C H A P T E R - III

JAYNES' MODEL AND IRREVERSIBILITY

CHAPTER - III

JAYNES' MODEL AND IRREVERSIBILITY

3.1 INTRODUCTION

Jaynes' stochastic model, based on the principle of maximum entropy estimation, was developed with a new point of view in which the information theory is placed in the basis of the statistical mechanics. Shannon (1948) defined entropy as a measure of uncertainity for a general stochastic model without any reference to thermodynamics. Though the occurrence of the same mathematical expression $-\Sigma p_i \log p_i$ for entropy in the statistical mechanics as well as in the information theory was known by the contemporary scientists, no deeper inter-relation was studied or duly stressed. Jaynes (1957a) was the first to develop the statistical mechanics in a general way starting with the information theoretic entropy. Maximising this entropy subject to some constraints, he calculated the probability distribution for a thermodynamic system, when only some partial information is supplied. Jaynes' theory is perfectly general, based on the statistical consideration irrespective of any branch of the physical sciences and proved its successful applicability in statistical mechanics. He set up a new method for solving problems of specification and estimation, with application in physical sciences.

Jaynes' model was set up mainly for equilibrium system, though its suitability was claimed subsequently (1957b) for the irreversible processes too.

In this chapter, in the above model, the effect on the probabilistic distribution due to the change in observed values of additive random variables, say X's in a prescribed manner is found out and the results are interpreted in terms of thermodynamics. The mathematical methods used are very simple. Linear relations between the small variations of X's with those of statistical parameters are obtained. Co-efficients therein satisfy Onsager (1931) principle of symmetry and also have the same statistical interpretation.

3.2 PROBABILITY DISTRIBUTION

Jaynes' approach regarding the formulation of probability distribution involves the subjective point of view.

A short resume of Jaynes' method of estimation might be helpful to make the basic idea of this present investigation more clear. If X_r $(r = 1,2,\ldots,k)$ denote observable entities of a thermodynamic system, n being an index of a state of the system, specified by an ordered set of values of X's , then Jaynes posed the problem as to how the corresponding probability distribution p_i may be calculated, when only the informations

$$\langle X_r \rangle = E_{p_i}(X_r) \qquad \ldots(3.1)$$

and

$$\sum_{i=1}^{n} p_i = 1 \qquad \ldots(3.2)$$

are supplied. Apparantly this problem is insolvable. For, (n-2) more conditions are required in order to solve for p_i's .

Jaynes started with the Shannon's information theoretic entropy

$$H(p_1, p_2, \ldots, p_n) = -k \sum p_i \log p_i \quad , (k > 0). \qquad \ldots(3.3)$$

Shannon (1948) defined entropy in the development of information theory, as a measure of uncertainity, apart from thermodynamic consideration. It was expressed as a function of discrete probability distribution. This function (3.3) is positive, increases with increasing uncertainity and additive for independent source of uncertainity. Now, the problem forwarded by Jaynes takes the form to obtain the most unbias probability distribution, agreeing with the information already given.

To estimate p_i, Jaynes maximised (3.3) subject to the constraints

$$\langle X_r \rangle = \sum_{i=1}^{n} p_i X_{i,r} \qquad \text{...(3.4)}$$

and

$$\Sigma p_i = 1 \qquad \text{...(3.4a)}$$

The justification of estimating p_i's corresponding to the maximum entropy value is that the smaller the entropy (i.e., the measure of uncertainity), the more information is conveyed by p_i.

On simple variations with respect to the distributions $\{p_i\}$ by the use of the method of Lagrange's multipliers, as done by Jaynes, one gets

$$p_i = e^{\alpha_0 + \sum_r \alpha_r X_{r,i}} \qquad \text{...(3.5)}$$

$$= \frac{e^{\sum_r \alpha_r X_{r,i}}}{\sum_i e^{\sum \alpha_r X_{r,i}}} \qquad \text{...(3.6)}$$

where the parameters are determined from

$$\langle X_r \rangle = \frac{\delta}{\delta \alpha_r} (\log Z) \qquad \text{...(3.7)}$$

$$\alpha_0 = - \log Z(\alpha_1, \alpha_2, \ldots, \alpha_k) \qquad \text{...(3.8)}$$

and

$$Z(\alpha_1,\ldots,\alpha_k) = \sum_i e^{\sum_r \alpha_r X_{r,i}} \qquad \ldots(3.9)$$

which is the corresponding partition function.

Correspondingly the entropy of the distribution (3.6) is

$$S_{max} = - \left[\alpha_o + \alpha_1 \langle X_1 \rangle + \ldots + \alpha_k \langle X_k \rangle \right] \qquad \ldots(3.10)$$

where k is set equal to unity.

3.3 APPLICATION TO STATISTICAL MECHANICS

Jaynes used this particular method of specification in statistical mechanics.

Let the energy levels of a system be $E_i(\beta_1,\beta_2,\ldots,\beta_k,\ldots)$, where the external parameters β_i may include volume, strain tensor, applied electric and magnetic fields, gravitational potentials etc.

If the average value of the energy $\langle E \rangle$ of the system which characterises the equilibrium of the system is known, then the previous method of maximum entropy estimation at once suggests that the distribution for the energy levels will be

$$p_i = \frac{e^{\alpha_1 E_i}}{Z(\alpha_1)} \qquad \ldots(3.11)$$

where α_1 is given by

$$\langle E \rangle = \frac{\delta}{\delta\alpha_1} \log Z(\alpha_1) \qquad \ldots(3.12)$$

and

$$Z(\alpha_1) = \sum_i e^{\alpha_1 E_i} \qquad \ldots(3.13)$$

which is the partition function.

This distribution corresponds to Gibbs' canonical distribution, when α_1 is interpreted as temperature.

3.4 PHYSICAL INTERPRETATION

The distribution parameters α_i's may be identified with thermodynamic intensive variables. If one of the X_i's be taken as the energy of the system, the corresponding α_i comes as $-\frac{1}{kT}$, where k is the Boltzmann constant and T is the temperature. Jaynes identified the parameters of distribution with physical quantities, just by comparing them with usual thermodynamic quantities, as Schrödinger (1949) had done.

3.5 INTRODUCTION OF AN NEW PARAMETER WITH BASIC POSTULATES

Let λ be any parameter different from X_r's , α_r's so that the state of the system depends on λ, such that $\langle X_r \rangle$'s vary when λ varies. The first and foremost problem of the present investigation is to discuss as to how the entire statistical and consequently the thermodynamic behaviour will change when dependence of $\langle X_r \rangle$ and λ is specified.

Basic postulates :

The following postulates are the basis of the entire discussion.

(1) The effect of the slight variation of λ does not change the intrinsic nature (dynamical or quantal) of the system, in other words, the set of possible values of every X_r, $r = 1,2,\ldots,k$, where X_r is the extensive variable giving significant information about the system under consideration, remains unaffected.

(2) This variation of λ does not change the form of the probability distribution in states, though the distribution parameters $\alpha_r(\lambda)$'s are maximum likelihood estimations (Dutta, 1953) at every value of λ .

In symbols,

$$E_{p_i(\lambda)}(X_r) = \langle X_r(\lambda) \rangle, \quad r = 1,2,\ldots,k \qquad \ldots(3.14)$$

$$\sum_i p_i(\lambda) = 1 \qquad \ldots(3.15)$$

$$p_i(\lambda) = \frac{e^{\sum_r a_r(\lambda) X_{r,i}}}{\sum_i e^{\sum_r a_r(\lambda) X_{r,i}}} \qquad \ldots(3.16)$$

Note : The above postulates appear to be equivalent to the postulates of variations of states very near to equilibrium, statistical or thermodynamic.

3.6 DERIVATION OF A RESULT SIMILAR TO THAT OF ONSAGER

Let $X_r(\lambda)$, $r = 1,2,\ldots,k$, be k physically measurable entities depending on the parameter λ, in the system deviated slightly from its equilibrium.

Then by Section 3.5,

$$\langle X_r(\lambda) \rangle = \sum_{i=1}^{n} X_{r,i}\, p_i(\lambda) \qquad \ldots(3.17)$$

where

$$p_i(\lambda) = \frac{e^{\sum_{r=1}^{k} a_r(\lambda) X_{r,i}}}{Z(a_1(\lambda),\ldots,a_k(\lambda))} \qquad \ldots(3.18)$$

$$Z(a_1(\lambda),\ldots, a_k(\lambda)) = \sum_i e^{\sum_r a_r(\lambda)X_{r,i}} \quad . \qquad \ldots(3.19)$$

For a small deviation in this system from its equilibrium during a small change in λ say $\Delta\lambda$, let the corresponding changes in $\langle X_r(\lambda)\rangle$'s, $a_r(\lambda)$'s be $\Delta\langle X_r(\lambda)\rangle$ and $\Delta a_r(\lambda)$.

Then by method of simple calculus,

$$\Delta\langle X_r(\lambda)\rangle = \sum_{j=1}^{k} \frac{\delta\langle X_r(\lambda)\rangle}{\delta a_j} \Delta a_j(\lambda) \text{ , } r = 1,2,\ldots,k \text{ .}$$

$$= \left[\langle (X_r(\lambda))^2\rangle - \langle X_r(\lambda)\rangle^2\right] \Delta a_r(\lambda)$$

$$+ \sum_{j,j\neq r} \left[\sum_i (X_{r,i}X_{j,i} - \langle X_r\rangle\langle X_j\rangle)\, p_i(\lambda)\right] \Delta a_j(\lambda) \qquad \ldots(3.20)$$

$$\therefore \frac{d\langle X_r(\lambda)\rangle}{d\lambda} = \left[\langle (X_r(\lambda))^2\rangle - (\langle X_r(\lambda)\rangle)^2\right] \frac{da_r(\lambda)}{d\lambda}$$

$$+ \sum_{j,j\neq r} \left[\sum_i (X_{r,i}X_{j,i} - \langle X_r\rangle\langle X_j\rangle)\, p_i(\lambda)\right] \frac{da_j}{d\lambda}$$

$$r = 1,2,\ldots,k \quad . \qquad \ldots(3.21)$$

The significance of this linear relations (3.21) lies in the symmetry of the coefficients i.e. r-jth co-efficient is equal to j-rth co-efficient.

When this parameter λ is identified with time t, (3.21) gives a linear relation between

$$\frac{d \langle X_r(t) \rangle}{dt} \quad \text{and} \quad \frac{d\, a_j(t)}{dt}$$

i,e, between the variation of the extensive variable like energy etc. and the rate of change of the intensive variables like temperature, showing a good resemblence with the results of Onsager regarding the symmetry of the co-efficient matrix.

Now if any spatial coordinates are taken for λ and the above calculations are repeated, similar results are obtained on the hypothesis of spatially local equilibrium.

3.7 CONCLUDING REMARKS

The study of the results (3.21) leads to few more results, which are significant from the standpoint that the changes occur in a thermodynamic system, when it passes from reversibility to irreversibility, when the parameter λ stands for time (t). This will be discussed in the subsequent chapter.