on equilibrium statistical thermodynamics, we shall apply this method to obtain a new and quantally correct derivation of the canonical density operator.

REFERENCES

1. E. T. Jaynes, *Phys. Rev.* **108**, 171 (1957).
2. A. Katz, *Principles of Statistical Mechanics* (W. H. Freeman, San Francisco, 1967).
3. A. Hobson, *Concepts in Statistical Mechanics* (Gordon and Breach, New York, 1971).
4. J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, transl. by R. T. Beyer (Princeton U. P., Princeton, N.J., 1955), p. 379.
5. J. W. Gibbs, *Elementary Principles in Statistical Mechanics* (1901), in *The Collected Works*, Vol. 2 (Longmans, Green, and Co., New York, 1928).
6. J. L. Park and W. Band, *Found. Phys.* **6**, 157 (1976).
7. E. T. Jaynes, in *Delaware Seminar in the Foundations of Physics*, ed. by Mario Bunge (Springer-Verlag, New York, 1967), p. 77.
8. J. L. Park and W. Band, *Found. Phys.* **1**, 211 (1971); **1**, 339 (1971).

9. W. Heisenberg, *Physics and Philosophy* (Harper and Row, New York, 1958).
10. H. Margenau, *The Nature of Physical Reality* (McGraw-Hill, New York, 1950).
11. K. R. Popper, in *Quantum Theory and Reality*, ed. by Mario Bunge (Springer-Verlag, New York, 1967).
12. A. Landé, *New Foundations of Quantum Mechanics* (Cambridge U. P., Cambridge, 1965); *Quantum Mechanics in a New Key* (Exposition Press, New York, 1973).
13. H. Margenau and J. L. Park, *Found. Phys.* **3**, 19 (1973).
14. E. T. Jaynes, *Found. Phys.* **3**, 477 (1973).
15. E. T. Jaynes, *IEEE Trans. Systems Sci. Cybernetics* **SSC-4**, 227 (1968).
16. S. Watanabe, *Knowing and Guessing* (Wiley, New York, 1969).
17. A. Hobson, *J. Stat. Phys.* **1**, 383 (1969).