Journal of Statistical Physics, Vol. 10, No. 4, 1974

Foundations of the Quantum Statistics of Systems in Equilibrium: A Measure-Theoretic Approach

Garrison Sposito[1]

Received March 16, 1973; revised August 8, 1973

The fundamental equations of equilibrium quantum statistical mechanics are derived in the context of a measure-theoretic approach to the quantum mechanical ergodic problem. The method employed is an extension, to quantum mechanical systems, of the techniques developed by R. M. Lewis for establishing the foundations of classical statistical mechanics. The existence of a complete set of commuting observables is assumed, but no reference is made a priori to probability or statistical ensembles. Expressions for infinite-time averages in the microcanonical, canonical, and grand canonical ensembles are developed which reduce to conventional quantum statistical mechanics for systems in equilibrium when the total energy is the only conserved quantity. No attempt is made to extend the formalism at this time to deal with the difficult problem of the approach to equilibrium.

KEY WORDS: Foundations; ergodic theory; quantum statistics; measure theory; complete set of observables.

1. INTRODUCTION

In this paper an approach to the foundations of equilibrium quantum statistical mechanics is presented. The results obtained are a refinement and

[1] Department of Physics, California State College, Rohnert Park, California.

extension of some earlier work[1] on the ergodic problem in quantum mechanics and are established in close correspondence with the elegant and general formulation of classical statistical mechanics given by Lewis.[2],2 Thus the point of view taken here is fundamental rather than pragmatic and holds to the basic tenet that the purpose of statistical mechanics, for physical systems in states of equilibrium, is to calculate time averages of experimentally measurable properties. For this reason no ideas of ensembles or probability are introduced a priori. Instead the ensemble theory appears solely as a useful mathematical device which evolves from a restricted, but physically important solution of the ergodic problem and which permits the evaluation of time averages without previous integration of the quantum mechanical equation of motion.

In attacking the nontrivial problem of finding correspondents of Lewis' theorems, we have been led to develop a mathematical structure and physical perspective which are significantly different from what has been considered previously for quantum mechanical systems.[4,5] The primitive notion is that of *state*, a ray in an appropriate Hilbert space, rather than *observable*, and the concomitant mathematical apparatus is measure theory[6] rather than operator algebra. This basic tack makes possible the development of equilibrium quantum statistical mechanics along lines of reasoning which are parallel with those employed for classical mechanical systems, after due attention has been paid to the differences between quantum and classical dynamics. A possible additional benefit of this method is that the purely dynamical analysis is not a sample of the theory of Banach algebras, but remains a part of the conventional mathematical structure of quantum mechanics.[7] It should be emphasized, however, that the operator-algebra approach can offer a valid alternative to our method; in particular the recent work of Moyal[8] achieves some results which are close in spirit to those we present in Section 2.

In the sections which follow we shall present a solution of the ergodic problem in quantum mechanics and give derivations of the principal ensemble distributions in statistical mechanics. The equations for the latter reproduce conventional expressions in the density matrix formulation of quantum statistical mechanics whenever the complete set of observables for a physical system comprises only the total energy. Otherwise the equations are valid quite generally for any finite complete set and, because of a convenient vector notation, can be made to appear formally similar to the conventional expressions.

It should not be forgotten that the results to be obtained here can apply strictly only to quantum mechanical systems containing a large but *finite*

2 A summary of Lewis' paper is given by Truesdell.[3]

number N of interacting particles included in a macroscopic but *bounded* volume V of physical space. This restriction is a necessary prerequisite to the use of a density matrix formalism and is essential as well if we are to eschew the nontrivial mathematical questions which pertain to systems conceived from the very beginning to be in the thermodynamic limit.[9] On the other hand, we should not expect the results we obtain to have a strict physical significance in thermodynamics without some consideration of their behavior as $N \to \infty$, $V \to \infty$, with N/V finite. The point of view to be taken here is in agreement with common practice and with the general procedure suggested by Emch[9]: The ergodic averages of observables are calculated for finite N and V, then the thermodynamic limit is taken and the existence of the ensemble distributions is ascertained. This method appears to serve quite well in a general study of equilibrium phenomena, although it may in fact not be adequate for certain broader investigations concerned with the explicit irreversible behavior of quantum systems. For these studies physical conditions may dictate that time averages of observables be calculated either simultaneously with the passage to the thermodynamic limit, or afterward. Therefore at present the point of view in this paper should be considered judiciously with respect to its possible application in other areas of quantum ergodic theory.

With regard to the latter parts of our program, we can point out immediately that the rigorous, positive results presented by Ruelle[10] on the thermodynamic limit for quantum ensembles apply as well to the generalized ensembles considered in this paper with only slight modifications. In particular, it is required that the complete invariant vector rather than just the Hamiltonian operator for an N-particle system be stable (in the sense of Ruelle), and that the particle interactions be tempered (again as defined by Ruelle), and that the microcanonical distribution deal not with trace pairs whose image under the Hamiltonian operator is stationary, but rather with those whose image under the complete invariant vector is stationary. These conditions will be seen to be quite straightforward to impose on our results. We shall always assume that they hold, with no loss of generality in our arguments, in order that the expressions we derive will have meaning for infinite quantum mechanical systems. For all the well-known mathematical details of the passage from the finite N, V case we consider here to the thermodynamic limit, the clear exposition in Ruelle's book[10] should be consulted.

2. ERGODIC THEOREM

An assembly of N interacting particles is associated in quantum mechanics[7] with a complex, separable Hilbert space $\mathscr{H}$ which comprises square-summable state vectors $\psi(\mathbf{r}^N, s^N)$ with support in R^{4N}. Each one of the N

pairs $(\mathbf{r}, s)$ in the argument of ψ locates a particle in physical space and prescribes a value of its spin projection, respectively. If the assembly consists of identical particles, a closed subspace of $\mathscr{H}$ must be chosen to represent the totality of possible states. This subspace will contain only eigenvectors of the transposition operator[11]

$$P_{(ij)}\psi(..., \mathbf{r}_i, s_i,..., \mathbf{r}_j, s_j,...) \equiv \psi(..., \mathbf{r}_j, s_j,..., \mathbf{r}_i, s_i,...)$$

belonging to the eigenvalue $+1$ or to the eigenvalue -1.

Strictly, the Hilbert space of quantum mechanical states is a family of equivalence classes, with equivalence being "equality almost everywhere" on R^{4N}. Besides this mathematical subtlety there is the important and well-known (see, e.g., Ref. 12) *physical* equivalence relation defined by

$$\psi \sim \varphi \qquad \text{if} \qquad |(\psi, \varphi)| = \|\phi\| \, \|\varphi\|, \qquad \varphi \in \mathscr{H}, \qquad \psi \in \mathscr{H} \tag{1}$$

for any pair of nonzero vectors in $\mathscr{H}$. Relation (1) partitions $\mathscr{H}$ into disjoint sets called *rays*. At any chosen instant of time a state of an assembly of interacting particles is represented uniquely by a ray in $\mathscr{H}$. We shall denote the ray containing the state vector ψ by the symbol Ψ. The class of all rays in $\mathscr{H}$ (or in a suitable closed subspace of $\mathscr{H}$) will be denoted $\mathscr{R}$.

Now let us put $\mathscr{R}' \equiv \varnothing \cup \mathscr{R}$, where $\varnothing$ is the empty set, and consider the class $\mathscr{A}$ of all subsets of $\mathscr{R}'$. $\mathscr{A}$ surely contains the empty set and is closed under complementation and the taking of countable unions. Thus $\mathscr{A}$ is a σ-algebra and, by definition, the pair $(\mathscr{R}', \mathscr{A})$ is a measurable space. It is a simple matter to show further that the pair $(\Omega, \mathscr{M}) \equiv (\mathscr{R}' \times \mathscr{R}', \mathscr{A} \times \mathscr{A})$ is also a measurable space (Ref. 6, §33, Theorem F). This space is to be the correspondent of the pair $(\Pi, \mathscr{L})$ which appears in classical statistical mechanics,[2] where Π is phase space and $\mathscr{L}$ is the class of Lebesgue-measurable sets in Π. The correspondent of Lebesgue measure on Π (the Liouville measure) can be presented after stating the following definition.

Definition 1. A nonempty element of Ω, (Ψ, Φ), is called a *trace pair* if, for any $\psi \in \Psi$ and any $\varphi \in \Phi$, we have $\psi \sim \varphi$.

Let μ be a function from the elements of $\mathscr{M}$ into the set of extended non-negative integers such that

$$\mu(A \times B) = \begin{cases} \text{cardinality of the subset of trace pairs in} \\ A \times B, \text{ if this subset is finite} \\ +\infty \text{ if the subset of trace pairs is infinite} \end{cases} \tag{2}$$

where $A \times B \in \mathscr{M}$. It is evident that $\mu \geqslant 0$ and $\mu(\varnothing) = 0$. Moreover, if $\bigcap_{n=1}^{\infty} (A \times B)_n = \varnothing$ for some sequence of rectangles in $\mathscr{M}$, then the rule for adding integers specifies that

$$\mu\left(\bigcup_{n=1}^{\infty} (A \times B)_n\right) = \sum_{n=1}^{\infty} \mu((A \times B)_n)$$

If one of the rectangles in the sequence contains an infinite subset of trace pairs, the above equality is necessarily satisfied. It follows that μ is a measure function on Ω and that $(\Omega, \mathscr{M}, \mu)$ is a measure space.

The triple $(\Omega, \mathscr{M}, \mu)$ is the correspondent of the measure space $(\Pi, \mathscr{L}, \zeta)$, where ζ is Lebesgue measure. The measured quantities in classical mechanics are represented by real-valued, measurable functions on $(\Pi, \mathscr{L}, \zeta)$ which are called "phase functions." The correspondent of the phase function on Π is the Hilbert function[1] on Ω. Similarly to the phase function, the Hilbert function in a given case is restricted to a subset of Ω established in close relation with the physical character of the assembly of particles which it is to describe. This character is determined by the Hamiltonian operator H_{op} which, through its spectral decomposition, induces a convenient partition of $\mathscr{H}$ into closed subspaces, generating an orthogonal basis for $\mathscr{H}$. This basis is in one-to-one correspondence with a certain subset of $\mathscr{R}$ which will be denoted $\mathscr{R}_{\text{H}}$. Thus $\mathscr{R}_{\text{H}} = \{\Psi_\lambda\}$, where Ψ_λ is the equivalence class belonging to the energy eigenvalue E_λ of H_{op}.[3] Bearing this in mind, we state the following.

Definition 2. A *Hilbert function* is an elementary function from the squares $\mathscr{R}_{\text{H}} \times \mathscr{R}_{\text{H}} \equiv \Omega_{\text{H}}$ into C^1, defined by

$$f_0((\Psi, \Phi)) = \sum_{\{k\}} O_k \chi_{M_k}((\Psi, \Phi)) \tag{3}$$

In Eq. (3)

$$\chi_{M_k}((\Psi, \Phi)) = \begin{cases} 0, & (\Psi, \Phi) \notin M_k \\ 1, & (\Psi, \Phi) \in M_k \subset \Omega_{\text{H}} \end{cases}$$

$$M_k = \{(\Psi_\lambda, \Psi_{\lambda'}) \in \Omega_{\text{H}}:\ f_0((\Psi_\lambda, \Psi_{\lambda'})) = O_k\}$$

$$O_k = [(\psi_\lambda, O_{\text{op}}\psi_{\lambda'})/\|\psi_\lambda\|\ \|\psi_{\lambda'}\|]_k, \qquad \psi_\lambda \in \Psi_\lambda, \quad \psi_{\lambda'} \in \Psi_{\lambda'}$$

and the index set $\{k\}$ identifies all distinct values of the Hilbert function f_0. The operator O_{op} is any quantum mechanical observable (including H_{op}) which has physical meaning for the system of particles under consideration. The complex-valued quantity O_k is computed with the help of elements of the orthogonal basis set generated by H_{op}. Thus the image of Ω_{H} under f_0 is included in the set of all matrix elements of O_{op} relative to the basis $\{\psi_\lambda\}$.

Every Hilbert function can be associated with a continuous parameter t, which represents the time variable, by specifying that

$$\psi_{\lambda t} = U_t \psi_\lambda, \qquad \psi_\lambda \in \Psi_\lambda \in \mathscr{R}_{\text{H}}, \quad t \in R^1 \tag{4}$$

[3] We do not assume that the E_λ are all distinct.

where U_t is a strongly continuous, unitary operator whose Hermitian generator is H_{op}. A natural generalization of Eq. (4) to the elements of Ω_{H} is obtained by writing

$$T_t(\Psi_\lambda, \Psi_{\lambda'}) \equiv (\{U_t\psi_\lambda\}, \{U_t\psi_{\lambda'}\}) \tag{5}$$

The set $\{T_t\}$ is, then, a single-parameter group of transformations from $\mathscr{R} \times \mathscr{R}$ onto itself and serves as the correspondent of the "solution operator" which maps phase space into itself in classical mechanics.[2] The existence of T_t and its group properties is guaranteed by the axioms of Schrödinger quantum mechanics (Ref. 7, Chapter IV). This, in turn, makes possible the following important lemma.

Lemma 1. $(\Omega, \mathscr{M}, \mu, T_t)$ is a measure-preserving space.

Proof. 1. For every fixed t, T_t^{-1} exists and $T_t^{-1}M = \{(\Psi, \Phi)\colon T_t(\Psi, \Phi) \in M\}$ is included in $\mathscr{M}$ if $M \in \mathscr{M}$ because T_t has the group property and is onto.

2. The measure μ is invariant under the transformations T_t. This assertion is true fundamentally because the axioms of Schrödinger quantum mechanics establish the existence of the unitary time-evolution operator U_t which necessarily preserves the scalar product on $\mathscr{H} \times \mathscr{H}$. Therefore $T_t(\Psi, \Phi)$ is a trace pair if and only if (Ψ, Φ) is a trace pair and the number of trace pairs must remain invariant under T_t.[4]

3. $T_t(\Psi, \Phi)$ is a measurable transformation from $(R^1 \times \Omega, L \times \mathscr{M})$ into $(\Omega, \mathscr{M})$, where L is the class of Lebesgue-measurable sets in R^1. Since $T_t(\Psi, \Phi)$ is a strongly continuous transformation from $(R^1 \times \Omega)$ onto Ω, it certainly is measurable from $(R^1 \times \Omega, \mathscr{B} \times \mathscr{M}, \zeta_{\mathrm{B}} \times \mu)$ into $(\Omega, \mathscr{M}, \mu)$, where $\mathscr{B} \subset L$ is the class of Borel-measurable sets in R^1 and ζ_{B} is the restriction of Lebesgue measure on the real line to $\mathscr{B}$ (Ref. 7, p. 81). Now consider the set

$$[T^{-1}(M)]_t \equiv [(T_t(\Psi, \Phi))^{-1}]_t = \{(\Psi, \Phi)\colon \quad T_t(\Psi, \Phi) \in M \in \mathscr{M}\}$$

which is the section (Ref. 6, §34), determined by t, of the set $T^{-1}(M) \subset R^1 \times \Omega$. Because $[T^{-1}(M)]_t = T_t^{-1}(M)$ and μ is invariant under T_t, every section has μ measure zero if $\mu(M) = 0$. It follows (Ref. 6, §34) that $\zeta_{\mathrm{B}} \times \mu(M) = 0$ and therefore[2] that $T_t(\Psi, \Phi)$ is a measurable transformation.

The three foregoing statements show that T_t is measure-preserving and measurable, which proves the lemma.

The ideas in this section have been introduced so as to prepare the way for an ergodic theorem. It is accepted at the outset that such a theorem is

[4] This result is the correspondent of the well-known invariance of Liouville measure under the classical mechanical solution operator (Liouville's theorem).

relevant to experiment in two ways. First, because the process of measurement does not produce instantaneous values of observables, but rather time-averaged ones, it is germane to equilibrium phenomena to study the infinite-time average

$$\overline{f_0((\Psi, \Phi))}^{\infty} \equiv \lim_{T\to\infty} (1/T) \int_0^T f_0(T_t(\Psi, \Phi))\, dt \tag{6}$$

Since the values of a Hilbert function are always of the form

$$\frac{(\psi_{\lambda t}, O_{\mathrm{op}}\psi_{\lambda' t})}{\|\psi_{\lambda t}\|\ \|\psi_{\lambda' t}\|} = \frac{(\psi_{\lambda}, O_{\mathrm{op}}\psi_{\lambda'})}{\|\psi_{\lambda}\|\ \|\psi_{\lambda'}\|} \exp \frac{-i(E_{\lambda'} - E_{\lambda})t}{\hbar}$$

where $\hbar$ is Dirac's constant, they are strictly periodic in the time and are easily integrated as in Eq. (6). The result is

$$\overline{f_0((\Psi, \Phi))}^{\infty} = \sum_{\{j\}} \bar{O}_j{}^{\infty} \chi_{\bar{M}_j}((\Psi, \Phi)) \tag{7}$$

where

$$\bar{O}_j{}^{\infty} = [(\psi_l{}^m, O_{\mathrm{op}}\psi_l{}^{m'})/\|\psi_l{}^m\|\ \|\psi_l{}^{m'}\|]_j, \qquad m, m' = 1,\ldots, L_l \tag{8}$$

$$\bar{M}_j = \{(\Psi_l{}^m, \Psi_l{}^{m'})\colon\ f_{\mathrm{H}}((\Psi_l{}^m, \Psi_l{}^{m'})) = E_l, f_0((\Psi_l{}^m, \Psi_l{}^{m'})) = \bar{O}_j{}^{\infty}\} \tag{9}$$

and the index set $\{j\}$ labels distinct values of the time average. Second, it is clear from Eqs. (8) and (9) that

$$T_t\overline{f_0((\Psi, \Phi))}^{\infty} \equiv \overline{f_0(T_t(\Psi, \Phi))}^{\infty} = \overline{f_0((\Psi, \Phi))}^{\infty}$$

Some of these invariant values of the infinite-time average are complex. However, they have no bearing on experiment since they constitute the image of a set of μ-measure zero.[5] This fact leads further to the conclusion that

$$\begin{aligned}\int_M \overline{f_0((\Psi, \Phi))}^{\infty}\, d\mu &= \sum_{\{j\}} \bar{O}_j{}^{\infty} \mu(M \cap \bar{M}_j) = \sum_{\{k\}} O_k \mu(M \cap M_k) \\ &= \int_M f_0((\Psi, \Phi))\, d\mu \end{aligned} \tag{10}$$

where M is any element of $\mathscr{M}$ having finite μ measure. Equations (7) and (10) are the principal results of this section, which we may state as the following:

[5] It is important to note here the often overlooked point that *any* solution of the ergodic problem will contain a residual statistical aspect; namely that the sets of zero measure are of no physical significance. This fact has been discussed in the context of classical dynamics by Jancel (Ref. 5, p. 15).

Ergodic Theorem. Let $(\Omega, \mathscr{M}, \mu, T_t)$ be a measure-preserving space and let $f_0((\Upsilon, \Phi))$ be a Hilbert function on $(\Omega, \mathscr{M})$. Then the infinite-time average

$$\overline{f_0((\Upsilon, \Phi))}^\infty = \operatorname*{Lim}_{T\to\infty} (1/T) \int_0^T f_0(T_t(\Upsilon, \Phi))\, dt$$

exists, is real-valued, and is invariant under T_t except on sets of μ-measure zero. If $\mu(M)$ $(M \in \mathscr{M})$ is finite, then f_0 is integrable and

$$\int_M \bar{f}_0^\infty\, d\mu = \int_M f_0\, d\mu$$

3. THE MICROCANONICAL DISTRIBUTION

A fundamental assumption in quantum mechanics is that, for any isolated assembly of interacting particles, there exists a finite set of commuting observables, called a complete set, in terms of which every compatible observable for the assembly can be expressed as a function (Ref. 7, Chapter IV, §5). This important notion may be brought into the present context through the following definition.

Definition 3. If $(\Omega, \mathscr{M}, \mu, T_t)$ is a measure-preserving space and $C_1,..., C_n$ is a set of measurable Hilbert functions which are invariant under T_t, then

$$\mathbf{C}((\Upsilon, \Phi)) = \{C_1,..., C_n\}$$

is a *complete invariant vector* if every measurable invariant Hilbert function is a measurable function from $(R^n, \mathscr{B}^n)$ to (R^1, B) of $C_1,..., C_n$ except on sets of μ-measure zero.

We shall henceforth assume that $\mathbf{C}$ has been determined for any physical system of interest. Then, noting that the infinite-time average of a Hilbert function is an invariant, we may invoke Definition 3 and write

$$\overline{f_0((\Upsilon, \Phi))}^\infty = F[\mathbf{C}((\Upsilon, \Phi))] \tag{11}$$

on sets of finite μ-measure, where F is a Borel-measurable function from $(R^n, \mathscr{B}^n)$ to (R^1, B). It follows from Eq. (10) that

$$\int_{\mathbf{C}^{-1}(B)} F[\mathbf{C}((\Upsilon, \Phi))]\, d\mu = \int_{\mathbf{C}^{-1}(B)} f_0((\Upsilon, \Phi))\, d\mu \tag{12}$$

where $B \in \mathscr{B}$ and $\mu(\mathbf{C}^{-1}(B)) < \infty$. If we now put

$$\nu(B) \equiv \mu(\mathbf{C}^{-1}(B)) \tag{13}$$

as the measure induced in R^n by μ on Ω, we have

$$\int_{\mathbf{C}^{-1}(B)} F[\mathbf{C}((\Upsilon, \Phi))]\, d\mu = \int_B F[\mathbf{C}]\, d\nu \tag{14}$$

since F is Borel measurable. Therefore, by Eqs. (12) and (14),

$$\int_B F[\mathbf{C}]\, d\nu = \int_{\mathbf{C}^{-1}(B)} f_0((\Psi, \Phi))\, d\mu \tag{15}$$

This expression may be used to obtain an equation for $F[\mathbf{C}]$ (and therefore $\bar{f}_0{}^{\infty}$) by the application of a result from the theory of nets.[2] Let

$$I_\delta[\mathbf{K}] = \{\mathbf{C}: \quad |C_i - K_i| \leqslant \delta_i\}, \qquad i = 1,..., n$$

where $\mathbf{K} = \{K_1,..., K_n\} \in R^n$, $\boldsymbol{\delta} = \{\delta_1,..., \delta_n\}$ is a very small interval in R^n, and only the real values of the C_i are considered. Then

$$F[\mathbf{K}] = \lim_{\delta \to 0} \left[\int_{I\boldsymbol{\delta}[\mathbf{K}]} F\, d\nu \Big/ \nu(I_{\boldsymbol{\delta}}[\mathbf{K}]) \right] \tag{16}$$

Physically $|\boldsymbol{\delta}|$ should exhibit a nonzero least value because of uncertainty relations (e.g., the energy–time uncertainty relation). However, we may consider Eq. (16) to apply in a limiting sense, just as is done in classical mechanics,[2] and derive from Eqs. (11), (13), and (15) the result

$$\overline{f_0(C^{-1}(K))}^{\infty} = \lim_{\delta \to 0} \left[\int_{M\boldsymbol{\delta}[\mathbf{K}]} f_0\, d\mu \Big/ \mu(M_{\boldsymbol{\delta}}[\mathbf{K}]) \right] \tag{17}$$

except on sets of μ-measure zero, where

$$\mathbf{C}^{-1}(\mathbf{K}) = \{(\Psi, \Psi): \quad \mathbf{C}(\Psi, \Psi) = \mathbf{K}\}$$

and

$$M_\delta[\mathbf{K}] = \{(\Psi, \Psi): \quad |C_i(\Psi, \Psi) - K_i| \leqslant \delta_i\}, \qquad i = 1,..., n$$

Equation (17) is a generalization of the microcanonical distribution in quantum statistical mechanics. It states that the infinite-time average of a Hilbert function, on sets of nonvanishing μ-measure, may be calculated as an average over an ensemble in which equal a priori probability has been assigned to every trace pair whose image under $\mathbf{C}$ lies in a vanishingly small interval about the value $\mathbf{K}$. Under the condition that $\mathbf{C} = f_{\mathrm{H}}$, the total energy, Eq. (17) reduces to the conventional microcanonical distribution. Whenever the complete invariant vector $\mathbf{C}$ possesses a classical mechanical analog Eq. (17) becomes, in the classical limit, Lewis' generalization of the microcanonical distribution,[2] as has been shown previously.[1]

4. THE CANONICAL DISTRIBUTION

A consideration of assemblies of interacting particles which are not isolated from one another poses essentially only problems of a technical nature in extending the results of Section 3. To begin a careful formulation of the theory of the canonical ensemble, we name the quintuple $(\Omega, \mathscr{M}, \mu, T_t, \mathbf{C})$ a *complete space*, following Lewis,[2] and prescribe the following.

Definition 4. The complete spaces $(\Omega', \mathscr{M}', \mu', T_t', \mathbf{C}')$ and $(\Omega'', \mathscr{M}'', \mu'', T_t'', \mathbf{C}'')$ are in *weak interaction* if there exists a complete space $(\Omega, \mathscr{M}, \mu, T_t, \mathbf{C})$ such that $\Omega = \Omega' \times \Omega''$, $\mathscr{M} = \mathscr{M}' \times \mathscr{M}''$, $\mu = \mu' \times \mu''$, and

$$\mathbf{C}((\Psi, \Phi)) \equiv \mathbf{C}((\Psi', \Phi', \Psi'', \Phi'')) = \mathbf{C}'((\Psi', \Phi')) + \mathbf{C}''((\Psi'', \Phi''))$$

In the special case that $\mathbf{C}'$ and $\mathbf{C}''$ each is simply the energy of the respective system of particles it describes, the definition of "weak interaction" above is quite the same as what is used conventionally to establish the relation between two systems in purely thermal contact (see, e.g., Ref. 13, Chapter 4). Generally, of course, one of these two systems must be a heat reservoir if the concept of temperature is to be given a precise meaning. This requirement suggests the need for the following definition.

Definition 5. $(\Omega'', \mathscr{M}'', \mu'', T_t'', \mathbf{C}'') \gg (\Omega', \mathscr{M}', \mu'. T_t', \mathbf{C}')$ if the two complete spaces are in weak interaction and if:

(a) There exist measure functions p' and p'' such that $(\Omega', \mathscr{M}', p', T_t', \mathbf{C}')$ and $(\Omega'', \mathscr{M}'', p'', T_t'', \mathbf{C}'')$ are also complete spaces in weak interaction, with $p'(\Omega_{\mathrm{H}}') = p''(\Omega_{\mathrm{H}}'') = 1$.
(b) There exists a vector $\mathbf{E}$ in R^n such that, except for sets in Ω_{H}' of p'-measure zero,

$$p''\{(\Psi'', \Phi''): \quad |C_i''((\Psi'', \Phi'')) + C_i' - E_i| \leqslant \delta_i\} = \text{const} \tag{18}$$

for sufficiently small $|\boldsymbol{\delta}|$.

According to Definition 5, a reservoir has the distinguishing characteristics of being in weak interaction with another system and of being "large" by comparison with that system in the sense that, for a fixed total invariant vector, the (probability) measure of the states of the reservoir is constant, regardless of what may be the state of the adjacent system. These features are consonant with the usual physical concept of a reservoir[13] and they provide for a straightforward application of Eq. (17) to Hilbert functions on the product space $(\Omega, \mathscr{M}, p, T_t, \mathbf{C})$, where $p = p' \times p''$. If we call the constant value of p'' in Eq. (18) $p''[\mathbf{C}'((\Psi', \Phi'))]$ for the sake of clarity, we have

$$\overline{f_0(\mathbf{C}^{-1}(\mathbf{E}))^\infty} = \lim_{\delta \to 0} \frac{\int f_0((\Psi', \Phi'))p''[\mathbf{C}'((\Psi', \Phi'))]\, dp'}{p(M_\delta[\mathbf{E}])} \tag{19}$$

by Eq. (17), where f_0 is a Hilbert function on Ω_{H}' alone,

$$M_\delta[\mathbf{E}] = \{(\Psi, \Psi): \quad |C_i - E_i| \leqslant \delta_i\}, \qquad i = 1,..., n$$

and the integral over the appropriate set in Ω_{H}'' has been carried out in the numerator of Eq. (17). It is to be understood that for small enough $|\boldsymbol{\delta}|$ the integral remaining in Eq. (19) is over the entirety of Ω_{H}'. It follows that $p''[\mathbf{C}'((\Psi', \Phi'))] = p$, then, and (19) becomes

$$\overline{f_0(\mathbf{C}^{-1}(\mathbf{E}))}^{\infty} = \int_{\Omega_{\mathrm{H}}'} f_0((\Psi', \Phi'))\, dp' \tag{20}$$

This expression provides for the calculation of the infinite-time average of a Hilbert function, for a system in contact with a reservoir, on the points (Ψ, Φ) for which $\mathbf{C}((\Psi, \Phi)) = \mathbf{E}$.

Now we must consider the problem of finding the explicit forms of the measures p' and p''. It is evident that Eq. (20) describes a generalization of what one expects from a canonical distribution and therefore that p' ultimately should be proportional to a generalization of the Boltzmann factor. To see that this indeed is the case, we may begin by citing a lemma due to Lewis[2]:

Lemma 2. Let $\{T_t\}$ be a family of measure-preserving transformations on an arbitrary measure space $(\Omega, \mathscr{M}, \mu)$. Let $\varphi(\gamma)$ be nonnegative and integrable on $(\Omega, \mathscr{M}, \mu)$. For every $M \in \mathscr{M}$ let

$$q(M) = \int_M \varphi(\gamma)\, d\mu$$

be a measure on $(\Omega, \mathscr{M})$. Then $\{T_t\}$ is also a family of measure-preserving transformations on $(\Omega, \mathscr{M}, q)$ if and only if φ is invariant under T_t on $(\Omega, \mathscr{M}, \mu)$.

It follows from Lemma 2 that if $(\Omega', \mathscr{M}', \mu', T_t', \mathbf{C}')$ and $(\Omega'', \mathscr{M}'', \mu'', T_i'', \mathbf{C}'')$ are complete spaces, the spaces $(\Omega', \mathscr{M}', q', T_t', \mathbf{C}')$ and $(\Omega'', \mathscr{M}'', q'', T_t'', \mathbf{C}'')$ likewise will be complete if and only if φ' and φ'', respectively, are invariant Hilbert functions on $(\Omega', \mathscr{M}', \mu')$ and $(\Omega'', \mathscr{M}'', \mu'')$ (except on sets of μ-measure zero). But this means

$$\varphi'((\Psi', \Phi')) = F'[\mathbf{C}'((\Psi', \Phi'))]$$
$$\varphi''((\Psi'', \Phi'')) = F''[\mathbf{C}''((\Psi'', \Phi''))]$$

where the right-hand sides are Borel-measurable functions on R^n. Therefore

$$\begin{aligned} q'(M')q''(M'') &= \int_{M'} F'[\mathbf{C}']\, d\mu' \int_{M''} F''[\mathbf{C}'']\, d\mu'' \\ &= \int_{M' \times M''} F'[\mathbf{C}']F''[\mathbf{C}'']\, d\mu \\ &= q(M' \times M'') = q(M) \end{aligned} \tag{21}$$

where μ is the product measure on the product space $(\Omega, \mathscr{M}) = (\Omega' \times \Omega'', \mathscr{M}' \times \mathscr{M}'')$ and $M' \times M''$ is any rectangle in $\Omega_H \equiv \Omega_H' \times \Omega_H''$ for which $q(M)$ is finite. Since the two systems do interact,[6] the density $\varphi((\Psi, \Phi))$ is the same function of $\mathbf{C}$ as are φ' and φ'' of $\mathbf{C}'$ and $\mathbf{C}''$:

$$q(M) = \int_M F[\mathbf{C}] d\mu \tag{22}$$

The necessary and sufficient condition that $(\Omega, \mathscr{M}, q, T_t, \mathbf{C})$ be a complete space comes from Eqs. (21) and (22) as

$$F'[\mathbf{C}'] \cdot F''[\mathbf{C}''] = F[\mathbf{C}] = F[\mathbf{C}' + \mathbf{C}'']$$

The most general solution of this equation is

$$F'[\mathbf{C}'] = a' \exp(-\boldsymbol{\beta} \cdot \mathbf{C}') \tag{23}$$

$$F''[\mathbf{C}''] = a'' \exp(-\boldsymbol{\beta} \cdot \mathbf{C}''), \qquad F[\mathbf{C}] = a'a'' \exp(-\boldsymbol{\beta} \cdot \mathbf{C})$$

where $\boldsymbol{\beta}$ is an arbitrary vector in R^n and the a's are arbitrary, nonnegative numbers. Now let us define

$$Z'(\boldsymbol{\beta}) \equiv \int_{\Omega_H'} \exp(-\boldsymbol{\beta} \cdot \mathbf{C}')\, d\mu'$$

$$Z''(\boldsymbol{\beta}) \equiv \int_{\Omega_H''} \exp(-\boldsymbol{\beta} \cdot \mathbf{C}'')\, d\mu''$$

on the subsets D' and D'', respectively, in R^n containing all the $\boldsymbol{\beta}$ for which the integrals above are finite. If we set

$$\begin{aligned} p'(M') &\equiv \frac{1}{Z'(\boldsymbol{\beta})} \int_{M'} \exp(-\boldsymbol{\beta} \cdot \mathbf{C}')\, d\mu', \qquad M' \subset \Omega_H' \\ p''(M'') &\equiv \frac{1}{Z''(\boldsymbol{\beta})} \int_{M''} \exp(-\boldsymbol{\beta} \cdot \mathbf{C}'')\, d\mu'', \qquad M'' \subset \Omega_H \end{aligned} \tag{24}$$

we have constructed measures which maintain $(\Omega', \mathscr{M}', p'. T_t', \mathbf{C}')$ and $(\Omega'', \mathscr{M}'', p'', T_t'', \mathbf{C}'')$ as complete spaces in weak interaction. Moreover, if $D'' \supset D' \neq \varnothing$, we can write

$$Z(\boldsymbol{\beta}) \equiv Z'(\boldsymbol{\beta}) Z''(\boldsymbol{\beta}) = \int_{\Omega_H} \exp(-\boldsymbol{\beta} \cdot \mathbf{C})\, d\mu$$

and

$$p(M) = [1/Z(\boldsymbol{\beta})] \int_M \exp(-\boldsymbol{\beta} \cdot \mathbf{C})\, d\mu \tag{25}$$

[6] The weak, but nonvanishing, interaction is essential to the validity of Eq. (22). Otherwise, we would have some $G[\mathbf{C}] \neq F[\mathbf{C}]$ as the density in $q(M)$ corresponding to the physical *independence* of the spaces $(\Omega', \mathscr{M}')$ and $(\Omega'', \mathscr{M}'')$.

Equations (24) permit Eq. (20) to be rewritten in the form

$$\overline{f_0(\mathbf{C}^{-1}(\mathbf{E}_\beta))}^\infty = [Z'(\boldsymbol{\beta})]^{-1} \int_{\Omega_H'} f_0((\Psi', \Phi')) \exp(-\boldsymbol{\beta}\cdot\mathbf{C}')\, d\mu' \tag{26}$$

Equation (26) is the desired generalization of the canonical distribution. It states that the infinite-time average of a Hilbert function describing a physical system in weak interaction with a reservoir may be calculated (on sets of nonzero μ-measure, such that $\mathbf{C} = \mathbf{E}_\beta$) as an average over an ensemble in which the probability p' given in Eqs. (24) has been assigned to every set in Ω_H' whose image under $\mathbf{C}$ is $\mathbf{E}_\beta$. The vector $\boldsymbol{\beta}$ can be shown[2] to provide a criterion for equilibrium between the reservoir and the system in contact with it. In the special case that $\mathbf{C}$ is equal to the total energy of the reservoir plus system, $\boldsymbol{\beta}$ reduces to the familiar temperature modulus $(kT)^{-1}$ of the Boltzmann factor. The function $Z(\boldsymbol{\beta})$ is then the canonical partition function.

5. THE GRAND CANONICAL DISTRIBUTION

In order to describe a physical system which can exchange matter with its surroundings, we must generalize the space Ω to the space $\Gamma = \bigcup_{n\in\mathscr{N}} \Omega_n$, where $\mathscr{N}$ is the space of all vectors $n = \{n_1,..., n_k\}$ denoting particular compositions of an assembly containing k different kinds of particle. The resulting measurable space is then $(\Gamma, \mathscr{P})$, where $\mathscr{P} = \{P: P \cap \Omega_n \in \mathscr{M}_n\}$ is the class of all subsets in Γ. For every $P \in \mathscr{P}$ we define

$$m(P) = \sum_{n\in\mathscr{N}} \mu_n(P \cap \Omega_n) \tag{27}$$

where μ_n is given by Eq. (2). Equation (27) defines a measure on the sets of $\mathscr{P}$ since

$$m\left(\bigcup_{l=1}^{\infty} P_l\right) = \sum_{n\in\mathscr{N}} \sum_{l} \mu_n(P_l \cap \Omega_n) = \sum_{l} m(P_l)$$

for all sets $\{P_l: \bigcap_{l=1}^{\infty} P_l = \varnothing\}$ in $\mathscr{P}$. In the same spirit we define the strongly continuous group of unitary transformations $\{S_t\}$ on the elements of Γ by setting $S_t \equiv T_{nt}$ for each $(\Psi_\lambda{}^n, \Psi_{\lambda'}{}^n) \in (\mathscr{R}_H \times \mathscr{R}_H)_n \equiv \Omega_{nH} \subset \Omega$. The proof that $(\Gamma, \mathscr{P}, m, S_t)$ is a measure-preserving space follows as a direct generalization of the proof of Lemma 1.

Now let us write $\mathbf{n}((\Psi, \Phi))$ as the vector-valued function which maps from Γ into $\mathscr{N}$ such that

$$\mathbf{n}((\Psi, \Phi)) = \begin{cases} \mathbf{n} & \Psi^n \sim \Phi^n \\ 0 & \text{otherwise} \end{cases} \tag{28}$$

Then we may define the vector

$$\mathbf{C}((\Psi^n, \Phi^n)) \equiv \{\mathbf{n}, \mathbf{C}_n\}, \qquad \forall n \in \mathscr{N} \tag{29}$$

where $\mathbf{C}_n$ is a complete invariant vector for an assembly of particles of composition n (Definition 3). It is clear that Eq. (29) prescribes a set of Hilbert functions, for each $\Omega_{n\mathrm{H}} \subset \Gamma$, which has the property of completeness in the sense of Definition 3. Therefore $(\Gamma, \mathscr{P}, m, S_t, \mathbf{C})$ is a complete space.

If $(\Gamma'', \mathscr{P}'', m'', S_t'', \mathbf{C}'') \gg (\Gamma', \mathscr{P}', m', S_t', \mathbf{C}')$, we may apply the reasoning of Section 4 to show that

$$\overline{f_0(\mathbf{C}^{-1}(\mathbf{E}_\beta, \mathbf{N}_\mu))}^\infty = [z'(\boldsymbol{\beta}, \boldsymbol{\mu})]^{-1} \sum_{n' \in \mathscr{N}} [\exp(-\boldsymbol{\mu}\cdot\mathbf{n}')] \times \int_{\Omega'_{n\mathrm{H}}} f_0((\Psi^{n'}, \Phi^{n'})) \exp(-\boldsymbol{\beta}\cdot\mathbf{C}_n')\, d\mu_n' \tag{30}$$

$$z'(\boldsymbol{\beta}, \boldsymbol{\mu}) = \sum_{n' \in \mathscr{N}} \exp(-\boldsymbol{\mu}\cdot\mathbf{n}') \times \int_{\Omega'_{n\mathrm{H}}} [\exp(-\boldsymbol{\beta}\cdot\mathbf{C}_n')]\, d\mu_n' \tag{31}$$

and $\boldsymbol{\mu} = \{\mu_1,..., \mu_k\}$ is an arbitrary vector in R^k. Equation (30) is a generalization of the grand canonical distribution. It states that the infinite-time average of a Hilbert function describing a physical system in weak interaction with a particle reservoir is equal (on sets of nonzero μ-measure such that $\mathbf{C}_n' + \mathbf{C}_n'' = \mathbf{E}_\beta$ and $\mathbf{n}' + \mathbf{n}'' = \mathbf{N}_\mu$) to an average over an ensemble in which the probability

$$p'(P') \equiv [1/z'(\boldsymbol{\beta}, \boldsymbol{\mu})] \sum_{n \in \mathscr{N}_{p'}} \exp(-\boldsymbol{\mu}\cdot\mathbf{n}') \times \int_{M_n'} \exp(-\boldsymbol{\beta}\cdot\mathbf{C}_n')\, d\mu_n' \tag{32}$$

has been assigned to every set in $\bigcup_n \Omega_{n\mathrm{H}}$ whose image under $\mathbf{C}$ is $\{\mathbf{E}_\beta, \mathbf{N}_\mu\}$. [Note that in Eq. (32) $P' = \bigcup_n M_n' \subset U_n \Omega'_{n\mathrm{H}}$ and $\mathscr{N}_{p'} = \{n' \colon M_n' \subset P'\}$.] In the particular case that $\mathbf{C}_n'$ is the total energy Eq. (31) is the same as the conventional grand partition function, with $\boldsymbol{\mu}$ comprising the chemical potentials of the k different constituents in the system under consideration.

REFERENCES

1. G. Sposito, *J. Math. Phys.* **9**:1452 (1968).
2. R. M. Lewis, *Arch. Ratl. Mech. Anal.* **5**:355 (1960).
3. C. Truesdell, Ergodic Theory in Classical Statistical Mechanics, in *Ergodic Theories*, ed. by P. Caldirola, Academic Press, New York (1961).
4. I. E. Farquhar, *Ergodic Theory in Statistical Mechanics*, John Wiley, New York, (1964).
5. R. Jancel, *Foundations of Classical and Quantum Statistical Mechanics*, Pergamon Press, New York (1969).
6. P. R. Halmos, *Measure Theory*, D. Van Nostrand, Princeton, New Jersey (1950).

7. E. Prugovečki, *Quantum Mechanics in Hilbert Space*, Academic Press, New York (1971).
8. J. E. Moyal, *J. Math. Phys.* **10**:506 (1969).
9. G. G. Emch, *J. Math. Phys.* **7**:1413 (1966).
10. D. Ruelle, *Statistical Mechanics*, W. A. Benjamin, New York (1969).
11. A. Messiah, *Quantum Mechanics*, North-Holland, Amsterdam (1961), Vol. II, Chapter XIV.
12. G. Sposito, *An Introduction to Quantum Physics*, John Wiley, New York (1970), Chapter 3.
13. C. Kittel, *Thermal Physics*, John Wiley, New York (1969), Chapter 4.