NOVEMBER 1, 1937 PHYSICAL REVIEW VOLUME 52

On Quantum Measurements and the Role of the Uncertainty Relations in Statistical Mechanics

WALTER M. ELSASSER
California Institute of Technology, Pasadena, California

(Received May 1, 1937)

We investigate here the concept of generalized complementarity introduced by Bohr. It is shown that it is not possible to ascertain the microscopic state of a system without, for the purpose of measurement, imposing upon it physical conditions so stringent that they preclude the application of the statistical description which was appropriate before the measurement. One may conclude that the statistical averages which appear in the formation of an ensemble should in quantum statistics be given an interpretation somewhat different from that of classical statistics. These results lead naturally to a critical survey of the foundations of quantum statistical mechanics. We discuss the physical interpretation of statistical matrices which are the quantum equivalent of the classical ensemble. By a straightforward generalization of classical methods one obtains a statistical representation of a system if a set of measurements are given. Let ρ designate the statistical matrix which represents the system; the method consists in making the diagonal sum of $\rho \log \rho$ a minimum, while the results of the performed measurements appear as conditions of the minimum. As an application we treat the composition of two separate systems and derive from this the general proof of the H theorem which was first given by Klein.

THE fact that in quantum mechanics each measurement results in a more or less profound change in the measured system gives rise to various important problems if we consider systems with many degrees of freedom. Quantum statistical mechanics has been treated by a number of authors; among them are Pauli,[1] Klein[2] and v. Neumann.[3] It might appear, especially from v. Neumann's theory, that a quantum statistical mechanics can be built up that maintains in each step a strict analogy to the corresponding classical concepts. A closer investigation shows however that this procedure is open to serious objections. Ideas connected with the interpretation of quantum statistical mechanics have been put forward by Bohr[4] and designated by him as a generalization of complementarity. They are based upon a consistent application of the uncertainty principle to the theory of statistical ensembles. In the present paper we investigate more closely the bearing of these ideas upon certain aspects of statistical mechanics. We might remark that our problem is intimately connected with the difficulties which, according to Schrödinger,[5] still persist in our understanding of quantum theory and which in a closer consideration appear to pertain especially to the realm of statistical mechanics.

1. MEASUREMENTS AND ENSEMBLES

In classical theory an ensemble is generated by considering a large collection of samples of the same system and assuming (or proving, as the case may be) that the representative points in phase space are distributed according to certain rules. We designate throughout by "ensemble" the mathematical representation which is produced by superposing individual representations, whereas the physical specimens to which this description refers are said to form a "collection." If only the ensemble is given, the position in phase space of an individual sample may be unknown, but it is *determined*, since in classical theory we suppose that it can be measured without appreciable perturbation. In constructing the ensemble we abstract from certain parameters of each individual sample. The algebraic superposition thus obtained does not give us information about the individuals of the collection; in order to obtain such information we have to recur to measurements to be performed on the individual samples.

Now consider quantum theory and suppose that a certain collection is given and that by some straightforward argument we have obtained a representation by an ensemble. It is

---

[1] W. Pauli, *Sommerfeld Festschrift* (Leipzig, 1928), p. 30.
[2] O. Klein, Zeits. f. Physik **72**, 767 (1931).
[3] J. v. Neumann, *Mathematische Grundlagen der Quantenmechanik* (Berlin, 1932), especially chapters IV–VI.
[4] N. Bohr, Nature **131**, 421, 457 (1933).
[5] E. Schrödinger, Naturwiss. **23**, 807, 823, 844 (1935).

well-known[3, 6] that the quantum equivalent of a classical ensemble is a statistical matrix. We shall in the next section review briefly the more important properties of statistical matrices and shall for the moment confine ourselves to considerations of a more qualitative character. There arises the following problem. Suppose we want to know what the precise microscopic state of each sample is (as far as defined in quantum mechanics) and for this purpose we carry out a measurement on each sample. Our knowledge of the individual system is of course restricted by the validity of the uncertainty principle. But we are now not so much interested in this feature as in the possibility of *superposing* the representations of the individual samples again in order to form an ensemble after the measurement has been performed. In most cases the systems will be modified by the measurement. If this does not hold, we meet other difficulties. We shall consider two typical cases of measurements.[7]

The first type comprises such measurements which do not alter the measured system. As an example we consider the measurement of the energy. It is well known that this measurement requires time, and if we want to distinguish between two eigenvalues of the energy which differ by the amount ΔE, the time required will be of the order $\hbar/\Delta E$. During that interval the system has to be isolated from external influences. For simple systems such a measurement presents no difficulty.[6] But it is readily seen that for systems with many degrees of freedom, such as are considered in statistical mechanics, an enormous length of time is required since the density of eigenvalues becomes extremely large. Such a type of measurement is therefore without significance for the purposes of statistical mechanics. To make this point clear it should be noted that this difficulty does not result from any lack of ingenuity in constructing our measuring instruments or from an inadequate choice of the used idealizations. The importance of using adequate idealizations in clearing up the limits of measurements has been strikingly demonstrated by Bohr and Rosenfeld.[8] It is evident, however, that the limit considered here is inherent in the very nature of quantum theory. This applies also to the typical examples to be considered later. If we study representations of systems in the sense of statistical mechanics, it is absurd to make use of measurements which require long periods, as compared with the time during which we are actually interested in the evolution of the system. These measurements can teach us nothing for our purpose. Quantum mechanics provides us with an operational scheme, not only in the mathematical sense of the word, but also in the physical sense. It prescribes definite rules of how to manipulate the measuring instrument in relation to the measured object in order to obtain a certain type of information. In fact, according to what particular measuring operation we perform, the result will be one of a complete set of pure states, the resulting *set* depending essentially upon the type of measurement chosen.[9] It can, however, actually become absurd to fulfill the requirements which would lead to the determination of a pure state; then we lose the possibility of obtaining a *unique* application of the formalism of quantum mechanics. Mathematically, it means that we are justified in representing our system by a pure state (a single wave function) whenever an appropriate type of measurement has been carried through. In all other cases we cannot give a preference to one definite pure state and we should therefore represent our system by a statistical matrix (an ensemble of wave functions) in a way which will appear more clearly later.

The same kind of argument as outlined here for the energy applies also to all quantities for which the measurement requires a very long time. The quantity to be measured need not always be a constant of the motion; if it changes in time we may choose our measuring instrument in such a way that it just counterbalances in each moment the change undergone by the quantity. More details about this latter topic, which we

---

[6] W. Pauli, *Handbuch der Physik*, Vol. 24, second edition, section 9.

[7] The distinction between measurements of the first and second kind has been introduced by Pauli, reference 6.

[8] N. Bohr and L. Rosenfeld, *Danske Vidensk. Selskab.* (1933).

[9] This fact has recently been discussed from a viewpoint close to ours by E. Schrödinger, Proc. Camb. Phil. Soc. **31**, 555 (1935); **32**, 446 (1936).

need not consider here, may be found in v. Neumann's book.[10]

We take now a second kind of measurement. Suppose we measure the position in configuration space of our system (i.e. we measure simultaneously the positions of all its constituents). The measurement may be such that following it the system is certainly to be found within a small volume element of the configuration space. This measuring operation is evidently of a perturbing type. The coupling between the measuring device and the parts of the system will in general be much stronger than the interaction between the parts of the system themselves. After the measurement has been performed on each sample of the collection, we may again superpose the individual wave representations so that in the ensemble (statistical matrix) thus obtained each part of the configuration space appears with its due weight. But this collection will physically be fundamentally different from what it was before the measurement. Indeed, as a result of the measurement, the constituents of the system will have obtained very high average velocities; the previous couplings will thus no longer hold and the system will burst into a number of parts within a comparatively short time after the measurement. There is, then, no physical sense in having a common representation of this collection by a statistical matrix. For if we follow the behavior of the samples even during a short time, they will completely diverge from each other as well as from the average behavior of the collection before the measurement. In this case, again, the use of an ensemble for describing and especially for predicting the average behavior of a collection of systems as found in nature and the precise knowledge of the state of the individual samples exclude each other.

In order to avoid the change in a system caused by performing a measurement on it, it has been proposed[10] to compensate this change by an appropriately chosen additional potential so as, for instance, to make the state after the measurement a stationary one. In our case we could, immediately following the measurement, surround each system by a sufficiently high potential barrier in configuration space which prevents it from escaping out of the position in which it has been found by the measurement. But this auxiliary potential would have to be different for each individual sample and it lacks evidently all physical sense to use an algebraic superposition of such states for describing an average behavior of systems; this statistical representation by an ensemble would become a mere mathematical fiction without any physical significance. Therefore, in this case also we spoil the physical sense of the collection by determining the states of the individual samples.

The examples just considered explain the sense of the generalized principle of complementarity.[4] We recognize that quantum mechanics may be applied uniquely only in the case of systems of comparatively small size. If systems with many degrees of freedom are involved, *the possibility of giving a unique quantum-mechanical representation of a system by a pure state and the possibility of leaving it in approximately the conditions under which it appears as sample of a given collection, will in general exclude each other.* As distinguished from classical mechanics, this condition cannot be overcome by sufficiently refined measuring devices.

In order to understand better the sense of such a statement, it should be made clear that the concept of a statistical ensemble has essentially an *a posteriori* character. There exists no general *a priori* principle of how to collect samples which can be represented by the same statistical matrix and there is no universal criterion for the degree of variation that would suffice to exclude a sample from a given collection. It is therefore not possible to ascertain in rigorous *a priori* terms the meaning of an appropriate statistical representation of a system. In individual cases we can frequently define it without difficulty.

We may now ask how to characterize measurements which do not vitiate the capacity of systems to serve as samples of a given collection. The performance of a measurement on each sample will frequently make it necessary to split a given collection in a number of sub-collections according to the respective results of the measurements. If, for instance, a measurement would determine a pure state of the system, then each pure state might give rise to a separate collection.

---

[10] v. Neumann, reference 3, p. 189.

If the measurement is less precise, the new collections can be more comprehensive. We will especially ask for the conditions which have to be fulfilled in order that the effect of the uncertainty relations brought about by the measuring process does not lead to a change in the essential properties of the collection which was given before the measurement. We shall call such measurements *nonperturbing* ones. This term can be understood to mean also that the measurement should not require an excessively long time during which the system would have to be isolated from all external influences. As already pointed out, a concept of this kind will have an essentially *a posteriori* character and can only be defined with respect to given concrete circumstances. Since in quantum mechanics each measurement produces *some* perturbation, there exist only approximately nonperturbing measurements. The mathematical expression of a measurement in terms of statistical matrices will be given later. It might nevertheless be useful to illustrate at this time our considerations by studying briefly a nonperturbing measurement in a simple special case.

Take a Brownian particle suspended in a gas or liquid. We propose to measure, for instance, the x coordinate of the particle at a given moment. In doing this we communicate to the particle an average momentum of at least the amount

$$\Delta p = \hbar/\Delta x, \tag{1}$$

if Δx is the margin of precision of our measurement. The particle has itself a certain average speed due to its temperature movement and we may assume that our measurement does not essentially perturb the system, if the momentum produced by the measurement does not exceed the average momentum under the temperature movement. According to (1) this gives immediately

$$\Delta x \geqq \hbar/(mkT)^{\frac{1}{2}}, \tag{2}$$

where m is the mass of the Brownian particle. A measurement of the type defined by (2) may here be called nonperturbing.

The Brownian particle will undergo a displacement which, by Einstein's formula, using Stokes' law for the resistance of small spheres, may be expressed as

$$\langle x^2 \rangle_{\mathrm{Av}} = kT \cdot t/3\pi\eta r, \tag{3}$$

where r is the radius of the sphere and η the viscosity of the gas or liquid. It may be seen by comparing (3) with (2) that after some time the particle will have left the area covered by the wave packet of the size Δx; therefore relevant statistical predictions about the behavior of our system can be drawn, even if only nonperturbing measurements are given. Assume now, for instance, that the equality sign holds in (2). Predictions of the type (3) can only be made for a limited time. Indeed, from (1) we know that the wave packet of the size Δx representing the result of our measurement will spread out in course of time and after a period which is not too short we may write

$$\langle |x| \rangle_{\mathrm{Av}} = \Delta p \cdot t/m = t \cdot (kT/m)^{\frac{1}{2}}.$$

Since this effect is proportional to t, whereas the Brownian displacement is proportional to $t^{\frac{1}{2}}$, the spreading out of the wave packet will always in course of time overtake the diffusion effects, and predictions based upon (3) then become invalid. This, of course, represents just the remainder of the unavoidable perturbation connected in quantum mechanics with each measurement.

The considerations given above may illustrate the different role which statistical mechanics plays in quantum theory as compared with classical physics. In classical theory we use methods of statistical mechanics in order to save the labor of detailed measurements and of calculations concerning features of the system in which we have no particular interest. It has however a definite physical sense to ascribe such features to the system, even if we should happen not to know them, since they can always be determined by measurements which do not produce an essential perturbation of the system. In quantum mechanics, on the other hand, there is no sense in speaking about a definite microscopic structure of a system, *i.e.*, *a definite pure state*, unless such a state has first been realized in the measuring operation. Indeed the state, more precisely the orthogonal set of states which will show up as a result of the measurement, depends essentially upon the particular choice of the measuring device.[9] From this viewpoint statistical mechanics gains a much greater significance than it seems to have in classical physics.

2. Statistical Matrices

We shall now review briefly some properties of statistical matrices.[3, 6] Suppose first that our system can be described by a single wave function $\psi(x)$, where x denotes as usual all coordinates of the system; such a system is said to be in a pure state and we can define a corresponding statistical matrix by

$$(x'|\rho_\psi|x'')=\psi^*(x')\psi(x''). \tag{4}$$

A generalized statistical matrix which represents an ensemble or "mixture" consists of a number of pure states $\psi_1\psi_2\cdots$, superposed with the respective probabilities $r_1r_2\cdots$. We suppose the ψ_n to be all orthogonal and normalized and we normalize the r_n by the condition $\sum_n r_n=1$. The statistical matrix representative of the mixture is then defined as a sum of matrices of the type (4):

$$(x'|\rho|x'')=\sum_n r_n\psi_n^*(x')\psi_n(x''). \tag{5}$$

It is convenient to have a special sign for the sum of the diagonal elements of a matrix. This quantity, usually called the "trace" of the matrix, will be designated by $\mathrm{Tr}(\rho)$ and we have by the normalization of the r_n and ψ_n:

$$\mathrm{Tr}(\rho)=1. \tag{6}$$

The trace is invariant with respect to a transformation of the type

$$X^{-1}\rho X=r. \tag{7}$$

Now suppose in particular X to be unitary and r to be the diagonal form of ρ. We can easily show that the eigenvalues of ρ are just the r_n of (5). Indeed

$$\rho\psi_m^*=\sum_n r_n\psi_n^*(x')\int\psi_n(x'')\psi_m(x'')dx''=r_m\psi_m^*$$

and the ψ_m^* are the corresponding eigenfunctions. Since the r_n are all positive by their definition as probabilities, we see that a statistical matrix can only have positive eigenvalues. A matrix with this property is called "definite." On the other hand, each definite Hermitian matrix may of course be written in the form (5).

For the sake of simplicity we confine ourselves in the following considerations to matrices which have only discrete indices. The diagonal elements of a definite Hermitian matrix are all positive in *any* representation of the matrix. Indeed, we have from (7), since X is unitary

$$\rho_m=\sum_n|x_{nm}|^2 r_n, \tag{8}$$

where ρ_m stands for the diagonal elements ρ_{mm}. Let ρ' be any other statistical matrix; its eigenfunctions may be different from the ψ_n^*. Consider the sum $\rho''=\alpha\rho+\beta\rho'$ with positive coefficients α and β. This is again a definite matrix. We have namely for the diagonal elements $\rho_m''=\alpha\rho_m+\beta\rho_m'$ and therefore ρ_m'' is positive by (8). Since this is true for all systems of reference, it is evidently true for the system in which ρ'' is diagonal. Hence ρ'' has only positive eigenvalues and is a definite matrix.

We prove now *Klein's lemma*:[2, 11] For every definite Hermitian matrix we have the inequality

$$\sum\rho_m\log\rho_m\leqq\sum r_n\log r_n, \tag{9}$$

where ρ_m are again the diagonal elements of ρ in any representation and r_n the eigenvalues. The equality sign holds if and only if ρ is diagonal so that the ρ_m are just a permutation of the r_n.

We know that the ρ_m as well as the r_n are all positive. Consider the auxiliary function

$$Q_{nm}=r_n(\log r_n-\log\rho_m)-r_n+\rho_m.$$

It is readily seen that this function vanishes for $\rho_m=r_n$ and is positive for all positive values of ρ_m and r_n. Now multiply Q_{nm} by $|x_{nm}|^2$ and sum with respect to n and m. Considering (8) and the relations

$$\sum_n|x_{nm}|^2=\sum_m|x_{nm}|^2=1$$

we obtain

$$\sum_{nm}Q_{nm}|x_{nm}|^2=\sum_n r_n\log r_n-\sum_m\rho_m\log\rho_m. \tag{10}$$

Since the left-hand side is essentially positive, our inequality (9) is proved. We have yet to ascertain in which case the equality sign holds.

---

[11] See also v. Neumann, reference 3, p. 202.

We shall here confine ourselves to the case of a matrix ρ with a finite number, say N, of rows and columns. If the equality sign holds in (9), the left-hand side of (10) must vanish and since all terms are essentially positive, we must for each pair of indices (n, m) have either $Q_{nm}=0$ or $x_{nm}=0$. But a unitary matrix in N dimensions has at least N nonvanishing elements one in each row and one in each column; therefore at least N of the Q_{nm} must be zero one for each value of n and m. This proves according to (10) that the ρ_m reproduce the r_n in some order. If we label the ρ_m conveniently, X can be taken as the unit matrix; otherwise X will correspond to some permutation of the r_n.

We come now to the physical properties of statistical matrices. If H is the Hamiltonian of the system, the statistical matrix ρ changes in time according to the Schrödinger equation:[3, 12]

$$i\hbar\partial\rho/\partial t = H\rho - \rho H. \tag{11}$$

It is easily seen from (11) that $\mathrm{Tr}(\rho)$ is a constant of the motion. The same is true for the trace of any function of ρ.

We consider, finally, measurements. Let F be any physical quantity of the considered system. If we measure its value on each sample of the collection, then the average value or "expectation value" of F in the collection is given by

$$\bar{F} = \mathrm{Tr}(\rho F). \tag{12}$$

We have of course $\mathrm{Tr}(\rho F) = \mathrm{Tr}(F\rho)$. We must give closer consideration to the physical significance of formula (12). In v. Neumann's interpretation of statistical mechanics it is assumed that each individual sample of a collection is in a definite pure state and that the performance of a measurement results in another pure state (usually different from the previous one). Now we have shown in the first section that it is necessary to build up statistical mechanics without such a specialized assumption. Usually we will not be authorized to assume that the individual system is in a pure state, neither before, nor after the measurement. We can and we shall therefore regard the statistical matrix as the appropriate physical description of the *individual* samples as well as of the collection as a whole. The relation between the samples and the ensemble is then the same as the relation between a physical system and its wave description in the theory of pure states: Unless new measurements are performed, the wave function or the statistical matrix represents the most precise description of the system which is available with the given information.

A quantity F will in general have a statistical dispersion in the ensemble ρ and we may for instance measure the dispersion by the square root of $[(F-\bar{F})^2]_{\mathrm{Av}}$. v. Neumann[3] has shown that in a statistical matrix which is not a pure state each quantity has a nonvanishing dispersion. Suppose now we have measured F on each sample of our collection. The measurement will have a certain margin of precision. Suppose this margin is smaller than the dispersion of F in the ensemble. The ensemble will then be called *separable* with respect to the measurement of F. We may divide the corresponding collection in a number of part collections in such a way that all samples of one part collection have the same value of F within the margins of the measurement. An ensemble in which the dispersion of F does not exceed the margin of precision of the measurement of F is called *nonseparable* with respect to this measurement. In the same way we can define nonseparability with respect to a set of several simultaneous measurements. A separable statistical matrix can be expressed as a sum with positive coefficients of nonseparable matrices.

The concept of separability will be helpful in formulating in a more precise way the results of the first section. It is evident that if an ensemble is nonseparable at a certain instant, this must not hold for all time. Especially if the system undergoes a strong perturbation, the nonseparability will rapidly be lost. It has been pointed out above that in such a case the statistical description is of little physical value. Only nonseparable ensembles which remain so for a reasonable length of time are useful in statistical mechanics. Nonperturbing measurements are such measurements which applied to a collection described by a nonseparable matrix do not, for a reasonable length of time, transform it into a collection described by a separable matrix.

---

[12] P. A. M. Dirac, Proc. Camb. Phil. Soc. 25, 62 (1929).

We can in the future confine ourselves to the consideration of nonseparable ensembles. The breaking up of a separable ensemble into nonseparable ones and the reciprocal superposition do not involve any features characteristic of quantum mechanics. These operations are within the realm of classical statistical mechanics. On the other hand, if a collection can be represented by a nonseparable ensemble, this means that the samples have no individual features to distinguish them from each other. We said above that the statistical matrix represents the individual sample as well as the whole collection. We can now state more precisely that this applies only to nonseparable matrices.

If a number of measurements performed on an *individual* system is given, we may ask for a description of this system by a statistical matrix. The statistical statements implied in the matrix will express the behavior of a nonseparable collection for which all the measured quantities have the given values. Under these conditions formula (12) represents also the results of a measurement on the collection. If therefore a number of measurements on a system are given and a representative statistical matrix has to be found, we may write *conditions* for this matrix in the form (12), where on the left-hand side we insert the results of the measurements. The method of constructing such a matrix will be considered in the following section.

It will be useful to ascertain that the definition (12) of a measurement is a consistent one. By this we mean a proof of the fact that the result of a physical measuring operation can be given in the form (12); this will be analogous to well-known considerations about measurements in the quantum theory of pure states.[6] We shall give a proof for arbitrary measuring devices, assuming however that the interaction due to the measuring process can be represented by a first-order perturbation. The measuring operation consists in a unique coupling between the quantity to be measured and a macroscopic quantity called the "pointer" of the measuring instrument. From the position of the pointer we infer the value of the quantity to be measured.

Consider first macroscopic quantities. From Ehrenfest's well-known theorem on the movement of wave packets it follows immediately that if formula (12) is applied to any dynamical variables, their average values will obey the classical Hamiltonian equations

$$d\bar{p}_k/dt = -\partial\bar{H}/\partial q_k, \quad d\bar{q}_k/dt = \partial\bar{H}/\partial\bar{p}_k$$

whenever it is possible to express the Hamiltonian $\bar{H}$ in a sufficient approximation as a function of the variables $\bar{p}_k$ and $\bar{q}_k$. Formula (12) is therefore an appropriate definition of the macroscopic variables of a system and is thus certainly justified when the measured quantities pertain to classical macroscopic mechanics. The general measurement to be considered here need not be a nonperturbing one. Let the Hamiltonian of the measuring instrument be $M(y)$, while the measured object has the Hamiltonian $H(x)$. Let V be the interaction potential and let the representation of the system before the measurement be $\rho = \rho_x{}^\circ \cdot \rho_y{}^\circ$. Inserting this into the Schrödinger equation (11) we have

$$i\hbar\partial\rho/\partial t = (H+M+V)\rho_x{}^\circ\rho_y{}^\circ - \rho_x{}^\circ\rho_y{}^\circ(H+M+V).$$

We are interested in the change brought about in the measuring instrument by the interaction. We find for this, if Tr_x designates the trace taken in the x space alone,

$$i\hbar\partial\rho_y/\partial t = M\rho_y{}^\circ - \rho_y{}^\circ M + Tr_x(V\rho_x{}^\circ\rho_y{}^\circ - \rho_x{}^\circ\rho_y{}^\circ V).$$

Let Y be a macroscopic variable of the measuring instrument, the one which characterizes the position of the pointer. For the sake of simplicity assume further that the pointer is at rest if no interaction with the measured object takes place; i.e.,

$$\mathrm{Tr}(YM\rho_y{}^\circ - Y\rho_y{}^\circ M) = 0.$$

We obtain for the displacement of the pointer

$$\begin{aligned} dY/dt &= \mathrm{Tr}_y(Y\partial\rho_y/\partial t) \\ &= (i\hbar)^{-1}\,\mathrm{Tr}_y(Y\mathrm{Tr}_x(V\rho_x{}^\circ\rho_y{}^\circ - \rho_x{}^\circ\rho_y{}^\circ V)) \\ &= (i\hbar)^{-1}\,\mathrm{Tr}_x(\mathrm{Tr}_y(YV\rho_y{}^\circ)\cdot\rho_x{}^\circ - \rho_x{}^\circ\cdot\mathrm{Tr}_y(Y\rho_y{}^\circ V)) \\ &= (i\hbar)^{-1}\,\mathrm{Tr}_x(\rho_x{}^\circ\cdot\mathrm{Tr}_y(YV\rho_y{}^\circ - Y\rho_y{}^\circ V)). \end{aligned}$$

The last expression is just of the form (12) with respect to the variables x of the measured object; on the other hand the left-hand side gives the change of a macroscopic quantity to which the expectation value thus determined is uniquely related. This shows that within the limits of a first-order approximation the result of a measuring operation may be expressed by (12).

We can in the same way write down the perturbation brought about in the measured object by the measuring process. As pointed out above, the latter perturbation determines whether a nonseparable ensemble remains nonseparable after the measurement. It seems appropriate to use in this case not only a first order approximation but rather the general integral of the Schrödinger equation say $\rho(t)$. The perturbation brought about in the measured object by the measuring instrument may then be written

$$\Delta\rho_x(t)=\mathrm{Tr}_y(\rho(t))-\rho_x(t)$$

where ρ and ρ_x refer to the Hamiltonians $H+M+V$ and H, respectively. For nonperturbing measurements this expression has to remain small for a considerable length of time.

3. The Extremum Principle

According to the considerations of the previous section we will now mainly deal with nonseparable ensembles. Our problem is the construction of a nonseparable statistical matrix which represents a given system, the system being characterized by a number of measurements. The results of the measurements will appear as conditions of the form (12) for the statistical matrix ρ. These conditions do not uniquely define ρ; the problem of determining ρ is *essentially* an indefinite one. In a system with many degrees of freedom there will be a large number of eigenvalues r_n of ρ which are all not negligibly small; there will on the other hand only be a restricted number of measurements. Thus there is a very large number of statistical matrices, all equally well compatible with the given conditions. We can however try to make the problem artificially a determined one by introducing suitable assumptions. Such assumptions are suggested by familiar ideas of classical statistical mechanics. A closer consideration of the latter will show that we may distinguish two main types of assumptions. The first type expresses the equal *a priori* weight of all cells in phase space, or, in quantum theory, of all nondegenerate pure states. We might try to assume that all pure states which are compatible with the given measurements appear with the same weight and all other states with weight zero. It would be difficult to obtain in this way a simple analytical expression for ρ and the expression would moreover depend upon the frame of pure states chosen and would not be invariant in functional space. Although a similar condition seems necessary, it ought to be introduced in an alternate way.

A second assumption which is usually made in statistical mechanics concerns the distribution at random of certain phases of the samples. It should be noticed that "distribution at random" is an indefinite concept. By this we mean that its significance is not expressible in terms of a constructive rule of operation.[13] Concepts of this type occur frequently in statistical mechanics as well as in probability theory in general and are responsible for most of the paradoxa and complexities which are so characteristic of this science. The very fact that, in spite of the indefiniteness of such assumptions, we arrive at definite analytical expressions for the probabilities shows that we have not so much used the assumptions in a constructive way, as that we have substituted analytical relationships which agree sufficiently with the commonsense meaning of these concepts.

Let us now recall that the problem of finding a statistical matrix representing the system is essentially an indefinite one. In view of the difficulties just mentioned, it seems appropriate to introduce from the beginning an *assumption of an analytical character* which yields definite values for all the probabilities. In quantum theory this can be done in a fairly simple way by means of a variational principle. If we introduce such a principle as a postulate, all statements about the system take the form of definite analytical expressions for the probabilities. It is readily seen that the expectation value of $\rho\log\rho$ is closely related to the entropy.[2, 3] For a canonical ensemble it will just be equal to the negative of the entropy divided by Boltzmann's constant. The quantity

$$\eta=\mathrm{Tr}(\rho\log\rho)=\sum r_n\log r_n \tag{13}$$

will be called the *mixture index* of the statistical matrix ρ. The r_n, being the eigenvalues of ρ, are

---

[13] It is for instance not possible to give a prescription of how to construct an infinite sequence in which the figures 0 and 1 appear with equal probability, but otherwise "at random." This however does not disprove the existence of such sequences.

all positive fractions smaller than unity and therefore η is an essentially negative quantity. We take now as the fundamental assumption

$$\eta = \mathrm{Tr}(\rho \log \rho) = \mathrm{Min.} \tag{14}$$

The minimum has to be chosen in such a way as to account for the given measurements. Let

$$\mathrm{Tr}(A_1\rho) = a_1, \quad \mathrm{Tr}(A_2\rho) = a_2, \quad \cdots \tag{15}$$

stand for the measurements where $a_1a_2\cdots$ are the respective results of them. The minimum (14) has to be found under the conditions (15) together with the normalization condition (6). We shall show later on that the problem has in many cases a unique solution. Therefore, by (14), all probability statements concerning our system become definite.

An important special case of (15) is the one in which all the operators $A_1A_2\cdots$ commute with each other. We shall prove the following *theorem:* If all the operators $A_1A_2\cdots$ commute with each other, the matrix ρ which has a minimum mixture index commutes with all of them. Since $A_1A_2\cdots$ commute, we can transform them simultaneously into diagonal matrices. In the corresponding system of reference the conditions (15) and (6) will involve only the diagonal elements ρ_m of ρ. On the other hand, from Klein's lemma (9) we have by (13)

$$\sum \rho_m \log \rho_m \leqq \mathrm{Tr}(\rho \log \rho). \tag{9$'$}$$

This shows that for any given set of diagonal elements the minimum mixture index is assumed if all nondiagonal elements vanish, for only in this case the equality sign will hold. This being true for any set of ρ_m, it holds also for the one which minimizes $\sum \rho_m \log \rho_m$ under the conditions (15) and (6). Thus our theorem is proved.

There are certain limits for the application of (14). According to the previous section, it is necessary that for all the measured quantities $A_1A_2\cdots$ the dispersions in the ensemble ρ obtained by (14) come out to be smaller than the respective margins of the measurements. Measurements refer primarily to individual systems; instead we can consider a collection of systems in which the measurements have all yielded the same results within their respective margins. Such a collection is represented by a nonseparable ensemble. The condition just stated means that the statistical matrix obtained by (14), in order to have a physical sense, must be nonseparable with respect to the given measurements. In this case the collection to which it refers does not contain samples which would deviate from the given measured values.

We may easily extend our method so as to include separable collections. The expressions (15) will then refer to average values of measurements on the collection. This is just the case of "classical" statistics. Let us now consider the superposition of two ensembles, say $\rho = \alpha\rho' + \beta\rho''$ where $\alpha + \beta = 1$ and the matrices are all normalized. If ρ' and ρ'' are determined by (14), their superposition ρ will in general be different from the matrix obtained when the two collections are first superposed and then (14) be applied with the expectation values of the combined collection substituted in (15). Simple relations will hold if ρ' and ρ'' differ only very slightly from each other. Since we deal with systems with many degrees of freedom, we may neglect the variation of $\log \rho$ if the variation of ρ is small and write

$$\rho \log \rho = (\alpha\rho' + \beta\rho'') \log \rho = \alpha\rho' \log \rho' + \beta\rho'' \log \rho''.$$

This formula refers to the superposition of two *collections* which are suitably described by ρ' and ρ'' and it has nothing to do with measurements. Let us on the other hand assume that ρ' and ρ'' have been obtained by (14) under conditions (15). For the sake of simplicity let $a_2a_3\cdots$ be the same in both ensembles and let only a_1 be different, its value being a' in the first and a'' in the second ensemble. The minimum mixture index can be considered as a function of a_1; now since a' and a'' differ only very slightly from each other, we have identically

$$\eta(\alpha a' + \beta a'') = \alpha\eta(a') + \beta\eta(a'').$$

This formula refers to the dependence of η (and thus of ρ) upon *measurements* formally given by (15) and states nothing about collections. Now $\alpha a' + \beta a''$ is just the expectation value of a_1 in the combined collection. Comparing the two last formulae we infer: If two collections differ only very slightly from each other, we obtain the same statistical matrix whether we apply (14) to the collections individually and then superpose the matrices obtained, or superpose the collections first and determine ρ from the expectation values of the combined collection.

We may close this section by a few remarks of a general character. The difficulty which we encountered above, namely to describe a system with a restricted amount of information is of course by no means characteristic of quantum theory. We have merely to deal, under a particular aspect, with a general problem of probability theory, namely the problem of inverse probabilities or of *inductive inference*. By definition, inductive inference starts from premises which are not sufficient to determine uniquely the results to be reached and its application has therefore to be based upon appropriate general principles.[14]

Now the fact that in our problem we deal with a form of inductive inference may so far be taken as a rather trivial point of terminology. However, in statistical mechanics one may give to the term "induction" a specific *physical* significance. It has been pointed out above that the quantum theory of pure states does not exhaust the problem of giving a quantum representation of complex systems. If we obtain such a representation by means of (nonseparable) statistical matrices *we proceed essentially from the macroscopic features of a system towards the microscopic features*. This is quite in contradistinction to the naive classical picture of an object where it is usually silently supposed that the macroscopic features of a system are determined in function of its microscopic ones. We may therefore term as "physical induction" the method outlined here in which all elements of the description are determined by their relation to the macroscopic quantities characterizing the system. The theory of measurements has been briefly treated above and it may now be seen that a measurement constitutes the limiting case in which inductive inference concerning the microscopic features tends towards certainty. In this case, of course, the state is uniquely determined without an additional assumption like the minimum principle (14). Quantum statistical description of a system represents more precisely a mixture of inductive and deductive elements, the latter arising from the fact that the Hamiltonian of the system is usually assumed as given *a priori*.

We may apply this idea to explain a specific aspect of quantum indetermination in the theory of statistical matrices. Approaching the situation naively we might be tempted to say that our lack of knowledge concerning the microscopic variables of a system is of two different types. The one is that *necessary* lack of knowledge which results immediately from the uncertainty relations and which cannot be eliminated from the quantum theory of pure states. The second type is the lack of knowledge of phases which is basic in the classical theory of ensembles and which is there *accidental*, since the knowledge can always be gained by measurements on the individual samples. In quantum theory, the corresponding circumstance has to be expressed as the lack of knowledge which is due to the superposition of several pure states in an ensemble. Both types of indetermination will enter into a prediction of the future behavior of the system. Now it ensues from our above analysis that in a nonseparable ensemble there will be no physical sense in distinguishing between the effects due to the specific quantum indetermination and those due to the accessory indetermination expressed by the superposition of several pure states in the ensemble.

4. APPLICATIONS

We shall confine ourselves here to the most elementary applications of our principle (14). As a first example we shall consider the canonical ensemble.[15] It results under the conditions

$$\mathrm{Tr}(H\rho)=E, \quad \mathrm{Tr}(\rho)=1.$$

From the theorem proved in the preceding section we see that ρ commutes with H and this means according to the Schrödinger equation (11) that the ensemble is stationary. Let W_n be the eigenvalues of H; for the eigenvalues r_n of ρ we have the differential equations

$$(d/dr_n)(r_n \log r_n+\alpha W_n r_n+\beta r_n)=0,$$

where α and β are Lagrangian parameters. Determining these from $\sum W_n r_n=E$ and $\sum r_n=1$ we

---

[14] For a discussion of this problem from the viewpoint of probability calculus see for instance R. A. Fisher, Proc. Camb. Phil. Soc. **26**, 528 (1929); **28**, 257 (1932).

[15] J. v. Neumann, *Göttingen Nachrichten*, 1927, p. 273.

obtain finally

$$r_n = e^{-W_n/\theta}/(\sum_n e^{-W_n/\theta}),$$

where the constant θ is defined by the equation

$$\sum_n W_n e^{-W_n/\theta} = E \sum_n e^{-W_n/\theta}.$$

We may also write the result in the invariant form[15]

$$\rho = e^{-H/\theta}/\mathrm{Tr}(e^{-H/\theta}).$$

The properties of the canonical ensemble are familiar and need not be considered here.

We pass now to a second application. Let two separate systems be given; we may then either have a common representation of them by a statistical matrix and may ask for the representation of one system which does not refer to the other; or we may be given a statistical matrix for each system and may be required to find a common representation of both.

We begin with the first problem. Let (x) and (y) designate the coordinates of the first and second system, respectively. According to v. Neumann a consistent definition of the "projection" of ρ upon the part spaces (x) and (y) is given by

$$\rho_x = \mathrm{Tr}_y(\rho), \quad \rho_y = \mathrm{Tr}_x(\rho). \tag{16}$$

The projection of a statistical matrix upon a part system is therefore equal to the expectation value of the unit operator in the complementary part system. From the fact, proved in the second section, that a sum of definite matrices is again definite it can be deduced that the projection of a definite matrix upon a part space is also a definite matrix. Suppose now the system is in a pure state. A wave function of the whole system will be a product of the wave functions of the part systems. If, however, the two systems have previously been interacting and have then been separated, the common wave function will no longer be a product. It can easily be verified from (16) and (5) that the projection of the pure state which represents the total system upon a part system will in general be a mixture.[16] Einstein, Podolsky, Rosen[17] and a number of other authors[18] have thoroughly investigated this case and have, in particular, brought out an interesting paradox concerning the influence which a measurement performed on one part system has upon the quantum mechanical representation of the other. For our present purpose we are not so much concerned with this difficulty as with the elementary fact underlying it, namely, that the systems after interaction have phase relations; measurements performed on one system will enable us to draw inferences about the state of the other.

Consider the inverse problem of the previous one: Given two separate systems and two statistical matrices ρ_x and ρ_y representing them; find a common representation by one statistical matrix ρ. It is easily seen that the problem is far from being uniquely determined, it has in general a large number of solutions. If, however, we choose as the representation that for which the common mixture index is a minimum under the conditions (16), the solution will become unique. If the mixture index is to be made a minimum, ρ will be the so-called outer product of ρ_x and ρ_y:

$$\rho = \rho_x \times \rho_y,$$

$$\text{meaning} \quad (m\mu|\rho|n\nu) = (m|\rho_x|n)(\mu|\rho_y|\nu). \tag{17}$$

Since ρ_x and ρ_y are defined in independent spaces and therefore commute we can transform them simultaneously into diagonal matrices. We can show that the matrix ρ which has a minimum mixture index is diagonal in the same system of reference. This follows in the same way as the theorem about commutative matrices given in the preceding section: By Klein's lemma (9′) the minimum mixture index compatible with a given set of diagonal elements is obtained if all nondiagonal elements are zero. These values fulfill identically the nondiagonal part of the matrix Eqs. (16), namely,

$$\sum_m (m\mu|\rho|m\nu) = \sum_\mu (m\mu|\rho|n\mu) = 0$$

for $\mu \neq \nu$ and $m \neq n$, respectively.

---

[16] This fact may for instance be used to obtain the canonical ensemble by projecting a pure state with respect to the energy of a very large system upon a small part system—quite in analogy to classical statistics.

[17] A. Einstein, B. Podolsky and N. Rosen, Phys. Rev. **47**, 777 (1935).

[18] N. Bohr, Phys. Rev. **48**, 696 (1935); E. Schrödinger, reference 9; W. M. Furry, Phys. Rev. **49**, 393 (1936); H. Margenau, Phys. Rev. **49**, 420 (1936).

Let $\rho_{m\mu}$ and ρ_{xm} and $\rho_{y\mu}$ designate the *diagonal* elements of ρ and ρ_x and ρ_y, respectively. The diagonal part of the conditions (16) reads now

$$\sum_m \rho_{m\mu} = \rho_{y\mu}, \quad \sum_\mu \rho_{m\mu} = \rho_{xm}. \tag{18}$$

The minimum is determined by the differential equations

$$(d/d\rho_{m\mu})(\rho_{m\mu} \log \rho_{m\mu} + \alpha_\mu \rho_{m\mu} + \beta_m \rho_{m\mu}) = 0$$

which have the solutions

$$\rho_{m\mu} = e^{-\alpha_\mu - \beta_m - 1}.$$

Inserting this into (18) we obtain

$$e^{-\alpha_\mu} = e^\gamma \rho_{y\mu}, \quad e^{-\beta_m} = e^{1-\gamma} \rho_{xm},$$

where γ is a real constant. Therefore, finally $\rho_{m\mu} = \rho_{xm}\rho_{y\mu}$. Now formula (17) is invariant if ρ_x and ρ_y undergo a transformation of the type (7) in their respective part spaces. We may therefore bring them back in their original form, if the latter was nondiagonal; thus (17) holds generally true when the mixture index is minimized.

If we measure the expectation value of any operator R in (x) and of any operator S in (y), then the simultaneous expectation value of both is according to (17) equal to $\bar{R} \cdot \bar{S}$. Now this is just the formula for the simultaneous occurrence of independent events; therefore no correlations between pairs of measurements performed separately on each of the part systems do exist. This excludes also the occurrence of the Einstein paradox referred to above. We can show on the other hand that if (17) does not hold, such correlations can always be found.

If again ρ_x and ρ_y are given, there will then be at least one element of ρ, say $(m'\mu'|\rho|n'\nu')$, which is not equal to $(m'|\rho_x|n')(\mu'|\rho_y|\nu')$. Let $E_{m'n'}$ be an operator in the x-space whose matrix element $m = m'$, $n = n'$ is unity, all other elements being zero, and let $E_{k'l'}$ be another operator of this kind. Consider the linear combination $\alpha E_{m'n'} + \beta E_{k'l'}$; in order that all expectation values in system (y) be independent of the particular choice of α and β in (x), we must have

$$\begin{aligned}(\mu'|\mathrm{Tr}_x(\rho(\alpha E_{m'n'} + \beta E_{k'l'}))|\nu') \\ = (\alpha \bar{E}_{m'n'} + \beta \bar{E}_{k'l'})(\mu'|\rho_y|\nu')\end{aligned}$$

for all values of α and β. From this we derive easily

$$\frac{\alpha}{\beta} = -\frac{(k'\mu'|\rho|l'\nu') - (k'|\rho_x|l')(\mu'|\rho_y|\nu')}{(m'\mu'|\rho|n'\nu') - (m'|\rho_x|n')(\mu'|\rho_y|\nu')},$$

where the denominator does not vanish according to hypothesis. This relationship can evidently not hold identically for all values of α and β and therefore, if a pair of measurements involves essentially the matrix elements $(m'|\rho_x|n')$ and $(\mu'|\rho_y|\nu')$, respectively, their results are not independent of each other.

We may thus state: There exist no correlations between measurements performed on two separate systems if and only if their common representation, subject to the conditions (16), is of the form (17) corresponding to a minimum mixture index. From this we see that a representation which is unique will remain unique if any number of dimensions are added to the phase space of the system, so that there is no interaction between the newly added part and the old one. We may therefore expect that the representation induced by the minimum principle (14) will in many cases be unique.

We make now an application of the last results. The mixture index of a combination of two systems fulfills the inequality

$$\mathrm{Tr}(\rho \log \rho) \geqq \mathrm{Tr}(\rho_x \log \rho_x) + \mathrm{Tr}(\rho_y \log \rho_y). \tag{19}$$

The case of equality follows immediately from (17) and (13); if (17) does not hold the inequality sign applies, since then η will be larger than its minimum value. We may use this relation to prove the quantum equivalent of Boltzmann's H theorem in a case which, although not the most comprehensive one, is rather general. The first application of Boltzmann's method to quantum theory has been given by Pauli[1] who used a semiclassical method. An entirely quantum mechanical treatment is due to Klein.[2] He considers the change in time of the diagonal sum $\sum \rho_m \log \rho_m$ in a fixed system of reference. However the selection of this quantity as representing the entropy might seem arbitrary, and we shall here give a slightly different presentation of Klein's result.

Let us first remark that, just as in classical mechanics, any representation ρ will in course of time come back arbitrarily close to its value for $t=0$. Indeed we have in the system of reference in which H is diagonal

$$\rho_{mn}(t)=\rho_{mn}{}^{\circ}e^{-i(W_m-W_n)t/\hbar},$$

where the $\rho_{mn}{}^{\circ}$ are constants. After a possibly very long time all the exponentials will simultaneously be again arbitrarily close to unity. Therefore, in order to obtain statements about irreversibility, it is, just as in classical theory, necessary to introduce "disorder" assumptions or an equivalent hypothesis. Now since such assumptions are implied in our variational principle, we may in many cases avoid a repeated introduction of them.

We know that $d\eta/dt=0$ for any closed system. But this is not true for the mixture index referring to the *projection* of a statistical matrix upon a part system. It seems therefore appropriate to use such projections in the statement of the H theorem, particularly since they have frequently an immediate physical significance. Let us write $\eta_x=\mathrm{Tr}(\rho_x \log \rho_x)$, etc. We presume that

$$(d/dt)\ (\eta_x+\eta_y+\eta_z+\cdots)\leqq 0, \qquad (20)$$

where (x), (y), $(z)\cdots$ refers to any division of the system in part systems. Whereas the mathematical statement can be made generally, the division, in order to be significant should be established by nonperturbing measurements. Since $\eta_x\eta_y\cdots$ come after a certain long time very closely back to their initial values an actual decrease will only take place if the period is shorter than this.

It will be sufficient to consider two part systems (x) and (y). We shall confine ourselves to the important special case that for $t=0$ the representation is of the form (17), η having its minimum value. By this we anticipate disorder assumptions. The equality sign holds then in (19) for $t=0$, but while the left-hand side is a constant of the motion, the right-hand side must decrease in time. This follows from the straightforward fact that after some time has elapsed, ρ will no longer be a product (17) and therefore the inequality sign applies in (19) and consequently in (20).

The mixture index η of a system is a function of the performed measurements (15). If we add new measurements to a given set of conditions (15), the result will in general be an increase of η. There is thus no sense in saying that the mixture index of a system will in general decrease in time. We have instead to refer to the *same* set of measurements made at two successive times. Then the mixture index resulting from the second set will be smaller than that resulting from the first set.

We may express our result in an alternate way. If by (14) we have obtained a representation of the system by a statistical matrix, we may calculate the representation $\rho(t)$ at a later moment by the Schrödinger equation (11). Inserting this in the conditions (15), the expectation values $a_1a_2\cdots$ become functions of the time. But under new values $a_1a_2\cdots$ the previous η will no longer be the minimum mixture index. We may state that these *expectation values change always in such a way that the minimum η to be calculated from them decreases in course of time.* This is a direct expression of the second law of thermodynamics.

The author is indebted to Professor F. London for a number of stimulating discussions which he had with him some time ago, and wants to express particular thanks to Professor J. R. Oppenheimer for his continued interest and most helpful advice and criticism.