AN ABSTRACT OF THE THESIS OF

Robert Edward Kvarda (Name) for the Ph. D. (Degree) in Mathematics (Major)

Date thesis is presented August 26, 1965

Title CANONICAL STATES IN QUANTUM STATISTICAL MECHANICS

Abstract approved [redacted]
Major professor

This report presents a characterization of the quantum mechanical analog of the Gibbs canonical density. The approach is based on a method developed by D. S. Carter for the case of classical statistical mechanics, which considers composite mechanical systems composed of mechanically and statistically independent components. After a brief introductory chapter, Chapter II outlines how the case of classical mechanics may be described in terms of the usual measure theoretic treatment of probability. The necessary statistical background of quantum mechanics is then discussed in Chapter III, relying on the classic treatment of J. von Neumann and the more recent work of G. W. Mackey. The basic idea of probability measure in quantum mechanics differs from that in classical measure theory, for the measure is defined on a non-Boolean lattice consisting of all closed linear subspaces of a Hilbert space. Because of this difference, the classical theory of product measures does not apply.

Chapter IV presents a detailed treatment of probability measures for composite quantum systems.

The analog of the Gibbs canonical density is characterized in Chapter V, by considering a large collection Q of noninteracting quantum systems, each of which is in an equilibrium statistical state. The set Q, the Hamiltonian operator for each system, and the equilibrium states are assumed to have certain properties which are given as axioms.

The axioms require each Hamiltonian operator to have a pure point spectrum. It is assumed, without loss of generality, that the lowest characteristic value of each Hamiltonian is zero. The set Q is assumed to be closed under the formation of pairwise mechanically independent composite systems. This implies that the set $\mathcal{S}$ of all Hamiltonian spectra is closed under addition. It is further assumed that $\mathcal{S}$ is closed under positive differences. The final requirement on the set Q is that it contain certain "harmonic oscillators". More precisely, for each positive $\lambda \in \mathcal{S}$, Q must contain a system whose Hamiltonian has the spectrum $\{n\lambda : n=0, 1, 2, \cdots\}$. The usual assumption is made that each density operator is a function of the system Hamiltonian. Finally, it is assumed that for each composite system in Q, with two mechanically independent components, the component systems are statistically independent.

It is shown that these assumptions imply that each member of

Q is in a canonical state at a temperature which is the same for all systems. The possibility of zero absolute temperature is included.

CANONICAL STATES IN QUANTUM STATISTICAL MECHANICS

by

ROBERT EDWARD KVARDA

A THESIS

submitted to

OREGON STATE UNIVERSITY

in partial fulfillment of
the requirements for the
degree of

DOCTOR OF PHILOSOPHY

June 1966

APPROVED:

[redacted]

Professor of Mathematics

In Charge of Major

[redacted]

Chairman of Department of Mathematics

[redacted]

Dean of Graduate School

Date thesis is presented August 26, 1965

Typed by Carol Baker

ACKNOWLEDGMENTS

The work on this dissertation was done while I was a resident graduate student at Oregon State University, on leave from the United States Navy Electronics Laboratory.

I am very much indebted to Professor David S. Carter for his direction and critical reading of this work. The rigorous development of the method of statistical independence was carried out by him for the case of classical statistical mechanics, and, although the extension to quantum mechanics requires a markedly different mathematical approach, the essential idea of the method remains unchanged.

I would like to thank the Department of Mathematics at Oregon State University for providing support under a U. S. National Science Foundation Grant in Applied Analysis. I also wish to thank the United States Navy Electronics Laboratory for their financial support.

R. E. K.

TABLE OF CONTENTS

Chapter		Page
I.	INTRODUCTION	1
II.	CLASSICAL STATISTICAL MECHANICS	5
III.	QUANTUM STATISTICAL MECHANICS	20
IV.	COMPOSITE QUANTUM SYSTEMS	39
V.	THE MAIN THEOREM	87
	BIBLIOGRAPHY	110
	APPENDIX	112

CANONICAL STATES IN QUANTUM STATISTICAL MECHANICS

CHAPTER I. INTRODUCTION

Statsitical Mechanics is the theory of matter in which each material object is regarded as a mechanical system, composed of vast numbers of particles or "molecules". The method of statistical mechanics is to suppose that at each instant of time the mechanical state of the system is not precisely known, but is "randomly distributed" over many possible values. More precisely, it is supposed that there is defined, at each time, a <u>probability measure</u> on the set of mechanical states. This probability measure may be called the <u>statistical state</u> of the system. It is through the statistical state that the essentially "non-mechanical" properties of matter, such as temperature and entropy, are brought into the theory.

Now the direct experimental observation of instantaneous mechanical states of a piece of matter is an entirely hopeless task, on practical grounds alone. The direct observation of statistical states is of course equally hopeless. It is therefore necessary, as part of the theory of statistical mechanics, to pick out appropriate probability measures, either by direct <u>ad</u> <u>hoc</u> hypotheses as to their forms, or by other more or less plausible postulates. In the case of classical equilibrium statistical mechanics, where the statistical state is independent of time, J. Willard Gibbs [6, Ch. 25] discovered that a particularly simple form for the probability measure served to explain the basic thermodynamic properties of matter in equilibrium. Gibbs called these special probability measures "canonical", and they are still referred to as "Gibbs canonical states" (see Chapters 2 and 3, respectively, for the form of the canonical states in classical and quantum statistical mechanics).

The Gibbs canonical states met with such complete success that many scientists felt there should be a deeper justification for them, beyond the brute fact that they work. There now exist a number of "derivations" of the canonical states, all based on fairly plausible statistical assumptions. An especially appealing justification, based on the statistical independence of noninteracting systems, has recently been pointed out by D.S. Carter [4] in the classical case (see also R. Kurth [14, p. 129]). The essential idea is as follows: Suppose that with each mechanical system there is associated a definite class of equilibrium states. Consider a mechanical system composed of two separate, noninteracting, component systems. When this composite system is in an equilbirium state, each component system is assumed to be in an equilibrium state. Finally, it is supposed that since the component systems are <u>mechanically independent</u> (i.e. noninteracting) they are also <u>statistically independent</u>, so that the knowledge of an event in one component does not affect the probability of an event in the other (in the language of measure theory, this means that the probability measure representing the state of the composite system is the direct product of the measures representing the states of the component systems). These assumptions essentially characterize the canonical states (see Chapter 2).

The purpose of this dissertation is to extend the above characterization of the Gibbs states from classical to quantum statistical mechanics. To provide the necessary background, Chapter 2 is devoted to the classical case. Following a brief introduction to the general theory, the canonical states are defined. The characterization is then stated precisely, and compared with two better-known "derivations" - the method of Khinchin and the maximum entropy method. Chapter 3 introduces the basic structure of quantum statistical mechanics. Here the basic notion of "probability measure" differs markedly from that of classical measure theory, for the measure is now defined on a non-Boolean lattice consisting of all closed linear subspaces of a Hilbert space, rather than a Boolean σ-algebra of sets. Because of this difference, the classical theory of product measures does not apply, and the necessary new theory seems as yet incomplete in the literature. Chapter 4 therefore presents a detailed treatment of composite quantum systems, where the probability measure is defined on the tensor product of Hilbert spaces.

The analogies between classical and quantum statistical mechanics are summarized and tabulated at the end of Chapter 4.

Finally, the desired characterization of the canonical states is presented in Chapter 5.

CHAPTER II. CLASSICAL STATISTICAL MECHANICS

In classical statistical mechanics, a mechanical system is represented mathematically by a phase space Γ, whose points z are the instantaneous mechanical states of the system. The phase space has the structure of a differentiable manifold, with a preferred class of local coordinate systems, called "generalized canonical coordinates" or "Hamiltonian coordinates", $(p_i, q_i : i = 1, 2, 3 \cdots 3n)$. The phase space also has the structure of a measure space, given locally by Lebesgue measure with respect to the canonical coordinates. The measure is independent of the particular coordinates because the transformation from one canonical coordinate system to another is a contact transformation, whose Jacobian determinant is identically equal to one [3, p. 92].

The Lebesgue measurable subsets of Γ are called events. If L is a Lebesgue measurable subset of Γ, it is identified with the event that the mechanical state z is included in L. The set of all events, denoted by $\mathcal{L}$, therefore forms a Boolean σ-algebra with respect to the usual operations of union, intersection, and complementation.

Real valued measurable functions on the phase space Γ are called observables. In particular, the canonical coordinates are observables. Moreover, every observable is given locally by a

Lebesgue measurable function of these coordinates.

A particularly important observable is the Hamiltonian H. This function, whose value at each point z is the total energy of the system when it is in the mechanical state z, determines the dynamics of the system through Hamilton's equations

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i}. \tag{2.1}$$

(Actually H is determined only up to an arbitrary additive constant.) The integral curves of these equations are the trajectories of a one-parameter group of contact tranformations on Γ, the parameter being the time. We denote the value of this transformation at time t by T_t, so that during a time interval of length t the mechanical state will change from a point $z \varepsilon \Gamma$ to the point $T_t z$. The transformation T_t, being continuous, maps $\mathcal{L}$ into itself, in the sense that for each $L \varepsilon \mathcal{L}$ the set

$$T_t[L] = \{T_t(z): z \varepsilon L\}$$

also belongs to $\mathcal{L}$. Moreover, since T_t is a contact transformation, it is measure preserving; that is,

$$\mu(L) = \mu(T_t[L]).$$

A statistical state of the system is defined to be a probability measure on the measurable space $(\Gamma, \mathcal{L})$. Thus, a system in a statistical state p is a probability space $(\Gamma, \mathcal{L}, p)$. In general, the statistical state is a function of the time. The dependence on time is determined by the condition that the probability of a "moving event" is unchanged as the event changes with the dynamical motion of the system. In symbols, let p_0 be the initial state (at zero time) and p_t be the state after a time interval t. Each event $L \varepsilon \mathcal{L}$ evolved during this interval from the initial event $T_{-t}[L]$, and we take

$$p_t(L) = p_0(T_{-t}[L]) . \tag{2.2}$$

The statistical state is usually, though not always, assumed to be absolutely continuous with respect to Lebesgue measure. In other words p_t is determined by a probability density D_t, which is a non-negative point function such that for all $L \varepsilon \mathcal{L}$,

$$p_t(L) = \int_L D_t(z)\, \mu(dz) . \tag{2.3}$$

Combining this with equation (2.2) yields

$$D_t(z) = D_0(T_{-t}(z)) \tag{2.4}$$

almost everywhere in Γ.

Equilibrium statistical mechanics deals with cases in which the statistical state is stationary, that is, constant in time. Then one may write

$$p_t = p_0 = p, \tag{2.5a}$$

or, if p is absolutely continuous,

$$D_t = D_0 = D. \tag{2.5b}$$

Combining these conditions with (2.2) and (2.4) yields the identities

$$p(L) = p(T_{-t}[L]) \tag{2.6a}$$

$$D(z) = D(T_{-t}\, z). \tag{2.6b}$$

These conditions represent very mild restrictions on the statistical states. Much more severe restrictions are needed to pick out those states which describe matter in equilibrium. Equation (2.6b), for example, merely requires that the probability density $D(z)$ be constant on trajectories of the system; in other words, $D(z)$ is an integral of the motion. Given any set $\{k_1(z), k_2(z), \cdots, k_n(z)\}$ of such "integrals" which are measurable but not necessarily positive, we can form probability densities by taking positive integrable functions of $k_1(z), \cdots, k_n(z)$, as follows: Let f be a positive measurable function of n (real) variables such that the composite function

$$F(z) = f(k_1(z), \cdots, k_n(z))$$

is defined almost everywhere in Γ, and such that F is integrable:

$$\int_\Gamma F(z)\mu(dz) < \infty .$$

Then the function D given by

$$D(z) = \frac{F(z)}{\int_\Gamma F(y)\mu(dy)} \tag{2.7}$$

is a probability density satisfying (2.4) and (2.6). Moreover, every probability density D satisfying (2.4) and (2.6) can be obtained in this way, for D is itself an integral of the motion and one may take $n = 1$, $k_1 = D$, f = identity function, to obtain $F = D$.

In general, the only integrals of the motion which are known "a priori", without solving the equations of motion, are functions of the Hamiltonian H - the "energy integral". A great restriction on the possible choice of probability densities is made by taking D to be a function of H, that is,

$$D(z) = \frac{f(H(z))}{\int_\Gamma f(H(y))\mu(dy)} . \tag{2.8}$$

This hypothesis has been partially justified on the basis of ergodic theory [13, p. 55]. Perhaps the best justification is on empirical grounds; if D should depend on other integrals besides H, this fact would show up in the experimental properties of matter. This actually occurs in the case of chemically active systems, where the number of particles of each substance is variable (or in quantum mechanics, systems composed of photons). Then the number N of particles is regarded as an integral of the motion, and D is taken to be a function of both H and N. We shall consider only "closed systems" where N does not change, and therefore restrict ourselves to probability densities of the form (2.8).

Even after restricting D to be a function of H alone, we are faced with a vast array of possible states. J. Willard Gibbs [6, Ch. 25] discovered that the appropriate choice for D is an exponential function of H, so that

$$(2.9) \qquad D(z) = \frac{e^{-\theta H(z)}}{\int_{\Gamma} e^{-\theta H(y)} \mu(dy)}, \quad 0 < \theta < \infty,$$

where the parameter θ determines the absolute temperature T of the system (more precisely, $\theta = \frac{1}{kT}$, where k is Boltzmann's constant). States given by these densities are known as "Gibbs canonical states". Although Gibbs did not explain his choice of states beyond showing that they work, various "derivations" have been advanced, each based on more or less plausible grounds. Three such derivations - the method of Khinchin, the maximum entropy method, and the method of statistical independence are reviewed briefly below. First, however, we must introduce the concept of a composite system.

Given two mechanical systems represented by phase spaces Γ_1, Γ_2, we may form the composite system, with phase space $\Gamma_{12} = \Gamma_1 \times \Gamma_2$. The points of Γ_{12} are the ordered pairs $(z_1, z_2): z_1 \varepsilon \Gamma_1, z_2 \varepsilon \Gamma_2$, and local Hamiltonian coordinates $(p_i^{(1)}, p_i^{(2)}, q_i^{(1)}, q_i^{(2)})$ are given in terms the coordinates $(p_i^{(1)}, q_i^{(1)})$, $(p_i^{(2)}, q_i^{(2)})$ for the respective systems. As a measure space, Γ_{12} is the direct product of the measure spaces Γ_1, Γ_2. Thus if $\mathcal{L}_1, \mathcal{L}_2$ are the σ-algebras of events in Γ_1, Γ_2, the corresponding set $\mathcal{L}_{12}$ of events in Γ_{12} is the σ-algebra generated by all sets of the form $L_1 \times L_2$: $L_1 \varepsilon \mathcal{L}_1$, $L_2 \varepsilon \mathcal{L}_2$. We interpret $L_1 \times L_2$ as the event that $z_1 \varepsilon L_1$ and $z_2 \varepsilon L_2$.

A composite system in a statistical state p_{12} is a probability space $(\Gamma_1 \times \Gamma_2, \mathcal{L}_1 \times \mathcal{L}_2, p_{12})$, where p_{12} is a probability measure on the measurable space $(\Gamma_1 \times \Gamma_2, \mathcal{L}_1 \times \mathcal{L}_2)$. As in the case of a single system, p_{12} will in general depend on the time.

Now the statistical state p_{12} determines not only the statistical properties of the composite system, but the statistical properties of the component systems as well, since the probabilities of all events related to one or the other components may be determined from p_{12}. That is, given the statistical state p_{12} of the composite system, we define the statistical states p_1 and p_2 of the component systems by

$$p_1(L_1) = p_{12}(L_1 \times \Gamma_2)$$

$$p_2(L_2) = p_{12}(\Gamma_1 \times L_2)$$

for all $L_1 \varepsilon \mathcal{L}_1$ and $L_2 \varepsilon \mathcal{L}_2$. It may be easily verified that p_1 and p_2 are probability measures. If p_{12} is absolutely continuous with respect to Lebesgue measure on $\Gamma_1 \times \Gamma_2$, then it is again determined by a probability density $D_{12}(z_1, z_2)$, that is, for all sets $A \varepsilon \mathcal{L}_1 \times \mathcal{L}_2$

$$p_{12}(A) = \int_A D_{12}(z_1, z_2)\mu(dz_1 dz_2) .$$

In particular, we have

$$p_1(L_1) = p_{12}(L_1 \times \Gamma_2) = \int_{L_1 \times \Gamma_2} D_{12}(z_1, z_2)\mu(dz_1 dz_2)$$

for all $L_1 \varepsilon \mathcal{L}_1$. By Fubini's theorem [9, p. 148] , this is

$$\int_{L_1} [\int_{\Gamma_2} D_{12}(z_1, z_2)\mu(dz_2)] \mu(dz_1),$$

where

(2.10)
$$\int_{\Gamma_2} D_{12}(z_1, z_2)\mu(dz_2) = D_1(z_1)$$

is a non-negative, integrable function defined almost everywhere in Γ_1. In a similar way, the probability that $z_2 \varepsilon L_2$ is given by

$$P_2(L_2) = p_{12}(\Gamma_1 \times L_2) = \int_{\Gamma_1 \times L_2} D_{12}(z_1, z_2)\mu(dz_1 dz_2) = \int_{L_2} [\int_{\Gamma_1} D_{12}(z_1, z_2)\mu(dz_1)] \mu(dz_2),$$

where

(2.11)
$$\int_{\Gamma_1} D_{12}(z_1, z_2)\mu(dz_1) = D_2(z_2)$$

is also a non-negative, integrable function and is defined almost everywhere in Γ_2. Moreover, Fubini's theorem states that

$$\int_{\Gamma_1} D_1(z_1)\mu(dz_1) = \int_{\Gamma_2} D_2(z_2)\mu(dz_2) = \int_{\Gamma_1 \times \Gamma_2} D_{12}(z_1, z_2)\mu(dz_1 dz_2) = 1.$$

Therefore $D_1(z_1)$ and $D_2(z_2)$ are also probability densities and they determine the induced statistical states p_1 and p_2.

In general, it will not be true that $D_{12}(z_1, z_2) = D_1(z_1)D_2(z_2)$. However, if the component systems are <u>statistically independent</u>, then the probability of any event of the form $L_1 \times L_2 \varepsilon \mathcal{L}_1 \times \mathcal{L}_2$ is the product $p_1(L_1)p_2(L_2)$, and there is one and only one probability measure on $(\Gamma_1 \times \Gamma_2, \mathcal{L}_1 \times \mathcal{L}_2)$, denoted by $p_1 \times p_2$, such that

$$(p_1 \times p_2)(L_1 \times L_2) = p_1(L_1)p_2(L_2)$$

for all rectangles $L_1 \times L_2 \varepsilon \mathcal{L}_1 \times \mathcal{L}_2$. If this is the case, then we also have

$$D_{12}(z_1, z_2) = D_1(z_1)D_2(z_2) .$$

Hamilton's equations for the composite system are

$$\text{(2.12a)} \qquad \frac{\partial q_i^{(1)}}{\partial t} = \frac{\partial H_{12}}{\partial p_i^{(1)}}, \quad \frac{\partial p_i^{(1)}}{\partial t} = -\frac{\partial H_{12}}{\partial q_i^{(1)}}$$

$$\text{(2.12b)} \qquad \frac{\partial q_i^{(2)}}{\partial t} = \frac{\partial H_{12}}{\partial p_i^{(2)}}, \quad \frac{\partial p_i^{(2)}}{\partial t} = -\frac{\partial H_{12}}{\partial q_i^{(2)}},$$

where H_{12} is the Hamiltonian function of the composite system. If H_{12} is expressible as the sum of the component system Hamiltonians, that is,

$$H_{12}(z_1, z_2) = H_1(z_1) + H_2(z_2) , \tag{2.13}$$

then equations (2.12) reduce to two independent systems of equations. This means that the trajectories in the phase spaces Γ_1 and Γ_2 are independent of each other, or in other words, the component systems are <u>mechanically independent</u> or <u>noninteracting</u>.

In his derivation of the canonical states, Khinchin uses the notion of a <u>microcanonical</u> state. Roughly speaking, a microcanonical state is one in which the system is restricted to a single energy surface

$$H(z) = E_0 = \text{constant}$$

in its phase space Γ, but uniformly distributed over the surface. Because of this restriction to a hypersurface, microcanonical states are not absolutely continuous with respect to Lebesgue measure in Γ (for a detailed discussion of microcanonical states see Khinchin [13, p. 110]).

To obtain the canonical states, Khinchin considers a system S composed of a large number N of component systems

$S_1, \cdots S_N$. The components are taken identical to each other, so that the component phase spaces $\Gamma_1, \cdots, \Gamma_N$, and the Hamiltonians $H_1, \cdots, H_N$, are all the same. Moreover, the component systems are supposed to be mechanically independent, so that the Hamiltonian H for S is the sum

$$H = \sum_{n=1}^{N} H_n .$$

Now suppose the large system S is in a microcanonical state p_0 of energy E_0. As N increases, let the energy E_0 increase in proportion to N, that is,

$$E_0 = NE_1,$$

where E_1 is the "energy per component". Khinchin shows that as $N \to \infty$, the state p_1 induced by p_0 in each of the component systems approaches the Gibbs canonical state with mean energy E_1. (In the physical interpretation, the large system S represents a heat bath with which each component system is weakly interacting. Khinchin's result is interpreted as meaning that the canonical state is obtained in the limit as the interaction vanishes and the heat bath becomes infinite).

Another characterization of the canonical state makes use of the ideas of information theory [11; 16, p. 48] . From this point of view, the canonical probability density D_0 represents a statistical state in which our ignorance of the system is maximal, except for knowing the mean value of the Hamiltonian function. In other words, the system is distributed over Γ in the most random possible way subject to the condition

$$(2.14) \qquad \langle H(z) \rangle = \int_\Gamma H(z) D_0(z) \mu(dz) = E_0 \text{ (constant).}$$

Information theory, which actually has its origins in statistical mechanics, suggests that the <u>entropy integral</u> be taken as a measure of the degree of ignorance (or "randomness"). If D is a probability density defined almost everywhere in Γ, then the entropy of D is defined by

$$S_D = -\int_\Gamma D(z) \log D(z) \mu(dz) .$$

It can be shown that if D is any probability density other than D_0, having the mean value E_0 of H given by (2.14), then the entropy satisfies $S_D < S_{D_0}$. Therefore the condition of maximum entropy characterizes the Gibbs canonical state.

The "method of statistical independence", as presented by D.S. Carter [4] , characterizes the Gibbs canonical states as follows: Suppose there is associated with each system a probability density D, expressible as a measurable function of the system Hamiltonian H. Suppose further that for each composite system with noninteracting components, the component systems are also statistically independent. Let $D_1 = f_1(H_1)$ and $D_2 = f_2(H_2)$ be the densities for the components of a composite system with density $D_{12} = f_{12}(H_{12})$. Then the equations of mechanical and statistical independence

$$H_{12}(z_1, z_2) = H_1(z_1) + H_2(z_2)$$

$$D_{12}(z_1, z_2) = D_1(z_1)D_2(z_2)$$

together yield the functional equation

$$f_{12}(H_1 + H_2) = f_1(H_1)f_2(H_2). \tag{2.15}$$

If reasonable assumptions are made on the Hamiltonians H_1 and H_2, one finds that all measurable solutions f_1, f_2, and f_{12} of equation (2.15) have the form

$$\begin{aligned} f_1(x) &= A_1 e^{-\theta x} \\ f_2(x) &= A_2 e^{-\theta x} \\ f_{12}(x) &= A_{12} e^{-\theta x} \end{aligned} \tag{2.16}$$

where A_1, A_2, and θ are constants, and $A_{12} = A_1 A_2$. This implies that the component systems, and also the composite system, are in canonical equilibrium states corresponding to the same value of the absolute temperature. The specific conditions imposed on the Hamiltonians are:

a) H is a continuous function of z which varies between the limits 0 and $+\infty$. (This means that equation (2.15) becomes

$$f_{12}(x+y) = f_1(x) f_2(y)$$

for almost all $x, y \geq 0$).

b) The so called "structure function",

$$\Omega(x) = \frac{d}{dx} \int_{0 \leq H(z) \leq x} \mu(dz)$$

is a strictly positive and continuous function of x.

c) The integral $\int_0^\infty e^{-\theta x} \Omega(x) \mu(dx)$ exists for all $\theta > 0$.

CHAPTER III. QUANTUM STATISTICAL MECHANICS

Many of the ideas of classical mechanics carry over to quantum mechanics, at least in an analogous form. Our discussion of quantum mechanics will be based on the description presented by J. von Neumann [17]. In this description the concepts of phase space, mechanical state, and observable, which in classical mechanics correspond respectively to a measure space $(\Gamma, \mathcal{L}, \mu)$, points in Γ, and measurable functions on Γ, correspond to a separable complex Hilbert space $\mathcal{H}$, unit vectors in $\mathcal{H}$, and self-adjoint operators on $\mathcal{H}$. In contrast to classical mechanics, where the mathematical structure is well known, it is generally recognized that the mathematical structure of quantum theory is incomplete. For example, it is conceivable that the phase space might have a structure other than that of a Hilbert space [2; 16, p. 71-74]. Furthermore, it is difficult to justify the choice of a complex scalar field, except that if this choice is made, certain formal features of classical mechanics also appear in quantum mechanics. In fact, the possibility of quaternionic Hilbert spaces has been introduced in quantum mechanics [5]. However, we shall not attempt to investigate these other possibilities and proceed from the conventional assumption of a complex Hilbert space.

If this is done, the quantum mechanical states are the unit vectors in $\mathcal{H}$. If ϕ is a unit vector and c is any complex number, of absolute value 1, quantum mechanics asserts that ϕ and $c\phi$ are identified as the same mechanical state. It is further asserted that the observables correspond in a one-to-one way with the self-adjoint operators on $\mathcal{H}$.

In general, there exists more than one way to construct a quantum mechanical description of a classical system. For example, if we are given a classical system of N particles, with Hamiltonian coordinates $(p_i, q_i: i = 1, 2, \cdots 3N)$, we obtain the coordinate or configuration representation as follows: We define the configuration space of the system as the space E^{3N} with rectangular coordinates $(q_i: i = 1, 2, \cdots 3N)$, and identify $\mathcal{H}$ with the space of all those complex-valued functions defined on E^{3N} which are square-summable with respect to Lebesgue measure; that is, the space $L^2(E^{3N}, \mu)$. However, except for the purpose of constructing examples, it will not be necessary to specify a particular representation. We shall regard $\mathcal{H}$ as an abstract Hilbert space.

With $\mathcal{H}$ known the system can then be described by a pair $(\mathcal{H}, H)$ where H is a self-adjoint operator, called the Hamiltonian operator, which plays the correpsonding role in quantum mechanics as the Hamiltonian function in classical mechanics. In particular, it is the quantum mechanical analog of the total energy of the system, and it completely determines the dynamics of the system through the

Schrödinger differential equation

$$\frac{d\phi}{dt} = -iH\phi,$$

where ϕ is the instantaneous quantum mechanical state.

The evolution of $(\mathcal{H}, H)$ with time may also be expressed by a transformation T_t as in the classical case. If $\phi_{t'}$ is the quantum mechanical state at time t', then $T_t\phi_{t'} = \phi_{t+t'}$ is the state after a time interval of length t. We shall be concerned only with systems which are reversible in the sense that T_t is invertible. Defining $T_t^{-1} = T_t$, one may identify the set $\{T_t\colon t \epsilon R\}$ as a continuous one-parameter group of unitary transformations called the dynamical group of the system. The infinitesimal generator of $\{T_t\}$ is just iH, and by Stone's theorem [19, p. 385; 22], one obtains $T_t = e^{-iHt}$ for all $t \epsilon R$. The stationary states of $(\mathcal{H}, H)$ are those unit vectors which are fixed in time in the sense that $T_t\phi_{t'} = \phi_{t'}$ for all t. Among them are included the characteristic vectors of $H : H\phi = \lambda\phi$.

As in classical mechanics, the Hamiltonian is arbitrary up to an additive constant, that is, we are free to renormalize the Hamiltonian as $H' = H - \lambda^0 I$, where λ^0 is a real number which may be chosen arbitrarily and I is the identity operator. If ϕ_k is a characteristic vector of H corresponding to the characteristic value λ_k, then ϕ_k remains a characteristic vector of H', but now corresponds to the characteristic value $\lambda_k - \lambda^0$. The renormalization of H translates the energy spectrum along the real axis by an amount λ^0, but this change causes no real difficulties. Of course, changing H to H' changes T_t to $T_t' = e^{i\lambda^0 t} T_t$, but if ϕ and ψ are mechanical states such that $T_t \phi = \psi$, then $T_t' \phi = e^{i\lambda^0 t}\psi$, and as we have stated previously ψ and $c\psi$, where $|c| = 1$, are identified as the same mechanical state. Therefore the dynamical group remains essentially unchanged.

Events in quantum mechanics, corresponding to the measurable subsets of Γ in classical mechanics, are the closed linear subspaces of $\mathcal{H}$. If M is a closed linear subspace of $\mathcal{H}$, it corresponds to the event that the representative quantum mechanical state of the system is a unit vector in M. One can then proceed to define a calculus of events for quantum mechanics. The intersection and closed linear sum of any two, and the orthogonal complement of any one closed linear subspace of $\mathcal{H}$ are themselves closed linear subspaces and therefore events. If one defines a relation of implication, "M_1 implies M_2" to mean that the subspace M_1 is also a subspace of M_2, then one obtains a calculus of events with three operations and a relation of implication. In this calculus, the closed linear subspaces of a Hilbert space form a complemented lattice. They need not, however, satisfy the distributive identity which is a law in classical mechanics. Hence in quantum mechanics one has a non-Boolean set of events (for a discussion of the logic of quantum mechanics, the reader is referred to G. Birkhoff and J. von Neumann

[2]).

A <u>statistical state</u> in quantum mechanics is a "probability measure" on the closed subspaces of $\mathcal{H}$, in the sense of

<u>Definition 3-1.</u> A probability measure on the closed subspaces of a Hilbert space is a function p which assigns to every closed subspace $M \subset \mathcal{H}$ a non-negative real number p(M), such that

(a) $p(\mathcal{H}) = 1, \quad p(\{0\}) = 0$

(b) $0 \leq p(M) \leq 1$, for all closed subspaces $M \subset \mathcal{H}$

(c) If $\{M_i\}$ is a countable collection of mutually orthogonal subspaces having closed linear span M, then

$$p(M) = \sum_{i=1}^{\infty} p(M_i) .$$

This definition of a statistical state is due to G. W. Mackey [16, p. 63].

In both classical and quantum mechanics we may therefore say that a statistical state is a probability measure on the set of all events. Whereas in classical mechanics a statistical state is <u>usually</u> determined by a probability density, it turns out that in quantum mechanics <u>every</u> statistical state is determined by a non-negative, self-adjoint operator of the trace class, called a <u>density operator</u>. To explain this, we must first introduce the trace class operators, which play a role in quantum mechanics analogous to integrable functions in classical mechanics. The trace of an operator in the trace class is the quantum mechanical analog of the integral of an integrable function in classical mechanics.

Let T^* denote the adjoint of the operator T. We have

Definition 3-2. The trace class (tc) consists of all bounded operators T such that the sum

$$\sum_{n=1}^{\infty} ((T^*T)^{1/2} \phi_n, \phi_n) \tag{3.2.1}$$

is convergent for some complete orthonormal sequence $\{\phi_n\}$.

Actually the sum (3.2.1) is independent of the choice of orthonormal sequence $\{\phi_n\}$. A proof of this may be found in R. Schatten [20, p. 42], where it is also shown that the sum

$$\sum_{n=1}^{\infty} |(T\phi_n, \phi_n)| \tag{3.2.2}$$

converges. Moreover, the sum

$$\sum_{n=1}^{\infty} (T\phi_n, \phi_n)$$

is independent of the choice of orthonormal basis $\{\phi_n\}$. This last property of the trace class leads to

Definition 3-3. For $T\varepsilon(tc)$, the finite number $\sum_{n=1}^{\infty} (T\phi_n, \phi_n)$ is called the trace of T, denoted by trace (T).

Our interest will center on products of operators of which at least one is in (tc). The following properties will be useful [20, p. 38].

(3.3.1) If $T\varepsilon(tc)$, and X is any bounded operator, then XT and TX are in (tc), and trace (XT)=trace (TX).

(3.3.2) If $T_1, T_2 \varepsilon(tc)$, then $(T_1 \pm T_2)\varepsilon tc$, and trace $(T_1 \pm T_2)$ = trace T_1 ± trace T_2.

A density operator may now be defined as follows.

Definition 3-4. An operator D on $\mathcal{H}$ is a density operator if and only if

(3.4.1) $D\varepsilon(tc)$ is self-adjoint and non-negative (i.e. $(D\phi, \phi) \geq 0$ for all $\phi\varepsilon\mathcal{H}$).

(3.4.2) trace D = 1.

The following theorem [7] establishes that every statistical state in quantum mechanics is determined by a density operator.

Theorem 3-5. (A. M. Gleason) Let m be a measure on the closed subspaces of a separable (real or complex) Hilbert space of dimension at least three. Then there exists a unique, non-negative, self-adjoint operator $T \varepsilon (t c)$ such that for all closed subspaces M of $\mathcal{H}$,

$$m(M) = \text{trace}\,(TP_M),$$

where P_M is the orthogonal projection of $\mathcal{H}$ onto M.

An important property of density operators is that they are completely continuous, in the sense of

Definition 3-6. An operator T defined on $\mathcal{H}$ is completely continuous if it transforms every weakly convergent sequence of vectors in $\mathcal{H}$ into a strongly convergent sequence.

Recall that a sequence of vectors $\{f_m\}$ is weakly convergent to f if

$$\lim_{m \to \infty} |(f_m, \eta) - (f, \eta)| = 0$$

for all vectors $\eta \varepsilon \mathcal{H}$. A sequence of vectors $\{g_m\}$ is strongly convergent to g if

$$\lim_{m \to \infty} ||g_m - g|| = 0.$$

Definition 3-6 is credited to Hilbert [10, note 4]. Other definitions of a completely continuous operator may be found in F.

Riesz and B. Sz-Nagy [19, p. 206] along with the proofs of their equivalence.

The fact that all density operators are completely continuous follows from a deeper result of R. Schatten [20, p. 41], to the effect that every operator of the trace class is completely continuous, even if the Hilbert space $\mathcal{H}$ is non-separable, and with a broader definition of trace class based on summability of $\sum((T^*T)^{1/2}\phi_n, \phi_n)$ rather than convergence. Because this fact is of fundamental importance to our discussion, a special proof is given in the Appendix.

It is well known [20, p. 16] that if a completely continuous operator T on a separable Hilbert space is also self-adjoint, then it admits a complete orthonormal sequence of characteristic vectors $\{\phi_k\}$. Its non-zero (necessarily real) characteristic values are of finite multiplicity and form either a finite or countably infinite sequence $\{\lambda_k\}$. Therefore (see equation (3.6.2a)) a density operator has the spectral representation

$$(3.6.1) \qquad D = \sum_{k=1}^{\infty} \lambda_k P_k, \quad 0 \leq \lambda_k \leq 1, \quad \sum_{k=1}^{\infty} m_k \lambda_k = 1,$$

where P_k is the projection onto the characteristic subspace, of dimension m_k, corresponding to the characteristic value λ_k. Here λ_k runs through the different characteristic values of D.

The analogy between probability densities in classical statistics and density operators in quantum statistics can be brought out in another way. If $f(z)$ is a bounded real-valued measurable function (an observable) on Γ, its mean value is

$$<f(z)> = \int_{\Gamma} D(z)f(z)\mu(dz),$$

where $D(z)$ is the probability density defining the statistical state. In quantum statistics, the mean value of a bounded self-adjoint operator A corresponding to an observable is given by

$$<A> = \text{trace}\ (DA),$$

where D is the density operator defining the statistical state.

Every observable in classical mechanics is given by a measurable function of the Hamiltonian observables $(p_i, q_i: i=1, 2, \cdots 3n)$. In quantum mechanics, one can conveniently define functions of observables by using the spectral theorem. If A is any self-adjoint operator on $\mathcal{H}$, then the spectral theorem [19, p. 320] states that A has the unique representation

(3.6.2) $$A = \int_{-\infty}^{\infty} \lambda \, dE_{\lambda}$$

everywhere in the domain of A, where $\{E_{\lambda}\}$ is a spectral family of projections uniquely determined by A. The integral is taken in the Lebesgue-Steltjes sense. If f is any Borel function on the real line, then one defines functions of A by the representation

$$(3.6.3) \qquad f(A) = \int_{-\infty}^{\infty} f(\lambda)\,dE_\lambda ,$$

where the domain of $f(A)$ consists of all elements $\phi \varepsilon \mathcal{H}$ for which the integral

$$(3.6.4) \qquad \int_{-\infty}^{\infty} |f(\lambda)|^2 d(E_\lambda \phi, \phi)$$

converges [19, p. 345]. If f is real-valued, then $f(A)$ defines a self-adjoint operator, hence an observable.

In the case where A has the pure point spectrum $S(A) = \{\lambda_k\}$, the spectral representations reduce to

$$(3.6.2a) \qquad A = \sum_{k=1}^{\infty} \lambda_k P_k$$

$$(3.6.3a) \qquad f(A) = \sum_{k=1}^{\infty} f(\lambda_k) P_k ,$$

where P_k is the projection onto the characteristic subspace of A corresponding to the characteristic value λ_k. The domain of $f(A)$ consists of all $\phi \varepsilon \mathcal{H}$ for which the sum

$$(3.6.4a) \qquad \sum_{k=1}^{\infty} |f(\lambda_k)|^2 (P_k \phi, \phi)$$

converges.

Summarizing, we are now able to state the fundamental assertion of quantum statistics: The events of a physical system $(\mathcal{H}, H)$ coincide with the set $\mathfrak{M}$ of all closed subspaces of the separable Hilbert space $\mathcal{H}$. Each statistical state of $(\mathcal{H}, H)$ is a probability measure p on $\mathfrak{M}$. Theorem 3-5 then identifies each state with a density operator in a one-to-one way such that if p and D correspond then the probability of an event $M \subset \mathfrak{M}$ is given by

$$p(M) = \text{trace}\,(DP_M).$$

In non-equilibrium statistical mechanics, the statistical state, and therefore the corresponding density operator, will depend on the time. It follows from Theorem 3-5 that for each time t there exists a density operator $D(t)$ defined on $\mathcal{H}$ such that for all closed subspaces $M \subset \mathcal{H}$

$$p_t(M) = \text{trace}\,(D(t)P_M).$$

As time evolves, $(\mathcal{H}, H)$ evolves according to the dynamical group $\{T_t\}$, where $T_t = e^{-iHt}$, $t \varepsilon R$. Since T_t is a unitary operator, we have a mapping of the set $\mathfrak{M}$ into itself given by

$$T_t(M) = \{T_t(\phi): \phi \varepsilon M\}$$

for all $M \varepsilon \mathfrak{M}$. If $\phi \varepsilon M$ at time t, then $\phi \varepsilon T_{-t}(M)$ at time zero.

Thus if the quantum mechanical state ϕ is observed to be in $T_{-t}(M)$ at time $t = 0$, then an observation at time t must yield that $\phi \varepsilon M$. As in the case of classical mechanics, we take

$$p_t(M) = p_0(T_t M) \tag{3.6.5}$$

for all closed subspaces $M \subset \mathcal{H}$. This equation determines how the statistical states change with time. The change in the corresponding density operator is given by the following theorem.

<u>Theorem 3-7.</u> For each time t, the density operator $D(t)$ is given by $D(t) = T_t D(0) T_{-t}$.

<u>Proof</u>: Let M be any closed subspace in $\mathcal{H}$. If $\{\phi_k\}$ is an orthonormal basis in M, then

$$p_t(M) = \text{trace}\,(D(t)P_M) = \sum_k (D(t)\phi_k, \phi_k).$$

Let $\psi_k = T_{-t}\phi_k$. Then since T_t is a unitary operator, we have

$$p_t(M) = \sum_k (D(t)T_t\psi_k,\ T_t\psi_k) = \sum_k ((T_{-t}D(t)\,T_t)\psi_k, \psi_k).$$

Since $D(t)\varepsilon(tc)$, by property (3.3.1), $D(t)T_t$ and therefore $(T_{-t}D(t)T_t)$ are in (tc). Consequently we may write

$$p_t(M) = \text{trace}\,((T_{-t}D(t)T_t)P_{T_{-t}M}).$$

Clearly this holds for all closed subspaces $M \subset \mathcal{H}$. By (3.6.5), we obtain

$$p_0(T_{-t}M) = \text{trace}\,((T_{-t}D(t)T_t)P_{T_{-t}M})$$

for all M. Hence for all closed subspaces $N \subset \mathcal{M}$,

$$p_0(N) = \text{trace}\,((T_{-t}D(t)T_t)P_N).$$

By Theorem 3-5, we also have

$$p_0(N) = \text{trace}\,(D(0)P_N)$$

for all N. Therefore, since the density operator determined by the probability measure p_0 is unique, we have

$$T_{-t}D(t)T_t = D(0)\,,$$

or equivalently

$$D(t) = T_t\,D(0)T_{-t}.$$

Q.E.D.

Thus the time evolution of $(\mathcal{H}, H)$ brings about a systematic and continuous change in the density operator defining the statistical state. If the statistical state p is an equilibrium state, that is, if

$$p_t(M) = p_0(M)$$

for all $M \subset \mathfrak{M}$, then the equilibrium density operator D must satisfy

$$D(t) = D(0)$$

for all time t. It follows from Theorem 3-7 that equilibrium density operators commute with the dynamical group $\{T_t\}$. In the case of bounded self-adjoint operators, this is the defining property of the <u>integrals of the motion</u> of a quantum system. The simplest integrals of the motion are obtained by taking functions of the system Hamiltonian. For instance, if f is bounded and continuous on the spectrum of H, then [19, p. 346-347] we have

$$f(H)T_t = T_t f(H)$$

for all $t \in R$.

As in classical mechanics, the equilibrium density operator is often restricted to be a function of H, the physical justification being the same in quantum as in classical mechanics (see Chapter II). Thus if f is any non-negative function defined on the spectrum of H such that $f(H)$ is of trace class and trace $f(H) > 0$, the operator

$$D = \frac{f(H)}{\text{trace } f(H)}$$

serves as an equilibrium density operator. In particular, the Gibbs canonical state is given by the one-parameter family

$$D = \frac{e^{-\theta H}}{\text{trace } e^{-\theta H}},$$

where the parameter θ may assume any value in the open interval $(0, \infty)$ and determines the absolute temperature T according to

$$\theta = \frac{1}{kT},$$

k being Boltzmann's constant. Therefore the Gibbs canonical state is a continuous function of the Hamiltonian operator in the sense defined previously.

In the case of zero absolute temperature, the equilibrium density operator D_0 may be obtained as a limiting case of the Gibbs canonical state. Suppose the Hamiltonian operator of the system $(\mathcal{H}, H)$ has a pure point spectrum $\{\lambda_n\}$, with multiplicities $\{m_n\}$. For convenience let H be normalized such that the smallest spectral value λ_1 is zero. Taking the limit as $T \to 0$ in the canonical state with density

$$D_T = \frac{e^{-\frac{H}{kT}}}{\text{trace } e^{-\frac{H}{kT}}} = \frac{\sum_{n=1}^{\infty} e^{-\frac{\lambda_n}{kT}} P_n}{\sum_{n=1}^{\infty} m_n e^{-\frac{\lambda_n}{kT}}} ,$$

we obtain

$$D_0 = \lim_{T \to 0} D_T = \frac{1}{m_1} P_1 .$$

Thus, D_0 corresponds to a limiting state p_0 in which the probability of the characteristic subspace M_1, corresponding to the smallest spectral value λ_1, is 1:

$$p_0(M_1) = \text{trace } D_0 P_1 = \frac{1}{m_1} \text{ trace } P_1 = 1.$$

It is convenient to include this limiting state in the family of canonical states, and we shall make this inclusion in the sequel.

The density operator defining a canonical state of zero absolute temperature also has the form of a <u>microcanonical density operator.</u> These density operators are defined as follows: Consider a system $(\mathcal{H}, H)$ whose energy is fixed at some definite value E (a characteristic value of H). This means that the quantum mechanical state of $(\mathcal{H}, H)$ must be confined to the finite dimensional characteristic subspace $M_E \subset \mathcal{H}$ corresponding to the characteristic value E. The microcanonical state of $(\mathcal{H}, H)$ is then defined by the density operator $\frac{1}{m} P_{M_E}$, where P_{M_E} is the projector of $\mathcal{H}$ onto M_E, and $m = \dim M_E$.

The characterizations of the classical Gibbs canonical state presented in Chapter II extend in a more or less satisfactory way to quantum statistics as well. If D is a density operator defining a statistical state, having characteristic values $\{\lambda_k\}$, with multiplicities $\{m_k\}$, the entropy of D is defined by

$$S = -\text{ trace } (D \log D) = -\sum_{k=1}^{\infty} m_k \lambda_k \log \lambda_k .$$

In maximizing the entropy in quantum mechanics, one seeks a density operator D such that $-\text{trace } (D \log D)$ is as large as possible subject to the condition that the mean value of the Hamiltonian operator has a certain value. Using techniques from the calculus of variations adapted to operators, it can be shown [16, p. 112] that the canonical state satisfies this extremal problem provided the Hamiltonian satisfies certain conditions which are not highly restrictive.

Khinchin [12, p. 172-177] has also extended his characterization of the Gibbs state to quantum statistics. He again considers a large system which is perfectly isolated and composed of a large number of identical noninteracting components. The large system is assumed to be in a microcanonical state corresponding to an energy

E. His approach is formulated in a way that makes use of the discrete limit theorems of the theory of probability. However, to do this, he finds it necessary to assume that the spectrum of each component system Hamiltonian may be represented by non-negative integers (Khinchin argues that any energy spectrum can be made to approximate integers as closely as desired by choosing a sufficiently small unit of energy. His use of the discrete limit theorems also requires the assumption that the Hamiltonian of the large system is the sum of the component system Hamiltonians).

Under these conditions Khinchin shows that in the limit, as the number of component systems increases in such a way that the energy of the large system is proportional to the number of components, the statistical state of each component is a canonical state.

As we shall show in Chapter V, the method of statistical independence applied to a sufficiently large collection of quantum systems yields canonical states and does not require integer-valued energies. First, however, we must develop the theory of composite quantum systems.

CHAPTER IV. COMPOSITE QUANTUM SYSTEMS

In this chapter we shall consider composite systems which are composed of two component systems. The concepts of phase space, quantum mechanical state, and event must be given precise mathematical definitions, and the statistical states and observables of the composite system must be related to the statistical states and observables of the component systems. Clearly it is desirable to have a development analogous to the theory of measures on product spaces in classical mechanics, but which recognizes the non-Boolean nature of the lattice of events. We shall begin with a discussion of the abstract Hilbert space $\mathcal{H}_{12}$ of a composite system.

Consider two systems $(\mathcal{H}_1, H_1)$ and $(\mathcal{H}_2, H_2)$ which are to be thought of as forming one composite system $(\mathcal{H}_{12}, H_{12})$. The Hilbert space $\mathcal{H}_{12}$ is defined to be the "tensor product" of $\mathcal{H}_1$ and $\mathcal{H}_2$ and is denoted by $\mathcal{H}_1 \otimes \mathcal{H}_2$. Since our discussion will make extensive use of these products, we shall develop some of their properties (see also J. von Neumann and F. Murray [18]). To motivate the definition of tensor products, we begin by discussing a concrete example from quantum mechanics:

Let $\mathcal{H}_1$ and $\mathcal{H}_2$ have the coordinate representations $L^2(E_1^{3n_1}, \mu_1)$ and $L^2(E_2^{3n_2}, \mu_2)$. Then the coordinate representation of $\mathcal{H}_{12}$ is the set of all square summable complex-valued functions defined on the configuration space $(E_1^{3n_1} \times E_2^{3n_2})$ of the composite system, that is $L^2(E_1^{3n_1} \times E_2^{3n_2}, \mu_{12})$ where μ_{12} is the $3n_1 + 3n_2$ dimensional Lebesgue measure. To every pair of quantum mechanical states $\phi \varepsilon L^2(E_1^{3n_1}, \mu_1)$ and $\psi \varepsilon L^2(E_2^{3n_2}, \mu_2)$, there corresponds a quantum mechanical state in $L^2(E_1^{3n_1} \times E_2^{3n_2}, \mu_{12})$ which is the product $\phi \cdot \psi$. It can then be shown that if $\phi \times \psi$ denotes the "tensor product" of ϕ and ψ (see Definition (4-2)), then the transformation

$$\phi \times \psi \rightarrow \phi \cdot \psi$$

can be extended to an isometric isomorphism (congruence) of $L^2(E_1^{3n_1}, \mu_1) \otimes L^2(E_2^{3n_2}, \mu_2)$ onto $L^2(E_1^{3n_1} \times E_2^{3n_2}, \mu_{12})$. We may therefore identify the coordinate representation of $\mathcal{H}_{12}$ as the tensor product of the coordinate representations of $\mathcal{H}_1$ and $\mathcal{H}_2$.

Now let $\mathcal{H}_1$ and $\mathcal{H}_2$ be Hilbert spaces.

Definition 4-1. An anti-bilinear functional on the cartesian product $\mathcal{H}_1 \times \mathcal{H}_2$ is a complex valued function Φ such that

(a) $\Phi(a_1\phi_1 + a_2\phi_2, \psi) = \bar{a}_1 \Phi(\phi_1, \psi) + \bar{a}_2 \Phi(\phi_2, \psi)$

(b) $\Phi(\phi, a_1\psi_1 + a_2\psi_2) = \bar{a}_1 \Phi(\phi, \psi_1) + \bar{a}_2 \Phi(\phi, \psi_2)$,

where a_1 and a_2 are any complex numbers, and ϕ, ϕ_1, $\phi_2 \varepsilon \mathcal{H}_1$

and $\psi, \psi_1, \psi_2 \varepsilon \mathcal{H}_2$.

Every member of $\mathcal{H}_1 \otimes \mathcal{H}_2$ may be identified with a certain anit-bilinear functional on $\mathcal{H}_1 \times \mathcal{H}_2$. In particular, we have

Definition 4-2. Let $\phi \varepsilon \mathcal{H}_1$ and $\psi \varepsilon \mathcal{H}_2$. The tensor product of ϕ and ψ, denoted by $\phi \times \psi$, is the anti-bilinear functional defined by

$$(4.2.1) \qquad (\phi \times \psi)(\theta_1, \theta_2) = (\phi, \theta_1)(\psi, \theta_2)$$

for all pairs $(\theta_1, \theta_2) \varepsilon \mathcal{H}_1 \times \mathcal{H}_2$.

It follows from Definition 4-2 that if a is any complex number, the formally distinct expressions $(a\phi) \times \psi$, $\phi \times (a\psi)$, and $a(\phi \times \psi)$ are equal at all points of their domain.

Let $(\mathcal{H}_1 \otimes \mathcal{H}_2)'$ denote the set of all finite linear aggregates of the form

$$(4.2.2) \qquad \Phi = \sum_{i=1}^{p} (\phi_i \times \psi_i), \quad p = 1, 2, \cdots,$$

where $\{\phi_1, \phi_2, \cdots, \phi_p\} \subset \mathcal{H}_1$, and $\{\psi_1, \psi_2, \cdots, \psi_p\} \subset \mathcal{H}_2$. Clearly the representation of Φ need not be unique. However, if

$\Phi, \Psi \varepsilon (\mathcal{H}_1 \otimes \mathcal{H}_2)'$, that is $\Phi = \sum_{i=1}^{p} (\phi_i \times \psi_i)$ and $\Psi = \sum_{j=1}^{q} (\xi_j \times \eta_j)$,

for finite p, q, then the inner product (Φ, Ψ) is defined by

$$(\Phi, \Psi) = \sum_{i=1}^{p} \sum_{j=1}^{q} (\phi_i, \xi_j)\,(\psi_i, \eta_j) \tag{4.2.3}$$

and is independent of the representation of Φ and Ψ. With this definition of an inner product, $(\mathcal{H}_1 \otimes \mathcal{H}_2)'$ is a pre-Hilbert space and can be metrized by defining

$$\text{distance}\,(\Phi, \Psi) = ||\Phi - \Psi||, \tag{4.2.4}$$

where $||\Phi|| = (\Phi, \Phi)^{1/2}$.

It is a theorem [18] that if $\{\Phi_r\}$ is a sequence in $(\mathcal{H}_1 \otimes \mathcal{H}_2)'$ for which

$$\lim_{r, s \to \infty} ||\Phi_r - \Phi_s|| = 0 \tag{4.2.5}$$

(that is a Cauchy sequence),then there exists a unique anti-bilinear functional Φ on $\mathcal{H}_1 \otimes \mathcal{H}_2$ such that

$$\lim_{r \to \infty} \Phi_r(\theta_1, \theta_2) = \Phi(\theta_1, \theta_2) \tag{4.2.6}$$

for all $(\theta_1, \theta_2) \varepsilon\, \mathcal{H}_1 \times \mathcal{H}_2$. The set of all anti-bilinear functionals for which there exists a sequence $\{\Phi_r\} \subset (\mathcal{H}_1 \otimes \mathcal{H}_2)'$ satisfying (4.2.5) and (4.2.6) is called the tensor product of $\mathcal{H}_1$ and $\mathcal{H}_2$ and is denoted by $\mathcal{H}_1 \otimes \mathcal{H}_2$. If $\Phi, \Psi \varepsilon\, \mathcal{H}_1 \otimes \mathcal{H}_2$ correspond to the Cauchy sequences $\{\Phi_r\}$ and $\{\Psi_r\}$ in the sense of (4.2.5) and (4.2.6), then the inner product (Φ, Ψ), defined by

$$(4.2.7) \qquad (\Phi, \Psi) = \lim_{r \to \infty} (\Phi_r, \Psi_r)$$

exists and is independent of the sequences $\{\Phi_r\}$ and $\{\Psi_r\}$. That is, if $\{\Phi'_r\}$ and $\{\Psi'_r\}$ are any other sequences which correspond to Φ, Ψ in the sense of (4.2.5) and (4.2.6), then

$$(4.2.8) \qquad \lim_{r \to \infty} (\Phi'_r, \Psi'_r) = \lim_{r \to \infty} (\Phi_r, \Psi_r).$$

With this definition of the inner product, $\mathcal{H}_1 \otimes \mathcal{H}_2$ is a Hilbert space.

Thus for each $\Phi \varepsilon \mathcal{H}_1 \otimes \mathcal{H}_2$, considered as an anti-bilinear functional on $\mathcal{H}_1 \times \mathcal{H}_2$, there exists a Cauchy sequence $\{\Phi_r\}$ in $(\mathcal{H}_1 \otimes \mathcal{H}_2)'$ such that $\{\Phi_r\}$ converges pointwise to Φ. That $\mathcal{H}_1 \otimes \mathcal{H}_2$ is the ordinary metric completion of $(\mathcal{H}_1 \otimes \mathcal{H}_2)'$ depends on a proof that the correspondence between Φ and $\{\Phi_r\}$ in the sense of (4.2.5) and (4.2.6) is equivalent to $\lim_{r \to \infty} ||\Phi - \Phi_r|| = 0$, so that every Cauchy sequence in $(\mathcal{H}_1 \otimes \mathcal{H}_2)'$ also converges in norm to an element of $\mathcal{H}_1 \otimes \mathcal{H}_2$. It follows that every Cauchy sequence in $\mathcal{H}_1 \otimes \mathcal{H}_2$ converges in norm to an element of $\mathcal{H}_1 \otimes \mathcal{H}_2$; that is, as a metric space $\mathcal{H}_1 \otimes \mathcal{H}_2$ is complete.

In addition, each $\Phi \varepsilon \mathcal{H}_1 \otimes \mathcal{H}_2$ has the property that for all pairs $(\theta_1, \theta_2) \varepsilon \mathcal{H}_1 \times \mathcal{H}_2$, (4.2.9) $$\Phi(\theta_1, \theta_2) = (\Phi, \theta_1 \times \theta_2).$$

With Definition 4-2, this implies that Φ is a continuous function of θ_1 and θ_2. Moreover,

(4.2.10) $$|\Phi(\theta_1, \theta_2)| \leq \|\Phi\| \|\theta_1\| \|\theta_2\|.$$

Thus Φ is a bounded bilinear functional.

If $\{\phi_m\}$ and $\{\psi_n\}$ are complete orthonormal sequences in $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, then (see Lemma 4-6) the sequence $\{\phi_m \times \psi_n\}$ is a complete orthonormal sequence in $\mathcal{H}_1 \otimes \mathcal{H}_2$. Therefore, we may express any $\Phi \in \mathcal{H}_1 \otimes \mathcal{H}_2$ as

$$\Phi = \sum_{m,n=1}^{\infty} (\Phi, \phi_m \times \psi_n)(\phi_m \times \psi_n).$$

By (4.2.9), this is

$$\Phi = \sum_{m,n=1}^{\infty} \Phi(\phi_m, \psi_n)(\phi_m \times \psi_n).$$

Therefore each $\Phi \in \mathcal{H}_1 \otimes \mathcal{H}_2$ is completely characterized by the complex numbers $\Phi(\phi_m, \psi_n) = a_{mn}$, and $\sum_{m,n=1}^{\infty} |a_{mn}|^2$ converges.

Moreover, if $\{a_{mn}\}$ is any sequence of complex numbers such that $\sum_{m,n=1}^{\infty} |a_{mn}|^2$ converges, then there exists an anti-bilinear functional

$\Phi \varepsilon \mathcal{H}_1 \otimes \mathcal{H}_2$ such that $\Phi(\phi_m, \psi_n) = a_{mn}$. Clearly then

$$(4.2.11) \qquad ||\Phi||^2 = \sum_{m,n=1}^{\infty} |\Phi(\phi_m, \psi_n)|^2 .$$

Having defined the Hilbert space of the composite system $(\mathcal{H}_{12}, H_{12})$ to be $\mathcal{H}_1 \otimes \mathcal{H}_2$, it will be important for our discussion of statistical mechanics to identify a certain class of subspaces of $\mathcal{H}_1 \otimes \mathcal{H}_2$, given by

Definition 4-3. If M and N are closed linear subspaces of $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, then $(M \otimes N)'$ denotes the subset of $(\mathcal{H}_1 \otimes \mathcal{H}_2)'$ consisting of all finite linear aggregates of tensor products of the form $\phi \times \psi$, where $\phi \varepsilon M$ and $\psi \varepsilon N$. The symbol $M \otimes N$ will denote the closure of $(M \otimes N)'$.

Clearly $M \otimes N$ is a closed subspace of $\mathcal{H}_1 \otimes \mathcal{H}_2$.

The next step is to define certain operators on $\mathcal{H}_1 \otimes \mathcal{H}_2$. Our interest of course is in those operators which define the statistical states and observables of the composite system. The following lemmas will be needed for this development.

Lemma 4-4. Let $\{\xi_k\}$ and $\{\eta_\ell\}$ be sequences in $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, which are summable in the sense that

$\sum_{k=1}^{\infty} \xi_k = \phi \varepsilon \mathcal{H}_1$ and $\sum_{\ell=1}^{\infty} \eta_\ell = \psi \varepsilon \mathcal{H}_2$. Then the sums

$$\text{(a)} \sum_{k=1}^{\infty} \sum_{\ell=1}^{\infty} (\xi_k \times \eta_\ell), \quad \text{(b)} \sum_{\ell=1}^{\infty} \sum_{k=1}^{\infty} (\xi_k \times \eta_\ell), \quad \text{(c)} \sum_{k,\ell=1}^{\infty} \xi_k \times \eta_\ell$$

exist and are equal to $\phi \times \psi$.

Proof: The proof follows from the fact that a tensor product is a continuous function of its factors. Consider the sum $\sum_{\ell=1}^{\infty} (\xi_k \times \eta_\ell)$. Since tensor multiplication is distributive over finite sums, we have

$$\xi_k \times \psi = \lim_{N\to\infty} [\xi_k \times (\sum_{\ell=1}^{N} \eta_\ell)] = \lim_{N\to\infty} \sum_{\ell=1}^{N} (\xi_k \times \eta_\ell) = \sum_{\ell=1}^{\infty} (\xi_k \times \eta_\ell).$$

Hence $\sum_{\ell=1}^{\infty} (\xi_k \times \eta_\ell)$ exists. Then

$$\phi \times \psi = \lim_{N\to\infty} [(\sum_{k=1}^{N} \xi_k) \times \psi] = \lim_{N\to\infty} \sum_{k=1}^{N} (\xi_k \times \psi) = \lim_{N\to\infty} \sum_{k=1}^{N} \sum_{\ell=1}^{\infty} (\xi_k \times \eta_\ell)$$

$$= \sum_{k=1}^{\infty} \sum_{\ell=1}^{\infty} (\xi_k \times \eta_\ell).$$

A similar argument establishes that the sums (b) and (c) exist and are equal to $\phi \times \psi$.

Q. E. D.

Lemma 4-5. Let $\{\psi_1, \psi_2, \cdots \psi_p\}$ be a finite orthogonal sequence in $\mathcal{H}_2$, where $\psi_n \neq 0$ for all $n = 1, 2, \cdots p$. If $\sum_{n=1}^{p} (\phi_n \times \psi_n) = 0$ for some set $\{\phi_1, \phi_2, \cdots \phi_p\}$ in $\mathcal{H}_1$, then $\phi_n = 0$ for all n.

Proof: If $\sum_{n=1}^{p} (\phi_n \times \psi_n) = 0$, then for all pairs $(\theta_1, \theta_2) \epsilon \mathcal{H}_1 \times \mathcal{H}_2$ we have

$$\left(\sum_{n=1}^{p} (\phi_n \times \psi_n) \right) (\theta_1, \theta_2) = \sum_{n=1}^{p} (\phi_n, \theta_1) (\psi_n, \theta_2) = 0 .$$

Let $\theta_2 = \psi_1$, then

$$0 = \sum_{n=1}^{p} (\phi_n, \theta_1) (\psi_n, \psi_1) = (\phi_1, \theta_1) (\psi_1, \psi_1) .$$

Hence $(\phi_1, \theta_1) = 0$ for all $\theta_1 \epsilon \mathcal{H}_1$, and $(\phi_1, \phi_1) = 0$ implies $\phi_1 = 0$. If we successively let $\theta_2 = \psi_n$, $n = 1, 2, \cdots p$, we obtain the result:

$$\phi_1 = \phi_2 = \phi_3 = \cdots = \phi_p = 0.$$

Q. E. D.

Lemma 4-6. Let the complete orthonormal sequences $\{\phi_k\} \subset \mathcal{H}_1$ and $\{\psi_\ell\} \subset \mathcal{H}_2$ be given. Then the sequence $\{\phi_k \times \psi_\ell\}$ forms a complete orthonormal sequence in $\mathcal{H}_1 \otimes \mathcal{H}_2$.

Proof: That the sequence $\{\phi_k \times \psi_\ell\}$ is orthonormal is clear since, by (4.2.2), $(\phi_k \times \psi\ \ , \phi_m \times \psi_\ell) = (\phi_k, \phi_m)(\psi_\ell, \psi_n)$. This is zero if $\ell \neq n$ or $k \neq m$ and is equal to one if $k = m$ and $\ell = n$.

To show completeness, consider any $\Phi \varepsilon (\mathcal{H}_1 \otimes \mathcal{H}_2)'$. Then Φ has the representation

$$\Phi = \sum_{m=1}^{p} \xi_m \times \eta_m$$

for some finite p, where the sequences $\{\xi_1, \xi_2, \cdots \xi_p\}$ and $\{\eta_1, \eta_2, \cdots \eta_p\}$ are in $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively. If ξ_m and η_m have the expansions

$$\xi_m = \sum_{k=1}^{\infty} a_{mk} \phi_k, \qquad \eta_m = \sum_{\ell=1}^{\infty} \beta_{m\ell} \psi_\ell ,$$

then

$$\Phi = \sum_{m=1}^{p} \left[\left(\sum_{k=1}^{\infty} a_{mk} \phi_k \right) \times \left(\sum_{\ell=1}^{\infty} \beta_{mk} \psi_\ell \right) \right].$$

By Lemma 4.4, this is

$$\sum_{m=1}^{p}\left[\sum_{k=1}^{\infty}\sum_{\ell=1}^{\infty}(\alpha_{mk}\phi_k \times \beta_{m\ell}\psi_\ell)\right] = \sum_{m=1}^{p}\left[\sum_{k=1}^{\infty}\sum_{\ell=1}^{\infty}(\alpha_{mk}\beta_{m\ell})(\phi_k \times \psi_\ell)\right]$$

$$= \sum_{k=1}^{\infty}\sum_{\ell=1}^{\infty}\left[\sum_{m=1}^{p}(\alpha_{mk}\beta_{m\ell})\right](\phi_k \times \psi_\ell) .$$

If we let $\sum_{m=1}^{p}(\alpha_{mk}\beta_{m\ell}) = a_{k\ell}$, then

$$\Phi = \sum_{k=1}^{\infty}\sum_{\ell=1}^{\infty} a_{k\ell}(\phi_k \times \psi_\ell),$$

where $\sum_{k,\ell=1}^{\infty} |a_{k\ell}|^2$ converges. Hence every $\Phi \varepsilon (\mathcal{H}_1 \otimes \mathcal{H}_2)'$ has an expansion in terms of the orthonormal sequence $\{\phi_k \times \psi_\ell\}$. Therefore the set spanned by $\{\phi_k \times \psi_\ell\}$ contains $(\mathcal{H}_1 \otimes \mathcal{H}_2)'$ which is dense in $\mathcal{H}_1 \otimes \mathcal{H}_2$. Hence $\{\phi_k \times \psi_\ell\}$ is maximal with respect to the property of being orthonormal.

Q. E. D.

Remark: Lemma 4-6 generalizes to the case of tensor products of an arbitrary but finite number of Hilbert spaces, $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_N$. Therefore a finite tensor product of separable Hilbert spaces always yields a separable Hilbert space, and the phase space of any composite system composed of N components has the structure of a separable Hilbert space. However, infinite tensor products of Hilbert spaces (of dimension greater than one!) are always nonseparable. Since separability is required in our description of quantum mechanics, the approach presented here cannot be extended to composite systems composed of an infinite number of components.

Definition 4-7. Let A_1 be a linear operator with domain $\mathfrak{D}(A_1)$ contained in $\mathcal{H}_1$ and A_2 be a linear operator with domain $\mathfrak{D}(A_2)$ contained in $\mathcal{H}_2$. The operator $A_1 \times A_2$, which is the tensor product of the operators A_1 and A_2, is defined as follows: If $\{\phi_1, \cdots, \phi_m\}$ are elements of $\mathfrak{D}(A_1)$ and $\{\psi_1, \cdots, \psi_m\}$ are elements of $\mathfrak{D}(A_2)$, then for finite m,

$$(A_1 \times A_2) \sum_{i=1}^{m} (\phi_i \times \psi_i) = \sum_{i=1}^{m} (A_1\phi_i \times A_2\psi_i).$$

Remark: The reader is cautioned that in the case of self-adjoint operators we shall later extend the domain of $A_1 \times A_2$ so that this tensor product will be self-adjoint (see discussions following Theorems 4-10 and 4-14).

Clearly $A_1 \times A_2$ is a linear operator with domain $(\mathfrak{D}(A_1) \otimes \mathfrak{D}(A_2))'$ provided the correspondence

$$\sum_{i=1}^{m} (\phi_i \times \psi_i) \rightarrow \sum_{i=1}^{m} (A_1\phi_i \times A_2\psi_i)$$

is single valued. This is proved as the next lemma.

Lemma 4-8. Let $\Phi \varepsilon (\mathfrak{D}(A_1) \otimes \mathfrak{D}(A_2))'$, then $(A_1 \times A_2)\Phi$ is independent of the representation of Φ.

Proof: If we have two ways of writing an element $\Phi \varepsilon (\mathfrak{D}(A_1) \otimes \mathfrak{D}(A_2))'$, we may suppose that these are

$$\Phi = \sum_{i=1}^{m} (\phi_i \times \psi_i) = \sum_{i=m+1}^{q} ((-\phi_i) \times \psi_i)$$

where $\psi_i \neq 0$ for $i = 1, 2, \cdots, q$. Hence $\sum_{i=1}^{q} (\phi_i \times \psi_i) = 0$. We shall prove that $\sum_{i=1}^{q} (A_1\phi_i \times A_2\psi_i) = 0$. Let Q be the subspace spanned by $\{\psi_1, \cdots, \psi_q\}$. Clearly $1 \leq \dim Q \leq q$. Let $\dim Q = p$, and let $\{\eta_\ell : \ell = 1, 2, \cdots p\}$ be an orthonormal basis in Q. Then for each $i = 1, 2, \cdots q$, we have an expansion

$$\psi_i = \sum_{\ell=1}^{p} a_{i\ell}\eta_\ell ,$$

and we may write

$$0 = \sum_{i=1}^{q} \phi_i \times \psi_i = \sum_{i=1}^{q} (\phi_i \times (\sum_{\ell=1}^{p} a_{i\ell}\eta_\ell)) .$$

Since tensor multiplication is distributive with respect to finite sums, we have

$$0 = \sum_{i=1}^{q} \sum_{\ell=1}^{p} (\phi_i \times a_{i\ell}\eta_\ell).$$

Interchanging the order of summation and using the distributive law again gives

$$0 = \sum_{\ell=1}^{p} \sum_{i=1}^{q} (\phi_i \times a_{i\ell}\eta_\ell) = \sum_{\ell=1}^{p} \sum_{i=1}^{q} (a_{i\ell}\phi_i \times \eta_\ell) = \sum_{\ell=1}^{p} ((\sum_{i=1}^{q} a_{i\ell}\phi_i) \times \eta_\ell).$$

By lemma 4-5, we must have $\sum_{i=1}^{q} a_{i\ell}\phi_i = 0$ for each $\ell = 1, 2, \cdots, p$.

Therefore $A_1 \sum_{i=1}^{q} a_{i\ell}\phi_i = 0$ for each ℓ, and

$$0 = (A_1 \sum_{i=1}^{q} a_{i\ell}\phi_i) \times A_2\eta_\ell = (\sum_{i=1}^{q} a_{i\ell}A_1\phi_i) \times A_2\eta_\ell = \sum_{i=1}^{q} (a_{i\ell}A_1\phi_i \times A_2\eta_\ell)$$

$$= \sum_{i=1}^{q} (A_1\phi_i \times a_{i\ell}A_2\eta_\ell).$$

Therefore we may sum over ℓ to obtain

$$0 = \sum_{\ell=1}^{p} \sum_{i=1}^{q} (A_1\phi_i \times a_{i\ell}A_2\eta_\ell) = \sum_{i=1}^{q} \sum_{\ell=1}^{p} (A_1\phi_i \times a_{i\ell}A_2\eta_\ell)$$

$$= \sum_{i=1}^{q} (A_1\phi_i \times (\sum_{\ell=1}^{p} a_{i\ell}A_2\eta_\ell)) = \sum_{i=1}^{q} (A_1\phi_i \times (A_2 \sum_{\ell=1}^{p} a_{i\ell}\eta_\ell))$$

$$= \sum_{i=1}^{q} (A_1\phi_i \times A_2\psi_i).$$

Q. E. D.

We shall be concerned with cases where A_1 and A_2 correspond to observables and are therefore self-adjoint. The operator $A_1 \times A_2$ as defined above may not be self-adjoint, but in all cases of interest we can extend $A_1 \times A_2$ so as to obtain a self-adjoint operator (an observable of the composite system). We shall treat the case of bounded and unbounded operators separately.

A bounded operator defined on a domain $\mathcal{D} \subset \mathcal{H}$ always has a linear extension to all of $\mathcal{H}$. In fact, if $\mathcal{D}$ is dense in $\mathcal{H}$, the bounded operator has a unique extension to all of $\mathcal{H}$. Thus in considering bounded operators A_1 and A_2, there is no loss in generality in regarding them as everywhere defined in $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively. The following lemma is based on a proof by J. von Neumann and F. Murray [18] .

<u>Lemma 4-9.</u> Let A_1 and A_2 be bounded linear operators on $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, with norms $||A_1||$ and $||A_2||$. Then $A_1 \times A_2$ is a bounded linear operator on $(\mathcal{H}_1 \otimes \mathcal{H}_2)'$ with norm $||A_1 \times A_2|| \leq ||A_1|| \, ||A_2||$.

<u>Proof:</u> It follows from Definition 4-7 that $A_1 \times A_2$ is defined for all $\Phi \varepsilon (\mathcal{H}_1 \otimes \mathcal{H}_2)'$ and is linear also. If $\{\phi_m\}$ and $\{\psi_n\}$ are complete orthonormal sequences in $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, then by (4. 2. 11) we have

$$||(A_1 \times A_2)\Phi||^2 = \sum_{m,n=1}^{\infty} |a_{mn}|^2 = \sum_{m,n=1}^{\infty} |((A_1 \times A_2)\Phi)(\phi_m, \psi_n)|^2 .$$

Since $\Phi \varepsilon (\mathcal{H}_1 \otimes \mathcal{H}_2)'$, we may represent it as the finite sum

$$\Phi = \sum_{i=1}^{p} (\xi_i \times \eta_i).$$

Then

$$((A_1 \times A_2)\Phi)(\phi_m, \psi_n) = [(A_1 \times A_2) \sum_{i=1}^{p} (\xi_i \times \eta_i)](\phi_m, \psi_n)$$

$$= (\sum_{i=1}^{p} (A_1 \xi_i \times A_2 \eta_i))(\phi_m, \psi_n) = \sum_{i=1}^{p} (A_1 \xi_i \times A_2 \eta_i)(\phi_m, \psi_n).$$

By definition 4-2, this is

(4. 9. 1)
$$\sum_{i=1}^{p} (A_1 \xi_i, \phi_m)(A_2 \eta_i, \psi_n) .$$

Since A_1 and A_2 are bounded operators, their adjoints A_1^* and A_2^* are everywhere defined (and bounded with the same norms). Therefore (4. 9. 1) becomes

$$\sum_{i=1}^{p} (\xi_i, A_1^* \phi_m)(\eta_i, A_2^* \psi_n),$$

and by definition (4-2), this is

$$\sum_{i=1}^{p} (\xi_i \times \eta_i)\,(A_1^* \phi_m, A_2^* \psi_n).$$

Therefore

(4.9.2) $$\|(A_1 \times A_2)\Phi\|^2 = \sum_{m,n=1}^{\infty} \left| \left(\sum_{i=1}^{p} (\xi_i \times \eta_i)\,(A_1^* \phi_m, A_2^* \psi_n) \right|^2 .$$

Now for each $n = 1, 2, \cdots$, define

(4.9.3) $$f_n(\theta_1) = \left(\sum_{i=1}^{p} (\xi_i \times \eta_i)\right)(\theta_1, A_2^* \psi_n)$$

for all $\theta_1 \varepsilon \mathcal{H}_1$. Then $f_n(\theta_1)$ is a bounded anti-linear functional on $\mathcal{H}_1$. Consequently the conjugate functional $\overline{f}_n$ is a linear functional on $\mathcal{H}_1$, and by the Riesz representation theorem for bounded linear functionals [8, p. 31], there exists a unique $\theta_{1n}^0 \varepsilon \mathcal{H}_1$ (which depends on ψ_n) such that

(4.9.4) $$\overline{f}_n(\theta_1) = (\theta_1, \theta_{1n}^0)$$

for all θ_1. Hence

$$\left(\sum_{i=1}^{p} (\xi_i \times \eta_i)\right)(A_1^* \phi_m, A_2^* \psi_n) = f_n(A_1^* \phi_m) = \overline{\overline{f}}_n (A_1^* \phi_n) = (\theta_{1n}^0, A_1^*, \phi_m).$$

Therefore (4.9.2) becomes

$$||(A_1 \times A_2)\Phi||^2 = \sum_{m,n=1}^{\infty} |(\theta^0_{1n}, A_1^* \phi_m)|^2 = \sum_{m,n=1}^{\infty} |(A_1\theta^0_{1n}, \phi_m)|^2$$

$$= \sum_{n=1}^{\infty} ||A_1\theta^0_{1n}||^2 \leq ||A_1||^2 \sum_{n=1}^{\infty} (\theta^0_{1n}, \theta^0_{1n})$$

$$= ||A_1||^2 \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} (\theta_{1n}, \phi_m)(\phi_m, \theta^0_{1n})$$

$$= ||A_1||^2 \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} |(\theta^0_{1n}, \phi_m)|^2 .$$

By (4. 9. 3), this is

$$||A_1||^2 \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} | f_n(\phi_m)|^2,$$

and by (4. 9. 2), we obtain

$$(4.9.5) \quad ||(A_1 \times A_2)\Phi||^2 \leq ||A_1||^2 \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} |(\sum_{i=1}^{p} (\xi_i \times \eta_i)(\phi_m, A_2^* \psi_n)|^2 .$$

Now for each fixed ϕ_m, $m = 1, 2, \cdots$, define

$$(4.9.6) \qquad g_m(\theta_2) = (\sum_{i=1}^{p} (\xi_i \times \eta_i)) (\phi_m, \theta_2)$$

for all $\theta_2 \varepsilon \mathcal{H}_2$. Then g_m is a bounded anti-linear functional on

$\mathcal{H}_2$, and, as before, by the Riesz representation theorem, there exists a unique $\theta^0_{2m} \varepsilon \mathcal{H}_2$ (which depends on ϕ_m) such that

$$(4.9.7) \qquad g_m(\theta_2) = (\theta^0_{2m}, \theta_2) .$$

Hence we have

$$\left(\sum_{i=1}^{p} (\xi_i \times \eta_i)\right)(\phi_m, A_2^* \psi_n) = g_m(A_2^*\psi_n) = (\theta^0_{2m}, A_2^*\psi_n).$$

Therefore (4.9.5) becomes

$$||(A_1 \times A_2)\Phi||^2 \leq ||A_1||^2 \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} |(\theta^0_{2m}, A_2^*\psi_n)|^2 = ||A_1||^2 \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} |(A_2\theta^0_{2m}, \psi_n)|^2$$

$$= ||A_1||^2 \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} |(A_2\theta^0_{2m}, \psi_n)|^2 = ||A_1||^2 \sum_{m=1}^{\infty} ||A_2\theta^0_{2m}||^2$$

$$\leq ||A_1||^2 ||A_2||^2 \sum_{m=1}^{\infty} ||\theta^0_{2m}||^2 = \|A_1\|^2 \|A_2\|^2 \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} |(\theta^0_{2m}, \psi_n)|^2 .$$

By (4.9.6) and (4.9.7), we obtain

$$\|(A_1 \times A_2)\Phi\|^2 \leq \|A_1\|^2 \|A_2\|^2 \sum_{m,n=1}^{\infty} \left|\left(\sum_{i=1}^{p} (\xi_i \times \eta_i)\right)(\phi_m, \psi_n)\right|^2 = \|A_1\|^2 \|A_2\|^2 \|\Phi\|^2 .$$

The last equality is obtained using (4.2.11). Therefore

$$\|(A_1 \times A_2)\Phi\| \leq \|A_1\| \; \|A_2\| \; \|\Phi\|$$

for all $\Phi \varepsilon (\mathcal{H}_1 \otimes \mathcal{H}_2)'$.

Q. E. D.

We are now able to prove the following theorem for bounded operators:

Theorem 4-10. Let A_1 and A_2 be bounded linear operators on $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively. Then there exists a unique bounded linear operator A_{12} on $\mathcal{H}_1 \otimes \mathcal{H}_2$ such that $A_{12}\Phi = (A_1 \times A_2)\Phi$ for all $\Phi \varepsilon (\mathcal{H}_1 \otimes \mathcal{H}_2)'$.

Proof: By Lemmas 4-8 and 4-9, $A_1 \times A_2$ is uniquely defined and bounded on $(\mathcal{H}_1 \otimes \mathcal{H}_2)'$. Since $(\mathcal{H}_1 \otimes \mathcal{H}_2)'$ is dense in $\mathcal{H}_1 \otimes \mathcal{H}_2$, $A_1 \times A_2$ has a unique bounded linear extension to all of $\mathcal{H}_1 \otimes \mathcal{H}_2$.

Q. E. D.

In view of Theorem 4-10, we shall, in the case of bounded operators, henceforth identify $A_1 \times A_2$ with its bounded extension A_{12}. In the case where the bounded operators A_1 and A_2 are also self-adjoint, the self-adjointness of $A_1 \times A_2$ now follows:

Theorem 4-11. If A_1 and A_2 are bounded self-adjoint operators on $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, then $A_1 \times A_2$ is a bounded self-adjoint operator on $\mathcal{H}_1 \otimes \mathcal{H}_2$.

Proof: Recall first that in the case of bounded everywhere defined linear operators, the self-adjoint operators are characterized by the relation

$$(Af, g) = (f, Ag),$$

which holds for all $f, g \varepsilon \mathcal{H}$.

Now by Theorem 4-10 and the discussion following, $A_1 \times A_2$ is bounded and everywhere defined in $\mathcal{H}_1 \otimes \mathcal{H}_2$. Consider any pair Φ, $\Psi \varepsilon (\mathcal{H}_1 \otimes \mathcal{H}_2)'$. Then for some finite p, q, we have the representations

$$\Phi = \sum_{j=1}^{p} (\phi_j \times \psi_j) \quad \text{and} \quad \Psi = \sum_{k=1}^{q} (\xi_k \times \eta_k).$$

Then

$$((A_1 \times A_2)\Phi, \Psi) = (\sum_{j=1}^{p} (A_1\phi_j \times A_2\psi_j), \sum_{k=1}^{q} (\xi_k \times \eta_k))$$

$$= \sum_{j=1}^{p}\sum_{k=1}^{q} (\phi_j, A_1\xi_k)(\psi_j, A_2\eta_k) = \sum_{j=1}^{p}\sum_{k=1}^{q} (\phi_j \times \psi_j, A_1\xi_k \times A_2\eta_k)$$

$$= (\sum_{j=1}^{p} (\phi_j \times \psi_j), \sum_{k=1}^{q} (A_1\xi_k \times A_2\eta_k))$$

$$= (\sum_{j=1}^{p} (\phi_j \times \psi_j), (A_1 \times A_2) \sum_{k=1}^{q} (\xi_k \times \eta_k)).$$

Therefore

$$((A_1 \times A_2)\Phi, \Psi) = (\Phi, (A_1 \times A_2)\Psi) \tag{4.11.1}$$

for all $\Phi, \Psi \varepsilon (\mathcal{H}_1 \otimes \mathcal{H}_2)'$. We shall show that this relation holds for all $\Phi, \Psi \varepsilon \mathcal{H}_1 \otimes \mathcal{H}_2$. Consider any pair (Φ, Ψ) in $\mathcal{H}_1 \otimes \mathcal{H}_2$. Then there exist Cauchy sequences $\{\Phi_r\}$ and $\{\Psi_r\}$ in $(\mathcal{H}_1 \otimes \mathcal{H}_2)'$ such that

$$\lim_{r \to \infty} \|\Phi - \Phi_r\| = 0 \quad \text{and} \quad \lim_{r \to \infty} \|\Psi - \Psi_r\| = 0.$$

Since $A_1 \times A_2$ is bounded, it is also continuous. Hence

$$\lim_{r \to \infty} \|(A_1 \times A_2)\Phi - (A_1 \times A_2)\Phi_r\| = 0$$

$$\lim_{r \to \infty} \|(A_1 \times A_2)\Psi - (A_1 \times A_2)\Psi_r\| = 0 .$$

Then, by (4.2.7),

$$\lim_{r \to \infty} ((A_1 \times A_2)\Phi_r, \Psi_r) = ((A_1 \times A_2)\Phi, \Psi)$$

$$\lim_{r \to \infty} (\Phi_r, (A_1 \times A_2)\Psi_r) = (\Phi, (A_1 \times A_2)\Psi).$$

By (4.11.1),

$$((A_1 \times A_2)\Phi_r, \Psi_r) = (\Phi_r, (A_1 \times A_2)\Psi_r)$$

for each $r = 1, 2, \cdots$. Hence

$$((A_1 \times A_2)\Phi, \Psi) = (\Phi, (A_1 \times A_2)\Psi)$$

for all $\Phi, \Psi \in \mathcal{H}_1 \otimes \mathcal{H}_2$.

Q. E. D.

We conclude our discussion of bounded operators with the following theorem, which will be needed later.

Theroem 4-12. Let $P_M, P_N,$ and $P_{M \otimes N}$ denote the projection operators onto the closed subspaces $M \subset \mathcal{H}_1$, $N \subset \mathcal{H}_2$, and $M \otimes N \subset \mathcal{H}_1 \otimes \mathcal{H}_2$. Then

$$P_M \times P_N = P_{M \otimes N} .$$

Proof: Consider first those functionals $\Phi \varepsilon \mathcal{H}_1 \otimes \mathcal{H}_2$ which are of the form $\phi \times \psi$ with $\phi \varepsilon \mathcal{H}_1$ and $\psi \varepsilon \mathcal{H}_2$. We have the unique decompositions

$$\phi = P_M\phi + \phi_1, \quad \text{where} \quad \phi_1 \perp M$$

$$\psi = P_N\psi + \psi_1, \quad \text{where} \quad \psi_1 \perp N .$$

Therefore

(4.12.1) $$\phi\times\psi = P_M\phi\times P_N\psi + P_M\phi\times\psi_1 + \phi_1\times P_N\psi + \phi_1\times\psi_1,$$

where $P_M\phi\times P_N\psi \in M\otimes N$, and $(P_M\phi\times\psi_1+\phi_1\times P_N\psi+\phi_1\times\psi_1) \perp M\otimes N$.

Since the decomposition (4.12.1) into the sum of a vector in $M\otimes N$ and a vector orthogonal to $M\otimes N$ is uniquely expressed as

$$\phi\times\psi = P_{M\otimes N}(\phi\times\psi) + (\phi\times\psi)_1, \quad \text{where } (\phi\times\psi)_1 \perp M\otimes N,$$

we have

$$P_{M\otimes N}(\phi\times\psi) = P_M\phi\times P_N\psi = (P_M\times P_N)(\phi\times\psi)$$

for all $\Phi\in \mathcal{H}_1\otimes\mathcal{H}_2$ of the form $\phi\times\psi$. Therefore $P_M\times P_N = P_{M\otimes N}$ everywhere in $(\mathcal{H}_1\otimes\mathcal{H}_2)'$. The continuity of both $P_{M\otimes N}$ and $P_M\times P_N$ establishes that

$$P_M\times P_N = P_{M\otimes N}$$

everywhere in $\mathcal{H}_1\otimes\mathcal{H}_2$.

Q.E.D.

The situation for unbounded operators A_1 and A_2 is completely different. We first summarize some of the theorems in unbounded operator theory [19, p. 305-309]. Recall that an unbounded self-adjoint operator A on a Hilbert space $\mathcal{H}$ cannot be everywhere defined since a self-adjoint operator is necessarily closed, and a closed everywhere defined linear operator is necessarily bounded. We shall say that A is symmetric if its domain $\mathcal{D}(A)$ is dense in $\mathcal{H}$ and the adjoint operator A^* is an extension of A. This definition is equivalent to saying that A is symmetric if and only if

$$(4.12.2) \qquad (Af, g) = (f, Ag)$$

for all $f, g \in \mathcal{D}(A)$ and $\mathcal{D}(A)$ is dense in $\mathcal{H}$. Moreover, every symmetric operator has a smallest closed linear extension, namely $(A^*)^*$. The operator A is essentially self-adjoint if $A^* = (A^*)^*$, and for symmetric operators this relation holds if and only if the spectrum of A is confined to the real axis [21, p. 144].

We shall show in Lemma 4-13 that $A_1 \times A_2$ is not bounded if either A_1 or A_2 is not bounded. Hence every closed linear extension of $A_1 \times A_2$ must be defined on a proper subset of $\mathcal{H}_1 \otimes \mathcal{H}_2$. We will consider the case where A_1 and A_2 have complete sequences of characteristic vectors (that is, pure point spectra) and show that $A_1 \times A_2$ is essentially self-adjoint, so that the adjoint operator $(A_1 \times A_2)^*$ is the smallest closed extension of $A_1 \times A_2$ and is self-adjoint.

Lemma 4-13. Let $A_1 \neq 0$ and $A_2 \neq 0$ have the domains $\mathcal{D}(A_1)$ and $\mathcal{D}(A_2)$ (not necessarily dense in $\mathcal{H}_1$ and $\mathcal{H}_2$). If either A_1 or A_2 is not bounded, then $A_1 \times A_2$ is not bounded.

Proof: Recall that the domain of $A_1 \times A_2$ is $(\mathcal{D}(A_1) \otimes \mathcal{D}(A_2))'$. Suppose A_1 is not bounded. Then there exists a sequence $\{\xi_r\} \subset \mathcal{D}(A_1)$ such that

$$\|A_1 \xi_r\| > r \|\xi_r\| .$$

Let $\{\eta_r\}$ be a constant sequence $\{\eta, \eta, \cdots\}$ in $\mathcal{D}(A_2)$ such that $\|A_2 \eta\| = M\|\eta\|$ and $M > 0$. Then the sequence $\{\xi_r \times \eta_r\} \subset (\mathcal{D}(A_1) \otimes \mathcal{D}(A_2))'$, and

$$\|(A_1 \times A_2)(\xi_r \times \eta_r)\| = \|A_1 \xi_r \times A_2 \eta_r\| = \|A_1 \xi_r\| \, \|A_2 \eta_r\|$$

$$> Mr\|\xi_r\| \, \|\eta_r\| = Mr\|\xi_r \times \eta_r\| .$$

Hence there exists a sequence $\{\Phi_r\} \subset (\mathcal{D}(A_1) \otimes \mathcal{D}(A_2))'$ such that the sequence $\{(A_1 \times A_2)\Phi\}$ is not bounded. The case where A_2 is not bounded is treated in a similar way.

Q. E. D.

Lemma 4-14. Let A_1 and A_2 be self-adjoint operators having complete sequences of characteristic vectors, whose domains $\mathcal{D}(A_1)$ and $\mathcal{D}(A_2)$ are dense in $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively. Then $A_1 \times A_2$ is an essentially self-adjoint operator whose domain $(\mathcal{D}(A_1) \otimes \mathcal{D}(A_2))'$ is dense in $\mathcal{H}_1 \otimes \mathcal{H}_2$.

Proof: Since A_1 and A_2 are linear operators, their domains $\mathcal{D}(A_1)$ and $\mathcal{D}(A_2)$ are linear subspaces (not necessarily closed). Since $\mathcal{D}(A_1)$ and $\mathcal{D}(A_2)$ are dense in $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, there exist complete orthonormal sequences $\{\phi_m\} \subset \mathcal{D}(A_1)$ and $\{\psi_n\} \subset \mathcal{D}(A_2)$. Consequently the sequence $\{\phi_m \times \psi_n\}$ is a complete orthonormal sequence contained in $(\mathcal{D}(A_1) \otimes \mathcal{D}(A_2))'$. It follows that $(\mathcal{D}(A_1) \otimes \mathcal{D}(A_2))'$ is dense in $\mathcal{H}_1 \otimes \mathcal{H}_2$.

Let $\Phi, \Psi \varepsilon (\mathcal{D}(A_1) \otimes \mathcal{D}(A_2))'$. Then, as in the proof of Theorem 4-11, we obtain

$$((A_1 \times A_2)\Phi, \Psi) = (\Phi, (A_1 \times A_2)\Psi)$$

for all $\Phi, \Psi \varepsilon (\mathcal{D}(A_1) \otimes \mathcal{D}(A_2))'$ Hence, by (4. 12. 2), $A_1 \times A_2$ is symmetric. To show that $A_1 \times A_2$ is essentially self-adjoint, we need only verify that its spectrum is confined to the real axis (see remarks preceding Lemma 4-13). We observe at once that $A_1 \times A_2$ admits a complete sequence of characteristic vectors: Let $\{\phi_m\}$ and $\{\psi_n\}$ be complete sequences of characteristic vectors of A_1 and A_2 respectively with the corresponding characteristic values $\{\lambda_{1m}\}$ and $\{\lambda_{2n}\}$. Therefore the sequence $\{\phi_m \times \psi_n\}$ is a complete orthonormal sequence in $(\mathcal{D}(A_1) \otimes \mathcal{D}(A_2))'$, and we have

$$(A_1 \times A_2)(\phi_m \times \psi_n) = A_1\phi_m \times A_2\psi_n = \lambda_{1m}\phi_m \times \lambda_{2n}\psi_n = \lambda_{1m}\lambda_{2n}(\phi_m \times \psi_n)$$

for all $m, n = 1, 2, \cdots$. Hence the characteristic vectors of $A_1 \times A_2$ form a complete set in $\mathcal{H}_1 \otimes \mathcal{H}_2$. This means [19, p. 361] that the spectrum of $A_1 \times A_2$ consists of the characteristic values $\{\lambda_{1m}\lambda_{2n}: m, n=1, 2, \cdots\}$ plus their points of accumulation. Hence the spectrum of $A_1 \times A_2$ is a subset of the real axis.

Q. E. D.

We now complete the definition of $A_1 \times A_2$ for self-adjoint operators by generalizing the remark following Theorem 4-10 to include unbounded operators. Hereafter, the symbol $A_1 \times A_2$ will denote the smallest self-adjoint extension $(A_1 \times A_2)^*$ of the operator defined in Definition 4-7.

Returning now to the discussion of quantum systems, the quantum mechanical states of $(\mathcal{H}_{12}, H_{12})$ are the unit vectors in the phase space $\mathcal{H}_1 \otimes \mathcal{H}_2$. The events related to the composite system are the closed subspaces of $\mathcal{H}_1 \otimes \mathcal{H}_2$. If M and N are events related to the component systems, then $M \otimes N$ is the event that the quantum mechanical states of the component systems are in M and N respectively. Since the quantum mechanical state of the second system is always in $\mathcal{H}_2$, the event $M \otimes \mathcal{H}_2$ is the event that the quantum mechanical state of the first system is in M. The event $\mathcal{H}_1 \otimes N$ is interpreted in a similar way.

Each statistical state p_{12} of the composite system determines not only the statistical properties of the composite system, but the statistical properties of the component systems as well. In other words, p_{12} induces statistical states p_1 in $(\mathcal{H}_1, H_1)$

and p_2 in $(\mathcal{H}_2, H_2)$ in such a way that the corresponding events M and $M \otimes \mathcal{H}_2$ have the same probability:

<u>Theorem 4-15.</u> Let p_{12} be a probability measure on the closed subspaces of $\mathcal{H}_1 \otimes \mathcal{H}_2$. Then the set function $p_{12}(M \otimes \mathcal{H}_2)$, defined for all closed subspaces $M \subset \mathcal{H}_1$, is a probability measure. Alternatively, the set function $p_{12}(\mathcal{H}_1 \otimes N)$, defined for all closed subspaces $N \subset \mathcal{H}_2$, is a probability measure.

<u>Proof:</u> We shall prove only the first of these consequences, the alternative case may be disposed of in a similar way.

Define $p_1(M) = p_{12}(M \otimes \mathcal{H}_2)$ for all $M \subset \mathcal{H}_1$. Clearly $0 \leq p_1(M) \leq 1$ for all $M \subset \mathcal{H}_1$. If $M = \mathcal{H}_1$, then $p_1(\mathcal{H}_1) = p_{12}(\mathcal{H}_1 \otimes \mathcal{H}_2) = 1$, and if $M = \{0\}$, then $p_1(\{0\}) = p_{12}(\{0\} \otimes \mathcal{H}_2) = p_{12}(\{0\}) = 0$. If $\{M_n\}$ is any countable sequence of mutually orthogonal closed subspaces in $\mathcal{H}_1$ having closed linear span M, then the sequence $\{M_n \otimes \mathcal{H}_2\}$ is a countable sequence of mutually orthogonal closed subspaces in $\mathcal{H}_1 \otimes \mathcal{H}_2$ having closed linear span $M \otimes \mathcal{H}_2$. Consequently

$$p_1(M) = p_{12}(M \otimes \mathcal{H}_2) = \sum_{n=1}^{\infty} p_{12}(M_n \otimes \mathcal{H}_2) = \sum_{n=1}^{\infty} p_1(M_n).$$

Therefore p_1 defines a probability measure on the closed subspaces of $\mathcal{H}_1$.

Q. E. D.

Definition 4-16. Let p_{12} be a statistical state of the composite system $(\mathcal{H}_{12}, H_{12})$ composed of two component systems. Then the statistical states p_1 and p_2 of $(\mathcal{H}_1, H_1)$ and $(\mathcal{H}_2, H_2)$, defined by

$$p_1(M) = p_{12}(M \otimes \mathcal{H}_2)$$

$$p_2(N) = p_{12}(\mathcal{H}_1 \otimes N)$$

for all $M \subset \mathcal{H}_1$ and $N \subset \mathcal{H}_2$, are called the induced statistical states of p_{12}.

Throughout this discussion we are assuming that there is a one-to-one correspondence between the observables of a system $(\mathcal{H}, H)$ and the self-adjoint operators defined on $\mathcal{H}$. Consequently everything that has been said concerning self-adjoint operators on $\mathcal{H}_1$, $\mathcal{H}_2$, and $\mathcal{H}_1 \otimes \mathcal{H}_2$ has an interpretation in terms of observables of the respective systems. In particular, every observable A_1 of the component system $(\mathcal{H}_1, H_1)$ is also an observable of $(\mathcal{H}_{12}, H_{12})$; the correspondence between the self-adjoint operators A_1 and $A_1 \times I_2$ is interpreted to mean that these operators represent the same observable. For example, "the energy of system $(\mathcal{H}_1, H_1)$" is an observable of $(\mathcal{H}_{12}, H_{12})$ as well as $(\mathcal{H}_1, H_1)$. The operators H_1 and $H_1 \times I_2$ both represent this observable. The same interpretation is made for the correspondence between the operators H_2 and $I_1 \times H_2$. The operators $H_1 \times I_2 + I_1 \times H_2$ also defines an observable, in view of

Theorem 4-17. Let H_1 and H_2 be Hamiltonian operators with domains $\mathcal{D}(H_1)$ in $\mathcal{H}_1$ and $\mathcal{D}(H_2)$ in $\mathcal{H}_2$, having complete orthonormal sequences of characteristic vectors $\{\phi_m\}$ and $\{\psi_n\}$ respectively. Let $S(H_1) = \{\lambda_{1m}\}$ and $S(H_2) = \{\lambda_{2n}\}$. Then

(a) The operator $H_1 \times I_2 + I_1 \times H_2$ is essentially self-adjoint, and its spectrum consists of the sequence $\{\lambda_{1m} + \lambda_{2n} : m, n = 1, 2, \cdots\}$ plus its accumulation points.

(b) If the smallest characteristic values of H_1 and H_2 are both zero, then the smallest characteristic value of $H_1 \times I_2 + I_1 \times H_2$ is also zero.

Proof: (a) Clearly the sequence $\{\phi_m \times \psi_n\}$ is simultaneously a complete orthonormal sequence of characteristic vectors of both $H_1 \times I_2$, $I_1 \times H_2$ and therefore of $H_1 \times I_2 + I_1 \times H_2$ also. For each $m, n = 1, 2, \cdots$, we have

$$
\begin{aligned}
(H_1 \times I_2 + I_1 \times H_2)(\phi_m \times \psi_n) &= (H_1 \times I_2)(\phi_m \times \psi_n) + (I_1 \times H_2)(\phi_m \times \psi_n) \\
&= (H_1\phi_m \times I_2\psi_n) + (I_1\phi_m \times H_2\psi_n) \\
&= \lambda_{1m}(\phi_m \times \psi_n) + \lambda_{2n}(\phi_m \times \psi_n) \\
&= (\lambda_{1m} + \lambda_{2n})(\phi_m \times \psi_n) .
\end{aligned}
$$

Consequently [19, p. 361] the spectrum of $H_1 \times I_2 + I_1 \times H_2$ consists of the sequence $\{\lambda_{1m} + \lambda_{2n} : m, n = 1, 2, \cdots\}$ plus its accumulation points. Now $\mathcal{D}(H_1 \times I_2 + I_1 \times H_2)$ contains all $\Phi \varepsilon \mathcal{H}_1 \otimes \mathcal{H}_2$ which are common to $\mathcal{D}(H_1 \times I_2)$ and $\mathcal{D}(I_1 \times H_2)$. Therefore

$$(\mathcal{D}(H_1) \otimes \mathcal{D}(H_2))' \subset \mathcal{D}(H_1 \times I_2 + I_1 \times H_2).$$

By Lemma 4-14, $(\mathcal{D}(H_1) \otimes \mathcal{D}(H_2))'$ is dense in $\mathcal{H}_1 \otimes \mathcal{H}_2$ and both $H_1 \times I_2$ and $I_1 \times H_2$ are self-adjoint. Therefore $\mathcal{D}(H_1 \times I_2 + I_1 \times H_2)$ is dense in $\mathcal{H}_1 \otimes \mathcal{H}_2$ and $H_1 \times I_2 + I_1 \times H_2$ is symmetric. The essential self-adjointness of $H_1 \times I_2 + I_1 \times H_2$ follows since its spectrum is confined to the real axis.

The proof of (b) follows as a simple corollary.

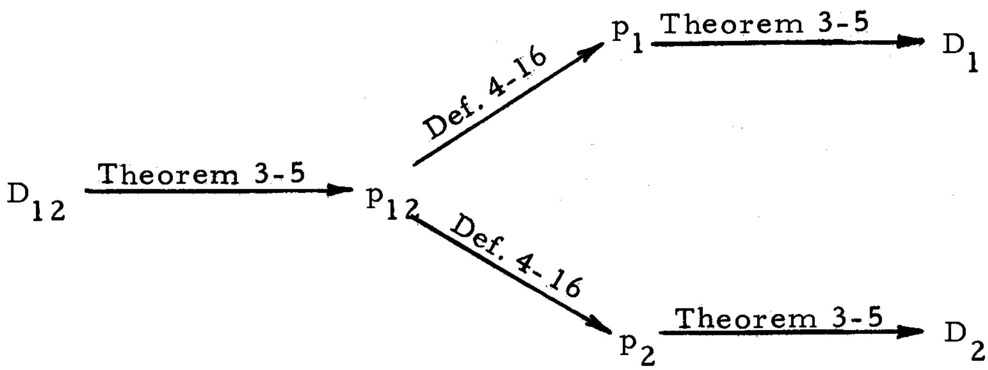
Q. E. D.

If we again identify $H_1 \times I_2 + I_1 \times H_2$ with the smallest closed extension, then we interpret this operator as the observable of the composite system $(\mathcal{H}_{12}, H_{12})$ which is "the sum of the energies of $(\mathcal{H}_1, H_1)$ and $(\mathcal{H}_2, H_2)$". In the general case, $H_{12} \neq H_1 \times I_2 + I_1 \times H_2$, the difference between them being the <u>interaction energy</u> of the combined system. For noninteracting systems we have

<u>Definition 4-18</u>. A composite system $(\mathcal{H}_{12}, H_{12})$ has

noninteracting components $(\mathcal{H}_1, H_1)$ and $(\mathcal{H}_2, H_2)$ if and only if $H_{12} = H_1 \times I_2 + I_1 \times H_2$.

If the composite system is considered as a single system, then each statistical state of $(\mathcal{H}_{12}, H_{12})$ is a probability measure on the closed subspaces of $\mathcal{H}_1 \otimes \mathcal{H}_2$. Theorem 3-5 then identifies each statistical state with a density operator D_{12} on $\mathcal{H}_1 \otimes \mathcal{H}_2$. The induced statistical states p_1 in $(\mathcal{H}_1, H_1)$ and p_2 in $(\mathcal{H}_2, H_2)$, given by Definition 4-16, are then in unique correspondence with density operators D_1 on $\mathcal{H}_1$ and D_2 on $\mathcal{H}_2$ respectively. The logical system is



The induced density operators D_1 and D_2 are the uniquely determined operators on $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, whose traces with all projection operators satisfy

$$\text{trace}\,(D_1 P_M) = \text{trace}\,(D_{12} P_{M \otimes \mathcal{H}_2})$$

$$\text{trace}\,(D_2 P_N) = \text{trace}\,(D_{12} P_{\mathcal{H}_1 \otimes N}) \, .$$

In keeping with the analogy between <u>the integral</u> in classical mechanics and <u>the trace</u> in quantum mechanics, it is appropriate to define two operator-valued functions $\mathrm{tr}_{\mathcal{H}_1}$ and $\mathrm{tr}_{\mathcal{H}_2}$ on the set of all composite density operators as follows (compare with (2. 11) and (2. 12):

$$D_1 = \mathrm{tr}_{\mathcal{H}_2} D_{12}$$

$$D_2 = \mathrm{tr}_{\mathcal{H}_1} D_{12} .$$

Although the density operator D_{12} of the composite system induces unique densities D_1 and D_2 on the components, it is not true that the induced states D_1 and D_2 determine D_{12} uniquely (see example below). However, if D_1 and D_2 are given, there is always at least one density operator D_{12} which induces D_1 and D_2; namely, the tensor product $D_1 \times D_2$:

<u>Theorem 4-19</u>. If D_1 and D_2 are density operators defined on $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, then the operator $D_1 \times D_2$ is a density operator on $\mathcal{H}_1 \otimes \mathcal{H}_2$. Moreover,

$$D_1 = \mathrm{tr}_{\mathcal{H}_2} (D_1 \times D_2)$$

$$D_2 = \mathrm{tr}_{\mathcal{H}_1} (D_1 \times D_2)$$

Proof: Since D_1 and D_2 are both bounded and self-adjoint, by Theorem 4-11 $D_1 \times D_2$ is a bounded self-adjoint operator defined everywhere in $\mathcal{H}_1 \otimes \mathcal{H}_2$. To show that $D_1 \times D_2$ is non-negative, let $\{\phi_m\}$ and $\{\psi_n\}$ be orthonormal bases in $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, composed of characteristic vectors of D_1 and D_2. Let $\{\lambda_{1m}\}$ and $\{\lambda_{2n}\}$ denote the corresponding sequences of characteristic values. Since D_1 and D_2 are each non-negative, $\lambda_{1m}, \lambda_{2n} \geq 0$ for all $m, n = 1, 2, \cdots$. Let $\Phi \varepsilon\, \mathcal{H}_1 \otimes \mathcal{H}_2$. Then for some sequence $\{a_{mn}\}$, we have

$$\Phi = \sum_{m,n=1}^{\infty} a_{mn}(\phi_m \times \psi_n),$$

and

$$(D_1 \times D_2)\Phi = \sum_{m,n=1}^{\infty} a_{mn}(D_1\phi_m \times D_2\psi_n) = \sum_{m,n=1}^{\infty} a_{mn}(\lambda_{1m}\phi_m \times \lambda_{2n}\psi_n)$$

$$= \sum_{m,n=1}^{\infty} a_{mn}\lambda_{1m}\lambda_{2n}(\phi_m \times \psi_n).$$

Therefore

$$((D_1 \times D_2)\Phi, \Phi) = \Big(\sum_{m,n=1}^{\infty} a_{mn}\lambda_{1m}\lambda_{2n}(\phi_m \times \psi_n), \sum_{k,\ell=1}^{\infty} a_{k\ell}(\phi_k \times \psi_\ell)\Big)$$

$$= \sum_{m,n=1}^{\infty}\sum_{k,\ell=1}^{\infty} a_{mn}\bar{a}_{k\ell}\lambda_{1m}\lambda_{2n}(\phi_m, \phi_k)(\psi_n, \psi_\ell) = \sum_{m,n=1} |a_{mn}|^2\lambda_{1m}\lambda_{2n} \geq 0.$$

Clearly this is true for all $\Phi \varepsilon \mathcal{H}_1 \otimes \mathcal{H}_2$; hence $D_1 \times D_2$ is non-negative. Furthermore,

$$\text{trace}\,(D_1 \times D_2) = \sum_{m,n=1}^{\infty} ((D_1 \times D_2)(\phi_m \times \psi_n), \phi_m \times \psi_n)$$

$$= \sum_{m,n=1}^{\infty} (D_1\phi_m \times D_2\psi_n, \phi_m \times \psi_n) = \sum_{m,n=1}^{\infty} (D_1\phi_m, \phi_m)(D_2\psi_n, \psi_n) .$$

The last equality follows from the definition of the inner product in $\mathcal{H}_1 \otimes \mathcal{H}_2$. Hence by (3.2.2),

$$\text{trace}\,(D_1 \times D_2) = \sum_{m=1}^{\infty} (D_1\phi_m, \phi_m) \sum_{n=1}^{\infty} (D_2\psi_n, \psi_n) = (\text{trace}\, D_1)(\text{trace}\, D_2) = 1.$$

Therefore $D_1 \times D_2$ is a non-negative, self-adjoint operator with trace $(D_1 \times D_2) = 1$; i.e., a density operator.

Moreover, if p_{12} is the statistical state defined by $D_1 \times D_2$ then the induced statistical state p_1' is defined by

(4.19.1) $$p_1'(M) = p_{12}(M \otimes \mathcal{H}_2) = \text{trace}\,((D_1 \times D_2)P_{M \otimes \mathcal{H}_2})$$

for all closed subspaces $M \subset \mathcal{H}_1$. If we denote the induced density operator by D_1', we have from Theorem 3-5,

(4.19.2) $$p_1'(M) = \text{trace}\,(D_1' P_M)$$

for all M. From (4.19.1) we obtain

$$p_1'(M) = \text{trace}\,((D_1 \times D_2)P_{M \otimes \mathcal{H}_2}) = \text{trace}\,(D_1 P_M \times D_2),$$

where we have used the relations $P_{M \otimes \mathcal{H}_2} = P_M \times P_{\mathcal{H}_2} = P_M \times I_2$. Evaluating trace $(D_1 P_M \times D_2)$ by means of the sequence $\{\phi_m \times \psi_n\}$, we obtain

$$p_1'(M) = \sum_{m,n=1}^{\infty} ((D_1 P_M \times D_2)(\phi_m \times \psi_n), \phi_m \times \psi_n) = \sum_{m,n=1}^{\infty} (D_1 P_M \phi_m \times D_2 \psi_n, \phi_m \times \psi_n)$$

$$= \sum_{m,n=1}^{\infty} (D_1 P_M \phi_m, \phi_m)(D_2 \psi_n, \psi_n) = \sum_{m=1}^{\infty} (D_1 P_M \phi_m, \phi_m) \sum_{n=1}^{\infty} (D_2 \psi_n, \psi_n)$$

$$= \sum_{m=1}^{\infty} (D_1 P_M \phi_m, \phi_m).$$

Therefore, for all $M \subset \mathcal{H}_1$,

(4.19.3) $$p_1'(M) = \text{trace}\,(D_1 P_M).$$

Since the density operator defined by the probability measure p_1' is unique, from (4.19.2) and (4.19.3) we obtain $D_1' = D_1$, or that

$$D_1 = \operatorname{tr}_{\mathcal{H}_2} (D_1 \times D_2).$$

In a similar way one obtains

$$D_2 = \operatorname{tr}_{\mathcal{H}_1} (D_1 \times D_2).$$

Q.E.D.

To show that D_1 and D_2 do not determine D_{12} uniquely, consider an operator of the form

$$D_{12} = t_1(D_1' \times D_2') + t_2 (D_1'' \times D_2'')$$

where D_1', D_2', D_1'', and D_2'' are density operators, and $t_1+t_2=1$. It is easily verified that D_{12} is a density operator and that the induced statistical states define the density operators

$$D_1 = t_1 D_1' + t_2 D_1''$$

$$D_2 = t_1 D_2' + t_2 D_2'' .$$

Then

$$D_1 \times D_2 = t_1^2 (D_1' \times D_2') + t_1 t_2 (D_1' \times D_2'') + t_2 t_1 (D_1'' \times D_2') + t_2^2 (D_1'' \times D_2'')$$

also induces the density operators D_1 on $\mathcal{H}_1$ and D_2 on $\mathcal{H}_2$ and clearly $D_{12} \neq D_1 \times D_2$.

This result is not surprising, for if our knowledge of the composite system consists in knowing only D_1 and D_2, then none of the statistical dependencies which may exist between the component systems are apparent. However, if the component systems are known to be statistically independent, this problem does not arise, for the statistical state of $(\mathcal{H}_{12}, H_{12})$ is then uniquely determined by those of $(\mathcal{H}_1, H_1)$ and $(\mathcal{H}_2, H_2)$. To show this, we must first define what is meant by statistically independent component systems in quantum mechanics.

Definition 4-20. A composite system $(\mathcal{H}_{12}, H_{12})$ in a statistical state p_{12} has statistically independent components $(\mathcal{H}_1, H_1)$ and $(\mathcal{H}_2, H_2)$ if and only if their respective statistical states satisfy the condition

$$p_{12}(M \otimes N) = p_1(M) \cdot p_2(N)$$

for all closed subspaces $M \subset \mathcal{H}_1$ and $N \subset \mathcal{H}_2$.

This definition is clearly motivated by the classical concept of statistically independent events. It applies equally well to our definition of events in quantum mechanics.

Statistical independence can be described alternatively in terms of the density operators. As shown in Theorem 4-22 below, the composite system has statistically independent components if and only if

$$D_{12} = D_1 \times D_2.$$

This is the justification of our assertion above that D_{12} is uniquely determined by D_1 and D_2 when the components are statistically independent.

To establish Theorem 4-22, we shall need an additional lemma on trace class operators.

Lemma 4-21. Let $A \varepsilon (tc)$ be a self-adjoint operator defined for all $\Phi \varepsilon\ \mathcal{H}_1 \otimes \mathcal{H}_2$. If trace $(AP) = 0$ for all projection operators P of form $P_{M \otimes N}$, where M and N are closed subspaces in $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, then $A = 0$.

Proof: Choose any $\phi_1 \varepsilon\ \mathcal{H}_1$ and $\psi_1 \varepsilon\ \mathcal{H}_2$ and extend them to orthonormal bases $\{\phi_m\}$ in $\mathcal{H}_1$ and $\{\psi_n\}$ in $\mathcal{H}_2$. Let $[\phi_1]$ and $[\psi_1]$ denote the one dimensional subspaces spanned by ϕ_1 and ψ_1 respectively. Evaluating

$$\text{trace}\ (A P_{[\phi_1] \otimes [\psi_1]})$$

using the orthonormal basis $\{\phi_m \times \psi_n\}$, we obtain

(4.21.1) $$(A(\phi_1 \times \psi_1),\ \phi_1 \times \psi_1) = 0\ .$$

Since ϕ_1, ψ_1 can be chosen arbitrarily, (4.21.1) holds for all $\phi_1, \psi_1 \varepsilon\ \mathcal{H}_1, \mathcal{H}_2$. We first show that for all $\phi_1 \varepsilon\ \mathcal{H}_1$ and $\psi_1, \psi_2 \varepsilon\ \mathcal{H}_2$

(4.21.2) $$(A(\phi_1 \times \psi_1),\ \phi_1 \times \psi_2) = 0\ .$$

Clearly

$$(A(\phi_1 \times \psi_1), \phi_1 \times \psi_2) = (A(\phi_1 \times \psi_1), \phi_1 \times \psi_1) + (A(\phi_1 \times \psi_1), \phi_1 \times (\psi_2 - \psi_1)).$$

The first term on the right is zero by (4.21.1). The second term may be written

$$(A[\phi_1 \times (\psi_1 - \psi_2)] , \phi_1 \times (\psi_2 - \psi_1)) + (A(\phi_1 \times \psi_2), \phi_1 \times (\psi_2 - \psi_1)),$$

where the first term is again zero by (4.21.1), and the second term is

$$(A(\phi_1 \times \psi_2), \phi_1 \times \psi_2) - (A(\phi_1 \times \psi_2), \phi_1 \times \psi_1) .$$

Since the first term is again zero, we obtain

(4.21.3) $$(A(\phi_1 \times \psi_1), \phi_1 \times \psi_2) = -(A(\phi_1 \times \psi_2), \phi_1 \times \psi_1).$$

Now replace ψ_1 by $i\psi_1$ to get

$$i(A(\phi_1 \times \psi_1), \phi_1 \times \psi_2) = i(A(\phi_1 \times \psi_2), \phi_1 \times \psi_1) ,$$

or

(4.21.4) $$(A(\phi_1 \times \psi_1), \phi_1 \times \psi_2) = (A(\phi_1 \times \psi_2), \phi_1 \times \psi_1)$$

Adding (4.21.3) and (4.21.4) gives (4.21.2):

$$(A(\phi_1 \times \psi_1), \phi_1 \times \psi_2) = 0 .$$

Next, we show that

$$(4.21.5) \qquad (A(\phi_1 \times \psi_1), \phi_2 \times \psi_2) = 0$$

Making repeated use of (4.21.2), we obtain

$$(A(\phi_1 \times \psi_1), \phi_2 \times \psi_2) = (A(\phi_1 \times \psi_1), \phi_1 \times \psi_2) + (A(\phi_1 \times \psi_1), (\phi_2 - \phi_1) \times \psi_2)$$

$$= (A[(\phi_1 - \phi_2) \times \psi_1], (\phi_2 - \phi_1) \times \psi_2) + (A(\phi_2 \times \psi_1), (\phi_2 - \phi_1) \times \psi_2)$$

$$= (A(\phi_2 \times \psi_1), \phi_2 \times \psi_2) - (A(\phi_2 \times \psi_1), \phi_1 \times \psi_2).$$

Therefore, since $(A(\phi_2 \times \psi_1), \phi_2 \times \psi_2) = 0$, we obtain

$$(4.21.6) \qquad (A(\phi_1 \times \psi_1), \phi_2 \times \psi_2) = -(A(\phi_2 \times \psi_1), \phi_1 \times \psi_2).$$

Now replace ϕ_1 by $i\phi_1$ to get

$$i(A(\phi_1 \times \psi_1, \phi_2 \times \psi_2) = i(A(\phi_2 \times \psi_1), \phi_1 \times \psi_2),$$

or

$$(4.21.7) \qquad (A(\phi_1 \times \psi_1, \phi_2 \times \psi_2) = (A(\phi_2 \times \psi_1), \phi_1 \times \psi_2)$$

Adding (4.21.6) and (4.21.7) gives (4.21.5):

$$(A(\phi_1 \times \psi_1), \phi_2 \times \psi_2) = 0 .$$

Now it follows that $(A\Phi, \Phi) = 0$ for all $\Phi \varepsilon (\mathcal{H}_1 \otimes \mathcal{H}_2)'$, that is,for all

$$\Phi = \sum_{n=1}^{p} (\xi_n \times \eta_n)$$

with finite p. By the continuity of the inner product, we have $(A\Phi, \Phi) = 0$ for all $\Phi \varepsilon\ \mathcal{H}_1 \otimes \mathcal{H}_2$. It is well known [8, p. 38] that this implies $A = 0$.

Q. E. D.

Theorem 4-22. Let D_1, D_2, and D_{12} be density operators on the separable Hilbert spaces $\mathcal{H}_1$, $\mathcal{H}_2$, and $\mathcal{H}_1 \otimes \mathcal{H}_2$ respectively, and let p_1, p_2, and p_{12} be the corresponding probability measures. Then $D_{12} = D_1 \times D_2$ if and only if $p_{12}(M \otimes N) = p_1(M) \cdot p_2(N)$ for all closed subspaces $M \subset \mathcal{H}_1$ and $N \subset \mathcal{H}_2$.

Proof: Suppose $D_{12} = D_1 \times D_2$. Then, by Theorem 4-12, for all subspaces $M \subset \mathcal{H}_1$ and $N \subset \mathcal{H}_2$, we have

$$p_{12}(M \otimes N) = \text{trace } (D_{12} P_{M \otimes N}) = \text{trace } [(D_1 \times D_2)(P_M \times P_N)] .$$

Choose any orthonormal basis $\{\phi_m \times \psi_n\}$ in $\mathcal{H}_1 \otimes \mathcal{H}_2$, where $\{\phi_m\}$ and $\{\psi_n\}$ are orthonormal bases in $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively. Then

$$\text{trace}\,[(D_1 \times D_2)(P_M \times P_N)] = \sum_{m,n=1}^{\infty} ((D_1 \times D_2)(P_M \times P_N)(\phi_m \times \psi_n), \phi_m \times \psi_n)$$

$$= \sum_{m,n=1}^{\infty} (D_1 P_M \phi_m \times D_2 P_N \psi_n, \phi_m \times \psi_n)$$

$$= \sum_{m,n=1}^{\infty} (D_1 P_M \phi_m, \phi_m)(D_2 P_N \psi_n, \psi_n)\ .$$

The last equality is obtained using (4. 2. 3) and gives

$$p_{12}(M \otimes N) = \sum_{m=1}^{\infty} (D_1 P_M \phi_m, \phi_m) \sum_{n=1}^{\infty} (D_2 P_N \psi_n, \psi_n)$$

by virtue of (3. 2. 2). Therefore

$$p_{12}(M \otimes N) = \text{trace}\,(D_1 P_M) \cdot \text{trace}(D_2 P_N) = p_1(M) \cdot p_2(N).$$

Now suppose $p_{12}(M \otimes N) = p_1(M) \cdot p_2(N)$ for all closed subspaces $M \subset \mathcal{H}_1$ and $N \subset \mathcal{H}_2$. Then

$$p_{12}(M \otimes N) = \text{trace}(D_1 P_M) \cdot \text{trace}(D_2 P_N) = \sum_{m=1}^{\infty} (D_1 P_M \phi_m, \phi_m) \sum_{n=1}^{\infty} (D_2 P_N \psi_n, \psi_n).$$

Using (4. 2. 3) again, this becomes

$$\sum_{m,n=1}^{\infty} (D_1 P_M \phi_m \times D_2 P_N \psi_n, \phi_m \times \psi_n) = \sum_{m,n=1}^{\infty} ((D_1 \times D_2)(P_M \times P_N)(\phi_m \times \psi_n), \phi_m \times \psi_n)$$

$$= \text{trace}\,((D_1 \times D_2)(P_M \times P_N))$$

$$= \text{trace}\,((D_1 \times D_2) P_{M \otimes N})\ .$$

Hence for all M, N,

$$p_{12}(M \otimes N) = \text{trace } ((D_1 \times D_2) P_{M \otimes N}) \, .$$

Since p_{12} corresponds to D_{12}, we also have

$$p_{12}(M \otimes N) = \text{trace } (D_{12} P_{M \otimes N})$$

for all M, N. Therefore

$$0 = \text{trace } ((D_1 \times D_2) P_{M \otimes N}) - \text{trace } (D_{12} P_{M \otimes N}),$$

or

$$\text{trace } ([(D_1 \times D_2) - D_{12}] \, P_{M \otimes N}) = 0$$

for all projection operators of the form $P_{M \otimes N}$. Since $D_1 \times D_2$ and D_{12} are both bounded self-adjoint operators of trace class, so is $(D_1 \times D_2) - D_{12}$, and it follows from Lemma 4-21 that

$$D_{12} = D_1 \times D_2 .$$

Q. E. D.

We conclude this chapter with a brief discussion of some analogies between classical and quantum statistics. Perhaps it is best to begin by stating again the main difference, namely, that in classical statistics the set of all events forms a complemented distributive lattice whereas the complemented lattice of events in quantum statistics need not satisfy the distributive law. However, by altering the definition of a probability measure, we are able to define a statistical state in both cases as a probability measure on the set of all events. Certain analogies then become apparent. Whereas in classical statistics a system in a statistical state p is a probability space $(\Gamma, \mathcal{L}, p)$, in quantum statistics a system in a statistical state p (according to Definition 3-1) is a triple $(\mathcal{H}, \mathfrak{M}, p)$, where $\mathfrak{M}$ is the set of all closed subspaces of the Hilbert space $\mathcal{H}$. For brevity let us call $(\mathcal{H}, \mathfrak{M}, p)$ a "Q-space". If $(\mathcal{H}_1, \mathfrak{M}_1, p_1)$ and $(\mathcal{H}_2, \mathfrak{M}_2, p_2)$ are two Q-spaces, we may denote by $\mathfrak{M}_1 \otimes \mathfrak{M}_2$ the set of all closed subspaces of $\mathcal{H}_1 \otimes \mathcal{H}_2$. We may then rephrase Theorem 4-22 in a form analagous to the theorem on product measures in classical probability theory:

Theorem 4-23. Let $(\mathcal{H}_1, \mathfrak{M}_1, p_1)$ and $(\mathcal{H}_2, \mathfrak{M}_2, p_2)$ be two Q-spaces. Then there exists one and only one probability measure on $\mathfrak{M}_1 \otimes \mathfrak{M}_2$, denoted by $p_1 \otimes p_2$, such that

$$(p_1 \otimes p_2)(M_1 \otimes M_2) = p_1(M_1) \cdot p_2(M_2)$$

for all closed subspaces $M_1 \otimes M_2$.

It is then trivial to define the tensor product of two Q-spaces:

$$(\mathcal{H}_{12}, \mathfrak{M}_{12}, p_{12}) = \prod_{i=1}^{2} \otimes (\mathcal{H}_i, \mathfrak{M}_i, p_i) = (\mathcal{H}_1 \otimes \mathcal{H}_2, \mathfrak{M}_1 \otimes \mathfrak{M}_2, p_1 \otimes p_2).$$

Therefore in quantum mechanics a composite system composed of two statistically independent components is the Q-space

$$(\mathcal{H}_1 \otimes \mathcal{H}_2, \ \mathfrak{M}_1 \otimes \mathfrak{M}_2, \ p_1 \otimes p_2) .$$

The analogous structure in classical mechanics is of course the direct product

$$(\Gamma_1 \times \Gamma_2, \ \mathcal{L}_1 \times \mathcal{L}_2, \ p_1 \times p_2).$$

In both cases the extension of the theory of composite systems to the case of an arbitrary but finite number of statistically independent components is now straightforward.

We summarize the analogies between classical and quantum statistics in the following table:

TABLE

	Classical Mechanics	Quantum Mechanics
Phase space	differentiable measure space Γ	Hilbert space $\mathcal{H}$
Mechanical state	point in Γ	unit vector in $\mathcal{H}$
Event	measurable subset in Γ	closed linear subspace in $\mathcal{H}$
Statistical state	probability measure on $\mathcal{L}$	probability measure on $\mathcal{M}$
Observable	real valued measurable function on Γ	self-adjoint operator on $\mathcal{H}$
Probability density	probability density $D(z)$ on Γ	density operator D on $\mathcal{H}$
Composite phase space	direct product $\Gamma_1 \times \Gamma_2$	tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2$
Composite statistical state	probability measure p_{12} on $\mathcal{L}_1 \times \mathcal{L}_2$	probability measure p_{12} on $\mathcal{M}_1 \otimes \mathcal{M}_2$
Composite probability density	probability density $D_{12}(z_1, z_2)$ on $\Gamma_1 \times \Gamma_2$	density operator D_{12} on $\mathcal{H}_1 \otimes \mathcal{H}_2$
Induced states	$p_1(A) = p_{12}(A \times \Gamma_2)$ $p_2(B) = p_{12}(\Gamma_1 \times B)$	$p_1(M) = p_{12}(M \otimes \mathcal{H}_2)$ $p_2(N) = p_{12}(\mathcal{H}_1 \otimes N)$
Induced densities	$D_1(z_1) = \int_{\Gamma_2} D_{12}(z_1, z_2)\mu(dz_2)$ $D_2(z_2) = \int_{\Gamma_1} D_{12}(z_1, z_2)\mu(dz_1)$	$D_1 = \operatorname{tr}_{\mathcal{H}_2} D_{12}$ $D_2 = \operatorname{tr}_{\mathcal{H}_1} D_{12}$
Statistical independence	$p_{12}(A \times B) = p_1(A) \cdot p_2(B)$ $D_{12}(z_1, z_2) = D_1(z_1) \cdot D_2(z_2)$	$p_{12}(M \otimes N) = p_1(M) \cdot p_2(N)$ $D_{12} = D_1 \times D_2$

CHAPTER V. THE MAIN THEOREM

Our purpose in this chapter is to present a characterization of the canonical equilibrium states for quantum mechanical systems. We shall work with a collection Q of quantum systems, such that each system $(\mathcal{H}, H)$ of Q is in an equilibrium statistical state p. The set Q, the Hamiltonian operators, and the statistical states are assumed to have certain properties listed below as axioms. These assumptions imply that each system is in a canonical state at a temperature which is the same for all systems. The possibility of zero absolute temperature is included (see Chapter III).

We shall view Q as a set in the mathematical sense. However, Q may be interpreted physically as a large collection of systems, each pair of which can be brought together into equilibrium at a common temperature. We do not treat the mechanism for attaining equilibrium; this presumably would require some mechanical interaction between the members of each pair. We deal instead with the limiting case of zero interaction, and postulate that each system is in a limiting state p, independent of the second system of the composite pair. It is not necessary to make an explicit assumption regarding a common temperature for all systems. The temperature simply appears as a free parameter in the class of canonical states.

The axioms fall into three separate categories. The first axiom restricts the collection Q to systems whose Hamiltonian operators have pure point spectra, and the next three axioms ensure that Q is a "sufficiently large" collection for our characterization to succeed. The last two axioms place natural restrictions on the equilibrium states.

Axiom 1. For each system in Q, the Hamiltonian operator H has a pure point spectrum S(H) consisting of zero and a sequence of real numbers increasing to infinity:

$$0 = \lambda_0 < \lambda_1 < \cdots < \lambda_n < \cdots,$$

where

$$\lim_{n \to \infty} \lambda_n = \infty .$$

This axiom is usually interpreted physically as restricting each system to a finite enclosure, or at least to an infinitely deep potential well, as in the case of a harmonic oscillator. The requirement that the smallest characteristic value λ_1 is zero is imposed as a matter of convenience; it merely requires that each Hamiltonian be "normalized" by a constant energy shift (see Chapter III).

Axiom 2. Q is closed under the pairwise formation of noninteracting composite systems. That is, if $(\mathcal{H}_1, H_1)$ and $(\mathcal{H}_2, H_2)$ are any two members of Q, then the noninteracting composite system

$(\mathcal{H}_1 \otimes \mathcal{H}_2, \quad H_1 \times I_2 + I_1 \times H_2)$ also belongs to Q.

This axiom provides the mathematical counterpart for the physical assertion that each pair of systems can be brought into equilibrium with each other in the limit of zero interaction. Notice that Axioms 1 and 2 are consistent in the sense that when two systems are normalized so that the lowest characteristic values of their Hamiltonians are zero, then their noninteracting composite system is also normalized in the same way (see Theorem 4-17). It also follows from Axiom 2 and Theorem 4-17 that the set $\mathcal{D}$ consisting of all Hamiltonian spectral values for systems in Q is closed under addition. The next axiom ensures that $\mathcal{D}$ is also closed under positive differences:

<u>Axiom 3.</u> Let $\mathcal{D}$ be the union of the Hamiltonian spectra of all systems in Q. Then $\mathcal{D}$ is closed under subtraction, in the sense that if $\lambda_1, \lambda_2 \varepsilon \mathcal{D}$ and $\lambda_2 \geq \lambda_1$, then $\lambda_2 - \lambda_1 \varepsilon \mathcal{D}$.

<u>Axiom 4.</u> For each $\lambda \varepsilon \mathcal{D}$, $\lambda \neq 0$, there is a system $(\mathcal{H}, H)$ in Q such that S(H) consists of zero and all positive integral multiples of λ; that is,

$$S(H) = \{n\lambda : n = 0, 1, 2, \cdots\}.$$

Harmonic oscillators and composite systems of two or more identical harmonic oscillators have spectra of the form required by Axiom 4.

For the sake of convenience, we shall call the systems guaranteed by Axiom 4 "harmonic oscillators", bearing in mind that true harmonic oscillators are not specifically required.

Axiom 5. Let $(\mathcal{H}, H)$ be a system in Q, p be the corresponding equilibrium state, and D the density operator for this state. Then there is a function $f(\lambda)$ defined on $S(H)$ such that

$$D = f(H).$$

The significance of this standard restriction is discussed in Chapters II and III. It ensures, among other things, that each state p qualifies as a true equilibrium state.

Our final axiom presents the main statistical assumption, which, rather loosely interpreted, asserts that the mechanical independence of any pair of systems in Q implies their statistical independence also:

Axiom 6. For a composite system $(\mathcal{H}_{12}, H_{12})$ composed of two noninteracting components $(\mathcal{H}_1, H_1)$ and $(\mathcal{H}_2, H_2)$ in Q, the component systems are statistically independent; that is, their statistical states satisfy

$$p_{12}(M \otimes N) = p_1(M) \cdot p_2(N)$$

for all closed linear subspaces $M \subset \mathcal{H}_1$ and $N \subset \mathcal{H}_2$.

It must be stressed that the axioms do not determine the set Q uniquely. For example, Q might consist only of those systems which can be built up by composition from a single harmonic oscillator, in which case all systems in Q would have the same spectrum: $\mathcal{B} = \{0, \lambda, 2\lambda, \cdots\}$. Therefore it is important to know that any collection Q_0 of systems which we may wish to consider can be enlarged to obtain a collection Q satisfying Axioms 1 through 4, provided only that Q_0 satisfies Axiom 1:

<u>Theorem 5-1.</u> Let Q_0 be any collection of systems satisfying Axiom 1. Then there is a second collection Q which contains Q_0 and satisfies Axioms 1 through 4.

<u>Proof:</u> Let $\mathcal{B}_0$ be the union of all Hamiltonian spectra of systems in Q_0, and let $\mathcal{B}_1$ be the smallest set containing $\mathcal{B}_0$ such that $\mathcal{B}_1$ is closed under addition, subtraction, and multiplication by integers. Then

$$\mathcal{B}_1 = \{n_1\lambda_1 + \cdots + n_k\lambda_k : k = 1, 2, \cdots ; n_j = 0, \pm 1, \pm 2, \cdots \text{ for all } j = 1, 2, \cdots k;\ \lambda_j \varepsilon \mathcal{B}_0\}$$

Let $\mathcal{B}_+ = \{\lambda \varepsilon \mathcal{B}_1 : \lambda \geq 0\}$. Then $\mathcal{B}_+$ is closed under addition, positive differences, and multiplication by positive integers. To show this, let $\lambda_1, \lambda_2 \varepsilon \mathcal{B}_+$. Then $\lambda_1, \lambda_2 \varepsilon \mathcal{B}_1$ and $\lambda_1 + \lambda_2 \varepsilon \mathcal{B}_1$. Since $\lambda_1 + \lambda_2 > 0$, $\lambda_1 + \lambda_2 \varepsilon \mathcal{B}_+$. Moreover, if $\lambda_2 \geq \lambda_1$, then $\lambda_2 - \lambda_1 \varepsilon \mathcal{B}_1$, and since $\lambda_2-\lambda_1 \geq 0$, $\lambda_2-\lambda_1 \epsilon \mathcal{D}_+$. Since $\mathcal{D}_1$ is closed under multiplication by positive integers, so is $\mathcal{D}_+$.

Now for each non-zero $\lambda \epsilon \mathcal{D}_+$ adjoin to Q_0 the harmonic oscillator whose Hamiltonian spectrum is $S_\lambda = \{n\lambda : n=0, 1, 2, \cdots\}$ and denote the enlarged collection by Q_1. Clearly for each $\lambda \epsilon \mathcal{D}_+$, $S_\lambda \subset \mathcal{D}_+$, so that the union of all Hamiltonian spectra of systems in Q_1 is just $\mathcal{D}_+$. Moreover, if Q_1' is the collection of systems obtained by noninteracting pairwise composition of members of Q_1, the union of all Hamiltonian spectra of the collection

$$Q_2 = Q_1 \cup Q_1'$$

is also $\mathcal{D}_+$. Now for each integer $m \geq 1$, let Q_m' be the collection of systems obtained by noninteracting pairwise composition of members of Q_m, and define

$$Q_{m+1} = Q_m \cup Q_m' .$$

Finally, let

$$Q = \bigcup_{m=1}^{\infty} Q_m .$$

We shall show that Q satisfies Axioms 1 through 4.

Clearly Axiom 1 is satisfied, by virtue of Theorem 4-17. Now let $(\mathcal{H}_1, H_1)$ and $(\mathcal{H}_2, H_2)$ be any two systems in Q.

Then for some $m_1, m_2 = 1, 2, \cdots$, $(\mathcal{H}_1, H_1) \varepsilon Q_{m_1}$ and $(\mathcal{H}_2, H_2) \varepsilon Q_{m_2}$. Suppose $m_2 \geq m_1$, then $(\mathcal{H}_1, H_1)$ and $(\mathcal{H}_2, H_2)$ together with their composition $(\mathcal{H}_{12}, H_{12})$ are all in $Q_{m_{2+1}}$, hence in Q. Consequently Q satisfies Axiom 2.

Let $\mathcal{D}$ be the union of all Hamiltonian spectra of members in Q. Then $\mathcal{D} = \mathcal{D}_+$, and since $\mathcal{D}_+$ is closed under positive differences, Q satisfies Axiom 3.

Finally, Axiom 4 is satisfied, for if $\lambda \varepsilon \mathcal{D}$, then $\lambda \varepsilon \mathcal{D}_+$. Hence the harmonic oscillator whose spectrum is $\{n\lambda : n=0, 1, 2 \cdots\}$ is in $Q_1 \subset Q$.

Q. E. D.

Turning now to the consequences of our axioms, we have

<u>Lemma 5-2.</u> Let $D = f(H)$ be the density operator for the equilibrium state p of a system in Q. Let $S(H) = \{\lambda_k\}$ be the spectrum of H, where the characteristic values λ_k have multiplicities m_k respectively. Then the function f satisfies the conditions

(5. 2. 1) $$f(\lambda_k) \geq 0 \quad \text{for all} \quad k = 1, 2, \cdots$$

(5. 2. 2) $$\sum_{k=1}^{\infty} m_k f(\lambda_k) = 1 .$$

Proof: Let P_k be the projection operator onto the characteristic subspace M_k (of dimension m_k) corresponding to the characteristic value λ_k of H. Evaluating

$$p(M_k) = \text{trace}\ (D P_k)$$

by means of any complete orthonormal sequence of characteristic vectors of H, we obtain

$$p(M_k) = m_k f(\lambda_k) .$$

Since $p(M_k)$ is necessarily positive, it follows that $f(\lambda_k)$ is positive. Evaluating $p(\mathcal{H}) = \text{trace } D = 1$ by the same orthonormal sequence, we obtain (5.2.2).

Q.E.D.

Lemma 5-3. Let $(\mathcal{H}_1, H_1)$, $(\mathcal{H}_2, H_2)$ be two systems in Q, and $(\mathcal{H}_{12}, H_{12})$ be their noninteracting composition. If $D_1 = f_1(H_1)$, $D_2 = f_2(H_2)$, and $D_{12} = f_{12}(H_{12})$ are the corresponding equilibrium density operators, then the domain of f_{12} is

(5.3.1) $$S(H_{12}) = \{\lambda_1 + \lambda_2 : \lambda_1 \varepsilon S(H_1), \lambda_2 \varepsilon S(H_2)\}$$

and the identity

(5.3.2) $$f_{12}(\lambda_1 + \lambda_2) = f_1(\lambda_1) f_2(\lambda_2)$$

holds for all $\lambda_1 \varepsilon S(H_1)$ and $\lambda_2 \varepsilon S(H_2)$.

Proof: Since the component systems are noninteracting, it follows from Axiom 6 and Theorem (4-22) that

$$D_{12} = D_1 \times D_2 ,$$

or

(5.3.3) $$f_{12}(H_{12}) = f_1(H_1) \times f_2(H_2)$$

everywhere in $\mathcal{H}_1 \otimes \mathcal{H}_2$.

From Definition 4-18, $H_{12} = H_1 \times I_2 + I_1 \times H_2$, and by Theorem (4-17), $H_1 \times I_2 + I_1 \times H_2$ is essentially self-adjoint and has the pure point spectrum (5.3.1). Therefore, by Axiom 5, f_{12} has the domain (5.3.1).

If we identify $H_{12} = H_1 \times I_2 + I_1 \times H_2$ with its smallest closed extension, then from (3.6.2a) it has the spectral representation

$$H_{12} = \sum_{m=1}^{\infty} \mu_m P_m ,$$

where P_m is the projection operator onto the characteristic subspace corresponding to the characteristic value μ_m of H_{12}.

The operator $f_{12}(H_{12})$ is uniquely defined by (3.6.3a) to be

(5.3.4) $$f_{12}(H_{12}) = \sum_{m=1}^{\infty} f_{12}(\mu_m) P_m ,$$

and from (5. 3. 2),

$$f_1(H_1) \times f_2(H_2) = \sum_{m=1}^{\infty} f_{12}(\mu_m) P_m \, .$$

Now let λ_1, λ_2 be any characteristic values of H_1, H_2 respectively, and let ϕ, ψ be any corresponding unit characteristic vectors. Then when $\mu_n = \lambda_1 + \lambda_2$, we have

$$P_m(\phi \times \psi) = \delta_{mn}(\phi \times \psi),$$

and we obtain

$$(f_1(H_1) \times f_2(H_2))\,(\phi \times \psi) = \sum_{m=1}^{\infty} f_{12}(\mu_m) P_m(\phi \times \psi),$$

or

$$[f_1(H_1)\phi] \times [f_2(H_2)\psi] = \sum_{m=1}^{\infty} f_{12}(\mu_m)\,\delta_{mn}(\phi \times \psi).$$

Hence

$$f_1(\lambda_1)\phi \times f_2(\lambda_2)\psi = f_{12}(\mu_n)(\phi \times \psi),$$

which gives

$$f_1(\lambda_1) f_2(\lambda_2)(\phi \times \psi) = f_{12}(\lambda_1 + \lambda_2)(\phi \times \psi).$$

Therefore

$$f_1(\lambda_1) f_2(\lambda_2) = f_{12}(\lambda_1 + \lambda_2).$$

Q. E. D.

As a consequence of Lemma 5-3, every pair of systems in Q yields a functional equation of the form (5.3.2) where the functions f depend on the systems involved. By considering each pair $(\mathcal{H}_m, H_m)$, $(\mathcal{H}_n, H_n)$ of systems in Q, one obtains a system of functional equations

$$f_{mn}(\lambda_m + \lambda_n) = f_m(\lambda_m) f_n(\lambda_n) \tag{5.3.5}$$

which holds for all $\lambda_m \varepsilon S(H_m)$ and all $\lambda_n \varepsilon S(H_n)$. The subscripts m, n index the systems in Q and range over a possibly non-denumerable set. Our object is to determine the equilbirium state of each member of Q by solving this system to obtain the equilibrium density operators. The solution is not as difficult as it may at first appear. In fact, the great profusion of systems in Q actually simplifies our task.

Lemma 5-4. Let $(\mathcal{H}, H)$ be any system in Q. If $D = f(H)$ is its equilibrium density operator, then $f(0) \neq 0$.

Proof: We first consider the harmonic oscillators of Axiom 4. Let $(\mathcal{H}, H)$ be a harmonic oscillator with $S(H) = \{n\lambda : n=0,1,2,\cdots\}$. Form the composite system $(\mathcal{H}_{12}, H_{12})$ consisting of the identical components $(\mathcal{H}, H)$ and $(\mathcal{H}, H)$. Then, by Lemma 5-3, we obtain the functional equation

$$f_{12}(n\lambda + m\lambda) = f(n\lambda)f(m\lambda)$$

for all $m, n = 0, 1, 2 \cdots$. Now suppose $f(0) = 0$, then for all $m = 0, 1, 2 \cdots$, we have

$$f_{12}(m\lambda) = f(0)f(m\lambda) = 0 .$$

Hence f_{12} is zero everywhere on $S(H_{12}) = \{n\lambda + m\lambda : m, n = 0, 1, 2, \cdots\} = \{n\lambda : n = 0, 1, 2, \cdots\}$. This is a contradiction since we must have

$$\text{trace } D_{12} = \sum_{n=1}^{\infty} m_n f_{12}(n\lambda) = 1 .$$

Therefore, if $(\mathcal{H}, H)$ is any harmonic oscillator, $f(0) \neq 0$.

Now let $(\mathcal{H}, H)$ be an arbitrary system in Q, with $D = f(H)$, and suppose that $f(0) = 0$. Then for some $\lambda' \varepsilon S(H), \lambda' > 0$, we have $f(\lambda') \neq 0$. If $(\mathcal{H}', H')$ is the harmonic oscillator having $S(H') = \{n\lambda' : n = 0, 1, 2, \cdots\}$ and $D' = f'(H')$, then by forming the composite system $(\mathcal{H}'_{12}, H'_{12})$ composed of $(\mathcal{H}, H)$ and $(\mathcal{H}', H')$, we obtain

$$f'_{12}(\lambda') = f'(0)f(\lambda') = f'(\lambda')f(0) .$$

Here $f'(\lambda')f(0)$ is zero, but $f'(0)f(\lambda')$ is not, which again is a contradiction.

Q. E. D.

Lemma 5-4 permits an important simplification. It allows the system (5.3.5) to be replaced by a single functional equation. This is contained in the next lemma.

Lemma 5-5. Consider the set $\mathcal{D}$ which is the union of the Hamiltonian spectra for all systems in Q. There is a function F defined on $\mathcal{D}$, satisfying the functional equation

$$F(\lambda_1+\lambda_2) = F(\lambda_1)F(\lambda_2) \tag{5.5.1}$$

for all $\lambda_1, \lambda_2 \varepsilon \mathcal{D}$ and the initial condition

$$F(0) = 1$$

such that for each system $(\mathcal{H}, H)$ in Q, with equilibrium density operator $D = f(H)$, the function f satisfies

$$f(\lambda) = f(0)F(\lambda) \tag{5.5.2}$$

for all $\lambda \varepsilon S(H)$.

Proof: Choosing any system in Q we may, in view of Lemma 5-4, define a function F on S(H) by the equation

$$F(\lambda) = \frac{f(\lambda)}{f(0)} .$$

Indeed, this serves to define F on all of $\mathcal{D}$ provided we obtain the same ratio $f(\lambda)/f(0)$ for any two systems whose Hamiltonian spectra share a common value λ. That this is the case follows by first setting $\lambda_1 = \lambda$, $\lambda_2 = 0$ and then $\lambda_1 = 0$, $\lambda_2 = \lambda$ in (5.3.2) to obtain

$$f_{12}(\lambda) = f_1(\lambda)f_2(0) = f_1(0)f_2(\lambda),$$

whence

$$\frac{f_1(\lambda)}{f_1(0)} = \frac{f_2(\lambda)}{f_2(0)} .$$

Taking $\lambda = 0$, we see at once that $F(0) = 1$. To derive the functional equation (5.5.1), substitute (5.5.2) into (5.3.2) and obtain

$$f_{12}(0)F(\lambda_1+\lambda_2) = f_1(0)f_2(0)F(\lambda_1)F(\lambda_2).$$

But with $\lambda_1 = \lambda_2 = 0$, we have

$$f_{12}(0) = f_1(0)f_2(0) \neq 0$$

and the result follows.

Q. E. D.

<u>Lemma 5-6.</u> Let $F(\lambda)$ be the function defined in Lemma 5-5. Then $0 \leq F(\lambda) < 1$ for all $\lambda \varepsilon \mathcal{D}$, $\lambda > 0$.

Proof: Let $\lambda \varepsilon \mathfrak{D}$ and $\lambda \neq 0$. Consider the harmonic oscillator $(\mathcal{H}, H)$ having $S(H) = \{n\lambda : n=0,1,2 \cdots\}$. Then from (5.5.1),

$$F([n+1]\lambda) = F(n\lambda)F(\lambda) ,$$

and it follows by induction that

$$F(n\lambda) = [F(\lambda)]^n . \qquad (5.6.1)$$

Hence, if $D = f(H)$, the characteristic values of D are given by

$$f(n\lambda) = f(0)\, F(n\lambda) = f(0)[F(\lambda)]^n$$

for all $n = 0, 1, 2 \cdots$. In order for

$$\text{trace } D = \sum_{n=0} m_n f(n\lambda) = f(0) \sum_{n=0} m_n [f(\lambda)]^n \geq f(0) \sum_{n=0} [F(\lambda)]^n$$

to converge, we must have $0 \leq F(\lambda) < 1$.

Q. E. D.

Lemma 5-7. If $F(\lambda) = 0$ for some $\lambda > 0$, then $F(\lambda) = 0$ for all $\lambda > 0$, and each system in Q is in its canonical state of zero absolute temperature.

Proof: The proof depends on the fact that $\mathfrak{D}$ is closed under subtraction. Let $F(\lambda_1) = 0$, and let λ_2 be any positive number in $\mathfrak{D}$. There are two cases:

Case I. If $\lambda_2 > \lambda_1$, then by Axiom 3, $\lambda_2 - \lambda_1 \varepsilon \mathfrak{D}$, and from (5. 5. 1) we obtain

$$F(\lambda_2) = F(\lambda_1)F(\lambda_2 - \lambda_1).$$

Since $F(\lambda_1) = 0$, we have $F(\lambda_2) = 0$.

Case II. If $\lambda_2 < \lambda_1$, then for some positive integer n we obtain $n\lambda_2 > \lambda_1$, and by case I, $F(n\lambda_2) = 0$. It follows from (5. 6. 1) that $F(\lambda_2) = 0$.

Therefore $F(\lambda) = 0$ for all $\lambda > 0$. Now if $(\mathcal{H}, H)$ is any system in Q, with equilibrium density operator $D = f(H)$, then

$$\text{trace } D = m_0 f(0) = 1,$$

where m_0 is the multiplicity of the characteristic subspace corresponding to $\lambda = 0$. From (5. 5. 2), we obtain

$$f(\lambda) = \begin{cases} \frac{1}{m_0} & \text{if } \lambda = \lambda_0 = 0 \\ 0 & \text{if } \lambda \neq 0 \end{cases}$$

This is the canonical state of zero absolute temperature.

Q. E. D.

In view of Lemma 5-6, we are able to distinguish two cases. As we have just shown, the case where $F(\lambda) = 0$ for all positive $\lambda \varepsilon \mathfrak{D}$ defines the situation where each system in Q is in a canonical state of zero absolute temperature. We shall show that the alternative case, that is, where $F(\lambda) \neq 0$ for all $\lambda \varepsilon \mathfrak{D}$, leads to a canonical state for each system in Q corresponding to a uniform absolute temperature $T > 0$. Here we have $F(0) = 1$, and $0 < F(\lambda) < 1$ for all positive $\lambda \varepsilon \mathfrak{D}$. It is convenient to define the function

$$y(\lambda) = \ell n\, F(\lambda),$$

so that $y(\lambda)$ satisfies

$$(5.7.1) \qquad y(\lambda_1+\lambda_2) = y(\lambda_1) + y(\lambda_2)$$

for all $\lambda_1, \lambda_2 \varepsilon \mathfrak{D}$. By defining

$$y(-\lambda) = -y(\lambda),$$

we may extend the definition of $y(\lambda)$ to negative numbers, that is, $y(\lambda)$ is now defined on the extended domain $\mathfrak{D} \cup -\mathfrak{D}$, where $-\mathfrak{D} = \{-\lambda : \lambda \varepsilon \mathfrak{D}\}$. We will show that every such extended solution of (5.7.1) satisfies

$$y(\lambda_1+\lambda_2) = y(\lambda_1) + y(\lambda_2)$$

for all λ_1, λ_2 in the extended domain $\mathcal{D} \cup -\mathcal{D}$. The following lemma is needed.

Lemma 5-8. $\mathcal{D} \cup -\mathcal{D}$ is closed under addition and under multiplication by integers.

Proof: Clearly $\mathcal{D}$ and $-\mathcal{D}$ are each closed under addition. Let $-\lambda_2 \varepsilon -\mathcal{D}$ and $\lambda_1 \varepsilon \mathcal{D}$. Then if $\lambda_1+(-\lambda_2) = 0$, we have $\lambda_1+(-\lambda_2) \varepsilon \mathcal{D} \cup -\mathcal{D}$. If $\lambda_1+(-\lambda_2) > 0$, then $\lambda_1+(-\lambda_2) \varepsilon \mathcal{D}$ since $\mathcal{D}$ is closed under subtraction. If $\lambda_1+(-\lambda_2) < 0$, then $\lambda_2+(-\lambda_1)$ is in $\mathcal{D}$. Hence $-(\lambda_2+(-\lambda_1)) = \lambda_1+(-\lambda_2) \varepsilon -\mathcal{D}$.

Since $\mathcal{D}$ and $-\mathcal{D}$ are each closed under multiplication by positive integers, then so is $\mathcal{D} \cup -\mathcal{D}$. Hence for any positive integer m and any $\lambda \varepsilon \mathcal{D} \cup -\mathcal{D}$, $m\lambda \varepsilon \mathcal{D} \cup -\mathcal{D}$. If $m\lambda \varepsilon \mathcal{D}$ then $(-m)\lambda \varepsilon -\mathcal{D}$, and if $m\lambda \varepsilon -\mathcal{D}$, then $(-m)\lambda \varepsilon \mathcal{D}$. Therefore $\mathcal{D} \cup -\mathcal{D}$ is closed under multiplication by negative integers also.

Q.E.D.

Lemma 5-9. Let $y(\lambda)$ be any function, defined on $\mathcal{D} \cup -\mathcal{D}$, which satisfies the identities

(5.9.1) $$y(\lambda_1+\lambda_2) = y(\lambda_1) + y(\lambda_2)$$

for all $\lambda_1, \lambda_2 \varepsilon \mathcal{D}$, and

(5.9.2) $$y(-\lambda) = -y(\lambda).$$

Then (5.9.1) holds for all $\lambda_1, \lambda_2 \varepsilon \mathcal{D} \cup -\mathcal{D}$. Moreover, if m is any integer, then

$$y(m\lambda) = my(\lambda)$$

for all $\lambda \varepsilon \mathcal{D} \cup -\mathcal{D}$.

<u>Proof:</u> Let $-\lambda_1, -\lambda_2 \varepsilon -\mathcal{D}$, then by Lemma 5-8, $y(-\lambda_1-\lambda_2)$ is defined, and

$$y(-\lambda_1-\lambda_2) = y(-(\lambda_1+\lambda_2)) = -y(\lambda_1+\lambda_2) = -(y(\lambda_1) + y(\lambda_2)) = -y(\lambda_1)-y(\lambda_2)$$

$$= y(-\lambda_1) + y(-\lambda_2).$$

Hence (5.9.1) holds for all $\lambda_1, \lambda_2 \varepsilon -\mathcal{D}$. Now let $\lambda_1 \varepsilon \mathcal{D}$ and $-\lambda_2 \varepsilon -\mathcal{D}$. Then $y(\lambda_1-\lambda_2)$ is defined, and if $\lambda_1-\lambda_2 = 0$,

$$y(\lambda_1-\lambda_2) = y(0) = 0 = y(\lambda_1) - y(\lambda_2) = y(\lambda_1) + y(-\lambda_2).$$

If $\lambda_1-\lambda_2 > 0$, then $\lambda_1-\lambda_2 \varepsilon \mathcal{D}$, and

$$y(\lambda_1-\lambda_2) + y(\lambda_2) = y(\lambda_1-\lambda_2+\lambda_2) = y(\lambda_1).$$

Hence

$$y(\lambda_1-\lambda_2) = y(\lambda_1) - y(\lambda_2) = y(\lambda_1) + y(-\lambda_2).$$

If $\lambda_1-\lambda_2 < 0$, then $\lambda_2-\lambda_1 > 0$ and $\lambda_2-\lambda_1 \varepsilon \mathcal{D}$. Hence

$$y(\lambda_2 - \lambda_1) + y(\lambda_1) = y(\lambda_2),$$

or

$$y(\lambda_2 - \lambda_1) = y(\lambda_2) - y(\lambda_1).$$

Then

$$-y(\lambda_2 - \lambda_1) = y(\lambda_1) - y(\lambda_2),$$

which gives

$$y(\lambda_1 - \lambda_2) = y(\lambda_1) + y(-\lambda_2).$$

Therefore (5.9.1) holds for all $\lambda_1, \lambda_2 \varepsilon \mathcal{D} \cup -\mathcal{D}$.

If n is any integer, then for all $\lambda \varepsilon \mathcal{D} \cup -\mathcal{D}$, $y(m\lambda)$ is defined, and since (5.9.1) now holds for all $\lambda_1, \lambda_2 \varepsilon \mathcal{D} \cup -\mathcal{D}$,

$$y([m+1]\lambda) = y(m\lambda) + y(\lambda).$$

It follows by induction that

$$y(m\lambda) = my(\lambda)$$

for all integers m and $\lambda \varepsilon \mathcal{D} \cup -\mathcal{D}$.

Q.E.D.

Lemma 5-10. Consider the case where $0 < F(\lambda) < 1$ for all positive $\lambda \varepsilon \mathcal{D}$. Then there is a positive constant θ such that

$$F(\lambda) = e^{-\theta\lambda}$$

and each system in Q is in the Gibbs canonical state with the density operator

$$D = \frac{e^{-\theta H}}{\text{trace } e^{-\theta H}}.$$

Proof: As above, define $y(\lambda) = \ln F(\lambda)$, and extend $y(\lambda)$ to $\mathfrak{D} \cup -\mathfrak{D}$ by the equation $y(-\lambda) = -y(\lambda)$. We shall show that $\frac{y(\lambda)}{\lambda} = -\theta$ for all positive $\lambda \varepsilon \mathfrak{D}$, where θ is a positive constant.

Suppose that $\frac{y(\lambda)}{\lambda}$ is not constant. Then for some $\lambda_1, \lambda_2 \varepsilon \mathfrak{D}$,

(5. 10. 1) $$\frac{y(\lambda_1)}{\lambda_1} \neq \frac{y(\lambda_2)}{\lambda_2} .$$

Let a and b be any two positive numbers, and consider the equations

$$\begin{aligned} \alpha y(\lambda_1) + \beta y(\lambda_2) &= a \\ \alpha\lambda_1 + \beta\lambda_2 &= b . \end{aligned}$$

By (5. 10. 1), the determinant $\lambda_2 y(\lambda_1) - \lambda_1 y(\lambda_2)$ is not zero; hence these equations can be solved for the real coefficients α and β. In general, α and β will be irrational numbers. However, we can always choose rational numbers $\frac{r_1}{s_1}$ and $\frac{r_2}{s_2}$, sufficiently close to α, β, so that

(5.10.2a) $$\frac{r_1}{s_1}\, y(\lambda_1) + \frac{r_2}{s_2}\, y(\lambda_2) > 0$$

(5.10.2b) $$\frac{r_1}{s_1}\lambda_1 + \frac{r_2}{s_2}\lambda_2 > 0$$

Moreover, we may suppose that the denominators s_1, s_2 are positive. Multiplying both sides of (5.10.2a) and (5.10.2b) by $s_1 s_2$, we obtain

$$m_1 y(\lambda_1) + m_2 y(\lambda_2) > 0$$

$$m_1\lambda_1 + m_2\lambda_2 > 0 ,$$

where m_1 and m_2 are integers (not necessarily positive). By Lemma 5-9, these inequalities become

$$y(m_1\lambda_1 + m_2\lambda_2) > 0$$

$$m_1\lambda_1 + m_2\lambda_2 > 0 .$$

Now $m_1\lambda_1 + m_2\lambda_2 \varepsilon \mathcal{D}$ and $\ell n\, F(m_1\lambda_1 + m_2\lambda_2) > 0$ contradict the fact that $0 < F(\lambda) < 1$ for all positive $\lambda \varepsilon \mathcal{D}$. Consequently, we must have

$$y(\lambda) = -\theta\lambda$$

for some constant θ. Here θ is positive since $0 < F(\lambda) < 1$.

Thus for each system $(\mathcal{H}, H)$ in Q, having the equilibrium density operator $D = f(H)$, we obtain, from (5. 5. 2),

$$f(\lambda) = f(0)e^{-\theta\lambda}$$

for all $\lambda \varepsilon S(H)$. It follows that

$$D = f(0)e^{-\theta H} .$$

Since trace $D = 1$, this is

$$D = \frac{e^{-\theta H}}{\text{trace } e^{-\theta H}} .$$

Q. E. D.

Our main theorem summarizes the results of Lemmas 5-7 and 5-10:

Theorem 5-11. Let Q be any collection of quantum systems, each in an equilibrium state, satisfying Axioms 1 through 6. Then each system is in a Gibbs canonical state at some absolute temperature $T \geq 0$. The temperature is arbitrary, but it is the same for all systems in Q.

BIBLIOGRAPHY

1. Aronsazajn, N. and K. T. Smith. Invariant subspaces of completely continuous operators. Annals of Mathematics 60(2): 345-350.

2. Birkhoff, G. and J. von Neumann. The logic of quantum mechanics. Annals of Mathematics 37:823-843. 1936.

3. Caratheodory, C. Variationsrechnung und partialle Differentialgleichungen erster ordnung. Leipzig, Teubner, 1935. 407 p.

4. Carter, D. S. Unpublished research notes on classical statistical mechanics. Berkeley, Calif. University of California, Mathematics Dept., 1960.

5. Finklestein, D. et al. Foundations of quaternion quantum mechanics. Journal of Mathematical Physics 3(2):207-220. 1962.

6. Gibbs, J. W. Elementary principles of statistical mechanics. Hew Haven, Yale University Press, 1902. 207 p.

7. Gleason, A. M. Measures on Hilbert space. Journal of Mathematics and Mechanics 6:885-893. 1953.

8. Halmos, P. R. Introduction to Hilbert space and the theory of spectral multiplicity. New York, Chelsea, 1957. 114 p.

9. ______________. Measure theory. New York, van Nostrand, 1950. 304 p.

10. Hilbert, D. Grundzuge einer allgemeinen Theorie der linearen Integral-gleichungen. Leipzig, Teubner, 1912. 282 p. (Reprinted by Chelsea, New York, 1952)

11. Jaynes, E. T. Information theory and statistical mechanics. Physical Review 106:620-630, 1957.

12. Khinchin, A. Y. Mathematical foundations of quantum statistics. Tr. by Irwin Schapiro. Albany, New York, Graylock Press, 1960. 232 p.

13. ______________. Mathematical foundations of statistical mechanics. New York, Dover, 1949. 179 p.

14. Kurth, R. Axiomatics of classical statistical mechanics. New York, Pergamon Press, 1960. 180 p.

15. Kolmogorov, A. N. and S. V. Formin. Elements of the theory of functions and functional analysis. Tr. by Hyman Kamel and Horace Komm. Albany, New York, Graylock Press, 1961. 2 vols.

16. Mackey, G. W. Mathematical foundations of quantum mechanics. New York, W. A. Benjamin, 1963. 137 p.

17. Neumann, J. von. Mathematical foundations of quantum mechanics. Tr. by Robert T. Beyer. Princeton, Princeton University Press, 1955. 445 p.

18. Neumann, J. von and F. Murray. On rings of operators. Annals of Mathematics 37(2):116-229. 1936.

19. Riesz, F. and B. von Sz-Nagy. Functional analysis. Tr. by Leo F. Boron. New York, Ungar, 1955. 467 p.

20. Schatten, R. Normed ideals of completely continuous operators. (ergebnisse der Mathematik und ihrer Grenzgebiete, n. s. no. 27). Berlin, Springer-Verlag, 1960. 81 p.

21. Stone, M. H. Linear transformations in Hilbert space and their applications to analysis. New York, 1932. 622 p. (American Mathematical Society. Colloquim Publications, vol. 15)

22. ______________. On one-parameter unitary groups in Hilbert space. Annals of Mathematics 33:643-648. 1932.

APPENDIX

APPENDIX

The fact that density operators are completely continuous is fundamental to our discussion (see the discussion following Definition 3-6), and we shall include here a proof of this theorem.

Completely continuous operators on a Hilbert space have been discussed rather thoroughly in the literature (see N. Aronsazajn and K. T. Smith [1], and the many texts on functional analysis such as F. Riesz and B. Sz-Nagy [19, p. 203]. For a recent discussion of spaces of completely continuous operators, the reader is referred to R. Schatten [20]). It is well known that a completely continuous operator on an infinite-dimensional Hilbert space need not admit a characteristic value other than zero. However, much more can be said if the operator is also assumed to be self-adjoint. On a finite-dimensional unitary space, every self-adjoint operator admits a basis in that space made up of its characteristic vectors. The infinite-dimensional extension of this theorem asserts that every self-adjoint completely continuous operator T on a Hilbert space $\mathcal{H}$ (separable or non-separable) admits a complete orthonormal family of characteristic vectors. Its non-zero (necessarily real) characteristic values are of finite multiplicity and form either a finite or denumerably infinite sequence $\{\lambda_k\}$. If the corresponding sequence of characteristic vectors $\{\phi_k\}$ is finite or denumerable but not complete, we may represent any $f \varepsilon \mathcal{H}$ as

$$f = \sum_{k} (f, \phi_k)\phi_k + g$$

where g is some vector in the null space of T. Thus

$$Tf = \sum_{k} (f, \phi_k)\lambda_k \phi_k .$$

Adjoining to $\{\phi_k\}$ any orthonormal basis in the null space of T then gives a complete orthonormal family of characteristic vectors of T. Of course if $\mathcal{H}$ is a separable space, this family is a sequence.

We can go one step further if zero is not a characteristic value of T. In fact, in order that the sequence $\{\phi_k\}$ corresponding to the non-zero characteristic values $\{\lambda_k\}$ form a complete orthonormal sequence in $\mathcal{H}$, it is necessary and sufficient that zero not be a characteristic value of T, that is $Tf \neq 0$ whenever $f \neq 0$. The characteristic values are still of finite multiplicity but denumerable in number. Thus for any $f \varepsilon \mathcal{H}$, we have the simple decompositions

$$f = \sum_{k=1}^{\infty} (f_1 \phi_k)\phi_k$$

$$Tf = \sum_{k=1}^{\infty} (f_1 \phi_k)\lambda_k \phi_k ,$$

where λ_k runs through the characteristic values of T each as many times as its multiplicity indicates. In this latter case, we necessarily have $\lambda_k \to 0$.

The following proof is based on a proof given by R. Schatten [20, p. 41].

Theorem. Let T be a non-negative, self-adjoint operator of the trace class defined on a separable Hilbert space $\mathcal{H}$. Then T is completely continuous.

Proof: Let $\{\phi_\ell\}$ be a complete orthonormal sequence in $\mathcal{H}$. Since $T \epsilon (tc)$ and is non-negative, we have

$\sum_{\ell=1}^{\infty} (T\phi_\ell, \phi_\ell) = M$, where M is a non-negative real number.

Then

$$\left[\sum_{\ell=1}^{\infty} (T\phi_\ell, \phi_\ell) \right]^2 = M^2 = \sum_{n=1}^{\infty} (T\phi_n, \phi_n) \sum_{\ell=1}^{\infty} (T\phi_\ell, \phi_\ell)$$

$$= \sum_{n, \ell=1}^{\infty} (T\phi_n, \phi_n)(T\phi_\ell, \phi_\ell).$$

Since T is also self-adjoint, it satisfies the inequality

$$|(Tg, f)| \leq [(Tf, f)(Tg, g)]^{1/2}$$

for all $f, g \varepsilon \mathcal{H}$ [17, p. 101]. Hence

$$M^2 \geq \sum_{n,\ell=1}^{\infty} |(T\phi_\ell, \phi_n)|^2 = \sum_{n,\ell=1}^{\infty} (T\phi_n, \phi_\ell)(\phi_\ell, T\phi_n) .$$

Since $\{\phi_\ell\}$ is a complete orthonormal sequence, we have

$$(f, g) = \sum_{\ell=1}^{\infty} (f, \phi_\ell)(\phi_\ell, g)$$

for all $f, g \varepsilon \mathcal{H}$. Therefore

$$M^2 \geq \sum_{n=1}^{\infty} (T\phi_n T\phi_n) = \sum_{n=1}^{\infty} \|T\phi_n\|^2 ,$$

and $\sum_{n=1}^{\infty} \|T\phi_n\|^2$ must converge since its terms are all non-negative.

Now let $\{f_m\}$ be a weakly convergent sequence: $f_m \xrightarrow{w} f$. Then [15, vol. 1, p. 90] for some constant c, we have $\|f_m\| \leq c$ for all m, and therefore $\|f\| \leq c$. Since $\sum_{\ell=1}^{\infty} \|T\phi_n\|^2$ converges, for each $\epsilon > 0$ there exists a positive integer n_0 such that

$$\sum_{n=n_0}^{\infty} \|T\phi_n\|^2 < \frac{\epsilon^2}{16c^2} .$$

Expanding $f_m - f$, we obtain

$$f_m - f = \sum_{n=1}^{\infty} (f_m - f, \phi_n)\phi_n,$$

and

$$Tf_m - Tf = \sum_{n=1}^{\infty} (f_m - f, \phi_n)T\phi_n .$$

For every positive integer m,

$$\|Tf_m - Tf\|^2 = \|\sum_{n=1}^{\infty} (f_m - f, \phi_n)T\phi_n\|^2 = \|\sum_{n=1}^{n_0 - 1}(f_m - f, \phi_n)T\phi_n + \sum_{n=n_0}^{\infty}(f_m - f, \phi_n)T\phi_n\|^2$$

$$\leq 2\|\sum_{n=1}^{n_0 - 1}(f_m - f, \phi_n)T\phi_n\|^2 + 2\|\sum_{n=n_0}^{\infty}(f_m - f, \phi_n)T\phi_n\|^2 .$$

The first term in the last expression is a finite sum, so for sufficiently large m, say $m > m_0$, we have

$$2\|\sum_{n=1}^{n_0 - 1}(f_m - f, \phi_n)T\phi_n\|^2 < \frac{\epsilon^2}{2} ,$$

since $f_m \xrightarrow{w} f$ implies $\lim_{m\to\infty} (f_m - f, \phi_n) = 0$. The second term is less than or equal to

$$2 \sum_{n=n_0}^{\infty} \|(f_m - f, \phi_n)T\phi_n\|^2 = 2\sum_{n=n_0}^{\infty} |(f_m - f, \phi_n)|^2 \|T\phi_n\|^2$$

$$\leq 2\sum_{n=n_0}^{\infty} |(f_m - f, \phi_n)|^2 \sum_{n=n_0}^{\infty} \|T\phi_n\|^2 .$$

Since

$$\sum_{n=1}^{\infty} |(f_m - f, \phi_n)|^2 = \|f_m - f\|^2 ,$$

this last expression is less than or equal to

$$2\|f_m - f\|^2 \sum_{n=n_0}^{\infty} \|T\phi_n\|^2,$$

and since $\|f_m\|$, $\|f\| \leq c$ implies $\|f_m - f\| \leq 2c$, we have

$$2\|f_m - f\|^2 \sum_{n=n_0}^{\infty} \|T\phi_n\|^2 \leq 2(2c)^2 \frac{\epsilon^2}{16c^2} = \frac{\epsilon^2}{2}$$

for all $m > m_0$. Therefore

$$2\| \sum_{n=n_0}^{\infty} (f_m - f, \phi_n) \, T\phi_n\|^2 \leq \frac{\epsilon^2}{2},$$

and consequently

$$\|Tf_m - Tf\|^2 < \epsilon^2$$

for all $m > m_0$. Thus we have shown that $f_m \xrightarrow{w} f$ implies $\|Tf_m - Tf\| \to 0$, that is, T is completely continuous.

Q. E. D.