

QUANTUM MECHANICS AND THERMODYNAMIC IRREVERSIBILITY*

By B. DAVYDOV

Physical Technical Institute, Academy of Sciences of the USSR

(Received October 10, 1946)

In the classical statistical mechanics the quasi-ergodic hypothesis enables one only to calculate the time averages for an isolated system. It is insufficient for determining the change in the state of a system during small time intervals and it does not lead, for instance, to the thermal conductivity equation. To obtain the laws of the variation with time the assumption of equal probabilities of all states compatible with the observed values of macroscopic quantities is necessary. The independence of the phases of gas molecules, leading to the usual collision law, is a special form of such an assumption. But being accepted for the instant t_1 the latter cannot be satisfied at other instants. A minimum at the time t_1 is thus obtained for the entropy instead of a monotonic increase. This is connected with the general contradiction between reversibility of classical mechanics and thermodynamic irreversibility.

The irreversibility of measurements in quantum mechanics leads to a permanent increase of entropy. Each measurement effects an averaging of phase differences between eigenfunctions corresponding to its different results. It enables one to prove the kinetic equation and the H -theorem, provided the measurements are repeated in short time intervals. At the instant of a measurement the entropy is not changed but each new measurement brings about the conditions of its subsequent increase.

Irreversibility of quantum mechanics is characterized by a "quantum entropy" $c = -\text{Spur} \rho \ln \rho$. It is not an observable quantity, thermodynamic entropy S being observable. It has been proved that $\Delta c / \Delta t \geq 0$ and $S \geq c$. Hence, if at the instant of a measurement $S = c$ then later measurements give $\Delta S / \Delta t \geq 0$ for any interval Δt between the measurements. The H -theorem is obtained under Pauli's assumption of equal probabilities of all states of any cell and of the independence of their phases. This assumption leads also to $S = c$. According to the quasi-ergodic hypothesis it is satisfied nearly always owing to the former interaction of the system with the surrounding bodies. In the course of time it continues to hold.

It is a widespread opinion that the complete statistical foundation of thermodynamics can be obtained already on the basis of the laws of classical mechanics. Even if there are some vagueness and some logical gaps in the existing proofs, it is believed that these gaps are of no importance and in some way or other they can be eliminated. This opinion is supported by a great many experimental facts confirming the laws of statistical mechanics.

In accordance with this standpoint, it is

* Presented to the Session of the Physical Mathematical Department of the Academy of Sciences of the USSR, on July 1, 1946.

quite natural to search for the foundation of quantum statistical mechanics along the same lines as for the classical one. It was repeatedly stated that quantum mechanics contributes nothing of essential import to this problem.

As we shall see, such an opinion cannot be justified theoretically. As a matter of fact, classical mechanics, owing to its perfect reversibility, does not enable one to obtain statistical irreversibility. Only in quantum mechanics there arise elements of irreversibility, the second law of thermodynamics being their macroscopic reflection.

1. Classical statistical mechanics

It is well known that all the attempts of the statistical foundation of thermodynamics based on classical mechanics have encountered serious difficulties (¹). The concept of probability itself is foreign to rational classical mechanics. Some postulates or hypotheses are, therefore, inevitable for its introduction. The fundamental Boltzmann H -theorem deals with a perfect gas whose molecules obey the laws of classical mechanics. The initial positions of gas molecules in the phase space are the only things at our disposal for the enunciation of a probability hypothesis. In order to prove the H -theorem the following assumption is necessary: at a certain instant t_1 the position of any molecule, *i. e.* its phase, is to be independent of the positions of all the other molecules. Then during a small interval of time t_1 to $t_1 + \Delta t$ there is realized, in the first approximation, the collision law leading to Boltzmann's kinetic equation and further to the H -theorem. According to the latter the entropy must increase during this time interval, provided the distribution function differs from the equilibrium one.

Classical mechanics is, however, quite reversible. Our probability assumption, *i. e.* the initial condition for the instant t_1 continues to hold if the velocities of all the molecules are replaced by their opposites without any alteration in the relative positions of the molecules. Therefore, such a replacement being performed, the result cannot be changed. The H -theorem requires accordingly at once that the entropy should decrease during the time interval $t_1 - \Delta t$ to t_1 . This does not, of course, contradict its increase after t_1 , one being an inevitable consequence of the other. If the assumption of the independence of the positions of gas molecules in the phase space were permanent in the course of time, we should be able to repeat the same argument for the instant $t_2 = t_1 + \Delta t$. Instead of the entropy increase during the interval t_1 to t_2 we should obtain its decrease. This shows that such an assumption cannot be maintained in time. Our consideration is an argument of contraries and it can easily be illustrated by simple examples. If at the instant t_1 the phases of all the molecules are independent, then after some time macroscopically distinguishable molecule groups will arise, their phases being not independent. Hence, the assumption of the inde-

pendence of the phases of gas molecules can be made for one instant t_1 only. Then at other instants of time it will, in general, be not satisfied.

The variation of entropy obtained above does not correspond at all to that prescribed by thermodynamics. Owing to the perfect reversibility of classical mechanics a minimum quite symmetrical with respect to the time is obtained at the point t_1 . Such a variation has nothing to do with the monotonic increase required by thermodynamics. In fact, we have obtained the reversible course of a thermal fluctuation about the most probable value of entropy $S_{\max} = \text{const}$, while thermodynamics deals with the irreversible variation of the most probable value, with neglect of spontaneous fluctuations. Let us suppose that at the instant t_1 some temperature difference is detected. Then this temperature difference must decrease in the past as well as in the future. Just this behaviour corresponds to the second law of thermodynamics. To our result, however, there corresponds the decrease of the temperature difference in the future and its increase in the past.

All instants of time are quite equivalent in classical mechanics. Hence, only such assumptions can have a physical meaning which can be satisfied at every instant of time. The assumption of the independence of the phases of gas molecules does not meet this requirement. Therefore, no real significance can be attributed to it.

The criticism directed against the H -theorem caused Boltzmann and Maxwell to pass on from the six-dimensional phase space for one molecule only to the many-dimensional phase space for the whole molecular system considered. In order to give the laws of statistical physics the character of a purely mechanical necessity, they put forward the ergodic hypothesis. The latter is no longer a condition concerning the distribution of initial values. It is an assumption referring to the structure of an isolated molecular system by itself, *i. e.* to the properties of the interaction forces.

In its initial form the ergodic hypothesis proved to be logically inconsistent. The consequent form of this hypothesis can be obtained as follows. According to Ehrenfest (²) the "coarse-grained density" in the phase space is to be introduced. It means that all the phase space must be divided up into sufficiently small but still finite cells. Instead

of a continuous probability distribution in the phase space we must consider the probabilities of our molecular system being in these phase cells. A layer in the phase space, for which the total energy of the system falls within given and also sufficiently narrow limits, is to be considered. Then the quasi-ergodic hypothesis states that the relative time of the system being in any specified phase cell of such a layer approaches the limit proportional to the cell extension, irrespective of the initial conditions, provided the time interval taken into account increases indefinitely.

The entropy of the system satisfying this quasi-ergodic hypothesis for the most part of any sufficiently long time interval must be near its greatest (microcanonical) value, under given external conditions and a given value of the total energy. It can deviate from this greatest value only as a result of fluctuations. In general, the macroscopically observable quantities for the system for the most part of time must lie near their equilibrium values, calculated by the averaging over the whole constant energy layer. In other words the quasi-ergodic hypothesis leads to the identification of the time averages with the phase averages. Then it is possible to speak about statistical equilibrium, temperature, etc.

This is, however, insufficient if, instead of averaging over long periods of time, we consider the behaviour of a molecular system in the course of time. Let at some instant t_1 the values of macroscopically observable quantities for our system be very different from the equilibrium ones. The irreversible macroscopic equations as, for instance, the thermal conductivity equation must follow from the statistical laws. Definite changes of the values of observable quantities in the course of time are prescribed by these equations beginning with sufficiently small time intervals. The latter are not to be regarded as infinitesimal like the phase cell extension in the definition of coarse-grained phase density. But they must certainly be considerably less than the relaxation time, *i. e.* the time necessary, on the average, to approach the state of thermal equilibrium.

In order to obtain the statistical laws for the variations of the state of a system during such small time intervals, we cannot use the averaging with respect to the time. A further hypothesis of the nature of the initial condi-

tion is here necessary. It can be enunciated as follows. At a given instant of time the system considered with equal probabilities may be in all the points of its quasi-ergodic trajectory compatible with the values of macroscopic quantities observed at this instant. In other words, the probability of the system being in any specified phase cell compatible with the observed values of macroscopic variables, is proportional to the cell extension.

Such an assumption of "equal probabilities of all permissible phase points" is absolutely necessary to obtain the statistical laws of the variations of macroscopic quantities in the course of time. If the value of the entropy of a molecular system is abnormally low at a given time instant then at the next one it may be higher or lower. Equal probabilities of all permissible phases make it only possible to consider the increase of entropy as its probable variation. The assumption of the independence of the phases of gas molecules on which the proof of the H -theorem is based is a special form of the assumption of equal probabilities of all permissible phases. The assumption of equal probabilities of all points of a certain small extension in the many-dimensional phase space is another form of it.

The principal difficulty of classical statistical mechanics met with in the discussion of the H -theorem is here encountered in a more general form. If we assume, as has just been said, that at a certain instant t_1 all permissible points of the phase trajectory are equally probable then in the future, at $t > t_1$, the entropy is almost certain to increase. This result is proved though for a perfect gas only; but one can believe the behaviour of more complicated molecular systems to be similar, under sufficiently general assumptions. Our initial condition for the instant t_1 is but not altered by the replacement of all the molecular velocities by their opposites. Hence in the past at $t < t_1$ the entropy should decrease owing to the reversibility of classical mechanics, *i. e.* there is almost certain to be a minimum at the point t_1 . Further, if all permissible phase points were assumed to be equally probable at the instant $t_2 = t_1 + \Delta t$ then the entropy, instead of increasing, should decrease during the interval from t_1 to t_2 . Hence, this assumption cannot be satisfied at the instant t_2 , being accepted for the instant t_1 .

All the contradictions which we have obtained with respect to the H -theorem result thus

here word for word. The assumption of equal probabilities of all permissible phases also cannot be preserved in the course of time. If it is satisfied at the instant t_1 then it must be violated under the influence of the laws of classical mechanics at other instants of time which, as a matter of fact, differ by nothing. Moreover, the irreversibility does not at all follow from this assumption. It gives a minimum on the reversible fluctuation curve only. It is impossible to assume all permissible phases as equally probable for all times; such an assumption contradicts the laws of classical mechanics.

The question is thus not at all whether some additional assumptions are necessary for the classical foundation of thermodynamics. As a matter of fact, there exist no assumptions which could remove the obvious contradiction between reversibility of classical mechanics and irreversibility of thermodynamics.

2. Irreversibility of quantum mechanics

The concept of probability is intimately connected with the fundamental notions of quantum mechanics. Therefore, it is natural that with its arrival there arises a hope to get rid of any probability assumption in the statistical foundation of thermodynamics. Indeed, a few papers have appeared at different times in which the fundamental laws of quantum statistical mechanics are obtained unconstrainedly and without any additional assumption. However, at nearer examination all these derivations have proved to be unconvincing. As shown by Schrödinger⁽²⁾ and Pauli⁽³⁾ the probability assumptions are necessary here as before. They appear in the form of the assumption of equal probabilities of different eigenfunctions and of independence of their phases.

The question, we are particularly interested in, is the question of irreversibility. Quantum mechanics is no longer entirely reversible, namely, the measurement process in it is irreversible.

The principal problem of quantum mechanics is to calculate the probability of any possible result of measurement at some time instant t_2 , provided the results of the preceding measurement at the instant t_1 are known. For this purpose the wave function $\psi(t_1)$ for the instant t_1 or, more exactly, for the instant

$t_1 + 0$ immediately following the instant t_1 is to be determined, at first, according to the result of the measurement referring to the instant t_1 , the measurement being treated as instantaneous. Schrödinger's equation must then be solved with this $\psi(t_1)$ as initial condition. The wave function $\psi(t_2 - 0)$ will thus be found for the instant of time directly preceding the second measurement. The transition from $\psi(t_1 + 0)$ to $\psi(t_2 - 0)$ is reversible. If the wave function $\psi = \psi^*(t_2 - 0)$ has been given for the instant $t_1 + 0$ we should obtain $\psi = \psi^*(t_1 + 0)$ for the instant $t_2 - 0$. The replacement of the wave function by its conjugate complex corresponds here to the replacement of all momenta by their opposites.

Further our wave function must be expanded in a series in terms of eigenfunctions φ_k of the operator corresponding to the physical quantity observed at the instant t_2 :

$$\psi(t_2 - 0) = \sum_k \xi_k \varphi_k. \quad (1)$$

The desired probability of any specified value of k will then be

$$p_k = |\xi_k|^2. \quad (2)$$

In the process of measurement the functions φ_k are in general changed into some other functions φ'_k . The transition from the function $\psi(t_2 - 0)$ to a certain function φ'_k is here irreversible. If we do not know what the result of the second measurement is, *i. e.* if we do not know what a function φ'_k has been resulted in, then we can still speak about the irreversible transformation of the pure state (1) with the density matrix

$$\rho_{kk'}(t_2 - 0) = \xi_k \xi_{k'}^* \quad (3)$$

into the mixed state

$$\rho_{kk'}(t_2 + 0) = |\xi_k|^2 \delta_{kk'}. \quad (4)$$

The first density matrix being written down in the φ_k -representation while the second one in the $\varphi_{k'}$ -representation.

Let us now consider the inverse problem. Suppose we are given the result of the measurement at the instant t_2 . Are we then in a position to find the probability of any possible, now unknown, result of the measurement at the former instant t_1 ?

We can, certainly, apply the same scheme of calculations in the inverse direction; it does not lead, however, to the right result. Indeed, if the result of the measurement at the instant t_2 is known then we must take the corresponding eigenfunction φ_k as the initial or, rather, as the "final" condition for solving Schrödinger's equation in the direction of the past. If we have the results of statistical observations for the instant t_2 , the density matrix $\rho_{kk'}$ for the mixed state, according to (4) must be taken as a final condition, being supposed in the form (4) in the φ_k -representation. But we thus get by no means the pure state $\psi(t_1)$ corresponding to the values of physical quantities actually observed at the instant t_1 . In order to obtain the wave function $\psi(t_1)$ we should proceed not from the state (4) but from the pure state (3).

To know the state (3) we must know the phase differences between different ξ_k in the expansion (1), *i. e.* we must also know the non-diagonal elements of the density matrix (3). Such an information cannot be obtained from the measurements of the quantities k at the instant t_2 . On the basis of these measurements we can proceed as from a "final" condition from the mixed state (4) only. Solving the equation of motion for the density matrix in the direction of the past, we obtain a mixed state for the instant $t_1 + 0$ as well. It will give the predictions for the measurement at the instant t_1 which will be in contradiction with the true result. We shall discover a discrepancy between our predictions and experiment after this result has been ascertained. Hence, the application of the scheme of calculations of quantum mechanics in the inverse direction is illegitimate. It leads to a disagreement with the experiment.

On the other hand, classical mechanics is quite reversible and it can be applied to searching for the behaviour of the system considered in the past as well as in the future with equal success. Classical mechanics, of course, is an approximate, limiting form of quantum mechanics. There arises the question how it is possible to apply classical mechanics in the direction of the past while this is impossible for the quantum theory, the latter being the exact form of mechanics.

Classical mechanics neglects in principle the disturbance of the state of the system observed, accompanying the process of measurement. Under this condition only we can speak

of the continuous trajectory. This means, we must neglect the difference between the states (3) and (4) in the classical approximation. But the state (3) is just the "final" state leading to the right predictions for the past, the state (4) being known from the measurements of the latter instant t_2 . Hence the predictions of the past are only in so far possible as it is permissible to neglect the irreversible quantum disturbance of the system during the process of measurement.

The predictions of the past are, therefore, always quite determinate, for the future quantum, *i. e.* statistical predictions being also possible. This fully agrees with our psychology. It is customary to speak about the probabilities of future events, but we never speak like that about the past. The past may be known or not, but it is always considered as being quite definite.

3. The kinetic equation and the H -theorem

It is natural to connect the irreversibility of thermodynamics with that of quantum mechanics. We shall proceed from Pauli's generalized form of the H -theorem⁽³⁾.

A macroscopic molecular system, *i. e.* the system with a sufficiently dense spectrum, is to be considered. Some approximately stationary states $\varphi_{k\alpha}$ of this system must be taken as a set of orthogonal functions. Certain values of the energy can be ascribed to these states, they will be diagonal elements of the complete energy matrix in the corresponding representation. The non-diagonal elements cause then transitions between our approximately stationary states. We assume that these transitions can be treated using perturbation theory. This implies certain restrictions on the choice of the orthogonal functions.

Let us select from our approximately stationary states those for which the energy falls within given sufficiently narrow limits. It is our layer of constant energy, in its turn being divided up into groups of states each of which corresponds to a phase cell of classical statistical mechanics. Let k, l, \dots be the suffices of these cells, α, β, \dots being suffices of states in any cell. The number of states in any cell will be g_k , being different for different cells. For the macroscopic measurements some cells are always considerably larger than others. In any case there must be $g_k \gg 1$ for all k .

Suppose the result of a macroscopic measurement at some instant t_1 is known. It does not mean that the state of our system at this instant is known exactly, *i. e.* its wave function can be found definitely. The macroscopic measurements are never complete, never give the values of all the variables whose simultaneous measurement is possible in principle. As a matter of fact, a small part of them is determined by the measurements. We suppose that the suffix of the cell k only is determined, the state within the cell remaining uncertain. There is no reason to assume the system to be after the measurement in a pure state, permitting the description by a wave function. The system will, in general, be in a mixed state. To describe such a state the density matrix (v. Neumann's statistical operator) must be used. All the calculations become then quite cumbersome. There is the second approximation with the density matrix which corresponds to the first approximation of the usual method of variation of parameters. We shall, therefore, suppose for the sake of convenience that the density matrix is transformed to its principal axes. The state of our system will be described as a superposition of non-coherent wave functions corresponding to these principal axes (eigenfunctions of the statistical operator).

Let us consider one of these functions. We can expand it in a series in terms of the approximately stationary functions $\varphi_{k\alpha}$. Let

$$\psi(t_1+0) = \sum_{\alpha} \xi_{k\alpha} \varphi_{k\alpha}. \quad (5)$$

The value of k corresponds here to the result of observation.

In classical mechanics we might assume that the observation is continuous, but in quantum mechanics measurements always are discrete. We still suppose that they are made frequently enough and during the interval between two successive observations the state of the system, *i. e.* our $\xi_{k\alpha}$, vary but slightly. In order to predict the result of the measurement at the instant t , we can then confine ourselves to the first approximation of perturbation theory.

We have the following equation for the variation of $\xi_{k\alpha}$ in the course of time,

$$\frac{\hbar}{i} \frac{d\xi_{k\alpha}}{dt} = \sum_{k,\alpha} E_{kl} \exp \left[\frac{i}{\hbar} (E_{k\alpha} - E_{l\beta}) t \right] \xi_{k\alpha}. \quad (6)$$

Here $E_{k\alpha}$ are the diagonal elements of the energy matrix, E_{kl} being the non-diagonal elements; the dependence of the latter on α and β is neglected.

The form of our functions $\psi(t_1+0)$ cannot be ascertained by the macroscopic measurements referring to the instant t_1 . The value of k only is definite, being the same for any of these functions. The coefficients $\xi_{k\alpha}$ remain indefinite, having only to satisfy the condition:

$$\sum_{\alpha} |\xi_{k\alpha}|^2 = 1 \text{ at } t = t_1. \quad (7)$$

Some assumption concerning the distribution of the values $\xi_{k\alpha}(t_1)$ is further necessary. According to Pauli⁽⁸⁾ we must assume that all $\xi_{k\alpha}$ are equal with respect to their moduli their phases being independent:

$$\overline{\xi_{k\alpha} \xi_{k\beta}^*} = \frac{\delta_{\alpha\beta}}{g_k}. \quad (8)$$

The bar denotes here the averaging over all the non-coherent wave functions. We thus suppose statistical equilibrium within the given phase cell. Owing to equation (6) this assumption is to be satisfied throughout all the time for every k being accepted for one instant only.

According to (5) only $\xi_{l\beta}$ for $l=k$ are different from zero at the instant t_1 . This condition being preserved on the right-hand side of equation (6), in the first approximation we get:

$$\begin{aligned} \xi_{l\beta}(t) &= \\ &= \sum_{\alpha} E_{kl} \frac{\exp \left[\frac{i}{\hbar} (E_{k\alpha} - E_{l\beta})(t - t_1) \right] - 1}{E_{k\alpha} - E_{l\beta}} \xi_{k\alpha}(t_1). \end{aligned} \quad (9)$$

It gives in virtue of the assumption (8):

$$\begin{aligned} \overline{|\xi_{l\beta}(t)|^2} &= \\ &= \sum_{\alpha} |E_{kl}|^2 \frac{4 \sin^2 \left[\frac{1}{2\hbar} (E_{k\alpha} - E_{l\beta})(t - t_1) \right]}{g_k (E_{k\alpha} - E_{l\beta})^2}. \end{aligned} \quad (10)$$

The summation here can be replaced by an integration in the usual way. We obtain in the probability of the system being within the cell l , if it is within the cell k at the instant t_1 :

$$w_{kl} = \sum_{\beta} \overline{|\xi_{l\beta}(t)|^2} = \frac{2\pi}{\hbar} \frac{g_l}{\Delta E} |E_{kl}|^2 (t - t_1). \quad (11)$$

Besides it must, of course, be $\sum_l w_{kl} = 1$.

Hence

$$w_{kk} = 1 - \sum_{k \neq l} w_{kl}. \quad (12)$$

We have obtained formulae (11) and (12) provided the result of the measurement at the instant t_1 is known. If it is not so, there can still be considered the probabilities of its different results. Let them be

$$p_k = g_k f_k(t) = \sum_{\alpha} |\xi_{k\alpha}|^2. \quad (13)$$

In this case the assumption (8) will be re-written:

$$\overline{\xi_{k\alpha} \xi_{k\beta}^*} = f_k(t) \delta_{\alpha\beta}. \quad (8')$$

Each measurement gives moreover a definite value of k . Hence, the states with different k are incapable of interfering directly after the measurement: they refer then to different principal axes of the density matrix. Therefore, at $t = t_1 + 0$

$$\overline{\xi_{k\alpha} \xi_{l\beta}^*} = 0 \text{ for } k \neq l. \quad (14)$$

Owing to this condition the probabilities will simply be added together, and according to (11) and (12) we obtain the prediction for the result of the next measurement at the instant t :

$$g_l [f_l(t) - f_l(t_1)] = \sum_k g_k w_{kl} [f_k(t_1) - f_l(t_1)]. \quad (15)$$

Finally, considering the interval $t - t_1$ as infinitesimal, we get a differential equation for the distribution function $f_l(t)$ — the kinetic equation:

$$g_l \frac{df_l}{dt} = \sum_k a_{kl} [f_k(t) - f_l(t)]. \quad (16)$$

Here

$$a_{kl} = a_{lk} = \frac{g_k w_{kl}}{t - t_1} = \frac{2\pi}{h} \frac{g_k g_l}{\Delta E} |E_{kl}|^2. \quad (17)$$

The second law of thermodynamics can now be derived. Let entropy be introduced:

$$S = - \sum_k g_k f_k \ln f_k = \sum_k p_k (\ln g_k - \ln p_k). \quad (18)$$

This expression coincides practically with the usual definition of entropy as the logarithm of the phase extension. Indeed, in the macroscopic measurements the probabilities always give a sharp maximum at a certain k . The terms in the sum (18) relating to the values of k which differ considerably from this most probable k are vanishingly small and can be neglected. In the terms, in which p_k is not too small, $\ln p_k$ can be neglected as compared with $\ln g_k$. Hence

$$S \cong \sum_k p_k \ln g_k = \overline{\ln g_k}.$$

Owing to the sharpness of the maximum of p_k it differs very little from the logarithm of the most probable g_k , the latter being, as a matter of fact, an observable quantity.

According to equation (16)

$$\frac{dS}{dt} = \sum_{k,l} a_{kl} (f_k - f_l) \ln f_k.$$

It gives after the usual symmetrization:

$$\frac{dS}{dt} = \frac{1}{2} \sum_{k,l} a_{kl} (f_k - f_l) (\ln f_k - \ln f_l), \quad (19)$$

whence it follows:

$$\frac{dS}{dt} \geq 0. \quad (20)$$

The equality sign appears only if all f_k are equal.

The law of increase of entropy is obtained here as a consequence of two assumptions. The first of them is Pauli's condition (8'), referring to the states of any phase cell separately. It is the probability hypothesis without which statistical mechanics cannot be founded. Being accepted for some instant t_1 it will be satisfied throughout all time for any value of k . This can be easily verified if the values of $\xi_{k\alpha}$ and $\xi_{k\beta}^*$ according to (9) are put into (8) and the summation is then replaced by integration.

The second assumption is the condition (14) prescribing the absence of interference between the states referring to different phase cells. Under this condition only equation (15) is obtained from (11) and (12). This assumption is, however, no longer arbitrary. Each measurement gives a definite value of k and the states with different k cannot interfere directly after a measurement. In the course of time the condition (14) is violated under the influence

of equation (6), but the repeated measurements bring it every time about, the conditions of the validity of the kinetic equation during the subsequent time interval being realized again.

In the same way in which we have obtained the increase of entropy after the measurement at the instant t_1 , we could find its variation before t_1 or, for instance, before the next measurement at the instant t_2 . If we proceeded from the assumptions (8') and (14) again, as from the initial or, rather, the final conditions we should obtain a decrease of entropy during the time interval t_1 to t_2 , instead of its increase: $dS/dt \leq 0$, in disagreement with (20). This follows from the reversibility of Schrödinger's equation. But, as a matter of fact, the condition (14) is not satisfied immediately before the measurement. It is satisfied only directly after each discrete measurement owing to the irreversible quantum transition, accompanying it. We cannot, therefore, make use of the condition (14) as a final condition in the calculation of the entropy variation before t_2 . It would be an illegitimate application of quantum mechanics to obtain the probabilities of the results of previous measurements by means of subsequent ones as mentioned in § 2. In order to obtain the correct results, we should proceed not from the functions $\psi(t_2+0)$ but from the functions $\psi(t_2-0)$ as from the final conditions, the form of these functions being not determined by the results of the measurement at the instant t_2 . The difference between the functions $\psi(t_2-0)$ and $\psi(t_2+0)$ corresponds to the difference between the density matrices (3) and (4).

Likewise, to find the variation of the entropy before the instant t_1 we must proceed, as from the initial condition, not from the results of measurement at the instant t_1 , but from the preceding measurement. Hence a monotonic increase of entropy actually results here, while classical statistical mechanics gives a minimum at the instant t_1 . This increase appears as a peculiar quantum effect. It should be noted that at the instant of measurement the entropy is not altered. The measurement merely realizes the conditions of its increase, the equation (14) being every time confirmed.

4. Premises of the second law of thermodynamics

As we have shown, the entropy of a molecular system can only increase, the macroscopic

observations being repeated. Our proof rests upon the kinetic equation. In its derivation we have assumed that the time intervals between the successive measurements are sufficiently small and hence we can confine ourselves to the first approximation of perturbation theory. We shall get rid of this restriction.

Irreversibility of quantum mechanics is characterized by a quantity suggesting entropy*:

$$\sigma = -\text{Spur } \rho \ln \rho, \quad (21)$$

ρ being the density matrix. It can be called quantum entropy.

The variation of quantum matrices in the course of time under the influence of Schrödinger's equation is a contact transformation. Spur being a contact invariant. Hence the quantum entropy for an isolated system does not vary with time. It increases, however, discontinuously at every measurement. Indeed, let the elements of the density matrix before the measurement be

$$\rho_{k\alpha, k\alpha'} = p_{k\alpha} \delta_{\alpha\alpha'}. \quad (3')$$

The matrix is supposed here to have been transformed to the diagonal form with respect to the internal variables α . Let us make it diagonal also with respect to the observable variables k :

$$\rho_{\lambda\lambda'} = p_{\lambda} \delta_{\lambda\lambda'}. \quad (22)$$

After the measurement the density matrix will be diagonal in the former k, α -representation:

$$\rho_{k\alpha, k'\alpha'} = p_{k\alpha} \delta_{kk'} \delta_{\alpha\alpha'}. \quad (4')$$

Before the measurement the quantum entropy

$$\sigma_1 = - \sum_{\lambda} p_{\lambda} \ln p_{\lambda}, \quad (23)$$

* In the paper by O. Klein (4) this quantity is identified with thermodynamic entropy. Any additional hypothesis in the foundation of the second law of thermodynamics can then be avoided. But this is only possible if complete measurements are performed every time, *i. e.* all the quantities are observed whose simultaneous measurement is permitted by quantum mechanics. The incompleteness of macroscopic measurements is, however, a characteristic feature of statistical mechanics. There would be no problem if complete measurements were carried out.

after the measurement

$$\sigma_2 = - \sum_{k, \alpha} p_{k\alpha} \ln p_{k\alpha}. \quad (24)$$

Following Gibbs we shall introduce the quantities [cf. O. Klein (4)]:

$$G_{k\alpha, \lambda} = p_{\lambda} (\ln p_{\lambda} - \ln p_{k\alpha}) - p_{\lambda} + p_{k\alpha}; \quad (25)$$

$G_{k\alpha, \lambda} \geq 0$ for $p_{k\alpha}$ and $p_{\lambda} \geq 0$.

Let $U_{k\alpha, \lambda}$ be the reciprocal of the unitary matrix transforming the density matrix $\rho_{k\alpha, k'\alpha'}$ to the diagonal form. We have:

$$\left. \begin{aligned} p_{k\alpha} &= \sum_{\lambda} |U_{k\alpha, \lambda}|^2 p_{\lambda}; \\ \sum_{k, \alpha} |U_{k\alpha, \lambda}|^2 &= 1; \\ \sum_{\lambda} |U_{k\alpha, \lambda}|^2 &= 1. \end{aligned} \right\} \quad (26)$$

Multiplying $G_{k\alpha, \lambda}$ by $|U_{k\alpha, \lambda}|^2$ and summing up with respect to all suffices we obtain the desired result:

$$\sum_{k, \alpha} \sum_{\lambda} |U_{k\alpha, \lambda}|^2 G_{k\alpha, \lambda} = \sigma_2 - \sigma_1 \geq 0. \quad (27)$$

We have, in effect, proved that to the density matrix with vanishing non-diagonal elements there corresponds the greatest value of quantum entropy among the matrices with given diagonal elements. The non-diagonal elements are just cancelled by the measurement, the diagonal ones being unaltered.

Thermodynamic entropy according to (13) and (18) is

$$S = \sum_k p_k (\ln g_k - \ln p_k), \quad (28)$$

where $p_k = \sum_{\alpha} p_{k\alpha}$.

Let us compare S with σ . For this purpose we introduce similarly to (25):

$$\begin{aligned} G_{k\alpha} &= p_{k\alpha} (\ln p_{k\alpha} - \ln p_k + \ln g_k) - \\ &- p_{k\alpha} + \frac{p_k}{g_k} \geq 0. \end{aligned} \quad (29)$$

Summing up with respect to k and α we find that after the measurement $S \geq \sigma_2$ and, therefore, also before the measurement $S \geq \sigma_1$.

Hence thermodynamic entropy in general cannot be less than the quantum one:

$$S \geq \sigma, \quad (30)$$

An exhaustive answer to the question about the premises of the second law of thermodynamics can be given now. Under Pauli's condition (8') $p_k = g_k p_{k\alpha}$ and after the measurement $S = \sigma_2$. At subsequent measurements σ can only increase. Hence, the condition (8') being satisfied at the instant of a measurement, the following measurements can only give higher values of the thermodynamic entropy. As a matter of fact, the quite general law of quantum mechanics is the increase of quantum entropy. Quantum entropy, is, however, not an observable quantity. On the ground of macroscopic measurements one cannot even say whether the system is in a pure state, for which $\sigma = 0$, or in a mixed one. Thermodynamic entropy only is an observable quantity. In those cases when it coincides, with a sufficient degree of accuracy, with quantum entropy, we obtain the second law of thermodynamics. Particularly, S and σ coincide if the condition (8') is satisfied.

We have been so far dealing with the results of observations on a molecular system. An absolutely isolated system must moreover be investigated theoretically, being the system which is not subjected to observations. The general properties of solutions of the wave equation for such a system are of a great importance.

The works of v. Neumann (5) and of Pauli and Fierz (6) deal with the problems of this kind. The deviation of the time average of the entropy of an isolated system from the greatest possible value of the entropy is evaluated there; the necessary and sufficient conditions of its being small have been found. If these conditions are satisfied and the total energy of the system is given then the state of the system for the greater part of time will be close to the microcanonical one. The v. Neumann conditions are the inequalities which must be satisfied for every stationary state of the system considered.

v. Neumann and Pauli and Fierz have given the statistical foundation of these conditions. They introduce the statistics for the operators corresponding to the macroscopic measurements, *i. e.* measurements for which only some averaged quantities are accessible. The v. Neumann inequalities are thus proved to be satisfied for almost all the operators,

and the probability to encounter an operator not satisfying these conditions is quite small, provided the number of degrees of freedom of the system considered is very great.

Such a foundation seems to be not conclusive. Nothing proves that the operators for the real measurements are the most probable in v. Neumann's sense. Generally it is a proof replacing one hypothesis by another more complicated and, in this particular case, less founded. The inequalities of v. Neumann should be interpreted rather as the definition of a macroscopic measurement.

The validity of v. Neumann's conditions, however, is hard to verify. For this purpose all the stationary states of the molecular system considered ought to be found. It is, therefore, reasonable to accept immediately the quasi-ergodic hypothesis, in the form desired for further argument as the probability assumption, also in quantum statistical mechanics. It says that the time average of the probability of the system, being in the states of any specified phase cell of the constant energy layer, is proportional to the number of these states, irrespective of the initial conditions:

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T p_k dt = \frac{g_k}{\sum_i g_i}. \quad (31)$$

The summation is extended over the whole constant energy layer, provided the total energy of the system is given.

This assumption being accepted, the molecular system isolated for a long time must, first of all, be on the average in the state of statistical equilibrium. This statement exhausts all that is accessible for classical statistical mechanics. It does not contain, however, any elements of irreversibility. If we take into consideration two time instants, then at the latter instant the value of entropy may be lower or higher than at the former one with equal success even if these values are abnormally low.

On the contrary, the law of irreversible increase of entropy appears only in quantum statistical mechanics. The monotonic increase is here a consequence of the irreversibility of quantum measurements. The increase of entropy S results from the increase of quantum entropy σ , the latter varying only at the instants of measurement.

To detect the change of entropy two measurements are necessary. These measurements themselves realize the conditions of the increase

of entropy and the measurement referring to the latter time instant must give a higher value of the entropy than that referring to the earlier one. There is no need of direct measurements precisely at the instants referred to. We often speak about the behaviour of the system not observed directly for the sake of convenience only. Actually the measurements connected indirectly with the system and belonging even to a different time may be involved and this is sufficient for the appearance of irreversibility.

We have obtained the increase of entropy when proceeding from Pauli's assumption (8'). Let us try to make clear its physical meaning. An important case of the increase of entropy is the following one. Suppose the external restrictions upon our system prescribe specified values of its macroscopic variables, *i. e.* the value of k before the instant t_1 . This value must thus be regarded as an external parameter. In a more general way the external conditions may be supposed to prescribe the probability distribution of different values of k , *i. e.* the form of the distribution function f_k .

Under these external restrictions our system must be in a free interaction with the surrounding bodies for a long time. The total system consisting of our system and the surrounding bodies can be treated as isolated. The internal variables α of our system can then be considered as the macroscopic variables of the total system, its internal variables being the quantities characterizing the state of the surrounding bodies. The entropy S for such a composite system depends only on the distribution of values of the variables α . According to the quasi-ergodic hypothesis it must be close to its greatest value nearly always, the distribution of values of k being given. At the same time our system must be in the state of statistical equilibrium with respect to the variables α . It can deviate from this state from time to time only as a result of fluctuations.

We can suppose that our observable system is a small part of the total system. If the thickness of the constant energy layer ΔE is considerably less than the modulus of the canonical distribution then all the values of α within the layer can be regarded as equally probable. Hence the condition (8') must be satisfied nearly all times. With neglect of the fluctuations we can assume that at the instant t_1 it is satisfied as well.

If now at the instant t_1 the external restric-

tions are removed then the macroscopic variables k can further vary freely. The former state of conditional statistical equilibrium will now be not a steady one and the following measurements will detect an increase of the entropy depending on the distribution of k 's, as has been shown above. The assumption (8') will then be permanent in time.

As follows from our consideration the concept of probability in statistical mechanics is of dual origin. On the one hand, it is connected with the statistical nature of the laws of quantum mechanics and with the random character of the results of quantum measurements, leading to the increase of quantum entropy σ .

On the other hand, to establish the connection of observable entropy S with σ we had to turn to the quasi-ergodic hypothesis. The transition from the averaging with respect to time to the assumption (8') for a specified instant t_1 implies the assumption of equal probabilities of all time instants. It is the second source of the concept of statistical probability.

It is possible that the future theory will allow us to get rid of any additional hypothesis in the foundation of statistical mechanics and will then eliminate the duality of the concept of probability.

REFERENCES

- 1 Cf. P. a. T. Ehrenfest, *Encycl. d. math. Wiss.*, IV, 32.
2 E. Schrödinger, *Ann. d. Phys.*, **83**, 956 (1927).
3 W. Pauli, *Probleme d. modernen Physik* (Sommerfeld-Festschrift), 30 1928.
4 O. Klein, *ZS. f. Phys.*, **72**, 767 (1931).
5 J. v. Neumann, *ZS. f. Phys.*, **57**, 30 (1929).
6 W. Pauli u. M. Fierz, *ZS. f. Phys.*, **106**, 572 (1937).