One-dimensional reaction coordinates for diffusive activated rate processes in many dimensions

Alexander Berezhkovskii^{a)}

Mathematical and Statistical Computing Laboratory, Division for Computational Bioscience, Center for Information Technology, National Institutes of Health, Bethesda, Maryland 20892

Attila Szabo

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892

(Received 25 May 2004; accepted 22 September 2004; published online 13 December 2004)

For multidimensional activated rate processes controlled by diffusive crossing of a saddle point region, we show that a one-dimensional reaction coordinate can be constructed even when the diffusion anisotropy is arbitrary. The rate constant, found using the potential of mean force along this coordinate, is identical to that predicted by the multidimensional Kramers–Langer theory. This reaction coordinate minimizes the one-dimensional rate constant obtained using a trial reaction coordinate and is orthogonal to the stochastic separatrix, the transition state that separates reactants from products. © 2005 American Institute of Physics. [DOI: 10.1063/1.1818091]

I. INTRODUCTION

The introduction of a low dimensional description is an important first step towards understanding the dynamics of complex systems. When can a multidimensional activated rate process be described in terms of Markovian dynamics along a single reaction coordinate or order parameter? Our interest in this problem is in the context of protein folding, where collective coordinates such as the fraction of native contacts are widely used, 1-5 but similar questions arise in many other physical situations. The usual answer is that a reaction coordinate exists when dynamics along all other coordinates is sufficiently fast so that local equilibrium along these coordinates is rapidly established. In this note we show that a one-dimensional reaction coordinate exists under significantly less restrictive conditions, namely, for any system whose transition rate is described by Langer's^{6,7} multidimensional generalization of the one-dimensional Kramers theory⁸ of diffusive barrier crossing. This result certainly does not imply that a one-dimensional reaction coordinate necessarily exists for complex processes such as protein folding, but it is reassuring that it always exists for a simple yet nontrivial multidimensional system. In general, the problem of finding a low dimensional description of the reaction mechanism is difficult even when one can "see" how the reaction happens in microscopic detail⁹ (i.e., one is given many simulated transition path trajectories).

Langer's theory is applicable when (1) the reactant and product wells are separated by a single high saddle point and (2) the rate is determined by the diffusion of the system across the saddle point region. Our criterium for the existence of a reaction coordinate is simply this. The rate constant calculated using one-dimensional Kramers' theory for diffusion on the potential of mean force along the reaction

coordinate must be identical to the multidimensional Langer rate constant. We will show here that such a reaction coordinate does exist for arbitrary values of the diffusion coefficients, and it is perpendicular to the stochastic separatrix. Starting from anywhere on this separatrix the system reaches the reactants and products for the first time with equal probability.

Ryter¹⁰ seems to have been the first who suggested the physically appealing definition of the transition state as the stochastic separatrix. This idea has been widely exploited^{11–15} and was introduced into the protein folding literature by Du *et al.*¹⁶ a few years ago. Hummer¹⁷ recently has shown that for diffusion the stochastic separatrix is also the set of points with the highest probability that trajectories passing through them are reactive (i.e., form transition paths between reactants and products). Our analysis reveals another interesting property of the stochastic separatrix, namely, that it is orthogonal to the one-dimensional reaction coordinate. Equivalently, for diffusive dynamics, the reaction coordinate turns out to be perpendicular to the transition state just as in gas phase chemical reactions.^{18,19}

The outline of this note is as follows: We begin by reviewing Kramers' theory⁸ of diffusive barrier crossing and its generalization to many dimensions by Langer.⁶ Next we introduce a tentative reaction coordinate q specified by a vector \mathbf{e} and calculate the rate constant using one-dimensional Kramers' theory for diffusive dynamics along this coordinate in the presence of the potential of mean force U(q). We then show that this rate constant has a minimum and becomes identical to Langer's multidimensional result when \mathbf{e} is perpendicular to the stochastic separatrix. We also propose an efficient numerical method for finding the one-dimensional reaction coordinate based on singular value decomposition. A simple example is presented to illustrate the dependence of the reaction coordinate on the diffusion anisotropy. Finally we summarize our key findings.

a)Permanent address: Karpov Institute of Physical Chemistry, Ulica Vorontsovo Pole 10, 103064 Moscow K-64, Russia.

II. THEORY

Given a one-dimensional potential of mean force U(q), Kramers showed that the rate constant for diffusive barrier crossing of a high barrier is given by

$$k_1 = \left[\left(\int_{\cup} e^{-\beta U(q)} dq \right) \left(\int_{\cap} \frac{e^{\beta U(q)}}{D(q)} dq \right) \right]^{-1}, \tag{1}$$

where $\beta = (k_B T)^{-1}$ and the two integrals are taken near the reactant well bottom (\cup) and the barrier top (\cap) . Langer generalized the Kramers' solutions to many dimensions in the special case when the reaction rate is controlled by the passage of the system through the saddle point region. In this region the multidimensional potential (or more generally free energy) surface is given by $\beta U(\mathbf{x}) = \mathbf{x}^T \mathbf{V} \mathbf{x}/2$, where V is the matrix of second derivatives of the potential at the saddle point located at x=0, det V<0. Near the reactant well bottom $\beta U(\mathbf{x}) = -\Delta + (\mathbf{x} - \mathbf{x}_R)^T \mathbf{V}_R(\mathbf{x} - \mathbf{x}_R)/2$, where Δ is the energy difference between the saddle point and the reactant well bottom measured in units of k_BT and V_R is the matrix of second derivatives of the potential at the well bottom located at $\mathbf{x} = \mathbf{x}_R$, det $\mathbf{V}_R > 0$. Langer showed that in N dimensions the rate constant k_N for transitions between reactants and products is given by the expression

$$k_N = \frac{1}{2\pi} \left(\frac{\det \mathbf{V}_R}{|\det \mathbf{V}|} \right)^{1/2} \lambda_+ \exp(-\Delta), \tag{2}$$

in which λ_+ is the only positive root of the equation $\det(\lambda \mathbf{I} + \mathbf{D}\mathbf{V}) = 0$, where \mathbf{D} is the matrix of diffusion coefficients and \mathbf{I} is the unit matrix. Equivalently, λ_+ is the only positive eigenvalue of the matrix $(-\mathbf{D}\mathbf{V})$,

$$\mathbf{DVf}_{+} = -\lambda_{+}\mathbf{f}_{+} \tag{3}$$

with corresponding eigenvector \mathbf{f}_+ .

Consider now the potential of mean force along a reaction coordinate q, which is a straight line directed along a unit vector \mathbf{e} . The potential of mean force is defined as

$$\exp[-\beta U_{\rm MF}(q)] = \frac{1}{L} \int \delta(q - \mathbf{e}^T \mathbf{x}) \exp[-\beta U(\mathbf{x})] d\mathbf{x}, \quad (4)$$

where the length L is given by

$$L = \int \delta(\mathbf{e}^T \mathbf{x}) \exp[-\beta U(\mathbf{x})] d\mathbf{x}.$$

Is it possible to find a vector \mathbf{e} so that the rate constant calculated using one-dimensional Kramers' theory for diffusion on $U_{\mathrm{MF}}(q)$ would be identical to the multidimensional expression in Eq. (2) for arbitrary diffusion anisotropy? The answer is "yes." To demonstrate this we substitute the relation in Eq. (4) into the expression for k_1 in Eq. (1). Using the identity

$$\int \delta(q - \mathbf{e}^{T} \mathbf{x}) \exp\left[-\frac{1}{2} (\mathbf{x} - \mathbf{x}_{0})^{T} \mathbf{A} (\mathbf{x} - \mathbf{x}_{0})\right] d\mathbf{x}$$

$$= \frac{\exp\left[-\frac{1}{2} \frac{(q - \mathbf{e}^{T} \mathbf{x}_{0})^{2}}{\mathbf{e}^{T} \mathbf{A}^{-1} \mathbf{e}}\right]}{\sqrt{(2\pi)^{(1-N)} (\mathbf{e}^{T} \mathbf{A}^{-1} \mathbf{e}) \det \mathbf{A}}}$$
(5)

that is valid when $(\mathbf{e}^T \mathbf{A}^{-1} \mathbf{e}) \det \mathbf{A} > 0$, we find that

$$k_1(\mathbf{e}) = \frac{1}{2\pi} \left(\frac{\det \mathbf{V}_R}{|\det \mathbf{V}|} \right)^{1/2} \frac{\mathbf{e}^T \mathbf{D} \mathbf{e}}{|\mathbf{e}^T \mathbf{V}^{-1} \mathbf{e}|} \exp(-\Delta).$$
 (6)

Here we have used the fact that in the saddle point region the diffusion coefficient along our reaction coordinate is independent of q and given by $\mathbf{e}^T \mathbf{D} \mathbf{e}$. It is important to note that the rate constant $k_1(\mathbf{e})$ depends only on the direction of the vector \mathbf{e} in the saddle point region.

In the spirit of variational transition state theory, let us consider $k_1(\mathbf{e}+\boldsymbol{\delta})$, where $\boldsymbol{\delta}$ is a small increment orthogonal to the vector \mathbf{e} . Expanding $k_1(\mathbf{e}+\boldsymbol{\delta})$ to the second order in $(\boldsymbol{\delta}^T\boldsymbol{\delta})^{1/2}$, we find that $k_1(\mathbf{e})$ has a minimum when $\mathbf{e}=\mathbf{e}_+$ that satisfies

$$\mathbf{VDe}_{+} = -\lambda_{+}\mathbf{e}_{+} , \qquad (7)$$

where λ_+ is the only positive eigenvalue of the matrix $(-\mathbf{V}\mathbf{D})$ [or equivalently $(-\mathbf{D}\mathbf{V})$]. Multiplying both sides of Eq. (7) by $\mathbf{e}_+^T \mathbf{V}^{-1}$, we find that

$$-\frac{\mathbf{e}_{+}^{T}\mathbf{D}\mathbf{e}_{+}}{\mathbf{e}_{+}^{T}\mathbf{V}^{-1}\mathbf{e}_{+}} = \frac{\mathbf{e}_{+}^{T}\mathbf{D}\mathbf{e}_{+}}{|\mathbf{e}_{+}^{T}\mathbf{V}^{-1}\mathbf{e}_{+}|} = \lambda_{+},$$
(8)

where we have used the fact that $\mathbf{e}_{+}^{T}\mathbf{V}^{-1}\mathbf{e}_{+}<0$. Substituting this into Eq. (6) and comparing the result with Eq. (2) we see that $k_{1}(\mathbf{e}_{+})=k_{N}$. Thus we have shown that for a one-dimensional potential of mean force specified by the vector \mathbf{e}_{+} , the rate constant is not only a minimum but is also exact in the sense that it is identical to the Langer result in Eq. (2).

What is the significance of the direction specified by the vector \mathbf{e}_+ ? One's first guess might be that \mathbf{e}_+ points along the direction of the reactive flux in the saddle point region. Interestingly, this is only true for isotropic diffusion (D =DI) because the reactive flux is directed along the vector \mathbf{f}_{+} defined in Eq. (3) while \mathbf{e}_{+} is the corresponding eigenvector of the matrix VD. In fact e_+ is directed along the only unstable diffusive normal mode in the saddle point region. One can see this starting with the noiseless equation of motion $\dot{\mathbf{x}} = -\mathbf{D}\mathbf{V}\mathbf{x}$ and noting that $\mathbf{Q}_{+} = \mathbf{e}_{+}^{T}\mathbf{x}$ satisfies $\dot{\mathbf{Q}}_{+}$ $=\lambda_+ \mathbf{Q}_+$. Because of symmetry one might expect that \mathbf{e}_+ is perpendicular to the stochastic separatrix. To prove this we start with the observation of Klosek-Dygas et al.²⁰ that the stochastic separatrix is determined by the (N-1) eigenvectors \mathbf{f}_{-} 's of the matrix $(-\mathbf{D}\mathbf{V})$ that correspond to the negative eigenvalues. Because the eigenvectors of the matrices VD and DV corresponding to different eigenvalues are orthogonal to each other, it follows that $\mathbf{e}_{+}\mathbf{f}_{-}=0$.

We now give a more constructive proof of the fact that the vector \mathbf{e}_+ is perpendicular to the stochastic separatrix. For potentials, for which Langer's theory is applicable, it can be readily shown that the splitting probability $\sigma(\mathbf{x})$, which is the probability to reach the reactant well before reaching the product well starting out at the point \mathbf{x} , is

$$\sigma(\mathbf{x}) = \frac{1}{\sqrt{2\pi}} \int_{\mathbf{a}^T \mathbf{x}}^{\infty} \exp\left(-\frac{1}{2}z^2\right) dz, \qquad (9)$$

where the vector a is given by

$$\mathbf{a} = \mathbf{e}_{+} / |\mathbf{e}_{+}^{T} \mathbf{V}^{-1} \mathbf{e}_{+}|^{1/2}. \tag{10}$$

The stochastic separatrix is the set of points, which satisfy $\sigma(\mathbf{x})=0.5$. It immediately follows from Eq. (9) that these points satisfy $\mathbf{e}_+^T\mathbf{x}=0$. Thus the stochastic separatrix or the transition state is a hyperplane that passes through the saddle point perpendicular to the vector \mathbf{e}_+ .

When Langer's theory is applicable, if one determines the set of points with $\sigma(\mathbf{x}) = 0.5$, then the vector perpendicular to the hyperplane formed by these points determines the direction of the reaction coordinate. Recently, based on a preprint of this work, Rhee and Pande²¹ showed how one can also use points with $\sigma(\mathbf{x}) \neq 0.5$ to determine the vector \mathbf{e}_+ and illustrated their procedure for a two-dimensional potential. Here we propose an efficient method for doing this based on Eq. (9) using singular value decomposition. ^{22,23}

Given M points \mathbf{x}_i , i = 1, 2, ..., M, in N-dimensional space with splitting probabilities $\sigma(\mathbf{x}_i) = p_{\text{fold}}(\mathbf{x}_i)$ (say between 0.4 and 0.6), what is the best estimate for the unit vector \mathbf{e}_+ that is perpendicular to the stochastic separatrix? Using Eq. (9) the splitting probability of a conformation \mathbf{x}_i can be converted²¹ into a particular value of $\mathbf{a}^T\mathbf{x}_i$ which we will call b_i . Let us construct a $M \times N$ matrix \mathbf{X} so that the ith row is the vector \mathbf{x}_i^T . Then to find vector \mathbf{a} (which when normalized becomes \mathbf{e}_+) we must solve

$$\mathbf{X}\mathbf{a} = \mathbf{b},\tag{11}$$

where **b** is an M-dimensional vector with components b_i . This can be done most efficiently by means of singular value decomposition (SVD).^{22,23} The SVD of the matrix **X** is

$$\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}^T \quad \mathbf{U}^T\mathbf{U} = \mathbf{V}^T\mathbf{V} = \mathbf{V}\mathbf{V}^T = \mathbf{I},\tag{12}$$

where **S** is the diagonal matrix containing the singular values s_k , k = 1,2,...,N, and **I** is the $N \times N$ unit matrix. The vector **a** is then given by

$$\mathbf{a} = \mathbf{V}\mathbf{S}^{-1}\mathbf{U}^T\mathbf{b}.\tag{13}$$

There are two cases of interest. When M>N there are more equations than unknowns and Eq. (13) gives the best estimate of $\mathbf{a} = \mathbf{e}_+ / |\mathbf{e}_+^T \mathbf{V}^{-1} \mathbf{e}_+|^{1/2}$ in the least square sense. Having found \mathbf{e}_+ in this way, one can calculate the potential of mean force along it, say by umbrella sampling. In addition, if a good estimate of the effective diffusion coefficient along this coordinate can be extracted from simulations, the rate constant can be obtained using one-dimensional Kramers theory, Eq. (1).

When N>M there are more unknowns than equations. In this case only relatively few singular values, say K, have significant magnitudes. The vector \mathbf{a} with the smallest norm, $\mathbf{a}^T\mathbf{a} = (\mathbf{e}_+^T\mathbf{V}^{-1}\mathbf{e}_+)^{-1}$, can be found from Eq. (13) by setting $(S^{-1})_{ij} = s_i^{-1} \delta_{ij} i < K$ and zero otherwise. The estimate of \mathbf{e}_+ found in this way can be improved by exploiting the variational principle discussed above. Specifically using the above \mathbf{e}_+ as an initial guess for the direction of the reaction coordinate, one can vary this direction so as to minimize the one-dimensional Kramers rate constant.

III. ILLUSTRATIVE CALCULATIONS

As a simple illustrative example consider two-dimensional diffusion $(D_{xy}=D_{yx}=0)$ on the potential

$$\beta U(x,y) = \beta U(x) + \Omega^2 (x-y)^2 / 2,$$
 (14a)

where

$$\beta U(x) = -\Delta + \omega^2 (x + x_0)^2 / 2 \quad x < -x_0 / 2$$

$$= -\omega^2 x^2 / 2 \quad -x_0 / 2 < x < x_0 / 2$$

$$= -\Delta + \omega^2 (x - x_0)^2 / 2 \quad x_0 / 2 < x,$$
(14b)

and $\Delta = \omega^2 x_0^2/4$ is the energy difference between the saddle point and the well bottom measured in units of k_BT . The potential $\beta U(x,y)$ is shown in Fig. 1(a) for $x_0 = 2.2$, $\omega^2 = 4$, and $\Omega^2 = 1.01\omega^2$.

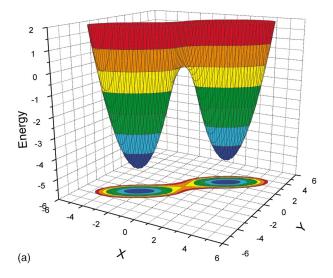
The potential of mean force along the x coordinate is just U(x). It is well known that when $D_y/D_x \rightarrow \infty$ the rate constant is given by Kramers' one-dimensional theory for a system diffusing on this potential with diffusion coefficient D_x . Using the results of this note we can extend this one-dimensional description to the case of arbitrary values of D_x and D_y , on condition that the rate constant is given by the Langer expression in Eq. (2) in two dimensions. It should be pointed out that even for the simple potential given in Eq. (14) Langer's formula is not always valid. Specifically it has been shown²⁴ that it does not work for large diffusion anisotropy when $\Omega^2 < \omega^2$. In this regime it is still an open question whether a reaction coordinate exists.

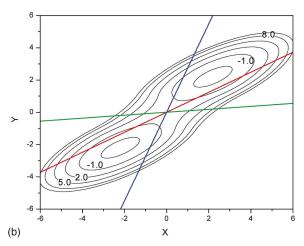
In Figs. 1(b) and 1(c) we present three one-dimensional reaction coordinates and corresponding potentials of mean force that lead to the same values of the rate constant as those predicted by Langer's theory at three different levels of the diffusion anisotropy for the two-dimensional potential shown in Fig. 1(a). The reaction coordinates shown in Fig. 1(b) are straight lines passing through the saddle point in directions parallel to the vectors \mathbf{e}_{+} found by solving Eq. (7) for $D_x = D_y$ (red), $D_x = 0.1D_y$ (green), and $D_x = 10D_y$ (blue). Note that these lines do not pass through the minima of reactant and product wells. The potentials of mean force along these coordinates are shown in Fig. 1(c). For each of these potentials the rate constants should be obtained using the effective diffusion coefficients given by $\mathbf{e}_{+}^{T}\mathbf{D}\mathbf{e}$ (D_{x} for $D_x = D_y$; 1.08 D_x for $D_x = 0.1D_y$; 0.21 D_x for $D_x = 10D_y$). It is of particular interest that the one-dimensional reaction coordinates depend on the dynamic parameters of the process, i.e., on the diffusion anisotropy. Finally we note that for high barriers the rate is independent of the direction of the reaction coordinate in the reactant and product wells. The system is in local equilibrium in both wells and the rate is determined by the dynamics in the saddle point region.

IV. SUMMARY

In this paper we have introduced a one-dimensional reaction coordinate for diffusive reaction dynamics, which has the following properties:

- (1) The rate constant for one-dimensional dynamics along this coordinate is identical to that given by the multi-dimensional Langer theory.⁶
- (2) It is a collective coordinate that is a linear combination of all coordinates of the multidimensional system.





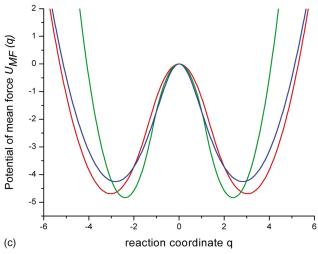


FIG. 1. (Color) Two-dimensional potential $\beta U(x,y)$ given in Eq. (14) [panel (a)], three one-dimensional reaction coordinates for $D_x = D_y$ (red), $D_x = 0.1D_y$ (green), and $D_x = 10D_y$ (blue) [panel (b)], and the corresponding potentials of mean force along these coordinates [panel (c)].

- (3) The reaction coordinate and, hence, the potential of mean force along it depend on the nature of the dynamics (i.e., diffusion anisotropy).
 - (4) In general the reaction coordinate is *not* parallel to

the reactive flux but is directed along the unstable diffusive normal mode in the saddle point region.

- (5) The reaction coordinate is perpendicular to the stochastic separatrix or the transition state. If the splitting probabilities or p_{fold} values are determined for a sufficient number of points on and near the stochastic separatrix from computer simulations, then the vector pointing along the reaction coordinate can be found from these p_{fold} values using singular value decomposition.
- (6) Finally, the rate constant for a multidimensional diffusive barrier crossing can be found by varying the reaction coordinate so that the corresponding rate constant calculated using the one-dimensional Kramers theory reaches a minimum. The hyperplane that passes through the saddle point perpendicular to the resulting reaction coordinate is the stochastic separatrix or the transition state. This variational principle may form the basis of a practical procedure to determine reaction coordinates and consequently transition states.

¹J. D. Bryngelson and P. G. Wolynes, J. Phys. Chem. **93**, 6902 (1989).

²A. Sali, E. Shakhnovich, and M. Karplus, Nature (London) **369**, 248 (1994).

³R. Zwanzig, A. Szabo, and B. Bagchi, Proc. Natl. Acad. Sci. U.S.A. 89, 20 (1995).

⁴N. D. Socci, J. N. Onuchic, and P. G. Wolynes, J. Chem. Phys. **104**, 5860 (1996).

⁵V. Munoz and W. A. Eaton, Proc. Natl. Acad. Sci. U.S.A. **96**, 11311 (1999).

⁶J. S. Langer, Ann. Phys. (Leipzig) **54**, 258 (1969).

 ⁷P. Hanggi, P. Talkner, and M. Borkovec, Rev. Mod. Phys. **62**, 251 (1990).
 ⁸H. A. Kramers, Physica (Amsterdam) **7**, 284 (1940).

⁹T. A. McCormick and D. Chandler, J. Phys. Chem. B **107**, 2796 (2003).

¹⁰D. Ryter, Physica A **142**, 103 (1987); J. Stat. Phys. **49**, 751 (1987).

¹¹ M. M. Klosek, B. J. Matkowsky, and Z. Schuss, Ber. Bunsenges. Phys. Chem. **95**, 331 (1991).

¹²E. Pollak, A. M. Berezhkovskii, and Z. Schuss, J. Chem. Phys. **100**, 334 (1994).

¹³ P. Talkner, Chem. Phys. **180**, 199 (1994); A. N. Drozdov and P. Talkner, Phys. Rev. E **54**, 6160 (1996).

¹⁴ P. L. Geissler, C. Delago, and D. Chandler, J. Phys. Chem. B **103**, 3706 (1999).

¹⁵ P. G. Bolhuis, D. Chandler, C. Delago, and P. L. Geissler, Annu. Rev. Phys. Chem. **53**, 291 (2002).

¹⁶R. Du, V. S. Pande, A. Yu. Grosberg, T. Tanaka, and E. Shakhnovich, J. Chem. Phys. **108**, 334 (1998).

¹⁷G. Hummer, J. Chem. Phys. **120**, 516 (2004).

¹⁸E. E. Nikitin, Theory of Elementary Atomic and Molecular Processes in Gases (Clarendon, Oxford, 1974).

¹⁹ P. Pechukas, in *Modern Theoretical Chemistry Part B*, edited by W. H. Miller (Plenum, New York, 1976), Vol. 2, p. 269.

²⁰ M. M. Klosek-Dygas, B. M. Hoffman, B. J. Matkowsky, A. Nitzan, M. A. Ratner, and Z. Schuss, J. Chem. Phys. 90, 1141 (1989).

²¹Y. M. Rhee and V. S. Pande (unpublished).

²²G. H. Golub and C. F. Van Loan, *Matrix Computations* (John Hopkins University Press, Baltimore and London, 1996).

²³C. L. Lawson and R. T. Hanson, Solving Least Squares Problems (Prentice-Hall, Englewood Cliffs, NJ, 1974).

²⁴ A. M. Berezhkovskii and V. Yu. Zitserman, Chem. Phys. Lett. **172**, 235 (1990); Chem. Phys. **157**, 141 (1991); Physica A **187**, 519 (1992); A. M. Berezhkovskii, E. Pollak, and V. Yu. Zitserman, J. Chem. Phys. **97**, 2422 (1992).