Theor Chem Acc (2006) 116: 373–382
DOI 10.1007/s00214-005-0033-9

REGULAR ARTICLE

Styliani Consta

Detecting reaction pathways and computing reaction rates in condensed phase

Received: 2 May 2005 / Accepted: 21 October 2005 / Published online: 26 January 2006

Abstract Methods for the computation of rate constants that characterize classical reactions occurring in the condensed phase are discussed. While microscopic expressions for these transport properties are well known, their computation presents challenges for simulation since reactive events often occur rarely, and the long time scales that are typical for reactive processes are not accessible using simple molecular dynamics methods. Furthermore, the underlying free energy surface is very complex with many saddle points that prevent sampling of possible reaction pathways. As a result, the reaction coordinate may be a complex many-body function of the system's degrees of freedom. Since there is not an a priori way to define a "good" reaction coordinate, methods are being developed to assist in a systematic construction of a reaction coordinate. These methods are reviewed and examples of non-trivial reaction coordinates are presented.

1 Introduction

Theoretical and experimental research in the study of reactions in the condensed phase has made remarkable progress in the understanding of the role of solvent on the mechanisms of a wide array of chemical reactions. Improvements in instrumentation and analysis has given possibilities of exploring reaction mechanisms in new environments such as surfaces [1] and clusters [2–4]. Advances in femtosecond optical techniques [5–7] stimulated studies on the role of the solvent in reaction pathways in solution [8] while various microscopies are used to investigate molecular details of reactions on surfaces [1]. Such techniques provide a great deal of insight into details of reaction pathways on microscopic scales and, together with advances in molecular simulation techniques [9], into limitations of phenomenological macroscopic descriptions of reaction rates [10]. Better understanding of the underlying phenomena lead to the development of many powerful theoretical models. Evolution of computational techniques developed for the past several decades has led to many advanced simulation methods for the study of reactions in condensed phase ranging from continuum modelling of the solvent [11,12] to molecular modelling where an extensive use of molecular dynamics (MD) techniques is made [13–15].

In this article, we provide an outline of the path of research on barrier crossing dynamics in condensed phase. Emphasis is given on the reactive flux method for the computation of rate constants as it is derived rigorously from fundamental principles using projection operator techniques. In Sect. 2 the use of the scheme in computer simulations is discussed. Due to the crucial role that a reaction coordinate (RC) plays in the computation of the reaction rate constants in Sect. 3 we review methods that are being developed to define RCs in a systematic way. Examples of non-trivial RCs are presented in Sect. 4 and the article closes with a summary of unresolved questions on methods and directions in the field.

Probably the most well-known phenomenological theoretical treatments of condensed phase reactions with a continuum treatment of the solvent are Smoluchowski's [16] diffusion equation approach for diffusion-limited reactions and Kramers' [17] Fokker–Planck equation treatment of barrier crossing dynamics. Both of these approaches account for solvent effects by introducing phenomenological diffusion and friction terms. Kramers [17] models a reaction by the Brownian motion of a particle that escapes from a potential well over a barrier. The minimum of the potential well is harmonic potential and the barrier an inverted parabola. The Langevin equation describing dynamics of a particle in viscous solution reads,

$$M\frac{\mathrm{d}u(t)}{\mathrm{d}t} = -\zeta u(t) - \frac{\partial W(\xi)}{\partial \xi} + f(t), \tag{1}$$

where u is the velocity of the Brownian particle in a fluid, M its mass, ξ is the RC, $W(\xi)$ the potential of mean force (PMF), ζ is the friction coefficient and $f(t)$ is a random force. The random force is taken to be a Gaussian random variable with

S. Consta
Department of Chemistry,
University of Western Ontario,
London, ON, N6A 5B7, Canada
E-mail: sconstas@uwo.ca

white noise spectrum and is related to the friction constant by the fluctuation–dissipation relation, $\langle f(t)f(t')\rangle = 2k_BT\zeta\,\delta(t-t')$, where T is the temperature, k_B Boltzmann's constant and $\delta(t)$ is the Dirac delta function. The solution of Kramers' model distinguishes three regimes, of high, moderate and low friction. The rate coefficients that correspond to the three regimes for irreversible passage across the barrier are given respectively by

$$k_f = \frac{\omega_0\omega' M}{2\pi\zeta}\exp(-\Delta W), \tag{2}$$

$$k_f = \frac{\omega_0}{2\pi}\exp(-\Delta W), \tag{3}$$

$$k_f = \frac{\zeta\beta\Delta W}{M}\exp(-\Delta W) \tag{4}$$

where ΔW is the difference in the PMF between the barrier top and the minimum, ω_0 and ω' are the frequencies of the harmonic potentials at the minimum and the barrier top of the PMF, respectively. The theory was generalized by Grote–Hynes for friction with memory in the high to moderate friction regime [18–21] and extensive studies have been made to compare Kramers' and Grote–Hynes models (see for example [22,23]). The Hamiltonian origin of Grote–Hynes model was investigated [24–29] by describing the effect of the solvent by a system of harmonic oscillators bi-linearly coupled with the RC. Many studies have also been devoted to find a unified expression for the rate coefficients that holds in all friction regimes (see for example [18,30,31]). A historical review on Kramers' theory and developments of methods on reaction rates are found in Refs. [11,15,32].

In parallel to the development of methods that describe the solvent effect macroscopically, microscopic expressions of chemical rate constants in terms of autocorrelation functions were developed. Such expressions were first derived by Yamamoto [33] in 1960 using linear response theory. These expressions have been re-derived in various ways and form the basis for numerical calculations of rate constants for condensed phase systems [34–39]. A general method for the derivation of rate laws and rate constants can be formulated using projection operator techniques [34,40]. Zwanzig [41, 42] and Mori [43,44] introduced projection operator techniques to provide microscopic derivation of equations that have the form of Eq. (1), generalized to any dynamical variable, and relaxed the condition that the random force correlations decay instantaneously as reflected by the δ-function correlations. The derivation of the generalized Langevin equation for a dynamical variable denoted by G is presented in Appendix A. The final expression is:

$$\frac{\mathrm{d}G(t)}{\mathrm{d}t} = i\Omega G(t) - \int_0^t \mathrm{d}\tau\, K(\tau)G(t-\tau) + F(t), \tag{5}$$

where $i\Omega = (iLG(0), G)(G, G)^{-1}$ is termed the frequency matrix; $F(t) = \mathrm{e}^{iL_Qt}iL_QG$ is the random force; $K(\tau) = -(F(\tau), F(0))(G, G)^{-1}$ is the memory kernel. $L_Q = QL$ where L is the Liouville operator and Q is the projection operator that projects onto the space of dynamical variables orthogonal to G. $(.,.)$ denotes an inner product defined as $(U, V) = \int UV^*\rho_0\mathrm{d}\Gamma$ where ρ_0 is the equilibrium distribution function and $\mathrm{d}\Gamma$ is the phase space volume element. To illustrate the derivation of the macroscopic rate law, consider the simple situation where two chemical species interconvert,

$$A \underset{k_r}{\overset{k_f}{\rightleftharpoons}} B, \tag{6}$$

and recall the phenomenological description of the rate of this reaction. Let the average number of A (B) molecules in the system at time t be $\bar{N}_A(t)$ ($\bar{N}_B(t)$) and their values at chemical equilibrium be $\bar{N}_A^{\mathrm{eq}}$ ($\bar{N}_B^{\mathrm{eq}}$). The mass action rate law for this system is

$$\begin{aligned}\frac{\mathrm{d}\overline{N}_A(t)}{\mathrm{d}t} &= -k_f\overline{N}_A(t) + k_r\overline{N}_B(t),\\ \frac{\mathrm{d}\delta\overline{N}_A(t)}{\mathrm{d}t} &= -(1+K_{\mathrm{eq}}^{-1})k_f\delta\overline{N}_A(t).\end{aligned} \tag{7}$$

In the second line of this equation, we used the fact that $\delta\bar{N}_B = -\delta\bar{N}_A$ and definition of the equilibrium constant and the detailed balance condition, $K_{\mathrm{eq}} = \dfrac{k_f}{k_r} = \dfrac{\overline{N}_B^{\mathrm{eq}}}{\overline{N}_A^{\mathrm{eq}}}$.

To derive this chemical rate law from the microscopic equations of motion, it is only necessary to choose microscopic variables in Eq. (5) that correspond to chemical species. The choice of microscopic chemical species variables can be a difficult task and some examples will be given below. Suppose we know the microscopic expressions for the dynamical variables N_A (N_B) which correspond to the number of molecules of A (B). These variables depend in general on the positions and momenta of the particles in the system and their average values correspond to the $\bar{N}_A(t)$ ($\bar{N}_B(t)$) that enter the phenomenological rate law. The deviations of these dynamical variables from their average values at equilibrium, $\overline{N}_A^{\mathrm{eq}}$ ($\overline{N}_B^{\mathrm{eq}}$), are denoted $\delta N_A = N_A - \overline{N}_A^{\mathrm{eq}}$ ($\delta N_B = N_B - \overline{N}_B^{\mathrm{eq}}$). Substituting δN_A and δN_B in Eq. (5) we obtain the generalized Langevin equation for the reactive system,

$$\frac{\mathrm{d}\delta N_A(t)}{\mathrm{d}t} = -(1+K_{\mathrm{eq}}^{-1})\int_0^t \mathrm{d}\tau\, K(\tau)\delta N_A(t-\tau) + F(t), \tag{8}$$

where $K = (F(\tau), iL\delta N_A)(\delta N_A, \delta N_A)^{-1} = (F(\tau), \delta\dot{N}_A)(\delta N_A, \delta N_A)^{-1}$. The frequency term in Eq. (5) is zero as a result of time reversal symmetry and the random force is given by $F(\tau) = \mathrm{e}^{iL_Q\tau}iL\delta N_A$.

In order to obtain the macroscopic chemical rate law from the generalized Langevin equation, we must first construct the equation for the average values of the chemical species numbers. In analogy with macroscopic chemical studies, we suppose that an initial non-equilibrium ensemble is constructed where only the chemical concentrations are specified

and all other degrees of freedom are assumed to be in equilibrium. The average of the dynamical variable at time t over this ensemble will be denoted by the same symbol as the average concentration in the phenomenological law, $\overline{N}_A(t)$. The average of the random force in this ensemble is zero, since it is orthogonal to the chemical species variables. Finally, if one assumes that the memory kernel decays much more rapidly than the species concentrations, an assumption whose validity rests on the fact that species concentrations are the only slow variables in the system, the chemical species number may be removed from under the integral and the integral extended to infinity. The result of these operations is the macroscopic chemical rate law,

$$\frac{\mathrm{d}\delta\bar{N}_A(t)}{\mathrm{d}t} = -(1 + K_{\mathrm{eq}}^{-1})\Big[\int_0^\infty \mathrm{d}\tau\, K(\tau)\Big]\delta\bar{N}_A(t). \tag{9}$$

By comparison with the phenomenological rate law in Eq. (7), one obtains the rate constant as the time integral of the random flux autocorrelation function,

$$k_f = \int_0^\infty \mathrm{d}\tau\, K(\tau). \tag{10}$$

The rigorous definition of the rate kernel involves the propagator e^{iQLt} that involves projected dynamics; however, one usually computes unprojected (natural) evolution where the rate kernel is given by the same expression as for projected dynamics, but the propagator e^{iQLt} is replaced by e^{iLt}. We denote the rate kernel using unprojected dynamics by $\tilde{K}$. Some care must be used in the evaluation of the rate kernel when this approximation is used since the infinite time integral of $\tilde{K}(\tau)$ is zero. (For a derivation of this result see Ref. [45]). The computation of the rate constant rests on the existence of a plateau value. If the time scale for the decay of the memory kernel is t_{mic} and that of the chemical concentrations is t_{chem} and these two time scales are widely separated, then one can define time t^* such that $t_{\mathrm{mic}} \ll t^* \ll t_{\mathrm{chem}}$. In such a circumstance, one may approximate the rate coefficient by

$$\begin{aligned} k_f \approx k_f(t^*) &= \int_0^{t^*} \mathrm{d}\tau\, \tilde{K}(\tau) \\ &= \int_0^{t^*} \mathrm{d}\tau\, (\delta\dot{N}_A(\tau), \delta\dot{N}_A)(\delta N_A, \delta N_A)^{-1}. \end{aligned} \tag{11}$$

2 Reaction coordinate and rate constants

In order to illustrate the use of this formula in a simple context, suppose a one-dimensional RC $\xi(q_1, q_2, \dots)$ is known and can be used to characterize the passage from A to B species. This variable may be a complicated function of the positions of the molecules in the system. The dynamical variable corresponding to species A is $N_A(t) = \theta(\xi(t) - \xi^\ddagger)$ so that when $\xi(t) < \xi^\ddagger$, we have species A and species B for $\xi \geq \xi^\ddagger$. Here $\theta(x)$ is the Heaviside function. In this case the time-dependent rate constant [Eq. (11) with $t^* \to t$] takes the form,

$$k_f(t) = \left(\frac{1}{\bar{N}_A^{\mathrm{eq}}} + \frac{1}{\bar{N}_B^{\mathrm{eq}}}\right)\langle\dot{\xi}\delta(\xi - \xi^\dagger)\theta(\xi(t) - \xi^\dagger)\rangle, \tag{12}$$

where we have performed the time integral to obtain this expression.

In computer simulations, the rate constant is usually computed in two steps [35,34]. First, the $t = 0$ value of the rate constant is calculated in the following way: suppose we consider a time $t = 0 + \epsilon \equiv 0+$ where ϵ is a small positive number. Equation (12) may be written as

$$k_f^{\mathrm{TST}} = k_f(0+) = \left(\frac{1}{\bar{N}_A^{\mathrm{eq}}} + \frac{1}{\bar{N}_B^{\mathrm{eq}}}\right)\langle\dot{\xi}\delta(\xi - \xi^\dagger)\theta(\dot{\xi})\rangle, \tag{13}$$

which is just the transition state theory expression, k_f^{TST}, for the rate constant [34,37]. The full time-dependent rate constant may then be expressed as,

$$k_f(t) = k_f^{\mathrm{TST}}\kappa(t), \tag{14}$$

Here $\kappa(t)$ $(0 < \kappa(t) \leq 1)$ is the transmission coefficient. The transition state of the reaction is found by computing the reversible work (free energy) $W(\xi)$ associated with the RC by

$$W(\xi) = -k_B T \ln \frac{P(\xi)}{P_u}, \tag{15}$$

where

$$P(\xi) = \langle\delta(\xi - \xi^\dagger)\rangle \tag{16}$$

is the probability density of finding the system with a particular value of the RC equal to ξ and P_u is a normalization factor. Distinction should be drawn, however, between the phenomenological definition of the free energy of activation and the reversible work $W(\xi)$ [46]. Definition of the latter (15) implies that its value depends not only on the value of ξ but also on the functional form of the variable according to the change of variable rules in probability theory. This correction is often called the entropic factor [23].

The computation of the free energy may be difficult since the transitions between reactants and products are rare events that cannot be studied by direct MD simulations because of insufficient sampling of transitions. The free energy surface can be constructed using non-trivial sampling techniques such as Monte Carlo methods in the context of umbrella sampling [47], MD using the "blue moon" ensemble [48–52] and metadynamics technique [53,54]. From a knowledge of the free energy at the barrier top and the equilibrium statistical averages in Eq. (14) one may calculate k_f^{TST}. One of the assumptions of TST is that the trajectories that arrive at the barrier top with velocities directed towards the product state will end in the product state: re-crossings of the free energy barrier are neglected. For reactions in the condensed phase,

this assumption often breaks down and the re-crossing corrections may change the value of the TST estimate. If the changes are too large, this signals a poor choice of RC unless it is diffusive barrier crossing. An example of such a breakdown is a nucleation process where changes in nucleus size are independent processes. The re-crossing correction in this case is given by the Zeldovich factor [55,56] and is a very small quantity. Sometimes the re-crossing can be reduced by a clever choice of RC [57], however, in many cases, such a choice is difficult.

Re-crossing corrections are taken into account by the transmission coefficient, $\kappa(t)$, which is computed by monitoring $\theta(\xi(t) - \xi^{\ddagger})$ for a large number of trajectories that evolve starting from the barrier top. The transmission coefficient will decay with time and if the time scale separation discussed above exists, the rate coefficient can be determined from the plateau value of $\kappa = \kappa(t^*)$ and the value of k_f^{TST}.

One of the difficulties in the study of the reaction rates is the search for a suitable RC. A 'good' RC may involve the solvent degrees of freedom in addition to the degrees of freedom of the reacting species. Often, it is a challenging problem in finding such an RC and methods have been developed [58] that may assist in finding out what effects are important in the reaction mechanism and therefore, help in constructing an RC.

3 Systematic ways of finding RCs

3.1 Transition path sampling

Transition path sampling (TPS) developed by Bolhuis, Chandler, Dellago, Geissler [59–61] is an alternative technique to find possible reaction paths. The approach is based on a simple and elegant technique to sample all possible states of a system subject to the constraints that the initial and final coordinates of the generated paths lie in the reactant and product states, respectively. The method is based on the scheme proposed by Pratt [62] where reactive paths constructed by a Markov chain of states are selected by Monte Carlo sampling in the space of trajectories. TPS has two parts: the sampling of reactive trajectories and the computation of the reaction rate using a type of thermodynamic integration.

For the sampling of the dynamic trajectories an initial reactive path is generated using constant energy deterministic or stochastic dynamics. Then, a state that corresponds to a randomly chosen time slice in the trajectory is modified by "shooting" moves or new states are generated by "shifting" moves. The new state(s) is accepted by an MC scheme that satisfies detailed balance provided that the new trajectory starts in the reactant and finishes in the product states. Otherwise, the trajectory is immediately rejected. By this scheme, the reactive trajectories are sampled and their analysis can provide knowledge of the reaction mechanism and give insight into ways of defining RC. The trajectories are integrated for short time since the reactive events are infrequent but once they occur, they are fast. However, the sampling may be difficult when the motion of the trajectory at the transition state is diffusive and when the reaction may happen through different channels.

The computation of the reaction rate is slightly more complicated. One may express the rate in a direct way as the number of successful trajectories n_{AB} from A to B within time t over the total number of trajectories, n_A, that start from A

$$k_{AB}(t) = \frac{n_{AB}(t)}{n_A} = \frac{\langle I_A(\xi(0)) I_B(\xi(t))\rangle}{\langle I_A(\xi(0))\rangle}. \tag{17}$$

$I_{A(B)}(\xi)$ is an indicator function defined as:

$$I_A(\xi) = \begin{cases} 1 & \text{if} \quad \xi \in A \\ 0 & \text{else} \end{cases}. \tag{18}$$

In the TPS method [60], the rate (17) can be computed using thermodynamic integration as follows: an order parameter (λ) is introduced such that the coordinate space is foliated into subsets

$$B(s) = \{q|\lambda(\xi) > s\}. \tag{19}$$

One considers an ordered sequence of values of the order parameter $\{s_0, s_1, \ldots, s_n\}$ and the corresponding subsets:

$$R^{3N} = B(s_0) \supset B(s_1) \supset \cdots \supset B(s_n) = B. \tag{20}$$

One may verify by direct substitution that rate expressed by Eq. (17) can be factorized as:

$$k_{AB}(t) = \prod_{i=1}^{n} \frac{\langle I_A(\xi(0)) I_{B_i}(\xi(t))\rangle}{\langle I_A(\xi(0)) I_{B_{i-1}}(\xi(t))\rangle}. \tag{21}$$

Each factor in the above product is the conditional probability for a trajectory to finish in subset B_k provided the trajectory reaches subset B_{k-1}. By refining the sequence B_k, one may expect to achieve a high degree of accuracy in the computation of the rate.

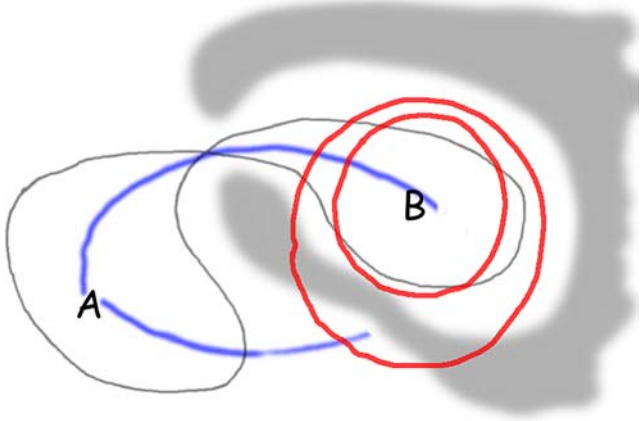

Fig. 1 Illustration of problems arising in refining the phase space in the course of thermodynamic integration. *Dashed blue lines* show the most probable paths corresponding to sets B_k and B_{k+1}. *Dashed area* signifies high values of the free-energy barrier. One observes singular changes in the most probable path connecting A and $B_{k,k+1}$ when the difference between B_k and B_{k+1} is small

In Fig. 1 we illustrate the importance of adequate description of the free energy surface by the order parameter. When this is not the case, ensemble of paths connecting A and B_i may undergo drastic changes for small changes in the value of the order parameter. This may directly affect the efficiency of the MC scheme. The fraction of paths contributing to the factor $k_{B_k B_{k+1}}$ is very small even for a small difference between sets posing a challenge for accurate sampling. The above example emphasizes the role of study of chemical mechanisms for accurate computations of the rates.

3.2 Minimal action path

A scheme proposed by Elber, Olender, Cárdenas, Ghosh and Shalloway [63–65] finds the most probable reaction path with fixed ends in the reactant and product states by minimization of an action:

$$S = \inf \int_{t_1}^{t_2} \mathrm{d}\tau \left\| \zeta \frac{\mathrm{d}\xi}{\mathrm{d}t} + \nabla W(\xi) \right\|^2. \tag{22}$$

Motivation for considering the above form of an action functional comes from Langevin equation in the large friction limit

$$\zeta \frac{\mathrm{d}\xi}{\mathrm{d}t} + \nabla W(\xi) = \mathbf{f}(t). \tag{23}$$

The most probable path (22) may reproduce the reaction mechanism when a reaction is determined by energetic factors and the majority of trajectories are found close to the minimum energy path. In high temperature where entropy effects are important, several trajectories that pass from different locations on the underlying free energy surface may be important for the reaction mechanism. A single trajectory that is found by this technique may not be representative of the reaction mechanism. Furthermore, multichannel reactions with not well-defined products may not be handled by the method. Computationally, determining the minimum of action in Eq. (22) is possible only for smooth energy surfaces. In rough energy landscapes, there would be many potential solutions that are local minima of the functional (22) making it impossible to determine the true solution.

Alternatively, instead of minimizing the action on a fixed interval, we may look for the steepest descent reaction paths $\xi(\lambda)$ such that

$$\frac{\mathrm{d}\xi(\lambda)}{\mathrm{d}\lambda} = \frac{\nabla W(\xi)}{\|\nabla W(\xi\|}. \tag{24}$$

In the above equation, we changed the parametrization of curve ξ from time t to a natural parameter λ. Differentiating the identity

$$\left\| \frac{\mathrm{d}\xi(\lambda)}{\mathrm{d}\lambda} \right\|^2 = 1 \tag{25}$$

and using Eq. (24), we deduce that for the steepest descent path the component of $\nabla W(\xi)$ perpendicular to curve ξ is zero:

$$\nabla W(\xi) \cdot \frac{\mathrm{d}\xi(\lambda)}{\mathrm{d}\lambda} = 0. \tag{26}$$

This condition

$$\nabla W(\xi)^{\perp} = 0 \tag{27}$$

is taken as the basis of minimal energy path approach [66]. For a smooth energy landscape, a solution of Eq. (27) connecting two local minima consists of a connected path between extrema of the energy profile $\nabla W(\xi) = 0$.

Ulitsky and Elber [66] proposed a parametrized path evolution technique to find piecewise steepest descent path by integrating the following equation:

$$\frac{\mathrm{d}\xi(\alpha, s)}{\mathrm{d}s} = -\nabla W(\xi)^{\perp}(\alpha, s). \tag{28}$$

In the evolution of Eq. (28) neither natural parametrization nor natural length of ξ is preserved. By including additional constraints, one arrives at the various numerical techniques of finding the steepest descent path (e.g. Nudged Elastic Band method [67,68]).

For a complex energy surface, steepest descent method produces many possible paths connecting the product and reactant states. The direct enumeration of the paths is impossible except in the simplest of cases. Direct integration of Eq. (28) will produce only one solution for a given initial condition. To overcome this problem, the finite temperature string method [58,69,70] was proposed where Eq. (28) is replaced with a stochastic differential equation

$$\frac{\mathrm{d}\xi(\alpha, s)}{\mathrm{d}s} = -\nabla W(\xi)^{\perp}(\alpha, s) + \mathbf{f}^{\perp}(\alpha), \tag{29}$$

where $\mathbf{f}^{\perp}(\alpha)$ is gaussian noise with zero component in the direction $\mathrm{d}\xi(\alpha, s)/\mathrm{d}\alpha$. This approach, while having no physical foundation, provides an efficient way for sampling multiple steepest descent pathways.

4 Applications to classical systems

4.1 Fragmentation of clusters charged with ions

Even though TPS and methods based on finding the most probable reaction path by minimization of the action are successful in unravelling reaction mechanisms, there are reactions where these approaches cannot be applied. Physical fragmentation of charged clusters is a multi-channel process without an a priori knowledge of the products. The study of fragmentation of charged liquid clusters is important for understanding fundamental questions of the physical and chemical processes in electrospray techniques [71,72] used in experiments. In our simulations, the clusters are composed of few hundreds of water molecules and several ions of the same sign and are in the liquid state. Fragmentation is controlled by two competing factors: the repulsive Coulomb interactions of ions with similar charge that tend to fragment the droplets and the opposing effect of the surface tension and the polarization of the solvent. For certain ratio of number of ions to solvent molecules, these effects give rise to a free energy barrier between the compact structure of the system that corresponds to the reactant state and the fragmented states

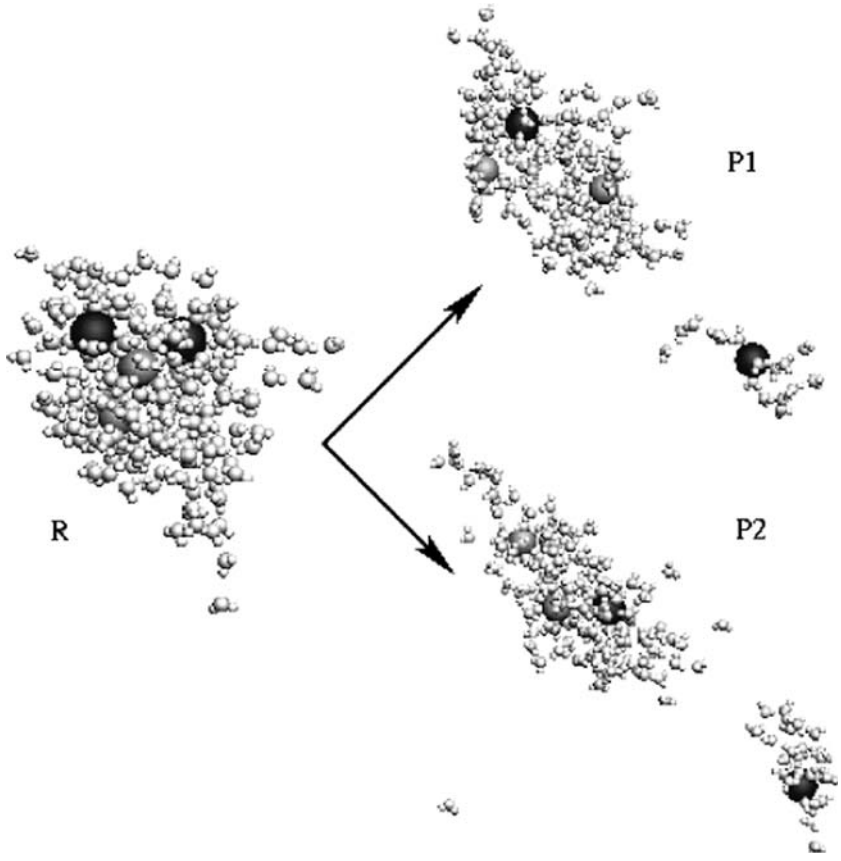

Fig. 2 Possible fragmentation channels of a cluster that is composed initially of 170 H_2O–$2Na^+$–$2K^+$ at $T = 338$ K. *Large black coloured spheres* represent K^+, *smaller dark grey*, Na^+, *small light grey spheres*, O, and *white spheres*, H. The spheres that represent K^+ and Na^+ are enlarged for visualization purposes. *R* denotes the reactant state, while P1 and P2 possible products or paths. In P1 path, the cluster fragments when it contains approximately 155 H_2O molecules and one K^+ detaches by carrying 17 H_2O molecules. In P2 path, the cluster fragments when it contains 159 H_2O molecules and one K^+ detaches carrying 26 H_2O molecules

that are the products. The products are not well defined since the ions may escape from the droplet with variable number of solvent molecules. An example of the process is shown in Fig. 2 where the snapshots are taken from direct constant energy MD simulations performed on a droplet that initially was composed of $170\ H_2O - 2Na^+ - 2K^+$.

The process is complicated because the fragmentation happens along with evaporation of the solvent. The reaction rate and the mechanism of fragmentation are found by constructing an appropriate RC, building the reversible work profile (free energy) along the RC and computing the reactive flux. Direct MD simulations for various system sizes, ions and different ways of modelling the interatomic interactions show that fragmentation occurs through configurations that look typically as the configuration shown in Fig. 3. Such bottlenecked configurations most likely correspond to the barrier top of the free energy profile.

Since the shape of the entire droplet changes during the process, the RC should be able to distinguish between the shape changes and to capture fluctuations that may lead to fragmentation. In articles [73,74], we discuss the behaviour of the radius of gyration of the droplet and dipole moment RCs and we proposed a new RC, called the transfer RC (TRC) that is able to distinguish important configurations that may lead to fragmentation. The TRC can be described in the following way: a fictitious particle is considered that is initially located on one of the ions. The fictitious particle travels from one ion to another through jumps to the intermediate solvent molecules. The fictitious particle jumps between molecules i, j found at distance r_{ij} with probability P_{ij} given by:

$$P_{ij} = A \exp(-r_{ij}^2/d^2), \tag{30}$$

where A and d are constants determined by trials (typical values used in the simulations are $A = 0.1$ and $d = \sigma_{OO}$). These parameters can be optimized so the barrier height of the free energy increases and the calculation of the rate constant becomes more efficient. Function (30) is arbitrary. One can define another expression provided that the function is less than one, its value is large for jumps to the molecules in the first coordination shell around a specific molecule and is close to zero, for jumps to the molecules in farther coordination shells. The probability of jumps to the closest molecules is large, so the fictitious particle can resolve the structure of the solvent by making small steps. The value of the RC is estimated by the average over all

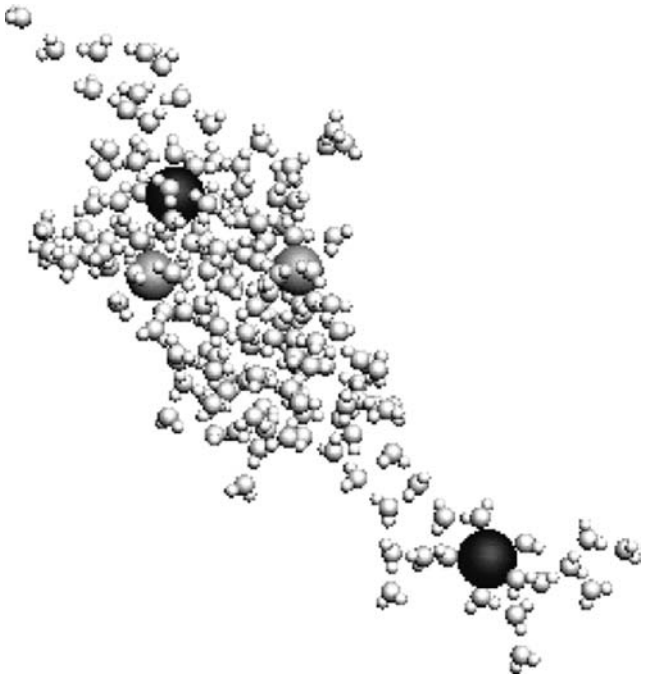

Fig. 3 Typical configuration at the bottleneck of the fragmentation that corresponds to the path P1 in Fig. 2. The colour coding of the atomic sites is the same as in Fig. 2

ion pairs of the mean first passage time for the fictitious particle to move between two ions. Hence, the RC (ξ) is defined by

$$\xi = \frac{m(m-1)}{2} \log\left(\frac{\sum_{\alpha,\beta} q_{\alpha\beta}}{m(m-1)}\right), \tag{31}$$

where α, β denote a pair of ions, m is the number of ions in the cluster and $q_{\alpha\beta}$ is the mean first passage time for the fictitious particle to move between the two ions. The TRC contains both ion and solvent spacial coordinates and it associates the motion of the ions, for example, the escape of the ion, with the location of the solvent molecules in the cluster. Using the TRC, cluster configurations with bottlenecks that separate ions are distinguished from spherical or elliptical configurations. When there are dumbbell configurations with ions found in different lobes, the fictitious particle that starts from an ion in one lobe makes many jumps to reach the ion in a different lobe since it needs to pass from the narrow bridge that connects the lobes. This yields a larger values of the RC relative to a spherical configuration.

Using the TRC, the free energy profile is estimated by umbrella sampling [47,73]. The free energy profile for a system of 180 water molecules and 4 Na^+ is shown in Fig. 4 along with typical configurations at the minima and the barrier top of the profile. In the free energy profile, we are not concerned with the fragmented configurations (products) since we are interested in the rate from the connected to fragmented states (forward process). Once the clusters fragment, they do not recombine since they are found in vacuum. The free energy profile assists us locate the dynamic bottleneck of the forward process. However, if we know the configurations that correspond to the barrier top, we can find the distribution of the products, by analyzing a large number of constant energy MD trajectories starting from the barrier top of the free energy profiles. The transition state theory estimate of the fragmentation rate is $2.5ns^{-1}$. The transmission coefficient is computed using Eq. (14) and it is shown in Fig. 5. A plateau is established after a transient time of 5 ps. The value of the transmission coefficient is very small, 0.05, which is indicative of diffusive crossing of the barrier. The full rate that includes the transmission coefficient is $1.25 \times 10^{-1} ns^{-1}$. Even though the transmission coefficient has a small value, the error bars indicate that a reliable estimate of the rate can be made.

4.2 Interconversion reactions in complexes of ions

Complexes of ions appear in aqueous clusters in the atmosphere where they may undergo many reactions [75] with implications in the reactivity of nitrogen oxides and other solutes that affect the quality of air. We chose to study the simplest complex of ions $[NaCl_2]^-$ in a cluster composed of several tens of water molecules. This ion complex undergoes quite complicated interconversion reactions between contact and solvent separated (SS) forms. The reaction can be described as:

$$\begin{aligned} Cl^-_{(1)} \mid Na^+ \mid Cl^-_{(2)} &\rightleftharpoons Cl^-_{(1)} \mid\mid Na^+ \mid Cl^-_{(2)} \\ &\rightleftharpoons Cl^-_{(1)} \mid\mid Na^+ \mid\mid Cl^-_{(2)}, \end{aligned} \tag{32}$$

where the single vertical line denotes contact ion (CI) forms while the two vertical lines SS forms. Equation (32) describes

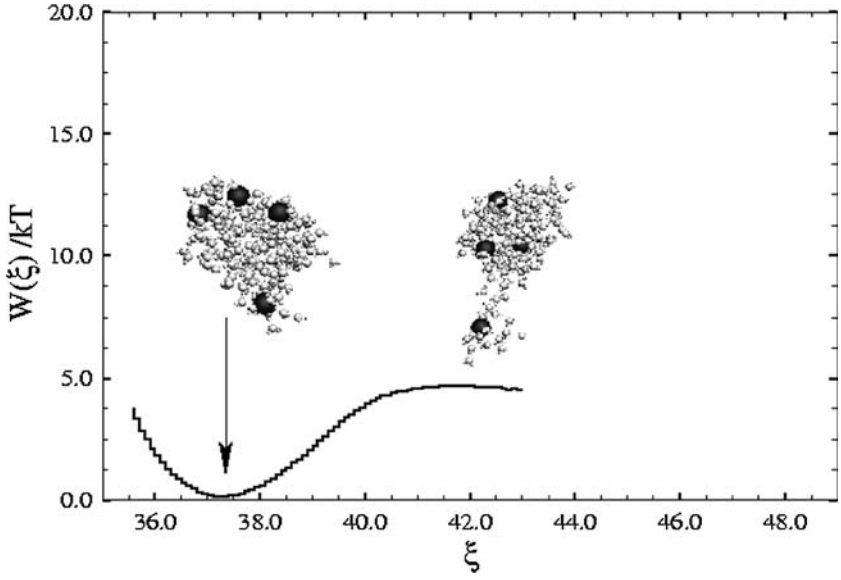

Fig. 4 Reversible work profile for a cluster containing 180 water molecules, 4 Na^+ and typical snapshots of the system at the reversible work minimum (reactant state) and barrier top. The staircase form of the graphs is due to the $P(\xi)$ histograms that enter Eq. (15). The colour coding of the atomic sites of water molecules is the same as in Fig. 2. The *large dark spheres* represent Na^+

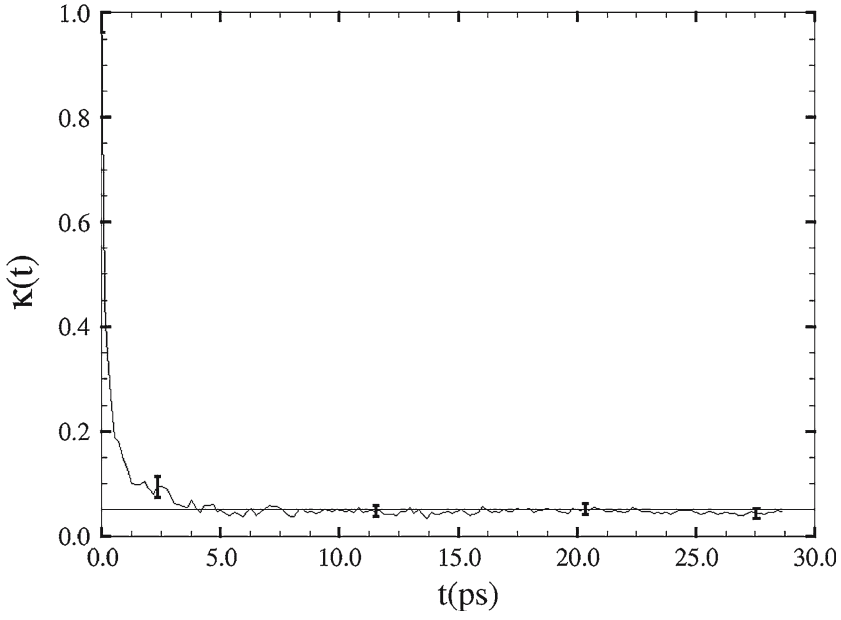

Fig. 5 Time dependent transmission coefficient for the escape of Na^+ from a cluster containing 180 water molecules and 4 Na^+

one possible channel of the interconversion process. An equivalent reaction channel exists, where $Cl^-_{(1)}$ | Na^+ | $Cl^-_{(2)}$ forms $Cl^-_{(1)}$ | Na^+ || $Cl^-_{(2)}$ in the first step and subsequently, $Cl^-_{(1)}$ || Na^+ || $Cl^-_{(2)}$, is formed in the second step of 32. The underlying free energy surface associated with the interionic distances $Cl^-_{(2)}$–Na^+ (r_1) and $Cl^-_{(2)}$–Na^+ (r_2) as RCs is shown schematically in Fig. 6 [75]. The topography of the free energy surface, motivates the use of $\xi(r_1, r_2) = (r_1^2 + r_2^2)^{1/2}$ as one-dimensional RC that takes into account both reaction channels. The reactive flux method can be used with the one-dimensional RC. It is found that the transmission coefficient has diffusive character and, therefore, there are considerable corrections to the TST estimate of the rate coefficient. Analysis of the trajectories shows that the shape fluctuations of the cluster and the diffusive motion of the surface solvent molecules contribute to the relaxation processes in the time-dependent transmission coefficient and this may possibly lead to the breakdown of simple first-order kinetics.

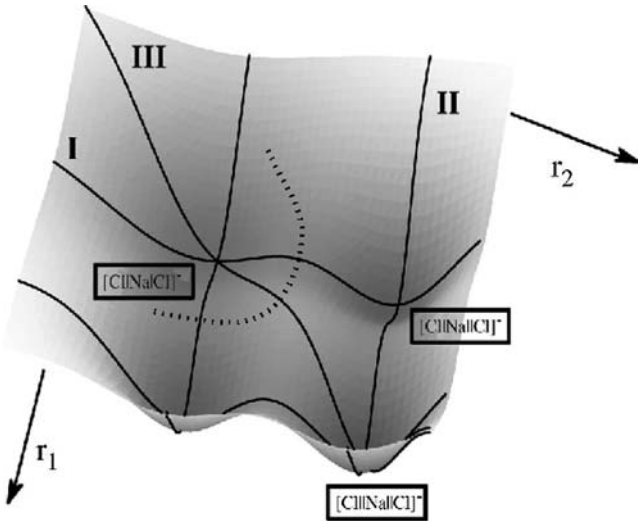

Fig. 6 Schematic representation of the three-dimensional free energy surface of $[NaCl_2]^-$ in the presence of 70 water molecules. The *solid lines* represent the sections of the free energy profile along which the free energy was computed using the blue moon ensemble. Paths I and II correspond to the first and second inter conversion reactions in Eq. (32), respectively. The diagonal path III passes from a global maximum energy state in the free energy surface and it is unlikely to be taken. The *dashed line* represents the region where the transition states may be found and motivates the construction of a one-dimensional RC

5 Conclusion

Development of methods for the computation of reaction rates in condensed phase is a field that is still in flux even though it has a history of several decades. Here, only aspects of the classical expressions for the rate coefficients and methods of determining the reaction mechanisms were presented. Because of the breadth of the field, important topics that are broad on their own, such as methods for computing the rate coefficient when the barrier crossing is diffusive, multidimensional transition state theory, quantum expressions for rate constants and their form when the system is treated by mixed quantum-classical dynamics were left out of this review. The rapidly growing computer technology allowed for the schemes presented here to be applied in realistic situations. For instance, applications of TPS has given insight into autoionization of bulk water [76] and interconversion reactions of ion pairs [77,78]. Minimization of the action schemes have been applied in the study of conformational changes of biological molecules [65]. Since the problems in nature are complex and the methodologies have limitations, it may be useful to use a combination of techniques and information from experiments to study reactions of high complexity. Treatment of multichannel reactions or reactions where the products are not known in advance is a challenging problem in the field. Furthermore, even though sampling of reactive paths or the most probable path may reveal the fluctuations that determine the reaction, it is still challenging to model these effects in an RC. The foundations for many of the numerical schemes we use today had been laid before the advent of ultrafast computers. Microscopic expressions of rate constants such as the reactive-flux scheme and kinetic models for reactions that take into account collisions and the effect of short-and long-range interactions in the colliding particles were established in the period between 1960 and 1980. Extensive calculations revealed the complexity of the chemical processes and showed the limitations of the existing schemes. This motivated the search for alternative methods that may focus on sampling of the reactive paths or finding the most probable reaction path. Since computer technology is improving rapidly, one may think of developing techniques that include the advantages of fast computing in their construction.

Appendix A: Generalized Langevin equation for a dynamical variable

The starting point for the derivation of a generalized Langevin equation for any dynamical variable is the Liouville equation for the evolution of a classical dynamical variable G,

$$\frac{\mathrm{d}G(t)}{\mathrm{d}t} = \left(\frac{\partial H}{\partial q_i}\frac{\partial G}{\partial p_i} - \frac{\partial H}{\partial p_i}\frac{\partial G}{\partial q_i}\right) = \{H, G\} = iLG(t), \quad (33)$$

where the second equality defines the Poisson bracket and the third equality defines the Liouville operator L. H is the Hamiltonian of the system while q_i, p_i denote the positions and momenta of the atoms. Since the Liouville operator does not depend on time, the solution of this equation is

$$G(t) = \mathrm{e}^{iLt}G(0). \quad (34)$$

In applications of the projection operator method, the dynamical variables of the system are separated into two sets: a set of slowly varying variables and a set of rapidly varying variables. A projection operator is constructed that projects the full classical dynamics onto subspace of the slow variables. Let G be the set of slow variables. A projection operator onto G can be constructed in the following way:

$$P = (\cdot, G)(G, G)^{-1}G, \quad (35)$$

When P acts on any dynamical variable, the result is proportional to G: it projects out of the variable, the slowly varying component that is parallel to G. The complement of the projection operator P is $Q = 1 - P$ that projects onto the space orthogonal to G.

The generalized Langevin equation is now easily written for the set G of dynamical variables using just two pieces of information. First, since by definition $P + Q = 1$, we may write the Liouville operator as $iL = iPL + iQL \equiv iL_P + iL_Q$. Second, for any operators C and D we have the operator identity [79],

$$\mathrm{e}^{(C+D)t} = \int_0^t \mathrm{d}\tau\ \mathrm{e}^{(C+D)\tau}C\mathrm{e}^{D(t-\tau)} + \mathrm{e}^{Dt}. \quad (36)$$

Then, using these results, we can write

$$\frac{\mathrm{d}G(t)}{\mathrm{d}t} = \mathrm{e}^{iLt} iLG = \mathrm{e}^{iLt} iL_P G + \mathrm{e}^{iLt} iL_Q G$$
$$= \mathrm{e}^{iLt} iL_P G + \int_0^t \mathrm{d}\tau\ \mathrm{e}^{iL\tau} iL_P \mathrm{e}^{iL_Q(t-\tau)} iL_Q G$$
$$+ \mathrm{e}^{iL_Q t} iL_Q G, \qquad (37)$$

where we let $C = iL_P$ and $D = iL_Q$. If we now substitute the definition of the projection operator in Eq. (35) we obtain the generalized Langevin equation Eq. (5).

Acknowledgements I thank the Natural Sciences and Engineering Research Council of Canada (NSERC) for funding this research through the AGENO grant.

References

1. Bonn M, Funk S, Hess CH, Denzler DN, Stampfl C Scheffler M Wolf M, Ertl G (1999) Science 285:1042
2. Neumark DM (2004) In: Yang X, Liu K (eds) In: Advanced series in physical chemistry, modern trends in chemical reaction dynamics (part 1). vol. 14. World Scientific Publishing, Singapore, p 453
3. Neumark DM (2005) Phys Chem Chem Phys 7(14):433
4. Folmer DE, Poth L, Wisniewski ES, Castleman AW (1998) J Chem Phys Lett 287:1
5. Zewail AH (1994) In: Femtochemistry: ultrafast dynamics of the chemical bond. World Scientific, Singapore
6. Stolow A, Bragg AE, Neumark DM (2004) Chem Rev 104:1719
7. Dermota TE, Zhong Q, Castleman AWJ (2004) Chem Rev 104:1861
8. Elles CG, Cox MJ, Barnes GL, Crim FF (2004) J Phys Chem A 108:10973
9. Frenkel D, Smit B (2001) In: Understanding molecular simulation: from algorithms to applications, 2nd edn. Elsevier, Amsterdam
10. Sachs C, Hildebrand M, Volkening S, Wintterlin J, Ertl G (2001) Science 293:1635
11. Hänggi P, Talkner P, Borkovec M (1990) Rev Mod Phys 62:251–341
12. Cramer CJ, Truhlar DG (1999) Chem Rev 99:2161–2200
13. Berne BJ, Ciccotti G, Coker DF (eds) (1998) In: Classical and quantum dynamics in condensed phase simulations. World Scientific Publishing, Singapore
14. Kapral R, Ciccotti G (2005) In: Dykstra C (ed) Theory and applications of computational chemistry. Elsevier BV, Amsterdam
15. Truhlar DG, Garrett BC, Klippenstein SJ (1996) J Phys Chem 100:12771–12800
16. Smoluchowski MV (1915) Ann Phys 48:1103
17. Kramers HA (1940) Physica 7:284
18. Grote F, Hynes JT (1980) J Chem Phys 73:2715
19. Grote F, Hynes JT (1981) J Chem Phys 74:4465
20. Grote F, Hynes JT (1981) J Chem Phys 75:2191
21. Adelman SA, Muralidhar R (1991) J Chem Phys 95:2752
22. Gertner BJ, Wilson KR, Hynes JT (1989) J Chem Phys 90:3537
23. Ciccotti G, Hynes JT, Kapral R (1990) J Chem Phys 93:7137
24. Zwanzig R (2001) Nonequilibrium statistical mechanics. Oxford University Press, New York
25. Zhou H-X, Zwanzig R (2002) J Phys Chem 106:7562
26. van der Zwan G, Hynes JT (1983) J Phys Chem 78:4174
27. van der Zwan G, Hynes JT (1984) Chem Phys 80:21
28. Dakhnovskii YI, Ovchinnikov AA (1985) Phys Lett 113A:147
29. Pollak E (1986) J Chem Phys 85:865
30. Hänggi P, Weiss U (1984) Phys Rev A 29:2265
31. Straub JE, Berne BJ (1986) J Chem Phys 85:2999
32. Talkner P, Hänggi P (eds) (1995) In: New trends in Kramers' reaction rate theory. Kluwer, Dordrecht
33. Yamamoto T (1960) J Chem Phys 33:281
34. Chandler D (1978) J Chem Phys 68:2959
35. Bennet CH (1977) In: Christofferson RE (ed) Algorithms for chemical computations. ACS Symp. Ser. No. 46. American Chemical Society, Washington, DC. p 63
36. Kapral R (1981) Adv Chem Phys 48:71
37. Chandler D (1987) In: Introduction to modern statistical mechanics. Oxford University Press, New York
38. Voth GA, Chandler D, Miller WH (1989) J Phys Chem 93:7009
39. Ciccotti G, Ferrario M, Laria D, Kapral R (1995) In: Manghi F, Reatto L (eds) Progress in computational physics of matter: methods, software and applications
40. Kapral R (1972) J Chem Phys 56:1842
41. Zwanzig R (1961) Phys Rev 124:983
42. Zwanzig R (1965) Ann Rev Phys Chem 16:67
43. Mori H (1965) Prog Theor Phys 34:399
44. Mori H (1965) Prog Theor Phys 33:423
45. Kapral R, Consta S, McWhirter L (1998) In: Berne BJ, Ciccotti G, Coker DF (eds) Classical and quantum dynamics in condensed phase simulations, pp 587
46. Schenter GK, Garrett BC, Truhlar DG (2003) J Chem Phys 119:5828–5833
47. Torrie GM, Valleau JP (1977) J Comp Phys 23:187
48. Carter EA, Ciccotti G, Hynes JT, Kapral R (1989) Chem Phys Lett 156:472
49. Sprik M, Ciccotti G (1998) J Chem Phys 109:7737
50. Sergi A, Ciccotti G, Falconi M, Desideri A, Ferrario M (2002) J Chem Phys 116:6329
51. Coluzza I, Sprik M, Ciccotti G (2003) Mol Phys 101:2885
52. Ciccotti G, Kapral R, Vanden-Eijnden E (2005) Phys Chem Chem Phys (to appear)
53. Laio A, Parrinello M (2002) Proc Natl Acad Sci USA 99:12562
54. Iannuzzi M, Laio A, Parrinello M (2003) Phys Rev Lett 90:238302
55. Zeldovich J (1942) J Expr Theor Phys (Russia) 12:525
56. ten Wolde PR, Ruiz-Montero MJ, Frenkel D (1996) J Chem Phys 104:9932
57. Truhlar DG, Gao JL, Garcia-Viloca M, Alhambra C, Corchado J, Sanchez ML, Poulsen TD (2004) Int J Quant Chem 100:1136–1152
58. Weinan E, Vanden-Eijnden E (2004) In: Attinger S, Koumoutsakos P (eds) Multiscale modelling and simulation, Lecture notes in computational science and engineering, vol 39. Springer, Berlin Heidelberg New York
59. Bolhuis PG, Dellago C, Chandler D (1998) Faraday Discuss Chem Soc 110:421
60. Dellago C, Bolhuis P, Geissler PL (2002) Adv Chem Phys 123:1
61. Bolhuis PG, Chandler D, Dellago C, Geissler PL (2002) Ann Rev Phys Chem 53:291
62. Pratt LR (1986) J Chem Phys 85:5045
63. Elber R, Meller J, Olender R (1999) J Phys Chem 103:899
64. Elber R, Shalloway D (2002) J Chem Phys 112:5539
65. Elber R, Ghosh A, Cárdenas A (2002) Acc Chem Res 35:396
66. Ulitsky A, Elber R (1990) J Chem Phys 92:1510
67. Jónsson H, Mills G, Jacobsen KW (1998) In: Berne BJ, Ciccotti G, Coker DF (eds) Classical and quantum dynamics in condensed phase simulations, pp 385
68. Henkelman G, Jónsson H (2000) J Chem Phys 113:9978
69. Weinan E, Ren W, Vanden-Eijnden E (2002) Phys Rev B 66:05230
70. Weinan E, Ren WQ, Vanden-Eijnden E (2005) J Phys Chem 109(14):6688
71. Kebarle P (2000) J Mass Spectrom 35:804
72. Fenn JB, Rosell J, Meng CK (1997) J Am Soc Mass Spectrom 8:1147
73. Consta S (2002) J Mol Struct Theochem 591:131
74. Consta S, Mainer KR, Novak W (2003) J Chem Phys 119:10125
75. Consta S, Kapral R (1999) J Chem Phys 111:10183
76. Geissler PL, Dellago C, Chandler D, Hutter J, Parrinello M (2001) Sience 291:2121
77. Geissler PL, Dellago C, Chandler D (1999) J Phys Chem B 103:3706
78. Truhlar DG, Garrett BC (2000) J Phys Chem B 104:1069–1072
79. McQuarrie DA (1976) In: Statistical mechanics. HarperCollins, New York