

## 5.4

# SIMULATING REACTIONS THAT OCCUR ONCE IN A BLUE MOON

Giovanni Ciccotti<sup>1</sup>, Raymond Kapral<sup>2</sup>, and Alessandro Sergi<sup>2</sup>

<sup>1</sup>*INFN and Dipartimento di Fisica, Università “La Sapienza”,  
Piazzale Aldo Moro, 2, 00185 Roma, Italy*

<sup>2</sup>*Chemical Physics Theory Group, Department of Chemistry,  
University of Toronto, Toronto, ON, Canada, M5S 3H6*

The computation of the rates of condensed phase chemical reactions poses many challenges for theory. Not only do condensed phase systems possess a large number of degrees of freedom so that computations are lengthy, but typically chemical reactions are activated processes so that transitions between metastable species states are rare events that occur on time scales long compared to those of most molecular motions. This time scale separation makes it almost impossible to determine reaction rates by straightforward simulation of the equations of motion. Furthermore, condensed phase reactions often involve collective degrees of freedom where the solvent participates in an important way in the reactive process. Consequently, the choice of a reaction coordinate to describe the reaction is often far from trivial task.

Various methods for determining reaction paths have been devised (see Refs. [1, 2] and references therein). These methods have the goal of determining how the system passes from one metastable state to another and thus finding the reaction path or reaction coordinate. In many situations one has some knowledge of how to construct a reaction coordinate (or set of reaction coordinates) for a particular physical problem. One example is the use of the many-body solvent polarization reaction coordinate to describe electron or proton transfer in solution. In almost all situations investigated to date the dynamics of condensed phase activated reaction rates can be described in terms of a small number of reaction coordinates (often involving collective degrees of freedom).

In this chapter, we describe how to simulate the rates of activated chemical reactions that occur on slow time scales, assuming that some set of suitable reaction coordinates is known. In order to compute the rates of rare reactive

events we need to be able to sample regions of configuration space that are rarely visited since the interconversion between reactants and products entails passage through such regions of low probability. We show that by applying holonomic constraints to the reaction coordinate in a molecular dynamics simulation we can force the system to visit unfavorable configuration space regions. Through such constraints we can generate an ensemble of configurations (the Blue Moon ensemble) that allows one to efficiently estimate the rate constant for activated chemical processes [3].

## 1. Reactive Flux Correlation Function Formalism

We begin with a sketch of the reactive flux correlation function formalism in order to specify the quantities that must be computed to obtain the reaction rate constant. In order to simplify the notation, we consider a molecular system containing  $N$  atoms with Hamiltonian  $H = K(\mathbf{p}) + V(\mathbf{r})$ , where  $K(\mathbf{p})$  is the kinetic energy,  $V(\mathbf{r})$  is the potential energy and  $(\mathbf{p}, \mathbf{r})$  denotes the  $6N$  momenta and coordinates defining the phase space of the system. A chemical reaction  $A \rightleftharpoons B$  is assumed to take place in the system. The reaction dynamics is described phenomenologically by the mass action rate law

$$\frac{d\overline{n_A}(t)}{dt} = -k_f\overline{n_A}(t) + k_r\overline{n_B}(t), \quad (1)$$

where  $\overline{n_A}(t)$  is the mean number density of species A. The task is to compute the forward  $k_f$  and reverse  $k_r = k_f K_{\text{eq}}^{-1}$  ( $K_{\text{eq}}$  is the equilibrium constant) rate constants by molecular dynamics simulation. (The formalism is easily generalized to other reaction schemes.) To this end, we assume that the progress of the reaction can be characterized on a microscopic level by a scalar reaction coordinate  $\zeta(\mathbf{r})$  which is a function of the positions of the particles in the system. A dividing surface at  $\zeta^\ddagger$  serves to partition the configuration space of the system into two A and B domains that contain the metastable A and B species. The microscopic variable corresponding to the fraction of systems in the A domain is  $n_A(\mathbf{r}) = \theta(\zeta^\ddagger - \zeta(\mathbf{r}))$ , where  $\theta$  is the Heaviside function. Similarly, the fraction of systems in the B domain is  $n_B(\mathbf{r}) = \theta(\zeta(\mathbf{r}) - \zeta^\ddagger)$ . The time rate of change of  $n_A(\mathbf{r})$  is

$$\dot{n}_A(\mathbf{r}) = -\dot{\zeta}(\mathbf{r})\delta(\zeta(\mathbf{r}) - \zeta^\ddagger). \quad (2)$$

The rate at which the A and B species interconvert can be determined from the well-known reactive flux formula for the rate constant [4–6]. Using this formalism the time-dependent forward rate coefficient can be expressed in terms of the equilibrium correlation function of the initial flux of A with the

A species density at time  $t$  as

$$k_f(t) = \frac{1}{n_A^{\text{eq}}} \langle \dot{n}_A(\mathbf{r}) n_A(\mathbf{r}, t) \rangle = \frac{1}{n_A^{\text{eq}}} \left\langle \left( \dot{\zeta} \delta(\zeta(\mathbf{r}) - \zeta^\ddagger) \right) \theta(\zeta(\mathbf{r}(t)) - \zeta^\ddagger) \right\rangle. \quad (3)$$

Here, the angular brackets denote an equilibrium canonical average,  $\langle \dots \rangle = Q^{-1} \int d\mathbf{r} d\mathbf{r}' \exp\{-\beta H\} \dots$ , where  $Q$  is the partition function and  $n_A^{\text{eq}}$  is the equilibrium density of species A. The forward rate constant can be determined from the plateau value of this time-dependent forward rate coefficient [6].

We can separate the static and dynamic contributions to the rate coefficient by multiplying and dividing each term on the right-hand side of Eq. (3) by  $\langle \delta(\zeta_a(\mathbf{r}) - \zeta^\ddagger) \rangle$  to obtain

$$k_f(t) = \left\{ \frac{\left\langle \left( \dot{\zeta} \delta(\zeta(\mathbf{r}) - \zeta^\ddagger) \theta(\zeta^\ddagger - \zeta(\mathbf{r}(t))) \right) \right\rangle}{\langle \delta(\zeta(\mathbf{r}) - \zeta^\ddagger) \rangle} \right\} \left\{ \frac{\langle \delta(\zeta(\mathbf{r}) - \zeta^\ddagger) \rangle}{n_A^{\text{eq}}} \right\}. \quad (4)$$

The equilibrium average  $\langle \delta(\zeta(\mathbf{r}) - \zeta^\ddagger) \rangle = P(\zeta^\ddagger)$  is the probability density of finding the value of the reaction coordinate  $\zeta(\mathbf{r}) = \zeta^\ddagger$ .

We may introduce the free energy  $\mathcal{W}(\zeta')$  associated with the reaction coordinate by the definition  $\mathcal{W}(\zeta') = -\beta^{-1} \ln(P(\zeta')/P_u)$ , where  $P_u$  is a uniform probability density of  $\zeta'$ . For an activated process the free energy will have the form shown schematically in Fig. 1. A high free energy barrier at  $\zeta = \zeta^\ddagger$  separates the metastable reactant and product states. The equilibrium density of species A is

$$\begin{aligned} n_A^{\text{eq}} &= \left\langle \theta(\zeta^\ddagger - \zeta(\mathbf{r})) \right\rangle = \int d\zeta' \theta(\zeta^\ddagger - \zeta') \langle \delta(\zeta(\mathbf{r}) - \zeta') \rangle \\ &= \int_{\zeta' < \zeta^\ddagger} d\zeta' P(\zeta'). \end{aligned} \quad (5)$$

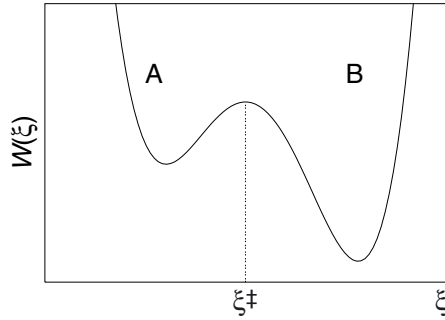


Figure 1. Sketch of the free energy versus  $\zeta$  showing the free energy maximum at  $\zeta = \zeta^\ddagger$  and specification of the A and B domains.

Using these results the expression for the time-dependent rate coefficient may be written as

$$k_f(t) = \left\langle \dot{\zeta} \theta(\zeta^\ddagger - \zeta(\mathbf{r}(t))) \right\rangle_{\zeta^\ddagger}^{\text{cd}} \left\{ \frac{e^{-\beta \mathcal{W}(\zeta^\ddagger)}}{\int_{\zeta' < \zeta^\ddagger} d\zeta' e^{-\beta \mathcal{W}(\zeta')}} \right\}, \quad (6)$$

where  $\langle \dots \rangle_{\zeta^\ddagger}^{\text{cd}}$  defines an average conditional on  $\zeta(\mathbf{r}) = \zeta^\ddagger$

$$\langle \dots \rangle_{\zeta'}^{\text{cd}} = \frac{\langle \dots \delta(\zeta(\mathbf{r}) - \zeta') \rangle}{\langle \delta(\zeta(\mathbf{r}) - \zeta') \rangle}. \quad (7)$$

Often it is useful to represent the results in terms of the time-dependent transmission coefficient  $\kappa(t)$  which is defined as  $\kappa(t) = k_f(t)/k_f^{\text{TST}}$ , where the transition-state theory value of the rate constant is given by the limit  $t \rightarrow 0_+$  of Eq. (3) as [5]

$$k_f^{\text{TST}} = \frac{1}{n_A} \left\langle \dot{\zeta} \delta(\zeta(\mathbf{r}) - \zeta^\ddagger) \theta(\dot{\zeta}) \right\rangle. \quad (8)$$

The transmission coefficient  $\kappa(t)$  measures the deviations from  $k_f^{\text{TST}}$  due to dynamical recrossing events.

From Eq. (6) we see that the computation of the rate coefficient requires the calculation of conditional averages depending on specified, rarely visited, values of the reaction coordinate. The ensemble of such configurations which are visited ‘‘once in a blue moon’’ is termed the blue moon ensemble [3]. In Section 2, we describe how the conditional averages which play an essential role in the computation of the rate constant may be determined from constrained molecular dynamics trajectories which allow one to efficiently sample rarely visited regions of configuration space.

## 2. The Blue Moon Ensemble

The presence of the delta function  $\delta(\zeta(\mathbf{r}) - \zeta')$  in the conditional equilibrium averages fixes the reaction coordinate to have the specified value  $\zeta'$ . We shall now show that such conditional averages of observables depending only on configuration space variables can be computed by applying holonomic constraints to the equations of motion. For simplicity, we assume that no other bond constraints are present but this more general case is easily treated [3]. To this end, we consider a system described by the Cartesian coordinates  $(\mathbf{r}, \mathbf{p})$  subject to a holonomic constraint  $\sigma(\mathbf{r}) = \zeta(\mathbf{r}) - \zeta' = 0$  on the reaction coordinate. When there are constraints one usually introduces a set of generalized coordinates  $\mathbf{q}$  and conjugate momenta  $\mathbf{p}^q$  such that  $\mathbf{r} = \mathbf{r}(\mathbf{q})$ . In general, it is not possible to invert this relation since there are more  $\mathbf{r}$  coordinates than  $\mathbf{q}$

coordinates. However, by adding the expression for the constraint there is an extra generalized coordinate and the one to one correspondence  $\mathbf{r} = \mathbf{r}(\mathbf{q}, \sigma)$  is recovered.

The statistical mechanics of the system may be formulated in terms of the generalized coordinates  $\mathbf{q}$ ; however, it is useful to have an equivalent formulation in terms of the original Cartesian coordinates. The dynamics of the system is described in Cartesian coordinates by the Lagrangian

$$\mathcal{L}(\mathbf{r}, \dot{\mathbf{r}}) = K(\dot{\mathbf{r}}) - V(\mathbf{r}) = \sum_{i=1}^N \frac{1}{2} m_i \dot{\mathbf{r}}_i^2 - V(\mathbf{r}), \quad (9)$$

to which we add the constraint  $\sigma = 0$ . The set of  $3N - 1$  generalized coordinates  $\mathbf{q}$  plus  $\sigma$  can be taken as a new set of equivalent coordinates denoted collectively by  $\mathbf{u}$ . We have  $\mathbf{r} = \mathbf{r}(\mathbf{q}, \sigma) = \mathbf{r}(\mathbf{u})$ , or  $\mathbf{u} = \mathbf{u}(\mathbf{r})$ . In the new variables the Lagrangian is given by

$$\mathcal{L}'(\mathbf{u}, \dot{\mathbf{u}}) = \frac{1}{2} \dot{\mathbf{u}}^T \mathbf{M} \dot{\mathbf{u}} - V'(\mathbf{u}), \quad (10)$$

where  $\mathbf{u}^T$  is the transpose of vector  $\mathbf{u}$  and  $\mathbf{M} = \mathbf{J}^T \mathbf{m} \mathbf{J}$  is the metric matrix with elements given by

$$M_{\mu\nu} = \sum_{i=1}^N m_i \frac{\partial \mathbf{r}_i}{\partial u_\mu} \cdot \frac{\partial \mathbf{r}_i}{\partial u_\nu}. \quad (11)$$

Here,  $\mathbf{J}$  is the Jacobian matrix of the transformation  $\mathbf{r} \leftrightarrow \mathbf{u}$  and  $\mathbf{m}$  is a diagonal matrix of the masses. The Lagrangian of the constrained motion is easily obtained by putting  $\sigma = \dot{\sigma} = 0$

$$\mathcal{L}_c(\mathbf{q}, \dot{\mathbf{q}}) \equiv \mathcal{L}'(\mathbf{q}, \sigma = 0, \dot{\mathbf{q}}, \dot{\sigma} = 0). \quad (12)$$

To derive the statistical mechanical ensemble we need the Hamiltonian description of the dynamical system. The Hamiltonian corresponding to the description with coordinates  $\mathbf{u}$  is given by

$$H'(\mathbf{u}, \mathbf{p}^u) = \frac{1}{2} \mathbf{p}^{uT} \mathbf{M}^{-1} \mathbf{p}^u + V'(\mathbf{u}), \quad (13)$$

where

$$\mathbf{p}^u = \frac{\partial \mathcal{L}'}{\partial \dot{\mathbf{u}}} = \mathbf{M} \dot{\mathbf{u}}. \quad (14)$$

The inverse of the metric matrix  $\mathbf{M}^{-1}$  can be written explicitly as

$$\left(\mathbf{M}^{-1}\right)_{\mu\nu} = \sum_{i=1}^N \frac{1}{m_i} \frac{\partial u_\mu}{\partial \mathbf{r}_i} \cdot \frac{\partial u_\nu}{\partial \mathbf{r}_i}. \quad (15)$$

To obtain the constrained motion we have to compute the Hamiltonian at  $\sigma = 0$  and  $p^\sigma$  satisfying the constraints  $\sigma = \dot{\sigma} = 0$ . Since

$$\dot{\sigma} = \left(\mathbf{M}^{-1} \mathbf{p}^u\right)_{3N} = \mathbf{E} \mathbf{p}^q + Z p^\sigma, \quad (16)$$

where the subscript  $3N$  denotes that element of the vector  $\mathbf{M}^{-1}\mathbf{p}^u$  and  $\mathbf{E}$  and  $Z$  are submatrices in the block form of  $\mathbf{M}^{-1}$

$$\mathbf{M}^{-1} = \begin{bmatrix} \Delta & \mathbf{E} \\ \mathbf{E}^T & Z \end{bmatrix}, \quad (17)$$

the above condition corresponds to taking  $p^\sigma = -\tilde{Z}^{-1}\tilde{\mathbf{E}}\mathbf{p}^q$ , where the tilde means that the matrices have to be evaluated at  $\sigma=0$ . The explicit form of  $Z$  is needed below and is given by

$$Z = \sum_{i=1}^N \frac{1}{m_i} \frac{\partial \sigma}{\partial \mathbf{r}_i} \cdot \frac{\partial \sigma}{\partial \mathbf{r}_i}. \quad (18)$$

The constrained Hamiltonian may now be written as

$$H_c(\mathbf{q}, \mathbf{p}^q) \equiv H'(\mathbf{q}, \sigma=0, \mathbf{p}^q, p^\sigma = -\tilde{Z}^{-1}\tilde{\mathbf{E}}\mathbf{p}^q). \quad (19)$$

Note that Eq. (16) implies  $p^\sigma + Z^{-1}\mathbf{E}\mathbf{p}^q = Z^{-1}\dot{\sigma}$ .

Letting  $\rho_c(\mathbf{q}, \mathbf{p}^q)$  be the probability density for the constrained dynamical system, we have

$$\begin{aligned} \rho_c(\mathbf{q}, \mathbf{p}^q) d\mathbf{q} d\mathbf{p}^q &= \rho'(\mathbf{u}, \mathbf{p}^u) \delta(\sigma) \delta(p^\sigma + Z^{-1}\mathbf{E}\mathbf{p}^q) d\mathbf{u} d\mathbf{p}^u \\ &= \rho(\mathbf{r}, \mathbf{p}^r) \delta(\sigma) \delta(Z^{-1}\dot{\sigma}) d\mathbf{r} d\mathbf{p}^r \equiv \rho_\xi(\mathbf{r}, \mathbf{p}^r). \end{aligned} \quad (20)$$

In the penultimate equality, we used the fact that the point contact transformation  $(\mathbf{u}, \mathbf{p}^u) \leftrightarrow (\mathbf{r}, \mathbf{p}^r)$  is a canonical, phase space volume conserving, transformation [7]. We may rewrite this probability density as

$$\rho_\xi(\mathbf{r}, \mathbf{p}^r) = \rho_\xi^r(\mathbf{r}) \rho_\xi^p(\mathbf{p}^r|\mathbf{r}), \quad (21)$$

where the configurational probability density,  $\rho_\xi^r(\mathbf{r})$ , is obtained by performing the momentum integration of the full probability density  $\rho_\xi(\mathbf{r}, \mathbf{p}^r)$  and the conditional probability density of the momenta given the configuration,  $\rho_\xi^p(\mathbf{p}^r|\mathbf{r})$ , is defined in Eq. (21). In the canonical ensemble, the configurational probability density is given by

$$\rho_\xi^r(\mathbf{r}) d\mathbf{r} = Q_c^{-1} |Z|^{1/2} e^{-\beta V(\mathbf{r})} \delta(\sigma) d\mathbf{r} = Q_c^{-1} |Z|^{1/2} e^{-\beta V(\mathbf{r})} \delta(\xi(\mathbf{r}) - \xi') d\mathbf{r}, \quad (22)$$

where  $Q_c$  is the partition function of the constrained system. The factor  $|Z|^{(1/2)}$  arises from performing the momentum integration of Eq. (20). The conditional probability density of the momenta given the configuration is

$$\rho_\xi^p(\mathbf{p}^r|\mathbf{r}) d\mathbf{p}^r = |Z|^{-1/2} e^{-\beta K} \delta(Z^{-1}\dot{\sigma}) d\mathbf{p}^r. \quad (23)$$

The physical interpretation of  $Z$  which enters this expression for the probability density has been discussed by several authors [8, 9] and has its origin in the restriction imposed in momentum space by the constraint  $\sigma = 0$  which, holding at all times, implies that the generalized velocity  $\dot{\sigma}$  must vanish.

The configurational probability density in Eq. (22) should be compared with the joint configurational probability density to be at  $\mathbf{r}$  and at  $\zeta = \zeta'$

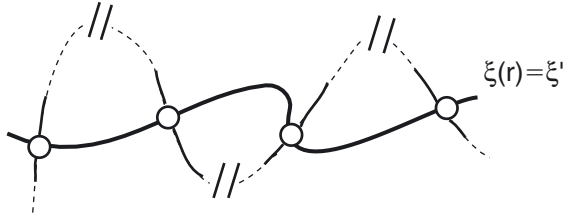
$$\rho'(\mathbf{r}) \delta(\zeta(\mathbf{r}) - \zeta') d\mathbf{r} = Q^{-1} e^{-\beta V(\mathbf{r})} \delta(\zeta(\mathbf{r}) - \zeta') d\mathbf{r}. \quad (24)$$

We may express the conditional average of any configurational property of our system in terms of the  $\zeta$ -constrained ensemble introduced above. While the value  $\zeta'$  we wish to sample is rare in the original ensemble, only configurations with  $\zeta = \zeta'$  are sampled in the  $\zeta$ -constrained ensemble. This feature is illustrated schematically in Fig. 2. By comparison of Eqs. (22) and (24) we may write

$$\frac{\langle O(\mathbf{r}) \delta(\zeta(\mathbf{r}) - \zeta') \rangle}{\langle \delta(\zeta(\mathbf{r}) - \zeta') \rangle} = \frac{\langle |Z|^{-1/2} O(\mathbf{r}) \rangle_{\zeta'}}{\langle |Z|^{-1/2} \rangle_{\zeta'}}, \quad (25)$$

where the observable  $O(\mathbf{r})$  is any function of the configuration space,  $\langle \dots \rangle$  denotes a canonical ensemble average and  $\langle \dots \rangle_{\zeta'}$  denotes an average over the constrained ensemble with  $\zeta = \zeta'$ . Equation (25) allows one to estimate the conditional average on the left-hand side in terms of averages in the constrained ensemble with evident statistical advantages.

The constrained ensemble we have constructed can be generalized to give the biased (whose bias may be removed) configurational sample determined above and the correct distribution of momenta. This *Blue Moon* ensemble can be easily obtained by multiplying the  $\zeta$ -constrained configurational probability



*Figure 2.* Schematic representation of the sampling procedure in the Blue Moon ensemble. The bold line depicts the constrained ( $\zeta(\mathbf{r}) = \zeta'$ ) dynamical evolution in phase space. The unconstrained natural evolution of the system is shown as a dashed line. The open circles represent common points in configuration space which are the initial conditions of the activated trajectory sampling. These points are not real crossings in phase space since the two trajectories differ in the momentum space. Interruptions of the dashed line denote lengthy segments in the natural trajectory and indicate that “crossings” are rare events. The dynamics represented by the solid line segments of the unconstrained trajectory in the vicinity of the crossing points provide the dynamical information needed to compute averages.

density times the correct (Maxwellian) conditional (most often the momenta and coordinates are independent) probability of the momenta

$$\rho_{\text{BM}}(\mathbf{r}, \mathbf{p}^r) = \rho_{\xi}^r(\mathbf{r}) \rho^p(\mathbf{p}^r | \mathbf{r}). \quad (26)$$

This ensemble provides the basis for a natural method that can be used to compute time correlation functions. Given two arbitrary dynamical variables  $O'(\mathbf{r}, \mathbf{p}^r)$  and  $O''(\mathbf{r}, \mathbf{p}^r)$  we may write

$$\begin{aligned} & \frac{\langle O'(\mathbf{r}, \mathbf{p}^r) O''(\mathbf{r}(t), \mathbf{p}^r(t)) \delta(\xi(\mathbf{r}) - \xi') \rangle}{\langle \delta(\xi(\mathbf{r}) - \xi') \rangle} \\ &= \frac{\langle |Z|^{-1/2} O'(\mathbf{r}, \mathbf{p}^r) O''(\mathbf{r}(t), \mathbf{p}^r(t)) \rangle_{\xi'}}{\langle |Z|^{-1/2} \rangle_{\xi'}}. \end{aligned} \quad (27)$$

The result in Eq. (27) allows one to compute the contribution to the time-dependent rate coefficient in the first set of brackets in Eq. (4) using the Blue Moon ensemble.

To complete the calculation of the time-dependent rate coefficient, we must be able to calculate  $P_{\xi}(\xi') = \langle \delta(\xi - \xi') \rangle = P_{\mathbf{u}} \exp(-\beta \mathcal{W}(\xi'))$  in an efficient manner. The free energy  $\mathcal{W}(\xi')$  is the reversible work needed to bring the system from a given reference state to  $\xi = \xi'$ . The associated thermodynamic force

$$F(\xi') = - \frac{d\mathcal{W}(\xi')}{d\xi'}, \quad (28)$$

can be expressed as the conditional average of a suitable observable and from the thermodynamic integration of  $F(\xi')$  over  $\xi'$ , we can obtain the potential of mean force

$$\mathcal{W}(\xi) = \int d\xi' \frac{d\mathcal{W}(\xi')}{d\xi'} = \int d\xi' \frac{\langle (\partial H / \partial \xi) \delta(\xi - \xi') \rangle}{\langle \delta(\xi - \xi') \rangle}. \quad (29)$$

The explicit form of the thermodynamic force can be obtained by performing the derivative in Eq. (28) to obtain

$$\begin{aligned} F(\xi') &= - \frac{\langle (\partial H / \partial \xi) \delta(\xi - \xi') \rangle}{\langle \delta(\xi - \xi') \rangle} = \frac{\langle (\beta^{-1} (\partial / \partial \xi) \ln |J| - (\partial V / \partial \xi)) \delta(\xi - \xi') \rangle}{\langle \delta(\xi - \xi') \rangle} \\ &\equiv \frac{\langle \hat{F} \delta(\xi - \xi') \rangle}{\langle \delta(\xi - \xi') \rangle} = \frac{\langle Z^{-1/2} \hat{F} \rangle_{\xi'}}{\langle Z^{-1/2} \rangle_{\xi'}}, \end{aligned} \quad (30)$$

where  $|J|$  is the Jacobian of the transformation  $\mathbf{r} \rightarrow \mathbf{u}$  resulting from the explicit integration over the momenta. The quantity  $\hat{F}$  whose conditional average determines the mean force is the sum of two terms: the first term,  $\beta^{-1} (\partial / \partial \xi) \ln |J|$ , represents the apparent forces acting on the system due to the use of generalized (non-inertial) coordinates, while the second term,

$-(\partial V/\partial \zeta)$ , corresponds to the component along the generalized coordinate  $\zeta$  of the force coming from the potential  $V$ . This result expresses the thermodynamic force as a conditional average which can be computed numerically by using the Blue Moon ensemble as indicated in the last line of Eq. (30).

It is possible to obtain another, more convenient, expression for the mean force that obviates the need to compute the Jacobian and the  $\zeta$  derivative of the potential, two quantities that are often difficult to determine [10]. The alternative form involves the Lagrange multiplier associated with the constraint on  $\zeta$ , one of the variables in the  $(\mathbf{u}, \mathbf{p}^u)$  representation of phase space

$$\dot{\mathbf{u}} = \frac{\partial H}{\partial \mathbf{p}^u}, \quad \dot{\mathbf{p}}^u = -\frac{\partial H}{\partial \mathbf{u}} - \lambda \delta_{\zeta} \mathbf{u}. \quad (31)$$

In this approach, one keeps the momentum dependent observable  $(\partial H/\partial \zeta)$  and computes the difference between the configurationally unbiased constrained average and the corresponding conditional average. The (negative of the) mean force can be written as

$$\begin{aligned} \frac{d\mathcal{W}(\zeta')}{d\zeta'} &= \frac{\langle (\partial H/\partial \zeta) \delta(\zeta - \zeta') \rangle}{\langle \delta(\zeta - \zeta') \rangle} \\ &= \frac{\langle Z^{-1/2} (-\lambda - \dot{p}_\zeta + (1/2\beta) (\partial \ln |Z|/\partial \zeta)) \rangle_{\zeta'}}{\langle Z^{-1/2} \rangle_{\zeta'}} \\ &= \frac{\langle Z^{-1/2} ((1/\beta)\mathcal{G} - \lambda) \rangle_{\zeta'}}{\langle Z^{-1/2} \rangle_{\zeta'}}, \end{aligned} \quad (32)$$

where

$$\mathcal{G} = \frac{1}{Z^2} \sum_{i,j=1}^N \frac{1}{m_i m_j} \frac{\partial \zeta}{\partial \mathbf{r}_i} \cdot \frac{\partial^2 \zeta}{\partial \mathbf{r}_i \partial \mathbf{r}_j} \cdot \frac{\partial \zeta}{\partial \mathbf{r}_j}. \quad (33)$$

The computations leading to this result are straightforward but lengthy [10]. This formula provides a much more convenient route for the computation of the mean force and, hence, the potential of mean force which uses quantities that are automatically provided by SHAKE [11, 12] in the constrained molecular dynamics simulation. From these considerations, we see that all quantities needed to estimate the rate coefficient may be determined efficiently in the Blue Moon ensemble.

It is a straightforward to include other constraints, such as bond constraints, in the formalism [3]. In particular, the expression for the correlation function takes the form

$$\begin{aligned} &\frac{\langle O'(\mathbf{r}, \mathbf{p}^r) O''(\mathbf{r}(t), \mathbf{p}^r(t)) \delta(\zeta(\mathbf{r}) - \zeta') \rangle}{\langle \delta(\zeta(\mathbf{r}) - \zeta') \rangle} \\ &= \frac{\langle D^{-1/2} O'(\mathbf{r}, \mathbf{p}^r) O''(\mathbf{r}(t), \mathbf{p}^r(t)) \rangle_{\zeta', M}}{\langle D^{-1/2} \rangle_{\zeta', M}}, \end{aligned} \quad (34)$$

where the subscript  $M$  refers to the other bond constraints and  $D = |\mathbf{Z}|/|Z|$  with  $\mathbf{Z}$  defined as in Eq. (18) but extended to include all constraints

$$Z_{mn} = \sum_{i=1}^N \frac{1}{m_i} \frac{\partial \sigma_m}{\partial \mathbf{r}_i} \cdot \frac{\partial \sigma_n}{\partial \mathbf{r}_i}, \quad (35)$$

while  $Z$  is defined by a similar equation with the restriction to only bond constraints.

To illustrate how the formalism can be applied we consider the adiabatic transfer of a proton in an  $[AHA]^-$  complex, with fixed internuclear separation between the  $A^-$  anions, solvated by a polar liquid [13]



Let  $\mathbf{r}_p$  be the quantum coordinate of the proton and  $\mathbf{R}$  the remainder of the complex and solvent classical degrees of freedom. The total potential energy of the system is  $V = V_{ps}(\mathbf{r}_p, \mathbf{R}) + V_s(\mathbf{R})$ , where  $V_{ps}$  and  $V_s$  are the proton–solvent and solvent–solvent interactions, respectively. In the adiabatic approximation, the proton wave function satisfies the following Schrodinger equation

$$\left[ -\frac{\hbar^2}{2m_p} \nabla_{\mathbf{r}_p}^2 + V_{ps}(\mathbf{r}_p, \mathbf{R}) \right] \Psi_n(\mathbf{r}_p; \mathbf{R}) = \epsilon_n(\mathbf{R}) \Psi_n(\mathbf{r}_p; \mathbf{R}), \quad (37)$$

where  $m_p$  is the mass of the proton,  $2\pi\hbar$  is Planck’s constant,  $\epsilon_n(\mathbf{R})$  is the  $n$ th adiabatic eigenvalue and the corresponding wave function is  $\Psi_n(\mathbf{r}_p; \mathbf{R})$ . The classical coordinates follow Newton’s equations of motion  $m_i \ddot{\mathbf{R}} = -\nabla_{\mathbf{R}}(V_{ps} + \epsilon_n(\mathbf{R}))$  on the  $n$ th adiabatic energy surface. In particular, the adiabatic dynamics on the ground-state surface can be calculated easily by solving the Schrodinger equation for each solvent configuration in order to obtain the ground-state energy  $\epsilon_0(\mathbf{R})$  and wave function  $\Psi_0$ . A convenient reaction coordinate for this problem is the solvent polarization,  $\zeta(\mathbf{R}) = \Delta E(\mathbf{R})$

$$\Delta E(\mathbf{R}) = \sum_i z_i \left( \frac{1}{|\mathbf{R}_i - \mathbf{u}|} - \frac{1}{|\mathbf{R}_i - \mathbf{u}'|} \right), \quad (38)$$

where  $z_i$  is the charge on site  $i$  and  $\mathbf{u}$  and  $\mathbf{u}'$  are two chosen reference positions. The Blue Moon expression for the time-dependent transmission coefficient is

$$\kappa_f(t) = \frac{\langle D^{-1/2}(\Delta \dot{E}) \theta(\Delta E(t) - \Delta E^\ddagger) \rangle_{\xi, M}}{\langle D^{-1/2}(\Delta \dot{E}) \theta(\Delta E) \rangle_{\xi, M}}. \quad (39)$$

Here,  $D^{-1/2}$  is the Blue Moon unbiasing factor with  $D = (2m)^{-1} \sum_i (\nabla_i \Delta E)^2$ , where the sum extends over all classical degrees of freedom assumed to have equal mass  $m$ . The time-dependent transmission coefficient  $\kappa(t)$  was calculated by constraining the system to  $\Delta E(t) = \Delta E^\ddagger$  to generate the Blue Moon ensemble and releasing the constraint to determine the time evolution of

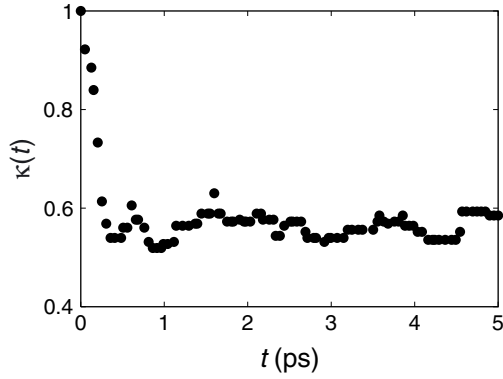


Figure 3. The time-dependent transmission coefficient for adiabatic proton transfer.

$\theta(\Delta E(t) - \Delta E^\ddagger)$  that appears in the correlation function expression for  $\kappa(t)$ . The results of this calculation are shown in Fig. 3 [13]. The simulations show that a plateau is reached on the order of 1 ps and that the forward rate constant  $k_f$  is about 0.6 of its  $k_f^{\text{TST}}$  value as a result of recrossings of the  $\Delta E(t) = \Delta E^\ddagger$  surface. The transition-state rate constant can also be determined in the Blue Moon ensemble using the expression

$$k_f^{\text{TST}} = (2\pi\beta)^{-1/2} \langle D^{-1/2} \rangle_{\xi, M}^{-1} \frac{e^{-\beta W(\Delta E^\ddagger)}}{\int_{\Delta E < \Delta E^\ddagger} d\Delta E e^{-\beta W(\Delta E)}}. \quad (40)$$

The free energy at the barrier top  $\Delta E^\ddagger$  was estimated from quadratic approximations to the free energy function near the metastable states and the expectation value of  $D^{-1/2}$  in the above formula was evaluated in the Blue Moon ensemble. The resulting value for the transition-state rate constant is  $k_f^{\text{TST}} = 4 \times 10^9 \text{ s}^{-1}$ . The direct molecular dynamics simulation of the rate constant would be a difficult task without the use of a rare event sampling technique in view of the activated nature of this proton transfer reaction.

### 3. Vectorial Reaction Coordinate

In some instances, the reaction path the system takes to go from reactants to products may not be simply related to a single scalar reaction coordinate which is chosen on physical grounds. An inappropriate choice of a reaction coordinate can lead to difficulties in the computation of the rate. The underlying structure of the reaction path may often be revealed by extending the description to vectorial reaction coordinates. For example, if the free energy surface as a function of two reaction coordinates  $\zeta_1$  and  $\zeta_2$  has the structure

shown in Fig. 4, then a description based on projections of the free energy along  $\xi_1$  will lead to misleading results. The discovery of the appropriate set of reaction coordinates, or more generally the reaction path, may not be a simple task for some systems.

To describe the free energy of such more complex situations, we suppose that the system can be characterized by a set of reaction coordinates which are a functions of the positions of the particles in the system,  $\xi(\mathbf{r}) = \{\xi_1, \xi_2, \dots, \xi_n\}$ . The probability density of finding  $\xi(\mathbf{r}) = \xi'$  is

$$P(\xi') = \langle \delta(\xi(\mathbf{r}) - \xi') \rangle = \left\langle \prod_{\alpha=1}^n \delta(\xi_{\alpha}(\mathbf{r}) - \xi'_{\alpha}) \right\rangle. \quad (41)$$

The free energy  $\mathcal{W}(\xi')$  associated with the vectorial reaction coordinate is

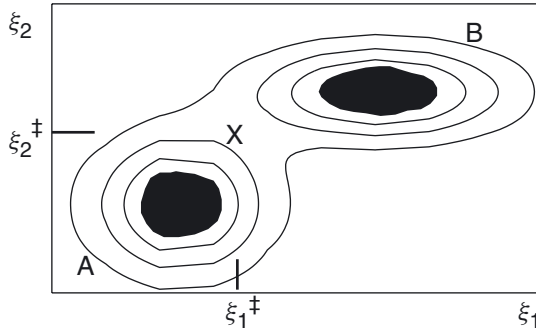
$$\mathcal{W}(\xi') = -\beta^{-1} \ln(P(\xi')/P_u), \quad (42)$$

where  $P_u$  is again a uniform probability density of  $\xi'$ . This free energy  $\mathcal{W}$  or reversible work needed to take the system from the vectorial reaction coordinate value  $\xi^a$  to  $\xi^b$  can be calculated by means of a  $n$ -dimensional line integral

$$\mathcal{W}(\xi^b) - \mathcal{W}(\xi^a) = \int_{C(\xi^a, \xi^b)} d\xi' \cdot \frac{\partial \mathcal{W}}{\partial \xi'}, \quad (43)$$

where  $C(\xi^a, \xi^b)$  is the path taken from  $\xi^a$  to  $\xi^b$ . Using Eqs. (41) and (42) and the properties of the delta function one may show that

$$-\frac{\partial \mathcal{W}}{\partial \xi'} = -\left\langle \frac{\partial \mathcal{H}}{\partial \xi} \right\rangle_{\xi'}^{\text{cd}} = \mathbf{F}_{\xi'}, \quad (44)$$



*Figure 4.* Sketch of a free energy surface showing two metastable regions depicted as shaded domains in the figure. To obtain such a plot, the free energy is computed for specified values of two reaction coordinates,  $\xi_1(\mathbf{r})$  and  $\xi_2(\mathbf{r})$ . The saddle point at  $(\xi_1^\ddagger, \xi_2^\ddagger)$  is indicated by a  $\times$  symbol.

where  $\mathbf{F}_\xi$  is the mean force associated with  $\xi$ . Following the procedure outlined in Section 2 we may write the (negative of the) mean force in the form [14]

$$\frac{\partial \mathcal{W}}{\partial \xi'} = \frac{\langle |\Xi|^{-1/2} [1/\beta \mathbf{G} - \lambda] \rangle_{\xi^\ddagger}}{\langle |\Xi|^{-1/2} \rangle_{\xi^\ddagger}}, \quad (45)$$

where

$$\Xi_{\alpha\beta} = \sum_k \frac{1}{m_k} \frac{\partial \xi_\alpha}{\partial \mathbf{r}_k} \frac{\partial \xi_\beta}{\partial \mathbf{r}_k}. \quad (46)$$

and we have defined the vector  $\mathbf{G}$  with elements

$$G_\alpha = \sum_{i,n} \frac{1}{m_i m_n} \sum_{\mu\gamma\nu} \Xi_{\mu\alpha}^{-1} \frac{\partial \xi_\mu}{\partial \mathbf{r}_i} \frac{\partial^2 \xi_\gamma}{\partial \mathbf{r}_i \partial \mathbf{r}_n} \frac{\partial \xi_\nu}{\partial \mathbf{r}_n} \Xi_{\gamma\nu}^{-1}, \quad \alpha = 1, \dots, n, \quad (47)$$

and  $\lambda$  is the vector of Lagrange multipliers appearing in the constrained equations of motion. The formalism for the most general treatment where there are both other constraints and vectorial reaction coordinate constraints has been given by [15].

The Blue Moon ensemble has been used to compute the free energy as a function of several reaction coordinates for ionization reactions of  $[\text{NaCl}_2]^-$  ion complexes in water clusters [16]. The multidimensional reaction coordinate formalism described above has also been applied to study the interaction between monomers in a superoxide dismutase protein, *Photobacterium leiognathi* [14], which we now briefly describe. This protein provides a good example of macromolecular recognition since the monomers are able to form the dimeric enzyme in water. Calculation of the binding force for different mutant proteins, obtained by substituting the amino acids at the monomer–monomer interface, and structural analysis could provide insight into the recognition process. In Fig. 5, we give a pictorial view of the protein as found in nature.

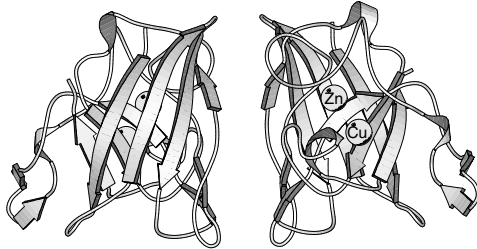


Figure 5. *Photobacterium leiognathi* Cu,Zn SOD structure. The ribbon shows the fold of the two identical subunits constituting the dimer. Arrows represent the  $\beta$ -strands while thin wires represent the random-coil structure and the turns. The copper and zinc ions are shown as dark and light gray labelled spheres, respectively.

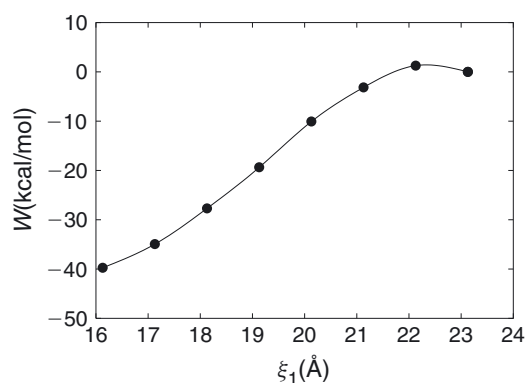


Figure 6. The graph shows the potential of mean force as a function of the separation.

The separation of the two monomers can be studied as a function of their relative displacement and orientation. This requires a six-dimensional reaction coordinate. A complete sampling of the phase space as a function of this six-dimensional coordinate is not feasible for the system under investigation (there are 2694 atoms in the proteins which were solvated by 9944 water molecules). However, the *principal* component of the binding force can be calculated by freezing the slow orientational modes of the two monomers and studying their separation at a fixed orientation. One can choose the initial orientation of the monomers to be that minimizing the energy of the system; hence, the mean force along the relative separation distance can be calculated. The result of such a calculation is shown in Fig. 6. While the separation path one obtains is not fully realistic, it can be used to perform a series of identical numerical experiments on different mutants of *Photobacterium leiognathi* to investigate the important structural and dynamical features in the recognition process.

## 4. Outlook

The description of condensed phase activated rate processes is a challenging problem. The choice of a suitable reaction coordinate or set of reaction coordinates is a central feature of such descriptions. Often this choice is made on physical grounds but schemes for determining reaction paths are needed to provide results when physical considerations are inadequate. Once such reaction coordinates are known the methods described in this chapter provide algorithms for the computation of reaction rates. As briefly described in the text, the methods presented here have been used to investigate also adiabatic quantum rate processes. Non-adiabatic reaction rates may also be treated using the techniques developed here [17].

## References

- [1] C. Dellago, P.G. Bolhuis, and P.L. Geissler, "Transition path sampling," *Adv. Chem. Phys.*, 123, 1–78, 2002.
- [2] E. Weinan and E. Vanden Eijnden, "Conformation dynamics and transition pathways in complex systems," To be published, 2004.
- [3] E. Carter, G. Ciccotti, C. Hynes, and R. Kapral, "Constrained reaction coordinate dynamics for the simulation of rare events," *Chem. Phys. Lett.*, 156, 472–477, 1989.
- [4] T. Yamamoto, "Quantum statistical mechanical theory of the rate of exchange chemical reactions in the gas phase," *J. Chem. Phys.*, 33, 281–289, 1960.
- [5] D. Chandler, "Statistical-mechanics of isomerization dynamics in liquids and transition-state approximation," *J. Chem. Phys.*, 68, 2959–2970, 1978.
- [6] R. Kapral, S. Consta, and L. McWhirter, "Chemical rate laws and rate constants," In: B. Berne, G. Ciccotti, and D. Coker (eds.), *Classical and Quantum Dynamics in Condensed Phase Systems*, World Scientific, Singapore, pp. 583–616, 1998.
- [7] H. Goldstein, *Classical Mechanics*, Addison-Wesley, Reading, MA, 1980.
- [8] M. Fixman, "Classical statistical-mechanics of constraints – theorem and application to polymers," *Proc. Natl. Acad. Sci. USA.*, 71, 3050–3053, 1974.
- [9] N.G. van Kampen and J.J. Lodder, "Constraints," *Am. J. Phys.*, 52, 419–424, 1984.
- [10] M. Sprik and G. Ciccotti, "Free energy from constrained molecular dynamics," *J. Chem. Phys.*, 109, 7737–7744, 1998.
- [11] J.P. Ryckaert, G. Ciccotti, and H.J.C. Berendsen, "Numerical-integration of Cartesian equations of motion of a system with constraints – molecular-dynamics of *n*-alkanes," *J. Comput. Phys.*, 23, 327–341, 1977.
- [12] G. Ciccotti and J.P. Ryckaert, "Molecular dynamics simulation of rigid molecules," *Comput. Phys. Rep.*, 4, 345–392, 1986.
- [13] D. Laria, G. Ciccotti, M. Ferrario, and R. Kapral "Molecular dynamics study of adiabatic proton transfer reactions in solution," *J. Chem. Phys.*, 97, 378–388, 1992.
- [14] A. Sergi, G. Ciccotti, M. Falconi, A. Desideri, and M. Ferrario, "Effective binding force calculation in a dimeric protein by molecular dynamics simulation," *J. Chem. Phys.*, 116, 6329–6338, 2002.
- [15] I. Coluzza, M. Sprik, and G. Ciccotti, "Constrained reaction coordinate dynamics for systems with constraints," *Mol. Phys.*, 101, 2885–2894, 2003.
- [16] S. Consta and R. Kapral, "Ionization reactions of ion complexes in mesoscopic water clusters," *J. Chem. Phys.*, 111, 10183–10191, 1999.
- [17] A. Sergi and R. Kapral, "Quantum-classical dynamics of non-adiabatic chemical reactions," *J. Chem. Phys.*, 118, 8566–8575, 2003.