

METHOD OF THE MOST PROBABLE PATH OF EVOLUTION IN THE THEORY OF STATIONARY IRREVERSIBLE PROCESSES

A. A. Filyukov and V. Ya. Karpov

Inzhenerno-Fizicheskii Zhurnal, Vol. 13, No. 6, pp. 798-804, 1967

UDC 536.75: 530.16

Nonequilibrium stationary thermodynamic systems are described by stationary random Markov processes with discrete time. For given macroscopic conditions, the condition of maximum path entropy determines the process.

In statistical thermodynamics one describes the properties of physical systems averaged over all possible initial states because the latter are impossible to ascertain in practice. The set of systems corresponding to all possible initial states of a given system comprises its Gibbs ensemble. Functions defined on the set forming the ensemble are called random quantities, and calculating their expectations (averaged over the ensemble) is the basic problem of statistical thermodynamics.

To calculate these averages, it is not at all necessary to know the random quantities themselves, but rather only their distributions, which are time-independent because of the stationary nature of the equilibrium system. Because the dynamical laws of evolution of the system are either unknown or too complicated to use directly, the distributions of thermodynamics are found by introducing the basic postulate according to which equilibrium distributions of random quantities maximize the entropy.

To describe nonequilibrium processes, it is not enough to know the distributions of the random quantities, because it is important to know not only the distributions over states, but also the rates of transition from one state to another, i. e., one must know the evolution of the random quantities in time of the so-called "random" process. To completely determine the random process, one must by some means know the probabilities of all possible trajectories, i. e., the sequence of states traversed by the system in the course of its evolution in time. The validity of the assumptions used as a postulative basis in choosing the trajectory probabilities can be estimated only by comparing experiment with the results given by the approximation in question. Our task is to describe the real behavior of a given thermodynamic system by a random process chosen according to the basic postulate.

A random process defined at n different instants in time may be regarded as an n -dimensional random quantity. This suggests the possibility of applying the second law of thermodynamics to the theory of random processes, since the effectiveness of its application in the case of one-dimensional random quantities has been demonstrated by the agreement between equilibrium theory and experiment.

As our basic a priori postulate we will assume that for given macroscopic conditions imposed on the sys-

tem, the entropy of trajectory probabilities (entropy of evolution) attains its absolute maximum when and only when the system enters the steady state. In brief, the entropy of evolution is a maximum under stationary conditions. The proposed postulate reduces to the second law of thermodynamics for the case of thermodynamic equilibrium.

If the conditions imposed on the system, which include the conservation laws known for the system, and the conditions maintaining the system in the nonequilibrium state are written in the form

$$F_i = 0, \quad i = 1, 2, \dots, k, \quad (1)$$

where k is the number of conditions, and the F_i are trajectory probability functions, then small changes dF_i in the imposed conditions are related to the change in entropy of evolution dH by the equation

$$dH = \sum_i X_i dF_i. \quad (2)$$

Here the X_i are Lagrange multipliers, which we will refer to as thermodynamic forces. The reasonableness of this definition will become clear in what follows, where it will be shown that near equilibrium the above-defined forces coincide with the usual definition of thermodynamic forces in quasi-equilibrium theory in terms of the rate of entropy production [1]. It should be noted that no conditions are imposed on the concept of nearness to equilibrium, so that the definition of thermodynamic forces introduced has a wider range of applicability than the usual definition.

Because the quantity dH , as defined above, is a total differential (H is defined as a single-valued function of the external conditions), it follows that

$$X_i = \frac{\partial H}{\partial F_i} \quad (3)$$

and

$$\frac{\partial X_i}{\partial F_k} = \frac{\partial^2 H}{\partial F_i \partial F_k} = \frac{\partial X_k}{\partial F_i}. \quad (4)$$

Following the established terminology, we shall refer to the quantities F_i as currents. For a linear dependence of the currents on the forces, which will be shown below to be the case near equilibrium, relation (4) and the linearity relations

$$X_i = \sum_j L_{ij} F_j \quad (5)$$

lead at once to the Onsager relations

$$L_{ij} = L_{ji}. \tag{6}$$

Together with the entropy of evolution $H = H(F_1, F_2, \dots, F_k)$, defined by the entire set of intensive parameters F_i , one can define Φ (the generalized Gibbs function), which is related to H by a Legendre transformation and which depends only on the thermodynamic forces

$$H(F_1, F_2, \dots, F_k) = -\Phi(X_1, X_2, \dots, X_k) + \sum_i X_i F_i. \tag{7}$$

For a stationary state, the quantity

$$d\Phi(X_1, X_2, \dots, X_k) = \sum_i F_i dX_i \tag{8}$$

is a total differential, in accordance with the introduction, and the extremal properties of Φ may be used to find the stationary state itself in X_i coordinates.

Entropy of evolution. For simplicity we will assume that the evolution of the system may be described by a Markov chain with discrete times and a finite number N of states. At first, we will be interested only in the energy states of the system (ϵ_i is the energy of the system in the i -th state, $i = 1, 2, \dots, N$). We denote by $p_{ij}(\tau)$ the conditional probability of a transition on the time interval τ from the state i to the state j , and let p_i denote the stationary (final) probability of the state i , which is found from the system of equations defining the stationary character of the Markov chain,

$$\sum_i p_i p_{ij} = p_j. \tag{9}$$

The conditional transition probabilities and the stationary probabilities satisfy the normalization conditions

$$\sum_j p_{ij} = 1, \tag{10}$$

$$\sum_i p_i = 1. \tag{11}$$

One can describe the trajectory of a Markov chain of length s (s equals the number of steps) by the sequence of states $C_s \doteq i_0 i_1 i_2 \dots i_{s-1} i_s$ with probability

$$P(C_s) = p_{i_0} p_{i_0 i_1} p_{i_1 i_2} \dots p_{i_{s-1} i_s}, \tag{12}$$

which follows from a basic property [2] of Markov chains— independence of the previous history.

If we let m_{ij} denote the number of transitions $i \rightarrow j$ encountered in the chain, then the probability of trajectory C_s is

$$P(C_s) = p_{i_0} \prod_i \prod_j (p_{ij})^{m_{ij}}. \tag{13}$$

If all states are connected by a nonzero transition probability ($p_{ij} > 0$ —ergodic chain), then for a suffi-

ciently long trajectory with high probability one should expect that the relative number m_{ij}/s of appearances of $i \rightarrow j$ transitions will be nearly equal to the probability $p_i p_{ij}$ of such a transition, so that

$$\ln P(C_s) = \ln p_{i_0} + \sum_i \sum_j m_{ij} \ln p_{ij} \tag{14}$$

is almost equal to

$$s \sum_i \sum_j p_i p_{ij} \ln p_{ij}. \tag{15}$$

This also involves neglecting the term $\ln p_{i_0}$, which is constant and whose contribution can be made arbitrarily small by choosing a sufficiently long chain. This also justifies the introduction of the quantity

$$H \equiv H(1) = - \sum_i \sum_j p_i p_{ij} \ln p_{ij}, \tag{16}$$

the entropy per step of the Markov chain.

From (14)–(16) it follows that the probability of a trajectory of length s is equal to

$$P(C_s) = e^{-sH}. \tag{17}$$

As previously indicated, one may regard a random discrete process as a multidimensional random quantity, i. e., as a random vector whose coordinates are random quantities. Therefore, the entropy of a path consisting of s steps is

$$H(s) = - \sum_{\{C_s\}} P(C_s) \ln P(C_s), \tag{18}$$

where the summation extends over all possible trajectories C_s of length s .

The basic properties of the path entropy which we will need are contained in two assertions proved by A. Ya. Khinchin [3], who has rigorously formulated a fact which is evident from relations (14)–(17). It turns out that sufficiently long trajectories can always be divided into two classes. All trajectories in the first class have equal probabilities, given by (17). The second result is also important, namely, that the sum of the probabilities of trajectories of the second class can be made arbitrarily small by choosing s sufficiently large. These assertions show that if we consider those properties of a Markov chain which have unit probability (i. e., the statistical properties), we can confine ourselves to a small portion (e^{sH}) of the very large number N^s of trajectories of length s , assigning the probability e^{-sH} to each trajectory of this class. In this approximation, the evolution entropy (18) is simply expressed in terms of the entropy per step and number of steps:

$$H(s) = - \sum_{\{C_s \in \text{class}\}} e^{-sH} (-sH) = sH. \tag{19}$$

The situation with which we are confronted is completely analogous to the statistical conclusions of the

equilibrium-state theory based on the so-called most probable distribution method [4]. In this method, all possible distributions of a system in equilibrium are also grouped into two classes. The first class contains relatively few distributions—those of highest probability. In the second class are all the remaining distributions, the sum of whose probabilities may be made negligibly small by arbitrarily increasing the set of systems forming the equilibrium Gibbs ensemble. The proof of this is not contained in the probable distribution method, but may be obtained either by using the limit theorems of probability theory [5], or by the mean-value method of Darwin and Fowler [4, 6]. In the latter method, it is shown that the inclusion of all distributions of low probability does not affect the asymptotic conclusions of the most probable distribution method. It can be shown that it is possible to extend the mean-value method to the case of a stationary Markov chain. This confirms the close analogy between the structures of the equilibrium and stationary nonequilibrium theories, which, in our view, should be of considerable heuristic value in developing the macroscopic concepts of the thermokinetics of stationary processes.

Principle of maximum entropy of evolution. Let us consider the stationary evolution of the system. A complete description of its properties is given by the stochastic matrix

$$M = \{p_{ij}\}, \quad (20)$$

and the probabilities of the states are given by (9). If the system is in thermodynamic equilibrium, its state is completely determined by the set $\{p_i\}$. The latter quantities are found by using the second law of thermodynamics, which says that at equilibrium the entropy of the system

$$S = - \sum_i p_i \ln p_i \quad (21)$$

has a conditional maximum.

In describing nonequilibrium systems it is no longer possible to confine oneself to the vector $\{p_i\}$, and one must consider the entire stochastic matrix M . By analogy with the equilibrium case, one may ask if there exists some function of the p_{ij} such that the condition for its extremum might serve to determine the elements of the matrix M .

It is possible to offer several arguments in behalf of our choice of basic postulate.

Suppose that as a result of making measurements on the system, we know the average values of a number of quantities (energy, heat flux, etc.). The observed evolution of the system corresponds to a certain Markov chain trajectory. The measurements correspond to a series of equations satisfied by the stochastic matrix:

$$F_i(p_1, p_2, \dots, p_N; p_{11}, p_{12}, \dots, p_{NN}) = 0. \quad (22)$$

Let us take another system of the ensemble, i. e., another trajectory of the Markov chain with the same

time averages. From the macroscopic standpoint, these systems are indistinguishable insofar as the experimental results are the same for both. We shall use the term "adequate" to describe those trajectories which lead to the same macroscopic results, and refer to the rest as "inadequate". We will choose a certain stochastic matrix that satisfies (22), but is arbitrary in other respects.

It is known that the Markov chain trajectories described by this matrix fall into two classes. All trajectories in the first class that are of sufficient length have nearly equal probabilities, while the sum of the probabilities of trajectories in the second class may be made arbitrarily small. For trajectories of the first class the time-averaged value is the same as the mathematical expectation of the corresponding quantity and hence, assuming that for the stochastic matrix selected the experimental conditions (22) are satisfied, these trajectories are adequate. Both adequate and inadequate trajectories are to be found in the second class. Choosing a different stochastic matrix of greater entropy, which also satisfies (22), results in some of the adequate trajectories of the second class going over into the first class. This is because the new choice of matrix must correspond to a larger number of trajectories in the first class, which does not admit inadequate trajectories. To the observer all adequate trajectories are the same. For them to be equivalent in relation to the process as well, we must choose a stochastic matrix such that all adequate trajectories fall into the first class. To do this, we must make the first class as broad as possible, and since the number of trajectories in it is e^{sH} (where H is the entropy of evolution per step, and s is the number of steps), then the stochastic matrix corresponding to transition of all adequate trajectories into the first class must have the greatest possible entropy of evolution per step.

We are thus led to the conclusion that M , the conditional transition probability matrix satisfying all the imposed conditions (22), must maximize the entropy of evolution per step:

$$H = - \sum_i \sum_j p_i p_{ij} \ln p_{ij}.$$

Of course, this requirement is applicable only to stationary systems whose time step τ is much less than the observation time, i. e., systems which go through a great many steps during the course of the experiment.

Limiting case of equilibrium. The criterion for maximum ordinary entropy must follow from the proposed criterion in the limiting case of equilibrium. According to the theory developed in the present article, the equilibrium case is different from other stationary situations in that it is completely determined by the state vector $\{p_i\}$. This means that the additional conditions impose constraints only on the state vector, as distinct from the general stationary case where they depend explicitly on the conditional transition probabilities.

At equilibrium conditions (22) become

$$F_k(p_1, p_2, \dots, p_N) = 0. \tag{23}$$

Here, the process of determining max H can be divided into two stages. For a given choice of $\{p_i\}$ satisfying (23), we find max H by varying p_{ij} ; then, having obtained p_{ij} as a function of p_i , we find its maximum as a function of p_i . First, we need to prove the inequality

$$-\sum_i q_i \ln q_i \leq -\sum_i q_i \ln p_i \tag{24}$$

for sets of nonnegative numbers normalized to unity. This inequality is a direct consequence of the well-known inequality $\ln x \leq -1 + x$ for any $x > 0$.

The proof of (24) follows, in fact, from the sequence of relations

$$\sum_i q_i \ln x_i \leq \sum_i q_i (-1 + x_i) = -\sum_i q_i + \sum_i p_i = 0,$$

where $x_i = p_i/q_i$.

We will next prove a fundamental inequality between the entropy of evolution and the entropy of the state

$$H \leq S. \tag{25}$$

For sets of numbers $\{p_j\}$ and $\{p_{ij}\}$ normalized to j, inequality (24) gives

$$-\sum_j p_{ij} \ln p_{ij} \leq -\sum_j p_{ij} \ln p_j. \tag{26}$$

By multiplying both sides of (26) by p_i , summing over i, and using the stationary condition (9), we immediately obtain

$$-\sum_i \sum_j p_i p_{ij} \ln p_{ij} \leq -\sum_j p_j \ln p_j, \tag{27}$$

which is a developed form of (25). The equality holds only for

$$p_{ij} = p_j. \tag{28}$$

In this case the entropy of evolution is an absolute maximum.

Thus, if no conditions are imposed on M, at the maximum $H = S$, and subsequent maximization of H with respect to p_i ordinarily leads to a Boltzmann form of probability distribution. Hence from (28) we obtain the following expression for the equilibrium probability of a transition $i \rightarrow j$ on the interval τ :

$$p_i p_{ij} = p_i p_j = [\exp\{(-\varepsilon_i - \varepsilon_j)/kT\}] \times \left[\sum_j \exp(-\varepsilon_j/kT) \right]^{-2}. \tag{29}$$

It is worth noting that (28) is stronger than the principle of detailed balance, which at equilibrium requires only the equality

$$p_i p_{ij} = p_j p_{ji}. \tag{30}$$

In order to determine the elements of the stochastic matrix M in the nonequilibrium case, with k conditions of the general form (22) imposed on the system, it is necessary to solve the variational equation

$$\delta \left(H + \sum_i^k X_i F_i \right) = 0. \tag{31}$$

Among the necessary conditions to be taken into account are the normalization conditions (10) and (11), the stationary condition (9), and also the relation determining the average energy and average heat flux through the system.

Elsewhere [7] we have given a detailed solution of Eq. (31) for the case of a system connected with two heat baths, and have found the explicit form of the stochastic matrix. We also intend to examine the quasi-equilibrium approximation of this solution and will obtain an explicit form of the Onsager coefficients for the case of thermal and particle currents flowing simultaneously through the system.

REFERENCES

1. I. Prigogine, Introduction to the Thermodynamics of Irreversible Processes [Russian translation], IL, 1960.
2. W. Feller, An Introduction to Probability Theory and Its Applications [Russian translation], IL, 1952.
3. A. Ya. Khinchin, UMN, vol. 8, no. 3(55), 1953.
4. E. Schrödinger, Statistical Thermodynamics [Russian translation], IL, 1948.
5. A. Ya. Khinchin, Mathematical Foundations of Statistical Mechanics [in Russian], Gostekhizdat, 1943.
6. K. Huang, Statistical Mechanics [Russian translation], Izd. Mir, 1966.
7. A. A. Filyukov and V. Ya. Karpov, IFZh [Journal of Engineering Physics], 13, 5, 1967.

30 January 1967

Institute of Applied Mathematics
AS USSR, Moscow