On the foundations of equilibrium statistical mechanics

On the Foundations of Equilibrium Statistical Mechanics

by **M. R. Schafroth*)**

School of Physics**) The University of Sydney, Sydney (Australia)

(10. V. 1959)

Zusammenfassung. Es wird betont, dass die Gesamtheiten der Quantenstatistik mit Hilfe idealisierter Experimente auf operationale und daher objektive Weise erzeugt werden können. Das Wärmegleichgewicht kann dann, wie in der Thermodynamik, als der Zustand definiert werden, aus dem keine Änderung mehr möglich ist. Der Ergodensatz gehört nach dieser Auffassung nicht in die statistische Mechanik der Gleichgewichtszustände; er erscheint als grundlegender Satz der statistischen Mechanik der Nicht-Gleichgewichtszustände.

1. Introduction

The question of the foundations of equilibrium statistical mechanics has recently been much discussed[1-3]. The question is as old as the subject itself, and the two founders of the subject, BOLTZMANN and GIBBS, are generally held responsible for the two viewpoints which have been taken:

(1) The approach named after BOLTZMANN is, briefly, the following: From the empirical fact that all macrosystems, when left alone, tend towards an equilibrium state, one defines the "thermal equilibrium" value of physical quantities as their time-average over long times ("long" hereby meaning "long compared to the relevant relaxation times").

Since such time-averages cannot, in practice, be worked out, one then relies on the validity of the ergodic theorem for the practical applications. This theorem states, in qualitative terms:

"Ergodic theorem". The time-average over sufficiently long times of quantity of interest for a physical system is equal to the average of the same quantity over the microcanonical ensemble.

It is well-known that this theorem has not been proved, and recent discussions[1] indicate that it cannot validly be proved. It appears from

---

these discussions that such a proof could only be based on the introduction of "random" observers, or of "random" perturbations of the system. It is then very obscure on what concept of probabilities this "randomness" is based.

(2) The approach ascribed to Gibbs is more radical: One *defines* thermal equilibrium by a suitable statistical ensemble, so that all quantities in equilibrium appear as averages taken over this ensemble. For the basic case of closed systems at a given energy, the appropriate ensemble is the microcanonical one, so that the practical results are the same as with the Boltzmann approach.

Here, then, a "randomness" concept is applied directly to the system under study: Thermal equilibrium is defined by "random" systems on the energy shell.

*   *
*

In both approaches, therefore, the difficulty appears to be the definition of "randomness", which seems to require the introduction of a-priori probabilities. Now a-priori probabilities are usually considered meaningless, or, at best, as a subjective concept with no place in science. (This view has been challenged recently by JAYNES[2]) who holds that a satisfactory foundation of statistical mechanics may be constructed, on the Gibbsian viewpoint, by the use of information theory, defining a-priori probabilities according to the knowledge of the observer).

The present paper is an attempt at showing instead that a-priori probabilities may be avoided entirely on the Gibbs view-point, so that the latter may be taken as an objective foundation for equilibrium statistical mechanics. This is based on the idea[4]) that there is one theory in which probabilities are objectively determined by laws of nature and not simply used, in a subjective sense, to cover up the ignorance of some observer: this is quantum theory. Quantum theory permits, as a thought experiment, the construction of an apparatus which produces, reproducibly, a given ensemble. Thus, ensembles are as well-defined as pure states, and, therefore, it becomes possible to define uniquely the equilibrium ensemble as the one which cannot be changed any more by operations which leave the rest of the world unchanged: the analogue of the thermodynamical definition of equilibrium. The equilibrium ensemble may also, in principle, be produced, in a reproducible way, by experimental manipulations and is thus objectively defined.

From this view-point, the Boltzmann approach is really concerned with a non-equilibrium problem, namely the approach to thermal equilibrium. It is not surprising that it leads to great difficulties, since non-equilibrium statistical mechanics has not, as yet, been properly formulated.

2. Probabilities in Physics

Probabilities have been used in theoretical physics in two essentially distinct ways:

(A) As a means of taking into account a lack of knowledge about some basically determined process. In this way, probabilities were first introduced in the kinetic theory of gases ("Unordnungsannahmen").

(B) As well-defined predictions which theory makes for given, experimental conditions. Such *objective* probabilities occur in quantum mechanics where the result of a well-defined experimental set up is predicted in terms of them.

The use of probabilities in the sense A ("Unordnungsannahmen") is not very satisfactory as a logical foundation of statistical mechanics, since the concept of probability or randomness in this sense depends on the observer. What appears random to one observer may exhibit regularities to another, differently equipped observer. To use an average (or "random") observer in order to avoid this, only shifts the problem, since now the randomness of observes cannot be defined objectively. Jaynes[2]) has argued that by introducing concepts of information theory, such subjective probabilities may be used as a satisfactory basis for statistical mechanics. Most authors, however, feel that a foundation of statistical mechanics which lacks the subjective element would be preferable.

The "operational" way (B), in which probabilities occur in quantum mechanics is entirely objective. The typical prediction of quantum mechanics is of the type:

"If a preparatory measurement M_1 (of a complete set of commuting observables Ξ) is performed on the system Σ, then Σ is in the pure state $|\xi>$ in which Ξ is diagonal and has the measured values ξ:

$$\Xi \mid \xi> = \xi \mid \xi> \tag{1}$$

The outcome of any other measurement M_2 of a quantity Z on Σ is then governed by the probability distribution

$$\varphi_\xi(\zeta) = \mid <\zeta \mid \xi> \mid^2 \tag{2}$$

over the eigenvectors $|\zeta>$ of Z."

This means that, if the preparatory measurement, M_1, is performed on a large number N of identical systems Σ and a selection made of the systems which yield the values ξ, (1), then the results of M_2 will be distributed according to (2) (with an accuracy which increases in a predictable fashion with N). This is a prediction as objective as one may have; the non-deterministic nature of quantum mechanics shows itself merely in the fact that accurate predictions can only be made for large

numbers of identical replicas of the system. The observer does not enter the picture, except for his classical function in choosing the experiments (M_1 and M_2) he wishes to perform.

For the transition to statistical mechanics, the consideration of pure states is not sufficient, one has to introduce "mixtures", i.e. incoherent superpositions of pure states. One might now feel that mixtures in quantum mechanics serve precisely the purpose A, of taking into account an incomplete knowledge of an intrinsically better-determined system (which "is really in one of the pure states"), and that, therefore, one is back at the difficulties of A.

However, notwithstanding any philosophical interpretations, mixtures, like pure states, may be defined in a purely objective manner, independent of the observer, by prescribing an idealized experiment, analogous to M_1 above, for constructing them.

Assume, for illustrative purpose, that we wish to construct, for the system Σ, a mixture P given by its weights p_ξ attached to the (complete and orthonormal) set of states $|\xi>$. First, by the classical procedure for pure states, one may by a measurement M_1' select the pure state

$$|p> = \sum_\xi \sqrt{p_\xi} \, |\xi> \tag{3}$$

By a second measurement M_1'', performed on $|p>$, of the complete set of observables Ξ, of which deliberately no record is taken, the system Σ is left in the mixture P; i.e. the subsequent measurement M_2 of the arbitrary quantity Z now yields the distribution

$$\varphi_P(\zeta) \equiv \sum_\xi p_\xi \, |<\zeta\,|\,\xi>|^2. \tag{4}$$

The statement, therefore, is: If the combined measurement $(M_1' + M_1'')$ is performed on a sytem Σ, the latter is left in the mixture P. If $(M_1' + M_1'')$ is performed on a large number N of identical systems Σ, there results the ensemble defined by the density operator

$$U = \sum_\xi |\xi> p_\xi <\xi|. \tag{5}$$

It is important to note that no corresponding operational procedure exists for creating an ensemble with given probability distribution in classical physics. *There is no provision for operationally defined, objective probabilities in a deterministic theory like classical physics.* For this reason, it appears impossible to found classical statistical mechanics on classical mechanics alone. Although the quantal structure of the system Σ may be of no relevance, one needs at least an auxiliary quantum mechanical system Σ' in order to introduce operationally well-defined probabilities

into the system $\Sigma : \Sigma'$ acts as a "quantum-mechanical dice*)". The simplest quantum-mechanical dice is, of course, the system Σ itself; after its quantum character has been used to define probabilities in an operational way, the transition to classical mechanics may be performed. In this way, densities in classical phase space (being the limiting case of quantum-mechanical density-matrices) may be thought of as operationally defined by the underlying quantum structure of the system.

3. The Equilibrium Mixture

After having given an operational and, therefore, objective definition of a mixture of states, together witht a hought-experiment capable of actually producing an ensemble of systems realizing this mixture, the particular mixture which corresponds to thermal equilibrium may be defined in the usual way.

The starting point is the observation that there exists a function $S(P)$, defined for all mixtures P, which may be used to decide whether one given mixture may be transformed into another one by measurements (in the widest sense used in section 2) or not. It has been proved by v. NEUMANN[5]) that the function

$$S(P) = - \text{Trace} \, (U_P \log U_P) \tag{6}$$

given in terms of the density operator U_P of the system Σ of the mixture P has the following basic properties:

(**1**) In the free mechanical development of the system, it remains constant (see sect. 4).

(**2**) When a measurement is performed on P, which transforms the mixture P into the mixture P', then S can thereby only increase:

$$S(P') \geqq S(P). \tag{7}$$

The particular mixture P_0 for which $S(P_0)$ is a maximum (under given external conditions) has, therefore, the fundamental property of thermal equilibrium, namely, that no manipulation, i.e. measurement, in the widest sense, which leaves the external conditions unchanged, can change P.

(This assumes that the maximum mixture P_0 is unique. If there were two mixtures P_0' and P_0'' with $S(P_0') = S(P_0'') > S(P)$ for all other P, then

---

*) An ordinary dice, being governed by classical laws, can only yield probabilities of type A, not operationally defined ones.

the mixture $\bar{P}_0$ given by the density matrix $U_{\bar{P}_0} = \frac{1}{2}(U_{P_0'} + U_{P_0''})$ would have even higher S: from (6) follows)

$$S(\bar{P}_0) > S(P_0') \tag{8}$$

which contradicts the assumption).

One is, therefore, justified in identifying the mixture P_0 given by the maximum of S, under given external conditions, with thermal equilibrium.

For a system Σ whose energy E is given within a margin Δ, the equilibrium mixture is the "microcanonical" one given by the density operator

$$U_0 = \delta_\Delta (\boldsymbol{H} - E) \tag{9}$$

where $\boldsymbol{H}$ is the hamiltonian of Σ, and

$$\delta_\Delta (x) = \begin{cases} 1 & \text{if } |x| \leqq \Delta \\ 0 & \text{if } |x| > \Delta \end{cases} \tag{10}$$

From the microcanonical ensemble, all of equilibrium statistical mechanics follows in a well-known manner.

That the special choice (6) for the function $S(P)$ does not introduce any arbitrariness can be seen by noticing that any mixture with fixed energy can, by a succession of suitable measurements, be transformed into the microcanonical one. It is sufficient to note that a mixture of two orthonormal states u_1 and u_2, with weigths p_1 and p_2, by an unrecorded measurement of the observables diagonal in

$$\left(\frac{u_1 + u_2}{\sqrt{2}} , \frac{u_1 - u_2}{\sqrt{2}} \right)$$

is reduced to a mixture with equal weights $p_1' = p_2' = (p_1 + p_2)/2$. Thus, the equilibrium definition is unique.

The identification of $S(P_0)$ with the thermodynamical entropy and the generalization of the entropy concept by (6) to cover also non-equilibrium situations is then straight-forward.

It should be noted that the definition of thermal equilibrium arrived at in this manner is the one actually used in thermodynamics, where also thermal equilibrium is defined as the state from which no further change is possible, and not as the state towards which the system evolves spontaneously.

The two definitions are by no means trivially identical: the theorem asserting their identity is the ergodic theorem which has not, as yet, been proved. It is perhaps worth remembering that systems are known which do not appear to tend spontaneously towards equilibrium, or at least only extremely slowly. The best-known example of such a system is a superconducting ring with its locked-in magnetic field.

4. The Ergodic Problem

From the present (Gibbsian) point of view, the ergodic theorem does not belong into the foundation of equilibrium statistical mechanics at all. It is rather a fundamental theorem on non-equilibrium processes and might be enunciated crudely as *"All physical systems evolve spontaneously towards thermal equilibrium"*, where thermal equilibrium is defined by the microcanonical ensemble.

The fact that this theorem is difficult not only to prove, but to discuss in its meaning, is but an indication of the fact that we have, as yet, not even an approximately satisfactory foundation for non-equilibrium statistical mechanics. Whereas (as outlined above) the quantum-mechanical mixture representing the thermal equilibrium state may be defined and constructed entirely without reference to the observer and his apparatus, it appears necessary to introduce him in an essential way for the study of non-equilibrium processes [5] [7]. This may be seen from the old observation that, if the system Σ, in a mixture U, is left to follow its quantum-mechanical course undisturbed, the entropy S is a constant:

$$
\begin{aligned}
S(t) &= - \, k \ \text{Trace} \, \{U(t) \log U(t)\} \\
&= - \, k \ \text{Trace} \, \{e^{iHt} \, U_0 \, \bar{e}^{iHt} \cdot e^{iHt} \log U_0 \, \bar{e}^{iHt}\} \\
&= - \, k \ \text{Trace} \, \{U_0 \log U_0\} = S(0).
\end{aligned}
\tag{9}
$$

Obviously, therefore, an increase in entropy can only arise from outside interference, i.e. by the actions of the observer and his apparatus. Furthermore, there are always idealized observers (those who observe quantities commuting with the given density matrix) whose measurements will not produce an approach towards equilibrium. It seems, therefore, unavoidable to introduce "random" observers, or other "randomization" processes into the concepts of non-equilibrium statistical mechanics. These then allow a proof of the ergodic theorem and its corollary, the H-theorem [1] [6].

What the significance of this randomization process in non-equilibrium statistical mechanics is, is as yet unclear. The purpose of this paper has been to point out that this unsolved difficulty in no way interferes with the much simpler problem of the foundation *of equilibrium* statistical mechanics.

References

[1]) M. FIERZ, Helv. Phys. Acta *28*, 705 (1955).

[2]) E. T. JAYNES, Phys. Rev. *106*, 620 (1957) and *108*, 171 (1957).

[3]) P. CORINALDESI and A. LOINGER, «Max-Planck-Festschrift» (in the press).

[4]) Ref. [1]), footnote p. 713.

[5]) v. NEUMANN, «Mathematische Grundlagen der Quantenmechanik» (Berlin 1932).

[6]) W. PAULI, «Probleme der Modernen Physik», p. 30 (Leipzig 1928).

[7]) W. PAULI and M. FIERZ, Z. Phys. *106*, 572 (1937).