QUANTIFYING DYNAMIC TRAPS IN HIGH-DIMENSIONAL PHASE SPACE OF MOLECULAR ENSEMBLES

V.B. Ryabov[1, a], D. Nerukh[2, b]
[1]Future University Hakodate
116-2 Kamedanakano-cho, Hakodate, Hokkaido, Japan 041-8655
[2]Non-linearity and Complexity Research Group
Aston University, Aston Triangle, Birmingham, B4 7ET, UK
E-mail: [a]*riabov@fun.ac.jp*, [b]*D.Nerukh@aston.ac.uk*

Abstract

Deterministic chaos in systems with multiple degrees of freedom like, e.g., an ensemble of water molecules, appears indistinguishable from noise by any standard statistical method. At the same time, the time evolution of trajectories in the phase space is described by deterministic differential equations; hence the signatures of deterministic Hamiltonian dynamics are contained in the time dependent coordinates and momenta of each atom. In this paper, we analyze scalar time series originating from a numerical simulation of a large ensemble of water molecules with the purpose of finding a way of distinguishing it from a purely stochastic signal. We consider information theory as a bridge between deterministic chaos in nonlinear dynamical systems with few degrees of freedom and apparently random trajectories in a high-dimensional phase space. Our analysis of statistical complexity of the experimental time series in terms of Markov chain theory with memory allows making a quantitative distinction between high-dimensional deterministic motion and random noise. The approach is based on the construction of ε-machine that decomposes the phase space into non-overlapping elementary areas, which, depending on their decay laws for Poincare recurrence times, can be further divided between two qualitatively different classes. By an analogy with Standard map, they can be attributed to the chaotic sea and quasiperiodic motions in the vicinity of sticky dynamic traps. The proposed method is thus based on identifying areas with sticky dynamics in a high-dimensional phase space that has far reaching implications in understanding the molecular transport, including the anomalous diffusion process.

1. Introduction

Newtonian equations of motion describe the trajectories of atoms and molecules in liquids. Therefore, the dynamics of any particle is deterministic and most probably chaotic [1] characterized by positive value of the largest Lyapunov exponent. On the other hand, it is known that snapshots of molecular liquids obtained either in numerical simulations or real experiments often demonstrate highly structured behavior, i.e. atoms form patterns like, e.g. famous tetrahedral conformation of clusters in bulk water. The immediate problem thus appears on the description of such patterns in terms of deterministically chaotic dynamics of particles, i.e. finding geometric counterparts of molecular patterns in the phase space, the space defined by all the coordinates and velocities of all the particles in the analyzed molecular ensemble. The problem of describing and predicting the appearance of such patterns is crucially important since they ultimately define the functionality of the molecular systems and dynamics of such processes as, for example, protein folding. There is, however, a fundamental difficulty in the dynamical picture of molecular systems due to high-

dimensionality of the corresponding phase space. Commonly used approaches from non-linear dynamics [2], such as Lyapunov exponents, dimensions, and entropies fail if the motions occur in the phase space of dimension higher than ~10. Therefore, new conceptually different methodologies have to be developed for such intrinsically high-dimensional systems.

One of the most difficult problems in the analysis of the high-dimensional molecular trajectories is the very definition of the notion of "cluster" in the phase space. We address this issue in a broad statistical sense considering deviations from the uniform phase space filling by a typical trajectory as clusters. The clusters appear in the phase space due to the presence of abundant resonances that arise as a result of nonlinear interactions between atoms. The borders of resonant areas are known to be "sticky" in a sense that any trajectory spends a long time in their vicinity in contrast to other, non-resonant areas where trajectories move randomly filling the phase space almost uniformly.

2. Symbolic dynamics

We analyze the trajectories of computer simulated bulk water at room temperature sampled at discrete time moments. Strictly speaking, we analyze the motion of one of the hydrogen atoms of a randomly selected molecule in the ensemble of 392 water molecules. In the case of normal diffusion, displacement of the atom along any of Cartesian coordinates is well approximated by a by a random walk model.

In order to avoid the influence of short time correlations caused by rotations of hydrogen atom around the axis along hydrogen bond we down-sample the initially three-dimensional trajectory by introducing a two-dimensional cross-section (Fig.1). This procedure approximately corresponds to sampling the data at a rate defined by the first minimum of the autocorrelation function, a value widely used in analyses of nonlinear dynamics. It further appears convenient to replace the real-valued signal with symbolic string obtained by using a few-symbol alphabet, for example, the three symbol one that we used in our studies {0; 1; 2}. Symbolizing the trajectory corresponds to introducing a partition of the phase space into non-overlapping areas, where each element of the partition corresponds to a symbolic sub-sequence (word). Considering symbolic words of finite length, we obtain the reconstruction of high-dimensional dynamics from the analysis of a scalar time series, similar to the procedure based on Takens embedding theorem [3].

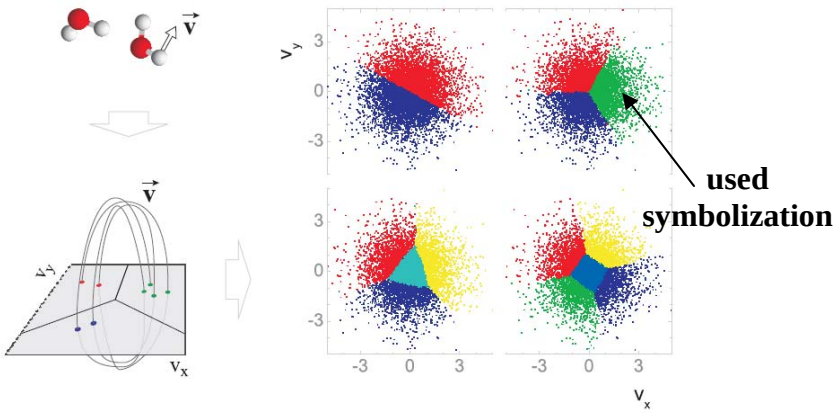

Fig. 1: Symbolization of 3-d velocity time series. From velocity (upper left) to the points on the cross section plane (lower left), and to symbolization with different alphabets containing from 2 to 5 symbols (right).

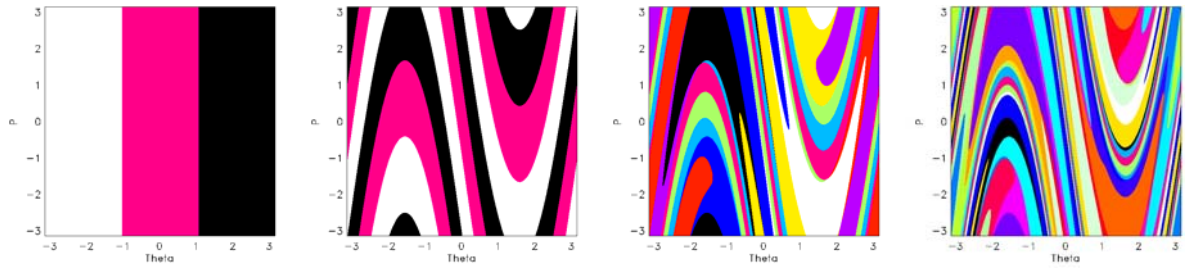

Fig.2. Partitions caused by considering symbolic words of different length in the case of the Standard map. Words partition the phase space into 3^n non-overlapping segments, where n is the word length. Cases of n=1,2,3,4 are shown from left to right.

3. Poincare recurrence

Quantitative description of the non-uniformity of the phase space can be achieved via the Poincare recurrence theory [4]. Consider a small element $\Delta\Gamma$ of the phase space Γ of a Hamiltonian system located around the point x. A trajectory wanders in the chaotic area visiting the element $\Delta\Gamma$ from time to time (recurring to it). Denoting the time between successive recurrences as τ, the probability distribution function of recurrence times $P(\Delta\Gamma, x, \tau)$ can be introduced. If the motion is ergodic the dependence of τ on the coordinates x becomes inessential and one can introduce the distribution function

$$P(\tau) = \lim_{\Delta\Gamma \to 0} P(\Delta\Gamma, \tau)/\Delta\Gamma \qquad (1)$$

For a typical chaotic trajectory the following asymptotic relation can be expected to hold

$$P(\tau) = \exp\left(-\tau/\langle\tau\rangle\right)/\langle\tau\rangle \qquad (2)$$

where $\langle\tau\rangle$ is the average recurrence time over the distribution $P(\tau)$. Eq. (2) then can be used for distinguishing areas with chaotic motion from those close to sticky areas by introducing a partition of the phase space into non-overlapping volumes and analyzing the distributions $P(\tau)$ for each of them. Note also, that the problem of choice of the sizes and shapes of the partition elements is not trivial one and, in the general case the distribution of the Poincare recurrence times, can depend on the location and shape of the area $\Delta\Gamma$.

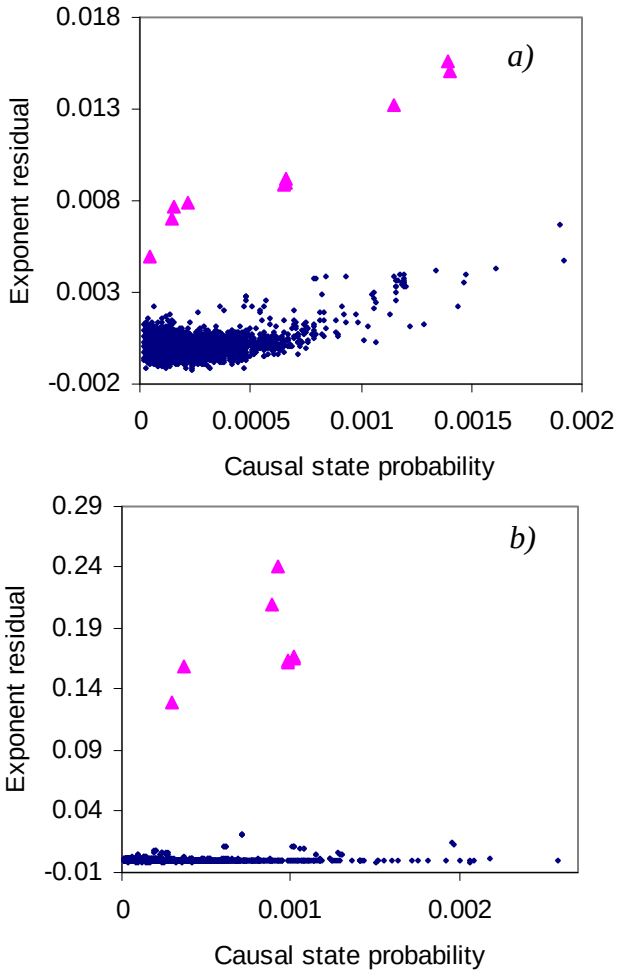

Fig. 3. Clustering of causal states for the hydrogen atom velocity time series (a) and the Standard map (b) into 'periodic' (diamonds) and 'chaotic' (triangles) classes. Parameter *D* is plotted vs. occurrence rates of the causal states.

3. ε-machine: a Markov chain with memory

We use Computational Mechanics, a promising new concept aimed at building a statistical and at the same time dynamical description of time series. It combines a well developed theory of Markov chains, with the concept of short-time predictability, characteristic to dynamical systems. As a result a model called ε-machine is built that provides a 'natural' partitioning in the phase space by means of causal states (groups of symbolic words). Every causal state can be considered as an element of the phase space that contains a set of Poincare cycles distributed in accordance with Eq. (2).

Moreover, considering the deviations from Eq. (2) allows identification of the two types of motion in the phase space, periodic and chaotic, i.e. separating the causal states with special properties that correspond to the resonance areas within chaotic sea in the phase space. Here we introduce a dimensionless parameter $D = \left(1/\lambda\langle\tau\rangle - 1\right)\sigma$,where λ is the exponent defining the shape of the distribution function $P(\tau) \propto \exp(-\lambda\tau)$ found numerically, $\sigma = \sqrt{p}$, where p is the occurrence of the given causal state. Parameter *D* is calculated for each of the causal states, and it equals to the normalized discrepancy between the decay exponent λ calculated from the histogram of recurrence times and its 'normal' value $1/\langle\tau\rangle$ defined by the Eq. (2). Large *D* values indicate strong discrepancy between the calculated value of the exponent λ and its expected value $1/\langle\tau\rangle$.

In Fig. 3 we plot the scatter diagrams representing the apparent clustering of the causal states into two classes with respect to the parameter *D*. The horizontal axis approximates the occurrence rate of the causal states. Every point in Fig.3 corresponds to a causal state.

The apparent clustering of points in Fig.3 evidences the presence of the areas in the phase space where the trajectory spends longer time compared to the rest of the accessible volume. Such areas can not be detected easily by other methods, most probably due to abundance of resonant areas in the high-dimensional phase space that makes difficult a clear distinction between chaotic and quasi-periodic motions.

From a different perspective, our method has a special importance for the problem of quantifying transport properties in high-dimensional molecular systems, since it reveals a (small) number of areas in the phase space playing crucial importance for particle motion. Finding such areas from the analysis of a scalar time series can be very useful in interpreting numerical experiments with large number of interacting particles that typically generate huge volumes of data. Extracting the essential information from the trajectory of a single test particle thus looks a promising approach, for example, in modeling the process of protein folding or dynamics of complex bio-molecules.

References

[1] P. Gaspard, M. E. Briggs, M. K. Francis, J. V. Sengers, R. W. Gammon, J. R. Dorfman & R. V. Calabrese, Experimental evidence for microscopic chaos. Nature 394, 865-868, 1998

[2] J.-P. Eckmann & D. Ruelle, Ergodic theory of chaos and strange attractors. Rev. Mod. Phys. 57, 617-656, 1985

[3] F. Takens, Detecting strange attractors in turbulence. In D. A. Rand and L.-S. Young, eds. Dynamical Systems and Turbulence, Lecture Notes in Mathematics, vol. 898. Springer-Verlag. pp. 366–381, 1981

[4] G. Zaslavsky, Chaos, fractional kinetics, and anomalous transport. Phys. Repts. 371, 461-580, 2002