The Gibbs-Bogoliubov inequality dagger

1968 J. Phys. A: Gen. Phys. 1 539

(http://iopscience.iop.org/0022-3689/1/5/305)

J. PHYS. A (PROC. PHYS. SOC.), 1968, SER. 2, VOL. 1. PRINTED IN GREAT BRITAIN

The Gibbs–Bogoliubov inequality†

A. ISIHARA

Statistical Physics Laboratory, Department of Physics, State University of New York, Buffalo, N.Y., U.S.A.

MS. received 12*th February* 1968, *in revised form* 14*th May* 1968

Abstract. A simple inequality, expressed in terms of two arbitrary distribution functions of the same normalization, is shown to be useful. By choosing various different forms for the distribution functions one can derive important results, such as the upper and lower bounds of the configurational free energy.

1. Introduction

In statistical mechanics variational methods based on minimizing the free energy have been used widely. For many-body systems several other variational techniques, such as those based on quantum-mechanical principles of minimizing the energy, have been successfully developed (Feynman and Cohen 1956, Bogoliubov 1958, Kohn 1964, Mermin 1965).

As one of such variational principles, Peierls (1938) presented an inequality which gave a rigorous lower bound to the exact partition function, namely an upper bound to the free energy. The inequality is of particular importance when one tries to replace the Hamiltonian by its diagonal elements. Since then several investigations on the inequality have been reported. For instance, Husimi (1940) showed that the Peierls inequality was based on the convexity of an exponential function. Griffiths (1964) gave more extensive discussions on the convexity of the free-energy function and derived what he called Bogoliubov inequality (Tolmachev 1960, Muehlschlegel 1960, Girardeau 1962) for the free energy in terms of the maximum and minimum eigenvalues of the interaction Hamiltonian. The Bogoliubov and Peierls inequalities are closely related. In fact, by simply setting all of the off-diagonal elements of the Hamiltonian to zero one obtains the Peierls inequality from the Bogoliubov inequality.

On the other hand, one recalls that the H theorem is one of the basic principles for variational theories. Actually, there are several different H theorems, such as that based on the Boltzmann or on the Gibbs canonical distribution function. We are concerned with the H theorem for a canonical distribution and also with the H theorem based on coarse graining. The former H theorem states that the canonical distribution corresponds to minimum H, and the latter implies a decrease in H due to coarse graining.

The purpose of this paper is to show that these H theorems are particular cases of a more general theorem expressed by a simple mathematical inequality. This inequality involves two arbitrary distribution functions of the same normalization. Therefore by choosing the distribution functions suitably one obtains various important results, including the H theorems and the Bogoliubov inequality. One can also show in a particular case that an important part of the entropy of a physical system is negative or zero, if it is defined in terms of a distribution function satisfying a certain condition. Moreover, this negative property is related to the Gibbs–Bogoliubov inequality to be discussed in this paper. The new mathematical inequality holds irrespective of quantum or classical statistics. Thus some of the variational principles in statistical mechanics will be discussed from a unified point of view.

2. Mathematical inequality

For the simplest presentation of our starting point let us introduce x for the phase-space variable, which can be momenta or spatial coordinates and can even be discrete. Let us define two phase-space distribution functions $f(x)$ and $g(x)$, which are both positive and

† This work was supported by the National Science Foundation.

satisfy the same normalization condition (Isihara and Wu 1968):

$$\int f(x)\,\mathrm{d}x = \int g(x)\,\mathrm{d}x. \tag{2.1}$$

For these distribution functions the following inequality holds:

$$\int f(x)\ln f(x)\,\mathrm{d}x \geqslant \int f(x)\ln g(x)\,\mathrm{d}x. \tag{2.2}$$

This result may look strange since $g(x)$ and $f(x)$ are arbitrary. However, its validity can be proved following the method which Gibbs used for the H theorem (Gibbs 1902):

$$\int f(x)\ln f(x)\,\mathrm{d}x - \int f(x)\ln g(x)\,\mathrm{d}x$$
$$= \int g(x)\left\{\frac{f}{g}\ln\left(\frac{f}{g}\right)g - \frac{f}{g}\ln g - \frac{f}{g}+1\right\}\mathrm{d}x \geqslant 0.$$

It is important to notice in (2.2) that both $f(x)$ and $g(x)$ are arbitrary. Thus if $f(x)$ and $(g)x$ are both functions of energy and give the same average energy:

$$\int f(x)H(x)\,\mathrm{d}x = \int g(x)H(x)\,\mathrm{d}x$$
$$= \langle H(x)\rangle$$

and if $g(x)$ is the canonical distribution:

$$g(x) = \exp\{\beta(F-H)\}$$

then the inequality (2.2) expresses the Gibbs H theorem for canonical ensemble. In these expressions $\beta = 1/kT$, F is the free energy and H is the Hamiltonian. On the other hand, if $g(x)$ is a coarse-grained distribution function obtained from a fine-grained distribution function $f(x)$:

$$g(x) = \int_{\Delta_i} \frac{f(x)\,\mathrm{d}x}{\Delta_i x} \qquad \text{for } x \text{ in } \Delta_i\text{, the } i\text{th cell in phase space}$$

then (2.2) is reduced to the H theorem due to coarse graining.

Further applications of the inequality (2.2) are presented in this paper. In spite of the simple form it yields many useful and interesting results.

3. Linked cluster expansion

In this section we shall consider thermodynamic functions evaluated by a perturbation method. As is well known, a linked cluster expansion expresses a physical quantity such as the free energy in terms of the contributions from connected graphs. Based on the inequality (2.2) we shall derive interesting results for such an expansion.

Let us start with the Hamiltonian of the following general structure:

$$H = H_0 + H_1. \tag{3.1}$$

Correspondingly, we shall write the Helmholtz free energy as follows:

$$F = F_0 + F_1. \tag{3.2}$$

We now use equation (2.2) for the following two distribution functions:

$$\begin{aligned} f &= \exp\{\beta(F_0 - H_0)\} \\ g &= \exp\{\beta(F-H)\}. \end{aligned} \tag{3.3}$$

Both distribution functions are normalized to 1. From this condition we find

$$\exp(-\beta F_1) = \frac{\int \exp\{-\beta(H_0+H_1)\}\,dx}{\int \exp(-\beta H_0)\,dx}$$

$$= \langle \exp(-\beta H_1)\rangle_0. \tag{3.4}$$

Equation (3.4) may be expressed in terms of the Thiele semi-invariants:

$$-\beta F_1 = \sum_s c_s \beta^s \tag{3.5}$$

where

$$\begin{aligned} c_1 &= \langle -H_1\rangle_0 \\ c_2 &= \tfrac{1}{2}(\langle H_1^2\rangle_0 - \langle H_1\rangle_0^2) \\ c_3 &= \frac{1}{3!}\{\langle -H_1^3\rangle_0 - 3\langle -H_1\rangle_0\langle H_1^2\rangle_0 - 2\langle H_1\rangle_0^3\} \end{aligned} \tag{3.6}$$

and

$$\langle H_1^s\rangle_0 = \frac{\int \exp(-\beta H_0) H_1^s\,dx}{\int \exp(-\beta H_0)\,dx}. \tag{3.7}$$

The mathematical inequality (2.2) now states

$$F_0 - \langle H_0\rangle \geqslant \int (F-H)\exp\{\beta(F_0-H_0)\}\,dx. \tag{3.8}$$

In terms of the invariants this expression assumes the next form:

$$\begin{aligned} \frac{\beta}{2}(\langle H_1^2\rangle_0 - \langle H_1\rangle_0^2) \geqslant{}& \frac{\beta^2}{3!}(\langle H_1^3\rangle_0 - 3\langle H_1\rangle_0\langle H_1^2\rangle_0 + 2\langle H_1\rangle_0^3) \\ & - \frac{\beta^3}{4!}(\langle H_1^4\rangle_0 - 4\langle H_1\rangle_0\langle H_1^3\rangle_0 - 3\langle H_1^2\rangle_0^2 \\ & + 12\langle H_1\rangle_0^2\langle H_1^2\rangle_0 - 6\langle H_1\rangle_0^4) + \dots . \end{aligned} \tag{3.9}$$

Therefore the first correlation or the second Thiele invariant is larger than the sum of the rest in the linked cluster expansion.

Inequality (3.9) is the result of (2.2) and equation (3.3). Generally, it may be expressed in the following form:

$$\langle H_1\rangle_0 \geqslant F_1. \tag{3.10}$$

According to Tolmachev (1960) this inequality is due to Bogoliubov (1958). Muehlschlegel and Zittartz (1963) and Girardeau (1962) developed variational theories based on the inequality.

So far our discussions have been based on the choice of the distribution functions given by equation (3.3). We could, however, switch f and g in equation (3.3) as follows:

$$\begin{aligned} f &= \exp\{\beta(F-H)\} \\ g &= \exp\{\beta(F_0-H_0)\}. \end{aligned} \tag{3.11}$$

Using these distribution functions in inequality (2.2), we arrive at another interesting result:

$$F_1 \geqslant \langle H_1 \rangle \tag{3.12}$$

where $\langle H_1 \rangle$ is defined by

$$\langle H_1 \rangle = \int \exp\{\beta(F-H)\}H_1 \, \mathrm{d}x. \tag{3.13}$$

In particular, if H_0 and H_1 do not have common variables, such as in the case of the kinetic and potential energies, we can express H_1 as follows:

$$\langle H_1 \rangle = \int \exp\{\beta(F_1 - H_1)\}H_1 \, \mathrm{dx}. \tag{3.14}$$

Namely, $\langle H_1 \rangle$ is the configurational internal energy. In bringing $\langle H_1 \rangle$ of equation (3.13) into the form of (3.14) integrations over momentum variables may be required. Therefore, although we have used the same symbolic phase-space variable x, dx in equation (3.13) may be different from that in equation (3.14). Because of the normalization imposed on the distribution functions $f(x)$ and $g(x)$, such a difference in dx does not cause any difficulty.

We may now summarize observations. Combining inequalities (3.10) and (3.12), we arrive at

$$\langle H_1 \rangle_0 \geqslant F_1 \geqslant \langle H_1 \rangle. \tag{3.15}$$

Similar inequalities have been obtained by Griffiths (1964) by a different method. From now on we shall call the entire relation (3.15) the Gibbs–Bogoliubov inequality. As we remarked, the first of the inequalities in (3.15) is apparently due to Bogoliubov.

We have arrived at our result (3.15) by choosing the distribution functions in suitable ways. The distribution functions have to satisfy the condition (2.1). Therefore the question remains whether such distribution functions can actually be introduced in treating many-body problems. In what follows we shall discuss this question for several interesting cases.

4. Ising lattice

The Hamiltonian of a spin lattice may be written as follows:

$$H = -\tfrac{1}{2} \sum_{i \neq j} J_{ij}\mu_i\mu_j. \tag{4.1}$$

Let us express this Hamiltonian in momentum space and treat the spin lattice under a random-phase approximation. The Hamiltonian may be expressed in the form

$$H = -\frac{1}{2N} \sum_{i \neq j}' J(q) \exp\{\mathrm{i}\boldsymbol{q} \cdot (\boldsymbol{r}_i - \boldsymbol{r}_j)\}\mu_i\mu_j \tag{4.2}$$

where N is the total number of lattice sites. To include the $i = j$ term it is customary to assume (Horwitz and Callen 1961, Brout 1964)

$$\sum_q J(q) = 0 \tag{4.3}$$

corresponding to

$$J(r) = 0 \text{ for } r = 0. \tag{4.4}$$

Thus, defining the Fourier transform

$$\sum_i \exp(\mathrm{i}\boldsymbol{q} \cdot \boldsymbol{r}_i)\mu_i = \mu(q) \tag{4.5}$$

we can express the Hamiltonian in the following way:

$$H = -\tfrac{1}{2} \sum_q J(q)|\mu(q)|^2. \tag{4.6}$$

Equation (4.5) states that $\mu(q)$ is a sum of random variables about zero with the average

$$\langle|\mu(q)|^2\rangle_0 = 1 \tag{4.7}$$

where the suffix corresponds to the non-interacting Ising lattice. Therefore, by the central limit theorem which is expected to hold for $N \to \infty$, we can introduce the distribution function

$$f(x, y) = \frac{\left[\prod_q{}' \exp\{-x(q)\}\right] \exp\{-\frac{1}{2}y^2(q)\}}{(\frac{1}{2}\pi)^{1/2}} \tag{4.8}$$

where

$$\begin{aligned} x(q) &= |\mu(q)|^2 \qquad (q \neq 0) \\ y(q) &= \mu_0. \end{aligned} \tag{4.9}$$

The distribution function gives the averages:

$$\begin{aligned} \langle|\mu(q)|^{2n}\rangle_0 &= n! \\ \langle{\mu_0}^{2n}\rangle_0 &= (2n-1)(2n-3)\dots 1 \end{aligned} \tag{4.10}$$

in agreement with equation (4.7) in the special case. The distribution function thus defined is normalized as follows:

$$\int_0^\infty \int_0^\infty f(x, y)\, \mathrm{d}x\, \mathrm{d}y = 1. \tag{4.11}$$

We now try to introduce another distribution function $g(x, y)$ for comparison. This distribution function must satisfy the same condition as (4.11):

$$\int_0^\infty \int_0^\infty g(x, y)\, \mathrm{d}x\, \mathrm{d}y = 1. \tag{4.12}$$

We note that, in terms of the expressions in the previous section, we are treating the case where $F_0 = 0$. Thus we choose

$$g(x, y) = \exp(\beta F) \prod_q{}' [\exp\{-x(q) + \beta J(q) x(q)\}]\, \frac{\exp\{-\frac{1}{2}y^2 + \frac{1}{2}\beta J(0) y^2\}}{(\frac{1}{2}\pi)^{1/2}}. \tag{4.13}$$

From the normalization condition (4.12) we find

$$\begin{aligned} \exp(-\beta F) &= \prod_q{}' \int_0^\infty \exp\{-x(q)\} \exp\{\beta J(q) x(q)\}\, \mathrm{d}x \int_0^\infty \frac{\exp\{-\frac{1}{2}y^2 + \frac{1}{2}\beta J(0) y^2\}}{(\frac{1}{2}\pi)^{1/2}}\, \mathrm{d}y \\ &= \prod_q{}' \langle \exp\{\beta J(q) x(q)\}\rangle \langle \exp\{\tfrac{1}{2}\beta J(0) y^2\}\rangle. \end{aligned} \tag{4.14}$$

Using equation (4.14) we arrive at

$$\begin{aligned} \beta F &= \sum_q{}' \ln\{1 - \beta J(q)\} + 2^{-1} \ln\{1 - \beta J(0)\} \\ &= \tfrac{1}{2} \sum_q \ln\{1 - \beta J(q)\}. \end{aligned} \tag{4.15}$$

The energy is obtained as follows:

$$U = -\tfrac{1}{2} \sum_q \frac{J(q)}{1 - \beta J(q)}. \tag{4.16}$$

The short-range correlations of the spins are taken into consideration in this expression. This result may be compared with a simple arithmetic average of equation (4.6)

$$\langle H_1\rangle_0 = -\tfrac{1}{2}\sum_q J(q). \tag{4.17}$$

We observe that the internal energy U under the random-phase approximation is lower than $\langle H_1\rangle_0$, and that

$$\langle H_1\rangle_0 \geqslant F \geqslant \langle H_1\rangle \tag{4.18}$$

in conformity with inequality (3.15). More explicitly, (4.18) is

$$-\tfrac{1}{2}\sum_q \beta J(q) \geqslant \tfrac{1}{2}\sum_q \ln\{1-\beta J(q)\} \geqslant -\tfrac{1}{2}\sum_q \frac{\beta J(q)}{1-\beta J(q)}. \tag{4.19}$$

This inequality holds for $\beta J(q) < 1$. The equality corresponds to the case $\beta = 0$, when the spin lattice is completely at random.

Equations (4.16) and (4.6) give

$$\langle|\mu(q)|^2\rangle = \frac{1}{1-\beta J(q)} \tag{4.20}$$

in contrast with equation (4.7). Thus the theory based on the random-phase approximation is an improvement on that which neglects the spin correlations.

5. Imperfect gas

As in the previous section, there are many physical systems which, under certain situations, obey the normal distribution. There are also many cases where the phase space extends to infinity. For the discussion of such cases let us choose the distribution function

$$g(x) = \left(\frac{\lambda}{\pi}\right)^{1/2} \exp(-\lambda x^2). \tag{5.1}$$

From the inequality (2.2) we derive

$$\int f \ln f \,\mathrm{d}x \geqslant -\lambda\langle x^2\rangle + 2^{-1}\ln\left(\frac{\lambda}{\pi}\right). \tag{5.2}$$

Here we have used the normalization

$$\int f(x)\,\mathrm{d}x = 1. \tag{5.3}$$

Therefore, if we use the entropy defined by

$$S = -k\int f\ln f\,\mathrm{d}x \tag{5.4}$$

we arrive at

$$S \leqslant \lambda k\langle x^2\rangle - \tfrac{1}{2}k\ln\left(\frac{\lambda}{\pi}\right). \tag{5.5}$$

The right-hand side is a minimum when λ satisfies

$$\lambda = \frac{1}{2\langle x^2\rangle}. \tag{5.6}$$

The inequality (5.5) for this case is

$$S \leqslant \tfrac{1}{2}k(1+\ln 2\pi\langle x^2\rangle). \tag{5.7}$$

Essentially the same result may be obtained if many variables are included in the distribution function. We find

$$S \leqslant \sum_i \tfrac{1}{2}k(1+\ln 2\pi\langle x_i^2\rangle). \tag{5.8}$$

We observe that the entropy is a maximum for

$$2\pi\langle x_i^2\rangle = 1.$$

This is the case when the equipartition law holds for the independent variables x_i.

So far we have used only the normal distribution (5.1) in (2.2). Let us now introduce another distribution function $f(x)$ by

$$f(x) = g(x)\exp\{\beta(F_c-\Phi)\} \tag{5.9}$$

where Φ is the total potential energy and F_c is the free energy. Then inequality (2.2) yields

$$\int (F_c-\Phi)f(x)\,dx \geqslant 0 \tag{5.10}$$

or

$$F_c \geqslant U_c \tag{5.11}$$

where c stands for configuration.

We now extend our consideration to a system composed of N interacting gas particles. First, we introduce the distribution function $g(x)$ corresponding to an ideal gas:

$$g(x) = \prod_i \left(\frac{\lambda_i}{\pi}\right)^{1/2} \exp(-\lambda_i x_i^2) \tag{5.12}$$

where

$$x_i = \begin{cases} \boldsymbol{p}_i/h & (\text{for } i = 1, 2, \ldots, N) \\ \boldsymbol{r}_i/V & (\text{for } i = N+1, \ldots, 2N). \end{cases} \tag{5.13}$$

Here $\boldsymbol{p}_i$ represent the momenta, h Planck's constant, $\boldsymbol{r}_i$ the spatial coordinates and V is the total volume. Furthermore, we introduce the notation which we have used before:

$$g(x) = \exp\{\beta(F_0-H_0)\} \tag{5.14}$$

where F_0 is the Helmholtz free energy and H_0 is the kinetic energy. The normalization condition for $g(x)$ is

$$\int g(x)\,dx_1 \ldots dx_{2N} = \int g(x)\,dx_1 \ldots dx_N = 1.$$

As we remarked in connection with equations (3.13) and (3.14), we have used here the same symbolic notation $g(x)$ on both sides. As in equation (3.4), we have the relation

$$\exp(-\beta F_c) = \int \exp(-\beta\Phi)\,dx_{N+1} \ldots dx_{2N}. \tag{5.15}$$

Except for the factor $N!$, this agrees with the standard definition of the configurational free energy. Equation (5.15) is the result of the normalization to 1 of the distribution function:

$$f(x) = g(x)\exp\{\beta(F_c-\Phi)\}. \tag{5.16}$$

In terms of this (total) distribution function we can define the average of the potential energy:

$$\int f(x)\Phi\,dx_1 \ldots dx_{2N} = \int \exp\{\beta(F_c-\Phi)\}\Phi\,dx_{N+1} \ldots dx_{2N}$$

$$= U_c. \tag{5.17}$$

Again, in these expressions the suffix c stands for configuration.

The configurational free energy and the internal energy have been evaluated in terms of the irreducible integrals. As is well known, the results are

$$F_c = N\sum_k \frac{1}{k+1}\left(-\frac{\beta_k}{\beta}\right) v^{-k}$$
$$U_c = N\sum_k \frac{1}{k+1}\left(-\frac{\partial\beta_k}{\partial\beta}\right) v^{-k} \tag{5.18}$$

where v is the specific volume and the β_k are irreducible integrals. Thus, applying our result (5.11), we find an interesting result:

$$\sum_k \frac{1}{k+1}\frac{\beta_k}{\beta} v^{-k} \leqslant \sum_k \frac{1}{k+1}\frac{\partial\beta_k}{\partial\beta} v^{-k}. \tag{5.19}$$

The inequality (5.19) is satisfied if

$$\frac{\beta_k}{\beta} \leqslant \frac{\partial\beta_k}{\partial\beta}. \tag{5.20}$$

This condition holds in the case when the irreducible integrals vary with temperature as β^ϵ when ϵ is larger than 1. This means that the irreducible integrals may have a non-linear β dependence.

Let us examine this observation in the particular case of a van der Waals gas. In this case only β_1 enters the theory. Thus we find

$$\frac{\beta_1}{\beta} \leqslant \frac{\partial\beta_1}{\partial\beta}. \tag{5.21}$$

In the standard units and notation the first irreducible integral β_1 for a van der Waals gas is

$$\beta_1 = 2(a\beta - b). \tag{5.22}$$

Therefore the condition (5.21) is reduced to

$$b > 0 \tag{5.23}$$

which is certainly correct because b is four times the molecular volume.

It is known that at high temperatures the second virial coefficient given by equation (5.22) is moderately close to experiments. At low temperatures β_1 increases more rapidly with increasing temperature than is indicated by the expression. The third virial coefficient β_2 seems generally to be positive and considerably larger than $\beta_1{}^2$, indicating that its temperature dependence is also strong. β_3 is larger and varies more rapidly with temperature than $\beta_1{}^3$ or $\beta_2{}^{3/2}$ which have the same dimension. Concerning higher irreducible coefficients and the irreducible series itself, not much is known. In this situation the relation (5.19) will serve as a general criterion which the irreducible series should satisfy. It may be used particularly when the virial series is evaluated by a certain approximate method.

6. Quantum statistics

Our considerations in the previous sections can easily be extended to quantum-mechanical many-body systems. For the discussion let us use the grand ensemble. We introduce the grand potential by

$$\exp(-\beta\Omega) = \operatorname{tr}\exp\{-\beta(H-\mu N)\} \tag{6.1}$$

and the normalized distribution f by

$$f = \exp\{\beta(\Omega - H + \mu N)\}. \tag{6.2}$$

Once again we consider the case

$$\begin{aligned} H &= H_0 + H_1 \\ \Omega &= \Omega_0 + \Omega_1. \end{aligned} \tag{6.3}$$

As is well known, the grand potential is obtained from

$$\exp(-\beta\Omega) = \mathrm{tr}[\exp\{-\beta(H_0 - \mu N)\}U(\beta)] \tag{6.4}$$

or from

$$\begin{aligned} \exp\{-\beta(\Omega - \Omega_0)\} &= \langle U(\beta) \rangle_0 \\ &= 1 + \sum_n \frac{(-)^n}{n!} \int_0^\beta \langle P\{H_1(x_1) \ldots H(x_n)\} \, \mathrm{d}x_1 \ldots \mathrm{d}x_n \rangle_0 \end{aligned} \tag{6.5}$$

where P is Dyson's chronological operator.

Using the notation of the second quantization, the grand potential of equation (6.5) may be expanded in a linked cluster series:

$$\begin{aligned} -\beta(\Omega - \Omega_0) = \sum_{n=1}^{\infty} \frac{(-)^n}{n!(2V)^n} \sum \prod_{i=1}^{n} (r_i s_i|\phi|r_i' s_i') \\ \times \int_0^\beta \langle P\{a_{r_1}{}^*(x_1) a_{s_1}{}^*(x_1) a_{s_1'}(x_1) a_{r_1'}(x_1) \ldots\} \rangle_0. \end{aligned} \tag{6.6}$$

Here ϕ is a pair potential, and the Hamiltonian has been taken in the following form:

$$\begin{aligned} H_0 &= \sum_r \epsilon_r a_r{}^* a_r \\ H_1 &= \frac{1}{2V} \sum_{\substack{rs \\ r's'}} (rs|\phi|r's') a_r{}^* a_s{}^* a_{s'} a_{r'}. \end{aligned} \tag{6.7}$$

By this notation we arrive at

$$\langle H_1 \rangle_0 \geqslant \Omega_1 \geqslant \langle H_1 \rangle \tag{6.8}$$

where

$$\begin{aligned} \langle H_1 \rangle_0 &= \frac{\mathrm{tr}\, H_1 \exp\{-\beta(H_0 - \mu N)\}}{\mathrm{tr} \exp\{-\beta(H_0 - \mu N)\}} \\ \langle H_1 \rangle &= \frac{\mathrm{tr}\, H_1 \exp\{-\beta(H - \mu N)\}}{\mathrm{tr} \exp\{-\beta(H - \mu N)\}}. \end{aligned} \tag{6.9}$$

According to the Gibbs theorem, the grand potential reaches its minimum for a grand canonical distribution function. If it is evaluated in terms of some other trial distribution function, we shall have

$$\Omega_{\mathrm{t}} \geqslant \Omega. \tag{6.10}$$

Together with the inequality (6.8), the relation (6.10) may be used for a variational determination of a correct distribution function.

As we remarked before, there is yet another theorem originally due to Peierls (1938), which gives a condition on the grand potential. The Peierls theorem states

$$\Omega_{\mathrm{D}} \geqslant \Omega \tag{6.11}$$

where

$$\begin{aligned} \exp(-\beta\Omega_{\mathrm{D}}) &= \mathrm{tr} \exp\{-\beta(H - \mu N)_{\mathrm{D}}\} \\ \exp(-\beta\Omega) &= \mathrm{tr} \exp\{-\beta(H - \mu N)\} \end{aligned} \tag{6.12}$$

where the suffix D stands for diagonal. Compared with the inequality (6.11), our result (6.8) is slightly more general. In the case when $\Omega_1 = \Omega - \Omega_0$ and $\langle H_1 \rangle_0 = 0$ the first of the inequalities in (6.8) coincides with what we expect from (6.11).

7. Concluding remarks

Starting with the simple inequality (2.2), we have derived various results. In concluding this article, we remark that one more simple but interesting result can be derived from the inequality. As we shall discuss, this result is not independent of some of our results in the preceding sections.

We recall that the distribution function $g(x)$ is arbitrary. Thus let us choose simply

$$g(x) = 1.$$

Then inequality (2.2) is reduced to

$$\int f \ln f \, dx \geqslant 0 \tag{7.1}$$

under the condition

$$\int f(x) \, dx = \int dx \tag{7.2}$$

in accordance with equation (2.1).

Equation (7.1) may be interpreted that the entropy of a system of a distribution function $f(x)$ satisfying (7.2) is negative or zero. This conclusion may seem strange at first sight since we are accustomed to the positiveness of entropy. However, it is the consequence of the condition (7.2).

One can easily confirm this property by considering simple examples, such as a dipole in a uniform field or a spin in a magnetic field. Moreover, one sees also that equation (3.12) or (5.11) is in agreement with this observation. Thus, if S_c is the configurational entropy related to the free energy by $F_c = U_c - TS_c$, then equation (3.12) or (5.1) gives

$$S_c \leqslant 0. \tag{7.3}$$

This property is simple to remember and may be used to check theoretical results. Furthermore, examining (3.15), one can conclude that the second of the inequalities coincides with what (7.3) states: namely, that part of the Gibbs–Bogoliubov inequality corresponds to the negativeness of the entropy.

Acknowledgments

The author thanks Professors H. Friedman and T. Y. Wu for helpful discussions. After the manuscript was submitted, Dr. S. Harris kindly informed the author of an interesting paper by Grad (1961) which proved (2.2) for the first time.

References

Bogoliubov, N. N., 1958, *Dokl. Akad. Nauk SSSR,* **119,** 244.
Brout, R., 1964, *Phase Transitions* (New York: W. A. Benjamin).
Feynman, R. P., and Cohen, M., 1956, *Phys. Rev.*, **102,** 1189.
Gibbs, J. W., 1902, *Elementary Principles in Statistical Mechanics* (Oxford: Oxford University Press), chap. XI, p. 129.
Girardeau, M. D., 1962, *J. Math. Phys.*, **3,** 131.
—— 1964, *J. Chem. Phys.*, **41,** 2945.
Grad, H., 1961, *Communs Pure Appl. Math,* **14,** 323.
Griffiths, R. B., 1964, *J. Math. Phys.*, **5,** 1215.
Horwitz, G., and Callen, H., 1961, *Phys. Rev.*, **124,** 1757.
Husimi, K., 1940, *Proc. Phys. Math. Soc. Japan,* **22,** 264.
Isihara, A., and Wu, T. Y., 1968, *Nuovo Cim.*, submitted for publication.
Kohn, W., 1964, *Phys. Rev.*, **136,** B864.
Mermin, N. D., 1965, *Phys. Rev.*, **137,** A1441.
Muehlschlegel, B., 1960, *Sber. Bayer Akad. Wiss., Math.–naturw. Kl.*, [3], **10.**
Muehlschlegel, B., and Zittartz, H., 1963, *Z. Phys.*, **175,** 553.
Peierls, R. E., 1938, *Phys. Rev.*, **54,** 918.
Tolmachev, V. V., 1960, *Dokl. Akad. Nauk SSSR,* **134,** 1324.