STATISTICAL MECHANICS AND ITS APPLICATIONS TO PHYSICS

VICTOR F. LENZEN

UNIVERSITY OF CALIFORNIA, BERKELEY

I

STATISTICAL MECHANICS as a field of theoretical physics owes its origin and development to theories of atomism. I use the term "atom" in a generalized sense to include molecules, atoms, electrons, quanta of energy. An atomic theory is one that resolves physical reality into discrete constituents, or particles. The most striking characteristic of contemporary physics is the general acceptance of the hypothesis that physical systems consist of particles, at least from a significant point of view. The idea goes back to Democritus and similarly minded Greek philosophers, who invented the concept of the atom in order to reconcile the rational demand for unity and permanence in reality with the mutability of observable phenomena. The atomic theory, however, did not make significant quantitative progress until the nineteenth century. Then Dalton proposed his atomic theory for chemical systems. Subsequently Clausius, Maxwell, and Boltzmann developed the kinetic theory of gases. Later, it was found that electrical theory required atoms of electric charge. Since the beginning of the present century atomistic conceptions have found their way into the theory of energy, and have become the instruments of a quantum theory which is based on discontinuity in natural processes. The physiochemical systems which contemporary science investigates are conceived to be composed of many particles, corpuscles, atoms.

The problem posited by atomism arises in the following way: The classical physicist investigated and described physical phenomena in terms of concepts which expressed quantities that were relatively directly determined by experiments. If, for example, he wished to describe the observable properties of a gaseous system he employed the concept of volume, which is measurable by operations performed with a standard rod; the concept of temperature, which is measurable by a thermometer; the concept of pressure, which is measurable by a manometer. Between these measurable quantities the physicist sought functional relations which would involve empirical constants. There was a period in the nineteenth century when an influential school of thought taught that physics should not go beyond the establishment of generalizations concerning the observable properties of physical systems. To this school of energetics, classical thermodynamics was the ideal of physical theory. The course of development, however, has been guided by physicists who were not content with such a restricted goal. These physicists sought a deeper, more detailed description and explanation of physical phenomena. They have explained large-scale, observable processes in terms of fine-scale processes which are beyond the range of ordinary perception. Thus a kinetic-molecular theory of gases was created, according to which the large-scale properties of gaseous systems are determined by the mass action of a large number of molecules. The pressure of a gas is the average time-rate of transfer of momentum per unit area; the temperature is correlated with the kinetic energy per degree of freedom. From the mechanics of a system of molecules, and correlations between measurable and molecular quantities, it was possible to derive the empirically discovered equation of state of a gas.

Now it was not possible to describe the motion of a system of molecules in terms of the mechanics of individual particles. A classical description of the motion would require specification of initial values of coördinates of position and components of momentum. The molecules, however, were too small to make such observations on all of them; and I might add, according to quantum mechanics the classically necessary observations could not have been completed in principle. Even if the necessary observations could have been made, the enormous number of molecules would have made the mathematical solution of the problem beyond the range of our present capacities. In this situation the physicist had recourse to statistical concepts. It proved possible to determine, for example, the fraction of the molecules in a gaseous system which on the average would have speeds between 400 and 401 meters per second. The recourse to statistics was not as great a renunciation as one might think at first sight. It is statistically defined quantities that determine the observable properties of matter. Indeed, it would be inappropriate to apply the concept of temperature to a system of only one particle. Thus statistical knowledge about systems of particles provided the basis for the explanation of large-scale processes in terms of fine-scale ones.

With the foregoing preliminary statement as background, I now set forth the program of statistical physics. The basic problem is to explain the observable properties of physical reality, especially the average properties in a state of equilibrium, in terms of the integrated action of an assembly of particles. Two modes of procedure have been devised. First, one may construct a model of the system under consideration and inquire into the distribution laws for the particles. A historic example is the model of a gas which was constituted of elastic spheres. The Maxwell-Boltzmann law expresses the distribution of energy over the system of molecules. It is $a_i = Ae^{-\epsilon_i/kT}$, where a_i is the number of molecules characterized by the energy ϵ_i at temperature T. For such a model one also derives expressions for the entropy and other thermodynamic quantities. Second, one may seek to give a dynamical explanation of large-scale phenomena in terms of general principles. Especially has the attempt been made to explain the second law of thermodynamics without recourse to a detailed model. Now, in the effort to solve these problems concerning assemblies of particles, use has been made of the statistical properties of ensembles of assemblies. Such an ensemble is an assembly of a higher order, and in general will be a fiction. In the strict sense of the term, statistical mechanics probably should be restricted to those investigations into distribution laws and the laws of thermodynamics which employ ensembles of systems. The earlier kinetic theory of gases considered problems of assemblies of molecules with the aid of a concept of probability. In statistical mechanics, statements about probabilities were replaced by statements concerning relative distributions within clearly defined ensembles. In this survey I give especial attention to these ensembles.

II

In order to describe the ensembles employed in statistical mechanics it is necessary to explain the methods of describing the state of systems that are studied. In classical dynamics the state of a system is specified in terms of coördinates of position, $q_1, \cdots, q_n$, and components of momentum, $p_1, \cdots, p_n$. A set of these quantities constitutes the instantaneous phase of the system. The time-rate of change of these canonical variables is determined by Hamilton's canonical equations,

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i}.$$

It is convenient to represent the phase of a system by a point, the coördinates of which in phase space are the q's and p's. Motion of a system is then represented by progress along a curve in phase space. Since physiochemical systems consist of many partial systems, two kinds of phase space are employed, the μ-space and the Γ-space. The μ-space is the phase space for the partial system, such as a molecule of a gaseous system. If a molecule has r degrees of freedom, its phase is represented by a point in a μ-space of $2r$ dimensions. If there are N molecules in the total system, its momentary state is represented by a constellation of N points in the μ-space. The Γ-space is the phase space for the total system. If the total system consists of N partial systems, each having r degrees of freedom, the Γ-space will have $2Nr$ dimensions. To the constellation of N points in μ-space corresponds a single point which represents the total system in Γ-space.

The statistical study of systems is based upon a distinction between microphysical state and macrophysical state. This distinction may be explained with the aid of the μ-space, which is conceived to be subdivided into a large number of volume elements. In classical physics this subdivision of the phase space of the partial system could be carried out in an arbitrary manner. A microphysical state is defined by the assignment of individual molecules to the various cells. If the N molecules of a gas are designated $a, b, c, \cdots$, a microphysical state would be specified by the statement that molecules a, b, c are in cell 1, molecules d, e are in cell 2, and so forth. Such a specific assignment of individual partial systems in a total system is also called a complexion. If the cells of μ-space are of equal volume ω, a volume ω^N in Γ-space corresponds to the complexion which defines the microphysical state. A macrophysical state for a gas is determined by the distribution of the molecules among the various cells of μ-space. A distribution is specified in terms of the numbers of molecules in the cells. Thus a distribution is exemplified by the statement that there are three molecules in cell 1, two molecules in cell 2, and so forth. The distribution of similar partial systems among cells determines the observable properties of the total system. In general terms, a macrophysical state is determined by a distribution $a_1, a_2, a_3, \cdots$. A distribution law expresses the numbers a_i as functions of the specification of the cells and the conditions of the system. As previously mentioned, the Maxwell-Boltzmann law expresses the distribution of the molecules of a gas with respect to energy. To a specific distribution there correspond many complexions. A given macrophysical state can be realized by many microphysical states. Assuming interchangeability of the partial systems, the number of complexions for a given distribution is $\frac{N!}{a_1!\,a_2!\cdots}$. This number has been fundamental in the statistics based on classical mechanics and on the classical quantum theory. The volume in Γ-space corresponding to a distribution is $\frac{N!}{a_1!\,a_2!\cdots}\,\omega^N$.

As we shall see, the foregoing concepts have been modified in order to take account of the discontinuity in physical processes which has been revealed by the quantum theory. The first form of the quantum theory imposed quantum conditions upon motions that were otherwise classically determined. The phase space of a quantum system was assigned a discrete structure, such that for a system of f degrees of freedom the volume of the cells was h^f. The systems studied, multiply periodic systems, admitted the introduction of angle and action variables, $w_1, w_2, w_3, \cdots$ and $J_1, J_2, J_3, \cdots$ respectively. The action variables were quantized in accordance with the condition $J_i = n_i h$. For quantum theory a phase space was also used in which $J_1, J_2, J_3, \cdots$ are the coördinates of a point in f-dimensional action space. The phase space of the action variable has Nr dimensions in the J's. In quantum statistics the assignment of a point to a cell in phase space may be interpreted to mean that the point is restricted to the center of the cell.

III

The study of assemblies of particles with the aid of representative spaces is based on the properties of ensembles. These are the time ensemble of the given real system, and virtual ensembles of similar systems. P. Hertz also used the term "space ensemble," but in place of it I use the term "assembly." I first take up the time ensemble, which may be defined as the aggregate of phases through which a single system passes in the course of time, together with the relative durations of the system in these phases. In order to represent the time ensemble, changes of a system during time may be indicated by progress along the axis of a time coördinate. The length of time that the system is in a specific phase Π_i is then represented by an interval on the axis. The probability of a phase is expressed by the relative length of time that the system is in that phase. Thus the time ensemble is the distribution of the system in phase with respect to time. The probability of a phase is defined in terms of relative time during which the system is in the phase, but this probability may be expressed in terms of the distribution of systems within an ensemble of similar systems. One is thus led to introduce an ensemble of co-existing independent systems, and to consider their distribution in phase. Such an ensemble is essentially fictitious, and therefore has been called "virtual" by P. Hertz. The term "virtual" suggests that the ensemble consists of the possibilities of phase of a single system which alone is real. P. Hertz, however, has imagined a natural virtual ensemble: Consider a roulette game in which a circular board is subdivided into 37 equal sectors, around which a ball traveling with constant speed will remain in the null field 1/37th of the time. This defines a phase and the probability of its occurrence in terms of the time ensemble. If in a gambling hall there were 3700 games going on independently we should expect that in 100 of the games the ball would be in the null sector. To the time ensemble of a single game there corresponds a virtual ensemble of games which can be imagined as real.

Historically, the first statistical inquiry was into the distribution of molecules in an assembly with respect to a property. Maxwell's distribution law of molecular velocities was the first important result obtained in the field. For the study of these problems the application of the μ-space was especially the work of Boltzmann. Statistical inquiries were next directed to the time ensemble, the study of which was initiated by Boltzmann in his effort to prove that the Maxwell-Boltzmann distribution is the only one which can remain steady. The explicit consideration of virtual ensembles was first presented by Boltzmann in 1871. Maxwell continued the investigation of such ensembles and applied the term "statistical" to the employment of them. The theory of virtual ensembles, or simply ensembles, was built into a system by Gibbs, who in his classical treatise published in 1902 restricted the term "statistical mechanics" to the theory of ensembles of independent systems; in this usage he has been followed by P. and T. Ehrenfest. However, the scope of contemporary statistical mechanics may include any consideration of distribution laws, as well as operations with ensembles of systems.

The most characteristic procedure of statistical mechanics is the employment of virtual ensembles, which are ensembles of systems similar in structure to a given real system but differing in initial conditions. A virtual ensemble represents the totality of possibilities of phase of a given real system. The individual systems are represented by points in the Γ-space, which is filled with the representative points of many systems as if it were a fluid. As the systems change in accordance with the equations of motion, the representative points describe phase paths in the Γ-space; a volume element containing particular points is deformed by the motions into a new volume element (see fig. 1). Liouville's theorem states that the volume of a given set of points traveling with the fluid remains constant during time; the representative points flow like an incompressible fluid. Now states mechanically connected with one another are to be considered equally probable. Since the representative points for mechanically connected phases occupy equal volumes in qp phase space, the probability of a phase may be expressed as proportional to the volume occupied by the phase points. Hence Liouville's theorem renders possible the definition of probability in terms of volume of phase space.

The density of systems in Γ-space is expressed as a function of the energy. It follows that the virtual ensembles are in statistical equilibrium, that is, the time-rate of flow of points into a fixed volume element in phase space is equal to the time-rate of flow out of it. Virtual ensembles may be divided into two classes: surface ensembles distributed over an energy surface in phase space, and ensembles distributed throughout a volume in phase space. The theoretical problems raised by the employment of ensembles may be explained in terms of the surface ensemble, called "microcanonical" by Gibbs.

The observable properties of a single real system are determined by the average in time of a function of the coördinates and momenta which specify the system. Let us have given an ensemble, the phase points of which lie on a given energy surface in the phase space of the total system. The ensemble is in statistical equilibrium. The function of the ensemble of similar systems is to enable one to find the time-average of a quantity for a single system by calculating the average over the ensemble. The operation depends on the representation of the time ensemble by the virtual ensemble of similar systems. In order to justify the method, Boltzmann and Maxwell assumed that the time-average is the same for all systems on the energy surface. This will hold if all phase curves coincide, that is, if a given phase curve passes through every point of the energy surface. This hypothesis has been called the ergodic hypothesis; the systems on the surface of constant energy, the ergodic surface, have been called ergodic systems. But it has been demonstrated that ergodic systems are impossible. Of substitute hypotheses, which would justify the application of ensembles in order to obtain the approximate results of experiment, the quasi-ergodic hypothesis may be mentioned. According to the latter hypothesis the phase curve passes as near as one pleases to every point of the energy surface. The investigations of Koopman, G. D. Birkhoff, E. Hopf, and J. von Neu-

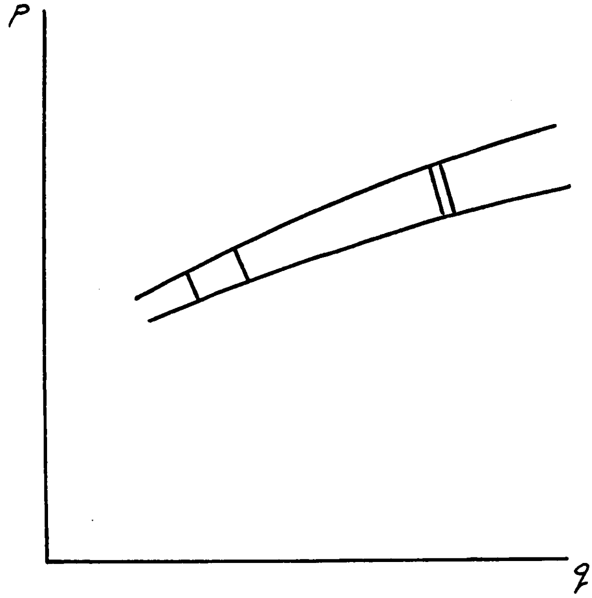

Fig. 1

mann, with the aid of the theory of linear transformations of Hilbert space, have shown that the great majority of non-integrable dynamical systems possess the quasi-ergodic property.

I now introduce virtual ensembles that are distributed throughout a volume in Γ-phase space. We have seen that the surface ensemble consists of systems, the phase points of which lie on a given energy surface, E_0 = const., in the Γ-space. Gibbs introduced an ensemble which consists of systems, the phase points of which are distributed with uniform density within a shell of phase space of thickness ΔE_0. A graph of this distribution is represented by the solid line in the accompanying diagram (fig. 2). Since ΔE_0 is infinitesimal, the distribution in the shell is equivalent for physical problems to the ergodic surface distribution of Boltzmann. The distribution in the thin shell is a discontinuous function of the energy. Gibbs also introduced a continuous distribution of phase points throughout all phase space. This canonical distribution is expressed by $\rho = Ne^{(\psi - E)/\theta}$, where ρ is the density of distribution, N the total number of systems, and ψ and θ are constants. The graph of the canonical distribution is represented by the dotted line in figure 2. On account of the large number of degrees of freedom of the systems studied, the canonical distribution has a steep maximum within the thin shell, so that the overwhelming majority of phase points fall within the thin shell. If one sets $P = e^{(\psi - E)/\theta}$, then $Pdq_1, \cdots, dq_n dp_1, \cdots, dp_n$ is the ratio of the number of systems falling within the element of extension-in-phase $dq_1, \cdots, dp_n$ to the whole number of systems. P was called by Gibbs the coefficient of probability of the phase considered. As a realization of a canonical ensemble for a gaseous system, Gibbs cited a set

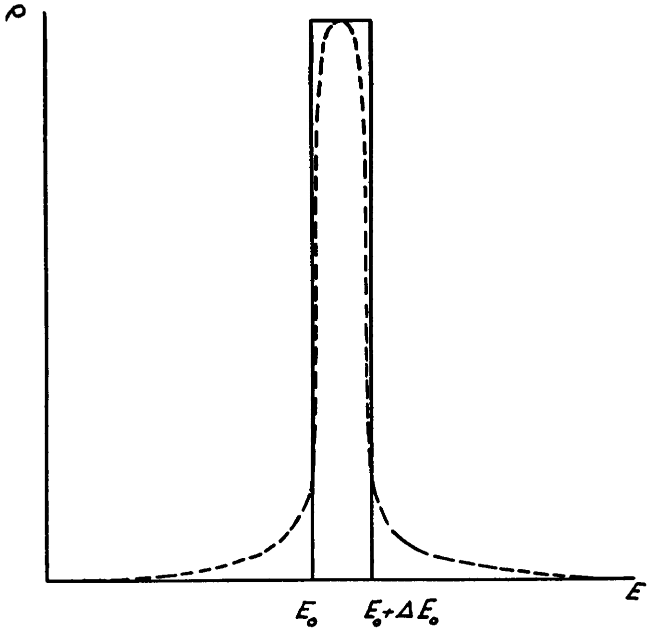

Fig. 2

of portions of gas, each in contact with an infinite reservoir of heat at a constant temperature. The energy of a gaseous system under these conditions would fluctuate about a mean value $\bar{E}$. This canonical ensemble of systems would represent the time ensemble of a single gaseous system in contact with an infinite source of heat at a given temperature. For systems of many degrees of freedom, such as a gaseous system, calculations with the canonical ensemble yield results that are equivalent for physical purposes to those obtainable with the ensemble of uniform density in a thin shell. The average value of a physical quantity taken over the canonical ensemble will be practically the same as the average over the systems in the shell, and hence the same as the average over an ensemble on the ergodic surface of constant energy.

IV

It is time now to present some of the applications of statistical mechanics to physical problems. As I have already said, statistical mechanics was created in order to apply atomism to the theory of matter. The distribution of molecular velocities in a gas was the first problem; the explanation of the laws of thermodynamics on mechanical principles was the second. In the hands of Gibbs, statistical mechanics was an instrument to explain thermodynamics as generally as possible. I consider this latter application of statistical mechanics first.

Thermodynamics is the study of transformations of systems in which changes of energy are correlated with changes of heat content and temperature. It is based upon two classical general principles, the first and second laws, to which has been added a principle or theorem about the entropy of substances at the absolute zero. The subject matter of thermodynamics is phenomenological; its concepts express directly measurable quantities such as pressure, temperature, volume, quantity of heat, electromotive force. In principle, no use is made of the concepts of atom, molecule, and other particles in thermodynamics.

Thermodynamics commands interest and admiration for the generality of its principles, the cogency of its deductions, the agreement of its consequences with experience. But thermodynamics was limited in its power to give a theoretical account of physical phenomena. For example, it supplied relations between specific heats at constant volume and at constant pressure, but was unable to determine the functional dependence of specific heat on temperature. The standpoint of this classical phenomenological theory had the advantage that its results were not affected by changes in theories of atomic and molecular structure, but it failed to give a theoretical account of the more intimate and detailed nature of phenomena. The theoretical impulse, therefore, insisted upon leaving the secure field of thermodynamics in order to analyze more intimately the mechanism of phenomena. Hence was developed the kinetic-molecular theory of matter; by the dynamics of systems of molecules it was proposed to interpret the second law of thermodynamics. As applied to reversible processes, the second law involves the concept of entropy, which is a function of the state of a system. Irreversible processes in isolated systems tend to go in a direction so as to increase the entropy of a system.

The kinetic theory, which was employed initially to explain the laws of thermodynamics mechanically, was based upon definite molecular models. Molecules were imagined as elastic spheres, and the deduction of Maxwell's distribution law by the method of collisions exemplifies the employment of this model. As conceived by Gibbs, statistical mechanics was developed in order to generalize the mechanical foundations of thermodynamics. Such a generalization was held to be desirable because, in many instances, there was lacking a knowledge of the structural and dynamical properties of matter; and even if such detailed knowledge existed one should not make use of it, for the laws of thermodyanamics, which are valid for systems of different structure, must be a consequence of the properties common to all bodies. The generality of the laws of thermodynamics demands a like generality in our mechanical discussion of them. The fundamental assumption of the statistical mechanical derivation of thermodynamics by Gibbs is that the systems are mechanical.

The canonical ensemble provides a basis for Gibbs's discussion of thermodynamics. I have already cited his expression for the coefficient of probability, $P = e^{(\psi - E)/\theta}$. $\eta = \log P = (\psi - E)/\theta$ is called the index of probability. E is the energy of the system and is a function of the phase and certain external parameters $r_1, r_2, r_3, \cdots$. The reaction of the system to external forces along the parameter r_h is given by $R_h = -(\partial E/\partial r_h)$. ψ corresponds to Helmholtz's free energy. θ is called the modulus and is equal to kT, where T is the absolute temperature in thermodynamics. The coupling of two ensembles of the same modulus is analogous to the contact of two bodies at the same temperature.

Of fundamental interest is the expression for changes of entropy in reversible processes, as a formulation of the second law of thermodynamics. Since the sum of all the probabilities is taken as unity, $\int e^{(\psi - E)/\theta} d\tau = 1$, where the integral is extended over all phase space. Considering changes of θ and r_h, we get

$$-d\bar{\eta} = \frac{d\bar{E} + \bar{R}_1 dr_1 + \cdots}{\theta}.$$

This, if we neglect the sign of averages, is identical in form with the thermodynamic equation

$$dS = \frac{d'Q}{T} = \frac{dE + R_1 dr_1 + R_2 dr_2 + \cdots}{T},$$

which is the mathematical expression of the second law of thermodynamics for reversible processes. Gibbs thus showed that $-k\bar{\eta}$ is analogous to the entropy S.

Gibbs also gave other definitions of entropy in addition to $S = -k\bar{\eta}$. A second definition is, when provided with a constant k, $S = k \log V$, where V is the volume of phase space enclosed by the ergodic surface in space over which the energy is E. A third definition is $S = k \log dV/dE$. For systems of many degrees of freedom these definitions are virtually equivalent.

Having exemplified the first application of statistical mechanics to physical problems by Gibbs's discussion of thermodynamics, I now take up the second problem, namely, to find the distribution of particles in an assembly in equilibrium. With this problem one abandons the generality of thermodynamics, and assumes the risks involved in a theory of the fine-scale structure of physical reality. For concreteness the system under consideration may be visualized as a gas consisting of N similar molecules. The distribution which is sought is expressed with the aid of the μ-space, the phase space for a molecule. In the present discussion I assume that the μ-space is subdivided into equal volume elements ω. A macrophysical state is determined by the distribution of the molecules, which is defined by a set of numbers $a_1, a_2, a_3, \cdots$. A microphysical state is determined by a complexion, which is a particular assignment of molecules to the various cells. Corresponding to a particular microphysical state there is a volume in Γ-space of magnitude ω^N. If molecules are interchanged between cells, new complexions are generated without changing the distribution. Thus, assuming the significance of interchangeability of molecules, to a given macrophysical state there correspond many microphysical states. Since to each microphysical state there corresponds an element of volume in Γ-space, there corresponds to a macrophysical state, as determined by a distribution, a set of volume elements, which in view of relations of symmetry between the members of the set has been called a "star." The number of complexions which will realize a given macrophysical state determined by the distribution $a_1, a_2, a_3 \cdots$ is $\dfrac{N!}{a_1!\, a_2!\, a_3! \cdots}$. All microphysical states are postulated to have an equal a priori probability, or, better, the same statistical weight. The total volume in Γ-space representing a given macrophysical state is then $\dfrac{N!}{a_1!\, a_2! \cdots}\omega^N$. The most probable macrophysical state, which is assumed to be the equilibrium state, is that for which the volume in phase space is a maximum. Accordingly, one seeks the distribution for which the number of complexions is a maximum, subject to the conditions

$$\begin{aligned}\Sigma a_i &= N \\ \Sigma a_i\epsilon_i &= E\,,\end{aligned}$$

where ϵ_i is the energy of the molecule in the ith cell in μ-space. One obtains the Maxwell-Boltzmann distribution law $a_i = Ae^{-\epsilon_i/kT}$.

V

So far in this survey of statistical mechanics I have formulated problems on the basis of classical mechanics. Quantum theory has brought to light discontinuities in physical processes which are symbolized by the quantum of action h ($h = 6.55 \times 10^{-27}$ erg sec.). The modification of statistical mechanics required by quantum theory has occurred in two stages. The earlier form of quantum theory, appropriately called the "classical quantum theory," introduced the quantum of action by imposing quantum conditions upon motions that were otherwise determined by classical dynamics. Consider, for example, the simple harmonic oscillator, the phase curves of which are ellipses in two-dimensional qp space (see fig. 3). The quantum condition $J = \oint pdq = nh$ restricts the possible phase curves to a discrete series of curves, between which the area is h. It is convenient to introduce angle and action variables, w, J; the action phase space for this problem of one degree of freedom is one-dimensional. The essence of the quantum theory is that the phase space of such a periodic system has a definite structure. The weight factor for regions of phase space is variable, because certain regions are prohibited to the system by the quantum conditions. In classical applications of statistical mechanics the subdivision of a phase space into cells was arbitrary. The quantum theory fixes the cells. For a system of two particles, each of one degree of freedom, the action Γ-space consists of two dimensions, along the axes of which are laid off values of $J_1 = n_1h$ and $J_2 = n_2h$. Since $J = nh$, the representative point of the system can occur only in certain discrete positions of this space. It is convenient to construct around each point a square having the side h and to think of the Γ-space as divided into elementary cells of area h^2; thus the space has the form of a two-dimensional cubical lattice. In the general case of f degrees of freedom, the cells are f-dimensional cubes of volume h^f. One speaks of points representing systems as "lying in certain cells," but in reality their positions are limited to the cell centers. All cells are equally probable, that is, have the same statistical weight. The weight of a degenerate state is given by the total volume of all cells which correspond to it in Γ-space.

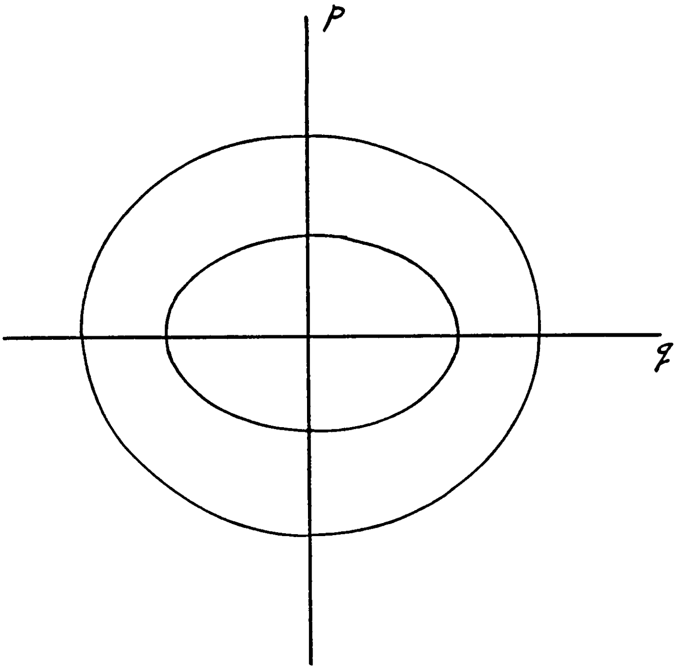

Fig. 3

The statistics of classical quantum theory, like that of classical mechanics, was based upon the principle that all complexions arising from the permutations of similar particles are equally probable a priori. The concept of complexion was based upon the assumption that one can distinguish between similar elements. The present form of the quantum theory, which may be called "wave mechanics" since it symbolizes states of systems by functions which satisfy a wave equation, abandons the principle of distinguishability, so that new forms of statistics result. However, even before the new wave mechanics was conceived, these new types of statistics had been discovered in the effort to create an adequate theory of radiation and to formulate the laws of the distribution of electrons in atoms.

The new forms of statistics discovered by physicists may be explained by considering the simple example of two particles, a and b, which may be distributed over two cells, 1 and 2. We can assign the two particles to the two cells in the four ways shown in figure 4. Each of these realizations is a complexion. In classical statistical mechanics every one of these complexions had the same statistical weight. Now a distribution is defined by the numbers a_1, a_2 of particles in the respective cells. Thus we may have three possible distributions of the two particles over the two cells:

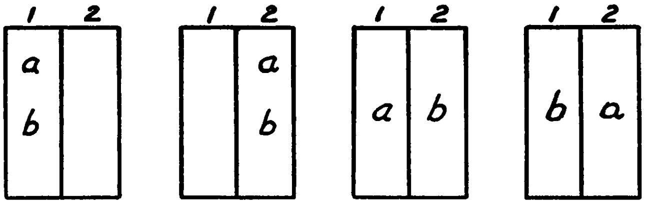

Fig. 4

$$\begin{array}{lll} \text{I.} & a_1 = 2\,, & a_2 = 0\,. \\ \text{II.} & a_1 = 0\,, & a_2 = 2\,. \\ \text{III.} & a_1 = 1\,, & a_2 = 1\,. \end{array}$$

Since two complexions realize distribution III, whereas one complexion realizes I and one complexion realizes II, in classical statistical mechanics distribution III was held to be twice as probable as either I or II. The new forms of statistics are based upon the principle that, instead of the complexion, the distribution is the unit of discussion. One mode of expression for the new statistics is that the cells of the phase space are distributed over the particles. If the μ-space is divided into M cells, and b_0 is the number of quantum cells containing no particles, b_1 the number of cells containing one particle, and so on, then the probability factor is $\frac{M\,!}{b_0\,!\,b_1\,!\,b_2\,!\,b_3\,!\,\cdots}$.

I first take up the new statistics of Bose-Einstein. This statistics was assumed by Bose in order to derive Planck's law for heat radiation by a statistical theory of light quanta, and was then applied by Einstein to ideal gases. For the

Bose-Einstein statistics it is assumed that a cell can have any number of particles. The three distributions I, II, and III are equally probable. In terms of our example of two particles which may be distributed over two cells, the two complexions in which there is one particle in each cell counts as one.

The second type of new statistics was introduced by Fermi. Pauli had previously discovered an exclusion principle for electrons in atoms: no two electrons in an atom can have the same set of quantum numbers. Fermi extended the Pauli *Verbot* to molecules of an ideal gas in which no two molecules can occupy the same cell. In our example of two particles and two cells, the distributions in which there is more than one particle, I and II, are excluded. Only distribution III is physically possible.

The difference between the Boltzmann-Planck statistics and the new forms may be expressed in terms of a difference in the definition of phase. Gibbs distinguished between a specific definition and a generic definition of phase. The basis of the generic definition was stated by Gibbs as follows: "If two phases differ only in that certain entirely similar particles have changed places with one another, are they to be regarded as identical or different phases? If the particles are regarded as indistinguishable, it seems in accordance with the spirit of the statistical method to regard the phases as identical." Classical theories of distributions of particles in systems were based upon a specific definition of phase, according to which the permutation of particles generates new phases. The new forms of quantum statistics are based upon a generic definition of phase; microphysical states which arise from a given state by interchanging similar particles are not considered to be new states.

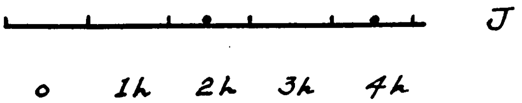

Fig. 5

The difference between the two definitions may be explained in terms of modes of representation in phase space. Consider a system of two one-dimensional particles, for each of which the possible states are $J = 0, 1h, 2h, 3h, 4h$. The action μ-space is a line which is divided into cells of volume h (see fig. 5). Let the assignment of the particles in the system be that particle a is in cell $J = 4h$, b in cell $J = 2h$. The distribution is defined by $a_2 = 1$, $a_4 = 1$, all other $a_i = 0$. The Γ-space of the system is two-dimensional with axes for J_1 and J_2 (fig. 6). The microphysical state defined by the given assignment is represented in Γ-space by the cell in which falls the point $J_1 = 4h$, $J_2 = 2h$. This cell is designated A in the diagram of the Γ-space. Another assignment of the two particles is obtained by interchanging them, so that $J_1 = 2h$, $J_2 = 4h$. This microphysical state is represented by the cell A'. Cells A and A' are symmetrical with respect to the bisectrix OD. In general, the permutation of similar particles produces in Γ-space, out of one cell, a star-shaped arrangement of cells symmetrical with respect to certain hyperplanes. If there are N particles the maximum number of cells in such a "star" is $N\,!$.

The classical statistics of physical atomism was based on the interchangeability of similar particles. All assignments which were related by permutations of particles were regarded as different microphysical states, and each was postulated to have the same probability, or statistical weight. The probability of all cells of Γ-space was the same, and every cell within a "star" had to be counted. This specific definition of phase was at the basis of classical quantum statistics, as well as Boltzmann's classical statistics. The application to the present example is that the macrophysical state determined by the distribution $a_2 = 1, a_4 = 1$ is represented by A and A'. The weight of the state is represented by the total area of the two cells A and A'. The number of possible assignments of particles in a given distribution is $\Phi_B = \frac{N!}{a_1!\,a_2!\,a_3!\cdots}$. In the example, $\Phi_B = \frac{2!}{1!1!} = 2$.

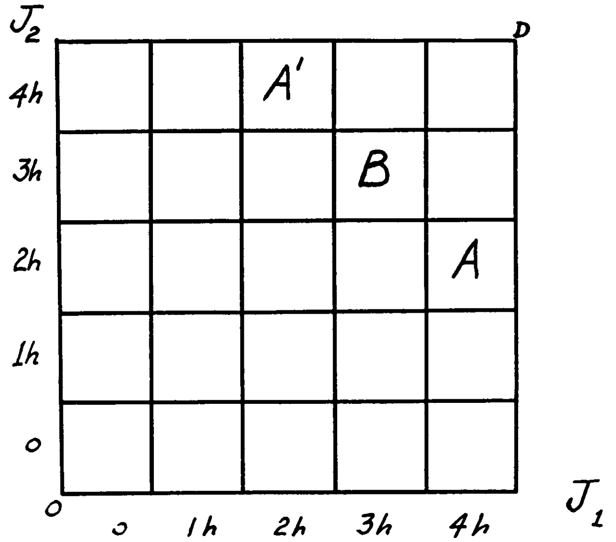

Fig. 6

The generic definition of phase is the basis of the new statistics. On this definition the totality of all equivalent phases resulting from the permutation of identical particles must be considered as the same phase, and counted as one. Only one cell in the star of cells related by permutation can be counted. In the example, the state $a_2 = 1$, $a_4 = 1$ is represented by the single cell A. Within the statistics founded on the generic definition of phase are distinguished two forms. These two forms of quantum statistics assign different roles to singular cells which occupy positions of symmetry. Consider a state of our two-dimensional illustration in which both particles have the same momenta, $J_1 = J_2 = 3h$. In the action Γ-space the state is represented by the cell B lying on the bisectrix. In this instance the permutation of particles does not generate a new cell. The Bose-Einstein statistics presupposes that any number of identical particles may be in the same quantum state. Singular cells, such as B, must be included in the representation of phase. The Fermi statistics excludes singular cells, since identical particles may not be in the same quantum state.

The statistical mechanical methods of Gibbs have been extended to quantum statistics. One expression for entropy is $S = k \log V$, where V is the volume of Γ-space enclosed by the ergodic surface of constant energy. In quantum theory the cells of the action Γ-space are of volume h^f. If Φ is the number of cells with centers in the portion of Γ-space limited by the energy surface, $\Phi = V/h^f$, so that the entropy can be expressed by $S = k \log \Phi$. In view of the generic definition of phase, however, only a fraction of the phase space must be considered. In the two-dimensional example, one takes into account only the half of the quadrant below the bisectrix OD. In general, one has to take as Φ in the equation for entropy the number of cells in a pyramid with its vertex in the origin filling the part $1/N!$ of the total volume. Since the Bose-Einstein statistics allows singular cells, whereas that of Fermi excludes them, the pyramid must be accommodated to the type of statistics.

Gibbs also defined entropy in terms of the canonical ensemble. For this ensemble $\int e^{(\psi - E)/\theta} d\tau = 1$, where the integral is extended over all phase space. The free energy is determined by $e^{-\psi/\theta} = \int e^{-E/\theta} d\tau$. The integral is called the integral of states. The entropy

$$\bar{S} = -\left(\frac{\partial\psi}{\partial T}\right)_v .$$

In quantum theory the integral of states is replaced by a sum of states taken over all the cells of Γ-space.

$$Z = \Sigma e^{-E_i/kT},$$

$$\psi = -kT \log Z,$$

$$\bar{S} = -\left(\frac{\partial\psi}{\partial T}\right)_v .$$

For the evaluation of the sum of states, Darwin and Fowler have devised methods based upon complex integration and the method of steepest descents.

VI

There remains as a final topic the physical basis which wave mechanics has provided for the new form of statistics. The new mechanics involves a drastic transformation in the concept of state of a dynamical system. In classical mechanics the state is defined in terms of values of coördinates of position and components of momentum. Wave mechanics, as interpreted under the leadership of Niels Bohr, is guided by the operational definition of physical concepts, and from this point of view significance depends on methods of determining the values of physical quantities. Now experimental arrangements for the measurement of canonically conjugated coördinates of position and components of momentum mutually interfere with one another. The determination of the value of one quantity disturbs in an uncontrollable manner a previously determined value of the conjugate quantity. Precise values of canonically conjugated variables cannot be assigned simultaneously, and hence the classical definition of state is inapplicable. In quantum mechanics the state of a system is symbolized by a function ψ, usually called a wave function, of a set of variables, such as the coördinates of position which specify the configuration of a system. The wave function transforms in time in accordance with an equation first expressed by Schroedinger, $\frac{h}{2\pi i}\frac{d\psi}{dt} + H\psi = 0$, where H is an operator derived from the Hamiltonian function of the classical equations. The characteristic values determined by Schroedinger's equation are the possible results of observation on the system. From characteristic wave functions one can calculate the probabilities of obtaining the possible results of observation.

The theory of wave mechanics for the problem of many bodies has been developed and applied to assemblies of particles, such as a gas which is an assembly of molecules, radiation in an enclosure which is an assembly of photons, an atom which is an assembly of electrified particles. The particles of these assemblies are assumed to be alike. The constituent particles in the assemblies cannot be discriminated observationally, and acknowledgment of this circumstance has led to important consequences for statistical physics.

Let us suppose that a solution of the wave equation for N particles is $\Psi = \psi_{n_1}(r_1)\psi_{n_2}(r_2)\cdots\psi_{n_N}(r_N)$, in which $r_1, r_2, r_3, \cdots$ designate the coördinates of the particles. An interchange of arguments for two particles leads to another solution of the wave equation, because a permutation of similar particles leaves the state of the system unaltered. By permuting the particles one can build up $\frac{N!}{a_1!\,a_2!\,a_3!\ldots}$ independent wave functions, a linear combination of which is again a solution. Of especial interest are the symmetrical and anti-symmetrical solutions. The symmetrical solution is formed by adding the $N!$ products obtained by permutation, $\Psi_s = \Sigma\psi_{n_1}(r_1)\psi_{n_2}(r_2)\cdots\psi_{n_N}(r_N)$. This solution remains unchanged by an interchange of particles .The anti-symmetrical solution changes its sign if two particles are interchanged. The simplest representation of this solution is the determinant

$$\Psi_a = \left|\psi_{n_i}(r_k)\right| , \qquad i, k = 1, 2, \cdots, N .$$

The anti-symmetrical solution does not exist when two or more of the subscripts n_i are identical. Anti-symmetric states of the system provide the basis for the Fermi statistics, according to which no two particles can be in the same quantum state, in other words, fall in the same cell of μ-space. In a system of N similar particles only states occur which can be expressed as a linear aggregate of characteristic functions of the same symmetric class. Symmetric and anti-symmetric states can never combine.

To conclude: Statistical mechanics was initially developed on the basis of classical mechanics. Classical statistical mechanics provided a general mechanical foundation for thermodynamics and provided means of finding the distribution of particles within an assembly. Quantum theory introduced discontinuity which was expressed in terms of a discrete structure of phase space. In classical discussions the permutation of similar particles was held to give rise to different states for statistical purposes. But the demands imposed by the statistical theories of quanta and of electrons brought about the introduction of statistics based upon indistinguishability of the particles. This statistics then found its physical foundation in the new wave mechanics which brought to light non-combining symmetric and anti-symmetric states. The symmetric states form the basis of the Bose-Einstein statistics which applies to light quanta; the anti-symmetric states form the basis of Fermi's statistics which applies to electrons. However, the general methods of classical statistical mechanics have been carried over to quantum theory. Especially have the general methods, which Gibbs so profoundly formulated for classical theory, admitted adaptation for the development of a quantum statistical mechanics.

REFERENCES

EHRENFEST, P., and T. EHRENFEST. *Encyklopaedie der Math. Wiss.*, vol. IV, p. 32, 1911.

EPSTEIN, P. "Critical appreciation of Gibbs' *Statistical Mechanics*," in *Commentary on the Scientific Writings of J. Willard Gibbs*, vol. II (ed. A. Haas). Yale University Press, 1936.

———. "Gibbs' methods in quantum statistics," *op. cit.*

GIBBS, J. WILLARD. *Elementary Principles in Statistical Mechanics*. Yale University Press, 1902.

HERTZ, P. "Statistische Mechanik," in Weber u. Gans, *Rep. der Physik*, vol. I, 1916.