Ring-polymer instanton theory

Review Article

Author(s):
Richardson, Jeremy O.

Ring-polymer instanton theory

JEREMY O. RICHARDSON[*]

Laboratory of Physical Chemistry, ETH Zurich, 8093 Zurich, Switzerland

(Dated: April 30, 2018)

Abstract

Instanton theory provides a simple description of a quantum tunnelling process in terms of an optimal tunnelling pathway. The theory is rigorously based on quantum mechanics principles and is derived from a semiclassical approximation to the path-integral formulation. In multidimensional systems, the optimal tunnelling pathway is generally different from the minimum-energy pathway and is seen to 'cut the corner' around the transition state. A ring-polymer formulation of instanton theory leads to a practical computational method for applying the theory to describe, simulate and predict quantum tunnelling effects in complex molecular systems. It can be used to compute either the rate of a tunnelling process leading to a chemical reaction or the tunnelling splitting pattern of a molecular cluster. In this review, we introduce a unification of theory's derivation and discuss recent improvements to the numerical implementation.

Keywords: instanton; ring polymer; tunnelling; reaction rate theory

_____

[*] jeremy.richardson@phys.chem.ethz.ch

CONTENTS

I. INTRODUCTION

Quantum tunnelling can significantly affect chemical reactivity [1–4] and can enable molecular rearrangements even at low temperatures which would otherwise be forbidden [5]. However the most common theoretical methods for describing chemical reactions, such as transition-state theory (TST) [6–8], treat the nuclear dynamics using classical principles, which neglect these important effects [9].

Instanton rate theory has been used to describe tunnelling dynamics in a wide range of scientific disciplines from subnuclear physics, through condensed-phase quantum dynamics to cosmology [10–13]. The same theory is also applicable to the study of molecular reactions and rearrangements [14–17] and can be rigorously derived as a semiclassical approximation to the exact rate constant [18, 19]. It defines an optimal tunnelling pathway and the quantum process is assumed to take place predominantly along this path. In this way, the theory provides an approximate quantum-mechanical generalization of classical TST for describing thermal rate constants of chemical reactions including the quantum-mechanical effects of tunnelling and zero-point energy.

Above a certain crossover temperature, no optimal tunnelling pathway can be found. The process is then well described by shallow tunnelling theory developed by Wigner based on a parabolic barrier approximation [1, 20]. Below the crossover temperature, the instanton pathway exists and describes a deep tunnelling process. Instanton theory thus also provides a simple definition for the onset of this regime.

Instanton rate theory describes incoherent tunnelling, which occurs for instance in a reaction with scattering boundary conditions or for a reaction coupled to a thermal environment. A related approach exists for describing coherent tunnelling leading to a tunnelling splitting of the ground state of an isolated molecule or cluster [10]. As these approaches have a similar derivation and can be implemented using related algorithms, we discuss them both in this review.

Recently it has become possible to apply instanton theory to complex molecular systems using the ring-polymer instanton method [17, 21–24]. This approach locates the optimal tunnelling pathway on the full potential-energy surface by searching for stationary points of the discretized action using multidimensional optimization techniques. In large complex systems, it is practically impossible to simulate exact quantum dynamics but the semiclassical instanton approach has a good balance between accuracy and efficiency. It has therefore become popular for studying polyatomic reactions in the gas phase [23, 25, 26] and rearrangements of molecules and clusters [22, 27–29]. Due to its remarkable efficiency, it has even been applied successfully to much larger problems of interest from reactions on surfaces [30–33] to hydrogen transfers in enzymes [34]. In this way the method has contributed to a better understanding in various fields including astrochemistry, biochemistry, surface science and high-resolution spectroscopy. Many more interesting applications will be studied with this approach in the future.

A molecular process can be described in terms of the shape of the potential-energy surface (PES) and of the quantum dynamics taking place on it. There are thus two problems to be solved which both require using approximations and are therefore a source of error in the final result. The first problem requires computationally expensive electronic-structure calculations to solve the electronic Schrödinger equation to obtain an *ab initio* PES. The second problem can be solved approximately using instanton theory which requires very little computational power in comparison to the electronic-structure calculations. The errors in the PES typically outweigh those of the instanton approximation as the result is exponentially dependent on features of the PES such as the barrier height. The main reason for the success of instanton theory in describing tunnelling in molecular systems is due to its simplicity. One only needs knowledge of a small part of the PES along a line which describes the optimal tunnelling pathway through the barrier. Electronic-structure calculations typically have to find a compromise between accuracy and efficiency and thus the fewer calculations that have to be performed, the higher the accuracy can be. As the instanton method requires only a small number of accurate *ab initio* calculations to be performed, this leads to a good overall description of the tunnelling process.

The ring-polymer instanton approach is closely related to other path-integral rate theories, including centroid-based quantum transition-state theory (QTST) [35, 36], ring-polymer molecular dynamics (RPMD) [37, 38] and the quantum instanton approach [39, 40]. This

relationship can be used to explain the success of RPMD in predicting deep-tunnelling rates, even for asymmetric systems, and shows how to improve the other approaches [21]. It is particularly important to have a clear understanding of the approximations involved in the derivation of the instanton approach if it is to be extended to new problems or if it is to be used as inspiration for obtaining improved path-integral QTSTs.

The steepest-descent approximation used in instanton theory is not valid for systems with labile modes. This could lead to an error for certain gas-phase reactions if floppy torsional modes are coupled to the tunnelling motion and renders the instanton method unable to treat liquids. Eyring TST is limited in a similar way. To describe tunnelling in a fluxional environment such as liquid water, it is advisable to dispense with instanton theory and use RPMD instead. On the other hand, for reactions in the gas-phase, in solids or on surfaces, the simplicity and efficiency of instanton theory makes it the better candidate, especially if this allows more accurate electronic-structure methods to be used.

In Sec. II a summary of semiclassical mechanics is given and using this we outline a number of different formulations and derivations of instanton theory in Sec. III. This formal theoretical derivation is not directly used in the numerical computation of instanton theory and the working equations used in most practical situations are much simpler and given in Sec. III C. Aspects important for an efficient numerical implementation are discussed in Sec. IV and examples given in Sec. V. In Sec. VI and Sec. VII further applications of instanton theory to electron-transfer reactions and tunnelling splitting calculations are discussed.

II. SEMICLASSICAL APPROXIMATION

The theory of instantons is based on a semiclassical approximation to the path-integral description of quantum mechanics. Before deriving the instanton expression for the rate constant, we first introduce semiclassical trajectories which appear as an asymptotic limit from the path-integral formulation. As the instanton equations involve only imaginary time trajectories, we will define all our quantities as functions of imaginary time. This leads to a more convenient formulation for our purpose, but is nonetheless slightly different from those typically presented in classical-mechanics or path-integral textbooks. The usual equations are related to ours by $t = \mathrm{i}\tau$.

A. Semiclassical trajectories

We consider a particle of mass m moving in a potential $V(x)$ from position x' to x'' in imaginary-time τ. According to Feynman's path-integral description of quantum mechanics [41], the relative probability of this event is $|K(\tau)|^2$ where $K(\tau)$ is the kernel, expressed in path-integral notation as

$$K(\tau) = \int_{x(0)=x'}^{x(\tau)=x''} e^{-S[x(t)]/\hbar} \, \mathcal{D}x(t), \tag{1}$$

where t is a dummy variable running from 0 to τ. The Euclidean action is

$$S[x(t)] = \int_0^\tau \left[\frac{m}{2}|\dot{x}|^2 + V(x) \right] dt, \tag{2}$$

where $\dot{x}$ is the rate of change of position with imaginary time. Note that as we are dealing with imaginary time, the exponent in the kernel is real and negative, which differs from the real-time expression. This kernel can also be written in Dirac-notation using $K(\tau) = \langle x'' | e^{-\hat{H}\tau/\hbar} | x' \rangle$, where $\hat{H}$ is the Hamiltonian, and can thus also be thought of as an element of the density matrix at a temperature defined by $\beta \equiv (k_{\mathrm{B}}T)^{-1} = \tau/\hbar$. This relationship between temperature and imaginary time is fundamental to many path-integral approaches including instanton theory.

All these formulae are valid for either a one- or a multidimensional system. In the latter case, x is a vector of f components and the usual rules of vector calculus apply such that second derivatives with respect to x are $f \times f$ Hessian matrices. Rather than explicitly dealing with different masses m_a for each atom, we assume that the Cartesian coordinates of atom a, $\vec{R}_a$, have been mass-weighted to give $\vec{x}_a = \sqrt{\frac{m_a}{m}} \vec{R}_a$ where m can be freely chosen as a reference mass (typically unity in whatever mass units are being used). We will continue to use this notation throughout this review.

Performing the path integral analytically is possible only for a limited number of systems and so in order to obtain a simple and practical formulation in the general case, we employ a semiclassical approximation. Formally, this approximation is performed by replacing a given function $A(\hbar)$ by $B(\hbar)$ if the asymptotic relation $A(\hbar) \sim B(\hbar)$ holds for $\hbar \to 0$; this is equivalent to the statement $\lim_{\hbar \to 0} A(\hbar)/B(\hbar) = 1$ [42]. An example of this approximation, which we will make much use of, is also known as the steepest-descent approximation,

$$\int e^{-\phi(x)/\hbar} \, dx \sim \sqrt{2\pi\hbar} \left(\frac{\partial^2 \phi}{\partial x^2} \right)^{-\frac{1}{2}} e^{-\phi(x)/\hbar}, \tag{3}$$

where on the right-hand side, x is defined such that $\frac{\partial \phi}{\partial x} = 0$.

The semiclassical approximation of the imaginary-time kernel is [43]

$$K(\tau) \sim \sum_{\text{traj.}} \sqrt{\frac{C}{(2\pi\hbar)^f}}\, e^{-S/\hbar}, \tag{4}$$

where $S \equiv S(x', x'', \tau)$ is the action along a trajectory and the prefactor is defined as the determinant of the second variation:

$$C = \left| -\frac{\partial^2 S}{\partial x' \partial x''} \right|. \tag{5}$$

Equation (4) is equivalent to the van-Vleck propagator [44] in imaginary time. Rather than requiring an integral over all possible paths between the end points, this simpler expression is defined as a sum over imaginary-time classical trajectories. These trajectories are defined such that the action functional is stationary and thus obey the Euler-Lagrange equation,

$$\frac{\mathrm{d}}{\mathrm{d}t} m\dot{x} = \frac{\partial V}{\partial x}. \tag{6}$$

The solutions of this equation are similar to standard (real-time) classical trajectories except that they move according to Newton's equations of motion in the upside-down potential [45].

As shown in Appendix A, first derivatives of the action have useful relations [43],

$$\frac{\partial S}{\partial x'} = -p' \qquad\qquad \frac{\partial S}{\partial x''} = p'' \qquad\qquad \frac{\partial S}{\partial \tau} = E, \tag{7}$$

where p' and p'' are the initial and final momenta and $E = V(x) - \frac{1}{2}m|\dot{x}|^2$ is the conserved energy along the trajectory.

Note further that within the semiclassical approximation, we can split a trajectory into two using $K(\tau) = K_a(\tau_a)K_b(\tau_b)$, where $\tau = \tau_a + \tau_b$ [43]. The kernels of the two shorter trajectories are defined as above using $S_a \equiv S_a(x', x_0, \tau_a)$ and $S_b \equiv S_b(x_0, x'', \tau_b)$, where the intermediate point, x_0, is defined such that $\frac{\partial S_a}{\partial x_0} + \frac{\partial S_b}{\partial x_0} = 0$, i.e. the momentum should be continuous. The prefactors are related by

$$C = C_a C_b \left| \frac{\partial^2 S_a}{\partial x_0 \partial x_0} + \frac{\partial^2 S_b}{\partial x_0 \partial x_0} \right|^{-1}. \tag{8}$$

The semiclassical approximation for a partition function of a vibrational well (with its

minimum potential at an energy of 0) is

$$Z = \int K(x, x, \beta\hbar) \, \mathrm{d}x \tag{9}$$

$$\sim C^{\frac{1}{2}} \left| \frac{\partial^2 S}{\partial x \partial x} \right|^{-\frac{1}{2}} \tag{10}$$

$$= \sum_{j=1}^{f} [2 \sinh \tfrac{1}{2}\beta\hbar\omega_j]^{-\frac{1}{2}}, \tag{11}$$

where we have used the semiclassical expression for the kernel and then integrated by steepest descent. Note that $S \equiv S(x, x, \beta\hbar)$ here. In this case, the classical trajectory is collapsed at the bottom of the well; *i.e.* it has the same position at all times. As the semiclassical formulae depend only on a second-order expansion of the potential around the trajectory, we have effectively made a harmonic approximation for the vibrational well, for which analytical expressions for S and its derivatives are known in terms of the normal mode frequencies ω_j (see Appendix B). This result is equal to the partition function of a set of quantum harmonic oscillators. Note, however, that in general the semiclassical expressions are not always equivalent to a harmonic approximation and can also describe processes such as tunnelling through an anharmonic barrier.

B. Discretized path integrals

The path integral lends itself best to numerical calculations in its discretized form. A path can be discretized into N segments of imaginary-time $\tau_N = \tau/N$, and represented by 'beads', x_i, each expressing one set of coordinates of the f-dimensional system. Then the action functional, Eq. (2), becomes the function

$$S_N(\boldsymbol{x}) = \sum_{i=1}^{N} \frac{m}{2\tau_N} |x_i - x_{i-1}|^2 + \tau_N \left[\tfrac{1}{2}V(x_0) + \sum_{i=1}^{N-1} V(x_i) + \tfrac{1}{2}V(x_N) \right], \tag{12}$$

where $\boldsymbol{x} = \{x_1, \ldots, x_{N-1}\}$ are the $N - 1$ free beads and $x_0 \equiv x'$ and $x_N \equiv x''$ are the fixed end points. This can be derived either from finite differences and a trapezium-rule integration of Eq. (2) or from a Trotter splitting of the Boltzmann operator. The kernel is [41]

$$K(\tau) = \left(\frac{m}{2\pi\tau_N\hbar} \right)^{Nf/2} \int e^{-S_N(\boldsymbol{x})/\hbar} \, \mathrm{d}\boldsymbol{x}. \tag{13}$$

8

This and all other formulae in this section tend to the continuous forms given in Sec. II A in the limit $N \to \infty$.

A steepest-descent approximation gives [46]

$$K(\tau) \sim \left(\frac{m}{2\pi\tau_N\hbar} \right)^{Nf/2} \sqrt{\frac{(2\pi\hbar)^{(N-1)f}}{\det \boldsymbol{\nabla}^2 S_N}} \, \mathrm{e}^{-S_N(\tilde{\boldsymbol{x}})/\hbar} \tag{14}$$

$$\sim \left(\frac{m}{2\pi\tau_N\hbar} \right)^{f/2} (\det \boldsymbol{J})^{-\frac{1}{2}} \, \mathrm{e}^{-S_N(\tilde{\boldsymbol{x}})/\hbar}, \tag{15}$$

where $\tilde{\boldsymbol{x}}$ is the stationary point of $S_N(\boldsymbol{x})$ and the matrix $\boldsymbol{J} = \frac{\tau_N}{m}\boldsymbol{\nabla}^2 S_N(\tilde{\boldsymbol{x}})$ has elements

$$J_{ii'} = -\delta_{i-1\,i'} + 2\delta_{ii'} - \delta_{i+1\,i'} + \delta_{ii'}\frac{\tau_N^2}{m}\nabla^2 V(\tilde{x}_i), \tag{16}$$

for $i \in \{1, \dots, N-1\}$. For multidimensional systems, these elements are $f \times f$ submatrices which build up the block matrix $\boldsymbol{J}$ [47]. By comparison with Eq. (4), we see that in the $N \to \infty$ limit,

$$(\det \boldsymbol{J})^{-1} = \left(\frac{\tau_N}{m} \right)^f C, \tag{17}$$

which can also be shown directly from the Gelfand-Yaglom formula [46].

This discretization scheme is valid for an open-ended path integral but is also strongly related to the more familiar ring polymer, which has become a popular simulation technique in computational chemistry [38, 48–54]. The ring-polymer representation will be discussed further in Sec. III C.

III. INSTANTON RATE THEORY

Instanton rate theory has been obtained independently by a number of people using different approaches. The two earliest papers introducing the theory are by Miller in 1975 [14] and by Coleman in 1977 [10, 55]. Despite being formulated in a different way, Althorpe has proved that these two approaches give the same result [47]. Later work has also obtained further formulations, all of which are equivalent [18, 23, 56, 57].

Coleman's approach is derived with the aim of describing tunnelling in quantum field theory. It is based on the so-called 'Im F' premise, which assumes that the low-temperature rate of decay of a metastable state, k, is related to the system's free energy, F, by $k \approx -(2/\hbar)\operatorname{Im} F$ [55, 58]. This formula is not an exact definition of the rate. In fact, it must

9

be modified above a certain crossover temperature and its application at intermediate temperatures is understood simply as an approximate interpolation between known low- and high-temperature limits [59]. The imaginary part of the free energy can only be defined at the semiclassical limit using an analytic continuation of a divergent integral [46, 60]. It is difficult to see how to see how this concept can apply rigorously to chemical reactions, which are described by Hermitian Hamiltonians and have therefore purely real free energies [35, 61].

Chemical reaction rates can, however, be rigorously defined using scattering boundary conditions [62]. This was the basis of Miller's semiclassical instanton formulation [14]. Here we follow a similar approach, and by taking asymptotic approximations to the exact quantum expression we present a unified derivation of instanton rate theory from first principles, and thus show how the various formulations are related.

A. First-principles derivation

The flux correlation function provides a rigorous definition for the exact rate constant, k, of a chemical reaction [62]. This gives

$$kZ_\mathrm{r} = \frac{1}{2} \int_{-\infty}^{\infty} \mathrm{Tr}\left[\hat{F}\mathrm{e}^{-\hat{H}(\tau_0+\mathrm{i}t)/\hbar}\hat{F}\mathrm{e}^{-\hat{H}(\tau_1-\mathrm{i}t)/\hbar}\right]\mathrm{d}t, \tag{18}$$

where Z_r is the reactant partition function and $\tau_0 + \tau_1 = \beta\hbar$. The choice of splitting this factor into τ_0 and τ_1 does not affect the rate in principle, although this flexibility is important to obtain the correct instanton formulation. The flux through a dividing surface $\sigma(x) = 0$ is

$$\hat{F} = \frac{1}{2m}\left(\delta[\sigma(\hat{x})]\hat{p}_\sigma + \hat{p}_\sigma^\dagger\delta[\sigma(\hat{x})]\right), \tag{19}$$

where $\hat{p}_\sigma = \frac{\partial\sigma}{\partial\hat{x}}\cdot\hat{p}$ is the momentum normal to the dividing surface. Although the exact rate is independent of the choice of dividing surface, for simplicity we will assume that it has been placed in the barrier region. Reactants are defined by $\sigma(x) < 0$ and products by $\sigma(x) > 0$.

By taking semiclassical approximations for the time integral and the quantum trace in Eq. (18), the instanton expression for the rate constant can be found. This derivation of instanton theory presented here does not rely on the 'Im F' premise and thus has similarities with Miller's original formulation [14] and the derivation given in Ref. [18]. The new

10

derivation presented here is more elegant as it directly approximates the thermal rate rather than first obtaining the microcanonical rate and thermalizing.

Expanding the trace in position states, we rewrite the exact rate as

$$kZ_{\mathrm{r}} = \frac{1}{8m^2} \iiint_{-\infty}^{\infty} A(x', x'', t)\, \delta[\sigma(x')]\delta[\sigma(x'')]\, \mathrm{d}x'\mathrm{d}x''\mathrm{d}t, \tag{20}$$

where we have defined the function

$$\begin{aligned}
A(x', x'', t) &= \langle x'|\hat{p}_\sigma\, \mathrm{e}^{-\hat{H}(\tau_0+\mathrm{i}t)/\hbar}|x''\rangle \langle x''|\hat{p}_\sigma\, \mathrm{e}^{-\hat{H}(\tau_1-\mathrm{i}t)/\hbar}|x'\rangle \\
&\quad + \langle x'|\hat{p}_\sigma\, \mathrm{e}^{-\hat{H}(\tau_0+\mathrm{i}t)/\hbar}\, \hat{p}_\sigma^\dagger|x''\rangle \langle x''|\mathrm{e}^{-\hat{H}(\tau_1-\mathrm{i}t)/\hbar}|x'\rangle \\
&\quad + \langle x'|\mathrm{e}^{-\hat{H}(\tau_0+\mathrm{i}t)/\hbar}|x''\rangle \langle x''|\hat{p}_\sigma\, \mathrm{e}^{-\hat{H}(\tau_1-\mathrm{i}t)/\hbar}\, \hat{p}_\sigma^\dagger|x'\rangle \\
&\quad + \langle x'|\mathrm{e}^{-\hat{H}(\tau_0+\mathrm{i}t)/\hbar}\, \hat{p}_\sigma^\dagger|x''\rangle \langle x''|\mathrm{e}^{-\hat{H}(\tau_1-\mathrm{i}t)/\hbar}\, \hat{p}_\sigma^\dagger|x'\rangle .
\end{aligned} \tag{21}$$

Using $\hat{p}\,|x\rangle = \frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x}\,|x\rangle$ and $\langle x|\,\hat{p} = -\frac{\hbar}{\mathrm{i}}\frac{\partial}{\partial x}\,\langle x|$, each term in Eq. (21) can be written as products of the kernels and their derivatives. We can use the semiclassical approximation to write each kernel in terms of a sum of all the imaginary-time trajectories travelling from x' to x''. As shown in Fig. 1, there are three trajectories which dominate, notated by $\mu \in \{0, -, +\}$. The first is direct, and the others bounce exactly once. Further trajectories also exist which bounce at least two times, but as these are longer, the action is larger and their contribution to the rate exponentially suppressed. They are thus ignored in the following.

Therefore the semiclassical approximation to Eq. (21) gives

$$A(x', x'', t) \sim \sum_{\mu\nu} A_{\mu\nu}(x', x'', t), \tag{22}$$

where the sum over μ and ν is over the three different trajectories $\{0, -, +\}$. The individual terms are given by

$$\begin{aligned}
A_{\mu\nu}(x', x'', t) = \frac{\hbar^2}{\mathrm{i}^2} \Bigg[&\frac{\partial K_\mu(x', x'', \tau_0+\mathrm{i}t)}{\partial x'_\sigma} \frac{\partial K_\nu(x'', x', \tau_1-\mathrm{i}t)}{\partial x''_\sigma} \\
&- \frac{\partial^2 K_\mu(x', x'', \tau_0+\mathrm{i}t)}{\partial x'_\sigma \partial x''_\sigma} K_\nu(x'', x', \tau_1-\mathrm{i}t) \\
&- K_\mu(x', x'', \tau_0+\mathrm{i}t)\frac{\partial^2 K_\nu(x'', x', \tau_1-\mathrm{i}t)}{\partial x'_\sigma \partial x''_\sigma} \\
&+ \frac{\partial K_\mu(x', x'', \tau_0+\mathrm{i}t)}{\partial x''_\sigma} \frac{\partial K_\nu(x'', x', \tau_1-\mathrm{i}t)}{\partial x'_\sigma} \Bigg],
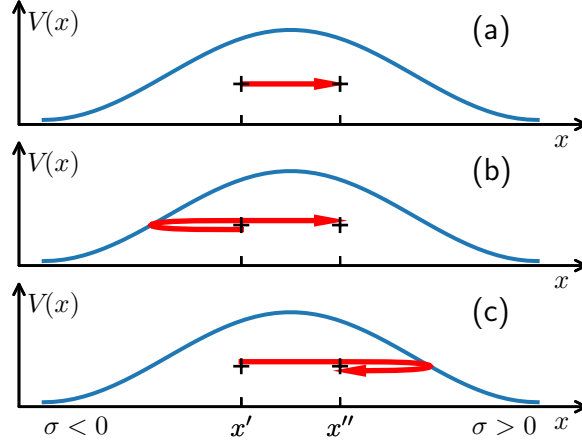\end{aligned} \tag{23}$$

FIG. 1. Three possible imaginary-time trajectories starting at x' and ending at x'': (a) $\mu = 0$, (b) $\mu = -$, (c) $\mu = +$. In the derivation, it is assumed that $\sigma(x) > 0$ on the right and $\sigma(x) < 0$ on the left.

where $\frac{\partial}{\partial x_\sigma} = \frac{\partial \sigma}{\partial x} \cdot \frac{\partial}{\partial x}$.

As expected from the relationship between the momentum operator and the derivative with respect to positions, within the semiclassical approximation, derivatives of the kernel return the kernel multiplied by a momentum term. For example, $\hbar \frac{\partial K_\mu(x', x'', t)}{\partial x'_\sigma} = -\frac{\partial S_\mu}{\partial x'_\sigma} K_\mu(x', x'', t) = s'|p'_\sigma| K_\mu(x', x'', t)$, where s' determines the sign, equal to $+1$ or -1 depending on whether the trajectory is moving to the right or left. The same relations hold for derivatives with respect to x''_σ. Using Eq. (7) and Fig. 1, it can be determined that for $\mu = \pm$, $s' = s'' = \pm 1$ but for $\mu = 0$, $s' = +1$ and $s'' = -1$.

By noting that many terms cancel, it can be shown that $A_{00} = A_{0\pm} = A_{\pm 0} = A_{\pm\pm} = 0$ and the only terms which are nonzero are

$$A_{\pm\mp}(x', x'', t) = 4|p'_\sigma||p''_\sigma|K_\pm(x', x'', \tau_0 + \mathrm{i}t)K_\mp(x'', x', \tau_1 - \mathrm{i}t). \tag{24}$$

Although the functions A_{+-} and A_{-+} are not identical (as they have stationary points in different places), their integrals over time are equal. Therefore we could compute the semiclassical rate in terms only of one of them as

$$kZ_\mathrm{r} \sim \frac{1}{8m^2} \iiint_{-\infty}^{\infty} 2A_{+-}(x', x'', t)\, \delta[\sigma(x')]\delta[\sigma(x'')]\, \mathrm{d}x'\mathrm{d}x''\mathrm{d}t \tag{25}$$

$$= \iiint_{-\infty}^{\infty} \frac{|p'_\sigma||p''_\sigma|}{m^2} \sqrt{\frac{C_+}{(2\pi\hbar)^f}} \sqrt{\frac{C_-}{(2\pi\hbar)^f}}\, \mathrm{e}^{-S/\hbar}\, \delta[\sigma(x')]\delta[\sigma(x'')]\, \mathrm{d}x'\mathrm{d}x''\mathrm{d}t, \tag{26}$$

12

where $S \equiv S_+(x', x'', \tau_+ + \mathrm{i}t) + S_-(x'', x', \tau_- - \mathrm{i}t)$ is the sum of the action along the specified trajectories and we have renamed the imaginary time variables $\tau_+ \equiv \tau_0 \equiv \tau$ and $\tau_- \equiv \tau_1 \equiv \beta\hbar - \tau$.

When performing the integral over time, it is possible to choose the value of τ to make S stationary such that the integrand is dominated by the region around $t = 0$. Note that in general it is not possible to choose value for τ such that the stationary point is at $t = 0$ for both A_{+-} and A_{-+} simultaneously. Exceptions to this rule are if the barrier is symmetric, when $\tau_0 = \tau_1 = \beta\hbar/2$ or in the classical, high-temperature, limit when τ_0 and τ_1 both tend to 0.

The quantum instanton approach [39] is derived from steepest-descent integrals of the flux correlation function directly [40]. As we have just shown, this is only a good approximation for a symmetric barrier, and for this reason, the quantum instanton approach, like a number of other quantum transition-state theories [63, 64], can give poor approximations for the rate of tunnelling through an asymmetric barrier [65]. In Sec. III C, we will discuss how quantum transition-state theory can be formulated to avoid this problem based on an understanding of instanton theory.

In order to treat multidimensional systems, we consider a transformation of coordinates from x to $(q, \boldsymbol{Q})$. The coordinate q is parallel to the trajectory, such that $|\dot{q}| = |\dot{x}|$, and chosen such that $q = 0$ at the dividing surface. The remaining $f - 1$ coordinates form an orthonormal set, $\boldsymbol{Q}$. The rate formula is seen to be independent of the dividing surface using $|p| = m|\dot{q}|$, $|p_\sigma| = \left|\frac{\partial\sigma}{\partial q}\right| |p|$, and $\delta[\sigma(x)] = \left|\frac{\partial\sigma}{\partial q}\right|^{-1} \delta(q)$. The integrals over q coordinates can then be performed easily due to the delta functions giving

$$kZ_{\mathrm{r}} \sim \iiint_{-\infty}^{\infty} |\dot{q}'\dot{q}''| \sqrt{\frac{C_+ C_-}{(2\pi\hbar)^{2f}}} \, \mathrm{e}^{-S/\hbar} \, \mathrm{d}\boldsymbol{Q}' \mathrm{d}\boldsymbol{Q}'' \mathrm{d}t, \tag{27}$$

where $q' = q'' = 0$.

Equation (27) provides us with an expression for the rate in terms of an integral which can be simply evaluated by the method of steepest descent. Depending on the order in which the integrals over coordinates and time are performed, we can obtain the different expressions for the rate which have previously appeared in the literature. However, all are formally equivalent.

For consistency, the reactant partition function, Z_{r}, should also be computed using the steepest-descent approximation, similarly to Eq. (11). As the rate constant is then defined

as a ratio between two steepest-descent approximations, at least some of the error can be expected to cancel which will slightly improve the prediction.

The first expression we consider is obtained by integrating over all remaining variables simultaneously:

$$kZ_{\mathrm{r}} \sim (2\pi\hbar)^{-\frac{1}{2}} \sqrt{\frac{\Sigma_- \Sigma_+}{-\Sigma}}\, \mathrm{e}^{-S/\hbar}, \tag{28}$$

where we have defined

$$\Sigma_{\pm} = \dot{q}^2 C_{\pm} = \begin{vmatrix} \frac{\partial^2 S_{\pm}}{\partial \boldsymbol{Q}' \partial \boldsymbol{Q}''} & \frac{\partial^2 S_{\pm}}{\partial \boldsymbol{Q}' \partial \tau_{\pm}} \\ \frac{\partial^2 S_{\pm}}{\partial \tau_{\pm} \partial \boldsymbol{Q}''} & \frac{\partial^2 S_{\pm}}{\partial \tau_{\pm}^2} \end{vmatrix} \tag{29}$$

$$\Sigma = \begin{vmatrix} \frac{\partial^2 S}{\partial \boldsymbol{Q}' \partial \boldsymbol{Q}'} & \frac{\partial^2 S}{\partial \boldsymbol{Q}' \partial \boldsymbol{Q}''} & \frac{\partial^2 S}{\partial \boldsymbol{Q}' \partial \tau} \\ \frac{\partial^2 S}{\partial \boldsymbol{Q}'' \partial \boldsymbol{Q}'} & \frac{\partial^2 S}{\partial \boldsymbol{Q}'' \partial \boldsymbol{Q}''} & \frac{\partial^2 S}{\partial \boldsymbol{Q}'' \partial \tau} \\ \frac{\partial^2 S}{\partial \tau \partial \boldsymbol{Q}'} & \frac{\partial^2 S}{\partial \tau \partial \boldsymbol{Q}''} & \frac{\partial^2 S}{\partial \tau^2} \end{vmatrix}, \tag{30}$$

and $\boldsymbol{Q}'$, $\boldsymbol{Q}''$ and τ are chosen according to the steepest-descent prescription $\frac{\partial S}{\partial \boldsymbol{Q}'} = \frac{\partial S}{\partial \boldsymbol{Q}''} = \boldsymbol{0}$ and $\frac{\partial S}{\partial \tau} = 0$.

This formulation of instanton theory is written in terms of two trajectories, one bouncing on the left and one on the right. It is similar to the expression presented in Sec. VI for the rate of an electron-transfer process [66]. However, other formulations of instanton theory in the literature have used a single periodic orbit. To show how these can be derived, a few more steps are necessary.

This time, again starting from Eq. (27), the integral is performed over t and $\boldsymbol{Q}''$ simultaneously to give

$$kZ_{\mathrm{r}} \sim \int |\dot{q}' \dot{q}''| \sqrt{\frac{C_+ C_-}{(2\pi\hbar)^f}} \begin{vmatrix} \mathrm{i}^2 \frac{\partial^2 S}{\partial \tau^2} & \mathrm{i}\frac{\partial^2 S}{\partial \tau \partial \boldsymbol{Q}''} \\ \mathrm{i}\frac{\partial^2 S}{\partial \boldsymbol{Q}'' \partial \tau} & \frac{\partial^2 S}{\partial \boldsymbol{Q}'' \partial \boldsymbol{Q}''} \end{vmatrix}^{-\frac{1}{2}} \mathrm{e}^{-S/\hbar}\, \mathrm{d}\boldsymbol{Q}', \tag{31}$$

with $\boldsymbol{Q}''$ and τ given by the steepest-descent conditions. The two trajectories can be combined into one which bounces both on the left and the right and which we will call the 'instanton' trajectory. The chain rule states that $\frac{\partial S}{\partial \tau} = \frac{\partial S}{\partial q}\dot{q}$ and thus the prefactors are combined as in Eq. (8) to give

$$C_{\mathrm{inst}} = C_- C_+ |\dot{q}''|^2 \begin{vmatrix} \frac{\partial^2 S}{\partial \tau^2} & \frac{\partial^2 S}{\partial \tau \partial \boldsymbol{Q}''} \\ \frac{\partial^2 S}{\partial \boldsymbol{Q}'' \partial \tau} & \frac{\partial^2 S}{\partial \boldsymbol{Q}'' \partial \boldsymbol{Q}''} \end{vmatrix}^{-1}. \tag{32}$$

Then the rate can be written in terms of the instanton trajectory as

$$kZ_r \sim \int \mathrm{Im}\, K_{\mathrm{inst}}(x, x, \beta\hbar)|\dot{q}|\mathrm{d}\boldsymbol{Q}, \tag{33}$$

where the variable $x' = (q', \boldsymbol{Q}')$ has been renamed $x = (q, \boldsymbol{Q})$ and

$$\mathrm{Im}\, K_{\mathrm{inst}}(x, x, \beta\hbar) = \sqrt{\frac{-C_{\mathrm{inst}}}{(2\pi\hbar)^f}}\, \mathrm{e}^{-S_{\mathrm{inst}}/\hbar}. \tag{34}$$

Note that $C_\pm > 0$ and $\frac{\partial^2 S}{\partial \tau^2} < 0$ so that $C_{\mathrm{inst}} < 0$. For this reason $K_{\mathrm{inst}}(x, x, \beta\hbar)$ is pure imaginary, but combines with the factor of i from Eq. (31) such that what appears in Eq. (33) is purely real. This also implies that the instanton is not a minimum of the action, but a first-order saddle point.

It is important to note that here only the instanton trajectory is included in the definition for the semiclassical kernel and there is no sum over other possible trajectories. Other trajectories also exist which are minima, including ones which bounce only once but these must not be allowed to contribute as they would dominate the integrand. This problem is illustrated in Ref. [67]. Our derivation above shows why only the instanton trajectory should be included—because the others cancel in Eq. (23) to give zero. For this reason no version of Eq. (33) has been found without using the semiclassical approximation, and hence it has not been possible to derive an exact imaginary-time path-integral Monte Carlo method for computing the rate.

Completing the final integral by steepest descent gives

$$kZ_r \sim (2\pi\hbar)^{-\frac{1}{2}}|\dot{q}||C_{\mathrm{inst}}|^{\frac{1}{2}} \left|\frac{\partial^2 S}{\partial \boldsymbol{Q}\partial \boldsymbol{Q}}\right|^{-\frac{1}{2}} \mathrm{e}^{-S/\hbar}, \tag{35}$$

where $S \equiv S(x', x'', \beta\hbar)$ and $x' = x'' = x$ such that $\frac{\partial^2 S}{\partial \boldsymbol{Q}\partial \boldsymbol{Q}} = \frac{\partial^2 S}{\partial \boldsymbol{Q}'\partial \boldsymbol{Q}'} + \frac{\partial^2 S}{\partial \boldsymbol{Q}'\partial \boldsymbol{Q}''} + \frac{\partial^2 S}{\partial \boldsymbol{Q}''\partial \boldsymbol{Q}'} + \frac{\partial^2 S}{\partial \boldsymbol{Q}''\partial \boldsymbol{Q}''}$. The steepest-descent approximation is made around a value of $\boldsymbol{Q}$ for which the action is stationary. This forces the ends of the instanton trajectory to have a continuous momentum and thus the instanton is a periodic orbit.

This result can be transformed into the instanton theory given by Miller [14] using a number of Legendre transformations (see Appendix A):

$$\left|\frac{\partial^2 S}{\partial \boldsymbol{Q}\partial \boldsymbol{Q}}\right| = \left|\frac{\partial^2 W}{\partial \boldsymbol{Q}\partial \boldsymbol{Q}}\right| \frac{\mathrm{d}^2 W}{\mathrm{d}E^2} \left(\frac{\partial^2 W}{\partial E^2}\right)^{-1}, \tag{36}$$

where

$$\frac{\mathrm{d}^2 W}{\mathrm{d}E^2} = \frac{\partial^2 W}{\partial E^2} - \frac{\partial^2 W}{\partial E \partial x}\left(\frac{\partial^2 W}{\partial x \partial x}\right)^{-1}\frac{\partial^2 W}{\partial x \partial E}, \tag{37}$$

$\frac{\partial}{\partial x} = \frac{\partial}{\partial x'} + \frac{\partial}{\partial x''}$, and

$$C_{\mathrm{inst}} = \left|-\frac{\partial^2 S}{\partial x' \partial x''}\right| = \left(\frac{\partial^2 W}{\partial E^2}\right)^{-1} D_{\mathrm{inst}}, \tag{38}$$

where [68]

$$D_{\mathrm{inst}} = (-1)^{f+1}\begin{vmatrix} \frac{\partial^2 W}{\partial x' \partial x''} & \frac{\partial^2 W}{\partial x' \partial E} \\ \frac{\partial^2 W}{\partial E \partial x''} & \frac{\partial^2 W}{\partial E^2} \end{vmatrix} = |\dot{q}|^{-2}\left|-\frac{\partial^2 W}{\partial \boldsymbol{Q}' \partial \boldsymbol{Q}''}\right|. \tag{39}$$

Therefore Eq. (35) can be transformed to Miller's expression [14],

$$kZ_{\mathrm{r}} \sim (2\pi\hbar^2)^{-\frac{1}{2}} Z^{\ddagger}\left(-\frac{\mathrm{d}E}{\mathrm{d}\beta}\right)^{\frac{1}{2}}\mathrm{e}^{-S/\hbar}, \tag{40}$$

where $-\hbar\left(\frac{\mathrm{d}E}{\mathrm{d}\beta}\right)^{-1} = \frac{\mathrm{d}^2 W}{\mathrm{d}E^2}$ and

$$Z^{\ddagger} = \left|-\frac{\partial^2 W}{\partial \boldsymbol{Q}' \partial \boldsymbol{Q}''}\right|^{\frac{1}{2}}\left|\frac{\partial^2 W}{\partial \boldsymbol{Q} \partial \boldsymbol{Q}}\right|^{-\frac{1}{2}} \tag{41}$$

$$= \prod_{j=1}^{f-1}[2\sinh(u_j/2)]^{-1}. \tag{42}$$

The final step uses a proof given by Gutzwiller [68] to write $Z^{\ddagger}$ in terms of u_j, the nonzero stability parameters [14].

Equations (35) and (40) provide two expressions for the rate according to the instanton approximation. They can be evaluated using the numerical methods described in Sec. IV. However, neither is the most commonly-used formulation which we will introduce in Sec. III C.

B. Connection to the 'Im F' premise

The approach used by Coleman [10] is based on the 'Im F' premise [58], which assumes that the rate can be written as

$$kZ_{\mathrm{r}} \approx \frac{2}{\beta\hbar}\,\mathrm{Im}\,Z, \tag{43}$$

16

where the total partition function is

$$Z = \int K(x, x, \beta\hbar) \, \mathrm{d}x. \tag{44}$$

A steepest-descent evaluation of this integral would be dominated by minima of the action functional, which as already explained do not correspond to instantons. Within the 'Im F' premise, the imaginary part of the partition function is defined using analytic continuation around the saddle point, which is the instanton trajectory.

This definition is valid only below the crossover temperature, and takes a slightly different form at higher temperatures to ensure that it matches with the classical rate [59]. There has been some discussion as to whether the 'Im F' premise is valid for describing finite-temperature chemical reactions [14, 18, 47, 57]. However despite the difference in the derivation, it can be shown that this formulation is exactly equivalent to Miller's rate expression, Eq. (40) [18, 47], and hence to all the expressions given above, which also do not make use of the 'Im F' premise.

The analytic continuation gives

$$\mathrm{Im}\, Z \sim \frac{1}{2} \int \mathrm{Im}\, K_{\mathrm{inst}}(x, x, \beta\hbar) \, \mathrm{d}x, \tag{45}$$

where the factor of half comes from analytically continuing the function only over half the complex plane [10]. As before, only the instanton pathway is considered, and not other trajectories such as those which bounce only once which would ruin the result [67]. The final integral cannot be performed directly by steepest descent because of the existence of a zero-frequency mode which describes cyclic permutation of the periodic trajectory. Instead we change first from x to $(q, \boldsymbol{Q})$ coordinates and then from q to imaginary-time before integrating this variable over the period of the trajectory:[1]

$$kZ_{\mathrm{r}} \approx \frac{1}{\beta\hbar} \int \int_0^{\beta\hbar} \mathrm{Im}\, K_{\mathrm{inst}}(x, x, \beta\hbar) \, |\dot{q}| \, \mathrm{d}\tau \, \mathrm{d}\boldsymbol{Q}, \tag{46}$$

which as the integrand is independent of τ, gives a rate expression exactly equivalent to that from Eq. (33).

We have shown that the 'Im F' premise gives the same formula as the first-principles instanton derivations given in Sec. III A. Therefore all the instanton expressions are equivalent including that of Miller, of 'Im F' and the new expressions presented in Sec. III A. This

---

[1] Because the instanton folds back on itself, one could say that the integral should only be over half the period. However, as the instanton can travel around its orbit in two different directions, we account for this degeneracy here.

shows that the 'Im F' premise is valid in this case, but does not mean that it is valid in general for other problems.

C. Ring-polymer discretization

The most common discretization scheme used for representing closed path integrals is that of the ring polymer [38, 48–54]. Here the path is represented by N beads, x_i, each of which is an image of the f-dimensional system. An alternative scheme (not used here) is to discretize the path using Fourier coefficients [41, 69, 70].

The ring-polymer potential is given by

$$U_N(\boldsymbol{x}) = \sum_{i=1}^{N} \frac{m}{2\beta_N^2 \hbar^2} |x_i - x_{i-1}|^2 + \sum_{i=1}^{N} V(x_i), \tag{47}$$

where $\beta_N = \beta/N$. This is defined similarly to Eq. (12) but for a cyclic path with N variable bead positions, $\boldsymbol{x} = \{x_1, \ldots, x_N\}$, and no fixed end points but with the equivalence $x_0 \equiv x_N$. The action of a ring-polymer path is given by $\hbar\beta_N U_N(\boldsymbol{x})$.

The standard ring-polymer instanton rate expression is most easily derived directly from the 'Im F' premise [21], although it should be noted that we have shown this to be equivalent to the first-principles derivations. In this representation, the partition function is

$$Z = \Lambda_N^{-Nf} \int e^{-\beta_N U_N(\boldsymbol{x})} \, \mathrm{d}\boldsymbol{x}, \tag{48}$$

where $\Lambda_N = \sqrt{\frac{2\pi\beta_N \hbar^2}{m}}$.

According to the 'Im F' premise, the rate is proportional to the imaginary part of the partition function which can be obtained by performing a steepest-descent integration around the saddle point of $U_N(\boldsymbol{x})$, notated $\tilde{\boldsymbol{x}}$. This is the definition of the instanton trajectory for which $\beta_N U_N(\tilde{\boldsymbol{x}}) = S/\hbar$. A second-order normal-mode expansion of the ring-polymer potential about this point is

$$U_N(\boldsymbol{x}) \approx U_N(\tilde{\boldsymbol{x}}) + \sum_k \tfrac{1}{2} m \eta_k^2 c_k^2. \tag{49}$$

The frequencies, η_k, are defined as the square root of the eigenvalues of the mass-weighted ring-polymer Hessian matrix, $\boldsymbol{\nabla}^2 U_N(\tilde{\boldsymbol{x}})/m$ and the displacements of the corresponding eigenmodes are c_k.

18

There will be one negative eigenvalue and the steepest-descent procedure is modified for this mode: the absolute value is taken and the integral multiplied by a factor of a half [10]. This gives

$$\text{Im}\, Z = \tfrac{1}{2}\Lambda_N^{-1}\prod_k{}'\left|\frac{1}{\beta_N\hbar\eta_k}\right|e^{-\beta_N U_N(\tilde{\boldsymbol{x}})}\int \mathrm{d}c_0. \tag{50}$$

Not only does $\boldsymbol{\nabla}^2 U_N/m$ have one negative eigenvalue, but it also has one, η_0, which is zero, corresponding to the permutational mode, c_0, which points along $\dot{\tilde{\boldsymbol{x}}}$. The steepest-descent integral is not valid for this mode and therefore (as indicated by the prime) the product is taken over all modes except the zero-frequency mode corresponding to permutation of the instanton.

The zero-frequency mode has the effect $\mathrm{d}x_i = \sqrt{\frac{\beta_N^2\hbar^2}{B_N}}\dot{\tilde{x}}_i\,\mathrm{d}c_0$ where $\dot{\tilde{x}}_i \approx (\tilde{x}_i - \tilde{x}_{i-1})/\beta_N\hbar$ and

$$B_N = \sum_i |\tilde{x}_i - \tilde{x}_{i-1}|^2. \tag{51}$$

It is normalized such that $\sum_i\left(\frac{\partial x_i}{\partial c_0}\right)^2 = 1$. Comparison with $\mathrm{d}x_i = \dot{\tilde{x}}_i\,\mathrm{d}t$ shows that the integral over this mode gives $\int \mathrm{d}c_0 = \int_0^{\beta\hbar}\sqrt{\frac{B_N}{\beta_N^2\hbar^2}}\,\mathrm{d}t = N\sqrt{B_N}$. Finally, using Eq. (43), we obtain

$$kZ_\mathrm{r} = \frac{\Lambda_N^{-1}}{\beta_N\hbar}\sqrt{B_N}\prod_k{}'\left|\frac{1}{\beta_N\hbar\eta_k}\right|e^{-\beta_N U_N(\tilde{\boldsymbol{x}})}. \tag{52}$$

This is the most commonly used ring-polymer instanton expression [21, 23, 24]. However, as shown in Sec. III A, this is not the only valid expression and equivalent formulae exist. It turns out that there can be an advantage in using the alternative formulae as unlike this 'Im F' expression, they do not require building and diagonalizing a large $Nf \times Nf$ matrix.

In order to predict the crossover temperature, we consider a ring polymer collapsed at the transition state to find out if it is a first-order saddle point. It has eigenvalues of $\eta_k^2 = \frac{4}{\beta_N^2\hbar^2}\sin^2 k\pi/N - \bar{\omega}_\mathrm{b}^2$ for $k \in \{-N/2,\ldots,N/2-1\}$, where $\bar{\omega}_\mathrm{b}$ is the imaginary part of the barrier frequency. For the lowest modes where $k \ll N$, this is given approximately by $\eta_k^2 \approx \frac{4k^2\pi^2}{\beta^2\hbar^2} - \bar{\omega}_\mathrm{b}^2$ and we see that there is only one negative eigenvalue if $\beta\hbar\bar{\omega}_\mathrm{b} < 2\pi$ but more than one if $\beta\hbar\bar{\omega}_\mathrm{b} > 2\pi$. Therefore the crossover temperature is defined as $\beta_\mathrm{c} = 2\pi/\hbar\bar{\omega}_\mathrm{b}$. Below the crossover temperature, there must be a first-order saddle point elsewhere which describes the instanton. This is the onset of deep tunnelling when the instanton cannot

19

remain collapsed at the barrier top and must stretch across the barrier. This analysis was based on a parabolic approximation to the shape of the barrier, and different definitions for the crossover temperature will be more appropriate for barriers that deviate significantly from this shape [71]. It is also possible for instantons to exist well above the crossover temperature [72], as in general the ring-polymer surface can have more than one first-order saddle point simultaneously.

The instanton approach is strongly connected to other path-integral rate theories, and in particular to ring-polymer molecular dynamics (RPMD) [37, 38]. The main advantage of this method over instanton theory is that it is also applicable to studying tunnelling effects in a liquid environment. The instanton pathway represents an optimized path-integral configuration describing the reaction whereas most path-integral rate theories are based on an equation of the form

$$kZ_{\mathrm{r}} = A \int \mathrm{e}^{-\beta_N U_N(\boldsymbol{x})} \, \delta[\sigma_N(\boldsymbol{x})] \, \mathrm{d}\boldsymbol{x}, \tag{53}$$

with various definitions of the prefactor, A, and constraint $\sigma_N(\boldsymbol{x})$. No steepest-descent approximation is used for the integrals over $\boldsymbol{x}$, such that all paths obeying the constraint $\sigma_N(\boldsymbol{x}) = 0$ contribute. The dominant path is the one which minimizes $U_N(\boldsymbol{x})$ whilst simultaneously obeying the constraint. If this dominant path is the same as the instanton, $\tilde{\boldsymbol{x}}$, then the rate would be proportional to $\mathrm{e}^{-\beta_N U_N(\tilde{\boldsymbol{x}})} = \mathrm{e}^{-S/\hbar}$ in the same way as with the instanton expression.

Centroid-based path-integral methods [35, 36] use a constraint of the form $\sigma_N(\boldsymbol{x}) = \frac{1}{N} \sum_{i=1}^{N} (x_i - x^\ddagger)$. This approximation often performs well for symmetric barriers, but can fail spectacularly in asymmetric systems [63]. This is best understood by considering the dominant path-integral configuration under the centroid constraint. For symmetric systems, it is equal to the instanton, but this is not true for asymmetric systems [21], even after variationally optimizing $x^\ddagger$ [73]. Centroid-based methods can therefore make an error in a part of the formula which is exponentiated and hence exhibit large errors in the rate. In contrast, the semiclassical instanton method gives good approximations for rates in the deep-tunnelling regime for both symmetric and asymmetric systems.

Ring-polymer transition-state theory (RPTST) is defined such that the constraint on the ring polymer ensures that the instanton remains the optimum configuration [21, 67, 74]. Therefore instanton theory is strongly related to a harmonic approximation to RPTST.

The prefactors of the two approaches are similar but not exactly the same, such that it is common for RPTST and instanton predictions to be within a factor of 2 of each other [21]. RPTST can also be derived from first principles without invoking the steepest-descent approximation by considering the instantaneous quantum flux through a generalized dividing surface [67, 75].

RPTST can be thought of as a TST approximation to RPMD itself and thus for a well-chosen dividing surface, the RPTST and RPMD give approximately the same result. This shows the connection between RPMD and instanton theory which confirms that RPMD is a valid theory for studying deep-tunnelling in molecular systems. RPMD has a distinct advantage over RPTST however, as its results are independent of the choice of dividing surface. This is achieved using real-time ring-polymer trajectories, which have two roles, one is to correct for a suboptimal choice of dividing surface, $\sigma_N(\boldsymbol{x})$, and the second is to provide information on classical recrossing dynamics. This makes RPMD a practical method for the study of tunnelling in liquid environments, for which the optimal dividing surface is generally unknown. However, for gas-phase reactions or for reactions on surfaces, the instanton approach is valid and is typically much more efficient than other path-integral rate theories, allowing larger systems to be studied with accurate *ab initio* potential-energy surfaces.

IV. NUMERICAL IMPLEMENTATION

In this section we discuss efficient implementation strategies for the instanton rate theories presented above. Instanton pathways are imaginary-time classical trajectories and therefore solutions of Newton's equations of motion on an upside-down potential. Historically the pathways were located computationally using an approach known as 'shooting', which propagates the equations of motion to generate trajectories and searches for initial conditions leading to the required periodic orbit solution [15, 16, 76]. This is a simple procedure for one-dimensional systems, but quickly becomes practically impossible to perform as the number of degrees of freedom increases. This is because when searching for an unstable orbit, the final position at the end of the trajectory changes exponentially fast as the initial condition is varied.

The disadvantages of the shooting method are clearly outlined in *Numerical Recipes* [77]

where the alternative method of 'relaxation' is recommend in its stead [69]. In this approach, the path is discretized into a set of points with the periodic boundary condition implicitly built in. All points are then optimized simultaneously by an optimization algorithm until the required solution is found [21, 23, 24]. It turns out that when searching for instantons, the relaxation method is far superior to the shooting method.

A.   Optimization of the instanton

We have introduced the ring-polymer discretization for the instanton pathway above. However, there is a simple trick which can be used to reduce the computational cost of this representation [23, 24]. The instanton trajectory is known to fold back on itself in order to complete its periodic orbit. It is therefore only necessary to locate one half of the trajectory and thus half of the path, $x(t)$ for $t \in [0, \beta\hbar/2]$, is discretized into beads x_i for $i \in \{1, \ldots, N/2\}$, where we are assuming that N is even. The half-ring-polymer potential is given by

$$U_{N/2}(\boldsymbol{x}) = \sum_{i=1}^{N/2-1} \frac{m}{2\beta_N^2\hbar^2} |x_{i+1} - x_i|^2 + \sum_{i=1}^{N/2} V(x_i). \tag{54}$$

The solution which describes the instanton trajectory is denoted $\tilde{\boldsymbol{x}}$. Once the positions of the half-ring-polymer beads are found, it is simple to find the ring-polymer instanton configuration by symmetry. The action of this instanton is given by $S/\hbar = 2\beta_N U_{N/2}(\tilde{\boldsymbol{x}})$. Note that for the fluctuations in Eq. (52), one must construct and diagonalize the full ring-polymer Hessian as described in Sec. IV B.

According to the variation principle, the first variation of S is zero for a classical trajectory. Also it was shown in Sec. III that the instanton has a single negative eigenvalue. Thus the instanton can be obtained by searching for first-order saddle points of $U_{N/2}(\boldsymbol{x})$. For locating saddle points, standard transition-state optimization algorithms such as quasi-Newton approaches [78, 79] can be applied. Kästner and coworkers have thoroughly tested quasi-Newton algorithms and found them to be very efficient at optimizing instantons in a small number of iterations [24, 80].

However, the problem of optimizing instantons is not exactly equivalent to searching for saddle points in general. The first derivatives of the half-ring-polymer potential have the

simple form

$$\frac{\partial U_{N/2}(\boldsymbol{x})}{\partial x_1} = \frac{m}{\beta_N^2 \hbar^2}(x_1 - x_2) + \nabla V(x_1) \tag{55a}$$

$$\frac{\partial U_{N/2}(\boldsymbol{x})}{\partial x_i} = \frac{m}{\beta_N^2 \hbar^2}(-x_{i+1} + 2x_i - x_{i-1}) + \nabla V(x_i) \tag{55b}$$

$$\frac{\partial U_{N/2}(\boldsymbol{x})}{\partial x_{N/2}} = \frac{m}{\beta_N^2 \hbar^2}(x_{N/2} - x_{N/2-1}) + \nabla V(x_{N/2}), \tag{55c}$$

for $i \in \{2, \ldots, N/2 - 1\}$. This means that, unlike for the general saddle-point search, the Hessian matrix, $\boldsymbol{H} = \boldsymbol{\nabla}^2 U_{N/2}(\boldsymbol{x})$, is banded. Also, extra flexibility is offered in the choice of an algorithm used for converging the number of beads, N.

It is especially important to take the banded nature of the Hessian into account for systems with a large number of atoms, where the Hessian can be reduced in size from a square $Nf/2 \times Nf/2$ matrix to $f + 1$ bands of length $Nf/2$. For a system with 100 atoms and $N = 128$, this reduces the memory requirement from 3 GB to 50 MB, thus ensuring that the computation is feasible on a standard computer. Many of the more sophisticated quasi-Newton saddle-point optimization algorithms use a procedure known as 'eigenvector following' which requires knowledge of the eigenvectors of the Hessian. Even if the Hessian is banded, the matrix of eigenvectors is not, and thus the memory requirement becomes once more too large. However, the simplest quasi-Newton method based on the Newton-Raphson iteration does not require eigenvectors and can thus take full advantage of the banded nature of the Hessian.

The Newton-Raphson optimization algorithm takes steps of $\boldsymbol{h}$ defined by solving the linear equations $(\boldsymbol{H} + \varepsilon\boldsymbol{I})\boldsymbol{h} = -\boldsymbol{G}$, where $\boldsymbol{H} = \boldsymbol{\nabla}^2 U_{N/2}(\boldsymbol{x})$ is the Hessian and $\boldsymbol{G} = \boldsymbol{\nabla} U_{N/2}(\boldsymbol{x})$ the gradient. A shift of $\varepsilon \approx 0.1\,E_{\mathrm{h}}/\text{Å}^2$ is added to the diagonal of $\boldsymbol{H}$ to avoid problems with zero modes [81]. A maximum step length is defined and the step is scaled down if it is larger than this. The Hessian need not be very accurate at this stage of the calculation. Therefore, rather than recomputing the Hessian after each step, the Powell update formula [82] is used to update the nonzero bands. The algorithm is iterated until the norm of $\boldsymbol{G}$ is less than a given tolerance.

There is of course a well-known problem with the Newton-Raphson algorithm which is that it can optimize to any stationary point and not only to first-order saddle points. This means that the initial guess must be made quite close to the saddle point, which explains why it is typically not used for standard transition-state searches. However, when optimizing

23

instantons, we are often in the situation where we are able to make a good initial guess. This is because we can first obtain a good approximation using a small number of beads and interpolate this result to give an initial guess with a larger number of beads. In this way, the next optimization often converges in fewer than 10 steps and nearly always to the correct stationary point.

An automatic procedure for obtaining the first initial guess is given by a straight-line interpolation around the transition state in the direction of the imaginary normal mode [24]:

$$x_i = x^{\ddagger} + \Delta \cos\left(\frac{2\pi i}{N}\right) z, \tag{56}$$

where z is the f-dimensional eigenvector corresponding to the imaginary mode at the transition state, $x^{\ddagger}$. The scalar variable $\Delta \approx 0.1\,\text{Å}$ is chosen to be smaller than the bond lengths in the system, but large enough to represent the typical spread of the instanton trajectory. It can be varied by trial and error until the initial guess is good enough for the optimization to converge.

The calculations need to be converged with respect to the number of ring-polymer beads N. This is done by computing the rate with increasingly large values of N, typically doubling each time, until the rate converges. Rather than starting each optimization from Eq. (56), the optimized configuration from the previous calculation is used where available. Once an N-bead path has been optimized, a set of f one-dimensional spline interpolations can be used to generate an initial guess for a ring polymer with a larger number of beads. A similar approach can be used to generate approximate Hessians of each bead.

Note that when using the Newton-Raphson algorithm, the beads may collapse at the barrier top, which below the crossover temperature is a higher-order saddle point [21]. Therefore for the first instanton optimization with a small number of beads, for which no good guess is available, eigenvector-following algorithms are used. Even if the full matrix of eigenvectors is built for a system of 100 atoms, this can be easily stored in computer memory with $N = 16$. Many such algorithms exist; one which we have used most often and generally found stable is that of Ref. [79]. In summary, eigenvector-following is used when stability is required, but Newton-Raphson for efficiency.

A typical work flow is summarized as follows:

1. Optimize the minimum or minima and perform normal-mode analysis.

2. Optimize the transition state and perform normal-mode analysis.

3. Use Eq. (56) to give an initial guess for an $N = 16$ configuration.

4. Compute Hessians of each bead. (This can be done approximately.)

5. Optimize $N = 16$ instanton using eigenvector following, computing gradients at each iteration, and using the Powell update formula on the ring-polymer Hessian.

6. Double number of beads by interpolation of positions and Hessians to give new guess configuration.

7. Optimize with Newton-Raphson, computing gradients at each iteration, and using the Powell update formula on the ring-polymer Hessian.

8. Recompute Hessians for each bead. (This must be done accurately.)

9. Evaluate the fluctuation terms and hence the rate, k.

10. Return to step 7 until the rate converges.

If rates are required at more than one temperature, it is recommended to start with those closest to the crossover temperature and cool sequentially, again using initial guesses from the previous optimizations. Following this procedure, the instanton can be located even for polyatomic reactions of many atoms. A few specific examples are discussed in Sec. V.

B.   Fluctuation factors

Once the instanton has been located by the optimization algorithms described in Sec. IV A, the exponential part of the rate expression is known. What remains to be computed is the prefactor, which is based on fluctuations around the instanton. We have given a number of expressions for the instanton rate in Sec. III, which differ in the way the prefactor is evaluated.

The most commonly-used approach for computing the instanton rate is based on Eq. (52). In order to evaluate this expression, the full ring-polymer Hessian is built and diagonalized to find the frequencies, η_k. One of the frequencies should be imaginary, as the instanton is a saddle point, and one should be zero corresponding to the permutational mode. Depending

25

on the symmetries of the Hamiltonian, there may also be other zero frequencies corresponding to translational and rotational modes. These should also be removed from the product and are treated instead using classical translational and rotational partition functions.

The resulting instanton rate theory can be written in the following form:

$$k = \frac{1}{2\pi\beta\hbar} \frac{Z^{\text{trans}} Z^{\text{rot}} Z^{\text{vib}}}{Z_{\text{r}}^{\text{trans}} Z_{\text{r}}^{\text{rot}} Z_{\text{r}}^{\text{vib}}} \, e^{-S/\hbar}, \tag{57}$$

where the reactant partition function, Z_{r}, has been defined in terms of translational, rotational and vibrational contributions from the separated reactant molecules.

Note that this expression is similar to that of Eyring TST if $S/\hbar$ were replaced by $\beta V^{\ddagger}$, where $V^{\ddagger}$ is the barrier height. In this way, it is possible to define a finite-N version of Eyring TST which tends to the usual formula in the $N \to \infty$ limit. This is useful for comparison with the finite-N instanton expression to obtain a tunnelling factor with some error cancellation. In this case, one should use a collapsed ring-polymer configuration for both the transition-state geometry as well as the minimum or minima. However, for an instanton rate calculation, one uses collapsed ring polymers for the reactants only and the instanton configuration for the partition functions in the numerator.

For a gas-phase reaction, there will be f_0 modes corresponding to translations and rotations of the instanton or transition state. There can be 5 or 6 of these depending on whether the transition state is linear or nonlinear. Their partition functions are defined using the usual classical expressions as if all the beads in the ring polymer made up a $Nf/3$-atom 'super-molecule' at reciprocal temperature β_N. They are normalized such that for a collapsed ring polymer, the factors of N cancel to give the usual expression.

The translational partition function for a molecule of total mass M is

$$Z^{\text{trans}} = N^{-3} \left(\frac{NM}{2\pi\beta_N\hbar^2} \right)^{\frac{3}{2}}. \tag{58}$$

For a gas-phase bimolecular reaction, there are two reactant molecules and $Z_{\text{r}}^{\text{trans}}$ is the product of two expressions of this form.

The moments-of-inertia tensor is given by

$$I = \sum_{i=1}^{N} \sum_{a=1}^{f/3} m_a \left[(\vec{r}_{i,a} \cdot \vec{r}_{i,a}) \mathbb{I} - \vec{r}_{i,a} \times \vec{r}_{i,a} \right], \tag{59}$$

where $\vec{r}_{i,a}$ is the displacement from the centre of mass of the ring polymer of the ath atom (with mass m_a) of the ith replica, and $\mathbb{I}$ the 3×3 identity matrix. The rotational contribution

to the ring-polymer partition function is then computed using the appropriate formula:[2]

$$Z^{\text{rot}} = 1 \qquad\qquad\qquad \text{for a single atom,} \tag{60a}$$

$$Z^{\text{rot}} = N^{-2} \frac{2I_{\text{B}}}{\beta_N \hbar^2} \qquad\qquad \text{for linear configurations,} \tag{60b}$$

$$Z^{\text{rot}} = N^{-3} \sqrt{\frac{8\pi \det I}{\beta_N^3 \hbar^6}} \qquad\qquad \text{for nonlinear configurations,} \tag{60c}$$

where I_{B} is the value of the nonzero eigenvalues of I (for linear configurations).

The vibrational contribution of a collapsed ring polymer to the partition function is given by [46]

$$Z^{\text{vib}} = {\prod_k}' \left[2 \sinh \tfrac{1}{2} \beta \hbar \tilde{\omega}_k \right]^{-1}, \tag{61}$$

where $\tilde{\omega}_k$ are the corrected frequencies for finite values of N, as defined in Appendix B. The prime denotes that only modes with nonzero real frequencies are included in the product.

The vibrational contribution to the instanton partition function is however given by a different expression, obtained from Eq. (52),

$$Z^{\text{vib}}_{\text{inst}} = N^{1+f_0} \sqrt{\frac{2\pi m B_N}{\beta_N \hbar^2}} {\prod_k}' \left| \frac{1}{\beta_N \hbar \eta_k} \right|. \tag{62}$$

This is the most commonly-used ring-polymer instanton approach. However, we have shown that there are many other equivalent formulations which may have numerical advantages over this approach.

Few numerical implementations of Eq. (40) have been performed [15, 76] because the stability parameters for the unstable instanton orbit have been shown exhibit poor numerical behaviour [83, 84]. This problem is related to the problem of obtaining the monodromy matrix needed for other semiclassical dynamics methods, for which further approximations are often necessary to reduce the numerical errors [85].

We next consider the expression Eq. (35). Once translations and rotations are correctly treated, this can also be expressed as in Eq. (57) using

$$Z^{\text{vib}}_{\text{inst}} = (2\pi\hbar)^{-\frac{1}{2}} |\dot{q}| \left| \det{}' \frac{\partial^2 S}{\partial x' \partial x''} \right|^{\frac{1}{2}} \left(\det{}' \frac{\partial^2 S}{\partial \boldsymbol{Q} \partial \boldsymbol{Q}} \right)^{-\frac{1}{2}}, \tag{63}$$

---

[2] Note that the symmetry number does not appear here but instead manifests itself in a number of identical transition states.

where we use the notation

$$\det{}'A = \prod{}' \operatorname{eig}(A) \tag{64}$$

to mean that the product is taken over all eigenvalues except those corresponding to translational and rotational modes. These are instead treated by their partition functions.

In order to identify the translational and rotational modes from the determinant, we can look for eigenvalues with particular numerical values. For a free particle, $S_{\text{free}}(x', x'', \tau) = m(x''-x')^2/2\tau$ and therefore $\frac{\partial^2 S_{\text{free}}}{\partial x' \partial x''} = -m/\tau$. This means that f_0 of the eigenvalues of $\frac{\partial^2 S}{\partial x' \partial x''}$ will be equal to $-m/\beta\hbar$.

For the case of the determinant of $\frac{\partial^2 S}{\partial \boldsymbol{Q} \partial \boldsymbol{Q}}$, the eigenvalues corresponding to translations and rotations are seen to be 0 by a similar argument. One can obtain this matrix by rotating $\frac{\partial^2 S}{\partial x \partial x}$ to the $(q, \boldsymbol{Q})$ coordinate system and projecting out the q direction. However, a simpler approach is to take the eigenvalues of $\frac{\partial^2 S}{\partial x \partial x}$ directly which are equivalent except that there is also a further 0 eigenvalue corresponding to the q direction which should be removed.

Formulae for computing derivatives of the action with respect to the end points are given explicitly in the appendix of Ref. [86]. Using these we can obtain expressions for $\frac{\partial^2 S}{\partial x' \partial x''}$ as well as $\frac{\partial^2 S}{\partial x' \partial x'} = \frac{\partial^2 S}{\partial x'' \partial x''}$. Note that this algorithm requires solving a set of linear equations with a banded matrix. We cannot use a Cholesky factorization as the matrix is not positive definite, but algorithms do exist for solving this using an LU factorization. The speed of the trajectory at the end point can be obtained using $|\dot{q}| = \frac{1}{m}\left|\frac{\partial S}{\partial x'}\right|$. Alternatively, it is given by $|\dot{q}| = \sqrt{2[V(x') - E]/m}$ where $E = \frac{\partial S}{\partial \tau}$ is the energy of the instanton pathway.

An alternative method for obtaining the determinant of $\frac{\partial^2 S}{\partial x' \partial x''}$ is to use the multidimensional Gelfand-Yaglom formula [46]:

$$\Psi_1 = 2\mathrm{I} + \epsilon^2 \nabla^2 V(x_1) \tag{65a}$$

$$\Psi_2 = \left(2\mathrm{I} + \epsilon^2 \nabla^2 V(x_2)\right)\Psi_1 - \mathrm{I} \tag{65b}$$

$$\Psi_k = \left(2\mathrm{I} + \epsilon^2 \nabla^2 V(x_k)\right)\Psi_{k-1} - \Psi_{k-2} \qquad \text{for } k \in \{3, \ldots, N-1\}, \tag{65c}$$

where $\epsilon = \beta_N \hbar$ and I is the $f \times f$ identity matrix. The required eigenvalues can be obtained using $\operatorname{eig}\left(\frac{\partial^2 S}{\partial x' \partial x''}\right) = (\epsilon \operatorname{eig} \Psi_{N-1})^{-1}$. A further option is to evaluate the determinant of the banded matrix $\boldsymbol{J}$ which is related to the product of eigenvalues by Eq. (17).

The choice of which method to use depends on the system under study. The standard ring-polymer instanton formula, Eq. (52), has been well tested, and does not seem to exhibit

28

numerical problems, even when an *ab initio* PES is used. The alternative methods will only become necessary for particularly large systems, where one would like to avoid storing and diagonalizing the ring-polymer Hessian matrix.

V. APPLICATIONS

The ring-polymer instanton approach has been applied to a number of gas-phase reactions using both fitted surfaces and on-the-fly calculations for the PES [23, 26, 70, 87–94]. Some of these calculations have been used to benchmark the instanton approach against more accurate theories, but others have given new insights into physical chemistry. In this section we present example calculations on simple systems for which exact benchmark results are available for comparison.

However, the ring-polymer instanton method is by no means limited to such small gas-phase reactions and calculations have also been performed in condensed-phase systems with a larger number of atoms. These include studies of chemical reactions on ice surfaces [32], as well as hydrogen diffusion on surfaces [30, 33] or in crystals [93]. The effect of tunnelling on proton transfer in an enzyme-catalysed reaction [34] has been computed, and gave kinetic isotope effects consistent with experimental observations. In order to study this liquid system, for which instanton theory is not formally valid, an approximation was made to freeze the environment at the transition-state geometry. In a similar way, there are also approaches used to treat amorphous surfaces with a set of different possible configurations [31].

In certain cases, further work is needed to extend the instanton approach to treat a new type of tunnelling dynamics, which appears for diffusion over broad-topped barriers [71]. Extensions are also needed for treating low-temperature bimolecular reactions with a pre-reactive complex [91, 92]. However, for the majority of the studies carried out, the largest source of error in the predicted rate is expected to come from the electronic-structure method, rather than the instanton method itself. For the largest systems, and in particular in the condensed phase, density-functional theory (DFT) is often the only practical method available. In this case, due to errors in the barrier heights, quantitative rates cannot always be trusted, and the results obtained should be considered qualitative for describing which tunnelling mechanism dominates. Even in these cases, the calculation of kinetic isotope ef-

fects benefits from much error cancellation and can give better comparison with experiment.

A. Eckart barrier

The standard test case of the Eckart barrier has been used to compare results from many different approximate rate theories. We have previously derived an analytic result for the instanton rate on the symmetric Eckart barrier [19], and shown it to be an asymptotic limit of the exact rate. Here we present a similar derivation for the more general asymmetric barrier:

$$V(x) = \frac{A}{1 + e^{-2x/a}} + \frac{B}{\cosh^2(x/a)}. \tag{66}$$

The exact reaction probability can be written [1, 95]

$$P(E) = \frac{\cosh \alpha(\sqrt{\eta} + \sqrt{\eta - \gamma}) - \cosh \alpha(\sqrt{\eta} - \sqrt{\eta - \gamma})}{\cosh \alpha(\sqrt{\eta} + \sqrt{\eta - \gamma}) + \cosh \sqrt{4\alpha^2 - \pi^2}}, \tag{67}$$

where $\eta = E/B$ is the reduced energy, $\alpha = \pi \sqrt{2ma^2 B}/\hbar$ is a measure of the width of the barrier and $\gamma = A/B$ is the measure of asymmetry.

In the limit $\alpha \to \infty$,

$$P(E) \sim \frac{e^{\alpha\sqrt{\eta}} e^{\alpha\sqrt{\eta - \gamma}}}{e^{\alpha(\sqrt{\eta} + \sqrt{\eta - \gamma})} + e^{2\alpha}} \tag{68}$$

$$= \frac{1}{1 + e^{W/\hbar}} \tag{69}$$

$$\sim e^{-W/\hbar}, \tag{70}$$

where

$$W(E) = \hbar\alpha(2 - \sqrt{\eta} - \sqrt{\eta - \gamma}). \tag{71}$$

This can be identified as the same function as the the abbreviated action [1, 95], defined as the following integral between the turning points $x_<$ and $x_>$:

$$W(E) = 2\int_{x_<}^{x_>} \sqrt{2m(V(x) - E)}\,\mathrm{d}x. \tag{72}$$

The exact thermal rate is defined by

$$kZ_{\mathrm{r}} = \frac{1}{2\pi\hbar} \int_0^\infty P(E)\,e^{-\beta E}\,\mathrm{d}E. \tag{73}$$

30

TABLE I. Tunnelling factors from three different methods: the numerically exact solution, Eq. (73), the instanton approximation, Eq. (74), and Wigner's second-order approximation, Eq. (75). The system is an asymmetric Eckart barrier with $A = -18/\pi$, $B = 13.5/\pi$, $a = 8/\sqrt{3\pi}$ and $\hbar = m = 1$, which has a crossover temperature of 2π.

β	exact	instanton	Wigner
8	26	28	4
10	252	232	5
12	4068	3689	7

The asymptotic approximation to the rate is obtained using Eq. (70) and a steepest-descent integration over energy to give

$$kZ_\mathrm{r} \sim (2\pi\hbar)^{-\frac{1}{2}} \left(\frac{\partial^2 W}{\partial E^2} \right)^{-\frac{1}{2}} \mathrm{e}^{-W/\hbar - \beta E}, \tag{74}$$

where in the second line, E is defined such that $\frac{\partial W}{\partial E} = 0$. This is also the expression for the instanton approximation to the thermal rate, as can be seen by comparison with Eq. (40) for the case of a one-dimensional system [19].

In Table I, we show how the instanton predictions for the asymmetric Eckart barrier compare with the exact results. In each case, the tunnelling factor is defined as the ratio of the rate with classical transition-state theory, $k_\mathrm{TST} Z_\mathrm{r} = \frac{1}{2\pi\beta\hbar} \mathrm{e}^{-\beta V^\ddagger}$, where $V^\ddagger$ is the barrier height. For comparison, the second-order Wigner approximation [20],

$$k_\mathrm{Wigner}/k_\mathrm{TST} = 1 + (\beta\hbar\bar{\omega}_\mathrm{b})^2/24, \tag{75}$$

is also given. This approximation is only valid for shallow tunnelling, and thus gives poor predictions for this low temperature regime. However, the instanton results are consistently of the correct order of magnitude.

B. H + H$_2$

The simplest thermally activated chemical reaction, H + H$_2$, exhibits a large deviation from the Arrhenius law and has large kinetic isotope effects due to both tunnelling and zero-point energy effects [87, 96, 97]. It therefore provides an excellent system for testing the instanton

TABLE II. Semiclassical instanton (SCI) and exact quantum-mechanical (QM) values of $H + H_2$ and $D + D_2$ rates in cm^3/s. The crossover temperatures are 345 K and 244 K for the reactions.

T / K	$H + H_2$		$D + D_2$	
	SCI	QM	SCI	QM
300	5.2(-16)	3.24(-16)	-	4.65(-17)
250	4.3(-17)	3.67(-17)	-	3.53(-18)
200	2.2(-18)	2.01(-18)	1.2(-19)	8.92(-20)
150	6.8(-20)		5.1(-22)	

approximation, in particular because these rates can be compared with exact benchmark results, obtained using the log-derivative method [98–100].

In Table II, we compare rates from the instanton approximation [101] with the exact results [40] for two isotopic variants of this reaction. The BKMP2 potential-energy surface [102] was used by all three methods. The instanton rates presented are converged with respect to the number of beads to two significant figures. It was necessary to use 512 beads at the lowest temperature studied. Note that the instanton rates have been multiplied by a factor of 2 to account for the fact that the reactant radical can attack either atom of the diatomic.

The instanton approach is seen to agree within about 20% except for the highest temperature studied in each case for which the rate is overestimated by a larger factor. It is a well-known effect which causes instanton theory to overestimate the rate near to the crossover temperature by about a factor of two [19]. This could be corrected by a number of proposed extensions to the approach [12, 56, 70, 103] including an explicit integration over microcanonical rates [19]. The related RPMD approach also avoids this problem.

C. $H + CH_4$

The ring-polymer instanton method has been implemented in the Molpro electronic-structure package [104, 105] which allows the rates to be computed on the fly, without the need for a fitted analytic potential-energy surface. This has been applied to the $H + CH_4$ reaction (see Fig. 2) using coupled-cluster methodology, showing how the efficiency of the instanton method allows the user to systematically converge the tunnelling rate with respect to the
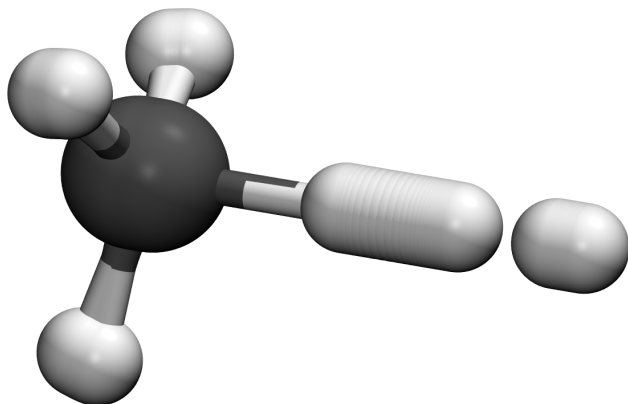
FIG. 2. Representation of the ring-polymer instanton describing proton tunnelling in the $H + CH_4$ reaction at 200 K. All atoms take part to some extent in the tunnelling process and become delocalized as they pass through the potential barrier.

TABLE III. Computed rates for $H + CH_4$ in cm^3/s. The *ab initio* semiclassical instanton (SCI) results are from Ref. 26, the SCI/CBE results from Ref. [101] and MCTDH/CBE results from Ref. 106.

T / K	SCI/*ab initio*	SCI/CBE	MCTDH/CBE
300	1.7(-19)	1.8(-19)	8.4(-20)
250	4.8(-21)	4.2(-21)	3.1(-21)
200	1.1(-22)	5.7(-23)	-
150	1.8(-24)	4.6(-25)	-

level of electronic-structure theory [26].

Testing RCCSD-F12a and RCCSD(T)-F12a with both cc-pVDZ and cc-pVTZ basis sets, barrier heights were found which varied from 63.21 kJ mol^{-1} with the most accurate combination to 67.51 kJ mol^{-1} for the least accurate. This variation in barrier height can cause more than a 10-fold difference in the predicted rate constant at 200 K, as is easily estimated from the Arrhenius equation. However, it is not just the barrier height which affects the quantum rate, but also the barrier shape. For instance, a thinner barrier is more conducive to tunnelling. It was also found that the tunnelling factor predicted by the lower level of theory was up to a factor of 4 larger than that predicted by the higher level. At 200 K, the best estimate of the tunnelling factor is about 700 and it is thus clear that the inclusion of quantum effects in low-temperature rate calculations cannot be ignored.

In Table III we compare the instanton rates computed on the fly with a high-quality RCCSD(T)-F12a/cc-pVTZ electronic structure calculations with rates on the CBE potential surface [107]. All instanton rates presented are converged with respect to the number of beads, for which it was necessary to use at least 256 for the lowest temperature. Also given for comparison are the rates found from exact quantum dynamics on this same surface [106] obtained using the multiconfigurational time-dependent Hartree method (MCTDH) [108].

First we compare the SCI results with the MCTDH approach on the CBE surface. As expected, the SCI result overestimates the rate at 300 K by about a factor of 2 as this is just below the crossover temperature of 327 K, but there is fairly good agreement ($\sim$35% error) at 250 K. This agreement would be expected to continue to lower temperatures but unfortunately no MCTDH results are available as the method becomes more computationally demanding in this regime.

A comparison between the *ab initio* instanton results and those on the fitted surface shows good agreement at high temperatures but the error increases as the temperature drops. Assuming that the *ab initio* PES is correct, this discrepancy can be assigned to errors in the fit of the CBE surface. A similar effect on the rate caused by these errors is seen for MCTDH calculations on using the more accurate WWM surface instead of CBE. These errors have a larger effect on the low-temperature instantons, which are more delocalized and sample a larger region around the barrier. In fact at 250 K, the error caused by the fit is similar in magnitude to that assigned to the instanton approximation.

It is not currently possible to perform exact quantum dynamics in full dimensionality using on-the-fly *ab initio* calculations for this system, although some progress has been made towards this goal [109]. Therefore in order to study such chemical reactions, it is necessary to make at least one approximation. This application shows that at low temperatures the accuracy of the *ab initio* SCI approach is similar to that of using exact quantum dynamics approaches with the CBE fitted PES. For systems for which an accurate fitted potential-energy surface is not available or for which exact quantum dynamics becomes too expensive, on-the-fly implementations of instanton theory are expected to give the most reliable estimates of the tunnelling rate.

VI. ELECTRON-TRANSFER RATES

Chemical reactions involving electron transfer (ET) occur in many different environments, from redox reactions to photosynthesis and the harvesting of light in solar cells [110]. In the simplest ET reactions, the charge is transferred from a donor to an acceptor; these entities can be anything from small solvated ions [111] to large protein complexes [112]. Thus there are at least two important electronic states involved in the reaction and typically the Born-Oppenheimer approximation breaks down, making it necessary to consider nonadiabatic dynamics in order to describe and predict the rate of the process.

The Hamiltonian describing an ET process can be represented in the diabatic representation as

$$\hat{H} = \hat{H}_0 \left|0\right\rangle \left\langle 0\right| + \hat{H}_1 \left|1\right\rangle \left\langle 1\right| + \Delta(\left|0\right\rangle \left\langle 1\right| + \left|1\right\rangle \left\langle 0\right|), \tag{76}$$

where $\left|0\right\rangle$ and $\left|1\right\rangle$ are the electronic states of the reactant and product which are coupled by Δ. Here we will assume that the Condon approximation holds, such that Δ is a constant. However, the instanton approach and its classical limit can be easily extended to describe a coordinate-dependent coupling, $\Delta(\hat{x})$.

The Hamiltonians $\hat{H}_0$ and $\hat{H}_1$ describe the nuclear degrees of freedom of each electronic state and are of the form

$$\hat{H}_n = \frac{|\hat{p}|^2}{2m} + V_n(\hat{x}), \qquad n \in \{0, 1\}, \tag{77}$$

where the functions $V_0(x)$ and $V_1(x)$ describe the reactant and product potential-energy surfaces.

In many cases, the rate can be considered to be in the golden-rule limit [113], that is where the coupling, Δ, between the electronic states is assumed to be weak and the ET itself is the bottleneck to the reaction. Fermi's golden rule thus provides a formula for the quantum rate constant of the process. It is obtained from perturbation theory with a lowest-order expansion in the coupling and gives a rate proportional to Δ^2. From the flux-flux correlation function [62], and expanding the trace in a coordinate-space representation defines the quantum rate

$$kZ_0 = \frac{\Delta^2}{\hbar^2} \iiint_{-\infty}^{\infty} K_0(x', x'', \beta\hbar - \tau + \mathrm{i}t) K_1(x'', x', \tau - \mathrm{i}t) \, \mathrm{d}x' \mathrm{d}x'' \mathrm{d}t, \tag{78}$$

where the quantum propagator K_n describes the dynamics for the Hamiltonian $\hat{H}_n$.

Nonetheless as the exact quantum dynamics are generally not known, approximations to Fermi's golden rule have to be made. Several theories have been proposed to tackle this kind of problem, most famously by Marcus [110, 114–117]. Its simple form and wide range of applications make Marcus theory a standard approach to treat ET. This theory is derived by applying a classical transition-state theory approximation to Fermi's golden rule and making a number of assumptions about the shapes of the free-energy curves involved.

Quantum nuclear effects are ignored by the standard Marcus theory. These effects allow for tunnelling of the nuclear coordinates and are expected to lead to a speed-up of the rate. One approach for including quantum effects into ET processes is to map the system onto a harmonic spin-boson model [118] and solve the resulting equations either using semiclassical approximations or numerically exactly [119–124]. This approach however cannot take account of anharmonicity, although certain generalized spin-boson systems can still be studied within these approaches [125] and anharmonic effects treated within a perturbative approach [126].

A method that promises to overcome these assumptions of Marcus theory is semiclassical instanton theory which has been extended to describe electron transfer in the nonadiabatic limit [66, 86]. A semiclassical approximation to Eq. (78) is found by replacing the quantum propagators by semiclassical propagators, and then by integrating over the end points, x' and x'', and over t by steepest descent. This gives the instanton expression for the ET rate,

$$kZ_0 \sim \sqrt{2\pi\hbar}\,\frac{\Delta^2}{\hbar^2}\sqrt{\frac{C_0 C_1}{-\Sigma}}\,\mathrm{e}^{-S/\hbar}, \tag{79}$$

where $S = S_0(x', x'', \beta\hbar - \tau) + S_1(x', x'', \tau)$ and

$$\Sigma = \begin{vmatrix} \frac{\partial^2 S}{\partial x' \partial x'} & \frac{\partial^2 S}{\partial x' \partial x''} & \frac{\partial^2 S}{\partial x' \partial \tau} \\ \frac{\partial^2 S}{\partial x'' \partial x'} & \frac{\partial^2 S}{\partial x'' \partial x''} & \frac{\partial^2 S}{\partial x'' \partial \tau} \\ \frac{\partial^2 S}{\partial \tau \partial x'} & \frac{\partial^2 S}{\partial \tau \partial x''} & \frac{\partial^2 S}{\partial \tau^2} \end{vmatrix}. \tag{80}$$

The standard model for electron-transfer reactions in the condensed phase is the spin-boson model [12, 118, 126]. The potentials of reactants and products are given by sets of

shifted harmonic oscillators,

$$V_0(x) = \sum_{j=1}^{f} \frac{1}{2} m\omega_j^2 (x_j + \xi_j)^2 \tag{81a}$$

$$V_1(x) = \sum_{j=1}^{f} \frac{1}{2} m\omega_j^2 (x_j - \xi_j)^2 - \epsilon, \tag{81b}$$

and the electronic coupling, Δ, is taken to be a constant.

The rate obtained by the instanton approach for the spin-boson model is [66]

$$k \sim \sqrt{2\pi\hbar} \frac{\Delta^2}{\hbar^2} \left(-\frac{\mathrm{d}^2 S}{\mathrm{d}\tau^2} \right)^{-\frac{1}{2}} \mathrm{e}^{-S/\hbar}, \tag{82}$$

where τ is chosen as the stationary point of the action function

$$S = -\epsilon\tau + \sum_{j=1}^{f} 2m\omega_j \xi_j^2 \left[\frac{1 - \cosh\omega_j\tau}{\tanh\beta\hbar\omega_j/2} + \sinh\omega_j\tau \right]. \tag{83}$$

For this system, the instanton rate is equivalent to results obtained by other semiclassical methods [12, 122, 127] and tends to standard Marcus theory in the high-temperature classical limit.

The ring-polymer instanton approach [86] approach is applicable in general to multi-dimensional anharmonic potential-energy surfaces. Only simple numerical algorithms including a saddle-point optimization are required to apply the method in practice and it is therefore computationally inexpensive. The implementation is similar to that described in Sec. IV, except that the instanton is defined as the saddle point of the discretized action, not just with respect to positions of the beads, but also for the value of τ. In the high-temperature limit, the instanton reduces to the classical rate, which is not necessarily equivalent to that of Marcus theory. Recent work applied both classical and semiclassical methods to an asymmetric system-bath model with anharmonic free-energy curves in order to explore the behaviour of the rates of various approaches with respect to both anharmonicity and tunnelling [128]. It was found that in these cases the rate predicted by the standard Marcus theory was in error by orders of magnitude but that the instanton approach was accurate to within 1%.

Applications of the approach are currently limited to model systems because of the inability of instanton theory to describe liquid systems, which are typically of particular interest

for electron-transfer reactions. Nonetheless, a thorough understanding of the instanton approach may lead to the development of novel nonadiabatic ring-polymer molecular dynamics methods which can be used to describe electron transfer in solution from atomistic simulations. The development of such methods is underway [129–134].

VII.   TUNNELLING SPLITTINGS

So far in this review, the instanton approach has been applied only to the calculation of rate processes. However, the rate constant is meaningful only if there is a decay into a continuum of product states [135]. The simplest case where a continuum of product states exists is that of a reactive scattering event as described in Sec. III. This is an example of *incoherent* tunnelling dynamics.

For finite bound-state systems, the energy-level structure is discrete and the concept of a rate constant formally breaks down. A typical example is given by a double-well system, for which it is known that the quantum dynamics of a wave function initially localized on a single well oscillates between one well and the other in a periodic motion. This *coherent* tunnelling dynamics cannot be described by an exponential decay, which is the role of the rate constant.

Note that we sometimes also discuss rates for intramolecular processes in large isolated molecules at finite temperature. In this case even though there is not a continuum of states, there are many accessible such that the behaviour, at least on experimental timescales, is similar to that of the decay process and we can assign a meaningful rate constant. However, as the temperature decreases towards zero, fewer states become accessible and the dynamics becomes oscillatory making the rate undefined.

There is however a well-defined quantity of interest for bound, finite systems with degenerate wells at zero temperature, which is the tunnelling splitting. This occurs when the degenerate vibrational states of each well interact to give a set of eigenstates closely spaced in energy. In the basis of localized vibrational states, the Hamiltonian can be written

$$\boldsymbol{H} = \begin{pmatrix} E_0 & -\hbar\Omega \\ -\hbar\Omega & E_0 \end{pmatrix}, \tag{84}$$

which has the eigenvalues $E_0 - \hbar\Omega$ and $E_0 + \hbar\Omega$ and thus the tunnelling splitting is given by $\Delta = 2\hbar\Omega$. A time-dependent analysis shows that if a particle is initially localized in one

well, its population oscillates like $\cos^2(\Omega t)$.

Instanton theory can also be used to compute tunnelling splittings in molecules and clusters within the semiclassical approximation. The derivation of the theory [10, 16] is different from that used to obtain the reaction rate. However, a similar ring-polymer version exists for evaluating the tunnelling splittings [22, 27] and the numerical implementation of the method is similar to that of rate theory.

The derivation starts from the definition of the partition function for the double-well system in the $\beta \to \infty$ limit, which is $Z = 2\mathrm{e}^{-\beta E_0} \cosh \beta \hbar \Omega$. In the absence of tunnelling the partition function would simply be $Z_0 = 2\mathrm{e}^{-\beta E_0}$. By computing instanton approximations to both Z and Z_0 and taking their ratio, an expression for Ω can be found. The Taylor series expression for the ratio is

$$\frac{Z}{Z_0} = \sum_{n=0,\text{even}}^{\infty} \frac{(\beta \hbar \Omega)^n}{n!}. \tag{85}$$

A. Instanton theory for tunnelling splittings

To derive the instanton theory, we require expressions for Z and for Z_0 within the semi-classical approximation. An expression for Z_0 can easily be obtained using Eq. (11). When contributions from tunnelling trajectories are also considered, the partition function includes further contributions to give

$$Z = Z_0 + Z_2 + Z_4 + \ldots, \tag{86}$$

where Z_n is the sum over all closed paths which tunnel through the barrier n times. Note that n must therefore be even for this double-well system.

In the following, we will assume that the coordinates of the minima of the left and right well are $x_<$ and $x_>$ and that they have potential energy 0. As will be shown, any kernel can be broken up into kernels describing either non-tunnelling paths, $K_0(\tau) \equiv K_0(x_<, x_<, \tau) \equiv K_0(x_>, x_>, \tau)$ or single barrier crossings, $K_1(\tau) \equiv K_1(x_<, x_>, \tau) \equiv K_1(x_>, x_<, \tau)$, known as 'kinks'. The former is given by

$$K_0(\tau) \sim \sqrt{\frac{C_0}{(2\pi\hbar)^f}}. \tag{87}$$

Next we consider the kernel for a single barrier crossing $K_1(\tau)$ in the $\tau \to \infty$ limit. There is an instanton pathway connecting the end points $x_<$ and $x_>$ which tunnels once through

the barrier. It has $E = 0$ and thus travels infinitesimally slowly near the end points but relatively quickly through the barrier; this in fact is the source of the name 'instanton'. However, the van-Vleck propagator Eq. (4) cannot be directly applied to compute $K_1(\tau)$ because the fluctuation factor C tends to infinity. This occurs because one of the eigenvalues of $\boldsymbol{J}$ tends to zero and corresponds to an eigenvector which is the permutational mode of the instanton. In this case, the permutation translates the time at which the instanton rushes through the barrier. In order to use the semiclassical formulae directly we split the trajectory into two pieces, using $\tau = \tau_a + \tau_b$, to give

$$K_1(x_<, x_>, \tau) = \int K_a(x_<, x, \tau_a) K_b(x, x_>, \tau_b)\, \mathrm{d}x. \tag{88}$$

The point at which the two trajectories meet is x. For now, we will force this point to be at a particular position in the barrier at a given imaginary time. We thus introduce a factor of $\delta(t(x) - \tau_a) = |\dot{q}|\delta(q)$ into the integrand, to obtain the modified kernel

$$K_1'(x_<, x_>, \tau) = \iint K_a(x_<, x, \tau_a)|\dot{q}|\delta(q) K_b(x, x_>, \tau_b)\, \mathrm{d}q\, \mathrm{d}Q \tag{89}$$

$$\sim K_a(x_<, x, \tau_a)|\dot{q}| K_b(x, x_>, \tau_b) \sqrt{\frac{(2\pi\hbar)^{f-1}}{\nabla_{\boldsymbol{Q}}^2 S}}, \tag{90}$$

where $\nabla_{\boldsymbol{Q}}^2 S = \frac{\partial^2 S_a}{\partial \boldsymbol{Q}'' \partial \boldsymbol{Q}''} + \frac{\partial^2 S_b}{\partial \boldsymbol{Q}' \partial \boldsymbol{Q}'}$. Note that this expression is independent of the choice of where the trajectory is split.

We now consider a trajectory for multiple barrier crossings. This can be broken into single kinks, e.g. $\beta\hbar = \tau_1 + \cdots + \tau_n$ with fixed centres at $t_1, \ldots, t_n$. A representation of such a trajectory is shown in Fig. 3. These are combined in pairs using the semiclassical integral (which is dominated by the region of x near to $x_>$):

$$\int K_1'(x_<, x, \tau_1) K_1'(x, x_<, \tau_2)\, \mathrm{d}x \sim K_1'(\tau_1)\Xi K_1'(\tau_2), \tag{91}$$

where $\Xi = (2\pi\hbar)^{f/2} \left| \frac{\partial^2 S(\tau_1)}{\partial x'' \partial x''} + \frac{\partial^2 S(\tau_2)}{\partial x' \partial x'} \right|^{-\frac{1}{2}}$.

For comparison, we similarly break up Z_0 into n pieces. The kinks are combined using the rule

$$\int K_0(x_<, x, \tau_1) K_0(x, x_<, \tau_2)\, \mathrm{d}x \sim K_0(\tau_1)\Xi K_0(\tau_2), \tag{92}$$

where Ξ is the same as in Eq. (91). This holds in the limit of $\tau \to \infty$ because then the time spent in the wells is much longer than the time spent during the barrier crossing.
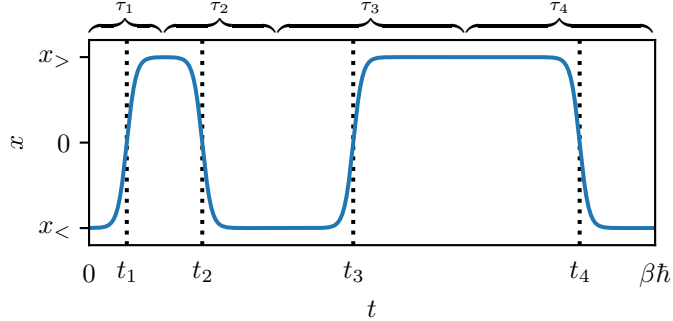
FIG. 3. A schematic of an $n = 4$ trajectory. This can be split into four single kinks, each of length τ_n and centred around t_n.

Thus, when we take the ratio of Z_n/Z_0, the Ξ factors will cancel leaving only ratios of kernels, given by

$$\theta = \lim_{\tau \to \infty} \frac{K'_1(\tau)}{K_0(\tau)} \tag{93}$$

$$= \sqrt{\frac{C_a C_b}{(2\pi\hbar)C_0|\nabla_{\boldsymbol{Q}}^2 S|}} \, |\dot{q}| \, e^{-S/\hbar}. \tag{94}$$

The Z_n partition function can then be found by integrating over the position of the centres:

$$\frac{Z_n}{Z_0} \sim \int_0^{\beta\hbar} dt_1 \int_{t_1}^{\beta\hbar} dt_2 \cdots \int_{t_{n-1}}^{\beta\hbar} dt_n \, \theta^n \tag{95}$$

$$= \frac{(\beta\hbar\theta)^n}{n!}. \tag{96}$$

By comparison with Eq. (85), we see that θ is the semiclassical approximation for the tunnelling matrix element which we require. This therefore provides the instanton formula for the tunnelling splitting, using $\Omega \sim \theta$.

B. Ring-polymer instanton theory for tunnelling splittings

The formula for the tunnelling splitting given in the previous section can be computed using the ring-polymer instanton method directly. However, an alternative derivation for a ring-polymer instanton method gives a different formula for θ, which we present below. Both formulations are in principle equivalent, but one may be more numerically stable than the other in practical applications.

41

Using path-integral discretization, Eq. (15), a non-tunnelling kernel is given by

$$K_0(x_<, x_<, \tau) \sim \left(\frac{m}{2\pi\tau_N\hbar}\right)^{f/2} (\det \boldsymbol{J_0})^{-\frac{1}{2}}, \tag{97}$$

where $\boldsymbol{J_0}$ is the matrix defined as in Eq. (16) with all beads located at $x_<$.

In order to compute a discretized path-integral version of $K_1(\tau)$, we consider the instanton path which tunnels through the barrier. This path, $\tilde{x}_i$, defines a matrix $\boldsymbol{J}$ which has a zero eigenvalue due to the permutational invariance of the instanton. Therefore the semiclassical van-Vleck formula must be modified to take account of this zero-mode.

We follow a similar procedure to that used in Sec. III C. The zero-mode eigenvector of $\boldsymbol{J}$ has elements $\mathrm{d}x_i = \sqrt{\frac{m\tau_N}{S_{\mathrm{kink}}}}\dot{\tilde{x}}_i \,\mathrm{d}c_0$, where

$$S_{\mathrm{kink}} = \int_0^\tau m|\dot{x}|^2\mathrm{d}t = \sum_i m\dot{\tilde{x}}_i^2\tau_N, \tag{98}$$

which is also equal to the action of the instanton as can be shown using the Legendre transformation described in Appendix A at $E = 0$. This eigenvector is normalized such that $\sum_i \left(\frac{\partial x_i}{\partial c_0}\right)^2 = 1$. By comparison with the relation $\mathrm{d}x_i = \dot{\tilde{x}}_i\mathrm{d}t$, we see that $\mathrm{d}c_0 = \sqrt{\frac{S_{\mathrm{kink}}}{m\tau_N}}\,\mathrm{d}t$.

By changing integration variables from c_0 to t, and fixing the value of t, we obtain the modified single-kink kernel:

$$K_1'(\tau) \sim \frac{1}{\tau_N}\left(\frac{m}{2\pi\tau_N\hbar}\right)^{f/2}\sqrt{\frac{S_{\mathrm{kink}}}{2\pi\hbar}}(\det{}'\boldsymbol{J})^{-\frac{1}{2}}\,\mathrm{e}^{-S_{\mathrm{kink}}/\hbar}, \tag{99}$$

where the prime indicates that the determinant is equal to the product over the nonzero eigenvalues only.

In the discretized version, the ratio of kernels is

$$\theta = \lim_{\tau\to\infty}\frac{K_1'(\tau)}{K_0(\tau)} \tag{100}$$

$$= \frac{1}{\Phi}\sqrt{\frac{S_{\mathrm{kink}}}{2\pi\hbar}}\,\mathrm{e}^{-S_{\mathrm{kink}}/\hbar}, \tag{101}$$

where

$$\Phi = \tau_N\left(\frac{\det{}'\boldsymbol{J}}{\det \boldsymbol{J_0}}\right)^{\frac{1}{2}}. \tag{102}$$

As before, the ratio of partition functions can be obtained by an integral over the centres of each kink. This provides a second method for obtaining θ which is the instanton approximation to Ω, which in turn defines the tunnelling splitting. It is this formulation, rather than Eq. (94), which has been tried and tested in previous studies [22, 27, 28, 136–139].

C. Extension to multi-well problems

A potential-energy surface exhibiting a double-well structure has tunnelling pathways passing from one well to the other and the tunnelling splitting for such a system is given by $\Delta = 2\hbar\Omega$. In order to make a study of the tunnelling splitting in more general molecular clusters it is necessary to extend the method to treat tunnelling between any number of wells [27]. The result is that the splitting pattern can be calculated as the eigenvalues of a Hückel-type matrix, each element of which corresponds to an instanton calculated as before between two wells.

The zero-point energy E_0 for a molecular system with G degenerate wells, between which tunnelling takes place, splits into a set of G levels, $\{E_\nu\}$, which cannot in general be described by a single tunnelling splitting Δ. Following this definition, the ratio of partition functions is

$$\lim_{\beta\to\infty} \frac{Z}{Z_0} = \lim_{\beta\to\infty} \frac{1}{G} \sum_{\nu=1}^{G} e^{-\beta(E_\nu - E_0)}. \tag{103}$$

As in the previous sections, we intend to take the steepest-descent approximation of the ring-polymer formulation of the partition functions. The low-temperature limit of the partition function of the tunnelling system, Z, evaluated by the method of steepest descent can be divided into contributions from each ring-polymer minimum. This set of minima, or periodic orbits, start and end in the same well but may take any number of kinks and pass through any number of the other wells on the way as they have an infinite amount of imaginary time available.

For enumerating these ring polymers, it can be useful to make an analogy with some standard results of graph theory [140], where the potential wells are depicted by vertices and kinks by edges. The sequence of kinks and wells visited by the ring-polymer minima are described by closed walks taken along the edges between the vertices. If we define $Z_{n,\nu}$ as the contribution of the closed paths of n kinks which start and finish at well ν, we may write

$$\lim_{\beta\to\infty} Z = \sum_{\nu=1}^{G} \sum_{n=0}^{\infty} Z_{n,\nu}. \tag{104}$$

We wish to count the total number of different sequences of n kinks, that is, in the parlance of graph-theory, the number of distinct walks which start and finish at a particular vertex, This is done using the *adjacency matrix* $\boldsymbol{A}$, which is defined such that $A_{\lambda\mu}$ is equal

to the number of edges between vertices λ and μ. Therefore the elements of the adjacency matrix, $A_{\lambda\mu}$, give the number of single-kink pathways which directly connect the wells λ and μ.[3] In the same way, it is seen that $\sum_{\kappa=1}^{G} A_{\lambda\kappa}A_{\kappa\mu}$ provides the number of sequences of two kinks which connect the wells λ and μ via any intermediate κ. And indeed $(\boldsymbol{A}^n)_{\lambda\mu}$ is the number of possible sequences connecting these wells using n kinks [141]. We require only pathways that start and end in the same well for the partition function and therefore use the formula $(\boldsymbol{A}^n)_{\nu\nu}$ to count them.

We can then factorize the contribution to $Z_{n,\nu}$ made by each n-kink orbit into a product of n weights $\theta_{\lambda\mu}$ each associated with a different kink in the sequence. The $\theta_{\lambda\mu}$ are defined analogously to θ in either Eq. (94) or Eq. (101). To obtain the total weight associated with all the n-kink orbits which contribute to $Z_{n,\nu}$, we define the *tunnelling matrix* $\boldsymbol{W}$ by

$$W_{\lambda\mu} = -A_{\lambda\mu}\theta_{\lambda\mu} \qquad \text{(no sum implied)}. \tag{105}$$

It then follows that

$$\frac{Z_{n,\nu}}{Z_{0,\nu}} = \frac{1}{n!}[(-\beta\hbar\boldsymbol{W})^n]_{\nu\nu}, \tag{106}$$

since the adjacency matrix element $A_{\lambda\mu}$ in $W_{\lambda\mu}$ ensures that the quantity $[(-\beta\hbar\boldsymbol{W})^n]_{\nu\nu}$ is the sum over the weights $\theta_{\lambda\mu}$ associated with all orbits contributing to $Z_{n,\nu}$. We thus obtain

$$\lim_{\beta\to\infty}\frac{Z}{Z_0} = \frac{1}{G}\sum_{n=0}^{\infty}\frac{1}{n!}\sum_{\nu=1}^{G}[(-\beta\hbar\boldsymbol{W})^n]_{\nu\nu}$$

$$= \frac{1}{G}\operatorname{tr}\left[\mathrm{e}^{-\beta\hbar\boldsymbol{W}}\right]. \tag{107}$$

Comparison with Eq. (103) reveals that the eigenvalues of $\hbar\boldsymbol{W}$ give the desired approximations to the energy splittings $E_\nu - E_0$.

D.   Implementation

Numerical implementation of this method is similar to that of ring-polymer instanton rate theory, except for the main difference that the instanton pathway is not a saddle point, but

---

[3] When $A_{\lambda\mu} > 1$, we shall assume that these kinks can be mapped onto one another by symmetry operations. The treatment could be generalized for systems where more than one kink joined a pair of wells along pathways unrelated by symmetry. However, it will probably be common that one kink will have a low action and will dominate, in which case the other higher-action kinks can be neglected.

the minimum of $S_N(\boldsymbol{x})$, Eq. (12). This function is typically optimized using the L-BFGS algorithm [142] to define the instanton and the action S_{kink}.

The Hessian, $\nabla^2 V(x_i)$, at the geometry of each bead was computed numerically and these were used to construct the mass-weighted Hessian or fluctuation matrix $\boldsymbol{J}$ of the linear-polymer. A banded-matrix linear algebra routine was used to compute the eigenvalues of which only one should be zero. In this case, the translational and rotational modes do not have zero eigenvalues, because the fixed end points break the symmetry.

Instanton calculations of ground-state tunnelling splittings are defined in the zero-temperature limit. In practice, the temperature is lowered until convergence is reached. Note that as the temperature is lowered, more ring-polymer beads are required, although simple approaches also exist for accelerating convergence of Φ with respect to the number of beads [137].

To reduce the number of beads needed, one could use larger time steps at the ends of the pathways [80]. A related approach which enforces the beads to be evenly spaced in position rather than time is a reformulation of the action in Hamilton-Jacobi form [143]. The difficulty here arises because one needs a penalty function to force the beads to be evenly spaced. However, the length of the path, and hence the spacing of the beads, is not known before optimization. Typically the penalty function is defined in terms of the average spacing of the current iteration and therefore couples all the beads together (rather than just nearest neighbours) which decreases the efficiency of the minimization algorithm. A recent algorithm seems to offer a promising alternative [138, 139]. It is based on a nudged-elastic-band approach [144] and provides a simple optimization algorithm which keeps the beads evenly spaced. It is also possible to parametrize the pathway using Fourier coefficients, allowing the points at which the potentials and gradients are evaluated more flexibly [145, 146].

E. Double-well model

In order to test the accuracy of the approach, numerical calculations [22] were performed on the one-dimensional double-well system described by the potential-energy surface

$$V(x) = V_0 \left(\frac{x^2}{x_0^2} - 1 \right)^2 , \qquad (108)$$

TABLE IV. Comparison of the instanton tunnelling splittings Δ (in reduced units) with the exact quantum splittings for different barrier heights V_0 in the model of Eq. (108) (with $x_0 = 5\sqrt{V_0}$, $\hbar = m = 1$). Note that the harmonic zero-point energy in the potential wells is 0.283 (reduced units). Data reproduced from Ref. [22].

	V_0			
	2	1	0.5	0.25
instanton	4.39(−8)	3.86(−4)	3.04(−2)	2.27(−1)
quantum	4.15(−8)	3.42(−4)	2.25(−2)	1.19(−1)
error	6%	13%	35%	91%

with mass m. The harmonic frequency in the wells is

$$\omega_{\mathrm{s}} = \sqrt{\frac{\nabla^2 V(x_0)}{m}} = \sqrt{\frac{8V_0}{mx_0^2}}. \tag{109}$$

For this particular double-well potential, the zero-temperature kink is known exactly [147]:

$$x(\tau) = \pm x_0 \tanh \frac{\omega_{\mathrm{s}}}{2}(\tau - \tau_c), \tag{110}$$

and the action is

$$S_{\mathrm{kink}} = \int_{-x_0}^{x_0} \sqrt{2mV(x)}\,\mathrm{d}x = \frac{4}{3}x_0\sqrt{2mV_0}. \tag{111}$$

The lowest eigenvalue is of course zero, the next largest is $3\omega_{\mathrm{s}}^2/4$ and the remainder form a continuum such that

$$\Phi = \sqrt{\frac{1}{12\omega_{\mathrm{s}}^2}}. \tag{112}$$

The instanton approximation to the tunnelling splitting is thus given by $\Delta \sim 2\hbar\theta$, where θ is defined by Eq. (101).

The exact tunnelling splitting for this system can also be obtained by numerical solution of the Schrödinger equation. Table IV shows the error implicit in the instanton approach caused by the steepest-descent approximation for a series of values of the barrier height. It is seen that for systems with high barriers the instanton method performs much better than for those with low barriers. This is because, in the latter case, the well is more anharmonic over a short range and, for example, fluctuations of the linear-polymer collapsed in a well may even include paths which access the top of the barrier and visit the other well. This

fluctuation is obviously poorly represented in the steepest-descent approximation. However, even for the lowest barrier tested, which is lower than the harmonic zero-point energy of the wells, the instanton splitting is still within a factor of 2 of the exact quantum result.

To go beyond the steepest-descent approximation, it is possible to use path-integral sampling schemes [148]. However, this is of course much less efficient than the instanton approach and loses the simplicity of interpretation provided by an optimal tunnelling pathway. This approach will then be expected to be of use only if very high accuracy is required.

F.   Molecular tunnelling splittings

Ring-polymer instanton theory has been applied to compute the tunnelling splitting resulting from hydrogen transfer in molecular systems. Here we consider the two benchmark systems of malonaldehyde and the formic acid dimer.

Malonaldehyde is known to rearrange via a single hydrogen transfer. A fully-dimensional PES fitted to CCSD(T) calculations has been constructed by Wang et al. [149]. This PES has a good balance between accuracy an efficiency and has been used with quantum-dynamics approaches including diffusion Monte Carlo (DMC) [150] and the MCTDH method, which in principle give exact numerical results. These calculations give benchmark results for comparison with the instanton approximation.

The formic acid dimer [151] is the prototypical example of a hydrogen-bonded complex which exhibits a double hydrogen transfer. It has recently become possible to obtain a full-dimensional potential-energy surface for the complex fitted to thousands of ab initio calculations [152]. Although the PES appears to be very accurate, reduced-dimensionality quantum dynamics calculations made tunnelling splitting predictions a factor of 2 larger than the observations [152]. Because the double hydrogen transfer leads to a much smaller tunnelling splitting than for malonaldehyde, the exact quantum dynamics approaches are no longer easily applicable and have not been performed successfully for this system. The experimental splitting is therefore used as our benchmark.

A summary of the results obtained by these calculations for the two systems is presented in Table V. The instanton approximation for the splitting in malonaldehyde is seen to be in good agreement with the benchmark calculations performed on the same accurate surface, and hence also with the experimental result. However, when using the less-accurate empirical

TABLE V. Tunnelling splittings obtained for malonaldehyde and the formic acid dimer from various theoretical and experimental studies. The potential-energy surface used in the theoretical calculations is specified.

molecule	PES	method	splitting (cm-1)
H-malonaldehyde	empirical [153]	instanton [22]	51
	Wang et al. [149]	instanton [138]	25
		MCTDH [154]	23.4
		MCTDH [155]	23.8
		DMC [149]	21–22
	—	experiment [156, 157]	21.6
D-malonaldehyde	Wang et al. [149]	instanton [138]	3.4
		DMC [149]	2–3
	—	experiment [158]	2.9
$(HCOOH)_2$	Qu & Bowman [152]	instanton [29]	0.014
		4-mode quantum [152]	0.037
	—	experiment [159, 160]	0.016

PES, the splitting is a factor of 2 too large. In fact, even with *ab initio* potentials, a similar error can be seen in the predictions, ranging from 4.5 cm-1 (QCISD) 77 cm-1 (MP2) [146]. This clearly shows how important it is to use an accurate surface for tunnelling splitting calculations.

The predictions from instanton theory agree with the experimental observations for the tunnelling splitting of the formic acid dimer within a 20% margin of error. The semiclassical approximation inherent in the instanton approach is therefore much less than the error caused by reducing the dimensionality of the model. A systematic study of the dependence of the tunnelling splitting on the number of degrees of freedom included in the calculation shows that the results are very sensitive and require almost all modes to be included before converging [29, 161].

It will always be necessary to make some approximations when simulating molecular processes and it is in most cases not possible to apply exact quantum dynamics in full dimensionality to an ab initio surface. These two applications reviewed here show that

better accuracy is achieved using the full-dimensional semiclassical instanton approach on an accurate surface than exact quantum dynamics within either a reduced-dimensionality framework or with less-expensive potential-energy surfaces.

A number of ring-polymer instanton studies have also been made of the tunnelling splitting patterns in various water clusters. These calculations made use of the many-body potential-energy surfaces fitted to thousands of *ab initio* points [162, 163]. The instanton approach reproduced known results for the water dimer [27] and octamer [136], but predicts new pathways in the trimer [27] and hexamer prism [28] which help explain the experimental splitting patterns. Of particular interest is that these pathways differ greatly from the minimum-energy pathways. This shows clearly that in certain cases, not only does instanton theory provide accurate quantitative results, but it is necessary to obtain the optimal tunnelling pathway in order to have a qualitative understanding of the tunnelling process.

VIII. CONCLUSIONS

We have shown that the instanton approximation can be derived from first principles and that there are a number of equivalent formulae which give the same result but may have particular advantages or disadvantages when implemented numerically. The ring-polymer instanton approach has proved itself to be an accurate and efficient method for simulating tunnelling effects in molecular systems [17]. Two distinct methods exist within this framework, for predicting either rates or tunnelling splittings. Because of its relative simplicity, compared with other quantum dynamics approaches, the method can be used in conjunction with expensive electronic-structure methods, leading to accurate predictions for gas-phase reactions.

Due to the steepest-descent approximation, the errors accrued by the instanton approach are expected to increase when it is applied to floppy molecules which have strongly-anharmonic vibrational modes, or strong coupling between vibrational and rotational modes. However, unless these modes are strongly coupled to the tunnelling coordinate, kinetic isotope effects may still be accurately predicted. The main disadvantage of the instanton approach is that it is not directly applicable to liquid systems. The reason for this is that many different but similar transition states exist corresponding to minor rearrangements of the outer solvation sphere. A more general method is required to study such systems, which

samples the different instantons, such as ring-polymer molecular dynamics (RPMD) [37, 38]. Nonetheless, a full understanding of the derivation of instanton theory has helped confirm the validity of RPMD, and may lead to the discovery of novel path-integral rate theories.

Path-integral simulations are becoming a more common theoretical approach for studying quantum effects in condensed-phase systems. In liquids, this remains the state-of-the art approach [53, 54], but is potentially unnecessary to study tunnelling in solids or on surfaces [164–166]. The instanton method is also derived from the theory of path integrals but does not require a full simulation involving sampling a large number of configurations. Instead the result is obtained from one single optimal tunnelling pathway and is thus much less computationally demanding and gives an easily interpretable mechanistic information.

There are a number of extensions of the instanton approach currently under development. Some suggestions have been given for avoiding the discontinuity at the crossover temperature [12, 56, 84, 103, 167], one of which involves computing the microcanonical rate and integrating over energy to get the thermal rate [19, 83]. There is a particular need for a reliable microcanonical theory in the atmospheric and astro-chemistry communities where complex reaction networks occur in low pressure environments such that the molecules do not react under a thermal equilibrium conditions. Instead, a master-equation solver is commonly used to describe the overall reaction progress based on a set of microcanonical reaction rates [168].

Further extensions to treat nonadiabatic reactions have been discussed in this review in the golden-rule limit only. Some suggestions have been given as to how instanton theory could be extended to go beyond this limiting case [169] but as they are based on the 'Im F' premise, they do not appear to give a reasonable classical high-temperature limit. Further work is needed to benchmark these approaches and to apply them to realistic problems.

IX. ACKNOWLEDGEMENTS

Appendix A: Imaginary-time classical mechanics

In this work, we have mostly employed the Lagrangian formulation of imaginary-time classical mechanics. The alternative Hamilton-Jacobi formulation can be obtained by the Legendre transformation,

$$S(x', x'', \tau) = W(x', x'', E) + E\tau, \tag{A1}$$

where W is the abbreviated action, given by an integral along the path

$$W(x', x'', E) = \int_{x'}^{x''} p(x, E) \, \mathrm{d}x, \tag{A2}$$

and $|p(x, E)| = \sqrt{2m[V(x) - E]}$. Therefore the following equations are obtained by differentiation:

$$E = \frac{\partial S}{\partial \tau} \qquad\qquad \tau = -\frac{\partial W}{\partial E} \tag{A3}$$

$$\frac{\partial S}{\partial x'} = \frac{\partial W}{\partial x'} = -p(x', E) \qquad\qquad \frac{\partial S}{\partial x''} = \frac{\partial W}{\partial x''} = p(x'', E). \tag{A4}$$

To transform between the formulations, the following equations hold [46]:

$$\frac{\partial^2 W}{\partial E^2} = -\left(\frac{\partial^2 S}{\partial \tau^2}\right)^{-1} \tag{A5}$$

$$\frac{\partial^2 W}{\partial x' \partial E} = \frac{\partial^2 S}{\partial x' \partial \tau} \left(\frac{\partial^2 S}{\partial \tau^2}\right)^{-1} \tag{A6}$$

$$\frac{\partial^2 W}{\partial x' \partial x''} = \frac{\partial^2 S}{\partial x' \partial x''} - \frac{\partial^2 S}{\partial x' \partial \tau} \left(\frac{\partial^2 S}{\partial \tau^2}\right)^{-1} \frac{\partial^2 S}{\partial \tau \partial x''}, \tag{A7}$$

and their equivalents with any exchange of x' and x''. Algorithms for evaluating second derivatives of S for discretized path integrals are given in the appendix of Ref. [86].

Appendix B: Harmonic oscillator

For the special case of a particle of mass m in a harmonic oscillator potential $V(x) = \frac{1}{2}m\omega^2(x - x_0)^2 + \epsilon$, the Euclidean action is known to be [41]

$$S(x', x'', \tau) = \frac{m\omega}{2 \sinh \omega\tau} \left[((x' - x_0)^2 + (x'' - x_0)^2) \cosh \omega\tau - 2(x' - x_0)(x'' - x_0)\right] + \tau\epsilon. \tag{B1}$$

Therefore

$$C = \frac{m\omega}{\sinh \omega\tau}.$$ (B2)

When the path is discretized into N segments, the fluctuation factor can be shown to be [46]

$$C_N = \frac{m\tilde{\omega}}{\sinh \tilde{\omega}\tau}$$ (B3)

where $\sinh \tilde{\omega}\tau_N/2 = \omega\tau_N/2$.

---

[1] R. P. Bell, *The Tunnel Effect in Chemistry* (Chapman and Hall, London, 1980).

[2] B. K. Carpenter, Science **332**, 1269 (2011).

[3] D. Ley, D. Gerbig, and P. R. Schreiner, Org. Biomol. Chem. **10**, 3781 (2012).

[4] J. Meisner and J. Kästner, Angew. Chem. Int. Ed. **55**, 5400 (2016).

[5] F. Hund, Z. Phys. **43**, 805 (1927).

[6] H. Eyring, Trans. Faraday Soc. **34**, 41 (1938).

[7] P. Pechukas, Annu. Rev. Phys. Chem. **32**, 159 (1981).

[8] D. G. Truhlar, W. L. Hase, and J. T. Hynes, J. Phys. Chem. **87**, 2664 (1983); D. G. Truhlar, B. C. Garrett, and S. J. Klippenstein, J. Phys. Chem. **100**, 12771 (1996).

[9] P. Hänggi, P. Talkner, and M. Borkovec, Rev. Mod. Phys. **62**, 251 (1990).

[10] S. Coleman, in *Proc. Int. School of Subnuclear Physics* (Erice, 1977) also in S. Coleman, *Aspects of Symmetry*, chapter 7, pp. 265–350 (Cambridge University Press, 1985).

[11] A. O. Caldeira and A. J. Leggett, Ann. Phys.–New York **149**, 374 (1983).

[12] U. Weiss, *Quantum Dissipative Systems*, 4th ed. (World Scientific, Singapore, 2012).

[13] G. W. Gibbons and S. W. Hawking, eds., *Euclidean Quantum Gravity* (World Scientific, Singapore, 1993).

[14] W. H. Miller, J. Chem. Phys. **62**, 1899 (1975).

[15] S. Chapman, B. C. Garrett, and W. H. Miller, J. Chem. Phys. **63**, 2710 (1975).

[16] V. A. Benderskii, D. E. Makarov, and C. A. Wight, *Chemical Dynamics at Low Temperatures*, Adv. Chem. Phys., Vol. 88 (Wiley, New York, 1994).

[17] J. O. Richardson, "Perspective: Ring-polymer instanton theory," (2018), *In preparation for* J. Chem. Phys.

[18] J. O. Richardson, J. Chem. Phys. **144**, 114106 (2016), arXiv:1512.04292 [physics.chem-ph].

[19] J. O. Richardson, Faraday Discuss. **195**, 49 (2016).

[20] E. Wigner, Z. Phys. Chem. B **19**, 203 (1932).

[21] J. O. Richardson and S. C. Althorpe, J. Chem. Phys. **131**, 214106 (2009).

[22] J. O. Richardson and S. C. Althorpe, J. Chem. Phys. **134**, 054109 (2011).

[23] S. Andersson, G. Nyman, A. Arnaldsson, U. Manthe, and H. Jónsson, J. Phys. Chem. A **113**, 4468 (2009).

[24] J. B. Rommel, T. P. M. Goumans, and J. Kästner, J. Chem. Theory Comput. **7**, 690 (2011).

[25] J. Kästner, WIREs Comput. Mol. Sci. **4**, 158 (2014).

[26] A. N. Beyer, J. O. Richardson, P. J. Knowles, J. Rommel, and S. C. Althorpe, J. Phys. Chem. Lett. **7**, 4374 (2016).

[27] J. O. Richardson, S. C. Althorpe, and D. J. Wales, J. Chem. Phys. **135**, 124109 (2011).

[28] J. O. Richardson, C. Pérez, S. Lobsiger, A. A. Reid, B. Temelso, G. C. Shields, Z. Kisiel, D. J. Wales, B. H. Pate, and S. C. Althorpe, Science **351**, 1310 (2016).

[29] J. O. Richardson, Phys. Chem. Chem. Phys. **19**, 966 (2017), arXiv:1611.04816 [physics.chem-ph].

[30] H. Jónsson, P. Natl. Acad. Sci. USA **108**, 944 (2011).

[31] T. Lamberts and J. Kästner, Astrophys. J. **846**, 43 (2017).

[32] J. Meisner, T. Lamberts, and J. Kästner, ACS Earth Space Chem. **1**, 399 (2017).

[33] B. Senevirathne, S. Andersson, F. Dulieu, and G. Nyman, Mol. Astrophys. **6**, 59 (2017).

[34] J. B. Rommel, Y. Liu, H.-J. Werner, and J. Kästner, J. Phys. Chem. B **116**, 13682 (2012).

[35] M. J. Gillan, J. Phys. C Solid State **20**, 3621 (1987).

[36] G. A. Voth, D. Chandler, and W. H. Miller, J. Chem. Phys. **91**, 7749 (1989).

[37] I. R. Craig and D. E. Manolopoulos, J. Chem. Phys. **123**, 034102 (2005).

[38] S. Habershon, D. E. Manolopoulos, T. E. Markland, and T. F. Miller III, Annu. Rev. Phys. Chem. **64**, 387 (2013).

[39] W. H. Miller, Y. Zhao, M. Ceotto, and S. Yang, J. Chem. Phys. **119**, 1329 (2003).

[40] J. Vaníček, W. H. Miller, J. F. Castillo, and F. J. Aoiz, J. Chem. Phys. **123**, 054108 (2005).

[41] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).

[42] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (McGraw-Hill, New York, 1978).

[43] M. C. Gutzwiller, *Chaos in Classical and Quantum Mechanics* (Springer-Verlag, New York, 1990).

[44] J. H. van Vleck, P. Natl. Acad. Sci. USA **14**, 178 (1928).

[45] W. H. Miller, J. Chem. Phys. **55**, 3146 (1971).

[46] H. Kleinert, *Path Integrals in Quantum Mechanics, Statistics, Polymer Physics and Financial Markets*, 5th ed. (World Scientific, Singapore, 2009).

[47] S. C. Althorpe, J. Chem. Phys. **134**, 114104 (2011).

[48] D. Chandler and P. G. Wolynes, J. Chem. Phys. **74**, 4078 (1981).

[49] M. Parrinello and A. Rahman, J. Chem. Phys. **80**, 860 (1984).

[50] D. Marx and M. Parrinello, J. Chem. Phys. **104**, 4077 (1996).

[51] M. E. Tuckerman, in *Quantum Simulations of Complex Many-Body Systems: From Theory to Algorithms,* Lecture Notes, NIC Series, Vol. 10, edited by J. Grotendorst, D. Marx, and A. Muramatsu (John von Neumann Institute for Computing, Jülich, 2002) pp. 269–298.

[52] I. R. Craig and D. E. Manolopoulos, J. Chem. Phys. **121**, 3368 (2004).

[53] M. Ceriotti, W. Fang, P. G. Kusalik, R. H. McKenzie, A. Michaelides, M. A. Morales, and T. E. Markland, Chem. Rev. **116**, 7529 (2016).

[54] T. E. Markland and M. Ceriotti, Nat. Rev. Chem. **2**, 0109 (2018).

[55] S. Coleman, Phys. Rev. D **15**, 2929 (1977).

[56] J. Cao and G. A. Voth, J. Chem. Phys. **105**, 6856 (1996).

[57] A. Andreassen, D. Farhi, W. Frost, and M. D. Schwartz, Phys. Rev. Lett. **117**, 231601 (2016).

[58] J. S. Langer, Ann. Phys.–New York **41**, 108 (1967); Ann. Phys.–New York **54**, 258 (1969); C. G. Callan, Jr and S. Coleman, Phys. Rev. D **16**, 1762 (1977).

[59] I. Affleck, Phys. Rev. Lett. **46**, 388 (1981).

[60] M. Stone, Phys. Lett. B **67**, 186 (1977).

[61] H. Aoyama, T. Harano, H. Kikuchi, I. Okouchi, M. Sato, and S. Wada, Prog. Theor. Phys. Supp. **127**, 1 (1997).

[62] W. H. Miller, S. D. Schwartz, and J. W. Tromp, J. Chem. Phys. **79**, 4889 (1983).

[63] G. A. Voth, D. Chandler, and W. H. Miller, J. Phys. Chem. **93**, 7009 (1989).

[64] M. Messina, G. K. Schenter, and B. C. Garrett, J. Chem. Phys. **103**, 3430 (1995).

[65] K. Karandashev, Ph.D. thesis, EPFL (2017).

[66] J. O. Richardson, R. Bauer, and M. Thoss, J. Chem. Phys. **143**, 134115 (2015), arXiv:1508.04919 [physics.chem-ph].

[67] T. J. H. Hele and S. C. Althorpe, J. Chem. Phys. **138**, 084108 (2013).

[68] M. C. Gutzwiller, J. Math. Phys. **12**, 343 (1971).

[69] J. D. Doll, T. L. Beck, and D. L. Freeman, Int. J. Quantum Chem. **36**, 73 (1989).

[70] M. Kryvohuz, J. Chem. Phys. **134**, 114103 (2011).

[71] W. Fang, J. O. Richardson, J. Chen, X.-Z. Li, and A. Michaelides, Phys. Rev. Lett. **119**, 126001 (2017), arXiv:1709.02752 [physics.chem-ph].

[72] S. Álvarez-Barcia, J. R. Flores, and J. Kästner, J. Phys. Chem. A **118**, 78 (2013).

[73] D. E. Makarov and M. Topaler, Phys. Rev. E **52**, 178 (1995).

[74] G. Mills, G. K. Schenter, D. E. Makarov, and H. Jónsson, Chem. Phys. Lett. **278**, 91 (1997).

[75] S. C. Althorpe and T. J. H. Hele, J. Chem. Phys. **139**, 084115 (2013); T. J. H. Hele and S. C. Althorpe, J. Chem. Phys. **139**, 084116 (2013).

[76] M. Ceotto, Mol. Phys. **110**, 547 (2012).

[77] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 3rd ed. (Cambridge University Press, Cambridge, 2007).

[78] J. Baker, J. Comput. Chem. **7**, 385 (1986).

[79] J. Nichols, H. Taylor, P. Schmidt, and J. Simons, J. Chem. Phys. **92**, 340 (1990).

[80] J. B. Rommel and J. Kästner, J. Chem. Phys. **134**, 184107 (2011).

[81] L. J. Munro and D. J. Wales, Phys. Rev. B **59**, 3969 (1999).

[82] R. Fletcher, *Practical Methods of Optimization*, 2nd ed. (John Wiley and Sons, Chichester, 1987).

[83] S. R. McConnell, A. Löhle, and J. Kästner, J. Chem. Phys. **146**, 074105 (2017).

[84] S. McConnell and J. Kästner, J. Comput. Chem. **38**, 2570 (2017).

[85] G. Di Liberto and M. Ceotto, J. Chem. Phys. **145**, 144107 (2016).

[86] J. O. Richardson, J. Chem. Phys. **143**, 134116 (2015), arXiv:1508.05195 [physics.chem-ph].

[87] R. Pérez de Tudela, Y. V. Suleimanov, J. O. Richardson, V. Sáez Rábanos, W. H. Green, and F. J. Aoiz, J. Phys. Chem. Lett. **5**, 4219 (2014).

[88] T. P. M. Goumans and J. Kästner, Angew. Chem. Int. Edit. **49**, 7350 (2010); S. Andersson, T. P. M. Goumans, and A. Arnaldsson, Chem. Phys. Lett. **513**, 31 (2011).

[89] J. Meisner, J. B. Rommel, and J. Kästner, J. Comput. Chem. **32**, 3456 (2011).

[90] J. Kästner, Chem. Eur. J. **19**, 8207 (2013).

[91] S. Álvarez-Barcia, M.-S. Russ, J. Meisner, and J. Kästner, Faraday Discuss. **195**, 69 (2016); T. Lamberts and J. Kästner, J. Phys. Chem. A **121**, 9736 (2017).

[92] J. Meisner and J. Kästner, J. Chem. Phys. **144**, 174303 (2016); T. Lamberts, P. K. Samanta, A. Köhn, and J. Kästner, Phys. Chem. Chem. Phys. **18**, 33021 (2016).

[93] V. Ásgeirsson, A. Arnaldsson, and H. Jónsson, J. Chem. Phys. **148**, 102334 (2018).

[94] M. Kryvohuz and R. A. Marcus, J. Chem. Phys. **137**, 134107 (2012); M. Kryvohuz, J. Chem. Phys. **137**, 234304 (2012); J. Phys. Chem. A **118**, 535 (2014).

[95] C. Eckart, Phys. Rev. **35**, 1303 (1930).

[96] R. Pérez de Tudela, F. J. Aoiz, Y. V. Suleimanov, and D. E. Manolopoulos, J. Chem. Phys. Lett. **3**, 493 (2012).

[97] Y. V. Suleimanov, R. P. de Tudela, P. G. Jambrina, J. F. Castillo, V. Sáez-Rábanos, D. E. Manolopoulos, and F. J. Aoiz, Phys. Chem. Chem. Phys. **15**, 3655 (2013).

[98] B. R. Johnson, J. Comput. Phys. **13**, 445 (1973).

[99] D. E. Manolopoulos, M. D'Mello, and R. E. Wyatt, J. Chem. Phys. **91**, 6096 (1989).

[100] D. Skouteris, J. F. Castillo, and D. E. Manolopoulos, Comput. Phys. Commun. **133**, 128 (2000).

[101] K. Karandashev, Z.-H. Xu, M. Meuwly, J. Vaníček, and J. O. Richardson, Struct. Dynam. **4**, 061501 (2017).

[102] A. I. Boothroyd, W. J. Keogh, P. G. Martin, and M. R. Peterson, J. Chem. Phys. **104**, 7139 (1996).

[103] Y. Zhang, J. B. Rommel, M. T. Cvitaš, and S. C. Althorpe, Phys. Chem. Chem. Phys. **16**, 24292 (2014).

[104] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Schütz, WIREs Comput. Mol. Sci. **2**, 242 (2012).

[105] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, *et al.*, "MOLPRO, version 2012.1, a package of ab initio programs," (2012).

[106] G. Schiffel, U. Manthe, and G. Nyman, J. Phys. Chem. A **114**, 9617 (2010).

[107] J. C. Corchado, J. L. Bravo, and J. Espinosa-Garcia, J. Chem. Phys. **130**, 184314 (2009).

[108] H.-D. Meyer, U. Manthe, and L. S. Cederbaum, Chem. Phys. Lett. **165**, 73 (1990).

[109] G. W. Richings, I. Polyak, K. E. Spinlove, G. A. Worth, I. Burghardt, and B. Lasorne, Int. Rev. Phys. Chem. **34**, 269 (2015).

[110] R. A. Marcus, Rev. Mod. Phys. **65**, 599 (1993).

[111] R. A. Kuharski, J. S. Bader, D. Chandler, M. Sprik, M. L. Klein, and R. W. Impey, J. Chem. Phys. **89**, 3248 (1988).

[112] J. Blumberger, Phys. Chem. Chem. Phys. **10**, 5651 (2008).

[113] D. Chandler, in *Classical and Quantum Dynamics in Condensed Phase Simulations*, edited by B. J. Berne, G. Ciccotti, and D. F. Coker (World Scientific, Singapore, 1998) Chap. 2, pp. 25–49.

[114] R. A. Marcus, J. Chem. Phys. **24**, 966 (1956).

[115] R. A. Marcus, Discuss. Faraday Soc. **29**, 21 (1960).

[116] R. A. Marcus, Annu. Rev. Phys. Chem. **15**, 155 (1964).

[117] R. A. Marcus and N. Sutin, Biochim. Biophys. Acta **811**, 265 (1985).

[118] A. Garg, J. N. Onuchic, and V. Ambegaokar, J. Chem. Phys. **83**, 4491 (1985).

[119] J. Ulstrup, *Charge Transfer Processes in Condensed Media* (Springer-Verlag, Berlin, 1979).

[120] P. Siders and R. A. Marcus, J. Am. Chem. Soc. **103**, 741 (1981).

[121] P. Siders and R. A. Marcus, J. Am. Chem. Soc. **103**, 748 (1981).

[122] J. S. Bader, R. A. Kuharski, and D. Chandler, J. Chem. Phys. **93**, 230 (1990).

[123] M. Topaler and N. Makri, J. Phys. Chem. **100**, 4430 (1996).

[124] H. Wang, D. E. Skinner, and M. Thoss, J. Chem. Phys. **125**, 174502 (2006).

[125] J. Tang, Chem. Phys. **188**, 143 (1994).

[126] J. Tang, Chem. Phys. **179**, 105 (1994); A. J. Leggett, S. Chakravarty, A. T. Dorsey, M. P. A. Fisher, A. Garg, and W. Zwerger, Rev. Mod. Phys. **59**, 1 (1987).

[127] P. G. Wolynes, J. Chem. Phys. **87**, 6559 (1987).

[128] J. Mattiat and J. O. Richardson, J. Chem. Phys. **148**, 102311 (2018), arXiv:1708.06702 [physics.chem-ph].

[129] J. O. Richardson and M. Thoss, J. Chem. Phys. **139**, 031102 (2013); J. O. Richardson, P. Meyer, M.-O. Pleinert, and M. Thoss, Chem. Phys. **482**, 124 (2017), arXiv:1609.00644 [physics.chem-ph].

[130] P. Shushkov, R. Li, and J. C. Tully, J. Chem. Phys. **137**, 22A549 (2012); P. Shushkov, J. Chem. Phys. **138**, 224102 (2013).

[131] A. R. Menzeleev, N. Ananth, and T. F. Miller, III, J. Chem. Phys. **135**, 074106 (2011); J. S. Kretchmer and T. F. Miller III, J. Chem. Phys. **138**, 134109 (2013); A. R. Menzeleev, F. Bell, and T. F. Miller III, J. Chem. Phys. **140**, 064103 (2014).

[132] N. Ananth, J. Chem. Phys. **139**, 124102 (2013); J. R. Duke and N. Ananth, Faraday Discuss. **195**, 253 (2016); T. J. H. Hele and N. Ananth, Faraday Discuss. **195**, 269 (2016); S. Pierre, J. R. Duke, T. J. Hele, and N. Ananth, J. Chem. Phys. **147**, 234103 (2017).

[133] S. N. Chowdhury and P. Huo, J. Chem. Phys. **147**, 214109 (2017), arXiv:1706.08403 [physics.chem-ph].

[134] J. E. Lawrence and D. E. Manolopoulos, J. Chem. Phys. **148**, 102313 (2018).

[135] R. Zwanzig, *Nonequilibrium Statistical Mechanics* (Oxford University Press, 2001).

[136] J. O. Richardson, D. J. Wales, S. C. Althorpe, R. P. McLaughlin, M. R. Viant, O. Shih, and R. J. Saykally, J. Phys. Chem. A **117**, 6960 (2013).

[137] T. Kawatsu and S. Miura, J. Chem. Phys. **141**, 024101 (2014); Chem. Phys. Lett. **634**, 146 (2015).

[138] M. T. Cvitaš and S. C. Althorpe, J. Chem. Theory Comput. **12**, 787 (2016).

[139] M. T. Cvitas, J. Chem. Theory Comput. **14**, 1487 (2018).

[140] R. J. Wilson, *Introduction to Graph Theory*, 4th ed. (Pearson Education, Harlow, 1996).

[141] J. A. Bondy and U. S. R. Murty, *Graph Theory with Applications* (Macmillan, London, 1976).

[142] D. C. Liu and J. Nocedal, Math. Program. **45**, 503 (1989).

[143] P. Faccioli, M. Sega, F. Pederiva, and H. Orland, Phys. Rev. Lett. **97**, 108101 (2006).

[144] D. M. Einarsdóttir, A. Arnaldsson, F. Óskarsson, and H. Jónsson, in *Applied Parallel and Scientific Computing*, Lecture Notes in Computer Science, Vol. 7134, edited by K. Jónasson, 10th International Conference, PARA 2010 (Springer-Verlag, Berlin, 2012) pp. 45–55.

[145] G. V. Mil'nikov and H. Nakamura, J. Chem. Phys. **115**, 6881 (2001).

[146] G. V. Mil'nikov, K. Yagi, T. Taketsugu, H. Nakamura, and K. Hirao, J. Chem. Phys. **120**, 5036 (2004).

[147] A. I. Vainshtein, V. I. Zakharov, V. A. Novikov, and M. A. Shifman, Sov. Phys. Uspekhi **25**, 195 (1982), also in *Instantons in Gauge Theories*, edited by M. Shifman, pp. 468 (Singapore: World Scientific, 1994).

[148] E. Mátyus, D. J. Wales, and S. C. Althorpe, J. Chem. Phys. **144**, 114108 (2016); E. Mátyus and S. C. Althorpe, J. Chem. Phys. **144**, 114109 (2016); C. L. Vaillant, D. J. Wales, and S. C. Althorpe, "Tunneling-splittings from path-integral molecular dynamics using a langevin thermostat," (2018), arXiv:1803.04433v1 [physics.chem-ph].

[149] Y. Wang, B. J. Braams, J. M. Bowman, S. Carter, and D. P. Tew, J. Chem. Phys. **128**, 224314 (2008).

[150] M. A. Suhm and R. O. Watts, Phys. Rep. **204**, 293 (1991).

[151] Ö. Birer and M. Havenith, Annu. Rev. Phys. Chem. **60**, 263 (2009).

[152] C. Qu and J. M. Bowman, Phys. Chem. Chem. Phys. **18**, 24835 (2016).

[153] T. D. Sewell, Y. Guo, and D. L. Thompson, J. Chem. Phys. **103**, 8557 (1995); Y. Guo, T. D. Sewell, and D. L. Thompson, Chem. Phys. Lett. **224**, 470 (1994).

[154] M. Schröder, F. Gatti, and H.-D. Meyer, J. Chem. Phys. **134**, 234307 (2011).

[155] T. Hammer and U. Manthe, J. Chem. Phys. **134**, 224305 (2011).

[156] D. W. Firth, K. Beyer, M. A. Dvorak, S. W. Reeve, A. Grushow, and K. R. Leopold, J. Chem. Phys. **94**, 1812 (1991).

[157] T. Baba, T. Tanaka, I. Morino, K. M. T. Yamada, and K. Tanaka, J. Chem. Phys. **110**, 4131 (1999).

[158] S. L. Baughcum, R. W. Duerst, W. F. Rowe, Z. Smith, and E. B. Wilson, J. Am. Chem. Soc. **103**, 6296 (1981).

[159] M. Ortlieb and M. Havenith, J. Phys. Chem. A **111**, 7355 (2007).

[160] K. G. Goroya, Y. Zhu, P. Sun, and C. Duan, J. Chem. Phys. **140**, 164311 (2014).

[161] K. Yagi, G. V. Mil'nikov, T. Taketsugu, K. Hirao, and H. Nakamura, Chem. Phys. Lett. **397**, 435 (2004).

[162] Y. Wang and J. M. Bowman, Chem. Phys. Lett. **491**, 1 (2010).

[163] V. Babin, G. R. Medders, and F. Paesani, J. Chem. Theory Comput. **10**, 1599 (2014).

[164] L. J. Lauhon and W. Ho, Phys. Rev. Lett. **85**, 4566 (2000); A. Yazdani, Nature **409**, 471 (2001).

[165] C. Drechsel-Grau and D. Marx, Phys. Rev. Lett. **112**, 148302 (2014).

[166] X. Meng, J. Guo, J. Peng, J. Chen, Z. Wang, J.-R. Shi, X.-Z. Li, E.-G. Wang, and Y. Jiang, Nat. Phys. **11**, 235 (2015); C. Drechsel-Grau and D. Marx, Nat. Phys. **11**, 216 (2015).

[167] M. Kryvohuz, J. Chem. Phys. **138**, 244114 (2013).

[168] D. R. Glowacki, C.-H. Liang, C. Morley, M. J. Pilling, and S. H. Robertson, J. Phys. Chem. A **116**, 9545 (2012).

[169] J. Cao, C. Minichino, and G. A. Voth, J. Chem. Phys. **103**, 1391 (1995).