

# THE PHYSICAL REVIEW

*A journal of experimental and theoretical physics established by E. L. Nichols in 1893*

SECOND SERIES, VOL. 107, No. 2

JULY 15, 1957

## Ergodic Theorem for Interacting Systems\*

H. EKSTEIN

Argonne National Laboratory, Lemont, Illinois

(Received March 18, 1957)

von Neumann has proved a quantum-mechanical ergodic theorem which deals with a closed macroscopic system and macroscopic observables. Certain unsatisfactory features of this theorem are probably due to the overly-strong idealization of a completely closed system.

Since the claims of statistical mechanics go beyond macroscopic systems, there is a need for a more general theorem. The present paper deals with a system of any number of degrees of freedom and with any observable, during interaction with a temperature bath. In order to avoid an assumption of randomness, the bath must be described explicitly and precisely by a time-independent Hamiltonian. The model of the bath is obtained by taking the Gibbs ensemble seriously, i.e., as a set of  $N$  identical systems interacting through a potential  $\lambda V$ .

The following theorem is proved. The time average of the quantum-mechanical expectation value of any observable with respect to any initial state is equal to its statistical average, in the double limit  $N \rightarrow \infty$ ,  $\lambda \rightarrow 0$ , for the overwhelming majority of all interaction potentials  $V$ .

### 1. INTRODUCTION AND STATEMENT

IT is well known that the ergodic theorem in quantum mechanics cannot be proved for all observables in a completely closed system. An ergodic theorem for macroscopic observables in a closed macroscopic system has been proved by von Neumann.<sup>1</sup>

However, the claims of statistical mechanics go beyond macroscopic systems and macroscopic observables.<sup>2</sup> The macrocanonical statistical theory claims that the average properties of a system with a Hamiltonian  $K(p, q)$  in loose interaction with a temperature bath of temperature  $T$  are given by the density operator

$$\rho = \exp(-K/kT) [\text{tr} \exp(-K/kT)]^{-1}, \quad (1.1)$$

so that the statistical mean value  $\langle A \rangle_{st}$  of any observable  $A$  associated with an operator  $A(p, q)$  is given by

$$\langle A \rangle_{st} = \text{tr}(A\rho). \quad (1.2)$$

To legitimize this claim, an ergodic theorem would have to prove that the time average of the quantum-mechanical expectation value of the Heisenberg

operator  $A(t)$ ,

$$\langle \bar{A} \rangle_t = \lim_{\tau \rightarrow \infty} (1/\tau) \int_0^\tau (\Psi(0), A(t)\Psi(0)) dt, \quad (1.3)$$

is equal to  $\langle A \rangle_{st}$  for any initial state  $\Psi(0)$  of the system.

Another motivation for seeking a more general ergodic theorem is the doubtful idealization of a completely closed system. Fierz<sup>3</sup> has pointed out that certain unsatisfactory assumptions on which von Neumann's theorem is based, probably arise from the undue idealization of the totally isolated system. Fierz conjectures that the introduction of random perturbations would remedy these defects. If these perturbations are not to be mathematical fictions, they must represent a temperature bath. They will, then, do more than merely establish equilibrium in a system of given energy; they will bring the system to the temperature of the bath. They will also cancel any conservation law, such as conservation of angular momentum, for the system.

It seems difficult and also arbitrary, to invent a system of perturbations which replaces the temperature bath; and it appears unavoidable that a full and precise description of the temperature bath must represent it

\* Work performed under the auspices of the U. S. Atomic Energy Commission.

<sup>1</sup> J. von Neumann, *Z. Physik* **57**, 30 (1929).

<sup>2</sup> E. Schrödinger, *Statistical Thermodynamics* (Cambridge University Press, Cambridge, 1948).

<sup>3</sup> M. Fierz, *Helv. Phys. Acta* **28**, 705 (1955).

as a system described by a Hamiltonian. Naturally, one has to choose a definite model for the temperature bath.

It is natural to adapt, for this purpose, Gibbs' canonical ensemble. Instead of merely thinking of juxtaposed systems, we take the model literally: The temperature bath consists of a large number  $N$  of identical systems with Hamiltonians  $K(p_i, q_i)$  for the  $i$ th system and an interaction  $\lambda V(p_1 q_1 \cdots p_N q_N)$ . The Hamiltonian of the total bath is

$$H \equiv H_0 + \lambda V = \sum_{i=1}^N K(p_i, q_i) + \lambda V(p_1 q_1 \cdots p_N q_N), \quad (1.4)$$

where  $p_i, q_i$  stand for the dynamical variables of the  $i$ th system. The spatial separation of the systems is expressed by boundary conditions which require the wave function to vanish in those parts of configuration space where particles belonging to the  $i$ th system are outside the boundaries of this system.

Measurements are thought of as being made only in one system (for example,  $i=1$ ) which will be referred to as "the system." The totality of all systems, described by the Hamiltonian  $H$ , will be called the bath (although it contains the system). Since we are interested only in the physical properties of the system, we shall consider only observables  $A$  which are defined with respect to the system. The corresponding operators  $A$  act only on the dynamical variables of the system, and are unit operators with respect to other dynamical variables.

The close analogy between the mental picture of the ensemble and our physical bath makes it convenient to express Eqs. (1.1) and (1.2) in the simpler form from which these equations are usually derived.<sup>2</sup> If  $\varphi_s(q_i)$  are the eigenfunctions of the  $i$ th system considered alone, i.e.,

$$K(p_i, q_i) \varphi_s(q_i) = \epsilon_s \varphi_s(q_i), \quad (1.5)$$

then a system of eigenfunctions of the unperturbed Hamiltonian  $H_0$  is given by

$$\Phi_n(q_1 \cdots q_N) = \prod_{i=1}^N (\varphi_s)_i(q_i). \quad (1.6)$$

Because of the spatial separation of the systems, no symmetrization of  $\Phi_n$  is necessary. Let

$$A_{nn} = (\Phi_n, A \Phi_n). \quad (1.7)$$

Then the fundamental hypothesis of statistical mechanics is that the average value of an observable  $A$  is obtained by assigning equal weights to all values  $A_{nn}$  which belong to the same energy  $E$  of the ensemble, so that

$$\langle A \rangle_{st} = \sum_{n=1}^R A_{nn} (1/R), \quad (1.8)$$

where  $R$  is the number of independent eigenfunctions

(1.6) belonging to an energy  $E$  of the ensemble energy  $E$ , i.e.,  $R$  is the degree of degeneracy of this eigenvalue. The summation in Eq. (1.8) extends over this degenerate subspace  $\Gamma$ . The Eqs. (1.1) and (1.2) are derived from the assumption (1.8) by counting the number of terms in (1.8) which have the particular value  $A_{nn}$ , and going to the limit  $N \rightarrow \infty$ .<sup>2</sup>

In this argument, the interaction between the systems is used only verbally, to make an exchange of energy between systems plausible, but it is not included in the Hamiltonian. Clearly, an ergodic theorem must make explicit use of the interaction, but it also is clear that the interaction must not appear explicitly in the final result, if it is to agree with Eq. (1.8). Hence, the ergodic theorem must deal with a bath described by a Hamiltonian  $H$  in the double limit  $\lambda \rightarrow 0, N \rightarrow \infty$ . For the validity of the ergodic theorem, two further conditions must be satisfied:

(a) The bath should not have any essential degeneracy, for if it did, an initial state of a given symmetry would always maintain this property and no equal distribution can be expected. In fact, we shall have to assume that the full Hamiltonian (1.4) is completely nondegenerate. In the absence of an essential degeneracy this is a plausible assumption. Since potential interactions are symmetric with respect to particle exchange, the removal of essential degeneracy requires that the interaction  $V$  contain also terms which act in different ways on particles of different systems: e.g., different electric fields within each system.

(b) Certain exceptional types of potential  $V$  must be excluded. For, if an eigenfunction  $\Psi_n$  of  $H$  should coincide with an eigenfunction  $\Phi_n$  of  $H_0$ , the bath would not be changed by the "switching on" of the interaction, i.e., in other words, the Hamiltonian would not properly represent a temperature bath. Fortunately, it will not be necessary to specify the properties of  $V$  explicitly. It will be possible to prove the ergodic theorem for the overwhelming majority of all interaction potentials, i.e., with the exception only of singular cases, and it will not be necessary to exclude them specifically.

We are now prepared to state the ergodic theorem: Given any initial state  $\Psi(0)$  of the bath, and that the "unperturbed" energy of  $\Psi(0)$  is  $E$  (i.e.,  $\Psi(0)$  is in the subspace  $\Gamma$ ), the time average  $\langle \bar{A} \rangle_t$  differs from its statistical average by an arbitrarily small number, as the number of systems  $N$  increases and the "strength" of the interaction  $\lambda$  decreases, for almost all interaction operators  $V$ .

## 2. PROOF OF THE THEOREM

The mathematical methods to be used are those developed by von Neumann<sup>1</sup> and by Pauli and Fierz<sup>4</sup> in connection with the microcanonical ergodic theorem. The following is an adaptation of their calculation to the present statement.

<sup>4</sup> W. Pauli and M. Fierz, *Z. Physik* **106**, 572 (1937).

If  $\Psi_n$  are the eigenfunctions of  $H$ , the initial state  $\Psi(0)$  is expanded as

$$\Psi(0) = \sum c_n \Psi_n. \quad (2.1)$$

The time average of the observable  $A$  is

$$\langle \bar{A} \rangle_t = \langle (\Psi(0), A(t)\Psi(0)) \rangle_t = \sum c_n^* c_m (\Psi_n, A\Psi_m) \times \langle \exp[i(E_n - E_m)t] \rangle_t, \quad (2.2)$$

where the operator  $A$  on the right-hand side is a Schrödinger operator. Because all energy eigenvalues belonging to different states are different by assumption, the nondiagonal elements are cancelled by the time integration, so that

$$\langle \bar{A} \rangle_t = \sum |c_n|^2 (\Psi_n, A\Psi_n). \quad (2.3)$$

Transformation to the system  $\Phi_n$  of Eq. (1.6) gives

$$\langle \bar{A} \rangle_t = \sum_{nlm} |c_n|^2 (\Psi_n, \Phi_m) (\Phi_l, \Psi_n) (\Phi_m, A\Phi_l). \quad (2.4)$$

One might think that it would be sufficient to consider, from here on, only the properties of the matrix  $\sum |c_n|^2 (\Psi_n, \Phi_m) (\Phi_l, \Psi_n)$ , since the operator  $A$  is not restricted to be macroscopic. Actually, the fact that  $A$  acts only on the variables of the system makes it highly degenerate in the  $H_0$  representation, and this is an essential part of the proof.

Consider now the transformation matrix

$$T_{nm} = (\Phi_n, \Psi_m). \quad (2.5)$$

It is known from the perturbation theory of degenerate Hermitian operators that the expansion of  $T$  in powers of the coupling constant  $\lambda$  has a zeroth term:

$$T_{nm} = T_{nm}^0 + \lambda T_{nm}^1 + \dots \quad (2.6)$$

This operator  $T^0$  maps the degenerate subspaces  $\Gamma$  of  $H_0$  into themselves, i.e., it forms the "correct linear combinations" of the  $\Phi_n$  belonging to a given eigenvalue  $E$ . Since we are ultimately interested in the limit  $\lambda=0$  and since, by assumption,  $\Psi(0)$  belongs to  $\Gamma$ , we may limit the summations in Eqs. (2.1), (2.2), (2.3), and (2.4) to the subspace  $\Gamma$ , i.e., to functions  $\Phi_n$  which belong to  $E$  and functions  $\Psi_n$  which are linear combinations of these. In the following, we may replace  $\Psi_n$  by  $\sum_{\Gamma} T_{nm}^0 \Phi_m$ .

The operator  $A$  acts, by assumption, as a unit operator on all factor functions  $(\varphi_s)_i$  of Eq. (1.6) except when  $i=1$ . Therefore all matrix elements  $(\Phi_m, A\Phi_l)$  vanish, unless all functions  $\varphi_{s_i}$  ( $i \neq 1$ ) are the same for both the products  $\Phi_m$  and  $\Phi_l$ . Furthermore, the two functions  $(\varphi_s)_1^{(m)}(q_1)$  and  $(\varphi_s)_1^{(l)}(q_1)$  belonging to the product functions  $\Phi_m$  and  $\Phi_l$ , respectively, must have the same energy eigenvalues  $\epsilon_s$ , because otherwise the energies  $\sum_{i=1}^N (\epsilon_s)_i = E$  could not be equal. If  $K(p_1, q_1)$  is nondegenerate, this shows that nondiagonal matrix elements  $(\Phi_m, A\Phi_l)$  are not contained in the sum (2.4).

If  $K(p_1, q_1)$  is degenerate, it is possible to choose the degenerate eigenfunctions  $\varphi_{s,\alpha}$  so that the nondiagonal elements  $(\varphi_{s,\alpha}, A\varphi_{s,\beta})$  vanish. With this choice, the only surviving matrix elements are diagonal with respect to all subscripts, and Eq. (2.4) reduces to

$$\langle \bar{A} \rangle_t = \sum_{i=1}^R |c_n|^2 A_{ii} |\langle \Phi_i, \Psi_n \rangle|^2, \quad (2.7)$$

where  $R$  is the number of functions  $\Phi_n$  in  $\Gamma$ .

Every numerical value  $A_{ii} = A_\nu$  occurs many times in the sum (2.7) because  $A_\nu$  depends only on the functions  $(\varphi_s)_1(q_1)$ , whereas numerous possible choices of the functions  $(\varphi_s)_i$  ( $i \neq 1$ ) are compatible with  $\sum (\epsilon_s)_i = E$ . If  $d_\nu$  is the number of different functions  $\Phi_i$  which lead to one value  $A_{ii} = A_\nu$ , and  $M$  the number of different values  $A_\nu$  in  $\Gamma$ ,

$$\sum_{\nu=1}^M d_\nu = R. \quad (2.8)$$

It is convenient to introduce double subscripts for  $\Phi_i \rightarrow \Phi_{\nu,\lambda}$ , so that the first subscript denotes the corresponding value of  $A_{ii} = A_\nu$ . Then Eq. (2.7) can be written

$$\langle \bar{A} \rangle_t = \sum_{\nu=1}^M A_\nu \sum_{n=1}^R |c_n|^2 \sum_{\lambda=1}^{d_\nu} |\langle \Phi_{\nu,\lambda}, \Psi_n \rangle|^2. \quad (2.9)$$

With the new notation, Eq. (1.8) reads

$$\langle A \rangle_{st} = \sum_{\nu=1}^M A_\nu (d_\nu/R). \quad (2.10)$$

To compare  $\langle A \rangle_{st}$  with  $\langle \bar{A} \rangle_t$ , we form the expression

$$\langle \langle A \rangle_{st} - \langle \bar{A} \rangle_t \rangle^2 = \left( \sum_{\nu=1}^M A_\nu \left[ \sum_n |c_n|^2 \sum_{\lambda=1}^{d_\nu} |\langle \Phi_{\nu,\lambda}, \Psi_n \rangle|^2 - (d_\nu/R) \right] \right)^2. \quad (2.11)$$

Since the normalization of  $\Psi(0)$  requires  $\sum |c_n|^2 = 1$ ,

$$\langle \langle \bar{A} \rangle_t - \langle A \rangle_{st} \rangle^2 = \left( \sum_{\nu=1}^M A_\nu (d_\nu/R)^{\frac{1}{2}} (R/d_\nu)^{\frac{1}{2}} \times \sum_{n=1}^R |c_n|^2 \{ (P_\nu \Psi_n, \Psi_n) - (d_\nu/R) \} \right)^2, \quad (2.12)$$

in which, for any function  $F$ ,

$$(P_\nu F, F) = \sum_{\lambda=1}^{d_\nu} |\langle \Phi_{\nu,\lambda}, F \rangle|^2. \quad (2.13)$$

Here,  $P_\nu$  is the operator of the projection into the subspace spanned by the  $d_\nu$  functions  $\Phi_{\nu,\lambda}$  which belong to  $(\Phi_{\nu,\lambda}, A\Phi_{\nu,\lambda}) = A_\nu$ . We now use Schwarz's inequality to

estimate the difference.

$$\langle\langle\bar{A}\rangle_t - \langle A \rangle_{st}\rangle^2 \leq \sum_{\nu=1}^M A_\nu^2 (d_\nu/R) \sum_{\nu=1}^M (R/d_\nu) \sum_{n=1}^R |c_n|^2 [(P_\nu \Psi_n, \Psi_n) - (d_\nu/R)]^2. \quad (2.14)$$

According to Eq. (2.10) the first sum is the statistical average  $\langle A^2 \rangle_{st}$  of the operator  $A^2$ . In order to simplify the expression further, we use Schwarz's inequality again on the second sum in Eq. (2.14), noting that  $\sum |c_n|^2 = 1$ :

$$\langle\langle\bar{A}\rangle_t - \langle A \rangle_{st}\rangle^2 \leq \langle A^2 \rangle_{st} \sum_{\nu=1}^M (R/d_\nu) \sum_{n=1}^R |c_n|^2 \times [(P_\nu \Psi_n, \Psi_n) - (d_\nu/R)]^2. \quad (2.15)$$

We cannot expect that this quantity is arbitrarily small, in general, since we have not eliminated the exceptional interactions  $V$  which would certainly make an ergodic theorem invalid. The averaging over potentials  $V$  will be done in the same manner in which von Neumann<sup>1</sup> averaged over systems of observed quantities. We average Eq. (2.15) over all possible interactions  $V$  in the following sense: A given  $V$  leads to a unitary transformation matrix  $T^0$  in  $R$ -dimensional space which rotates the system  $\Psi_n$  with respect to that of the  $\Phi_n$ . The averaging consists in assigning successively all possible values to  $T^0$ , the unitary matrix in  $R$ -dimensional space. The quantity of interest is, then, the geometric average of

$$x = \sum_{\nu=1}^M (R/d_\nu) \sum_{n=1}^R |c_n|^2 [(P_\nu \Psi_n, \Psi_n) - (d_\nu/R)]^2. \quad (2.16)$$

Von Neumann<sup>1</sup> and Pauli and Fierz<sup>4</sup> have shown that this average is bounded by the inequality

$$\{x\}_{\lambda\kappa} \leq \int_0^a W(u) u du + \exp[-\kappa a (R/M)^{\frac{1}{2}} + \kappa + 2 \ln R]. \quad (2.17)$$

Here,  $a$  is a positive number

$$a > 2M(R-2)^{-1}, \quad (2.18)$$

$\kappa$  is defined by

$$\kappa = 1 - \ln 2, \quad (2.19)$$

and  $W(u)$  is the geometric probability that  $x$  is between  $u$  and  $u+du$ .

In order to estimate  $M$ , the number of different values  $A_\nu$ , we note that it certainly cannot be larger than the number of different system eigenstates  $\varphi_s$  compatible with  $\epsilon_s \leq E$ .

In the course of the limiting process where the number of systems,  $N$ , increases indefinitely, the mean energy,  $U = (E/N)$  remains fixed. At the same time, the number of accessible system states increases. The asymptotic value will depend only on the number of system states whose energies approach the total energy  $NU$ . We consider a system of  $n$  particles at very high energies  $\epsilon \sim NU$ . At sufficiently high energies, all physical systems behave like ideal gases, i.e., the potential energy becomes unimportant. For ideal gases, the number  $\Omega$  of states of energy  $\epsilon$  is given by<sup>2</sup>

$$\ln \Omega = \frac{3}{2} n \ln \epsilon + \dots, \quad (2.20)$$

in which the omitted terms do not contribute to the asymptotic dependence on  $N$ . The total number  $M'$  of states of a perfect gas for energies from 0 to  $\epsilon = E$  is

$$M' \leq \int_0^E \epsilon^{\frac{3}{2}n} d\epsilon = \frac{2}{3} n N U^{1+\frac{3}{2}n}, \quad (2.21)$$

constants independent of  $N$  again being omitted. Since our idealization of the system as a perfect gas is valid only for very high energies, the actual number of accessible states  $M$  will be

$$M = M' + c, \quad (2.22)$$

where  $c$  is independent of  $E$ , if  $E$  is only sufficiently large. On the other hand, the total number  $R$  of states  $\Phi_n$  is, according to well-known results of statistical mechanics,<sup>2</sup> given by

$$k \ln R = NS, \quad (2.23)$$

where  $S$  is the (statistical) entropy of the system which is, like  $U$ , independent of  $N$ .

Since  $M$  increases like a finite power of  $N$ , the exponential increase of  $R$  with  $N$  dominates, and the exponent in the second term of Eq. (2.17) becomes negatively infinite for any positive number  $a$ . But, for the same reason, the number  $a$  in the inequality (2.18) may be chosen arbitrarily small, so that the first term in Eq. (2.17) can also be made to vanish as  $N$  increases.

We have shown that the average over all potentials vanishes in the limit:

$$\lim_{N \rightarrow \infty, \lambda \rightarrow 0} \{ \langle\langle\bar{A}\rangle_t - \langle A \rangle_{st} \rangle^2 \}_{\lambda\kappa} = 0. \quad (2.24)$$

Since the averaged quantity is positive, we may say that it is zero for the overwhelming majority of all potentials  $V$ . This concludes the proof of the theorem.

#### ACKNOWLEDGMENTS

The author is indebted to Peter G. Bergmann for critical comments, and to F. E. Thow for a review of the manuscript.