Anti-H-theorem in Markov processes

Chang-an Hu
Physics Department, Hupeh Medical College, Wuchang 811495, Hubei, People's Republic of China

(Received 18 November 1985)

In contrast to the H theorem in information theory, proved by Guiaşu and Watanabe, a new theorem is proposed in this paper and called the I theorem. The I theorem can be used to determine of which type is the ergodic Markov process characterized by the decrease of entropy; this is useful for the quantitative study of living systems.

I. INTRODUCTION

Boltzmann's famous H theorem, which gives the increase of the entropy in a thermodynamically isolated system, inspired the study of physical processes at the microscopic level, and this in turn introduced the concepts of Markov chains and Markov processes.[1] Does the H theorem hold in Markov process? Using Shannon's entropy, Guiaşu[2] and Watanabe[3] proved that for a large class of stochastic evolutions of the Markov type, the H theorem holds (without reference to the boundary of the system in which the stochastic evolutions proceed); since then, in view of the H theorem, Guiaşu suspected that Markov chains would be of use in the study of processes characterized by a decrease of entropy, but the original meaning of the sentence quoted from Guiaşu's monograph "...suspected that Markov chains would be of use in the study of processes characterized by..." is that he did not recognize Markov chains can be used to study the processes characterized by a decrease of entropy, so he did not prove that an anti-H-theorem exists in Markov-type evolutions.[4] In this paper, in accordance with the development of nonequilibrium thermodynamics,[5] we shall establish that, for another class of stochastic evolutions of the Markov type, an anti-H-theorem holds, which we call the I theorem, for evolutions in which the entropy decreases monotonically.

To clarify the significance of the hitherto-assumed terms entropy and information, we denote the quantity Shannon entropy by H, which can be derived from the Boltzmann equation,[6] and denote the quantity information by I, which means negentropy, displacement from randomness, or decrease of entropy.[7,8] We express this relation as

$$I=-\Delta H=-(H_e-H_0)=H_0-H_e \ , \tag{1}$$

where H_0 denotes the initial entropy and H_e the final entropy of the system in the transition process.

II. I THEOREM

Let Ω be a finite set; an element $\omega\in\Omega$ represents a state of an open or closed system. Let

$$t_0<t_1<t_2<\cdots<t_n<t_{n+1}<\cdots$$

be an increasing time sequence. The initial probability distribution on the set Ω is

$$P_{t_0}(\omega)\geq 0, \ \ \sum_{\omega} P_{t_0}(\omega)=1 \ \ (\omega\in\Omega) \tag{2}$$

and a transition stochastic matrix family is given as follows:

$$P_{t_n,t_{n+1}}(\omega'\,|\,\omega)>0, \ \ \sum_{\omega'} P_{t_n,t_{n+1}}(\omega'\,|\,\omega)=1$$
$$(\omega'\in\Omega, \ \ \omega\in\Omega; \ \ n=0,1,2,\ldots) \ . \tag{3}$$

Here $P_{t_0}(\omega)$ represents the probability of the state ω at the moment t_0, while $P_{t_n,t_{n+1}}(\omega'\,|\,\omega)$ represents the transition probability from the state ω at the moment t_n into the state ω' at the moment t_{n+1}.

If the successive probabilities of the different states at different moments are given according to the following Markov-type evolution:

$$P_{t_{n+1}}(\omega')=\sum_{\omega} P_{t_n}(\omega)P_{t_n,t_{n+1}}(\omega'\,|\,\omega) \ \ (\omega'\in\Omega, \ \omega\in\Omega) \tag{4}$$

and if at every moment t_n, the entropy on the set Ω of the system's state is given by Shannon's discrete entropy

$$H_{t_n}=-\sum_{\omega} P_{t_n}(\omega)\log P_{t_n}(\omega) \ , \tag{5}$$

where here the logarithms are taken with respect to an arbitrary base greater than unity, then we have the following theorem.

I theorem. If for every moment t_n, the transition probabilities of the matrix $P_{t_n,t_{n+1}}(\omega'\,|\,\omega)$ satisfy the inequalities

$$\lambda_{\omega'}\sum_{\omega} P_{t_n}(\omega)[P_{t_n,t_{n+1}}(\omega'\,|\,\omega)-\delta_{\omega'\omega}]$$
$$=\lambda_{\omega'}[P_{t_{n+1}}(\omega')-P_{t_n}(\omega')]\geq 0 \ \ (\omega'\in\Omega, \ \omega\in\Omega) \tag{6}$$

and if by relabeling the subscripts of the elements [this does not change the corresponding value of the entropy H_{t_n} (Ref. 9)], the probability distribution at any moment can be rearranged into the sequence

$$1>P_{t_n,\max}=P_{t_n}(\omega_1)\geq\cdots\geq P_{t_n}(\omega_{j-1})\geq P_{t_n}(\omega_j)$$
$$\geq\cdots\geq P_{t_n}(\omega_m)>0 \ , \tag{7}$$

then we have

$$H_{t_n}\geq H_{t_{n+1}} \ \ (n=0,1,2,\ldots) \ , \tag{8}$$

that is, the entropy of the system decreases monotonically in this evolution.

Here, $\delta_{\omega'\omega}$ is the Kronecker δ function, defined as

$$\delta_{\omega'\omega}=\begin{cases}1 & \text{if } \omega'=\omega\\ 0 & \text{if } \omega'\neq\omega\ ,\end{cases} \tag{9}$$

and $\lambda_{\omega'}$ is a sign factor, defined as

$$\lambda_{\omega'}=\begin{cases}+1 & \text{if } \omega'=\omega_1,\omega_2,\ldots,\omega_{j-1}\\ -1 & \text{if } \omega'=\omega_j,\omega_{j+1},\ldots,\omega_m\ .\end{cases} \tag{10}$$

In the sequence of (7), $P_{t_n,\max}$ denotes the maximum of $P_{t_n}(\omega)$; $P_{t_n}(\omega_1)$, $P_{t_n}(\omega_2),\ldots,P_{t_n}(\omega_{j-1})$ are those probabilities increased while $P_{t_n}(\omega_j),P_{t_n}(\omega_{j+1}),\ldots,P_{t_n}(\omega_m)$ are those decreased in the evolution; $P_{t_n}(\omega_j)$ is the maximum of the decreased probabilities, m is the number of elements in the set.

Proof. According to (4), (9), and (10), formula (6) may be expanded into the following expressions: in cases where $\omega'=\omega_1,\omega_2,\ldots,\omega_{j-1}$, $\lambda_{\omega'}=+1$, we get

$$P_{t_{n+1}}(\omega')=\sum_{\omega}P_{t_n}(\omega)P_{t_n,t_{n+1}}(\omega'\mid\omega)\geq P_{t_n}(\omega')\ , \tag{11a}$$

and in cases where $\omega'=\omega_j,\omega_{j+1},\ldots,\omega_m$, $\lambda_{\omega'}=-1$, we get

$$P_{t_{n+1}}(\omega')=\sum_{\omega}P_{t_n}(\omega)P_{t_n,t_{n+1}}(\omega'\mid\omega)\leq P_{t_n}(\omega')\ . \tag{11b}$$

This means that during the transition the state probabilities of a system change as follows: those $P_{t_n}(\omega')$ less than $P_{t_n}(\omega_{j-1})$ get less and less, while the $P_{t_n}(\omega')$ greater than $P_{t_n}(\omega_j)$ get larger and larger. Using the symmetric convex property of Shannon's entropy function $H(P_t(\omega_1),P_t(\omega_2),\ldots,P_t(\omega_{j-1}),P_t(\omega_j),\ldots,P_t(\omega_m))$, we can prove the I theorem holds. It is well known that Shannon's entropy function is continuous in each variable $P_t(\omega)$ in the interval (0,1), and is a symmetrical, single-valued function in all variables. The total differential at any time t is

$$dH_t=\frac{\partial H_t}{\partial P_t(\omega_1)}dP_t(\omega_1)+\cdots+\frac{\partial H_t}{\partial P_t(\omega_{j-1})}dP_t(\omega_{j-1})+\frac{\partial H_t}{\partial P_t(\omega_j)}dP_t(\omega_j)+\cdots+\frac{\partial H_t}{\partial P_t(\omega_m)}dP_t(\omega_m)\ . \tag{12}$$

Since the probability distribution is subject to the condition $\sum_\omega P_t(\omega)=1$, there are only $m-1$ independent variables. For the proof of the I theorem, if we select $P_t(\omega_j)$, the maximum of the decreased probabilities, as a dependent variable, namely

$$P_t(\omega_j)=1-[P_t(\omega_1)+\cdots+P_t(\omega_{j-1})+P_t(\omega_{j+1})+\cdots+P_t(\omega_m)]\ , \tag{13}$$

then the partial derivative of H_t with respect to $P_t(\omega_k)$ is

$$\begin{aligned}\frac{\partial H_t}{\partial P_t(\omega_k)}&=\sum_{i=1}^{m}\frac{\partial H_t}{\partial P_t(\omega_i)}\frac{\partial P_t(\omega_i)}{\partial P_t(\omega_k)}\\ &=-\frac{\partial}{\partial P_t(\omega_k)}[P_t(\omega_k)\log P_t(\omega_k)]-\frac{\partial}{\partial P_t(\omega_j)}[P_t(\omega_j)\log P_t(\omega_j)]\frac{\partial P_t(\omega_j)}{\partial P_t(\omega_k)}\\ &=-[\log e+\log p_t(\omega_k)]+[\log e+\log P_t(\omega_j)]\\ &=-\log\frac{P_t(\omega_k)}{P_t(\omega_j)}\ .\end{aligned} \tag{14}$$

In cases $\omega_k=\omega_1,\omega_2,\ldots,\omega_{j-1}$, we have

$$\frac{\partial H_t}{\partial P_t(\omega_1)}=-\log\frac{P_t(\omega_1)}{P_t(\omega_j)}\leq 0$$

since $P_t(\omega_1)\geq P_t(\omega_j)$,

$$\frac{\partial H_t}{\partial_t(\omega_2)}=-\log\frac{P_t(\omega_2)}{P_t(\omega_j)}\leq 0 \tag{15a}$$

since $P_t(\omega_2)\geq P_t(\omega_j)$, and so on up to

$$\frac{\partial H_t}{\partial P_t(\omega_{j-1})}=-\log\frac{P_t(\omega_{j-1})}{P_t(\omega_j)}\leq 0$$

since $P_t(\omega_{j-1})\geq P_t(\omega_j)$; and in cases $\omega_k=\omega_j,\omega_{j+1},\ldots,\omega_m$, we have

$$\frac{\partial H_t}{\partial P_t(\omega_j)}=-\log\frac{P_t(\omega_j)}{P_t(\omega_j)}=0\ ,$$

$$\frac{\partial H_t}{\partial P_t(\omega_{j+1})}=-\log\frac{P_t(\omega_{j+1})}{P_t(\omega_j)}\geq 0\ , \tag{15b}$$

since $P_t(\omega_{j+1})\leq P_t(\omega_j)$, and so on, up to

$$\frac{\partial H_t}{\partial P_t(\omega_m)}=-\log\frac{P_t(\omega_m)}{P_t(\omega_j)}\geq 0\ ,$$

since $P_t(\omega_m) \le P_t(\omega_j)$. Let

$$\Delta P_{t_{n+1}}(\omega) = P_{t_{n+1}}(\omega) - P_{t_n}(\omega) , \tag{16}$$

$$P_{t_\theta}(\omega) = P_{t_n}(\omega) + \theta \Delta P_{t_{n+1}}(\omega) \quad (0 < \theta < 1) \tag{17}$$

and according to (7) we have

$$1 > P_{t_\theta}(\omega_1) \ge P_{t_\theta}(\omega_2) \ge \cdots \ge P_{t_\theta}(\omega_j) \ge \cdots \ge P_{t_\theta}(\omega_m) > 0 , \tag{18}$$

so the formulas (15a) and (15b) hold for $P_{t_\theta}(\omega)$. On the other hand, from (11a) and (11b), we get

$$\Delta P_{t_{n+1}}(\omega_1) \ge 0 , \Delta P_{t_{n+1}}(\omega_2) \ge 0, \ldots , \Delta P_{t_{n+1}}(\omega_{j-1}) \ge 0 ; \tag{19a}$$

$$\Delta P_{t_{n+1}}(\omega_j) \le 0 , \Delta P_{t_{n+1}}(\omega_{j+1}) \le 0, \ldots , \Delta P_{t_{n+1}}(\omega_m) \le 0 . \tag{19b}$$

Hence, according to the mean value theorem, we have

$$\begin{aligned} \Delta H_{t_{n+1}} &= H_{t_{n+1}} - H_{t_n} \\ &= \frac{\partial H_t}{\partial P_t(\omega_1)}\bigg|_{P_{t_\theta}(\omega_1)} \Delta P_{t_{n+1}}(\omega_1) + \frac{\partial H_t}{\partial P_t(\omega_2)}\bigg|_{P_{t_\theta}(\omega_2)} \Delta P_{t_{n+1}}(\omega_2) + \cdots \\ &\quad + \frac{\partial H_t}{\partial P_t(\omega_{j-1})}\bigg|_{P_{t_\theta}(\omega_{j-1})} \Delta P_{t_{n+1}}(\omega_{j-1}) + \frac{\partial H_t}{\partial P_t(\omega_{j+1})}\bigg|_{P_{t_\theta}(\omega_{j+1})} \Delta P_{t_{n+1}}(\omega_{j+1}) + \cdots \\ &\quad + \frac{\partial H_t}{\partial P_t(\omega_m)}\bigg|_{P_{t_\theta}(\omega_m)} \Delta P_{t_{n+1}}(\omega_m) \le 0 . \end{aligned} \tag{20}$$

That is,

$$H_{t_n} \ge H_{t_{n+1}} \tag{21}$$

or

$$H_{t_0} \ge H_{t_1} \ge H_{t_2} \ge \cdots \ge H_{t_i} \ge H_{t_{i+1}} \ge \cdots . \tag{22}$$

Here, the meaning of the bars and subscripts is that after the differentiations are carried out, we replace $P_t(\omega)$ by $P_{t_\theta}(\omega)$ in all terms, namely

$$\begin{aligned} \frac{\partial H_t}{\partial P_t(\omega_k)}\bigg|_{P_{t_\theta}(\omega_k)} &= -\log \frac{P_t(\omega_k)}{P_t(\omega_j)}\bigg|_{P_{t_\theta}(\omega_k)} \\ &= -\log \frac{P_{t_\theta}(\omega_k)}{P_{t_\theta}(\omega_j)} , \end{aligned} \tag{23}$$

which satisfy the inequalities (15a) and (15b). Thus, we come to the conclusion that the entropy H decreases monotonically as the point changes from $[P_{t_0}(\omega_1), \ldots , P_{t_0}(\omega_m)]$ to $[P_{t_{n+1}}(\omega_1), \ldots , P_{t_{n+1}}(\omega_m)]$.

According to the definition of information (1), we may express the information (negentropy) in terms of entropy as

$$I_{t_1} = H_{t_0} - H_{t_1} ,$$

$$I_{t_2} = H_{t_0} - H_{t_2} ,$$

and so on up to

$$I_{t_{n+1}} = H_{t_0} - H_{t_{n+1}} . \tag{24}$$

Then, instead of (22), we have

$$I_{t_1} \le I_{t_2} \le \cdots \le I_{t_{n+1}} . \tag{25}$$

This means the information (negentropy), or the order of the system, increases monotonically during the Markov-type evolution satisfying the conditions (6) and (7), so we call it the I theorem, meaning the information-increasing theorem.

Eventually, we have a theorem with which we can deal with a system (open or closed, such as a living system) changing from the equilibrium state [the corresponding maximum entropy is $H(1/m, 1/m, \ldots , 1/m)$] toward any nonequilibrium state in Markov-type stochastic evolution, in which the entropy will decrease monotonically. However, there is still a practical problem remaining to be solved, and this will be discussed in the next section.

III. *I* CRITERION

The important practical problem in studying the information-processing ability of a living system is of which type is the ergodic Markov process characterized by the decrease of entropy and how to determine this Markov-type process. The answers would be clear by studying the condition that the transition probabilities $P_{t_n, t_{n+1}}(\omega' \mid \omega)$ at different moments are all equal and equal to $P(\omega' \mid \omega)$—the transition probabilities of an ergodic Markov chain. In this case, the inequalities (6) and (7) would hold when $P(\omega'\,|\,\omega)$ satisfies the following formulas:

$$[P(\omega_i\,|\,\omega)-P(\omega_{i+1}\,|\,\omega)]\geq 0\;, \tag{26}$$

$$\lambda_{\omega'}[P(\omega'\,|\,\omega_i)-P(\omega'\,|\,\omega_{i+1})]\geq 0\;, \tag{27}$$

$$\lambda_{\omega'}[P(\omega'\,|\,\omega_m)-P_{t_0}(\omega')]\geq 0\;, \tag{28}$$

where $\omega,\omega'=\omega_1,\omega_2,\ldots,\omega_m$, $i=1,2,\ldots,m-1$. All of these inequalities are together called the I criterion; here ω or ω' will be given as $\omega_1,\omega_2,\ldots,\omega_m$ successively; and to every given ω or ω', ω_i will be given as $\omega_1,\omega_2,\ldots,\omega_{m-1}$ in a similar way. The I criterion can be used to determine which type of Markov processes are characterized just by the decrease of entropy. This may be demonstrated as follows.

At first, let us consider the inequalities (7). According to Eq. (4)

$$\begin{aligned}P_{t_{n+1}}(\omega')&=\sum_{\omega}P_{t_n}(\omega)P_{t_n,t_{n+1}}(\omega'\,|\,\omega)\\&=\sum_{\omega}P_{t_n}(\omega)P(\omega'\,|\,\omega)\;,\end{aligned} \tag{29}$$

we have

$$P_{t_{n+1}}(\omega_i)=\sum_{\omega}P_{t_n}(\omega)P(\omega_i\,|\,\omega)\;,$$

$$P_{t_{n+1}}(\omega_{i+1})=\sum_{\omega}P_{t_n}(\omega)P(\omega_{i+1}\,|\,\omega)\;,$$

inasmuch as the inequalities (26), we get

$$\begin{aligned}&[P_{t_{n+1}}(\omega_i)-P_{t_{n+1}}(\omega_{i+1})]\\&\quad=\sum_{\omega}P_{t_n}(\omega)[P(\omega_i\,|\,\omega)-P(\omega_{i+1}\,|\,\omega)]\geq 0\;,\end{aligned} \tag{30}$$

that is, the inequality series (7) holds.

Next, for the proof of the inequalities (6), we select the probability of the last element $P_{t_n}(\omega_m)$ as a dependent variable, that is

$$P_{t_n}(\omega_m)=1-\sum_{\omega''=\omega_1}^{\omega_{m-1}}P_{t_n}(\omega'')\;, \tag{31}$$

and substituting it into (6), we obtain

$$\begin{aligned}\lambda_{\omega'}\sum_{\omega}P_{t_n}(\omega)[P_{t_n,t_{n+1}}(\omega'\,|\,\omega)-\delta\omega'\omega]&=\lambda_{\omega'}\sum_{\omega}P_{t_n}(\omega)[P(\omega'\,|\,\omega)-\delta\omega'\omega]\\&=\lambda_{\omega'}\left[\sum_{\omega}P_{t_n}(\omega)P(\omega'\,|\,\omega)-P_{t_n}(\omega')\right]\\&=\lambda_{\omega'}\left[\sum_{\omega''=\omega_1}^{\omega_{m-1}}P_{t_n}(\omega'')P(\omega'\,|\,\omega'')+P_{t_n}(\omega_m)P(\omega'\,|\,\omega_m)-P_{t_n}(\omega')\right]\\&=\lambda_{\omega'}\left[\sum_{\omega''=\omega_1}^{\omega_{m-1}}P_{t_n}(\omega'')P(\omega'\,|\,\omega'')+\left[1-\sum_{\omega''=\omega_1}^{\omega_{m-1}}P_{t_n}(\omega'')\right]P(\omega'\,|\,\omega_m)-P_{t_n}(\omega')\right]\\&=\lambda\omega'\left[\sum_{\omega''=\omega_1}^{\omega_{m-1}}P_{t_n}(\omega'')[P(\omega'\,|\,\omega'')-P(\omega'\,|\,\omega_m)]+[P(\omega'\,|\,\omega_m)-P_{t_n}(\omega')]\right]\\&=\lambda_{\omega'}\left[\sum_{\omega}P_{t_n}(\omega)[P(\omega'\,|\,\omega)-P(\omega'\,|\,\omega_m)]+[P(\omega'\,|\,\omega_m)-P_{t_n}(\omega')]\right]\geq 0\;.\end{aligned} \tag{32}$$

Here, we insert a zero term $P_{t_n}(\omega_m)[P(\omega'\,|\,\omega_m)-P(\omega'\,|\,\omega_m)]$ in the last step. Obviously, if the terms in the braces satisfy the conditions

$$\lambda_{\omega'}[P(\omega'\,|\,\omega)-P(\omega'\,|\,\omega_m)]\geq 0\;, \tag{33}$$

$$\lambda_{\omega'}[P(\omega'\,|\,\omega_m)-P_{t_n}(\omega')]\geq 0\;, \tag{34}$$

the inequalities (32) and then the inequalities (6) hold.

However, according to the conditions (27) and (28)

$$\lambda_{\omega'}[P(\omega'\,|\,\omega_i)-P(\omega'\,|\,\omega_{i+1})]\geq 0\;,$$

$$\lambda_{\omega'}[P(\omega'\,|\,\omega_m)-P_{t_0}(\omega')]\geq 0$$

of the I criterion, (34), or (32) hold only for $P_{t_n}(\omega')=P_{t_0}(\omega')$. So we have to prove that, if (32) holds for $P_{t_n}(\omega')=P_{t_0}(\omega')$, they hold automatically for $P_{t_n}(\omega')$ of arbitrary moment. This may be proved in a somewhat different way.

According to (4) and (16) we have

$$\Delta P_{t_{n+1}}(\omega')=P_{t_{n+1}}(\omega')-P_{t_n}(\omega')$$
$$=\sum_{\omega} P_{t_n}(\omega)P(\omega' \mid \omega)-P_{t_n}(\omega')$$
$$=\sum_{\omega} [\Delta P_{t_n}(\omega)+P_{t_{n-1}}(\omega)]P(\omega' \mid \omega)-P_{t_n}(\omega')$$
$$=\sum_{\omega} \Delta P_{t_n}(\omega)P(\omega' \mid \omega) . \qquad (35)$$

Hence, we can rewrite inequalities (32) into a new form

$$\lambda_{\omega} \sum_{\omega} P_{t_n}(\omega)[P(\omega' \mid \omega)-\delta\omega'\omega]$$
$$=\lambda_{\omega'} \sum_{\omega} [P_{t_{n+1}}(\omega')-P_{t_n}(\omega')]$$
$$=\lambda_{\omega'} \Delta P_{t_{n+1}}(\omega')$$
$$=\lambda_{\omega'} \sum_{\omega} \Delta P_{t_n}(\omega)p(\omega' \mid \omega) \geq 0 . \qquad (36)$$

We have noticed that, in the case of $P_{t_n}(\omega')=P_{t_0}(\omega')$, the inequality (32), now in a new form, holds:

$$\lambda_{\omega'} \Delta P_{t_1}(\omega')=\lambda_{\omega'}[P_{t_1}(\omega')-P_{t_0}(\omega')] \geq 0 . \qquad (37)$$

On the basis of (37), we shall prove that (36) holds for any moment.

Taking into account expression (2),

$$\sum_{\omega} P_{t_0}(\omega)=1 ,$$

we have

$$\sum_{\omega} \Delta P_{t_1}(\omega)=0 . \qquad (38)$$

Let us substitute

$$\Delta P_{t_1}(\omega_j)=-[\Delta P_{t_1}(\omega_1)+ \cdots +\Delta P_{t_1}(\omega_{j-1}) +\Delta P_{t_1}(\omega_{j+1})+ \cdots +\Delta P_{t_1}(\omega_m)]$$

into the left side of (36) for $t_{n+1}=t_2$. We get

$$\lambda_{\omega'} \Delta P_{t_2}(\omega')=\lambda_{\omega'} \sum_{\omega} \Delta P_{t_1}(\omega)P(\omega' \mid \omega)$$
$$=\lambda_{\omega'} \sum_{\omega''(\neq\omega_j)} \Delta P_{t_1}(\omega'')[P(\omega' \mid \omega'')-P(\omega' \mid \omega_j)]$$
$$=\lambda_{\omega'} \sum_{\omega} \Delta P_{t_1}(\omega)[P(\omega' \mid \omega)-P(\omega' \mid \omega_j)] . \qquad (39a)$$

Here the expression $\omega''(\neq\omega_j)$ under the sigma denotes the summation of all the terms with the exception of $\Delta P_{t_1}(\omega_j)$, and in the last step we insert a zero term $\Delta P_{t_1}(\omega_j)[P(\omega' \mid \omega_j)-P(\omega' \mid \omega_j)]$. Now, we split Eq. (39a) into the following expressions:

$$\lambda_{\omega'} \Delta P_{t_2}(\omega')=\lambda_{\omega'} \left[\sum_{\omega^+} \Delta P_{t_1}(\omega^+)[P(\omega' \mid \omega^+)-P(\omega' \mid \omega_j)]+ \sum_{\omega^*} \Delta P_{t_1}(\omega^*)[P(\omega' \mid \omega^*)-P(\omega' \mid \omega_j)] \right]$$
$$= \left[\sum_{\omega^+} \Delta P_{t_1}(\omega^+)\lambda_{\omega'}[P(\omega' \mid \omega^+)-P(\omega' \mid \omega_j)]+ \sum_{\omega^*} \Delta P_{t_1}(\omega^*)\lambda_{\omega'}[P(\omega' \mid \omega^*)-P(\omega' \mid \omega_j)] \right] . \qquad (39b)$$

Here $\omega^+=\omega_1,\omega_2,\ldots,\omega_{j-1}$, $\omega^*=\omega_j,\omega_{j+1},\ldots,\omega_m$.

According to (37)

$$\lambda_{\omega'} \Delta P_{t_1}(\omega') \geq 0 ,$$

we have

$$\Delta P_{t_1}(\omega^+) \geq 0 , \qquad (40a)$$

$$\Delta P_{t_1}(\omega^*) \leq 0 . \qquad (40b)$$

On the other hand, from the conditions (27)

$$\lambda_{\omega'}[P(\omega' \mid \omega_i)-P(\omega' \mid \omega_{i+1})] \geq 0 ,$$

we get

$$\lambda_{\omega'}[P(\omega' \mid \omega^+)-P(\omega' \mid \omega_j)] \geq 0 , \qquad (41a)$$

$$\lambda_{\omega'}[P\omega' \mid \omega^*)-P(\omega' \mid \omega_j)] \leq 0 . \qquad (41b)$$

Substituting (40a), (40b), (41a), and (41b) into (39b), we obtain

$$\lambda_{\omega'} \Delta P_{t_2}(\omega') \geq 0 . \qquad (42)$$

That is, (32) holds for $t_{n+1}=t_2$. Likewise, we can prove that (32) holds for arbitrary moments t_{n+1}, namely, (6) holds,

$$\lambda_{\omega'} \Delta P_{t_{n+1}}(\omega') \geq 0 . \qquad (43)$$

Consequently, we have the conclusion that the I theorem is true as the inequalities (26), (27), and (28) hold.

Thus, when an ergodic Markovian transition-probability family $P(\omega' \mid \omega)$ is known, with the help of I criterion, we can determine whether or not the evolution is a process characterized by the decrease of entropy. Furthermore, for the moment, the Markov-type transition probabilities cannot yet be determined experimentally and theoretically; however, the I criterion sheds light on the resolution of this key problem in living systems; we discuss this in the next section.

IV. DISCUSSION

The I theorem is a theorem of phenomenological microscopic stochastic theory concerned with the processes characterized by the decrease of entropy; it can be used to describe the changes of a system from the equilibrium state toward the steady nonequilibrium state in living organisms, in which the steady-state or homeostatic systems, perhaps the most characteristic of biological systems, are not at equilibrium.[10] It is well known in biological theory that life is a temporary reversal of a universal trend toward maximum disorder brought about by the production of information mechanisms.[11] A good example of such an information mechanism is the "Markov machine," since it is much more easily constructed and maintained, and it tends to be less upset by minor injures.[12]

An example of such a mechanism is active transport across a cell membrane. Ryan,[13] following Spanner,[14] introduced a model of it as a completely isolated system, divided into two equal halves A and B, by a semipermeable membrane C, containing water and a number of solutes. Suppose that for the system to be in equilibrium that all of its constituents must be uniformly distributed originally, and then a disturbance in the equilibrium may be caused by an activity of the membrane which causes a redistribution of the solutes within the system; in the transition processes a solute may be in any of three locations: side A, side B, or attached to a carrier located on the membrane C. If the transition probabilities were known, the associated entropy of the system in the steady nonequilibrium state could be calculated from the space of 2-tuples, denoted as Ω^2, the corresponding probabilities are

$$P_{t_n}(\omega)P_{t_n,t_{n+1}}(\omega' \mid \omega)=P_{t_n}(\omega)P(\omega' \mid \omega) , \tag{44}$$

that is,

$$H_{t_{n+1}}(\Omega^2)=-\sum_{\omega}\sum_{\omega'}P_{t_n}(\omega)P(\omega' \mid \omega)\log[P_{t_n}(\omega)P(\omega' \mid \omega)] , \tag{45}$$

when $t_{n+1}=t_\infty$, where $\omega,\omega'=\omega_A,\omega_B,\omega_C$, the elements in location A,B,C; the corresponding maximum entropy $H_{t_0}(\Omega^2)$, of the original equilibrium state associated with this system may be determined when all $P_{t_0}(\omega)$ and $P(\omega' \mid \omega)$ are equiprobable, and the information of the system processed by the membrane activity is

$$I_{t_{n+1}}(\Omega^2)=H_{t_0}(\Omega^2)-H_{t_{n+1}}(\Omega^2) . \tag{46}$$

Thus, we could calculate the average rate of the flow of information in such an active transport model and use it to estimate the information-processing capacity of a living cell if the $H_{t_\infty}(\Omega^2)$ were known.

A main problem involved in this process is that, though the final distribution of solutes, namely, the $P_{t_\infty}(\omega)$, can be determined experimentally,[15] the transition probabilities $P(\omega' \mid \omega)$ so far can neither be determined experimentally nor theoretically.

In theory, according to the properties of ergodic Markov chains, we have[16]

$$P_{t_\infty}(\omega')=\sum_{\omega}P_{t_\infty}(\omega)P(\omega' \mid \omega) , \tag{47}$$

$$\sum_{\omega'}P_{t_\infty}(\omega')=1 . \tag{48}$$

For the known experimentally $P_{t_\infty}(\omega')$, the expression (47) denotes m simultaneous equations with $m(m-1)$ unknown quantities $P(\omega' \mid \omega)$ in Eq. (3)

$$\sum_{\omega'}P(\omega' \mid \omega)=1 . \tag{49}$$

Now, in the light of I criterion, we can determine the Markovian transition probabilities $P(\omega' \mid \omega)$ by trial and error method from expression (47). This way, the set of transition probabilities satisfying the expression (47) is not unique, however, it is sufficient to evaluate the information-processing capacity of a living cell.

ACKNOWLEDGMENT

Projects supported by the Science Fund of the Chinese Academy of Sciences.

[1] P. Ehrenfest and T. Ehrenfest, Z. Phys. **8**, 311 (1907).

[2] S. Guiaşu, C.R. Acad. Sci. (Paris) **261**, 1179 (1965).

[3] S. Watanabe, *Knowing and Guessing. A Quantitative Study of Inference and Information* (Wiley, New York, 1969).

[4] S. Guiaşu, *Information Theory with Applications* (McGraw-Hill, London, 1977), p. 390.

[5] I. Prigogine, *Introduction to Thermodynamics of Irreversible Process* (Thomas, Springfield, Ill., 1955), p. 74.

[6] P. W. Atkins, *Physical Chemistry*, 2nd ed. (Oxford University Press, Oxford, 1982), p. 684.

[7] J. Rothstein, Science **114**, 171 (1951).

[8] R. Ash, *Information Theory* (Wiley, New York, 1965), p. 21.

[9] P. M. Reza, *Introduction to Information Theory* (McGraw-Hill, New York, 1961), p. 16.

[10] E. Ackerman, L. B. M. Ellis, and L. E. Williams, *Biophysical Science*, 2nd ed. (Prentice-Hall, New Jersey, 1979), p. 413.

[11] S. W. Angrist, L. G. Hepler, *Order and Chaos: Laws of Energy and Entropy* (Penguin, Hermondsworth, 1973).

[12] W. R. Ashby, *An Introduction to Cybernetics* (Chapmann & Hall, London, 1957), p. 231.

[13] J. P. Ryan, J. Theor. Biol. **84**, 31 (1980).

[14] D. C. Spanner, Symp. Soc. Exp. Biol. **VIII**, 76 (1954).

[15] H. Ryan, J. P. Ryan, and W. H. O'Connor, J. Biochem. **125**, 1081 (1971).

[16] See Ref. 9, p. 377.