The Journal of Chemical Physics

Complexity of classical dynamics of molecular systems. II. Finite statistical complexity of a water–Na^+ system

Cite as: J. Chem. Phys. **117**, 9618 (2002); https://doi.org/10.1063/1.1518011
Submitted: 17 July 2002 . Accepted: 09 September 2002 . Published Online: 12 November 2002

Dmitry Nerukh, George Karvounis, and Robert C. Glen


View Online

Export Citation

ARTICLES YOU MAY BE INTERESTED IN

Complexity of classical dynamics of molecular systems. I. Methodology
The Journal of Chemical Physics **117**, 9611 (2002); https://doi.org/10.1063/1.1518010

AIP Publishing

© 2002 American Institute of Physics.

Complexity of classical dynamics of molecular systems. II. Finite statistical complexity of a water–Na^+ system

Dmitry Nerukh, George Karvounis, and Robert C. Glen
Unilever Centre for Molecular Informatics, Department of Chemistry, Cambridge University, Cambridge CB2 1EW, United Kingdom

(Received 17 July 2002; accepted 9 September 2002)

The computational mechanics approach has been applied to the orientational behavior of water molecules in a molecular dynamics simulated water–Na^+ system. The distinctively different statistical complexity of water molecules in the bulk and in the first solvation shell of the ion is demonstrated. It is shown that the molecules undergo more complex orientational motion when surrounded by other water molecules compared to those constrained by the electric field of the ion. However the spatial coordinates of the oxygen atom shows the opposite complexity behavior in that complexity is higher for the solvation shell molecules. New information about the dynamics of water molecules in the solvation shell is provided that is additional to that given by traditional methods of analysis. © *2002 American Institute of Physics.* [DOI: 10.1063/1.1518011]

I. INTRODUCTION

Molecular systems are Hamiltonian nonlinear dynamical systems. They are also large systems consisting of small, relatively simple interacting parts. Systems with these characteristics are capable of exhibiting self-organizing complex behavior. It has been recognized that this type of behavior is quite general for a wide class of nonlinear dynamical systems. In view of this analysis, the emergent behavior of physical systems and their information-theoretical content has become an active area of research. Methods for estimating the complexity of physical systems have attracted close attention. However, very few investigations are devoted to this kind of analysis of molecular systems. It is therefore a promising line of research to focus on the details of the dynamical complexity of molecular systems.

To achieve this we need first to choose a relevant complexity measure among the variety of existing theoretical approaches. Second, a specific implementation of the methodology should be devised, in particular, a practical algorithm for symbolization of physical data. And, finally, the method should be applied to a particular characteristic of a molecular system. The results on the first two stages of the study are reported in our accompanying paper[1] while the details on this particular application to a molecular system are presented here.

It has been shown in Ref. 1 that computational mechanics by Crutchfield *et al.*[2–4] can be effectively used as a measure of complexity of continuous physical trajectories. This approach is a well-developed methodology nicely combining fundamental ideas such as Shannon entropy and Kolmogorov–Chaitin algorithmic complexity providing at the same time practical methods for computing complexity. One of the advantages of this approach is that it is based on informatic-theoretical analysis of the dynamical evolution of the system and opens up the possibility to explore details of the key events in the systems behavior.

The particular implementation of computational mechanics based on the method reported in Ref. 5 is discussed in Ref. 1. It is shown that it can be consistently used for calculating dynamic complexity—the finite statistical complexity in this case. Particular attention has been given to a method of converting the physical trajectory into a symbolic sequence. This is important because the complexity algorithms are mostly applied to abstract mathematical models that normally have analytical representation. Also, even though there is a rigorous mathematical formulation of the symbolization procedure, it is far from obvious how to apply it to a general continuous multidimensional trajectory.

Studying complexity is becoming an active field of research, nevertheless, its application to real physical systems is still quite scarce.

There is evidence of chaotic character in molecular motions. Various MD trajectories of proteins have been carefully analyzed in Refs. 6 and 7 and it is shown that they have positive Lyapunov exponents. The latter is a strong indication of the presence of chaos in the dynamics of the system. The existence of chaos can be rigorously verified for a simple three-atomic molecular model.[8] The existence of transversal homoclinic points in a Poincaré map of this system rigorously proves (via the theorem of Smale) its chaoticity. There is even the experimental indication of the existence of chaos in the dynamics of a Brownian particle.[9]

Lyapunov exponents and Kolmogorov–Sinai entropy are used in studying transport and reaction-rate coefficients in a classical fluid.[10] The connection of Kolmogorov–Sinai entropy with thermodynamic entropy is discussed in Refs. 11, 12. The informational aspect of Kolmogorov–Sinai entropy and, most interesting, its use as a measure of complexity is studied in Ref. 13.

Most theoretical work in connection with complexity uses ''symbolic dynamics.'' That is, the dynamics of a system is represented as a sequence of symbols from an alphabet of possible values. Besides formal mathematical investigations of symbolic dynamics there are applications to

 © 2002 American Institute of Physics

physical systems. Among them the studying of algorithmic complexity of solar spike events[14] and investigation of the important phenomena of temporal irreversibility in real physical phenomena.[15]

While the Kolmogorov–Chaitin algorithmic complexity is applied to physical systems there are few applications of computational mechanics and its measure of complexity. However it would be very advantageous to focus on this particular approach because it clearly provides more opportunities for understanding complex behavior. To the best of our knowledge there is no application of either algorithmic complexity or computational mechanics to the dynamics of real molecular systems. In this paper we would like to present the first attempt at investigating dynamic complexity of a classical molecular system.

The diversity of molecular properties arising from the dynamical structure of water to the role of water in the mechanism of protein folding is likely to be the result of the emergent, self-organizing character of these systems. Clearly, the extent of self-organization and consequently, the complexity is different for different systems. The power of analysis methods of the emergent behavior would, probably, be greatest in the most complex systems. However, even the simplest examples should shed a new light on these extensively studied systems.

As the first attempt in understanding complexity we have chosen to analyze the classical dynamics of water molecules in the Na^+–water system. We hypothesized that the dynamics of water molecules in the solvation shell of the ion and in the bulk would show significant differences in complexity. In addition, some other features, not immediately obvious from more traditional analysis could be revealed. The methodology of classical molecular dynamics of this type of system is rather well developed and an ample amount of literature sources are available for comparison.[16–22]

II. SIMULATION AND NUMERICAL MODEL

We collected the coordinates from the molecular dynamics simulation of a Na^+ ion in cubic box filled with water. The simulation was performed using the GROMACS 3.0 package (for details and documentation, see Ref. 23). The box included 216 molecules of SPC water[24] and periodic boundary conditions were applied. There was a 0.85 nm cutoff distance for both van der Waals and Coulomb potentials. When introducing the Na^+ ion in the middle of the cubic box (dimensions: 1.875 67 nm), energy minimization using a conjugate gradient algorithm was performed followed by a steepest descent algorithm for further minimization. The leapfrog algorithm for integrating equations of motion and the GROMOS-96 (Ref. 23) force field to implement the potential functions were used. The time step was 0.002 ps and we collected 11 000 time steps. The initial velocities were generated by a Maxwellian distribution at 300 K.

It is important to choose a relevant dynamic characteristic that reflects the key property of the system. We have analyzed the orientation of the water molecule with respect to the ion. The angle α is calculated as shown in Fig. 1, where $\mathbf{a}_1=\mathbf{r}_O-\mathbf{r}_{Na^+}$ and $\mathbf{a}_2=(\mathbf{r}_{H_1}-\mathbf{r}_O)+(\mathbf{r}_{H_2}-\mathbf{r}_O)$. Two

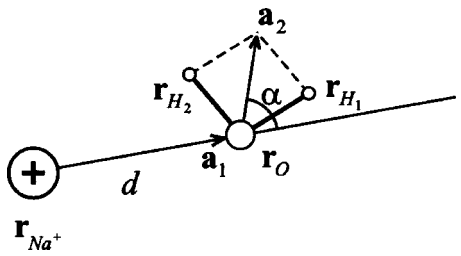

FIG. 1. The dynamical characteristic, $\cos\alpha$ and d, used to calculate complexity.

main characteristics were used as dynamic trajectories for calculating complexity: $\cos\alpha$ and the distance $d=|\mathbf{a}_1|$. In addition the raw three-dimensional coordinates of the oxygen atoms were also used to estimate the complexity of the atomic motion.

III. RESULTS AND DISCUSSION

There are a number of works devoted to MD simulation of the water–Na^+ system. The methods range from classical MD (Refs. 16–20) via mixed quantum-classical approaches[21] to *ab initio* molecular dynamics.[22] We will not focus on the structure and dynamics of the system here. It is worth noting that the particular event of a migrating water molecule in and out of the first solvation shell (ss) of the ion and the details of the dynamics is of great interest. This particular process is key to understanding chemical reactions in liquid water.

The trajectories analyzed are presented in Figs. 2 and 3. The 3D coordinates of the oxygen atoms were also used as dynamic trajectories. The $\cos\alpha$ and d trajectories are the characteristics of the water molecules as a whole while the space coordinate motion is a description at the single atom level. Nine representative water molecules were chosen and labeled by numbers from 1 to 9.

The distance time dependence was used to distinguish between the water molecules in the first solvation shell those from the bulk water. From Fig. 2 it is clear that molecules

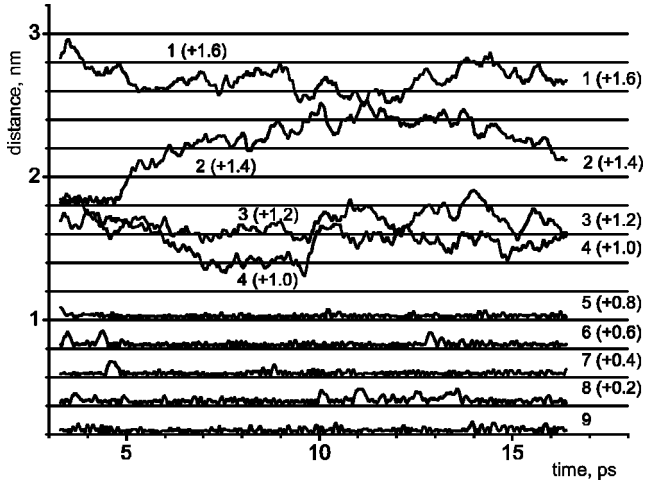

FIG. 2. The distances d of water molecules number 1–9 from the ion. The bulk molecule trajectories are marked with numbers corresponding to the molecule numbers (see text).

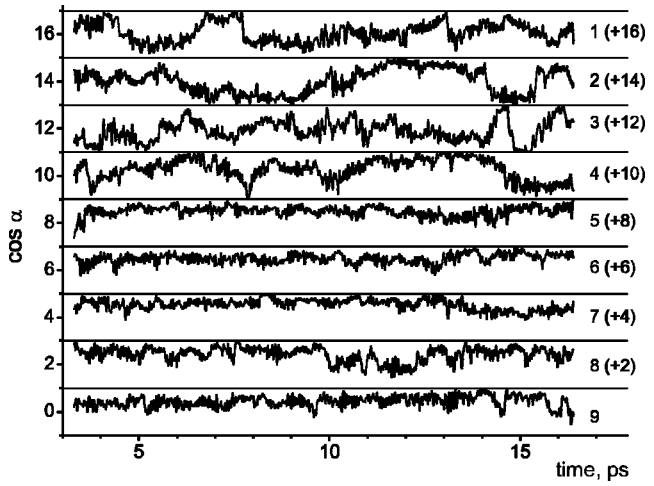

FIG. 3. cos α of the water molecules used in the analysis. Water molecule numbers are in the column on the right.

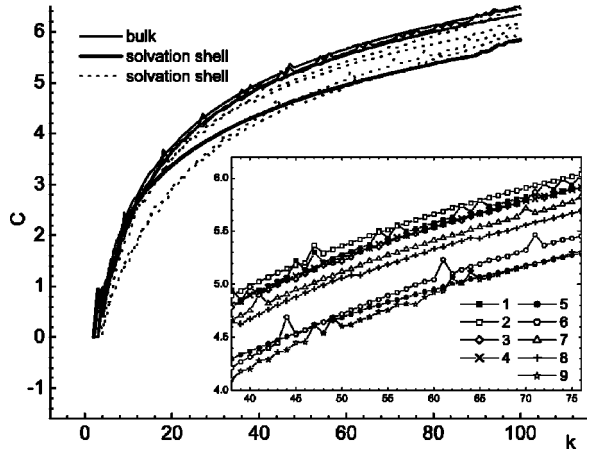

FIG. 4. The k dependence of the orientational dynamics complexity (C) of water molecules at various distances from the ion.

5–9 spend all their time in the ss region while the others, in the bulk. The molecules in the ss change their distance to the ion very insignificantly. Overall they behave very differently from molecules in the bulk, however they also differ between each other in small details. Molecules 6 and 8 exhibit noticeably bigger fluctuations.

In order to obtain consistent results in complexity its dependence on the algorithm's parameters must be analyzed. As discussed in Ref. 1 the parameters of our algorithm are l the length of left and right subsequences, and k the number of partitions used to symbolize original data. Also, the dependence on the number of data points, or time interval used to collect the trajectories, should be checked to insure that enough data has been collected.

The complexity of two typical water molecules are almost identical for all l in the range 2–10. For consistency we used $l=3$ for all our calculations.

Similar to the test cases[1] the k dependence of the complexity roughly follows the binary logarithm function. As was discussed, this dependence may come from two sources. First, the incomplete sampling of data may produce unnatural equivalence classes leading to an artificial increase in complexity. Second, the theoretical limit of the Shannon entropy at $k\rightarrow\infty$ has a logarithmic dependence. The former source would be indicated by the strong influence of l value and number of data points on the resulting curves. The latter, on the other hand, is an intrinsic property of the systems and would not be dependent on any parameters of the numerical algorithm, provided that we have enough data points.

The number of data points (or the overall time interval) does not influence the resulting values of complexity. The results are practically the same for all various number of data points down to an 8 ns interval.

The weak dependence of complexity on both the l value and the number of data points implies that the k dependence of the complexity comes primarily from the natural limiting value of Shannon entropy for continuous signals.

The complexities of the cos α and d trajectories for analyzed water molecules are plotted in Figs. 4 and 5.

Overall there is a well-defined difference in complexity for orientational and translational dynamics of water molecules in ss and away from the ion at all values of k. The molecules surrounded by the other water molecules (not in ss) show distinctively higher complexity than those molecules in the ss. Because the motion of the ion is slower than the water molecules, α can be considered for the distant water molecules as an orientational characteristic with respect to some point fixed in space. In other words, this characteristic serves as an estimate to the orientational dynamics of water molecules with respect to their immediate neighbors. Similarly, d can be used as a measure of the translational motion. Higher values of complexity imply that the dynamics of a water molecule exhibits richer behavior, processes more information and requires more informational memory when it is in a ''cage'' of other water molecules. On the contrary, when a molecule is constrained by the electric field of the ion, even though it undergoes as big local fluctuations as bulk molecules do, its motion is noticeably less complicated, there is less of a pattern in its time evolution.

Interestingly, different results are produced for different accuracies of symbolization. For one molecule (number 5,

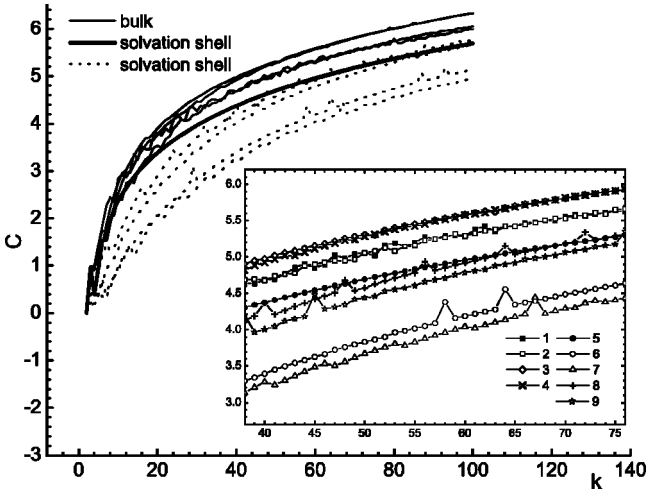

FIG. 5. The k dependence of the translation dynamics complexity (C) of water molecules at various distances from the ion.

TABLE I. Average $\overline{\cos\alpha}$ and distances $\overline{d}$ from the ion for various water molecules.

Molecule number	$\overline{\cos\alpha}$	$\overline{d}$
	Bulk	
1	0.07	0.78
2	0.00	1.14
3	−0.10	0.49
4	0.28	0.55
	Solvation shell	
5	0.49	0.23
6	0.49	0.23
7	0.54	0.23
8	0.42	0.24
9	0.43	0.23

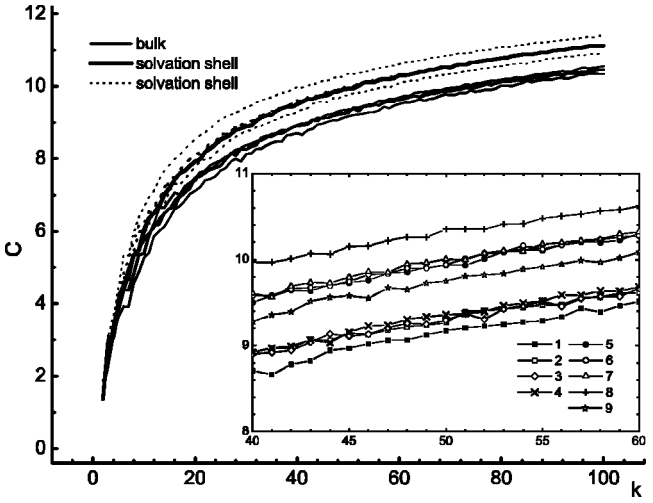

FIG. 6. The k dependence of dynamics complexity (C) for three-dimensional coordinates of oxygen in water molecules at various distances from the ion.

depicted by a bold solid line in Figs. 4 and 5) a high accuracy (high number of partitions k) is required to extract all information from the continuous trajectory. Therefore, the k -dependence of the complexity for this molecule shows a transition from higher to lower values with increasing k.

The complexity of the translational dynamics (Fig. 5) is more sensitive to the distance to the ion than the orientational one. However, for both characteristics the bulk molecules exhibit far less variations in the value of complexity than the molecules in ss.

To emphasize the new information that can be learned from this sort of complexity analysis, we present the averaged values of $\cos\alpha$ and distances from the ion (Table I). According to these data the dynamics of molecules numbered 5–9 is almost indistinguishable by both $\overline{\cos\alpha}$ and $\overline{d}$. The finite statistical complexity however, shows distinct differences in the dynamics. Interestingly, some details of the motion of molecules 7 and 8 make them follow more complex reorientations. As for the translations, the trajectories of molecules 5, 8, and 9 are more complex than the other two molecules in ss. These differences may be an indication of the start of the exchange event. We do not have enough evidence for a more definite conclusion yet and this transitional behavior is the subject of our current research.

In contrast to this, the average $\overline{\cos\alpha}$ of the bulk molecule differs from each other significantly. Nevertheless the complexity of all three characteristics of their dynamics (reorientation, translation, and space trajectories) is virtually the same.

The complexities calculated from the three-dimensional spatial coordinates of the oxygen atoms of the same molecules are presented in Fig. 6. Again, the distinctively different values of complexity for the atoms in ss and in the bulk are found. However, when comparing the complexity of reorientation between ss and bulk molecules we observe the opposite trend. The atoms from the ss molecules have higher complexity than those in the bulk. This emphasizes that the most significant features of the dynamics of bulk molecules come from their motion as a whole and not from the dynamics of individual atoms.

IV. CONCLUSIONS

It is shown that computational mechanics can be applied to the analysis of the dynamical complexity of molecular motion. It clearly demonstrates the sensitivity of the statistical complexity to the details of water molecule dynamics in the solvation shell of the ion and in the bulk. The reorientational and translational motion of the bulk molecules shows distinctively more complex character then those in the field of the ion. Complexity analysis gives more details of the motion of the molecules in the vicinity of the ion. It provides information about the differences of seemingly indistinguishable trajectories. The nature of these differences is currently under investigation.

In the future we plan to improve the algorithm which we hope will be even more sensitive to subtle differences in the dynamics. Also, studying the events of water exchange in the solvation shell may bring valuable insight into the understanding of emergent behavior. The same is true for bulk molecules. More developed algorithms may shed light on the details of water movements in the cage of surrounding water molecules.

The next step in our research will be the consideration of the trajectory in its phase space, i.e., the inclusion of the velocities into the analysis. Also, taking into account groups of molecules of various size will most probably give new valuable information. Currently the algorithm is being tested for high-dimensional problems (up to 20) and it demonstrates robustness. This also suggests that this methodology may uncover some of the mysteries of high-dimensional dynamics in the condensed phase.

Finally, it will be extremely promising to apply computational mechanics analysis to large molecular systems and in particular proteins. We believe that this novel approach will bring new insight into the process of protein folding and other complex self-organizing phenomena.

ACKNOWLEDGMENT

The work is supported by the Isaac Newton Trust and Unilever.

[1] D. Nerukh, G. Karvounis, and R. C. Glen, J. Chem. Phys. **117**, 9611 (2002), preceding paper.
[2] J. P. Crutchfield and K. Young, Phys. Rev. Lett. **63**, 105 (1989).
[3] J. P. Crutchfield and K. Young, in *Entropy, Complexity, and Physics of Information, SFI Studies in the Sciences of Complexity*, edited by W. Zurek (Addison–Wesley, Reading, 1990), Vol. VIII, pp. 223–269.
[4] J. P. Crutchfield, Physica D **75**, 11 (1994).
[5] N. Perry and P.-M. Binder, Phys. Rev. E **60**, 459 (1999).
[6] M. Braxenthaler, R. Unger, D. Auerbach, J. A. Given, and J. Moult, Proteins: Struct., Funct., Genet. **29**, 417 (1997).
[7] Huai-bei Zhou, J. Phys. Chem. **100**, 8101 (1996).
[8] T. Rage, A. Neumaier, and C. Schlier, Phys. Rev. E **50**, 2682 (1994).
[9] R. Gaspard, M. E. Briggs, M. K. Francis, J. V. Sengers, R. W. Gammon, J. R. Dorfman, and R. V. Calabrese, Nature (London) **394**, 865 (1998).
[10] J. R. Dorfman and P. Gaspard, Phys. Rev. E **51**, 28 (1995).
[11] V. Latora and M. Baranger, Phys. Rev. Lett. **82**, 520 (1999).
[12] M. Dzugutov, E. Aurell, and A. Vulpiani, Phys. Rev. Lett. **81**, 1762 (1998).
[13] P. Allegrini, J. F. Douglas, and S. C. Glotzer, Phys. Rev. E **60**, 5714 (1999).
[14] U. Schwarz, A. O. Benz, J. Kurths, and A. Witt, Astron. Astrophys. **227**, 215 (1993).
[15] C. S. Daw, C. E. A. Finney, and M. B. Kennel, Phys. Rev. E **62**, 1912 (2000).
[16] G. Toth, J. Chem. Phys. **105**, 5518 (1996).
[17] D. G. Bounds, Mol. Phys. **54**, 1335 (1985).
[18] S. Obst and H. Bradaczek, J. Phys. Chem. **100**, 15677 (1996).
[19] J. Chandrasekhar and W. Jorgensen, J. Chem. Phys. **77**, 5080 (1982).
[20] L. Degreve and F. da Silva, J. Chem. Phys. **110**, 3070 (1999).
[21] A. Tongraar, K. R. Liedl, and B. M. Rode, J. Phys. Chem. A **102**, 10340 (1998).
[22] Jody A. White, Eric Schwegler, Giulia Galli, and François Gygi, J. Chem. Phys. **113**, 4668 (2000).
[23] A. R. van Buuren, S. J. Marrink, and H. J. C. Berendsen, J. Phys. Chem. **97**, 9206 (1993); A. E. Mark, S. P. van Helden, P. E. Smith, L. H. M. Janssen, and W. F. van Gunsteren, J. Am. Chem. Soc. **116**, 6293 (1994); W. F. van Gunsteren and H. J. C. Berendsen, GROMOS-87 manual (Biomos BV Nijenborgh 4, 9747 AG Groningen, The Netherlands, 1987).
[24] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans, in *Intermolecular Forces*, edited by B. Pullman (Reidel, Dordrecht, 1981), pp. 331–342.