

# On the Statistical Description of Quantum Systems

Stig Stenholm<sup>1</sup>

*Received March 24, 1971*

---

It is emphasized that the use of statistical methods is a consequence of the incomplete isolation of physical systems. The description of a partially specified quantum state by a density operator is shown to be equivalent to a description in terms of an ensemble of systems. Several definitions of the statistical operator are shown to be identical.

---

**KEY WORDS:** Definition of the density matrix; ensemble; preparation device; reduction of the state; influence of the surroundings.

## 1. INTRODUCTION

The use of statistical ensemble theory in physics derives from our incomplete knowledge about reality. The traditional approach<sup>(1)</sup> to quantum statistical mechanics recognizes the fact that a macroscopic description never specifies the microstate of a system uniquely. When making macroscopic predictions, one has to incorporate contributions from all microscopic states compatible with our knowledge, with equal weight. This statement is introduced as the basic postulate of statistical physics by Tolman.<sup>(1)</sup> In practice, some macroscopic values occur in the overwhelming majority of microscopic states, and these are the most probable predictions of the values of the macroscopic variables. For large systems, the deviations are small enough to make the method rigorous.

In some systems, e.g., the classical gas, an ergodic theorem is needed to justify the use of ensembles, whereas in other systems, especially solids, the measured quantities are obviously averages over a large number of nearly equal parts and the representative ensemble is realized in Nature.

In this paper, we want to justify the method of statistical physics from the following point of view: The closed models of physics represent real systems only

---

<sup>1</sup> Department of Technical Physics, Helsinki University of Technology, Otaniemi, Finland.

approximately, as the rest of the world always acts as a perturbation. These surroundings constitute the heat and particle baths occurring in thermal physics. This view is well recognized by Gibbs,<sup>(2)</sup> who shows that energy conservation in the combined system leads to a canonical distribution in the subsystem.

The subsystem under consideration together with its surroundings may be described by a quantum mechanical state. If the description is formulated in terms of the variables of the subsystem only, it was recognized by Landau<sup>(3)</sup> that a more general description in terms of a statistical or density matrix is needed.

Each member of an ensemble of systems is described quantum mechanically by a state vector in Hilbert space. The "end points" of the vectors form a point set describing the ensemble in the same way as a classical ensemble is described by a point set in phase space. Von Neumann<sup>(4)</sup> showed that the ensemble is represented by a statistical operator in quantum mechanics. Independently, Dirac introduced both the quantum mechanical density operator<sup>(5)</sup> and the statistical one.<sup>(6)</sup> The equivalence of the two definitions seems intuitively clear, but is not transparent. To this author's knowledge, the equivalence is discussed only by ter Haar,<sup>(7)</sup> and his proof is too short to provide real insight into the physical assumptions behind the equivalence. It is the purpose of this paper to show, in a nonrigorous way, how the picture of physical systems embedded in unobserved surroundings leads to the description of systems by means of statistical operators.

The paper ends by displaying the equivalence of several usual definitions of the statistical density operator.

## 2. QUANTUM MECHANICAL DESCRIPTION OF A SUBSYSTEM

We consider a dynamic subsystem embedded in its surroundings. When the subsystem and the surroundings are uncoupled, its state can be described by the complete set  $\{|\psi_m\rangle\}$  and the surroundings can be characterized by an array of quantum numbers  $\mu$ . The symbol  $m$  as well as  $\mu$  stands for a set of quantum numbers, with enough members to characterize the states uniquely. The state of the combined system  $|\Psi\rangle$  may then be expanded in the states of the subsystem and the surroundings:

$$|\Psi\rangle = \sum_{\mu m} \langle \mu m | \Psi \rangle |\mu\rangle |\psi_m\rangle \quad (1)$$

The expectation values of operators  $A$  which act only on the variables of the subsystem can be written

$$\langle \Psi | A | \Psi \rangle = \sum_{mm'} \langle \psi_m | A | \psi_{m'} \rangle \sum_{\mu} \langle \mu m' | \Psi \rangle \langle \Psi | \mu m \rangle = \text{Tr } A\rho \quad (2)$$

where  $A$  and  $\rho$  are operators on the subsystem only with the matrix elements

$$\begin{aligned} A_{mm'} &= \langle \psi_m | A | \psi_{m'} \rangle \\ \rho_{m'm} &= \sum_{\mu} \langle \mu m' | \Psi \rangle \langle \Psi | \mu m \rangle \end{aligned} \quad (3)$$

If we define the pure state density operator as  $|\Psi\rangle\langle\Psi|$ , we see that  $\rho_{m'm}$  is the reduced density matrix obtained by tracing over the unobserved states of the surroundings.

As an example, we take the quantum numbers  $m$  to be the coordinates  $x_i$  which are observed and the numbers  $\mu$  to be unobserved coordinates  $q_j$ . The component

$$\langle q_j, x_i | \Psi \rangle \equiv \Psi(q_j, x_i) \tag{4}$$

is then the wave function of the combined systems. The relation (3) becomes

$$\rho(x_i, x_i') = \int dq_j \psi(q_j, x_i) \psi^*(q_j, x_i') \tag{5}$$

which is the conventional definition of the quantum mechanical density matrix in atomic and molecular physics.<sup>(7)</sup>

According to the conventional interpretation of quantum mechanics, the probability of occurrence of a given pair  $(\mu, m)$  in the state (1) is  $|\langle \mu m | \Psi \rangle|^2$ . The probability for a given value  $\mu$  irrespective of  $m$  is

$$P(\mu) = \sum_m |\langle \mu m | \Psi \rangle|^2 \tag{6}$$

We now take the conventional interpretation of quantum mechanics to imply that, if we measure the values of  $\mu$  and  $m$  on an ensemble of systems all in the state  $|\Psi\rangle$ , we will obtain a distribution determined by the probabilities above. After the measurement, the systems are generally not in the same state any more. We now assume the existence of a preparing apparatus which can separate the systems according to the value of  $\mu$  without measuring it. The simplest example of such a preparing device is the Stern–Gerlach spin filter.

The distribution of systems over the various channels is given by (6). If we establish the presence of a system in a given channel by measuring  $\mu$ , the state in this channel is reduced to

$$|\psi(\mu)\rangle = \left( \sum_m \langle \mu m | \Psi \rangle |\psi_m\rangle \right) / \left( \sum_m |\langle \mu m | \Psi \rangle|^2 \right)^{1/2} \tag{7}$$

obtained by normalizing the component of  $|\Psi\rangle$  on the vector  $|\mu\rangle$ .

We now consider the reduced density operator of Eq. (3),

$$\begin{aligned} \rho &= \sum_{m'm} |\psi_{m'}\rangle \rho_{m'm} \langle \psi_m| \\ &= \sum_{m'm} |\psi_{m'}\rangle \left[ \sum_{\mu} \langle \mu m' | \Psi \rangle \langle \Psi | \mu m \rangle \right] \langle \psi_m| \end{aligned} \tag{8}$$

Taking into account the definition (7) of the filtered states and the probabilities of their occurrence (6), we write (8) as

$$\begin{aligned} \rho &= \sum_{\mu} \left[ \sum_{m'} |\psi_{m'}\rangle \langle \mu m' | \Psi \rangle \right] \left[ \sum_m \langle \Psi | \mu m \rangle \langle \psi_m| \right] \\ &= \sum_{\mu} |\psi(\mu)\rangle P(\mu) \langle \psi(\mu)| \end{aligned} \tag{9}$$

This is the definition of the statistical operator given by Dirac.<sup>(8)</sup>

The conclusion is: Calculating expectation values of operators acting on the subsystem only using the state  $|\Psi\rangle$ , one obtains the same results as performing the following: (a) take an ensemble of systems in the state  $|\Psi\rangle$ ; (b) pass them through a filter separating them according to the value of  $\mu$ ; (c) verify by measurement that systems with the values  $\mu$  are present; the wave function of each is then reduced to (7); (d) form an ensemble of all the systems that have passed the filter; (e) calculate measurable quantities of the subsystem as averages over this ensemble.

This set of rules gives an operational prescription for the construction of an ensemble corresponding to an unobserved probability distribution in the surroundings.

### 3. THE STATISTICAL OPERATOR

If we have a total of  $N$  systems, we find that the number of systems in the channel  $\mu$  is

$$N_\mu = P(\mu)N \quad (10)$$

Introducing this into (9) gives

$$\rho = (1/N) \sum_\mu |\psi(\mu)\rangle N_\mu \langle\psi(\mu)| \quad (11)$$

If we number the systems by  $i = 1, 2, \dots, N$ , the state of the  $i$ th system is  $|\psi(\mu_i)\rangle$  and the sum in (11) can be written as a sum over the systems, instead of a sum over the states  $\mu$ ,

$$\rho = (1/N) \sum_i |\psi(\mu_i)\rangle \langle\psi(\mu_i)| \quad (12)$$

where the state  $|\psi(\mu_i)\rangle$  occurs  $N_{\mu_i}$  times in the summation. If we choose an arbitrary complete set of states

$$\sum_a |\varphi_a\rangle \langle\varphi_a| = 1 \quad (13)$$

in the Hilbert space of the subsystem, the density matrix in this representation is

$$\begin{aligned} \rho_{ab} &= \langle\varphi_a | \rho | \varphi_b\rangle = (1/N) \sum_i \langle\varphi_a | \psi(\mu_i)\rangle \langle\psi(\mu_i) | \varphi_b\rangle \\ &= \overline{c_a(i) c_b^*(i)} \end{aligned} \quad (14)$$

where we have expanded the state of the  $i$ th system in the basis  $\{|\varphi_a\rangle\}$ :

$$|\psi(\mu_i)\rangle = \sum_a c_a(i) |\varphi_a\rangle \quad (15)$$

The expressions (9), (11), (12), and (14) are equivalent expressions for the statistical density operator. The component form (14) of the definition is used by Tolman.<sup>(1)</sup>

### 4. DISCUSSION

In this paper, I have put forward a simple statistical interpretation of the quantum mechanical density matrix. A quantum mechanical system which is only incompletely described cannot be represented by a wave function but requires a density matrix (see the discussion by Landau and Lifshitz<sup>(9)</sup>). This approach is, in spirit, equivalent to their view of statistical physics<sup>(10)</sup> as a description of systems only approximately isolated from their surroundings. They define the probability distribution for the quantum mechanical states in such a way that the ergodic theorem is trivially satisfied. As I have pointed out, there is often reason enough for the introduction of ensembles even without the use of an ergodic theorem. This possibility is clearly indicated by Gibbs<sup>(2)</sup> and emphasized by Tolman.<sup>(1)</sup> In my opinion, it is futile to attempt a justification of statistical physics for closed systems, because its use in actual physical applications clearly derives from neglect of interactions between the systems involved.

If we only specify the total energy of the combined system, we have to include contributions from all states (1) with the given total energy. According to the principle of equal *a priori* probability, each system has the same weight and we obtain the micro-canonical distribution. It is then shown by Gibbs<sup>(2)</sup> that the subsystem is described by a canonical distribution (see also Ref. 10).

In case the surroundings consist of a classical gas of  $N$  free particles, this result can easily be derived explicitly.<sup>(11)</sup> The total energy is

$$E = \langle H_{\text{sub}} \rangle + \sum_i (1/2m) \langle p_i^2 \rangle \tag{16}$$

and the states  $|\mu\rangle$  are

$$\langle r_1, r_2, \dots, r_N | p_1, \dots, p_N \rangle = V^{-3N/2} \exp \left[ i \sum_i \mathbf{p}_i \cdot \mathbf{r}_i / \hbar \right] \tag{17}$$

If  $|\psi_n\rangle$  are the eigenstates of  $H_{\text{sub}}$  with eigenenergy  $E_n$ , the reduced density operator for given total energy  $E$  is

$$\begin{aligned} \rho &= \int d^{3N}r \sum_n \sum_{p_1 \dots p_N} |\psi_n\rangle \langle r_1, \dots, r_N | p_1, \dots, p_N \rangle^2 \delta \left( E - E_n - \sum_i p_i^2/2m \right) \langle \psi_n | \\ &= \sum_n |\psi_n\rangle P(n) \langle \psi_n | \end{aligned} \tag{18}$$

where, in the limit of a large system, we have

$$\begin{aligned} P(n) &= (2\pi)^{-3N} \int d^{3N}p \delta \left( E - E_n - \sum_i p_i^2/2m \right) \\ &= (2\pi)^{-3N} (1 - E_n/E)^{-1} \int d^{3N}p \delta \left\{ E - \sum_i [p_i^2/2m(1 - E_n/E)] \right\} \\ &= (2m/4\pi^2)^{3N/2} (1 - E_n/E)^{(3N-2)/2} \int d^{3N}x \delta \left( E - \sum_i x_i^2 \right) \end{aligned} \tag{19}$$

Setting  $E = N\epsilon$  and assuming  $E_n \ll E$ , we find for the dependence on  $E_n$  that

$$\begin{aligned} P(n) &= \text{const} \times (1 - E_n/N\epsilon)^{(3N-2)/2} \\ &\approx \text{const} \times e^{-3E_n/2\epsilon} \end{aligned} \quad (20)$$

which gives a canonical distribution with the temperature determined by equipartition

$$\epsilon = \frac{3}{2}kT \quad (21)$$

in the gas. As a classical gas may be used to define the temperature scale, this shows that our derivation is quite general.

In the filter discussed in Section 2, we assumed that the different values  $\mu$  were separated on an equal basis. When we use it as a preparing device for real measurements, we often want to transmit systems with different  $\mu$  values differently. Then, the weight function  $P(\mu)$  of (9) will depend both on the incoming state and on the filter. This is very close to the approach taken when density matrices are used in spin-polarization experiments.<sup>(12)</sup> The filter prepares the ensemble to be measured by affecting the spin states differently. The weight functions are here determined by the experimental setup only and no thermal considerations are normally involved.

## REFERENCES

1. R. C. Tolman, *The Principles of Statistical Mechanics*, Oxford University Press, Oxford (1938).
2. J. W. Gibbs, *Elementary Principles in Statistical Mechanics*, Yale University Press, New Haven (1902) [Dover Publications, New York (1960)], p. 183.
3. L. D. Landau, *Z. Physik* **45**:430 (1927) [*Collected Papers*, Pergamon Press, Oxford (1965), p. 8].
4. J. von Neumann, *Göttinger Nachr.* No. 3, **245**:273 (1927) [*Collected Papers*, Pergamon Press, Oxford (1961), Vol. I, p. 208].
5. P. A. M. Dirac, *Proc. Cambridge Phil. Soc.* **26**:376 (1930); **27**:240 (1931).
6. P. A. M. Dirac, *Proc. Cambridge Phil. Soc.* **25**:62 (1929); **26**:361 (1930).
7. D. ter Haar, in *Physical Chemistry*, Vol. 2, H. Eyring, D. Henderson, and W. Jost, eds., Academic Press, New York (1967), Chapter 2.
8. P. A. M. Dirac, *The Principles of Quantum Mechanics*, 4th ed., Oxford University Press, Oxford (1957), p. 132.
9. L. D. Landau and E. M. Lifshitz, *Quantum Mechanics* (Course of Theoretical Physics, Vol. 3), Pergamon Press, London (1958), pp. 35–38.
10. L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Course of Theoretical Physics, Vol. 5), Pergamon Press, London (1959), pp. 13–19.
11. P. C. Hemmer, unpublished lecture notes.
12. J. Hamilton, *The Theory of Elementary Particles*, Oxford University Press, Oxford (1959), pp. 375ff.