PHYSICAL REVIEW VOLUME 111, NUMBER 3 AUGUST 1, 1958

Ergodic Theorem in Quantum Mechanics

P. BOCCHIERI AND A. LOINGER
Istituto Nazionale di Fisica Nucleare, Sezione di Milano, Italy and Istituto di Fisica dell'Università, Pavia, Italy
(Received April 9, 1958)

Landsberg and Farquhar have shown that the von Neumann-Fierz approach to the quantum ergodic problem does not give any criterion of ergodicity.

In this paper it is shown that von Neumann's procedure of averaging over the macro-observers is sufficient by itself to "justify" at every time the use of the quantum microcanonical ensemble. It is then proved that the time behavior of the system is completely irrelevant to the demonstration of the von Neumann and von Neumann-Fierz ergodic theorems. It is concluded that the quantum ergodic problem must be attacked on entirely new lines.

I

THE ergodic problem of classical mechanics was attacked in a series of papers by von Neumann, Birkhoff, Hopf, and others.[1] The fundamental researches of these authors started from the "general dynamics" and from the formulation of classical mechanics in Hilbert space, due to Koopman[2] and von Neumann.[3] The condition of ergodicity required for the validity of Birkhoff's theorem is metric transitivity of the dynamical system. The important problem of singling out the class of the metrically transitive systems, whose existence is not forbidden by topological reasons,[4] is still open.

Shortly before the development of the classical ergodic theory, von Neumann[5] investigated the quantum-mechanical ergodic problem. In his approach the condition of ergodicity was the absence of degeneracies and resonances in the energy spectrum. By averaging over all macro-observers and making some qualitative assumptions on the density of quantum states in the "phase-cells," von Neumann was able to prove the ergodic theorem (and the H theorem).

Pauli and Fierz[6] simplified von Neumann's proof and evaluated explicitly the probability of finding a non-thermodynamical observer, i.e., an exceptional macro-observer for which the ergodic and H theorems do not hold.

Recently Fierz[7] objected to von Neumann's definition of entropy and suggested a somewhat different definition; according to this the H theorem can be proved to follow from a weaker ergodic theorem than von Neumann's. This ergodic theorem was derived by Fierz without the assumption of no resonances in the energy spectrum. In his paper the analog of the classical metric transitivity was the absence of degeneracies. He strongly emphasized, however, that the assumption of *equal a priori* probability of all macro-observers is completely unfounded from a *physical* point of view and consequently proposed to abandon von Neumann's approach.

Later, Landsberg and Farquhar[8] removed from the von Neumann-Fierz theorem also the assumption of no degeneracies, thus reaching the important conclusion that the von Neumann-Fierz approach fails to give any criterion of ergodicity. Consequently every system would be ergodic, provided that its quantum states are sufficiently numerous in every "phase-cell."

In this paper we prove that the von Neumann and von Neumann-Fierz ergodic theorems follow solely from the averaging over the macro-observers, while time evolution of the system is completely irrelevant to the validity of the theorems and could even not be governed by the quantum equations of motion. In other words, the averaging over the macro-observers is a kind of statistical procedure, which only "accounts" for the ergodic theorems.

The inadequacy of the von Neumann's approach is therefore mathematically proved. A satisfactory proof of the quantum ergodic theorem is still lacking.

II

Let us consider an isolated dynamical system characterized by a Hamiltonian H. Let $\Psi(t)$ be the state-vector of the system in the Schrödinger picture. We shall denote by "macro-observer" an observer who can make measurements with limited accuracy only. The statistics of the energy determined by a macro-observer is well represented by the macro-energy operator $\mathcal{H}=\sum_a \mathcal{E}_a P_{M_a}$. In this formula $\mathcal{E}_a$ is the energy value found with certainty by a macro-observer when the state-vector of the system belongs to the S-dimensional manifold M_a ("energy shell"), whose projection operator is P_{M_a}.

We shall assume, for the sake of simplicity only, that the state-vector belongs to the manifold M_a. The measurement of a set of commuting macro-observables made by a macro-observer results in projecting the vector $\Psi(t)$ in a s_ν-dimensional subspace of the energy

[1] E. Hopf, *Ergodentheorie* (Springer-Verlag, Berlin, 1937).
[2] B. O. Koopman, Proc. Natl. Acad. Sci. U. S. **17**, 315 (1931).
[3] J. von Neumann, Ann. Math. **33**, 587 (1932).
[4] I. Oxtoby and S. Ulam, Ann. Math. **42**, 874 (1941). In this paper is also recalled a plausibility argument according to which "almost every" measure-preserving continuous transformation is metrically transitive.
[5] J. von Neumann, Z. Physik **57**, 30 (1929).
[6] W. Pauli and M. Fierz, Z. Physik **106**, 572 (1937).
[7] M. Fierz, Helv. Phys. Acta **28**, 705 (1955).
[8] I. E. Farquhar and P. T. Landsberg, Proc. Roy. Soc. (London) **A239**, 134 (1957).

shell ($1<s_\nu<S$). This subspace is called a "phase cell," and all the cells corresponding to the possible eigenvalues of all the commuting macro-observables constitute the S-dimensional unitary space, which we have called the energy shell. Let us suppose that the energy shell contains N cells; evidently

$$\sum_{\nu=1}^{N} s_\nu = S.$$

This subdivision into cells of the energy shell characterizes a macro-observer. We shall refer the ν cell to an orthonormal basis $\{\omega_{\nu,j}\}$ $(j=1,\cdots s_\nu)$. All the vectors $\omega_{\nu,j}$ $(\nu=1,\cdots N;\, j=1,\cdots s_\nu)$ constitute a basis of the energy shell.

Let us consider now a second macro-observer, who makes measurements on a different set of commuting macro-observables, with the same accuracy as the first. We shall obtain a second subdivision of the energy shell into N cells, the ν cell containing s_ν quantum states. Again we refer every cell to an orthonormal basis, obtaining in this way another basis $\{\omega'_{\nu,j}\}$ $(\nu=1,\cdots N;\, j=1,\cdots s_\nu)$ for the energy shell; and so on for all possible macro-observers.

III

We expand the state-vector $\Psi(t)$ with respect to a basis $\{\omega_{\nu,j}\}$ of the energy shell, which diagonalizes all the macro-variables measured by a certain macro-observer. We obtain:

$$\Psi(t)=\sum_{\nu=1}^{N}\sum_{j=1}^{s_\nu}\tau_{\nu,j}(t)\omega_{\nu,j}\equiv\sum_{k=1}^{S}\tau_k(t)\omega_k. \qquad (1)$$

When this macro-observer makes a maximal macro-measurement, he finds the system in the ν cell with a probability:

$$u_\nu(t)=\sum_{j=1}^{s_\nu}|\tau_{\nu,j}(t)|^2=\sum_{j=1}^{s_\nu}|(\Psi(t),\omega_{\nu,j})|^2. \qquad (2)$$

We shall now prove that *the probability at a given time t, $u_\nu(t)$, equals the microcanonical value s_ν/S for most macro-observers.*

To this end we recall that by "average over all macro-observers, considered to be equally likely" is meant an average over all bases $\{\omega_{\nu,j}\}$—favoring no basis over another. Now, given a basis $\{\omega_{\nu,j}^{(0)}\}$ of the energy shell, any other basis $\{\omega_{\nu,j}\}$ can be obtained by transforming the first with a unitary operator U. The average over all macro-observers, i.e., over all bases $\{\omega_{\nu,j}\}$, is therefore equivalent to an average over all U transformations.[9] As we have

$$(\Psi(t),U\omega_{\nu,j}^{(0)})=\sum_{k,k'=1}^{S}\tau_k^{(0)}(t)(\omega_k^{(0)},\omega_{k'}^{(0)})a_{k'}^{(\nu,j)*}$$
$$=\sum_{k=1}^{S}\tau_k^{(0)}(t)a_k^{(\nu,j)*}, \qquad (3)$$

the average $\mathfrak{M}$ over the macro-observers can be written as follows:

$$\mathfrak{M}[u_\nu(t)]=\sum_{j=1}^{s_\nu}\sum_{k,k'=1}^{S}\tau_k^{(0)}(t)\tau_{k'}^{(0)*}(t)\mathfrak{M}[a_k^{(\nu,j)*}a_{k'}^{(\nu,j)}]. \qquad (4)$$

If we split the numbers $a_k^{(\nu,j)}$ into their real and imaginary parts, $\alpha_k^{(\nu,j)}$ and $\beta_k^{(\nu,j)}$ respectively, and make use of the relation

$$\sum_{k=1}^{S}|a_k^{(\nu,j)}|^2=\sum_{k=1}^{S}[(\alpha_k^{(\nu,j)})^2+(\beta_k^{(\nu,j)})^2]=1, \qquad (5)$$

we can evaluate the expression $\mathfrak{M}[a_k^{(\nu,j)*}a_{k'}^{(\nu,j)}]$ by integrating over the spherical surface of radius 1 defined by Eq. (5). Thus we obtain

$$\mathfrak{M}[a_k^{(\nu,j)*}a_{k'}^{(\nu,j)}]=(1/S)\delta_{kk'}, \qquad (6)$$

from which it follows that

$$\mathfrak{M}[u_\nu(t)]=\sum_{j=1}^{s_\nu}\sum_{k=1}^{S}|\tau_k^{(0)}(t)|^2\frac{1}{S}=\frac{s_\nu}{S}. \qquad (7)$$

It should be noticed that this result does not require any hypothesis on the magnitude of the quantities s_ν.

In order to prove that the relation $u_\nu=s_\nu/S$ holds for the greatest part of the macro-observers, at any given time, it is sufficient to show that

$$\frac{\mathfrak{M}[(u_\nu-s_\nu/S)^2]}{s_\nu^2/S^2}=\frac{\mathfrak{M}(u_\nu^2)-s_\nu^2/S^2}{s_\nu^2/S^2}\ll 1. \qquad (8)$$

This might be done in a way similar to that used in obtaining Eq. (7). We prefer, however, to transform the average over the macro-observers into an average over all the vectors $U^{-1}\Psi(t)$, making use of the relation

$$(\Psi(t),U\omega_{\nu,j}^{(0)})=(U^{-1}\Psi(t),\omega_{\nu,j}^{(0)}). \qquad (9)$$

If we put

$$U^{-1}\Psi(t)=\sum_{k=1}^{S}\vartheta_k^{(0)}(t)\omega_k^{(0)}, \qquad (10)$$

we have

$$\mathfrak{M}(u_\nu^2)=\mathfrak{M}\left[\left(\sum_{j=1}^{s_\nu}|\vartheta_{\nu,j}^{(0)}|^2\right)^2\right]. \qquad (11)$$

[9] The Schur-Weyl method of group integration [see, e.g., F. D. Murnaghan, *The Theory of Group Representations* (The Johns Hopkins Press, Baltimore, 1938), Chap. VIII] gives a unique way to perform this average. In the following we shall use the equivalent geometrical technique developed by von Neumann.[5]

It is not difficult to show that[10]

$$\mathfrak{M}\left[\left(\sum_{j=1}^{s_\nu} |\vartheta_{\nu,j}^{(0)}|^2\right)^2\right]=\frac{s_\nu(s_\nu+1)}{S(S+1)}, \tag{12}$$

from which Eq. (8) follows, *provided that* $s_\nu \gg 1$ $(\nu=1,\cdots N)$.

We notice now that relations (7) and (8) remain obviously valid when averaged over time. But time-averaging M and averaging $\mathfrak{M}$ over the macro-observers are commuting operations when applied to $u_\nu^2(t)$.[11] It follows that

$$\mathfrak{M}M[u_\nu(t)]=\frac{s_\nu}{S};\quad \frac{\mathfrak{M}M[(u_\nu-s_\nu/S)^2]}{s_\nu^2/S^2}\ll 1. \tag{13}$$

These relations constitute von Neumann's ergodic theorem.

It is now evident that von Neumann's method of justifying the microcanonical ensemble is not acceptable, because his result is a mathematical consequence *only* of the averaging over the macro-observers. In other words, von Neumann's ergodic theorem holds independently of the time evolution of the system, i.e., it holds for any "trajectory" of the extreme of the state-vector $\Psi(t)$ in the energy shell, *even if this "trajectory" does not satisfy the Schrödinger equation.*

Lastly we remark that the inequality

$$\frac{\mathfrak{M}\{[M(u_\nu)-s_\nu/S]^2\}}{s_\nu^2/S^2}\ll 1 \tag{14}$$

is a consequence of the second of relations (13). In fact, owing to the Schwarz inequality, we have

$$M[(u_\nu-s_\nu/S)^2]\geqslant[M(u_\nu)-s_\nu/S]^2. \tag{15}$$

We have recalled formula (14) because it is a basic formula in the von Neumann-Fierz approach. Actually, inequality (14) is sufficient to prove the H-theorem, provided that we assume Fierz's definition of entropy instead of von Neumann's. But our procedure shows that also formula (14) holds independently of the "trajectory" of the extreme of the state-vector, being a pure consequence of the averaging over the macro-observers.

IV

If the extreme of $\Psi(t)$ filled densely and uniformly in its time evolution the *whole* surface K defined by Eq. (5), the following relations would hold:

$$M(u_\nu)=\frac{s_\nu}{S};\quad \frac{M[(u_\nu-s_\nu/S)^2]}{s_\nu^2/S^2}\ll 1,\quad (s_\nu\gg 1). \tag{16}$$

In this case, time-averaging would coincide with the averaging over the macro-observers, and consequently the latter would be quite superfluous.

It is well known, however, that in no case will the extreme of the state-vector show the above behavior. To prove this, it is sufficient to expand $\Psi(t)$ in the energy eigenvectors φ_n and to notice that the existence of the S "constants of motion" $|(\Psi(t),\varphi_n)|^2$ $(n=1,\cdots S)$ prevents the extreme of $\Psi(t)$ from filling densely the whole surface K.

The averaging over the macro-observers "corrects" the "wrong" time behavior of the vector $\Psi(t)$. This averaging is in fact equivalent to an average performed on all the vectors of the energy shell and is *by itself* sufficient to "justify" the use of the microcanonical ensemble. *Von Neumann's approach is therefore unable to put statistical mechanics on a purely mechanical basis. The quantum ergodic problem should be approached in an entirely new way.*

As a conclusion we might remark that the mathematical classical analog of the averaging over the macro-observers would be as follows. Let $(p(t),q(t))$ be a phase configuration at a time t and $f(p,q)$ any function of the variables p, q. Let us average $f(p,q)$ over all the transformations which map (p,q) into all the points of the microcanonical ensemble. Such a procedure is equivalent to an average over the microcanonical ensemble and leads therefore to the laws of statistical mechanics, but obviously it cannot give a proof of the ergodic theorem.

ACKNOWLEDGMENT

It is a pleasure for us to thank Professor M. Fierz and Professor P. T. Landsberg for their kind remarks.

[10] Equations (6) and (12) may be also derived with the methods of probability theory, as Professor Landsberg has kindly pointed out to us. We have preferred to stick to the geometrical treatment since this will make more intuitive the further considerations of the paper.

[11] In fact, $u_\nu^2(t)$ is given by the following expression:

$$u_\nu^2(t)=\sum_{j,j'=1}^{s_\nu}\sum_{h,k,h',k'=1}^{S}(\tau_h^{(0)}(t)\tau_k^{(0)*}(t)\tau_{h'}^{(0)}(t)\tau_{k'}^{(0)*}(t))\times[a_h^{(\nu,j)*}a_k^{(\nu,j)}a_{h'}^{(\nu,j)*}a_{k'}^{(\nu,j)}];$$

therefore, since M acts on $(\cdots)$ and $\mathfrak{M}$ on $[\cdots]$, the final result is the same if one performs first M and then $\mathfrak{M}$ or vice versa.