Topography and Dynamics of Multidimensional Interatomic Potential Surfaces

R. Stephen Berry
Department of Chemistry, The University of Chicago, Chicago, Illinois 60637

Ralph Breitengraser-Kunz*
Institut für Theoretische Physik, Technische Universität Berlin, Sekretariat PN 7-1, Hardenbergstrasse 36, D-10623 Berlin, Germany
(Received 21 November 1994)

A statistically based characterization of the topography of a multidimensional potential surface classifies not only local minima and saddles but entire basins containing many minima, and divides separating basins and monotonic sequences of local minima within each basin. The data fold readily into the formalisms of chemical kinetic isomerization theory and master equations to provide a connection between that topography and the dynamics on the surface. This analysis permits an interpretation of the glass-forming or "focusing" character of the surface. The method is illustrated with a model system derived from the 19-atom Lennard-Jones cluster.

PACS numbers: 36.40.–c, 82.20.Db, 82.20.Wt

Clusters of tens or hundreds of atoms, for the most part, have regular geometries in their low-energy forms, e.g., icosahedral if the interatomic forces are predominantly central and attractive with ranges of ordinary chemical bonds or van der Waals forces, and rocksaltlike if the forces are those of typical ionic salts. However, even these simple systems have potential surfaces in their ground electronic states that contain a great many local minima corresponding to amorphous structures. This occurs with rare gas clusters [1] and with alkali halides [2,3]. In the latter study, the $(KCl)_{32}$ cluster exhibited amorphous, locally stable structures outnumbering the rocksaltlike structures by roughly $10^{12}:1$. Nevertheless, a simulated cluster of $(KCl)_{32}$, initially liquid, whose energy is reduced at 5×10^{12} K/s or slower, finds its way to a rocksaltlike structure. Only if the energy is removed at a cooling rate of over 10^{13} K/s can such clusters fall into the potential wells of amorphous structures [3,4]. This apparent paradox is much akin to the Levinthal paradox in protein structure: the "wrong" structures so outnumber the "correct" or physiologically active structures of a protein that organisms could not exist if random search were the way proteins folded to their active structures. In fact crystal formation by random search is probably more unlikely, statistically, than protein folding, insofar as the constraints on the protein—maintenance of the integrity of the chemical bonds of the polymer chain—rule out many structures available to the freely moving atoms on their way to forming a solid.

This example illustrates one pressing issue in the general problem of understanding how the multidimensional potential surfaces of polyatomic molecules, clusters, and nanoscale particles govern the dynamics and phaselike behavior of these systems. It is now practical to explore completely a given potential surface for a system of 8 or perhaps 10 atoms, in the sense of finding all the local minima, all the important saddles linking those minima, and the topology, the connectivity, of all the minima [5]. However, the number of geometrically distinct minima grows approximately exponentially with the number of particles N in the system [6,7], and the number of permutational isomers of each of these increases roughly factorially with N. Hence the complexity of potential surfaces increases so rapidly with N that we are most unlikely to wish to create catalogues of all the minima and important saddles for clusters of more than about 15 atoms. For such systems we must use statistical samples, not complete data sets, and we need conceptual frames and questions appropriate to such statistical data. Statistical methods have been used before to study complex kinetics [8] but not with the emphasis on the relation between topography and kinetics that is invoked here.

This Letter presents a way to generate and manipulate such a data set for a many-dimensional potential surface, and to use it to analyze the dynamics on that surface. The steps of the approach are as follows: (1) development, from simulation and quenching, of a statistical sample of minima and the saddles that connect them; (2) organization of the data in the statistical sample through a hierarchy of steps described below, to provide a pattern of the topography of the surface; (3) estimation of the matrix of well-to-well transition rates and incorporation of those rates into a master equation which is then solved; (4) analysis of the eigenvectors of the master equation in terms of the pattern of the topography; and, finally, if desired, (5) application of the master equation to address controlled annealing of the system.

Our statistical database was generated by carrying out constant-energy molecular dynamics of a liquidlike 19-particle Lennard-Jones system, with regular quenching of the energy [9] to sample the potential wells visited by the system. The saddles were determined by a hybrid of the eigenvector-following method [10,11] and the method of slowest slides [12]. Our sample, so determined, consisted of 299 different minima and 461 saddles that linked a pair

0031-9007/95/74(20)/3951(4)$06.00

of minima in the set of 299. Logarithmic extrapolation of the data for Lennard-Jones clusters of 7 through 13 atoms [7] implies $\sim 0.5 \times 10^6$ geometrically distinct minima on this surface. Hence we must be cautious in imputing validity of our sample as a representation of the real 19-particle rare gas cluster.

The second phase is the organization and analysis of these data, beginning with listing and ordering the minima by energy. The saddles are then cataloged by three numbers, the energies of the two minima that each saddle links and the saddle energy. Next is the compiling and ordering of all the minimum-saddle-minimum (min-sad-min, for convenience) sets. We order these according to the energy of the lower minimum. These triples can be displayed schematically as shown in the inset of Fig. 1; this seemingly unintelligible tangle contains all the min-sad-min triples in our sample, plotted with the lower-energy minima on the left. (The 8 minima in the set of 299 that were not linked to others were dropped from the sample.) These data show that the barrier heights of this system *on average* increase monotonically with decreasing energy of the local minima [13]. Hence the potential surface of Ar_{19} is not strongly "focusing," although it is not necessarily strongly "antifocusing" either [14].

In the triples of the inset of Fig. 1, many of the energies of minima on the right side, the upper wells of the min-sad-min triples, appear also on the left, as the lower wells of other min-sad-min triples. With the energy resolution of our simulation, there are no degeneracies among the minima, apart from permutational isomers; each energy ϕ_i defines a unique, locally stable structure. Hence a minimum at an energy on the left can be taken as identical with a minimum at the same energy on the right. This, in turn, means we can unfold the triples into entire sequences of min-sad-min-sad-$\cdots$ points as shown in Fig. 1. In fact the sequences of Fig. 1 are only those that connect by

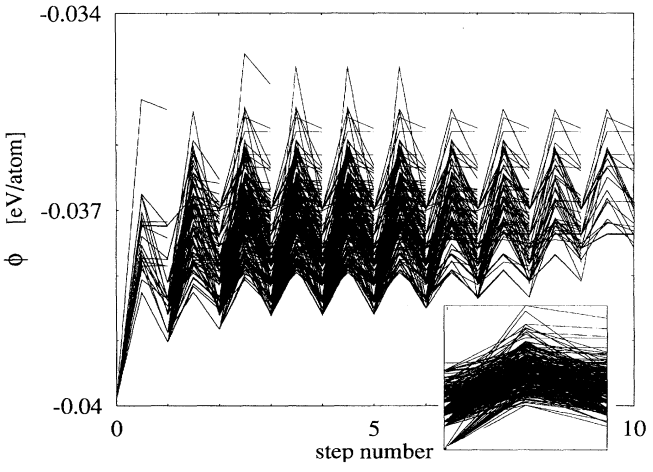

FIG. 1. Primary monotonic sequences in the sample database for the Lennard-Jones 19-particle cluster; inset: all minimum-saddle-minimum triples in the database.

a *monotonic* sequence of minima to the global minimum of the potential surface, which has the geometry of the double icosahedron. In our statistical sample, 245 of the $m = 291$ linked minima lie on some monotonic sequence shown in this figure. These can be thought of as being local minima in one large basin, which we call the *primary basin*. Note that the minima, not the saddles, define these sequences.

The saddle energies along the sequences of minima in general do not form monotonic sequences although the saddle heights generally increase along the downward direction of the sequences of minima [13]. This trend along monotonic sequences may be a more useful criterion for the antifocusing character of the surface than a correlation of barrier height with the energy of the upper well. Local minima not on the monotonic sequences leading to the global minimum do lie on other monotonic sequences. These 46 all lie on sequences separated from the primary basin by one saddle. We can immediately define the regions containing these as secondary (and, if they occurred, tertiary, etc.) basins. The saddles which separate basins, we call *divides*. A monotonic sequence of minima leading to the global minimum is then naturally called a primary monotonic sequence (PMS). Likewise a monotonic sequence lying in a secondary basin is a secondary monotonic sequence or SMS, a monotonic sequence in a tertiary basin is a tertiary monotonic sequence, etc. We call a divide separating a primary basin from a secondary basin a primary divide. Figure 2 shows a schematic representation of this way to describe a multidimensional potential surface.

We have just used the expression "a primary basin" rather than "the primary basin" deliberately. The reason is not apparent in the example of the Ar_{19} cluster that is our principle vehicle here, but it is in the case of the $(KCl)_{32}$ cluster [2,3]. In that example, it proved useful to distinguish not only the global minimum but all the locally stable forms that have rocksalt structures, even with a few defects. These constitute a class of low-energy forms that define a set of basins that can be classified together as primary. This generalization from a single primary basin for the global minimum to a set of primary basins becomes more and more relevant as the clusters

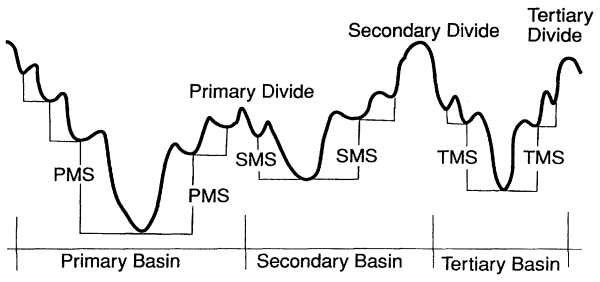

FIG. 2. Schematic representation of primary, secondary and tertiary basins, and monotonic sequences and divides.

become larger, where interest is in whether the system has found some fcc, hcp, icosahedral, or other type of structure, rather than in whether it has reached the global minimum. Even in the protein folding case [15,16], it is useful for purposes of classification to allow systems that have reached a state with the "right" physiological activity to have structural differences in irrelevant parts of the system. All such structures can be said to lie in primary basins.

This categorization of the surface into basins and wells within them allows us to study dynamics on the surface without treating explicitly the rapid intrawell vibrations. Instead, by using transition state theory in any of its forms, we in effect average over these high-frequency modes and concentrate on the interwell and interbasin transitions. To implement this at the present stage, we have used the RRKM method to estimate interwell transition probabilities, a method that gives reliable rates for isomerization of clusters [17]. It is now quite feasible to determine the eigenvalues of the Hessian matrix at the minimum of each well, either numerically or analytically [11,18], to use those eigenvalues to carry out the RRKM rate calculations in the harmonic approximation [17], and even to include temperature-dependent anharmonic corrections [2,19].

Our matrix $W_{jj'}(T)$ of transition probabilities, per unit time, at temperature T, from well j' to well j has the form

$$W_{jj'} = \sum_{l_{jj'}} \frac{k_B T}{h} \frac{Q_{jj'}^{l_{jj'}\neq}}{Q_{j'}} \exp(-\Delta\phi_{jj'}^{l_{jj'}\neq}/k_B T), \qquad (1)$$

where k_B and h are, respectively, Boltzmann's and Planck's constants, $\Delta\phi_{jj'}^{l_{jj'}\neq}$ is the saddle height the system must attain to pass from well j' to well j via the saddle denoted $l_{jj'}$, and $Q_{j'}$ and $Q_{jj'}^{l_{jj'}\neq}$ are, respectively, the partition functions of the system in well j' and at the saddle $l_{jj'}$. The fundamental assumption of transition state theory is the persistence of the system in the region of that saddle long enough for $Q_{jj'}^{l_{jj'}\neq}$ to be well defined. If a harmonic approximation is made for the vibrations (with n degrees of freedom), then $Q_{jj'}^{l_{jj'}\neq}/Q_{j'}$ becomes $h\prod_{i=1}^{n}\nu_{j',i}/k_B T\prod_{i=1}^{n-1}\nu_{jj',i}^{l_{jj'}\neq}$, in terms of $\nu_{j',i}$ and $\nu_{jj',i}^{l_{jj'}\neq}$, the vibrational frequencies or eigenvalues of the Hessian matrices of the j'th well and the corresponding saddle $l_{jj'}$.

With the matrix $\underline{\underline{W}}$ in hand, we can readily construct a master equation for the time evolution of the probability distribution $\underline{P}(t) \equiv (P_1(t), P_2(t), \ldots, P_m(t))^T$, in which $P_j(t)$ is the probability that the system be in well j at time t. The equation is

$$\frac{d}{dt}P_j = \sum_{j'=1}^{m} w_{jj'}P_{j'}, \qquad (2)$$

where we have let $w_{jj'} \equiv W_{jj'} - \delta_{jj'}\sum_{j''} W_{j''j'}$ to simplify the notation. The time evolution of $\underline{P}(t)$, based on some initial condition $\underline{P}(0)$, is the flow of the probability distribution from its initial set of values toward the fully relaxed equilibrium state at that temperature. We assume that the entire phase space of our system is accessible at the high energy of the molecular dynamics simulation, where the cluster is clearly liquidlike, i.e., the energy shell there is compact. This implies that the matrix $\underline{\underline{w}}$ is neither decomposable nor splitting, i.e., that it is not, nor can it be put into block form. In this case there is only one equilibrium state, the Boltzmann distribution $\underline{P}^{\rm eq}$. As a consequence of detailed balance, the master equation (2) can be symmetrized by introducing a new dependent variable $u_j = P_j/\sqrt{P_j^{\rm eq}}$ [20]. The solution of (2) can then be expanded in a complete set of eigenvectors $\underline{\tilde{u}}^k$ to the symmetric matrix $\tilde{w}_{jj'} = \sqrt{P_{j'}^{\rm eq}/P_j^{\rm eq}}\, w_{jj'}$:

$$P_j(t) = \sqrt{P_j^{\rm eq}} \sum_{k,k'=1}^{m} \tilde{u}_j^k e^{\lambda_k t}\tilde{u}_{k'}^k \frac{P_{k'}(0)}{\sqrt{P_{k'}^{\rm eq}}}. \qquad (3)$$

The largest eigenvalue of $\underline{\underline{\tilde{w}}}$ is $\lambda_1 = 0$ corresponding to the equilibrium solution $\tilde{u}_j^1 = \sqrt{P_j^{\rm eq}}$, the other eigenvalues λ_k are negative.

For a computational illustration of the development presented here, we chose to use an Einstein model for the wells, and assume that they all have the same, single-frequency spectra of intrawell vibrations. For this test case we constructed the matrix $\underline{\underline{w}}$ and found eigenvalues and eigenvectors of the master equation for our 291-well potential surface. The individual wells are identified according to basin and could also be categorized by distance, in a Hamming [21] sense, for example, or in the related sense used by Bryngelson and Wolynes [15], from the minimum point of its basin. Hence the results of this evaluation, apart from yielding the zero eigenvalue and Boltzmann-distribution eigenvector, show which eigenvectors correspond to intrabasin flows and which to interbasin flows. Moreover, the eigenvalues are essentially the inverses of the time constants for these flows. An eigenvector with nonzero components in both primary and secondary basins is associated with flow between these basins. To see whether the *net* flow for that eigenvector is nonzero and into or out of the primary basin, we sum the components of that eigenvector corresponding to wells in the primary basin only. The magnitude of this *net flow index* determines how important that flow is, and the sign indicates whether flow is into (negative) or out of (positive) the primary basin.

The dark areas in Fig. 3 (right scale) show the distributions of (nonzero) eigenvalues of the master equation, normalized to the Einstein frequency ν_0, for the lowest and highest temperatures we examined. The entire set of eigenvalues is distributed over 7 orders of magnitude at 30 K, over 5 orders at 60 K, over 4 at 120 K, and 3 at 300 K. This figure also shows the eigenvalues of the eigenvectors dominating the interbasin flows (left scale), corresponding to the 30 largest negative (lightly shaded areas) and positive (solid lines) net flow indices, i.e., flows into and out of the primary basin. These are approximately 100-fold slower than the average of all flows at 30 K, about 10-fold slower at 60 K, less than a power of 10 at 120 K,

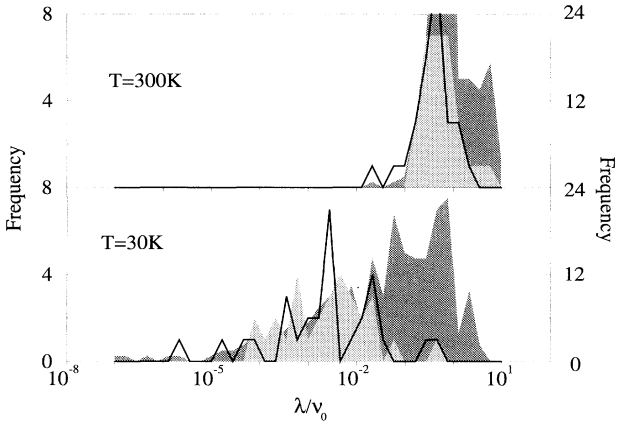

FIG. 3. Binned distributions of normalized eigenvalues λ/ν_0 of the master equation for 30 and 300 K. Areas shaded dark: all nonzero eigenvalues (right scale); light areas and solid lines: 30 eigenvalues with the largest net flow into and out of the primary basin, respectively (left scale).

and indistinguishable at 300 K [22]. The present analysis provides more precise information on the relation between topography and dynamics than did, for example, previous studies of Ar_{19} clusters, aimed at finding low-lying minima [23] and at quenching and phase changes [14]. This method shows that the physical basis for solid-liquid coexistence in clusters is the time scale separation between interbasin and intrabasin transitions (cf. Fig. 3, $T = 30$ K) and the not-strongly-focusing character of the surface.

We have taken this analysis one step further, to use the master equation to carry out two alternative optimized cooling strategies to bring the system to a preselected mean energy, and thereby to or near a preselected distribution of structures—including, for example, to bring it with high probability to the global minimum, or to a glass [13].

We have presented a procedure for studying topography and dynamics on multidimensional potential surfaces. It begins with a descriptive categorization of regions of a given surface, based on the analysis of a statistical sample of saddles and minima on that surface. The special significance of this method is its practicability, as a result of the efficient methods now available for locating minima and saddles on multidimensional potentials. The method then goes on to use transition state theory, at whatever level of accuracy is desired, to compute a matrix of well-to-well transition probabilities, and from this matrix, constructs a master equation whose solutions reveal the intrabasin and interbasin dynamics of the system moving on this surface.

Work in progress is exploring tests of the reliability of the statistical sample, and of the influence of the density of saddles, the distribution of saddle heights along monotonic sequences of minima and the distribution of monotonic sequences among the basins on the glass-forming or focusing character of the surface.

R. S. B. acknowledges the hospitality of Professor H. Baumgärtel and the Freie Universität Berlin, and the Alexander von Humboldt-Stiftung for its support. R. B. K. thanks Professor E. Schöll for the opportunity to pursue this work. We would like to acknowledge Tanya Astakhova's [24] construction of the statistical database. Support for the study was provided in part by the National Science Foundation.

*Name changed from Ralph E. Kunz.

[1] A. Rahman, M. J. Mandell, and J. P. McTague, J. Chem. Phys. **64**, 1564 (1976).

[2] J. P. Rose and R. S. Berry, J. Chem. Phys. **98**, 3246 (1993).

[3] J. P. Rose and R. S. Berry, J. Chem. Phys. **98**, 3262 (1993).

[4] H.-P. Cheng and U. Landman, Science **260**, 1304 (1993).

[5] R. S. Berry, Chem. Rev. **93**, 2379 (1993).

[6] M. R. Hoare, Adv. Chem. Phys. **40**, 49 (1979); P. A. Braier, R. S. Berry, and D. J. Wales, J. Chem. Phys. **93**, 8745 (1990).

[7] C. J. Tsai and K. D. Jordan, J. Phys. Chem. **97**, 11 227 (1993).

[8] J. C. Light, J. Chem. Phys. **40**, 3221 (1988); M. G. Sceats, Adv. Chem. Phys. **70**, Vol. 2, 357 (1988); A. Nitzan, *ibid.* 489 (1988).

[9] F. H. Stillinger and T. A. Weber, Kinam 3 **Serie A**, 159 (1981); F. H. Stillinger and T. A. Weber, Phys. Rev. A **25**, 978 (1982).

[10] J. Pancik, Coll. Czech. Chem. Comm. **40**, 1112 (1975); R. L. Hilderbrandt, J. Comput. Chem. **1**, 179 (1977); C. J. Cerjan and W. H. Miller, J. Chem. Phys. **75**, 2800 (1981).

[11] D. J. Wales, J. Chem. Phys. **91**, 7002 (1989).

[12] R. S. Berry, H. L. Davis, and T. L. Beck, Chem. Phys. Lett. **147**, 13 (1988); H. L. Davis, D. J. Wales, and R. S. Berry, J. Chem. Phys. **92**, 4473 (1990).

[13] R. Breitengraser-Kunz and R. S. Berry, J. Chem. Phys. (to be published).

[14] T. L. Beck and R. S. Berry, J. Chem. Phys. **88**, 3910 (1988); T. L. Beck, Ph.D. thesis, The University of Chicago, 1987.

[15] J. D. Bryngelson and P. G. Wolynes, Proc. Nat. Acad. Sci. U.S.A. **84**, 7524 (1987); J. D. Bryngelson and P. G. Wolynes, J. Phys. Chem. **93**, 6902 (1989); J. D. Bryngelson and P. G. Wolynes, Biopolymers **30**, 177 (1990).

[16] H. S. Chan and K. A. Dill, J. Chem. Phys. **100**, 9238 (1994); N. D. Socci and J. N. Onuchic, J. Chem. Phys. **101**, 1519 (1994).

[17] J. P. Rose and R. S. Berry, J. Chem. Phys. **96**, 517 (1992).

[18] D. J. Wales, J. Chem. Soc. Faraday Trans. **86**, 3505 (1990); D. J. Wales and R. S. Berry, J. Chem. Phys. **92**, 4283 (1990).

[19] F. H. Stillinger and T. A. Weber, J. Chem. Phys. **81**, 5095 (1984).

[20] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1981).

[21] R. Hamming, *Coding and Information Theory* (Prentice Hall, Englewood Cliffs, 1980).

[22] Note that these temperatures are simply characteristics of our Einstein model, and not of the real Ar_{19} cluster.

[23] M. R. Hoare and P. Pal, Adv. Phys. **20**, 161 (1971).

[24] Present address: Institute of Chemical Physics, Moscow, 117334, Russia.