What is special about how roaming chemical reactions traverse their potential surfaces? Differences in geodesic paths between roaming and non-roaming events

D. Vale Cofer-Shabica and Richard M. Stratt

Department of Chemistry, Brown University, Providence, Rhode Island 02912, USA

(Received 15 March 2017; accepted 17 May 2017; published online 6 June 2017)

With the notable exception of some illustrative two-degree-of-freedom models whose surprising classical dynamics has been worked out in detail, theories of roaming have largely bypassed the issue of when and why the counterintuitive phenomenon of roaming occurs. We propose that a useful way to begin to address these issues is to look for the *geodesic* (most efficient) pathways through the potential surfaces of candidate systems. Although roaming manifests itself in an unusual behavior at asymptotic geometries, we found in the case of formaldehyde dissociation that it was the pathways traversing the parts of the potential surface corresponding to highly vibrationally excited reactants that were the most revealing. An examination of the geodesics for roaming pathways in this region finds that they are much less tightly defined than the geodesics in that same region that lead directly to dissociation (whether into closed-shell products or into radical products). Thus, the broader set of options available to the roaming channel gives it an entropic advantage over more conventional reaction channels. These observations suggest that what leads to roaming in other systems may be less the presence of a localized “roaming transition state,” than the existence of an entire region of the potential surface conducive to multiple equivalent pathways. *Published by AIP Publishing.* [<http://dx.doi.org/10.1063/1.4984617>]

I. INTRODUCTION

Newton’s first law of motion evidently carries a substantial amount of weight in creating our intuitions about how atoms ought to behave in chemical reactions. It was certainly surprising when it was discovered that instead of proceeding in a manifestly economical and Newtonian fashion along an available barrier-free path towards separated radicals, photodissociated atoms sometimes seemed to reverse course, abstract an extra atom from their fellow radical, and head off towards molecular products.^{1–4} When it was subsequently noticed that this *roaming* mechanism was not only not limited to the formaldehyde example in which it was first observed,^{5,6} but was actually a fairly common occurrence,^{4,7–10} the desire to develop a theoretical understanding of the phenomenon grew still deeper.^{9–18} This paper is yet another attempt at doing so.

The first key step in this direction was the discovery that one could see roaming events (and correctly predict the unusual distributions of product internal states) by simply running purely classical trajectories on a single, electronic-ground-state potential surface.^{6,19,20} Neither quantum dynamics²¹ nor the multi-electronic-surface character of actual photodissociation experiments,^{22–24} though undoubtedly present, were in any way essential to the process. With that realization, much of the theoretical effort turned to trying to develop predictive models for the branching ratios between the three basic reaction channels: direct formation of molecular products, molecular products derived from roaming, and radical products. Many (though not all) roaming systems seemed to have distinguishable direct and roaming saddle points in their potential surface,^{10,24,25} suggesting that one could get at these ratios via a multiple-saddle-point-based transition-state theory.

But the relevant parts of the potential surface were invariably too flat, and the energetic separations between the different routes far too small, to make such a naïve approach viable.^{12,26}

Much more promising, it seemed, was to regard roaming as a kind of dynamically frustrated radical dissociation. One could, for example, ignore the question of *how* this frustration occurs and concentrate on the relative equilibrium preferences for the roaming-derived-molecular and radical products, the strategy embodied in the phase-space theory of Andrews *et al.*⁹ This approach does quite well with the branching ratio it attempts to calculate. Alternatively, one could formulate a multi-path transition state theory, but choose the dividing surfaces so as to represent the dynamical effects rather than just mirror the potential energy surface features. The statistical theory proposed by Klippenstein and co-workers falls into this category.¹⁰ It too achieves some quantitative success, though it does not seem to work once the energy exceeds the radical dissociation threshold.

Our goals in this paper are rather different. We are less interested in what the various branching ratios are than what distinguishes the unusual dynamics encountered in roaming. In that sense, our approach has more in common with the recent efforts by the Bristol-Cornell-Crete collaboration.^{13–18} Those authors’ compelling vision of roaming trajectories as being trapped between certain phase-space structures (the dividing surfaces bounded by normally hyperbolic invariant manifolds)^{13,15} and as being guided by still other phase-space structures (“sticky” resonant periodic orbits)¹⁶ may be getting at much of the “why” of roaming. On the other hand, the analysis associated with this approach tends to be highly specific to the details of each new problem and is limited in practice

(though not in principle)¹³ to low-dimensional (indeed, usually 2-degree-of-freedom) caricatures of the true dynamics.

The approach described in this paper starts by agreeing fully that the presence and location of critical points on the system's potential surface (first or second order saddles, for example) should not be the primary focus of a theory of roaming, and that understanding roaming does require an intrinsically dynamical perspective.^{4,18} However, we suggest that, contrary to some comments in the literature, an emphasis on dynamics does not necessarily imply thinking only about phase space (about positions and momenta together).^{13–18} We show here that one can still ask profitable questions posed in configuration space (questions about where atoms are on the potential surface), provided we are asking about special *paths* rather than special points.

Our starting assumption is that roaming is a surprisingly commonplace, classical-mechanical phenomenon defined mainly by energies that are large on the scale of the potential surface. Moreover, while it certainly seems useful to see what low-dimensional models can teach us, we want a theory that allows for the possibility that roaming may be even more prevalent in situations with many degrees of freedom. To put it another way, the high-energy character of roaming may mean that it has a significant entropic component: the difficulty the system faces in forming products may be more of finding a narrow window in a multi-dimensional space than of accumulating enough energy to surmount a local high spot on the potential. The greater role of roaming in acetaldehyde than in formaldehyde, for example,⁹ might arise as much from the higher dimensionality of the former as it does from the small differences observed between the two potential energy surfaces.

In addition to developing a theory that prioritizes entropic, as opposed to energetic concerns, we note that the high energies associated with roaming might free us to take the configuration-space perspective we alluded to. At much lower energies, molecules oscillate about well-defined structures, often making it natural to describe their dynamics in terms of such collective, non-Cartesian coordinates as bond-bending and dihedral angles. That kind of description automatically places a significant portion of their nonlinearities into their kinetic energies²⁷—suggesting the usefulness of a phase-space perspective. However, at the kinds of energies relevant to roaming, there are no longer privileged molecular structures, so one might expect specialized collective coordinate combinations to have less of a role. The Cartesian coordinates of each atom might, in fact, provide the best “basis,” eliminating the kinetic-coupling justification for a phase space analysis.

The suggestion then is that it is worth examining what is special about the prototypical coordinate-space pathways taken on (what may be) high-dimensional potential surfaces of systems that roam. A potential-landscape approach that has had some recent success in looking at such *inherent* paths in the context of slow liquid-state dynamics^{28–33} offers a novel possibility for understanding some features of roaming. Instead of looking for transition-state surfaces through which some classical-trajectory flux is a maximum, this approach looks instead for paths through the potential landscape that are, in

some sense, the most efficient. Within this framework, the question we need to ask is then simply what is different about these most efficient paths when the system roams?

There are, in fact, a number of different ways of formulating the concept of an optimum path through a landscape.³⁴ We focus in this paper on a particular one, the *geodesic-path* approach.²⁹ We review this method in Sec. II and discuss why it might be especially appropriate to apply in the context of small-molecule roaming. But two examples of recent applications to liquid dynamics might help illustrate the possibilities offered by the method: In one instance, that of a mixture of infinitely hard spheres of two different sizes,³² the question was how to understand the sharp drop in the diffusion constant with increasing density. There was clearly no benefit to looking to create a transition-state theory of diffusion centered on the passage over potential saddle points—the entire potential energy was horribly non-analytic; it had no stationary points to find. Nonetheless, it still proved possible to find geodesic (minimum-length) paths for the entire N-body system and to show that the growing lengths of those shortest paths with increasing density quantitatively predicted the decline in diffusion constants. By analogy, one would expect that geodesics navigating a conical intersection on a potential surface exhibiting roaming^{12,22} would have no more difficulty in coping with that intersection than they would with a second order saddle point.

In a second example, we were concerned with why preferential solvation dynamics (the potential-energy relaxation caused by the rearrangement of a mixture of two kinds of solvent molecules around a solute) was so much slower than the comparable potential-energy relaxation associated with either pure species.³⁰ The issue there was what the change in *mechanism* was between the pure and mixed solvents. By regarding the respective N-body geodesic paths as representative of the inherent dynamics—the classical trajectories stripped of vibrational noise—we were able to discover that mixed solvents had to undergo an extra step to relax: exchanging slower species in the solute's first solvation shell for faster species. Since even the geodesics, the most efficient paths exhibiting that process, could only exchange a single solvent at a time, it soon became clear why these exchanges were the culprits. In much the same way, we might anticipate learning whether specific degrees of freedom have different roles in direct dissociation reactions than they do in the much slower roaming processes.

The rest of this paper will be organized as follows: Section II will review our geodesic formalism, emphasizing the benefits and limitations of that formalism in helping analyze a non-many-body problem. Section III will then describe our models and methods for applying this development to formaldehyde dissociation. We present our numerical results in Sec. IV and conclude in Sec. V with some general comments.

II. GEODESIC PATHS AND HIGH-ENERGY INTRAMOLECULAR DYNAMICS

A. The differences between high- and low-energy perspectives on potential energy landscapes

While the language of potential-energy landscapes has become a standard tool in interpreting (classical limit) reaction

dynamics,²⁴ it is worth realizing that most applications of landscape perspectives tacitly assume that the relevant energies are very low on the scale of the potential-surface topography. It is usually taken as given that reactants and products have well-defined structures prescribed by potential minima, and that understanding the passage between them is largely a matter of locating saddle points (the structures that define the highest points on the lowest energy path). When systems are at relatively high energies, though, as in photodissociation reactions, the centrality of such stationary points is not nearly so obvious.

Perhaps an analogy with liquid dynamics would not be out of place. The typical translational and rotational energies available to molecules in liquids are only on the order of $k_B T$ per degree of freedom (200 cm^{-1} at room temperature), but as low as that seems, it is far above the well depths of the vast majority of the minima on intermolecular potential surfaces. One could argue that for ordinary liquids, at least, the bulk of the stationary points are irrelevant to understanding dynamics.^{35,36}

That does not imply, however, that the geometry of the potential energy landscape is unimportant to these systems. Long distance molecular motion in liquids is invariably diffusive, and since diffusion can be described by path integrals,^{37,38} it is possible to quantify the relative weights of competing pathways through a given potential landscape, and even be precise in specifying what the most efficient molecular pathways are. As we noted in Sec. I, the phrase “most efficient” is not uniquely defined, but for liquids, a useful definition arises from considering only those sequences of many-molecule configurations $\mathbf{R} = (\mathbf{r}_1, \dots, \mathbf{r}_N)$ that never allow the potential energy of the system $V(\mathbf{R})$ to exceed a prescribed value—the *landscape energy* E_L . Within this so-called “landscape-ensemble” perspective,²⁸ diffusion can be regarded as completely force-free in the allowed regions; the variations in the potential surface within those regions have no effect on the dynamics. The crucial role of the potential surface is not, of course, neglected. What the formalism does is simply transform its influence into a set of intricate, and possibly quite stringent, boundary conditions limiting where the permissible paths can wander.^{28,29}

The starting point for this development is that the landscape-ensemble probability density for diffusing from one molecular arrangement $\mathbf{R}_0$ to another $\mathbf{R}_t$ in a time t is given by a path integral over all the paths $\mathbf{R}(\tau)$ connecting those $\mathbf{R}$ points ($0 \leq \tau \leq t$) that never violate the $V[\mathbf{R}(\tau)] \leq E_L$ condition,

$$G(\mathbf{R}_0 \rightarrow \mathbf{R}_t; t) = \int_{\text{allowed paths}} \mathcal{D}\mathbf{R}(\tau) \exp\left(-\frac{1}{4D} S[\mathbf{R}(\tau)]\right), \quad (2.1)$$

$$S[\mathbf{R}(\tau)] = \int_0^t d\tau T[\dot{\mathbf{R}}(\tau), \mathbf{R}(\tau)]. \quad (2.2)$$

Here D is the diffusion constant and T the kinetic energy. Given these expressions, it is clear that when the dynamics is slow (D is small), the most efficient paths are just those lying within the allowed portions of the potential surface that yield the smallest value of the force-free action S .²⁹

But those same paths also minimize the total kinematic path length³⁹

$$\ell_{\text{total}} = \int_0^t d\tau \sqrt{2T[\dot{\mathbf{R}}(\tau), \mathbf{R}(\tau)]}, \quad (2.3)$$

meaning that they are also the *geodesic* (shortest energy-allowed) paths through the landscape.²⁹ Hence the most efficient paths within this formalism can be regarded as the Fermat’s-principle (least-time) paths. Moreover, since the kinetic energy has components associated with individual physical degrees of freedom such as vibration and rotation, we can define component lengths, as well, allowing us to ascertain how much these most efficient dynamical routes rely on each of the different kinds of degrees of freedom.^{31,33}

All of the published applications of these ideas, however, have been within the context of condensed matter problems, leaving us with the question of how relevant these particular most-efficient paths through the potential energy landscape are to high-energy intramolecular dynamics. While both kinds of applications have energies well above much of their potential surfaces’ topography, the much higher dimensionality of many-body problems provides an additional measure of justification for using the landscape ensemble. In the thermodynamic limit, the landscape ensemble is just one more in a list of completely equivalent statistical mechanical ensembles; fixing the landscape energies in liquids is completely equivalent to choosing the canonical temperatures or the microcanonical energies.^{28,40} Fewer degree-of-freedom problems have more approximate equivalences between ensembles.

Nonetheless, we suggest that especially in near-dissociation-threshold problems featuring *roaming*, this same landscape-ensemble perspective and these same geodesic paths might very well describe the essence of the few-body dynamics. For one thing, there is often a diffusive character to energy transfer processes and to passage through bottleneck regions of potential surfaces.^{41–43} More to the point of the present work, the *sine qua non* of a roaming behavior is the presence of a special potential surface feature—the existence of an energetically accessible, flat, but non-asymptotic, region. However, the simple existence of such a region may not be all we need to think about in order to understand the prevalence of roaming in acetaldehyde and its relative scarcity in formaldehyde.⁹ There may be strict geometric requirements for entering and leaving that particular portion of the multi-dimensional configuration space—narrow channels and sharp turns, for example. What geodesics excel at is in determining the consequences for the dynamics of such complex geometric requirements.

B. Finding intramolecular geodesic paths

In the case of formaldehyde, the quantity specifying the instantaneous locations of all 4 atoms is the laboratory-frame 12-dimensional configurational vector $\mathbf{R} = (\mathbf{r}_C, \mathbf{r}_O, \mathbf{r}_{H1}, \mathbf{r}_{H2})$.⁴⁴ Thus the geodesics we are looking for are the shortest paths $(\mathbf{r}_C(\tau), \mathbf{r}_O(\tau), \mathbf{r}_{H1}(\tau), \mathbf{r}_{H2}(\tau))$, ($0 \leq \tau \leq t$), that (1) start and end at prescribed configurations, (2) avoid all of the energy-forbidden regions ($V(\mathbf{R}) > E_L$), and (3) obey whatever other constraints are necessary to limit ourselves to the experimentally relevant pathways.

We will discuss in Sec. III how we choose initial and final configurations, but we should mention how it is that one can find the shortest paths connecting them. The basic mathematical problem of locating paths which minimize a functional such as Eq. (2.3) while avoiding the obstacles posed by forbidden regions is a classic example of an inequality-constrained optimization. Solutions to such problems must obey the Kuhn-Tucker theorem,⁴⁵ which requires that the path be a combination of two kinds of segments: unconstrained portions between the obstacles that unconditionally minimize the functional and portions constrained to lie along the boundary of the forbidden regions.

For an N -atom intramolecular dynamics problem, the total kinetic energy can always be written in laboratory-frame Cartesian coordinates as

$$T_{total} = \frac{1}{2} \mathbf{\dot{R}}^T \cdot \mathbf{m} \cdot \mathbf{\dot{R}}, \quad (\mathbf{m})_{j\mu,k\nu} = (m_j/M) \delta_{jk} \delta_{\mu\nu}, \quad M = \sum_{j=1}^N m_j, \quad (2.4)$$

with j, k labeling the atoms, μ, ν Cartesian coordinates (x, y, z), m_j the mass of the j th atom, and M the total molecular mass. The *between-obstacle paths* that minimize Eq. (2.3) for a kinetic energy of this form are then simply the straight-line paths²⁹

$$\mathbf{R}(\tau) = \mathbf{R}_0 \left(1 - \frac{\tau}{t}\right) + \mathbf{R}_t \left(\frac{\tau}{t}\right), \quad (0 \leq \tau \leq t). \quad (2.5)$$

Part of what makes geodesic paths suitable as exemplars of the full classical dynamics of a given problem is that we can always define revised geodesics so that, in addition to minimizing the kinematic length, they satisfy selected constraints obeyed by the full dynamics. In the case of the paths of Eq. (2.5), though, these geodesics already mirror a number of properties of the classical trajectories between the same endpoints. If the beginning and ending configurations of such paths are selected from classical trajectories obeying a conservation law, for example, then all the intermediate points along this geodesic obey that same conservation law. In particular, if the two termini have the same center-of-mass location, then this geodesic will automatically preserve that same center of mass.

A simple way to verify this fact relies on defining a $3 \times 3N$ mass matrix $\mathbf{M}$,

$$(\mathbf{M})_{\mu,k\nu} = (m_k/M) \delta_{\mu\nu}, \quad (k = 1, \dots, N), \quad (\mu, \nu = x, y, z). \quad (2.6)$$

Given any $3N$ -dimensional configuration vector $\mathbf{R}$, the corresponding 3-dimensional center of mass vector $\mathbf{r}_{COM}$ can be extracted by taking the product with this $\mathbf{M}$ matrix,

$$\mathbf{r}_{COM} = \sum_{k=1}^N (m_k/M) \mathbf{r}_k = \mathbf{M} \cdot \mathbf{R}. \quad (2.7)$$

So, if $\mathbf{r}_{COM}(0) = \mathbf{M} \cdot \mathbf{R}_0 = \mathbf{r}_{COM}(t) = \mathbf{M} \cdot \mathbf{R}_t$, then Eq. (2.5) implies

$$\mathbf{r}_{COM}(\tau) = \mathbf{M} \cdot \mathbf{R}(\tau) = \mathbf{r}_{COM}(0), \quad (0 \leq \tau \leq t). \quad (2.8)$$

The sections of the geodesic *along the obstacle boundaries* are not governed by any evolution equations as simple

as Eq. (2.5), but we still need these sections to preserve the system's center of mass. For example, if one ever had two different paths along the $V(\mathbf{R}) = E_L$ boundary identical except that one translated the center of mass and one did not, only the latter could be a geodesic: the former is guaranteed to be longer. Because the total kinetic energy of Eq. (2.4) can always be broken into positive-definite, independent, center-of-mass and internal components,

$$T_{total} = T_{COM} + T_{int} = \frac{1}{2} M \dot{\mathbf{r}}_{COM}^2 + \frac{1}{2} \sum_{j=1}^N m_j (\dot{\mathbf{r}}_j^{(int)})^2, \\ \mathbf{r}_j^{(int)} \equiv \mathbf{r}_j - \mathbf{r}_{COM},$$

minimizing the kinematic length in Eq. (2.3) requires that the center-of-mass portion vanishes.

In computational practice, the location of the $V(\mathbf{R}) = E_L$ boundaries is unknown, so geodesics are derived by taking small free-propagation steps towards the final endpoint in accord with Eq. (2.5), checking with each step that the $V(\mathbf{R}) < E_L$ condition remains satisfied.²⁹ Whenever an attempted step leads to a configuration $\mathbf{R}_i$ with too high a value of the potential energy, an “escape step” is instituted by finding the nearest configuration $\mathbf{R}_f$ such that $V(\mathbf{R}_f) = E_L$, and the attempt at free propagation is repeated from the new configuration. It is easy to see that for concave obstacles (potential barriers), this procedure will lead to repeated errant steps followed by escapes, with the sequence of resulting $\mathbf{R}_f$ configurations generating a path along the obstacle boundary—precisely as required by the Kuhn-Tucker theorem.²⁹

Previous geodesic-finding applications have found it convenient to implement the escape steps $\mathbf{R}_i \rightarrow \mathbf{R}_f$ via a Newton-Raphson (NR) root search in the direction of the (negative) potential gradient,

$$\Delta \mathbf{R} = \mathbf{R}_f - \mathbf{R}_i = \Delta \mathbf{R}^{NR}, \quad (2.9)$$

$$\Delta \mathbf{R}^{NR} \equiv \frac{\Delta V}{|\nabla V(\mathbf{R}_i)|^2} \nabla V(\mathbf{R}_i), \quad (2.10)$$

$$\Delta V \equiv E_L - V(\mathbf{R}_i) < 0, \quad (2.11)$$

iterating, if necessary, until the boundary is reached.^{29-31,33} Unlike the free-propagation steps, however, these particular steps are not guaranteed to leave the center of mass inviolate. Still, one can easily write a projected version of Eqs. (2.9)–(2.11) which steps along precisely the desired constraint surface.

Suppose we define $\mathbf{I}$ to be the $3N \times 3N$ identity matrix and $\mathbf{E}$ to be a $3N \times 3$ counterpart,

$$(\mathbf{E})_{j\mu,\nu} = \delta_{j\mu\nu}, \quad (j = 1, \dots, N), \quad (\mu, \nu = x, y, z).$$

Then, if we simply replace the escape step of Eq. (2.9) by

$$\Delta \mathbf{R} = \mathbf{R}_f - \mathbf{R}_i = (\mathbf{I} - \mathbf{E} \cdot \mathbf{M}) \cdot \Delta \mathbf{R}^{NR}, \quad (2.12)$$

we can use the properties of the $\mathbf{M}$ matrix defined in Eq. (2.6),

$$\mathbf{M} \cdot \mathbf{E} = \mathbf{e}, \quad \mathbf{e} \cdot \mathbf{M} = \mathbf{M}$$

(with $\mathbf{e}$ the 3×3 unit matrix) to show that such revised steps manifestly preserve the center of mass, Eq. (2.7),

$$\Delta \mathbf{r}_{COM} = \mathbf{M} \cdot \Delta \mathbf{R} = (\mathbf{M} - \mathbf{M} \cdot \mathbf{E} \cdot \mathbf{M}) \cdot \Delta \mathbf{R}^{NR} = \mathbf{0}.$$

Moreover, since the displacement subtracted in going from Eqs. (2.9)–(2.12),

$$(\mathbf{E} \bullet \mathbf{M} \bullet \Delta \mathbf{R}^{NR})_{j\mu} = \sum_{k=1}^N (m_k/M) (\Delta \mathbf{r}_k^{NR})_{\mu},$$

is the same for every atom j , the difference between Eqs. (2.9) and (2.12) is just a uniform back-translation of the entire system. Inasmuch as the potential energy is invariant to any such translation, Eq. (2.12) is guaranteed to fare precisely as well as Eq. (2.9) in satisfying the root search condition $V(\mathbf{R}_f) = E_L$.

One can make other choices for projected escape steps that will also preserve the center of mass, but Eq. (2.12) suffices for our purposes here.⁴⁶

III. MODELS AND COMPUTATIONAL METHODS

A. Potential surface and overall energetics

As we noted in the Introduction, all of the analysis in this paper is based on examining the classical dynamical possibilities afforded by motion on a single (ground-state) potential surface. The particular S_0 potential surface we use for formaldehyde, that of Zhang *et al.*,⁴⁷ was derived from a fit in internal coordinates to over 10^5 high-level, single-point, quantum chemistry calculations. The global minimum of this potential, the calculated equilibrium geometry for H_2CO , lies 1931.2 cm^{-1} lower than the H_2 and CO molecular products and $33\,238.7 \text{ cm}^{-1}$ lower than the threshold for dissociation into H and HCO radicals. The saddle point along the principal route from H_2CO to molecular products resides at $30\,583.7 \text{ cm}^{-1}$.⁴⁸

In an effort to mimic the conditions produced by the kinds of (ca. 330 nm) photodissociation experiments that lead to roaming,⁶ we restrict ourselves to nuclear geometries obtained by microcanonically sampling the zero-overall-angular-momentum phase space permitted by an internal energy $E = 37\,600.0 \text{ cm}^{-1}$. (After subtracting the 5844 cm^{-1} calculated H_2CO zero-point energy,⁴⁷ this value corresponds to a photodissociation wavelength of 315 nm.) Molecular dynamics studies on this potential carried out at this energy not only exhibit classical trajectories that manifestly roam, they also lead to the same kind of low-energy shoulder in the J_{CO} product rotational state distribution frequently taken as a spectroscopic signature of H_2CO roaming.^{1,6} Consistent with other molecular-dynamics-based investigations into formaldehyde,³ we find roaming occurs in just a small fraction of the trajectories dissociating to molecular products.

Masses of the relevant atoms are listed in Table I.

TABLE I. Atomic masses used in this paper.

Element	Mass (a.u.) ^a
H	1 837.15
C	21 874.66
O	29 156.95

^aMasses reported in atomic units.

B. Specifying the itineraries of geodesic paths

The principal results of this paper are the differences found in geodesic paths when we consider the endpoints (or, in the case of roaming, way-station points) associated with different reaction channels. Geodesic endpoints are completely specified by the beginning and ending atomic coordinates—momenta are irrelevant. However correct microcanonical sampling of the reactant and product configurations is most easily accomplished by sampling phase space (encompassing both the coordinates and the momenta), which automatically gives us appropriately weighted sets of coordinates.

Reactant phase-space points were sampled using the standard Monte Carlo procedure of Peslherbe, Wang, and Hase.⁴⁹ The total energy E is deposited into the harmonic normal modes associated with the global potential-energy minimum of our system and an anharmonic correction applied in such a way as to microcanonically sample the zero center-of-mass momentum/zero angular momentum phase space. A corresponding set of molecular and radical *product* configurations was then created by employing molecular dynamics to propagate those initial phase-space points forward in time. Trajectories deemed reactive (*vide infra*) which terminated with an H–H separation of less than $6 a_0$ ($3.18 \text{ \AA}$) were considered to have produced molecular products; the remaining reactive trajectories were classified as having led to radical products. Of the 39 999 trajectories we ran, 29 236 were assigned to the radical product channel, 7393 to the molecular product channel, and the remaining 2790 trajectories did not react.

To enable our geodesic pathway analysis to look explicitly at the phenomenon of roaming, a special variety of molecular dissociation, we need an analogous, purely coordinate-space, definition of this phenomenon. Inasmuch as roaming is an involved sequence of dynamical events, how one should create such a definition is far from obvious, but here we simply propose a *roaming region*: a finite volume of configuration space (as opposed to a reduced-dimensional manifold in either configuration or phase space) constructed so that the vast majority of classical trajectories passing through it manifestly roam on their way to forming molecular products.⁵⁰ Our corresponding roaming geodesics are then the shortest allowed pathways that start in one of our reactant configurations, pass through one of our roaming-region configurations, and end in one of our molecular-product configurations.

Given that roaming in formaldehyde requires that a single hydrogen atom separate itself from the other 3 atoms by a sizeable amount and then return, a plausible (though hardly unique) configuration-space definition of the roaming region is that (1) the distance between the two hydrogens, r_{HH} , is large enough to exclude vibrationally excited, but bound, H_2CO , and small enough to exclude the dissociated radical products,

$$d_{\text{min}} < r_{\text{HH}} < d_{\text{max}} \quad (3.1)$$

and (2) the restoring force attempting to return the roaming hydrogen to its HCO core is insufficient to direct a prompt return. That is, $F(\mathbf{H}_r)$, the component of the force on $\mathbf{H}_r$, the roaming hydrogen, in the direction of the center of mass of the

TABLE II. Parameter values defining the formaldehyde roaming region.

Parameter	Value
$F_{\max}$	$-103 \text{ cm}^{-1}/\text{\AA}$
$d_{\min}$	$3.10 \text{ \AA}$
$d_{\max}$	$4.76 \text{ \AA}$

entire system ($\mathbf{r}_{\text{COM}}$) must be below a certain threshold,

$$F(\mathbf{H}_r) = (-\nabla_{\mathbf{H}_r} V) \cdot \hat{\mathbf{r}}_{\mathbf{H}_r \rightarrow \text{COM}} < F_{\max}, \quad (3.2)$$

$$\hat{\mathbf{r}}_{\mathbf{H}_r \rightarrow \text{COM}} = (\mathbf{r}_{\text{COM}} - \mathbf{r}_{\mathbf{H}_r})/|\mathbf{r}_{\text{COM}} - \mathbf{r}_{\mathbf{H}_r}|.$$

The $d_{\min}$, $d_{\max}$, and $F_{\max}$ parameters we used to define the roaming region are given in Table II. Some 381 of our 7393 molecular-product-generating molecular dynamics trajectories (4.78%) were classified as roaming by virtue of their having passed through this region. Interestingly, adopting the more standard, experimentally motivated, definition of roaming, that the final CO products have to be rotationally cold ($J_{\text{CO}} < 16$), would have yielded a similar fraction, but would have missed trajectories that clearly roam but produce rotationally hot CO products.⁵¹

Adequate sampling of configurations for geodesic purposes requires more points in the roaming region than are provided by these few trajectories, so we generate those points by a Monte Carlo scheme. We take configurations in the roaming region from each roaming trajectory and use each as a seed for a 10^5 -step random walk through the portion of configuration space that simultaneously satisfies our landscape-energy and roaming-region criteria. Trial moves are generated by first displacing a single randomly chosen atom j by some randomly chosen vector $\vec{\Delta}_j$ and then restoring the position of the center of mass by applying a uniform back translation to all four atoms. The net trial move for each atom k was thus

$$\vec{\Delta}r_k = \vec{\Delta}_j \left(\delta_{jk} - \frac{m_j}{M} \right), \quad (k = 1, \dots, 4). \quad (3.3)$$

The vectors $\vec{\Delta}_j$ were sampled from a uniformly distributed set of random-length single-atom moves directed along randomly chosen, laboratory-frame, Cartesian axes. In particular, we took each $\vec{\Delta}_j = \pm x \delta \hat{\mu}$, with x selected from $[0, 1]$, $\delta = 0.145 a_0$ the step size, and $\hat{\mu}$ a unit vector in the x , y , or z direction. Moves were accepted if, and only if, they met both our energy and roaming criteria (giving us an acceptance probability of 0.84). A set of 50 000 statistically independent roaming-region configurations was then obtained by selecting configurations from our random walks.

C. Calculating the geodesics themselves

In the thermodynamic limit, consistency with the energies we have been discussing would require that we use a potential-energy landscape energy E_L equal to the microcanonically averaged potential energy associated with our $37\,600 \text{ cm}^{-1}$ total energy.²⁸ We are not precisely in that limit of course, but we nonetheless set our landscape energy to a thermodynamic value for roaming molecular dynamics trajectories launched from our initial conditions,

$$E_L = 34\,928.7 \text{ cm}^{-1}.$$

This value is (roughly) the most probable potential energy for trajectory configurations obeying the H/H-distance roaming criterion defined in Sec. III B. Repeating our analysis with landscape energies set as low as the $33\,238.7 \text{ cm}^{-1}$ radical dissociation threshold (the lowest value permitting a comparison of geodesics leading to molecular and radical-dissociation products) led to virtually identical results.⁵²

The desired geodesics connecting each pair of initial and final configurations ($\mathbf{R}_1, \mathbf{R}_2$) were calculated using the expressions provided in Sec. II. At each location along the path, we used Eq. (2.5) to attempt small free-propagation steps of length $\delta_s = 0.05 a_0 = 0.0265 \text{ \AA}$ in the 12-dimensional configuration space (with $\mathbf{R}_0$ taken as the current configuration and $\mathbf{R}_i = \mathbf{R}_2$). Analyzing the convergence of the total geodesic path length [Eq. (2.3)] between a sample pair of endpoints confirmed the appropriateness of this particular choice of step length (Table III).

Steps were accepted as long as they were energetically allowed, ($V(\mathbf{R}) < E_L$), but when an end-of-step configuration violated that condition, that configuration, $\mathbf{R}_i$, was translated towards the energetically allowed region using the mass-weighted gradient displacement given in Eq. (2.12). If that displacement sufficed to escape the forbidden region, the free steps toward $\mathbf{R}_2$ were resumed; if not, Eq. (2.12) was applied repeatedly until the potential energy boundary was reached (to within a fractional tolerance of 10^{-6}), and the free propagation re-initiated from that point. The overall geodesic search was terminated as soon as an accepted step took the system within a 12-dimensional Euclidean distance $2\delta_s$ of $\mathbf{R}_2$.

When this combined free-plus-escape-step algorithm failed to locate a path despite accumulating 100 times more steps than that required by an energy-unconstrained path between the same endpoints, those endpoints were regarded as being disconnected. The paths that were located, though, manifestly satisfied the requisite Kuhn-Tucker theorem (by construction).²⁹ However, since the theorem provides a necessary, but not sufficient, criterion for a minimum-length path, the candidate paths were then refined in two stages. We first simply reversed the initial and final configurations and used our standard path-finding algorithm to find a potentially different geodesic path. We then took whichever of this pair of paths was shorter and subjected it to a more local optimization procedure.²⁹

TABLE III. Convergence of geodesic path lengths with step size.^a

δ_s (a.u.)	ℓ_{total} (a.u.)
5	1067.32
5×10^{-1}	1098.27
5×10^{-2}	1102.08
5×10^{-3}	1102.39
5×10^{-4}	1102.42

^aConvergence of the unoptimized total kinematic path length ℓ_{total} [Eq. (2.3)] with step sizes δ_s used in the free-propagation steps (with both quantities reported in atomic units). The data shown here are for direct-channel geodesic paths between a single pair of formaldehyde/diatomic-molecular-products endpoints, and, unlike the results shown in Figs. 3 and 4, report the lengths of entire geodesics between reactant and asymptotic products. The endpoint pair chosen is one of the more computationally challenging direct-channel pairs. The straight-line kinematic distance separating this pair is 1015.34 a.u.

To implement this second stage, small random displacements were applied to a randomly selected configuration $\mathbf{R}$ along each $\mathbf{R}_1 \rightarrow \mathbf{R}_2$ path, generating a trial new configuration $\mathbf{R}'$. These displacements were carried out in precisely the same fashion as our previously described Monte Carlo sampling [using Eq. (3.3)] but with the randomly directed single-atom displacements $\hat{\Delta}_j$ of a fixed length δ_s rather than of a variable length. Geodesic paths $\mathbf{R}_1 \rightarrow \mathbf{R}'$ and $\mathbf{R}' \rightarrow \mathbf{R}_2$ were then computed and the combined path length compared to the previous path length. Whenever this total length was shorter, the $\mathbf{R}_1 \rightarrow \mathbf{R}' \rightarrow \mathbf{R}_2$ path was deemed to replace the previous path; otherwise the previous path was reinstated. This procedure was repeated until 10^3 consecutive displacements failed to shorten the path, so optimizing a geodesic typically required computing on the order of 10^4 paths. In most cases, optimization only led to the order of 5% reduction in path length, although a few more substantial reductions were also observed.

The important point from our perspective is that the geodesics we study are genuine indicators of the geometry of our potential surface, and not simply reflections of the classical trajectories used to define our endpoints. To help ensure that that dynamics plays no role in our results, all of the geodesics results reported here (Table IV) were computed from endpoint pairs $(\mathbf{R}_1, \mathbf{R}_2)$ selected randomly from our sets of beginning and ending configuration-space points, without regard to whether the pairs were connected by an actual classical trajectory. However this precaution may not have been necessary. A comparison calculation of direct-dissociation geodesics computed from dynamically connected endpoints gave a distribution of geodesic lengths with a virtually identical mean and standard deviation to that from our dynamically uncorrelated endpoints.⁴⁶

D. Geodesic path lengths and additional details

In addition to the total geodesic path lengths specified by Eq. (2.3), we also study the component path lengths for various degrees of freedom, α , of the diatomic products,

$$\ell_\alpha = \int_0^t d\tau \sqrt{2T_\alpha(\dot{\mathbf{R}}, \mathbf{R})} \quad (3.4)$$

$$(\alpha = \text{H}_2(\text{vib}), \text{H}_2(\text{rot}), \text{CO}(\text{vib}), \text{CO}(\text{rot})),$$

TABLE IV. Number of geodesic paths employed in our calculations.^a

Channel	Number of geodesics
Radical	49 981
Direct	50 000
Roaming	37 313

^aFor each channel, we looked for geodesics between 50 000 different, randomly chosen, reactant-product endpoint pairs. Dynamically uncorrelated endpoints for the two molecular-product channels (direct and roaming) were chosen by randomly pairing initial and final configurations from a set of 1000 direct-dissociation-channel molecular dynamics trajectories. An analogous random pairing with final configurations from 6938 radical-dissociation-channel trajectories was used to generate the radical channel endpoints. Candidates for roaming geodesics were further constrained by requiring that they also pass through a randomly chosen point in the roaming region. The figures report the number of geodesics actually found, all of which were used in constructing the figures presented in this paper.

with T_α the kinetic energy for that degree of freedom. For each diatomic A-B, those kinetic energies can be written as

$$T_{\text{vib}}^{AB} = \frac{1}{2} \mu_{AB} \left(\frac{dr_{AB}}{d\tau} \right)^2, \quad T_{\text{rot}}^{AB} = \frac{1}{2} \mu_{AB} r_{AB}^2 \left(\frac{d\hat{r}_{AB}}{d\tau} \right)^2, \quad (3.5)$$

$$\mathbf{r}_{AB} \equiv \mathbf{r}_B - \mathbf{r}_A \equiv r_{AB} \hat{r}_{AB}$$

in terms of the respective reduced masses μ_{AB} , and the τ (path-progress-variable) dependences of the interatomic distances r_{AB} and orientational unit vectors $\hat{r}_{AB}$.

Numerically computed locations along a path $\mathbf{R}(\tau)$ are only available at discrete, unevenly spaced, τ values, $\mathbf{R}_i \equiv \mathbf{R}(\tau_i)$. But since in each case, the kinetic energy has the basic structure,

$$T_\alpha = \frac{1}{2} \left(\frac{d\mathbf{R}(\tau)}{d\tau} \right)^T \cdot \mathbf{g}_\alpha(\mathbf{R}(\tau)) \cdot \left(\frac{d\mathbf{R}(\tau)}{d\tau} \right), \quad (3.6)$$

with each $\mathbf{g}_\alpha$ a 12×12 matrix, using the lowest-order forward-finite-difference expressions for the derivatives in Eq. (3.6), and the corresponding finite-difference expressions for the integral in Eq. (3.4), allows us to evaluate each length. For an n -step path, that length is just the sum of n mass-weighted Euclidean step lengths,

$$\ell_\alpha = \sum_{i=0}^{n-1} \sqrt{(\mathbf{R}_{i+1} - \mathbf{R}_i)^T \cdot \mathbf{g}_\alpha(\mathbf{R}_i) \cdot (\mathbf{R}_{i+1} - \mathbf{R}_i)}. \quad (3.7)$$

When needed, molecular dynamics simulations were performed by using the 4th-order Runge-Kutta integrator⁵³ from the GNU Scientific Library⁵⁴ to propagate Hamilton's equations. The time step employed, δt , was 5 atomic time units ($\delta t = 0.121$ fs) and configurations were recorded every 20 δt . Trajectories were launched from reactant initial conditions sampled as described in Sec. III B and terminated when either the H_2 -CO center of mass separation became larger than 12 a_0 (6.350 Å), denoting dissociation, or when the elapsed time exceeded $10^5 \delta t$ (12.1 ps), whichever occurred first.

Random numbers used in our Monte Carlo sampling were created via the MT19937 variant of the Mersenne Twister algorithm implemented in the GNU Scientific Library.⁵⁴ Gradients of the potential were computed by taking numerical partial derivatives along the 12 Cartesian laboratory-frame coordinates using a 5-point central-difference formula with an inter-abscissa spacing of $10^{-4} a_0$.⁵³

IV. RESULTS

A. Molecular dynamics

To help orient our thinking about the relevant portions of configuration space in this problem, we begin in Fig. 1 by projecting three representative examples of dissociative molecular dynamics trajectories into the H-H distance/ H_2 -CO distance plane. Each trajectory starts in the general vicinity of the global potential minimum, oscillates for quite some time within a tangled region which we will call the “bird’s nest,” and then chooses one of the three dissociative channels: dissociating promptly into H + HCO radical products, dissociating directly into H_2 + CO molecular products, or taking a more protracted, roaming, route to the same molecular products. In point of fact, while we have called these trajectories “representative,” there

FIG. 1. Representative molecular dynamics trajectories for formaldehyde dissociation. The trajectories plotted here are projections onto the two-dimensional surface prescribed by the $\text{H}_2\text{-CO}$ center-of-mass separation ($r_{\text{H}_2\text{-CO}}$) and the H-H separation ($r_{\text{H-H}}$) coordinates. Examples are shown for radical dissociation (blue dashes), direct molecular dissociation (red dots), and roaming molecular dissociation (solid black line). For reference, we also show the projections of a number of important special points (open symbols) on the formaldehyde potential surface we use: The circle, square, and star are, respectively, the global minimum, the first-order saddle point associated with the tight transition state leading to direct molecular dissociation, and the second-order saddle point associated with the S_0/S_1 conical intersection. The open triangle marks the location of a putative “roaming saddle point” reported in the literature—which, we note, does not seem to correspond to any stationary points of the potential employed here. In addition to these noteworthy points, we indicate the locations of what we define to be the “roaming region” and “bird’s nest” region of the potential surface (*vide infra*). For this and all succeeding figures, dimensioned quantities are given in atomic units.

is actually a substantial diversity in the trajectories that roam. We illustrate this feature in Fig. 2 by showcasing six randomly selected roaming trajectories in the same kind of plot.

FIG. 2. Six representative molecular dynamics trajectories that exhibit roaming, all projected onto the same kind of plot used in Fig. 1. The various noteworthy points and regions of the potential surface are described in the caption of Fig. 1. Notice that while all six of these trajectories display the high levels of H_2 vibrational excitation characteristic of roaming (as indicated by the large $r_{\text{H-H}}$ oscillations around $2a_0$), at least half of the trajectories avoid passing anywhere near the roaming saddle point.

These two figures also display projections of some noteworthy special points:⁴⁷ the global minimum of the potential we are using, the “tight” saddle point one would have expected to control the exit to molecular products, a second-order saddle point (which is quite close to the computed site of the S_0/S_1 conical intersection in formaldehyde),²² and most notably for our purposes, a point placed at the location that has been proposed as a roaming saddle point (a gateway to roaming-derived molecular products).^{25,26,48}

Two features stand out in these results. One is that the system’s decision whether to roam or not is apparently being made early on, inside the bird’s nest area, not in the vicinity of the putative roaming saddle. Indeed, while the roaming trajectory shown in Fig. 1 eventually seems to pass through the heart of this putative saddle (or at least through its 2-dimensional projection), an examination of Fig. 2 reveals that many instances of roaming do not do so. What different roaming paths do have in common seems to lie within the same region that also leads to radical and direct-molecular dissociations.

The second point, which quickly follows, is that it is all but impossible to tease out the proclivities towards the different pathways from the behavior of the classical trajectories in this jumbled region. What distinguishes the tendency towards one kind of pathway or another is buried under the noise of the vibrational dynamics—which is why we need to compare the geodesic (*inherent*) pathways.^{31–33}

B. Geodesics

Within the geodesic formalism, the key indicator of the character of a geodesic path is its kinematic length, Eq. (2.3). The greater the lengths of these (presumably) most-efficient pathways, the longer we expect the actual dynamics has to take to move between the endpoints.⁵⁵ In much the same way, the kinematic lengths of the individual degrees of freedom, Eqs. (3.4) and (3.5), are a measure of the time we expect to be spent in moving the respective coordinates.^{31,33}

Even so, there would be no particular purpose served in comparing the geodesic lengths that the different channels traverse before they reach the asymptotic locations of their products. We already know that roaming involves far more wandering than the other channels do. Suppose, instead, we ask about the relative kinematic lengths of just those portions of the geodesics that lie within the shared, bird’s nest, region shown in Fig. 1. That is, suppose we find geodesic paths appropriate for each channel, but only compute their kinematic length integrals up until the first time the system leaves the bird’s nest. There is no *a priori* reason to expect some geodesics within this region to be noticeably longer than any other geodesics within the region, so any observations to the contrary can be read as an indication of different dynamical propensities.

We need to define the bird’s nest to have $r_{\text{HH}} < 5.867 a_0$, so that it is bounded above by the roaming region, and $r_{\text{H}_2\text{-CO}_2} < 3.420 a_0$, so that it includes the smallest region containing the bulk of the configurations the system occupies during its original oscillatory incubation period. The latter choice turns out to coincide reasonably well with the ridge connecting the tight saddle point and the second-order saddle.^{47,48} With these choices, we can now look at the distribution of

kinematic lengths traversed by the different kinds of geodesic paths (Fig. 3).

Despite the fact that the geodesics cannot peer into the future to see what routes they will take once they leave the bird's nest, geodesics that will eventually roam have a remarkably broad length distribution even within the nest. Unlike the pathways that will end up following the radical channel, which seem to be somewhat narrowly constrained, and in even starker contrast to those destined to follow the direct-molecular channel, which are still more tightly constrained, there are evidently a wide variety of different pathways capable of positioning a trajectory to roam.

A look at the components of geodesics within the bird's nest affiliated with a variety of different molecular-product degrees of freedom (Fig. 4) tells much the same story.⁵⁶ Radical dissociation does take place in a wider variety of ways than direct molecular dissociation does (especially from the perspective of the H₂ coordinates), but both channels continue to appear much more constrained than the roaming channel, regardless of which coordinate one looks at.

It is worth emphasizing that the component kinematic lengths being reported here are measures of the potential-surface's propensity to focus kinetic energy into *individual degrees of freedom during the oscillatory incubation period*. We are not only not looking at correlations between degrees of freedom (does CO rotation cool down when H₂ vibration heats up?), we are not discussing anything about the eventual product-state distributions.⁹ At this stage, we are simply seeing that the onset of roaming is unique among the dissociative-channel onsets in not precluding having large amounts of kinetic energy spread among the diatomic-product-molecule degree of freedom. In short, the onset of roaming seems to be more random, making it a more entropically favorable option than other routes.

So, why is roaming so unspecific a process? We can turn to another property of geodesics to gain some insight into that

FIG. 3. Probability distributions for the portions of the *total* kinematic lengths of formaldehyde dissociation geodesics that lie within the bird's nest region. There is no *a priori* reason to expect that any of the dissociation channels should traverse the bird's nest region any differently than any other channel. Yet, the probability distribution for the (bird's nest) lengths of geodesics that roam on the way to molecular products (*roaming*) is substantially broader than that for either geodesics leading to radical products (*radical*) or geodesics leading directly to molecular products (*direct*).

FIG. 4. Probability distributions for the portions of the *component* kinematic lengths of formaldehyde dissociation geodesics that lie within the bird's nest. This figure is essentially a more detailed look at Fig. 3. As with Fig. 3, Fig. 4 compares geodesic kinematic lengths within the bird's nest for different reaction channels, but here the lengths reported are those associated with specific combinations of atomic coordinates. Note that these combinations (the H–H distance and the C–O orientation, for example) are only being assayed in the bird's nest region, before any dissociation has occurred, but they only take on the physical significance indicated by our labels (“H₂ vibration,” “CO rotation,” etc.) when (and if) diatomic molecular products have been formed.

question: the fraction of the geodesic length that lies along the potential energy boundary. As we discussed in Sec. II, if a system has landscape energy E_L , then every geodesic path consists entirely of curvilinear pieces winding along the boundaries of the potential-energy-allowed region ($V(\mathbf{R}) = E_L$), and straight-line segments traversing portions of the path residing within the allowed region ($V(\mathbf{R}) < E_L$).²⁹ The kinds of values we see for the fraction of the path length lying within some immediate neighborhood of a boundary

$$\langle \Theta \rangle = \frac{\int_0^t d\tau \sqrt{2T[\dot{\mathbf{R}}(\tau), \mathbf{R}(\tau)]} \Theta[\mathbf{R}(\tau); E_L]}{\int_0^t d\tau \sqrt{2T[\dot{\mathbf{R}}(\tau), \mathbf{R}(\tau)]}}, \quad (4.1)$$

$$\Theta[\mathbf{R}; E_L] = \begin{cases} 1 & (E_L - V(\mathbf{R}))/E_L < \delta \\ 0 & \text{otherwise} \end{cases},$$

thus give us a sense of how rugged the potential energy terrain is in the relevant regions of configuration space.

In Fig. 5, we plot the distribution of these boundary-fraction values obtained for each of our three different reaction channels, looking again, at just those portions of the geodesics that have yet to leave the bird's nest region.⁵⁷ The striking result here is that while the same region of the potential surface would seem to govern all three channels, the distributions of $\langle \Theta \rangle$ are dramatically different. Radical and direct molecular dissociation have 52% and 74% of their boundary-fraction values identically zero (respectively), whereas roaming has less than 9% in the category, and a most probable (nonzero) value of $\langle \Theta \rangle = 0.75$. Remembering that our landscape energy is quite high (almost 35 000 cm⁻¹, putting it 1700 cm⁻¹ above the radical dissociation threshold) leaves us with a picture of the pathways that will eventually roam as those that compelled to skirt the obstacles in the potential-energy landscape [regions

FIG. 5. Probability distributions for the fraction of each dissociation geodesic that follows the potential surface boundaries while it is passing through the bird's nest region of the potential, Eq. (4.1). The presence of geodesic segments spanning the regions between the boundaries implies that all three dissociation channels feature at least some delta function contribution at zero (which appears here as a finite-width step owing to our non-zero histogram bin size). However the roaming channel evidently has a relative paucity of such spanning segments.

with $V(\mathbf{R}) > E_L$] instead of flying over the valleys between them, as both the radical and direct-molecular dissociation paths seem to do.

This kind of result is certainly consistent with the longer geodesic path lengths that we see with roaming events. Geodesic paths can only become lengthy when they become convoluted—when they wind around obstacles. More than that, the pinball-machine character of the dynamics that this finding suggests may explain the origins of the essential randomness of the roaming process.

V. CONCLUDING REMARKS

Perhaps the most surprising result of our analysis is that there is information to be gleaned from the seemingly erratic pathways that atoms take in about-to-dissociate molecules in the interval *before* they leave their reactant potential well. The key is to look at the paths defined by the geodesics rather than the trajectories themselves, a step that reveals the “inherent dynamics.” If, for example, all one had was the kind of elementary, nearly integrable, dynamics seen with multiple weakly coupled oscillators, the configuration space routes traced by the classical trajectories would still be strikingly intricate unless the oscillator frequencies happened to have rational ratios.⁵⁸ But, even in the irrational cases, the landscape-ensemble geodesics would just be straight lines between the potential surface boundaries prescribed by the landscape energy. Both the radical dissociation and the direct molecular dissociation of formaldehyde display almost this level of simplicity in their reactant-well geodesics.

Roaming, by contrast, evidently arises when the geodesics follow the twists and turns of the potential surface boundaries in the reactant well. Geodesics of that sort appear random, with a variety of different close decisions between alternative paths leading to a broad array of possible geodesic features. Seeing such broad distributions is reminiscent of how chemical reactions proceeding through orbiting complexes

lead to a broad distribution of exit angles.⁵⁹ The remarkable feature of our results, though, is that we are seeing a broad distribution not just of the possible classical trajectories, but of their centroids. Indeed, from the geodesic perspective, *the essential difference between trajectories that will eventually roam and those that proceed via either of the more direct channels is how much more freedom a system has if it chooses to roam, and how much more narrowly prescribed non-roaming paths are.*

This take on roaming has a number of implications for some of the published approaches to understanding the phenomenon. Much of the previous work has focused, quite sensibly, on understanding the experimental branching ratios between radical and molecular products.^{4,9,10} But such an analysis misses a central point revealed by the geodesic results: the radical and direct-molecular channels have much more in common with each other than the direct route to molecular products has in common with the roaming route to the same products. We suggest that the essentially stochastic character of the roaming pathways sets them apart.

We do note the success that phase-space theory,⁹ itself a statistical perspective, has had in explaining experimental branching ratios in roaming systems, but it is important to realize that that success is largely orthogonal to our observation. Phase-space theory working here tells us that observables such as the molecular-product/radical-product branching ratio are largely independent of system's detailed dynamics, a not unexpected result given the near-threshold energies being investigated. Our finding, by contrast, is that a particular mechanistic pathway to one of these two kinds of products (a minor one in the case of formaldehyde) is even more statistical than one would have expected; in effect a statement that the system's behavior must be sensitive to at least some particulars of the dynamics.

This sensitivity is entirely consistent with the extensive effort being devoted to creating detailed portraits of the invariant phase space structures guiding roaming processes.^{14–18} Indeed, the stochastic nature of roaming is just what one might have predicted on discovering that there are special phase space “sinks” that roaming trajectories find themselves trapped within.^{15–17} However, the practical difficulties of investigating these kinds of phase space intricacies have limited almost all of the analysis to date to either two-coordinate model systems or to two-dimensional fragments of higher-dimensional problems. What our geodesic-based treatment has shown is that by considering just the coordinate-space aspects of roaming, it is possible to understand some of its uniqueness while doing justice to the higher dimensional contexts in which it actually happens.

Perhaps better still, this treatment may make it possible for us to understand some of the universality of roaming. The invariant phase-space structure approach seems to rely on painstakingly mapping out each of the noteworthy invariants for each new problem, making it difficult to see why roaming should be so commonplace. By contrast, from the landscape geodesic perspective, it seems likely that all one needs to ensure the presence of roaming-like (i.e., broadly distributed) geodesics are a flat enough asymptotic region of the potential surface and a high enough dimensionality to create

an entropic challenge out of the choice among multiple possible routes to that asymptotic region. From this viewpoint, it is hardly a surprise to see roaming as a recurring phenomenon.

It would certainly be advantageous to be able to forge a more detailed connection between this kind of geodesic analysis and the roaming literature, but to do so will require that we start examining a variety of more detailed questions. Roaming classical trajectories evidently appear over a limited energy window and have a specific energy dependence within that window.^{60,61} Do what we have identified as roaming geodesic signatures show up in that same window? If these same geodesic methods are applied to higher dimensional problems such as acetaldehyde dissociation,^{7,8} are the geodesic signatures indeed the same as in fewer degrees of freedom problems? And, finally, how closely can one tie this picture to more conventional views of reaction dynamics? In particular, in more traditional language, one often characterizes roaming as an instance of frustrated radical dissociation,⁹ meaning that something is impeding the intramolecular energy transfer needed to allow an individual radical species (an H atom or methyl radical, for example) to leave on its own. In condensed matter contexts, geodesic path lengths have been shown to have quantitative connections with diffusion constants.^{29,32} It would be interesting to see if one could ever use geodesic features in small molecule systems in an analogous fashion, as probes of the times needed to accomplish the requisite energy transfer.

SUPPLEMENTARY MATERIAL

See [supplementary material](#) for details concerning the potential surface used in this work.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation under Grant Nos. CHE-1265798 and CHE-1565540. We thank Joel Bowman for illuminating discussions and for supplying us with the formaldehyde potential surface his group created. We also thank Peter Weber for sharing his experimental perspectives and Layne Frechette for helpful comments.

¹A. G. Suits, *Acc. Chem. Res.* **41**, 873–881 (2008).

²N. Herath and A. G. Suits, *Phys. Chem. Lett.* **2**, 642–647 (2011).

³J. M. Bowman and B. C. Shepler, *Annu. Rev. Phys. Chem.* **62**, 531–553 (2011).

⁴J. M. Bowman, *Mol. Phys.* **112**, 2516–2528 (2014).

⁵R. D. van Zee, M. F. Foltz, and C. B. Moore, *J. Chem. Phys.* **99**, 1664–1673 (1993).

⁶D. Townsend, S. A. Lahankar, S. K. Lee, S. D. Chambreau, A. G. Suits, X. Zhang, J. Rheinecker, L. B. Harding, and J. M. Bowman, *Science* **306**, 1158–1161 (2004).

⁷P. L. Houston and S. H. Kable, *Proc. Natl. Acad. Sci. U. S. A.* **103**, 16079–16082 (2006).

⁸B. R. Heazlewood, M. J. T. Jordan, S. H. Kable, T. M. Selby, D. L. Osborn, B. C. Shepler, B. J. Braams, and J. M. Bowman, *Proc. Natl. Acad. Sci. U. S. A.* **105**, 12719–12724 (2008).

⁹D. U. Andrews, S. H. Kable, and M. J. T. Jordan, *J. Phys. Chem. A* **117**, 7631–7642 (2013).

¹⁰S. J. Klippenstein, Y. Georgievskii, and L. B. Harding, *J. Phys. Chem. A* **115**, 14370–14381 (2011).

¹¹L. B. Harding, Y. Georgievskii, and S. J. Klippenstein, *J. Phys. Chem. A* **114**, 765–777 (2010).

¹²L. B. Harding, S. J. Klippenstein, and A. W. Jasper, *J. Phys. Chem. A* **116**, 6967–6982 (2012).

¹³H. Waalkens and S. Wiggins, *Regular Chaotic Dyn.* **15**, 1–39 (2010).

¹⁴F. A. L. Maugière, P. Collins, G. S. Ezra, S. C. Farantos, and S. Wiggins, *Chem. Phys. Lett.* **592**, 282–287 (2014).

¹⁵F. A. L. Maugière, P. Collins, G. S. Ezra, S. C. Farantos, and S. Wiggins, *J. Chem. Phys.* **140**, 134112 (2014).

¹⁶F. A. L. Maugière, P. Collins, G. S. Ezra, S. C. Farantos, and S. Wiggins, *Theor. Chem. Acc.* **133**, 1507 (2014).

¹⁷F. A. L. Maugière, P. Collins, Z. C. Kramer, B. K. Carpenter, G. S. Ezra, S. C. Farantos, and S. Wiggins, *J. Phys. Chem. Lett.* **6**, 4123–4128 (2015).

¹⁸F. A. L. Maugière, P. Collins, Z. C. Kramer, B. K. Carpenter, G. S. Ezra, S. C. Farantos, and S. Wiggins, *J. Chem. Phys.* **144**, 054107 (2016).

¹⁹X. Zhang, J. Rheinecker, and J. M. Bowman, *J. Chem. Phys.* **122**, 114313 (2005).

²⁰B. C. Shepler, B. J. Braams, and J. M. Bowman, *J. Phys. Chem. A* **112**, 9344–9351 (2008).

²¹A. Li, J. Li, and H. Guo, *J. Phys. Chem. A* **117**, 5052–5060 (2013).

²²M. Araujo, B. Lasorne, M. J. Bearpark, and M. A. Robb, *J. Phys. Chem. A* **112**, 7489–7491 (2008).

²³M. P. Grubb, M. L. Warter, H. Xiao, S. Maeda, K. Morokuma, and S. W. North, *Science* **335**, 1075–1078 (2012).

²⁴S. Maeda, T. Taketsugu, K. Ohno, and K. Morokuma, *J. Am. Chem. Soc.* **137**, 3433–3445 (2015).

²⁵L. B. Harding, S. J. Klippenstein, and A. W. Jasper, *Phys. Chem. Chem. Phys.* **9**, 4055–4070 (2007).

²⁶B. C. Shepler, Y. Han, and J. M. Bowman, *J. Phys. Chem. Lett.* **2**, 834–838 (2011).

²⁷E. L. Siebert III, J. T. Hynes, and W. P. Reinhardt, *J. Phys. Chem.* **87**, 2032–2037 (1983).

²⁸C. Wang and R. M. Stratt, *J. Chem. Phys.* **127**, 224503 (2007).

²⁹C. Wang and R. M. Stratt, *J. Chem. Phys.* **127**, 224504 (2007).

³⁰C. N. Nguyen and R. M. Stratt, *J. Chem. Phys.* **133**, 124503 (2010).

³¹D. Jacobson and R. M. Stratt, *J. Chem. Phys.* **140**, 174503 (2014); L. Frechette, D. Jacobson, and R. M. Stratt, *ibid.* **141**, 209902 (2014).

³²Q. Ma and R. M. Stratt, *Phys. Rev. E* **90**, 042314 (2014).

³³L. Frechette and R. M. Stratt, *J. Chem. Phys.* **144**, 234505 (2016).

³⁴For a review, see W. E and E. Vanden-Eijnden, *Annu. Rev. Phys. Chem.* **61**, 391–420 (2010). See also the discussion in Ref. 29.

³⁵T. Keyes and J. Chowdhary, *Phys. Rev. E* **64**, 032201 (2001); J. Kim and T. Keyes, *J. Chem. Phys.* **121**, 4237 (2004).

³⁶G. M. Hocky and D. R. Reichman, *J. Chem. Phys.* **138**, 12A537 (2013).

³⁷F. W. Wiegand, *Introduction to Path-Integral Methods in Physics and Polymer Science* (World Scientific, Singapore, 1986), Chaps. I and II.

³⁸L. S. Schulman, *Techniques and Applications of Path Integration* (Wiley-Interscience, New York, 1981), Chap. 9.

³⁹This integral is proportional to the contour length of the path the system takes because kinetic energies T are proportional to $(ds/d\tau)^2$, with s the distance traveled. Its units, however, are $(mass)^{1/2}(length)$.

⁴⁰Inasmuch as the landscape ensemble describes only configurations (atomic positions), and not momenta, the equivalence referred to here is that between the respective probability densities of the different ensembles in configuration space.

⁴¹P. Hänggi, P. Talkner, and M. Borkovec, *Rev. Mod. Phys.* **62**, 251–341 (1990).

⁴²R. Zwanzig, *Nonequilibrium Statistical Mechanics* (Oxford University, Oxford, 2001), Sec. 4.3.

⁴³A. J. Lichtenberg and M. A. Leiberman, *Regular and Chaotic Dynamics*, 2nd ed. (Springer-Verlag, New York, 1992), Sec. 5.4.

⁴⁴Although we list the components of $\mathbf{R}$ within a line of text here in order to save space, in the equations that follow, we shall regard $\mathbf{R}$ as a 12-dimensional column vector.

⁴⁵P. R. Aaby and M. A. H. Dempster, *Introduction to Optimization Methods* (Chapman and Hall, London, 1974), Chap. 5.

⁴⁶D. Vale-Cofer Shabica, Ph.D. dissertation (Brown University, 2018).

⁴⁷X. Zhang, S. Zou, L. B. Harding, and J. M. Bowman, *J. Phys. Chem. A* **108**, 8980–8986 (2004).

⁴⁸The locations and harmonic frequencies of the principal stationary points of this surface are listed in the supplementary material. This supplementary material section also describes our evidence that this potential surface does *not* have a roaming saddle near the expected location. While we cannot comment on whether a more accurate portrayal of the formaldehyde S_0

surface would exhibit a roaming saddle, our observation that the surface we use produces roaming dynamics suggests that such a saddle is not a necessary ingredient in that dynamics.

⁴⁹G. Peslherbe, H. Wang, and W. Hase, *Adv. Chem. Phys.* **105**, 171–201 (1999).

⁵⁰We should emphasize that this choice of “roaming region” is quite different from what has been given that same label in the literature (Refs. 10 and 17). Our region is a portion of configuration space whose passage seems to be predictive of the roaming that will eventually happen, rather than descriptive of some region traversed during the roaming. Moreover, in order to help us clarify what is unique about the protracted wanderings characteristic of roaming, we need to ensure that we minimize the likelihood of misclassifying any non-roaming paths leading to molecular products. Our goal is not to arrive at an inclusive coordinate-space definition of roaming, but to stringently limit our candidate set of roaming pathways, even if it means erring on the side of excluding possible molecular-product pathways. We note, incidentally, that classical trajectories passing through this region will often lead to radical products.

⁵¹Under our conditions, and with our definitions, 12% of roaming trajectories lead to $J_{CO} \geq 45$, an observation not inconsistent with the literature findings for formaldehyde. Figure 3 in Ref. 20, for example, shows at least some $J_{CO} > 40$ states resulting from roaming.

⁵²We note that both of these choices put our landscape energy above the $33\,224.6\text{ cm}^{-1}$ value that this potential has at the location that has been suggested for a roaming saddle point. See Ref. 26.

⁵³M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1972), Secs. 25.3 and 25.5.

⁵⁴M. Galassi, J. Davies, J. Theiler, B. Gough, G. Jungman, P. Alken, M. Booth, and F. Rossi, *GNU Scientific Library Reference Manual*, 3rd ed. (Network Theory, 2009), www.gnu.org/software/gsl.

⁵⁵Indeed, in diffusion applications, diffusion constants can be shown to scale inversely with the square of the kinematic length (Ref. 29). We expect qualitatively similar behavior here, but have no evidence that the same quantitative relationships should hold.

⁵⁶Note that the designations of rotational and vibrational CO and H₂ coordinates remain well defined via Eq. (3.5) regardless of whether there are distinct CO and H₂ species present.

⁵⁷This plot is computed using a fractional boundary width $\delta = 10^{-4}$; we see negligible changes in the statistics of the boundary fraction when we switch to $\delta = 10^{-5}$.

⁵⁸V. I. Arnold, *Mathematical Methods of Classical Mechanics* (Springer-Verlag, New York, 1980), Sec. 5.

⁵⁹P. J. Kuntz, in *Dynamics of Molecular Collisions, Part B*, edited by W. H. Miller (Plenum, New York, 1976), Sec. 6.5.

⁶⁰S. A. Lahankar, S. D. Chambreau, X. Zhang, J. M. Bowman, and A. G. Suits, *J. Chem. Phys.* **126**, 044314 (2007); S. A. Lahankar, V. Goncharov, F. Suits, J. D. Farnum, J. M. Bowman, and A. G. Suits, *Chem. Phys.* **347**, 288–299 (2008).

⁶¹P. L. Houston, R. Conte, and J. M. Bowman, *J. Phys. Chem. A* **120**, 5103–5114 (2016).