

## On the Quantum Correction For Thermodynamic Equilibrium

By E. WIGNER

*Department of Physics, Princeton University*

(Received March 14, 1932)

The probability of a configuration is given in classical theory by the Boltzmann formula  $\exp[-V/kT]$  where  $V$  is the potential energy of this configuration. For high temperatures this of course also holds in quantum theory. For lower temperatures, however, a correction term has to be introduced, which can be developed into a power series of  $h$ . The formula is developed for this correction by means of a probability function and the result discussed.

1

IN classical statistical mechanics the relative probability for the range  $p_1$  to  $p_1+dp_1$ ;  $p_2$  to  $p_2+dp_2$ ;  $\dots$ ;  $p_n$  to  $p_n+dp_n$  for the momenta and  $x_1$  to  $x_1+dx_1$ ;  $x_2$  to  $x_2+dx_2$ ;  $\dots$ ;  $x_n$  to  $x_n+dx_n$  for the coordinates is given for statistical equilibrium by the Gibbs-Boltzmann formula

$$P(x_1, \dots, x_n; p_1, \dots, p_n) dx_1 \dots dx_n dp_1 \dots dp_n = e^{-\beta\epsilon} dx_1 \dots dx_n dp_1 \dots dp_n \quad (1)$$

where  $\epsilon$  is the sum of the kinetic and potential energy  $V$

$$\epsilon = \frac{p_1^2}{2m_1} + \frac{p_2^2}{2m_2} + \dots + \frac{p_n^2}{2m_n} + V(x_1 \dots x_n) \quad (2)$$

and  $\beta$  is the reciprocal temperature  $T$  divided by the Boltzmann constant

$$\beta = 1/kT. \quad (3)$$

In quantum theory there does not exist any similar simple expression for the probability, because one cannot ask for the simultaneous probability for the coordinates and momenta. Moreover, it is not possible to derive a simple expression even for the relative probabilities of the coordinates alone—as is given in classical theory by  $e^{-\beta V(x_1 \dots x_n)}$ . One sees this by considering that this expression would give at once the square of the wave function of the lowest state  $|\psi_0(x_1 \dots x_n)|^2$  when  $\beta = \infty$  is inserted and on the other hand we know that it is not possible, in general, to derive a closed formula for the latter.

The thermodynamics of quantum mechanical systems is in principle, however, given by a formula of Neumann,<sup>1</sup> who has shown that the mean value of any physical quantity is, (apart from a normalizing constant depending only on temperature), the sum of the diagonal elements of the matrix

$$Qe^{-\beta H} \quad (4)$$

where  $Q$  is the matrix (operator) of the quantity under consideration and  $H$  is the Hamiltonian of the system. As the diagonal sum is an invariant under

<sup>1</sup> J. von Neumann, Gött. Nachr. p. 273, 1927.

transformations, one can choose any matrix or operator-representation for the  $Q$  and  $H$ . In building the exponential of  $H$  one must, of course, take into account the non-commutability of the different parts of  $H$ .

## 2

It does not seem to be easy to make explicit calculations with the form (4) of the mean value. One may resort therefore to the following method.

If a wave function  $\psi(x_1 \cdots x_n)$  is given one may build the following expression<sup>2</sup>

$$P(x_1, \cdots, x_n; p_1, \cdots, p_n) = \left(\frac{1}{h\pi}\right)^n \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} dy_1 \cdots dy_n \psi(x_1 + y_1 \cdots x_n + y_n)^* \psi(x_1 - y_1 \cdots x_n - y_n) e^{2i(p_1 y_1 + \cdots + p_n y_n)/h} \quad (5)$$

and call it the probability-function of the simultaneous values of  $x_1 \cdots x_n$  for the coordinates and  $p_1 \cdots p_n$  for the momenta. In (5), as throughout this paper,  $h$  is the Planck constant divided by  $2\pi$  and the integration with respect to the  $y$  has to be carried out from  $-\infty$  to  $\infty$ . Expression (5) is real, but not everywhere positive. It has the property, that it gives, when integrated with respect to the  $p$ , the correct probabilities  $|\psi(x_1 \cdots x_n)|^2$  for the different values of the coordinates and also it gives, when integrated with respect to the  $x$ , the correct quantum mechanical probabilities

$$\left| \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \psi(x_1 \cdots x_n) e^{-i(p_1 x_1 + \cdots + p_n x_n)/h} dx_1 \cdots dx_n \right|^2$$

for the momenta  $p_1, \cdots, p_n$ . The first fact follows simply from the theorem about the Fourier integral and one gets the second by introducing  $x_k + y_k = u_k; x_k - y_k = v_k$  into (5).

Hence it follows, furthermore, that one may get the correct expectation values of any function of the coordinates or the momenta for the state  $\psi$  by the normal probability calculation with (5). As expectation values are additive this even holds for a sum of a function of the coordinates and a function of the momenta as, e.g., the energy  $H$ . In formulas, it is

$$\begin{aligned} & \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} dx_1 \cdots dx_n dp_1 \cdots dp_n [f(p_1 \cdots p_n) + g(x_1 \cdots x_n)] \\ & P(x_1 \cdots x_n; p_1 \cdots p_n) \\ & = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \psi(x_1 \cdots x_n)^* \left[ f\left(\frac{h}{i} \frac{\partial}{\partial x_1}, \cdots, \frac{h}{i} \frac{\partial}{\partial x_n}\right) \right. \\ & \quad \left. + g(x_1 \cdots x_n) \right] \psi(x_1 \cdots x_n) dx_1 \cdots dx_n \end{aligned} \quad (6)$$

for any  $\psi, f, g$ , if  $P$  is given by (5).

<sup>2</sup> This expression was found by L. Szilard and the present author some years ago for another purpose.

Of course  $P(x_1, \dots, x_n; p_1, \dots, p_n)$  cannot be really interpreted as the simultaneous probability for coordinates and momenta, as is clear from the fact, that it may take negative values. But of course this must not hinder the use of it in calculations as an auxiliary function which obeys many relations we would expect from such a probability. It should be noted, furthermore, that (5) is not the only bilinear expression in  $\psi$ , which satisfies (6). There must be a great freedom in the expression (5), as it makes from a function  $\psi$  of  $n$  variables one with  $2n$  variables. It may be shown, however, that there does not exist any expression  $P(x_1 \dots x_n; p_1 \dots p_n)$  which is bilinear in  $\psi$ , satisfies (6) and is everywhere (for all values of  $x_1, \dots, x_n, p_1, \dots, p_n$ ) positive, so (5) was chosen from all possible expressions, because it seems to be the simplest.

If  $\psi(x_1, \dots, x_n)$  changes according to the second Schrödinger equation

$$i\hbar \frac{\partial \psi}{\partial t} = - \sum_{k=1}^n \frac{\hbar^2}{2m_k} \frac{\partial^2 \psi}{\partial x_k^2} + V(x_1, \dots, x_n) \psi \tag{7}$$

the change of  $P(x_1, \dots, x_n; p_1, \dots, p_n)$  is given by

$$\frac{\partial P}{\partial t} = - \sum_{k=1}^n \frac{p_k}{m_k} \frac{\partial P}{\partial x_k} + \sum \frac{\partial^{\lambda_1 + \dots + \lambda_n} V}{\partial x_1^{\lambda_1} \dots \partial x_n^{\lambda_n}} \frac{(h/2i)^{\lambda_1 + \dots + \lambda_n - 1}}{\lambda_1! \dots \lambda_n!} \frac{\partial^{\lambda_1 + \dots + \lambda_n} P}{\partial p_1^{\lambda_1} \dots \partial p_n^{\lambda_n}} \tag{8}$$

where the last summation has to be extended over all positive integer values of  $\lambda_1, \dots, \lambda_n$  for which the sum  $\lambda_1 + \lambda_2 + \dots + \lambda_n$  is odd. In fact we get for  $\partial P/\partial t$  by (5) and (7)

$$\begin{aligned} \frac{\partial P}{\partial t} &= \frac{1}{(h\pi)^n} \int \dots \int dy_1 \dots dy_n e^{2i(p_1 y_1 + \dots + p_n y_n)/\hbar} \\ &\cdot \left\{ \sum_k \frac{i\hbar}{2m_k} \left[ - \frac{\partial^2 \psi(x_1 + y_1, \dots, x_n + y_n)^*}{\partial x_k^2} \psi(x_1 - y_1, \dots, x_n - y_n) \right. \right. \\ &+ \psi(x_1 + y_1, \dots, x_n + y_n)^* \frac{\partial^2 \psi(x_1 - y_1, \dots, x_n - y_n)}{\partial x_k^2} \\ &+ \frac{i}{\hbar} [V(x_1 + y_1, \dots, x_n + y_n) \\ &\left. \left. - V(x_1 - y_1, \dots, x_n - y_n)] \psi(x_1 + y_1, \dots, x_n + y_n)^* \psi(x_1 - y_1, \dots, x_n - y_n) \right\}. \end{aligned} \tag{9}$$

Here one can replace the differentiations with respect to  $x_k$  by differentiations with respect to  $y_k$  and perform in the first two terms one partial integration with respect to  $y_k$ . In the last term we can develop  $V(x_1 + y_1, \dots, x_n + y_n)$  and  $V(x_1 - y_1, \dots, x_n - y_n)$  in a Taylor series with respect to the  $y$  and get

$$\begin{aligned} \frac{\partial P}{\partial t} &= \frac{1}{(\pi\hbar)^n} \int \dots \int dy_1 \dots dy_n e^{2i(p_1 y_1 + \dots + p_n y_n)/\hbar} \\ &\cdot \left\{ \sum_k \frac{p_k}{m_k} \left[ - \frac{\partial \psi(x_1 + y_1, \dots, x_n + y_n)^*}{\partial y_k} \psi(x_1 - y_1, \dots, x_n - y_n) \right. \right. \end{aligned}$$

$$\begin{aligned}
& + \psi(x_1 + y_1, \dots, x_n + y_n) \left. \frac{\partial \psi(x_1 - y_1, \dots, x_n - y_n)}{\partial y_k} \right] \\
& + \frac{i}{h} \sum_{\lambda} \frac{\partial^{\lambda_1 + \dots + \lambda_n} V}{\partial x_1^{\lambda_1} \dots \partial x_n^{\lambda_n}} \frac{y_1^{\lambda_1} \dots y_n^{\lambda_n}}{\lambda_1! \dots \lambda_n!} \psi(x_1 + y_1, \dots, x_n + y_n)^* \\
& \cdot \psi(x_1 - y_1, \dots, x_n - y_n) \left. \right\}, \tag{10}
\end{aligned}$$

which is identical with (8) if one replaces now the differentiations with respect to  $y_k$  by differentiations with respect to  $x_k$ . Of course, (8) is legitimate only if it is possible to develop the potential energy  $V$  in a Taylor series.

Eq. (8) shows the close analogy between the probability function of the classical mechanics and our  $P$ : indeed the equation of continuity

$$\frac{\partial P}{\partial t} = - \sum_k \frac{p_k}{m_k} \frac{\partial P}{\partial x_k} + \sum_k \frac{\partial V}{\partial x_k} \frac{\partial P}{\partial p_k}$$

differs from (8) only in terms of at least the second power of  $h$  and at least the third derivative of  $V$ . Expression (8) is even identical with the classical when  $V$  has no third and higher derivatives as, e.g., in a system of oscillators.

There is an alternative form for  $\partial P/\partial t$ , which however will not be used later on. It is

$$\frac{\partial}{\partial t} P(x_1, \dots, x_n; p_1, \dots, p_n) = - \sum_k \frac{p_k}{m_k} \frac{\partial}{\partial x_k} P(x_1, \dots, x_n; p_1, \dots, p_n) \tag{11}$$

$$+ \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} dj_1 \dots dj_n P(x_1, \dots, x_n; P_1 + j_1, \dots, P_n + j_n) J(x_1, \dots, x_n; j_1, \dots, j_n)$$

where  $J(x_1, \dots, x_n; j_1, \dots, j_n)$  can be interpreted as the probability of a jump in the momenta with the amounts  $j_1, \dots, j_n$  for the configuration  $x_1, \dots, x_n$ . The probability of this jump is given by

$$\begin{aligned}
& J(x_1, \dots, x_n; j_1, \dots, j_n) \\
& = \frac{i}{\pi^n h^{n+1}} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} dy_1 \dots dy_n [V(x_1 + y_1, \dots, x_n + y_n) \\
& \quad - V(x_1 - y_1, \dots, x_n - y_n)] e^{-(2i/h)(y_1 j_1 + \dots + y_n j_n)} \tag{11a}
\end{aligned}$$

that is, by the Fourier expansion coefficients of the potential  $V(x_1, \dots, x_n)$ . This form clearly shows the quantum mechanical nature of our  $P$ : the momenta change discontinuously by amounts which would be half the momenta of light quanta if the potential were composed of light.<sup>2a</sup> To derive (11) one can insert both for  $P$  and  $J$  their respective values (5) and (11a) on the right hand side of (11). In the first term one can replace  $p_k e^{2i(p_1 y_1 + \dots + p_n y_n)/h}$  by

<sup>2a</sup> Cf. F. Bloch, Zeits. f. Physik **52**, 555 (1929).

$(\hbar/2i)(\partial/\partial y_k)e^{2i(p_1y_1+\dots+p_ny_n)/\hbar}$  and then perform a partial integration with respect to  $y_k$ . Then one can replace the differentiation with respect to  $y$  by differentiation with respect to  $x$ , upon which some terms cancel and the rest goes over to

$$\sum_k \frac{\hbar}{2im} \int \dots \int dy_1 \dots dy_n \left[ \frac{\partial^2 \psi(x_1+y_1, \dots, x_n+y_n)^*}{\partial x_k^2} \psi(x_1-y_1, \dots, x_n-y_n) - \psi(x_1+y_1, \dots, x_n+y_n) \frac{\partial^2 \psi(x_1-y_1, \dots, x_n-y_n)}{\partial x_k^2} \right] e^{2i(p_1y_1+\dots+p_ny_n)/\hbar} \quad (12)$$

which is just what we need for the left side of (11). By integrating the second term on the right side of (11)

$$\begin{aligned} & \int \dots \int dy_1 \dots dy_n \psi(x_1+y_1 \dots x_n+y_n)^* \psi(x_1-y_1 \dots x_n-y_n) \\ & \cdot \int \dots \int dj_1 \dots dj_n e^{2i(\hbar)^{-1}[(p_1+i_1)y_1+\dots+(p_n+i_n)y_n]} \\ & \cdot \frac{i}{\pi^n \hbar^{n+1}} \int \dots \int dz_1 \dots dz_n [V(x_1+z_1 \dots x_n+z_n) \\ & - V(x_1-z_1 \dots x_n-z_n)] e^{-2i(z_1i_1+\dots+z_ni_n)/\hbar} \end{aligned}$$

with respect to  $z$  and  $j$  one gets because of the Fourier theorem<sup>3</sup>

$$(i/\hbar) \int \dots \int dy_1 \dots dy_n \psi(x_1+y_1 \dots x_n+y_n)^* \psi(x_1-y_1 \dots x_n-y_n) e^{2i(p_1y_1+\dots+p_ny_n)/\hbar} [V(x_1+y_1 \dots x_n+y_n) - V(x_1-y_1 \dots x_n-y_n)] \quad (12a)$$

and this gives the second part of the left side of (11).

3

So far we have defined only a probability function for pure states, which gives us the correct expectation values for quantities  $f(p_1 \dots p_n) + g(x_1 \dots x_n)$ . If, however, we have a mixture,<sup>4</sup> e.g., the pure states  $\psi_1, \psi_2, \psi_3, \dots$  with the respective probabilities  $w_1, w_2, w_3, \dots$  (with  $w_1+w_2+w_3+\dots=1$ ) the normal probability calculation suggests a probability function

$$P(x_1, \dots, x_n, p_1, \dots, p_n) = \sum_{\lambda} w_{\lambda} P_{\lambda}(x_1, \dots, x_n, \dots, p_n) \quad (13)$$

where  $P_{\lambda}$  is the probability function for  $\psi_{\lambda}$ . This probability function gives obviously the correct expectation values for all quantities, for which (5) gives correct expectation values and therefore will be adopted.

For a system in statistical equilibrium at the temperature  $T=1/k\beta$  the relative probability of a stationary state  $\psi_{\lambda}$  is  $e^{-\beta E_{\lambda}}$  where  $E_{\lambda}$  is the energy of  $\psi_{\lambda}$ . Therefore the probability function is a part from a constant

<sup>3</sup> Cf. e. g., R. Courant und D. Hilbert, Methoden der mathematischen Physik I. Berlin 1924. p. 62, Eq. (29).

<sup>4</sup> J. v. Neumann, Gött Nachr. 245, 1927. L. Landau, Zeits. f. Physik 45, 430 (1927).

$$\begin{aligned}
P(x_1 \cdots x_n; p_1 \cdots p_n) \\
= \sum_{\lambda} \int \cdots \int dy_1 \cdots dy_n \psi_{\lambda}(x_1 + y_1 \cdots x_n + y_n)^* \\
e^{-\beta E_{\lambda}} \psi(x_1 - y_1 \cdots x_n - y_n) e^{2i(p_1 y_1 + \cdots + p_n y_n)/\hbar}. \quad (14)
\end{aligned}$$

Now

$$\sum_{\lambda} \psi_{\lambda}(u_1 \cdots u_n)^* f(E_{\lambda}) \psi_{\lambda}(v_1 \cdots v_n)$$

is that matrix element of the operator  $f(H)$ , ( $H$  is the energy operator) which is in the  $u_1 \cdots u_n$  row and  $v_1 \cdots v_n$  column. Therefore (14) may be written as

$$\begin{aligned}
P(x_1 \cdots x_n; p_1 \cdots p_n) \\
= \int_{-\infty}^{\infty} \cdots \int dy_1 \cdots dy_n e^{i[(x_1 + y_1)p_1 + \cdots + (x_n + y_n)p_n]/\hbar} [e^{-\beta H}]_{x_1 + y_1 \cdots x_n + y_n; x_1 - y_1 \cdots x_n - y_n} \\
\cdot e^{-i[(x_1 - y_1)p_1 + \cdots + (x_n - y_n)p_n]/\hbar}. \quad (15)
\end{aligned}$$

so that we have under the integral sign the  $x_1 + y_1 \cdots x_n + y_n$ ;  $x_1 - y_1 \cdots x_n - y_n$  element of the matrix  $e^{-\beta H}$  transformed by the diagonal matrix  $e^{i(p_1 x_1 + \cdots + p_n x_n)/\hbar}$ . Instead of transforming  $e^{-\beta H}$  we can transform  $H$  first and then take the exponential with the transformed expression. By transforming  $H$  we get the operator (the  $p$  are numbers, not operators!)

$$H = e^{i(x_1 p_1 + \cdots + x_n p_n)/\hbar} \left( - \sum_k \frac{\hbar^2}{2m_k} \frac{\partial^2}{\partial x_k^2} + V(x_1 \cdots x_n) \right) e^{-i(x_1 p_1 + \cdots + x_n p_n)/\hbar}$$

which is equal to

$$\tilde{H} = \epsilon + \sum_{k=1}^n \left( \frac{i\hbar p_k}{m_k} \frac{\partial}{\partial x_k} - \frac{\hbar^2}{2m_k} \frac{\partial^2}{\partial x_k^2} \right) \quad (16)$$

where

$$\epsilon = \sum_{k=1}^n \frac{p_k^2}{2m_k} + V(x_1, \cdots, x_n). \quad (17)$$

So we get for (15)

$$\begin{aligned}
P(x_1, \cdots, x_n; p_1, \cdots, p_n) \\
= \int \cdots \int dy_1 \cdots dy_n [e^{-\beta \tilde{H}}]_{x_1 + y_1 \cdots x_n + y_n; x_1 - y_1 \cdots x_n - y_n}. \quad (18)
\end{aligned}$$

By calculating the mean value of a quantity  $Q = f(p_1, \cdots, p_n) + g(x_1, \cdots, x_n)$  by (18) one has to obtain the same result as by using the original expression (4) of Neumann.

If we are dealing with a system, the behavior of which in statistical equilibrium is nearly correctly given by the classical theory, we can expand (18) into a power of  $\hbar$  and keep the first few terms only. The term with the zero power of  $\hbar$  is  $\sum_r (-\beta)^r \epsilon^r / r!$  Now  $\epsilon^r$  is the operator of multiplication with the  $r$

power of (17). Its  $x_1+y_1, \dots, x_n+y_n; x_1-y_1, \dots, x_n-y_n$  element is consequently

$$\epsilon(x_1 + y_1, \dots, x_n + y_n)^r \delta(x_1 + y_1, x_1 - y_1) \dots \delta(x_n + y_n, x_n - y_n).$$

As  $\delta$  (also  $\delta', \delta'', \dots$ ) only depends on the difference of its two arguments, one can write  $\delta(-2y_1) \dots \delta(-2y_n)$  for the last factors and perform the integration by introducing  $-2y_1, \dots, -2y_n$  as new variables. The terms with the zero power of  $\hbar$ , arising from the first part of (16) only, give thus

$$(1/2^n) \sum_r (-\beta)^r \epsilon(x_1, \dots, x_n)^r / r! = e^{-\beta\epsilon} / 2^n \tag{19}$$

which is just the classical expression.

The higher approximations of the probability function can be calculated in a very similar way. The terms of  $e^{-\beta\tilde{H}}$ , involving the first power of the second part of  $\tilde{H}$  only, are

$$\sum_{r=0}^{\infty} \frac{(-\beta)^r}{r!} \sum_{\rho=1}^r e^{\rho-1} \sum_k \left( \frac{i\hbar p_k}{m_k} \frac{\partial}{\partial x_k} - \frac{\hbar^2}{2m_k} \frac{\partial^2}{\partial x_k^2} \right) \epsilon^{r-\rho} \tag{20}$$

By replacing all operators by symbolic integral-kernels one gets for the  $x_1+y_1, \dots, x_n+y_n; x_1-y_1, \dots, x_n-y_n$  element of the operator (20)

$$\begin{aligned} & \sum_r \frac{(-\beta)^r}{r!} \sum_{\rho=1}^r \epsilon(x_1 + y_1, \dots, x_n + y_n)^{\rho-1} \\ & \cdot \sum_k \left[ \frac{i\hbar p_k}{m_k} \delta(-2y_1) \dots \delta'(-2y_k) \dots \delta(-2y_n) \right. \\ & \left. - \frac{\hbar^2}{2m_k} \delta(-2y_1) \dots \delta''(-2y_k) \dots \delta(-2y_n) \right] \epsilon(x_1 - y_1, \dots, x_n - y_n)^{r-\rho}. \end{aligned}$$

Now

$$\sum_{\rho=1}^r \epsilon_+^{\rho-1} \epsilon_-^{r-\rho} = \sum_{\rho=1}^r \epsilon_-^{r-1} \left( \frac{\epsilon_+}{\epsilon_-} \right)^{\rho-1} = \frac{\epsilon_+^r - \epsilon_-^r}{\epsilon_+ - \epsilon_-}$$

so that the summation over  $\rho$  and  $r$  can be performed in (21). By introducing again new variables  $w_1, \dots, w_n$  for  $-2y_1, \dots, -2y_n$  and performing the integration one has

$$\frac{1}{2^n} \sum_k \left[ \frac{i\hbar p_k}{m_k} \frac{\partial}{\partial w_k} - \frac{\hbar^2}{2m_k} \frac{\partial^2}{\partial w_k^2} \right] \frac{e^{-\beta\epsilon(x_1, \dots, x_k - w_k/2, \dots, x_n)} - e^{-\beta\epsilon(x_k, \dots, x_k + w_k/2, \dots, x_n)}}{\epsilon(x_1, \dots, x_k - \frac{1}{2}w_k, \dots, x_n) - \epsilon(x_1, \dots, x_k + \frac{1}{2}w_k, \dots, x_n)}$$

where  $w_k=0$  must be inserted after differentiation. The first differential quotient vanishes at  $w_k=0$ , as the expression to be differentiated is an even function of  $w_k$ . The second part gives

$$\frac{e^{-\beta\epsilon}}{2^n} \sum_k \frac{h^2}{m_k} \left( -\frac{\beta^2}{8} \frac{\partial^2 \epsilon}{\partial x_k^2} + \frac{\beta^3}{24} \left( \frac{\partial \epsilon}{\partial x_k} \right)^2 \right). \quad (21)$$

In principle it is possible to calculate in the same way the terms involving the higher powers of the second part of  $\tilde{H}$  also, the summation over  $r$  and the quantities corresponding to our  $\rho$  can always be performed in a very similar way. In practice, however, the computation becomes too laborious. Still it is clear, that if we develop our probability function for thermal equilibrium in a power series of  $h$

$$P(x_1, \dots, x_n; p_1, \dots, p_n) = e^{-\beta\epsilon} + hf_1 + h^2 f_2 + \dots \quad (22)$$

(we can omit the factor  $1/2^n$  before  $e^{-\beta\epsilon}$ , as we are dealing with relative probabilities anyway) all terms will be quite definite functions of the  $p$ ,  $V$  and the different partial derivatives of the latter. Furthermore it is easy to see, that  $f_k$  will not involve higher derivatives of  $V$  than the  $k$ -th nor higher powers of  $p$  than the  $k$ -th. These facts enable us to calculate the higher terms of (22) in a somewhat simpler way, than the direct expansion of (18) would be.

The state (22) is certainly stationary, so that it would give identically  $\partial P/\partial t = 0$  when inserted into (8). By equating the coefficients of the different powers of  $h$  in  $\partial P/\partial t$  to zero one gets the following equations:

$$\sum_k -\frac{p_k}{m_k} \frac{\partial e^{-\beta\epsilon}}{\partial x_k} + \sum_k \frac{\partial V}{\partial x_k} \frac{\partial e^{-\beta\epsilon}}{\partial p_k} = 0 \quad (23, 0)$$

$$\sum_k -\frac{p_k}{m_k} \frac{\partial f_1}{\partial x_k} + \sum_k \frac{\partial V}{\partial x_k} \frac{\partial f_1}{\partial p_k} = 0 \quad (23, 1)$$

$$\begin{aligned} \sum_k -\frac{p_k}{m_k} \frac{\partial f_2}{\partial x_k} + \sum_k \frac{\partial V}{\partial x_k} \frac{\partial f_2}{\partial p_k} - \sum_k \frac{\partial^3 V}{\partial x_k^3} \frac{h^2}{24} \frac{\partial^3 e^{-\beta\epsilon}}{\partial p_k^3} \\ - \sum_{k \neq l} \frac{\partial^3 v}{\partial x_k^2 \partial x_l} \frac{h^2}{8} \frac{\partial^2 e^{-\beta\epsilon}}{\partial p_k^2 \partial p_l} = 0 \end{aligned} \quad (23, 2)$$

and so on. The first of these equations is an identity because of (17), as it must be; (23, a), (23, 2),  $\dots$  will determine  $f_1, f_2, \dots$  respectively. All Eqs. (23, a) are linear inhomogeneous partial differential equations for the unknown  $f$ . From one solution  $f_a$  of (23, a) one obtains the general solution by adding to it the general solution  $F$  of the homogeneous part of (23, a), which is always

$$\sum_k -\frac{p_k}{m_k} \frac{\partial F}{\partial x_k} + \sum_k \frac{\partial V}{\partial x_k} \frac{\partial F}{\partial p_k} = 0.$$

This equation in turn is the classical equation for the stationary character of the probability distribution  $F(x_1, \dots, x_n; p_1, \dots, p_n)$ . It has in general only one solution which contains only a finite number of derivatives of  $V$ , namely

$$F(x_1, \dots, x_n; p_1, \dots, p_n) = F \left( \sum_k \frac{p_k^2}{2m_k} + v(x_1 \dots x_n) \right) = F(\epsilon).$$

In fact, if it had other integrals, like



$$F(p_1, \dots, p_n; V, \partial V/\partial x_1, \partial V/\partial x_2, \dots) \tag{24}$$

then all mechanical problems would have in addition to the energy-integral further integrals of the form (24) which, of course, is not true.

One solution of (23, 1) is  $f_1=0$  and the most general we have to consider is therefore  $f_1=F(\epsilon)$ . We have to take however  $F(\epsilon)=0$  as  $f_1$  has to vanish for a constant  $V$ . So we get  $f_1=0$ , as we know it already from the direct expansion of (18). The same holds consequently for  $f_3, f_5, \dots$ , as the inhomogeneous part of the equation for  $f_3$  only contains  $f_1$ , the inhomogeneous part of the equation for  $f_5$  only  $f_1$  and  $f_3$ , and so on.

For  $f_2$  one easily gets

$$f_2 = e^{-\beta\epsilon} \left[ \sum_k \left( -\frac{\beta^2}{8m_k} \frac{\partial^2 V}{\partial x_k^2} + \frac{\beta^3}{24m_k} \left( \frac{\partial V}{\partial x_k} \right)^2 \right) + \sum_{k,l} \frac{\beta^3 p_k p_l}{24m_k m_l} \frac{\partial^2 V}{\partial x_k \partial x_l} \right] \tag{25}$$

as a solution of (23, 2) and it is also clear, that this is the solution we need. The first two terms of  $f_2$  we have already directly computed (21), the third arises from terms with the second power of the second part of  $\tilde{H}$ . Similarly  $f_4$  is for one degree of freedom ( $n=1$ )

$$\begin{aligned} 64m^2\beta^{-2} e^{\beta\epsilon} f_4 = & H_4(q) [\beta^2 V''^2/72 - \beta V''''/120] \\ & + H_2(q) [\beta^3 V'^2 V''/18 - 2\beta^2 V''^2/15 - \beta^2 V' V''''/15 + \beta V''''/15] \\ & + H_0(q) [\beta^4 V'^4/18 - 22\beta^3 V'^2 V''/45 + 2\beta^2 V''^2/5 + 8\beta^2 V' V''''/15 \\ & - 4\beta V''''/15] \end{aligned} \tag{26}$$

where  $H_r$  is the  $r$ -th Hermitean polynomial and  $q = \beta^{1/2} \dot{p} / (2m)^{1/2}$ .

It does not seem to be easy to get a simple closed expression for  $f_k$ , but it is quite possible to calculate all of them successively. A discussion of Eqs. (23) shows, that the  $g$  in

$$P(x_1, \dots, x_n; p_1, \dots, p_n) = e^{-\beta\epsilon} (1 + h^2 g_2 + h^4 g_4 + \dots) \tag{27}$$

are rational expressions in the derivatives of  $V$  only (do not contain  $V$  itself) and all terms of  $g_k$  contain  $k$  differentiations and as functions of the  $p$  are polynomials of not higher than the  $k$ -th degree. The first term in (27) with the zero power of  $h$  is the only one, which occurs in classical theory. There is no term with the first power, so that if one can develop a property in a power series with respect to  $h$ , the deviation from the classical theory goes at least with the second power of  $h$  in thermal equilibrium. One familiar example for this is the inner energy of the oscillator, where the term with the first power of  $h$  vanishes just in consequence of the zero point energy. The second term can be interpreted as meaning that a quick variation of the probability function with the coordinates is unlikely, as it would mean a quick variation, a short wave-length, in the wave functions. This however would have the consequence of a high kinetic energy. The quantum mechanical probability is therefore something like the integral of the classical expression  $e^{-\beta\epsilon}$  over a finite range of coordinates of the magnitude  $\sim h/\bar{p}$  where  $\bar{p}$  is the mean momentum  $\sim (kTm)^{1/2}$ . The correction terms of (27) have, among other effects,

the consequence that the probability for a particle being in a narrow hole is smaller than would be in classical statistics. From now on we will keep only the first two terms of (27).

4

From (25) one easily calculates the relative probabilities of the different configurations by integration with respect to the  $p$ :

$$\begin{aligned} & \int \cdots \int dp_1 \cdots dp_n P(x_1 \cdots x_n; p_1 \cdots p_n) \\ &= e^{-\beta V} \left[ 1 - \frac{\hbar^2 \beta^2}{12} \sum_k \frac{1}{m_k} \frac{\partial^2 V}{\partial x_k^2} + \frac{\hbar^2 \beta^3}{24} \sum_k \frac{1}{m_k} \left( \frac{\partial V}{\partial x_k} \right)^2 \right]. \end{aligned} \quad (28)$$

Hence the mean potential energy is

$$\begin{aligned} \bar{V} &= \frac{\int V e^{-\beta V} dx}{\int e^{-\beta V} dx} + \frac{\hbar^2 \beta^2}{24} \frac{\int \sum_k \frac{1}{m_k} \frac{\partial^2 V}{\partial x_k^2} e^{-\beta V} dx \int V e^{-\beta V} dx}{\left( \int e^{-\beta V} dx \right)^2} \\ &+ \frac{\hbar^2 \beta}{24} \frac{\int \sum_k \frac{1}{m_k} \frac{\partial^2 V}{\partial x_k^2} (1 - \beta V) e^{-\beta V} dx}{\int e^{-\beta V} dx} \end{aligned} \quad (29)$$

where  $dx$  is written for  $dx_1 \cdots dx_n$  and the higher power terms of  $\hbar$  are omitted. Similarly the mean value of the kinetic energy is

$$\sum_k \frac{\overline{p_k^2}}{2m_k} = \frac{n}{2\beta} + \frac{\hbar^2 \beta}{24} \frac{\int \sum_k \frac{1}{m_k} \frac{\partial^2 V}{\partial x_k^2} e^{-\beta V} dx}{\int e^{-\beta V} dx}. \quad (30)$$

This formula also is correct only within the second power of  $\hbar$ ; in order to derive it one has to perform again some partial integrations with respect to the  $x$ . Eqs. (28), (29), (30) have a strict quantum mechanical meaning and it should be possible to derive them also from (4). One sees that the kinetic energy is in all cases larger than the classical expression  $\frac{1}{2}nkT$ .

5

One fact still needs to be mentioned. We assumed that the probability of a state with the energy  $E$  is given by  $e^{-\beta E}$ . This is not true in general, since the Pauli principle forbids some states altogether. The corrections thus introduced by the Bose or Fermi statistics even give terms with the first power of  $\hbar$ , so that it seems, that as long as one cannot take the Bose or Fermi statistics into account, Eq. (25) cannot be applied to an assembly of identical par-

ticles, as, e.g., a gas. There is reason to believe however, that because of the large radii of the atoms this is not true and the corrections due to Fermi and Bose statistics may be neglected for moderately low temperatures.

The second virial coefficient was first calculated in quantum mechanics by F. London on the basis of his theory of interatomic forces.<sup>5</sup> He also pointed out that quantum effects should be taken into account at lower temperatures. Slater and Kirkwood<sup>6</sup> gave a more exact expression for the interatomic potential of He and Kirkwood and Keyes<sup>7</sup> calculated on this basis the classical part of the second virial coefficient of He. H. Margenau<sup>8</sup> and Kirkwood<sup>9</sup> performed the calculations for the quantum-correction. The present author also tried to calculate it by the method just outlined. He got results, which differ from those of Margenau and Kirkwood in some cases by more than 100 percent.<sup>10</sup> It does not seem however to be easy to compare these results with experiment, as the classical part of the second virial coefficient is at low temperatures so sensitive to small variations of the parameters occurring in the expression of the interatomic potential, that it changes by more than 20 percent if the parameter in the exponential (2.43) is changed by  $\frac{1}{2}$  percent and it does not seem to be possible to determine the latter within this accuracy.

<sup>5</sup> F. London, *Zeits. f. Physik* **63**, 245 (1930).

<sup>6</sup> J. C. Slater and J. G. Kirkwood, *Phys. Rev.* **37**, 682 (1931).

<sup>7</sup> J. G. Kirkwood and F. G. Keyes, *Phys. Rev.* **38**, 516 (1931).

<sup>8</sup> H. Margenau, *Proc. Nat. Acad.* **18**, 56, 230 (1932). Cf. also J. C. Slater, *Phys. Rev.* **38**, 237 (1931).

<sup>9</sup> J. G. Kirkwood, *Phys. Zeits.* **33**, 39 (1932).

<sup>10</sup> I am very much indebted to V. Rojansky for his kind assistance with these calculations. The reason for the disagreement between our results and those of Margenau and Kirkwood may be the fact that they did not apply any corrections for the continuous part of the spectrum.

In a paper which appeared recently in the *Zeits. f. Physik* (**74**, 295 (1932)) F. Bloch gets results which are somewhat similar to those of the present paper. (*Note added at proof.*)