Maximum path information and the principle of least action for chaotic system

Qiuping A. Wang

Maximum path information and the principle of least action for chaotic system

Q.A. Wang
Institut Supérieur des Matériaux et Mécaniques Avancés,
44, Avenue F.A. Bartholdi, 72000 Le Mans, France

Abstract

A path information is defined in connection with the different possible paths of chaotic system moving in its phase space between two cells. On the basis of the assumption that the paths are differentiated by their actions, we show that the maximum path information leads to a path probability distribution as a function of action from which the well known transition probability of Brownian motion can be easily derived. An interesting result is that the most probable paths are just the paths of least action. This suggests that the principle of least action, in a probabilistic situation, is equivalent to the principle of maximization of information or uncertainty associated with the probability distribution.

PACS number : 05.45.-a;05.70.Ln;02.50.-r;05.45.-a

1 Introduction

The aim of this work is to investigate the probability distributions attributed to the different paths (path probability distribution) of chaotic system moving between two points in the phase space Γ. As usual, the phase space Γ of a system is defined such that a point in it represents a state of the system. For a N-body system moving in three dimensional ordinary configuration space, the Γ-space is of $6N$ dimension ($3N$ positions and $3N$ momenta) if it can be smoothly occupied. A system at equilibrium state is represented by a fixed point in Γ space.

Now we look at a nonequilibrium system moving in the Γ-space between two points, a and b, which are in two elementary cells of a given partition of the phase space (Figure 1). We will use the concept of trajectory or path of classical mechanics. If the motion of the system is regular, there will be only one possible trajectory between the two points, or, in other words, there will be only a fine bundle of paths which track each other between the initial and the final cells. These trajectories must be the path of least action according to the principle of Maupertuis[1] and have unitary probability. Any other path must have zero probability.

For a system in chaotic motion, the things can be different. Two points indistinguishable in the initial cell will separate from each other exponentially. Normally, these two points, after their departure from the initial cell, will never meet each other in a final cell in the phase space. However, it is possible that they pass through a same cell at two different times. So between two given phase cells represented in Figure 1 by a and b, respectively, there may be many possible paths labelled by k (k=1,2,...w) with different travelling time $t_{ab}(k)$ of the system and different probability $p_{ab}(k)$ for the system to take the path k (path probability distribution).

The path probability distribution is defined as follows. Suppose an ensemble of a large number L of identical systems all moving in the phase space from cell a to cell b with w possible paths. We observe L_k systems travelling on the path k ($k = 1, 2...w$). The probability $p_{ab}(k)$ that the system take the path k is thus defined as usual by $p_{ab}(k) = \frac{L_k}{L}$. We naturally have $\sum_{k=1}^{w} p_{ab}(k) = 1$. By definition, $p_{ab}(k)$ is a transition probability from state a to state b.

In this paper, the path probability distribution will be studied in connection with information theory and the principle of least action[1]. Inspired by this universal principle from which almost all the physics (classical mechanics, quantum mechanics, Maxwell equations, optics ...) can be derived, we suppose that the different paths between a and b are differentiated by their action defined by

$$A_{ab}(k) = \int_{t_{ab}(k)} L_k(t)dt \tag{1}$$

where $L_k(t)$ is the Lagrangian of the system at time t along the path k and is defined by $L_k(t) = U_k(t) - V_k(t)$ where $U_k(t)$ is the total kinetic energy

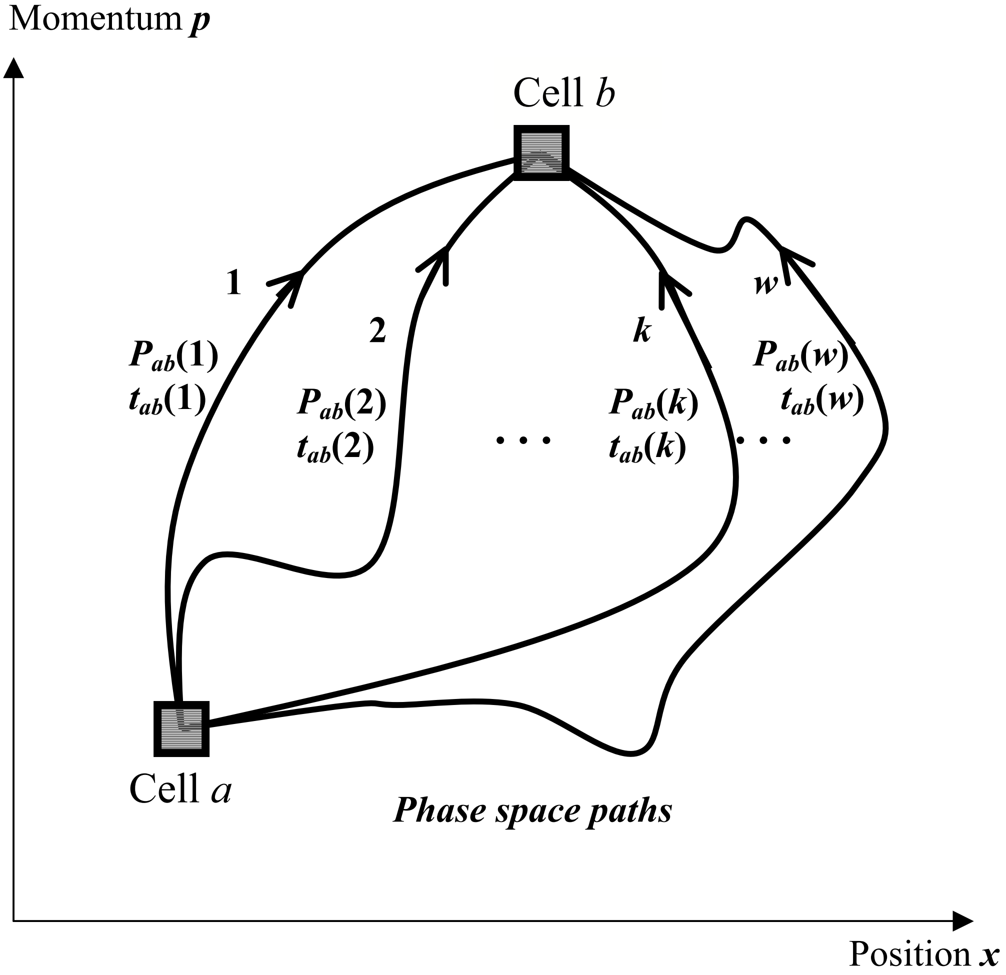

Figure 1: Possible phase space paths ($k = 1, 2...w$) for a chaotic system to go from the points in the phase cell a to the points in the phase cell b during the time $t_{ab}(k)$ (along the path k).

and $V_k(t)$ the total potential energy of the system. The integral is carried out along a path k during $t_{ab}(k)$, the travelling time of the system along the path. This is an essential assumption of this work. If the paths can be identified only by their actions, then it will be possible to study their probability distributions with the information concept and the method of maximum information of Jaynes[2] in connection with our knowledge about action. This approach will leads us to a probabilistic interpretation of the mechanical principle of Maupertuis and a probability distribution depending on action. As an application, the obtained path probability distribution will be used to derive the transition probability of Brownian motion.

2 A path information

The information we address here is our ignorance about the system under consideration. More we know about the system, less there is information. According to Shannon[3], this information can be measured by the formula $S = -\sum_i p_i \ln p_i$ where p_i is certain probability attributed to the situation i. We usually ask $\sum_i p_i = 1$ with a summation over all the possible situations.

Now for our ensemble of w possible paths in Figure 1, a Shannon information can be defined as follows :

$$H(a, b) = -\sum_{k=1}^{w} p_{ab}(k) \ln p_{ab}(k). \tag{2}$$

$H(a, b)$ is a *path information* and should be interpreted as the missing information necessary for predicting which path a system of the ensemble takes from a to b.

According to our starting assumption, the quantity that differentiates the paths and their probability of occurrence is the Lagrangian action. In what follows, a statistics is developed on the basis of this assumption.

3 Probability distribution of maximum information

We consider the ensemble containing a large number of the studied system moving from a to b. These systems are distributed over the w paths according to $p_{ab}(k)$ in connection with the action $A_{ab}(k)$. An expectation of the action over all the possible paths can then be calculated by

$$A_{ab} = \sum_{k=1}^{w} p_{ab}(k) A_{ab}(k). \tag{3}$$

On the other hand, the path information $H(a, b)$ in Eq.(2) is concave as a function of normalized probability $p_{ab}(k)$. According to the principle of Jaynes, in order to get the optimal distribution, $H(a, b)$ can be maximized under the constraints associated with our knowledge about the system and the relevant random variables, i.e., with the normalization of $p_{ab}(k)$ and the expectation A_{ab} :

$$\delta \left[-H(a, b) + \alpha \sum_{k=1}^{w} p_{ab}(k) + \eta \sum_{k=1}^{w} p_{ab}(k) A_{ab}(k) \right] = 0 \tag{4}$$

This leads to the following distribution

$$p_{ab}(k) = \frac{1}{Q} \exp[-\eta A_{ab}(k)] \tag{5}$$

Putting this distribution into $H(a, b)$ of Eq.(2), we get

$$H(a, b) = \ln Q + \eta A_{ab} \tag{6}$$

where Q is given by $Q = \sum_{k=1}^{w} \exp[-\eta A_{ab}(k)]$.

4 Stability of the path probability distribution

Now we shall show that the above distribution is stable with respect to the fluctuation of action. Suppose that each path is cut into two parts 1 (the segments on the side of the cell a) and 2 (the segments on the side of b). All the segments 1 are contained in the group 1, and all the segments 2 in the group 2. Each group has a path information $H_1 = H_2 = H$ and an average action $A_1 = A_2 = A$. The total information is then $H(a, b) = H_1 + H_2 = 2H$ and total average action is $A(a, b) = A_1 + A_2 = 2A$. If now we consider a small variation of the division of the paths with virtual changes in the two groups such that $\delta A_1 = \delta A = -\delta A_2$, the total information will be changed and can be written as

$$H'(a, b) = H(A + \delta A) + H(A - \delta A) \tag{7}$$

In view of the fact that the distribution Eq.(5) and the relationship Eq.(6) are consequence of maximum information, the stability condition requires that the information does not increase with the virtual changes of the two groups. We must have

$$\delta H = H'(a, b) - H(a, b) \leq 0, \tag{8}$$

i.e.,

$$H(A + \delta A) + H(A - \delta A) - 2H(A) \leq 0 \tag{9}$$

which means

$$\frac{\partial^2 H}{\partial A^2} \leq 0. \tag{10}$$

Now let us see if this stability condition is always fulfilled. From Eq.(6), one gets $\frac{\partial^2 H}{\partial A^2} = \frac{\partial \eta}{\partial A}$. Then considering the definition of average action Eq.(3), we straightforwardly calculate

$$\frac{\partial A}{\partial \eta} = -\sigma^2, \tag{11}$$

which implies

$$\frac{\partial^2 H}{\partial A^2} = -\frac{1}{\sigma^2} \leq 0 \tag{12}$$

where the variance $\sigma^2 = \overline{A^2} - \overline{A}^2 \geq 0$ characterizes the fluctuation of the action A. This proves the stability of the maximum information distribution Eq.(5) with respect to the action fluctuation of the paths.

5 Application to Brownian motion

What is the parameter η in the path probability distribution Eq.(5)? A possible physical meaning of η can be found with a special example : Brownian motion. Suppose a certain path in Figure 1 along which a Brownian particle moves from a to b via an intermediate point or cell k. Between the three cells a, k, and b, the particle is free. The action $A_{ab}(k)$ of the particle from a to b can be calculated to be[5]

$$A_{ab}(k) = \frac{m(x_k - x_a)^2}{2(t_k - t_a)} + \frac{m(x_b - x_k)^2}{2(t_b - t_k)}. \tag{13}$$

Then from Eq.(5), we have

$$p_{ab}(k) = \frac{1}{Q} \exp\left[-m\eta \frac{(x_k - x_a)^2}{2(t_k - t_a)}\right] \exp\left[-m\eta \frac{(x_b - x_k)^2}{2(t_b - t_k)}\right] \tag{14}$$

On the other hand, it is known[4] that, as a solution of the diffusion equation, the transition probability for the particle to go from a to b via k is

$$\begin{aligned} p_{ab}(k) &= p_{ak} p_{kb} \\ &= \frac{1}{[4\pi D(t_k - t_a)]^{d/2}} \exp\left[-\frac{(x_k - x_a)^2}{4D(t_k - t_a)}\right] \\ &\times \frac{1}{[4\pi D(t_b - t_k)]^{d/2}} \exp\left[-\frac{(x_b - x_k)^2}{4D(t_b - t_k)}\right] \end{aligned} \tag{15}$$

where D is the diffusion coefficient supposed constant everywhere in the phase space, t_a, t_k and t_b are the time and x_a, x_k and x_b the position coordinates of the particle at a k and b, respectively, d is the dimension of the ordinary configuration space, p_{ak} and p_{kb} are the transition probabilities of the particle from a to k and from k to b, respectively. A comparison of Eq.(14) and Eq.(15) leads to

$$\eta = \frac{1}{2mD} \tag{16}$$

where m is the mass of the particle. We see that the parameter η is related to the diffusion coefficient D.

For a system containing a large number of Brownian particles, the above result is still valid. The only difference is that, in this case, there are just more intermediate points and the total action will be calculated over all particles each having a large number of intermediate points. The above result can be used for many thermodynamic systems modelled with Brownian particles.

If we suppose that the self diffusion coefficient can be approximated by Stokes-Einstein relation $D = \frac{k_B T}{6\pi\eta_0 R}$ of a self-diffusion of the Brownian particles in a dilute medium, we get

$$\eta = \frac{6\pi\eta_0 R}{2mk_B T} \tag{17}$$

where η_0 is the viscosity of the medium, R is the radius of the Brownian particles, k_B is the Boltzmann constant and T the temperature of the medium. In this case, the path probability distribution Eq.(5) becomes

$$p_{ab}(k) = \frac{1}{Q} \exp[-\frac{\Lambda_{ab}(k)}{k_B T}]. \tag{18}$$

where $\Lambda_{ab}(k) = \frac{6\pi\eta_0 R}{2m} A_{ab}(k)$. This shows a temperature dependence of path probability distribution.

6 The principle of maximum information and the principle of least action

Now we turn our attention to the connection between maximum path information and least action. It can be shown (see below) that the paths of least action are the most probable paths provided $\eta = \frac{\partial H(a,b)}{\partial \overline{A}_{ab}} > 0$. In fact, from Eq.(5), positive η obviously implies that the paths of smaller action are statistically more probable than the paths of larger action. So the most probable paths must minimize the action.

This property of the distribution Eq.(5) can be mathematically discussed in the same manner as the stability of the distribution proved in section 4. We still consider the two groups 1 and 2 of path segments with $H_1 = H_2 = H$ and $A_1 = A_2 = A$. The total information is then $H(a,b) = 2H$ and the total average action is $A(a,b) = 2A$. Now suppose that the two groups are deformed such that $\delta H_1 = \delta H = -\delta H_2$. The total average action after the group deformation can be written as

$$\begin{aligned} A'(a,b) &= A_1(H_1 + \delta H_1) + A_2(H_2 + \delta H_2) \\ &= A(H + \delta H) + A(H - \delta H). \end{aligned} \tag{19}$$

If the distribution Eq.(5) and the relationship Eq.(6) correspond to least action, the total average action after the group deformation can not decrease: $\delta A = A'(a,b) - A(a,b) \geq 0$, i.e.,

$$A(H + \delta H) + A(H - \delta H) - 2A(H) \geq 0 \tag{20}$$

which means

$$\frac{\partial^2 A}{\partial H^2} \geq 0. \tag{21}$$

On the other hand, with the help of Eq.(6), we can prove:

$$\frac{\partial^2 A}{\partial H^2} = -\frac{1}{\eta^2}\frac{\partial \eta}{\partial H} = -\frac{1}{\eta^3}\frac{\partial \eta}{\partial A}. \tag{22}$$

Now considering $\frac{\partial \eta}{\partial A} = -\frac{1}{\sigma^2}$, we get

$$\frac{\partial^2 A}{\partial H^2} = \frac{1}{\sigma^2 \eta^3}. \tag{23}$$

We see that if Eq.(21) is true, we have

$$\eta \geq 0. \tag{24}$$

In other words, the positivity of η implies that the principle of maximum path information is intrinsically connected with the principle of least action:

the most probable paths given by the distribution of maximum information are just the paths of least action.

In view of the relationship Eq.(16), the positivity of η is ensured by the positivity of diffusion coefficient D. We can also invert this statement: if the most probable paths derived from the probability distribution Eq.(5) minimize the action, then the diffusion coefficient must be positive.

7 Concluding remarks

It is hoped that this work may contribute to the study of the behaviors of chaotic systems. If there is no chaos, the path information will vanish and there will be between two phase cells only a fine bundle of parallel paths which are the paths of least action with an unitary probability of occurrence. More the system under consideration is chaotic, more there are possible paths with different actions, and larger the information is. So we conjecture that the path information $H(a, b)$ may be used as a measure of chaos, like Kolmogorov-Sinai entropy (KSE)[6]. It should be noted that there is an important difference between $H(a, b)$ and KSE. $H(a, b)$ is an information associated with different paths relating two phase cells but having arbitrarily different travelling time. In the language of discrete time iteration, the different paths have arbitrarily different number of step of iteration. On the other hand, KSE can be defined as an information associated with different paths that leave the initial cell with arbitrary destination but same travelling time, i.e., same number of step of iteration[6]. Further investigation is necessary in order to clarify the relationship between these two information measures for chaotic systems.

Summarizing, a path information is defined for an ensemble of possible paths of chaotic systems moving between two cells in phase space. It is shown that, if we suppose that the different paths are identified by their actions, the maximization of the path information leads to a path probability distribution as a function of the action. As a special example, the transition probability of Brownian particles is derived. In this case, we show that the most probable paths derived from maximum information minimize the action. This suggests that the principle of least action, in probabilistic situation, is equivalent to the principle of maximization of information or uncertainty associated with the probability distribution. This result may be considered as an argument supporting this method of "unbiased guess"[2] for nonequilibrium systems.

References

[1] For a brief introduction to this principle, see for example $http : //perso.club - internet.fr/jac_leon/$

[2] E.T. Jaynes, *The evolution of Carnot's principle*, The opening talk at the EMBO Workshop on Maximum Entropy Methods in x-ray crystallographic and biological macromolecule structure determination, Orsay, France, April 24-28, 1984; Reprinted in Ericksen & Smith, Vol 1, pp. 267-282

[3] C.E. Shannon, A mathematical theory of communication, *Bell system Technical Journal,* **27**(1948)pp.379-423, 623-656

[4] R. Kubo, M. Toda, N. Hashitsume, *Statistical physics II, Nonequilibrium statistical mechanics,* Springer, Berlin, 1995

[5] R.P. Feynman, *Quantum mechanics and path integrals,* McGraw-Hill Publishing Company, New York, 1965

[6] J.R. Dorfmann, *An introduction to Chaos in nonequilibrium statistical mechanics*, Cambridge University Press, 1999