THE INFORMATION APPROACH TO THE THEORY OF IRREVERSIBLE QUASI-EQUILIBRIUM PROCESSES

A. A. Filyukov

Inzhenerno-Fizicheskii Zhurnal, Vol. 14, No. 3, pp. 493–499, 1968

UDC 536.75

The principles of the thermodynamics of irreversible quasi-equilibrium processes are derived by means of the formalism of information theory. The quasi-equilibrium analogies of familiar thermostatic concepts are formulated.

§1. **Information formalism.** The use of information theory for the formulation of equilibrium statistical thermodynamics was first applied in [1, 2] and is known in the literature as the Jaynes formalism. This approach is superior to the traditional one in that for a closed formulation and for the derivation of virtually the entire apparatus of statistical thermodynamics no use is made of the concept of phase space, original theorems, nor of the postulates of statistical mechanics, while the very problem of statistical thermodynamics becomes a special problem among the others with incomplete information. In this class of problems with incomplete statistical information we can also include the problem of describing irreversible quasi-equilibrium processes.

The subject of the present paper is the application of the information approach to the formulation of the basic statements in the thermodynamics of quasi-equilibrium processes, thus making it possible to introduce a number of analogs of well-known concepts from equilibrium statistical thermodynamics and to formulate the mathematical apparatus constructed in the manner of equilibrium statistical thermodynamics. This apparatus operates with concepts whose physical significance is different from the corresponding equilibrium analogs.

For completeness, we will briefly outline the information formalism, which should be defined more exactly as a statistical method of analysis that is based on premises corresponding to maximum entropy.

Information as to the specific problem under consideration is provided in the form of a set of known average values of $\overline{f}_k$ for certain random functions f_k which depend on the discrete random variable ε_i and, moreover, may possibly depend on one or more external parameters α

$$\langle f_k(\varepsilon_i, \alpha)\rangle = \overline{f}_k, \tag{1}$$

where $\langle f\rangle \equiv \sum_i p_i f(\varepsilon_i, \alpha)$; $\{p_i\}$ is some discrete distribution function which satisfies the following requirements:

1. The information entropy of the Shannon distribution $\{p_i\}$

$$S = -\sum_i p_i \log p_i \tag{2}$$

exhibits the maximum possible value commensurate with the value of the known average quantities

$$\sum_i p_i f_k = \overline{f}_k \quad (k = 1 - m).$$

2. It satisfies the normalization condition

$$\sum_i p_i = 1. \tag{3}$$

These requirements lead to the so-called canonical form of the distribution function

$$p_i = \frac{\exp\left[-\sum_k \lambda_k f_k(\varepsilon_i, \alpha)\right]}{Z(\{\lambda_k\})}, \tag{4}$$

where the function of the parameters λ_k of the form

$$Z(\{\lambda_k\}) = \sum_i \exp\left[-\sum_k \lambda_k f_k(\varepsilon_i, \alpha)\right] \tag{5}$$

is usually identified as a statistical sum while the parameters λ_k (k = 1 − m) themselves are introduced to account for the additional conditions in the form of Eq. (1).

Choosing the distribution function which corresponds to the maximum value of the Shannon entropy ensures that there will be no nonrandom information except that which is implicit in conditions (1) and forms the essence of the statistical analysis method applicable to any problem in which information is specified in this form. In addition to the above-cited papers by Jaynes, a detailed justification for the application of the information approach can be found in [3–6], while for convenience in the remaining discussion we will note that from the formal mathematical standpoint all of the expressions cited above, as well as all those which follow from their corollaries, are identical to the familiar formulas of the statistical thermodynamics of the equilibrium state [7], although their meaning may be entirely different, depending on the form of the problem and the significance of the random functions.

§2. **The quasi-equilibrium approximation.** The initial assumption of the quasi-equilibrium method in the theory of irreversible processes involves the fact that in a state close to the equilibrium, but different from that state, the average value of any random function characterizing the system can be presented in the form of an expansion in moments of distribution for the corresponding equilibrium function. The expansion-smallness parameter characterizes the degree of deviation from the state of equilibrium, and we can neglect all of the terms in the expansion higher than the linear.

We will denote the average value of the random function in the nonequilibrium state by $\langle f^+\rangle$, while the one in the equilibrium state will be denoted by $\langle f\rangle$. We will denote the expansion parameter by λ^+, and its equilibrium value will be denoted by λ, so that the quasi-equilibrium approximation of the nonequilibrium average value is given by

$$\langle f^+ \rangle = [\langle f^+ \rangle]_{\lambda^+=\lambda} + (\lambda^+ - \lambda)\left[\frac{d\langle f^+ \rangle}{d\lambda^+}\right]_{\lambda^+=\lambda}. \quad (6)$$

It is clear that

$$[\langle f^+ \rangle]_{\lambda^+=\lambda} = \langle f \rangle \quad (7)$$

and

$$\left[\frac{d\langle f^+ \rangle}{d\lambda^+}\right]_{\lambda^+=\lambda} = \frac{d\langle f \rangle}{d\lambda}. \quad (8)$$

If we simultaneously consider several nonequilibrium functions characterizing the system $\{f_k^+\}$, the expansion is naturally generalized as follows:

$$\langle f_k^+ \rangle = \langle f_k \rangle + \sum_l (\lambda_l^+ - \lambda_l)\frac{\partial\langle f_k \rangle}{\partial\lambda_l}, \quad (9)$$

where the subscripts l and k extend over identical set of values.

We will treat relationship (9) as all of the available information with regard to some nonequilibrium process; in the quasi-equilibrium approximation the information regarding the average value of the nonequilibrium quantity $\langle f_k^+\rangle$ is thus equivalent to the information with respect to its average value $\langle f_k\rangle$ in equilibrium, and it is also equivalent to the information about the variance and correlation of the random functions, since the derivatives in (9) under the summation sign have precisely this statistical significance [3],

$$\frac{\partial\langle f_k \rangle}{\partial\lambda_l} = \langle f_k f_l \rangle - \langle f_k \rangle\langle f_l \rangle, \quad (10)$$

$$\frac{\partial\langle f_k \rangle}{\partial\lambda_k} = \langle f_k^2 \rangle - \langle f_k \rangle^2, \quad (11)$$

and, moreover, the information about the values of the parameters λ_k^+ and λ_k.

It is then possible to formulate the problem of describing the nonequilibrium steady state from the position of information formalism: to find the nonequilibrium distribution function $\{p_i^+\}$ which satisfies the known information (9) and guarantees the absence of any other additional assumptions, nonrandom in nature.

The solution of this problem is possible with the aid of the formal entropy logic indicated in §1. Indeed, the problem pertains precisely to the class described there: there are certain averages of $\langle f_k^+\rangle$ with regard to which we know that they are equal to the equilibrium analogs of $\langle f_k\rangle$, plus certain additions proportional to the variances and correlations of the corresponding quantities.

The analysis method based on the maximization of entropy leads immediately to the conclusion that the nonequilibrium distribution function is again of canonical form:

$$p_i^+ = \frac{\exp\left[-\sum_k \Lambda_k f_k^+(\varepsilon_i, \alpha)\right]}{Z^+(\{\Lambda_k\})}, \quad (12)$$

where Λ_k is the Lagrange multiplier by means of which we take into consideration the known information in the form of (9), while Z^+ is a nonequilibrium statistical sum which is a function of these Lagrange parameters.

Let us find the relationship between the parameters Λ_k with the expansion parameters $\lambda_k^+ - \lambda_k$. For this we will expand the average value of the random function f_k—weighted with respect to the nonequilibrium distribution (12)—in the parameters Λ_k, assuming these to be infinitesimals:

$$\langle f_k^+ \rangle = [\langle f_k^+ \rangle]_{\Lambda=0} +$$
$$+ \sum_l \Lambda_l\left[\frac{\partial(1/Z^+)}{\partial\Lambda_l}\sum_i f_k^+ \exp\left[-\sum_k \Lambda_k f_k^+\right]\right]_{\Lambda=0} +$$
$$+ \sum_l \Lambda_l\Big[(1/Z^+)\sum_i f_k^+(-f_l^+)\times$$
$$\times\exp\Big[-\sum_k \Lambda_k f_k^+\Big]\Big]_{\Lambda=0} + \ldots =$$
$$= [\langle f_k^+ \rangle]_{\Lambda=0} +$$
$$+ \sum_l \Lambda_l[\langle f_k^+ \rangle\langle f_l^+ \rangle - \langle f_k^+ f_l^+ \rangle]_{\Lambda=0}. \quad (9^*)$$

We can therefore assume that (compare (9*) with (9))

$$\Lambda_k = \lambda_k^+ - \lambda_k. \quad (13)$$

For the remainder of the discussion we will also need a number of relationships which associate the above-introduced characteristics and whose validity is easily proved directly, using the definitions of (1), (2), (4), and (5):

$$\langle f_k^+ \rangle = -\frac{\partial\log Z^+(\{\Lambda_l\})}{\partial\Lambda_k}, \quad (14)$$

$$\Lambda_k = \frac{\partial S^+(\{\langle f_l^+ \rangle\})}{\partial\langle f_k^+ \rangle}, \quad (15)$$

$$\frac{\partial\langle f_k^+ \rangle}{\partial\Lambda_l} = \frac{\partial\langle f_l^+ \rangle}{\partial\Lambda_k}, \quad (16)$$

$$S^+(\{\langle f_k^+ \rangle\}) = Z^+(\{\Lambda_k\}) + \sum_k \Lambda_k\langle f_k^+ \rangle \quad (17)$$

(the Lagrange transform).

With the random functions dependent on the common parameter α, the statistical sum also becomes a function of that parameter,

$$Z^+(\{\Lambda_k\};\, \alpha) = \sum_i \exp\left[-\sum_k \Lambda_k f_k^+(\varepsilon_i, \alpha)\right], \quad (18)$$

while the logarithmic derivative with respect to this parameter is given by

$$\frac{\partial\log Z^+(\{\Lambda_k\};(\alpha)}{\partial\alpha} = -\sum_k \Lambda_k\left\langle\frac{\partial f_k^+}{\partial\alpha}\right\rangle. \quad (19)$$

The derivative of the entropy with respect to the external parameter α, in view of the implicit relationship, is given by

$$\frac{dS(\{\langle f_k^+(\varepsilon_i, \alpha)\rangle\})}{d\alpha} = \sum_k \frac{\partial S^+}{\partial \langle f_k^+\rangle} \frac{\partial \langle f_k^+\rangle}{\partial \alpha} = \sum_k \Lambda_k \frac{\partial \langle f_k^+\rangle}{\partial \alpha}.$$

§3. Microscopic reversibility. In addition to the original assumption of (9) with regard to the form of the average values for the nonequilibrium functions, the familiar principle of microscopic reversibility [8, 9] is employed in the method of quasi-equilibrium thermodynamics for the derivation of the Onsager symmetry relationships. Let us examine the role of this principle in the information approach.

The principle of microscopic reversibility presupposes that any sequence of events (the values of random functions), defined on an equilibrium ensemble, and occurring in forward time, has already occurred symmetrically in reverse time.

This principle can be expressed mathematically by the equality

$$f_i(t)\langle f_j(t+\tau)\rangle = f_i(t)\langle f_j(t-\tau)\rangle, \qquad (20)$$

where $f_i(t)$ is the value of some random function at an arbitrarily chosen instant t, while the averaging $\langle\ldots\rangle$, generally speaking, of some other random function f_j is carried out over all instants t in which $f_i(t)$ is equal to the chosen fixed value of $\bar{f}_i$.

For the purposes of the thermodynamics of the quasi-equilibrium state, the principle of microscopic reversibility is usually employed in a considerably weaker form, and namely, in the average sense [8] of

$$\langle f_i(t)\cdot f_j(t+\tau)\rangle = \langle f_i(t)\cdot f_j(t-\tau)\rangle, \qquad (21)$$

where the averaging is accomplished over all times t, rather than over the selected instants t for which $f_i(t) = \bar{f}_i$.

Bearing in mind the steadiness of the random functions, we can shift the time origin by τ, for example, in the right-hand portion of (21), and this will yield

$$\langle f_i(t) f_j(t+\tau)\rangle = \langle f_i(t+\tau) f_j(t)\rangle. \qquad (22)$$

Subtracting $\langle f_i(t) f_j(t)\rangle$ from both members, dividing by τ, and assuming the existence of an average limit

$$\left\langle \lim_{\tau\to 0} \frac{f(t+\tau)-f(t)}{\tau}\right\rangle = \left\langle \frac{\partial f}{\partial t}\right\rangle, \qquad (23)$$

we obtain the following form for the expression of the principle of microscopic reversibility in the average sense:

$$\left\langle f_i(t)\frac{\partial f_j(t)}{\partial t}\right\rangle = \left\langle f_j(t)\frac{\partial f_i(t)}{\partial t}\right\rangle. \qquad (24)$$

The principle of microscopic reversibility can also be formulated in similar fashion in a rigorous sense. Its expression under the assumption of (23) has the following form:

$$\left\langle (f_i(t)-\langle f_i(t)\rangle)\left(\frac{\partial f_j(t)}{\partial t} - \left\langle\frac{\partial f_j(t)}{\partial t}\right\rangle\right)\right\rangle = \left\langle (f_j(t)-\langle f_j(t)\rangle)\left(\frac{\partial f_i(t)}{\partial t} - \left\langle\frac{\partial f_i(t)}{\partial t}\right\rangle\right)\right\rangle. \qquad (24^*)$$

Up to this point we have not introduced the concept of time in the description of information formalism. This concept is conveniently introduced into the consideration as a general external parameter of the random functions $\alpha \doteq t$ (in analogy with the introduction of the volume parameter V in thermostatics).

Thus all of the random functions with the exception of the dependence on the random discrete variable ε_i will now also be a function of the general time parameter (t). With (19) we immediately derive the relationship between the average values of the velocities and the statistical sum

$$\frac{\partial \log Z^+}{\partial t} = -\sum_k \Lambda_k \left\langle \frac{\partial f_k^+}{\partial t}\right\rangle. \qquad (25)$$

If we are interested in a certain average value for the rate of change in the nonequilibrium function f_k^+, this can be found with (25) for a known statistical sum in the form

$$\left\langle \frac{\partial f_k^+}{\partial t}\right\rangle = -\frac{\partial^2 \log Z^+}{\partial \Lambda_k \partial t}. \qquad (26)$$

On the other hand, from (14) we can obtain an expression for the mixed derivative of the logarithm of the statistical sum, with an inverse sequence on the order of differentiation,

$$-\frac{\partial^2 \log Z^+}{\partial t \partial \Lambda_k} = \frac{\partial \langle f_k^+\rangle}{\partial t}, \qquad (27)$$

and as shown by a comparison of (26) and (27), the second mixed derivatives of log Z^+ with a different order of differentiation exhibit different statistical sense. Let us find this difference. For this we will examine the structure of the rate of change in the average value:

$$\frac{\partial \langle f_k^+\rangle}{\partial t} = \frac{\partial}{\partial t}\sum_i p_i^+(t) f_k^+(\varepsilon_i, t) = \sum_i p_i^+ \frac{\partial f_k^+}{\partial t} + \sum_i f_k^+ \frac{\partial p_i^+}{\partial t} = \left\langle \frac{\partial f_k^+}{\partial t}\right\rangle + \sum_i f_k^+ \frac{dp_i^+}{dt}. \qquad (28)$$

Thus the difference between the rate of change in the average value and the average value for the rate of change in the random functions is associated with the implicit relationship between the distribution $p_i^+(t)$ and time:

$$\frac{\partial \langle f_k^+\rangle}{\partial t} - \left\langle \frac{\partial f_k^+}{\partial t}\right\rangle = \sum_i f_k^+ \frac{dp_i^+}{dt}. \qquad (29)$$

Simple calculation shows that

$$\frac{dp_i^+}{dt} = -p_i^+ \sum_l \Lambda_l \left[\frac{\partial f_l^+}{\partial t} - \left\langle \frac{\partial f_l^+}{\partial t} \right\rangle\right]. \tag{30}$$

Substitution of (30) into (29) yields an expression for the difference under consideration:

$$\frac{\partial \langle f_k^+ \rangle}{\partial t} - \left\langle \frac{\partial f_k^+}{\partial t} \right\rangle =$$
$$= - \sum_l \Lambda_l \sum_i p_i^+ f_k^+ \frac{\partial f_l^+}{\partial t} +$$
$$+ \sum_l \Lambda_l \sum_i p_i^+ f_k^+ \left\langle \frac{\partial f_l^+}{\partial t} \right\rangle =$$
$$= - \sum_l \Lambda_l \left[\left\langle f_k^+ \frac{\partial f_l^+}{\partial t} \right\rangle - \langle f_k^+ \rangle \left\langle \frac{\partial f_l^+}{\partial t} \right\rangle\right] =$$
$$= - \sum_l \Lambda_l \langle (f_k^+ - \langle f_k^+ \rangle) \times$$
$$\times \left(\frac{\partial f_l^+}{\partial t} - \left\langle \frac{\partial f_l^+}{\partial t} \right\rangle \right) \rangle . \tag{31}$$

The requirement of equality for the mixed second derivatives of $\log Z^+$—which follows from the definition of the statistical sum as a function of state—naturally leads to the principle of microscopic reversibility in the rigorous sense.

§4. **The linear law.** Let us analyze the problem of the linear relationship between flows and thermodynamic forces; this relationship is also one of the postulates of quasi-equilibrium thermodynamics. It turns out in our examination that the linear relationship is a consequence of one of two possible assumptions.

In the first assumption

$$\left\langle \frac{\partial f_k^+}{\partial t} \right\rangle = 0. \tag{32}$$

From Eq. (31) we then automatically have the linear law relating the flow with force:

$$\frac{\partial \langle f_k^+ \rangle}{\partial t} = - \sum_l \Lambda_l \left\langle f_k^+ \frac{\partial f_l^+}{\partial t} \right\rangle \tag{33}$$

with the phenomenological coefficients

$$L_{lk} = \left\langle f_k^+ \frac{\partial f_l^+}{\partial t} \right\rangle, \tag{34}$$

satisfying the symmetry relationships in view of the microscopic reversibility in the average sense of (21).

The second assumption—also leading to a linear relationship between flow and force—reduces to the equation

$$\left\langle \frac{\partial f_k^+}{\partial t} \right\rangle = - \sum_l \Lambda_l \langle f_k^+ \rangle \left\langle \frac{\partial f_l^+}{\partial t} \right\rangle. \tag{35}$$

Substitution of (35) into (20) immediately yields the linear law (33) with the same phenomenological coefficients.

Entry into the nonlinear region [10] is obviously associated with the rejection of such assumptions as (32) or (35), as well as with the violation of the principle of microscopic reversibility (31). This interesting area goes beyond the framework of the basic assumptions which we have employed and requires particular investigation. It is to be hoped that the information approach will also provide this field with the characteristic clarity of the origin premises and ease in achieving concrete results.

REFERENCES

1. E. T. Jaynes, Phys. Rev., **106**, 620, 1957.
2. E. T. Jaynes, Phys. Rev., **108**, 171, 1957.
3. E. T. Jaynes, Information Theory and Statistical Mechanics, Notes by the lecturer, Washington University, 1964.
4. M. J. Tribus, Appl. Mech., March, 1961.
5. A. Katz, Il. Nuovo Cimento XXXIII, no. 6, 1544, 1964.
6. A. Katz, Il. Nuovo Cimento XXXIII, no. 6, 1553, 1964.
7. E. Schredinger, Statistical Thermodynamics, Cambridge, 1946.
8. J. C. M. Li, J. Chem. Phys., **29**, no. 4, 1958.
9. L. Onsager, Phys. Rev., **37**, 405; **38**, 2265, 1931.
10. Th. A. Bak, Advances in Chemical Physics, vol. III. (ed. Prigogine), New York.

5 June 1967

Institute of Applied Mathematics AS USSR, Moscow