Journal of Statistical Physics, Vol. 5, No. 1/2, 1972

Statistical Mechanics and the Gauss Principle

M. Dutta[1]

Received July 7, 1971

In an essentially statistical approach to statistical mechanics, it is seen that the Gauss principle of the arithmetic mean may be taken as the starting point. The equations from which the subject can be built up are deduced from the Gauss principle of the arithmetic mean.

KEY WORDS: Gauss principle of arithmetic mean; most probable value; best estimate; location parameter.

1. INTRODUCTION

The problems of statistical mechanics can be solved by several methods with equal success.(1) A physicist, to solve his problem, chooses one conveniently according to this training and practice. Schrödinger(2) and Born(3) discussed two of these methods in some detail. These methods are (1) the method of most probable values of Boltzmann and Planck, and (2) the method of mean values of Darwin and Fowler.

With the frequency interpretation of probability and also with regard to laws of large number,(4) the basic idea of the most probable value is assumed to be that the observed values are most probable values. In the other method, the mean values of a quantity are actually observed. The reconciliation of these two basic ideas can be seen in the Gauss princicple, which may be stated as, "For a set of observations of an unknown quantity, the probability distribution would be such that the probability would be maximum when the

[1] Centre of Advanced Study in Applied Mathematics, Calcutta University, Calcutta, India.

best estimate of a location parameter, used in the definition of probability, would be the arithmetic mean." In other words, the probability distribution associated with an assembly is such that the mean value of an observable quantity is the most probable value (the best estimate). The Gauss principle was first formulated and used in the theory of errors of observations.[5] Its use and significance as a method of estimation are well known.[6]

Here, it is shown that the usual formulas[1–3,7] of statistical mechanics can be obtained in the usual form.

2. PRELIMINARY OBSERVATIONS

For the application of the Gauss principle, the location parameters need to be chosen suitably. It is now clear[7] that if some quantity other than the energy or the like be chosen, the usual statistical mechanics cannot be developed. This point has also been emphasized in an axiomatic development of thermodynamics.[9] If, as in earlier, essentially statistical approaches[8,10] to investigating problems of statistical mechanics, energy and number of particles are chosen as basic variables (or variates) and the Gauss principle is applied, the usual results of statistical mechanics are obtained.

3. DESCRIPTION OF SYSTEM (ASSEMBLY)

Let ϵ_1, ϵ_2,... be the energies and n_1, n_2,... the numbers of particles of the system at different states and let the state defined by (ϵ_i, n_j) be with frequencies f_{ij} and probabilities p_{ij}. After the postulate of statistical independence of this probability of states of the entire system (assembly), consisting of particles in repeated observations, the probability of the entire system being in a number of states is

$$P = \prod_{ij} p_{ij}^{f_{ij}} = \prod_{ij} \{p(\epsilon_i, n_j, \alpha_1, \alpha_2)\}^{f_{ij}} \tag{1}$$

α_1 and α_2 being location parameters (to be identified with the mean energy and mean number of particles). Now, the first task in the essential statistical approaches is to determine the form of a function of $p(\epsilon, n; \alpha_1, \alpha_2)$ along with that to determine the parameters in terms of sample data, here, the observed values of ϵ and n taken as known. Here, this will be done from the Gauss principle.

4. CALCULATIONS

According to the Gauss principle,

$$P = \prod_{ij} p_{ij}^{f_{ij}} \tag{2}$$

is maximum, i.e.,

$$\log P = \sum_{ij} f_{ij} \log p(\epsilon_i , n_j; \alpha_1 , \alpha_2) \tag{3}$$

is maximum along with the conditions

$$\sum_{ij} f_{ij}(\epsilon_i - \alpha_1) = 0 \tag{4}$$

$$\sum_{ij} f_{ij}(n_j - \alpha_2) = 0 \tag{5}$$

where α_1 and α_2 are location parameters of the distribution signifying the mean values $\bar{\epsilon}$ and $\bar{n}$ of ϵ's and n's.

On the assumption that the partial derivatives up to the second order exist, the above conditions can be written as

$$\sum f_{ij}\psi_1(\epsilon_i , n_j ; \alpha_1 , \alpha_2) = 0 \tag{6}$$

$$\sum f_{ij}\psi_2(\epsilon_i , n_j ; \alpha_1 , \alpha_2) = 0 \tag{7}$$

along with the restrictions given by (4) and (5), where

$$\psi_1 = \partial(\log p)/\partial\alpha_1 \quad \text{and} \qquad \psi_2 = \partial(\log p)/\partial\alpha_2 \tag{8}$$

Now, by application of the method of undetermined multipliers for the first variations of ψ_1 and ψ_2 with respect to ϵ_i and n_j , we get

$$\sum f_{ij}[\{\psi_{k\epsilon} - \lambda_{k1}\} \Delta\epsilon_i + \{\psi_{kn} - \lambda_{k2}\} \Delta n_j] = 0, \qquad k = 1, 2 \tag{9}$$

where λ_{k1} and λ_{k2} are independent of ϵ's and n's. Since now $\Delta\epsilon_i$ and Δn_j may be taken as independent, we get

$$\psi_{k\epsilon} - \lambda_{k1} = 0 \qquad \text{and} \qquad \psi_{kn} - \lambda_{k2} = 0 \tag{10}$$

From (10), on account of (9) and (5), we get

$$\partial(\log p)/\partial\alpha_k = \psi_k = \lambda_{k1}(\epsilon - \alpha_1) + \lambda_{k2}(n - \alpha_2) \tag{11}$$

On the assumption that for p,

$$\partial^2(\log p)/\partial\alpha_1\, \partial\alpha_2 = \partial^2(\log p)/\partial\alpha_2\, \partial\alpha_1 \tag{12}$$

is satisfied, we get

$$\frac{\partial\lambda_{11}}{\partial\alpha_2}(\epsilon - \alpha_1) + \frac{\partial\lambda_{12}}{\partial\alpha_2}(n - \alpha_2) - \lambda_{12}$$
$$= \frac{\partial\lambda_{21}}{\partial\alpha_1}(\epsilon - \alpha_1) + \frac{\partial\lambda_{22}}{\partial\alpha_1}(n - \alpha_2) - \lambda_{21} \tag{13}$$

As $\epsilon - \alpha_1$ and $n - \alpha_2$ can assume values independently, we get

$$\partial\lambda_{11}/\partial\alpha_2 = \partial\lambda_{21}/\partial\alpha_1 , \qquad \partial\lambda_{12}/\partial\alpha_2 = \partial\lambda_{22}/\partial\alpha_1 \tag{14}$$

and

$$\lambda_{12} = \lambda_{21} \tag{15}$$

The conditions (14) leads to the existence of the functions ν_1 and ν_2 such that

$$\lambda_{11} = \partial\nu_1/\partial\alpha_1 , \quad \lambda_{21} = \partial\nu_1/\partial\alpha_2 , \quad \lambda_{12} = \partial\nu_2/\partial\alpha_1 , \quad \lambda_{22} = \partial\nu_2/\partial\alpha_2 \tag{16a}$$

with

$$\nu_k(\alpha_{10} , \alpha_{20}) = 0, \qquad k = 1, 2 \tag{16b}$$

and the condition (15) leads to the existence of a function Π such that

$$\nu_1 = \partial\Pi/\partial\alpha_1 \qquad \text{and} \qquad \nu_2 = \partial\Pi/\partial\alpha_2 \qquad \text{with} \quad \Pi(\alpha_{10} , \alpha_{20}) = 0 \tag{17}$$

where α_{10} and α_{20} are initial values of α_1 and α_2 . Thus, from (11), we get

$$\partial(\log p)/\partial\alpha_k = (\partial/\partial\alpha_k)\{\nu_1(\epsilon - \alpha_1) + \nu_2(n - \alpha_2) + \Pi\} \tag{18}$$

or

$$p(\epsilon, n; \alpha_1 , \alpha_2) = \omega(\epsilon, n) \exp\{\nu_1(\epsilon - \alpha_1) + \nu_2(n - \alpha_2) + \Pi\} \tag{19}$$

Now, denoting

$$Z(\nu_1 , \nu_2) = \sum_{ij} \omega(\epsilon_i , n_j) \exp\{\nu_1(\epsilon_i - \alpha_1) + \nu_2(n_j - \alpha_2)\} \tag{20}$$

$$= \exp(\nu_1\alpha_1 + \nu_2\alpha_2 - \pi) \tag{21}$$

as

$$\sum_{ij} p(\epsilon_i , n_j; \alpha_1 , \alpha_2) = 1 \tag{22}$$

We can easily see that

$$\bar{\epsilon} = \alpha_1 = \partial[\log Z(\nu_1 , \nu_2)]/\partial\nu_1 \tag{23}$$

$$\bar{n} = \alpha_2 = \partial[\log Z(\nu_1 , \nu_2)]/\partial\nu_2 \tag{24}$$

Now, it is easy to recognize $Z(\nu_1 , \nu_2)$ as the *Zustandsumme* of Planck[1] or the partition function of Darwin and Fowler.[7] Then, the parameters ν_1 and ν_2 may be interpreted as usual and all other usual results of statistical thermodynamics[7] can be deduced from these results as they are done in earlier work.[10]

5. CONCLUDING REMARKS

The above method for the deduction of the basic equation of statistical thermodynamics from the Gauss principle can be extended easily and directly to the case of more complicated systems when they consist of different types of particles or when considerations of other quantities such as charge or momentum are necessary. In earlier work,[10] basic equations have obtained in two phases, first, by taking the form of the probability functions with the help of the Bayes principle or some other suitable principle, and second, by estimating the parameters of distributions by the principle of maximum likelihood. However, one of the main advantages, of the present method is that the form of the probability distribution-function with the proper parameters is obtained at the same time as in the method of maximum entropy estimation.[8]

ACKNOWLEDGMENTS

The author expresses his thanks to his colleague Mr. P. K. Ghosh and his student Mr. C. G. Chakraborty for going through the manuscript.

REFERENCES

1. M. Dutta, *Indian J. Phys.* **44**:218 (1968); *Statistical Physics (Foundations)*, World Press, Calcutta (in press).
2. E. Schrödinger, *Statistical Thermodynamics*, Cambridge Univ. Press (1947).
3. M. Born, *Natural Philosophy of Cause and Chance*, Dover Publications, New York (1966).
4. H. Cramer, *Mathematical Methods of Statistics*, Princeton Univ. Press (1946).
5. R. von Mises, *Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik*, Mary Rosenberg, New York (1943). M. Dutta and S. Pal, *Introduction to Mathematical Theory of Probability and Statistics*, World Press, Calcutta (1963).
6. J. M. Keynes, *A Treatise on Probabilities*, Macmillan, London (1921). M. G. Kendall, The Advanced Theory of Statistics, Vol. II, 3rd. Ed., Charles & Co., London (1960). M. Dutta, *Sankhya, A* **28**:319–328 (1966). L. L. Campbell, *Ann. Math. Statistics* **41**:1011–1019 (1970).
7. R. H. Fowler, *Statistical Mechanics*, Cambridge Univ. Press, (1929).
8. E. F. Jaynes, "Information Theory and Statistical Mechanics," in *Statistical Physics*, G. E. Uhlenbeck *et al.*, eds. Benjamin, New York (1963).
9. R. Giles, *Mathematical Foundation of Thermodynamics*, Pergamon Press, New York (1964).
10. M. Dutta, *Bull. Indian Stat. Inst.* **33**:285 (1953); *Proc. Nat. Inst. Sci. (India)* **19**:109 (1953); *Proc. Nat. Inst. Sci. (India)* **21**:373 (1955); *Proc. Summer Inst. of Theor. Phys.*, Mussorie, India, p. 313 (1959); *Z. Phys. Chem.* **228**:380 (1965); *Indian J. Phys.* **40**:85 (1966).