Variational layer expansion for kinetic processes

Giorgio J. Moro*
Dipartimento di Chimica Fisica, Università di Padova, via Loredan 2, 35131 Padova, Italy

Franco Cardin†
Dipartimento di Matematica Pura e Applicata, Università di Padova, via Belzoni 7, 35131 Padova, Italy
(Received 29 July 1996; revised manuscript received 14 November 1996)

Often the analysis of the Fokker-Planck (FP) operator near the saddle point is sufficient to characterize the activated processes. However, there are also situations where the kinetic processes are controlled by the dynamics far away from the saddle points. Correspondingly, the knowledge of FP kinetic modes in all the phase space is required in order to describe accurately the activated processes. To this aim we propose a variational method for approximating the site-localizing functions that are defined as linear combinations of the FP slow eigenfunctions and describe the stable-state populations. The starting point is the layer expansion method that has been developed by Matkowsky and Schuss [Siam J. Appl. Math. **33**, 365 (1977); **36**, 604 (1979); **40**, 242 (1981)], which we apply to the covariant form of the FP equation. Error-function profiles across the separatrix are derived in this way for the site-localizing functions. The same kind of profile is found in the numerical solutions of a bistable two-dimensional Smoluchowski equation, but about a line (the so-called stochastic separatrix) that is, in general, different from the deterministic separatrix. Thus the layer expansion has to be generalized by considering the separatrix as a parametric function to be optimized according to a variational criterion for the decay rates. After discretization along the separatrix of the integral relation for the rate, the variational problem is solved numerically, with satisfactory agreement with the exact numerical results. [S1063-651X(97)08803-X]

PACS number(s): 05.45.+b

I. INTRODUCTION

Often kinetic events are controlled by the crossing of a saddle point of the energy function for the reactive system. In these cases a detailed picture of a transition can be derived by applying the Kramers theory [1] and its multidimensional generalization by Langer [2] to the Fokker-Planck (FP) equation for time-dependent distributions in the phase space. It allows a straightforward identification of the system features controlling the kinetic process: the energetic factors through the activation energy and the frictional coupling in the transmission factor that corrects the transition-state theory (TST) result for the rate coefficient.

A more complex phenomenology emerges when the kinetic event is driven by the system dynamics far away from saddle points. A well-studied case, because of its relation to the Kramers turnover, is the one-dimensional motion in the low friction limit, where the energy diffusion representation can be applied [1,3,4]. Correspondingly, the relaxation of the slow energy variable near its critical value controls the transition process independently of the proximity to the saddle point. The same control by the energy variable opens the possibility of multibarrier jumps that are absent in the overdamped regime [5–7].

Transition processes independent of the saddle-point crossing are not confined to the low friction limit of the FP equation. In fact, they are found also in the two-dimensional overdamped motions described by the FP equation of Smoluchowski type, as shown in the pioneering works by Berezhkovskii and Zitserman [8]. In the presence of a highly anisotropic diffusion matrix with properly oriented principal directions such that the saddle-point crossing is driven by its fast component, the population relaxation is controlled by the motion along the slow diffusion component far away from the saddle point. Another example has been found recently in a simple chain model with two bistable elements [9], where an anomalous type of transitions arises when the frictional coupling favors localized motions. No saddle point can be associated with such a kinetic process, which can be assimilated to crank-shaft conformational transitions of polymers [10,11].

Transitions that are not controlled by saddle-point crossing should be considered as a particular category of kinetic events requiring methods of analysis more general than the Kramers-Langer theory. As a contribution to this line of research, we intend to present a variational method capable of describing the effects of diffusion anisotropy on the two-dimensional FP equation of Smoluchowski type. It should be recalled that Berezhkovskii and Zitserman have already analyzed the main features of the transition rate in the presence of large diffusion anisotropies by projecting out the coordinate for the displacement along the fast component of the diffusion matrix [8] (see also [12] and references therein). This method, however, is not suited to recover the transition rate in the entire range of the diffusion anisotropy. Moreover, its results depend on the representation adopted for the time evolution operator. In fact, nonorthogonal transformations of the coordinates change the principal directions of the diffusion tensor and therefore also the results of the projection procedure. In the effort to overcome these difficulties an in-

*Electronic address: moro@pdchfi.chfi.unipd.it
†Electronic address: cardin@galileo.math.unipd.it

variant form of the projection procedure was proposed on the basis of the normal modes at the saddle point [13]. But this choice might not be adequate to eliminate fast motions far away from the saddle point.

We shall follow a completely different strategy based on the evaluation of the FP kinetic modes. By taking the proper linear combinations of the FP kinetic eigenfunctions, one can define a set of site-localizing functions for the populations of the stable states [9]. Then the transition rates are recovered by projecting the FP equation onto the subspace of site-localizing functions. The shape of these functions near the saddle point can be easily derived by employing the normal mode analysis of the Kramers-Langer theory [13]. Such a method, however, is not sufficient when dealing with kinetic events driven by motions far away from the saddle point since accurate approximations in the entire domain of the stochastic variables is required.

Some years ago Matkowsky and Schuss developed the boundary layer expansion originally as a tool for analyzing mean exit time problems [14,15]. The same type of procedure can provide the shape of the site-localizing functions in the entire phase space. The fundamental ingredient is the separatrix line, which is derived from the deterministic approximation of the FP equation. It determines the layer of the phase space where the local expansion should be performed. Error function profiles are obtained for the characteristic functions. Our implementation of the method is based on the covariant form of the FP equation, with the metric tensor derived from the diffusion matrix [16–19]. This ensures the invariance of the results with respect to alternative representations of the stochastic variables. Moreover, it supplies a natural choice for the distance from the separatrix to be employed in the layer expansion.

In order to test the method, a comparison will be made with the exact numerical solutions for the prototype system of a bistable quartic potential coupled to a harmonic degree of freedom [12,13,20]. In this way it becomes evident that the layer expansion about the deterministic separatrix might not provide reliable results in the presence of large diffusion anisotropies. The numerical solutions have an error-function profile, but about a line called the stochastic separatrix [21–23], which is, in general, different from the deterministic separatrix. This conclusion is supported also by recent calculations of Drozdov and Talkner on an equivalent system [24].

The main reason for this failure is precisely the choice of the deterministic separatrix as the locus for the layer expansion. In order to recover the correct transition rate in the entire range of diffusion anisotropies, a method for the determination of the stochastic separatrix should be found. One can benefit from the information that the error-function shape is preserved in the numerical solutions, by formulating on this basis a variational method for the calculation of the stochastic separatrix. The basic criterion is the minimization of the decay rate of each stable state. In our treatment also the metric tensor as a function of the position along the stochastic separatrix is included among the variational parameters in order to optimize the choice of the distance from the separatrix.

The paper is organized as follow. In Sec. II the FP equation of Smoluchowski type is introduced together with the site-localizing function method, which allows the derivation of the kinetic equations. In Sec. III the covariant formalism for the FP equation is summarized with the purpose of defining the deterministic separatrix that is independent of the stochastic variable representation. In Sec. IV the layer expansion about the deterministic separatrix is performed in order to derive the asymptotic form of the site-localizing functions. Moreover, a comparison is made with the exact numerical solutions for a model system in order to make evident the failure of this type of layer expansion for large diffusion anisotropies. In Sec. V the variational layer expansion is presented together with its numerical implementation. Also the typical results are illustrated in the same section, which is followed, in Sec. VI, by the general conclusions of this work.

II. TRANSITION RATES FROM THE FOKKER-PLANCK EQUATION

Let us first summarize the formal description of the stochastic problem in a domain $\Omega \subset R^N$ with coordinates $x \equiv (x^1, x^2, \ldots, x^N)$. The function $p^{\mathrm{eq}}(x)$ will denote the equilibrium distribution (probability density), which allows the calculation of the static average $\overline{f}$ of any observable $f(x)$ according to the equation

$$\overline{f} = \int_\Omega f(x) p^{\mathrm{eq}}(x) dx \equiv \int_\Omega f(x) p^{\mathrm{eq}}(x) \left(\prod_{i=1}^{N} dx^i \right), \quad (2.1)$$

with $p^{\mathrm{eq}}(x)$ being normalized accordingly. Often the actual physical problem determines a mean-field potential from which the equilibrium probability density is derived as the Boltzmann distribution

$$p^{\mathrm{eq}}(x) \propto \exp\{-V(x)/k_B T\}. \quad (2.2)$$

Our analysis of the stochastic dynamics will be confined to time-dependent distributions $p(x,t)$ with conserved norm at all times t,

$$\int_\Omega p(x,t) dx = 1, \quad (2.3)$$

and decaying asymptotically to equilibrium

$$\lim_{t \to +\infty} p(x,t) = p^{\mathrm{eq}}(x). \quad (2.4)$$

The appropriate Fokker-Planck equation of Smoluchowski type (i.e., the overdamped limit of the Kramers-Klein equation [17,25]) is written as

$$\partial p(x,t)/\partial t = -\Gamma p(x,t), \quad (2.5)$$

with the time evolution operator Γ specified according to a positive-definite diffusion tensor $D^{ij}(x)$, which, in general, depends on the coordinates x,

$$\Gamma = -\frac{\partial}{\partial x^i} D^{ij}(x) p^{\mathrm{eq}}(x) \frac{\partial}{\partial x^j} p^{\mathrm{eq}}(x)^{-1} \quad (2.6)$$

(throughout the paper we use the Einstein convention for the summation of repeated indices). Suitable boundary conditions should be provided in order to enforce the norm con-

servation Eq. (2.3) [26]. Notice that equilibrium distribution $p^{\mathrm{eq}}(x)$ is the only stationary solution of Eq. (2.5) under suitable ellipticity conditions. The solution of Eq. (2.5) allows the calculation of the time dependence of an observable $f(x)$,

$$\overline{f(t)}=\int_{\Omega} f(x)p(x,t)dx, \tag{2.7}$$

or of the corresponding time correlation function that is required in the analysis of spectroscopic observations [27].

In the presence of a mean potential $V(x)$ with well-defined minima separated by large barriers (in k_BT units), one expects that the system evolution can be approximately described in simple kinetic terms as an ensemble of unimolecular processes [28] for the site populations (or concentrations in the chemical language) $P_\alpha(t)$ of stable states (or sites), that is,

$$\partial P_\alpha(t)/\partial t=\sum_{\beta\neq\alpha}\left[P_\beta(t)w(\beta\rightarrow\alpha)-P_\alpha(t)w(\alpha\rightarrow\beta)\right], \tag{2.8}$$

where $w(\alpha\rightarrow\beta)$ is the rate coefficient for the transition $\alpha\rightarrow\beta$. The relation between a continuum representation such as the FP equation and the discrete stochastic process of the master equation (2.8) has been intensively studied, starting from the seminal work of Kramers [1], who clearly stated the requirement of the time-scale separation

$$1/\tau_{\mathrm{kin}}\ll 1/\tau_{\mathrm{leq}}, \tag{2.9}$$

where τ_{leq} denotes the typical time for the local equilibration of a distribution around a stable state, while τ_{kin} is the time scale of kinetic processes overcoming the energetic barriers. Only when condition (2.9) is satisfied, kinetic equations (2.8) become effective in reproducing the long-time behavior (i.e., in the time scale of τ_{kin}) of a system. Much work has been dedicated to the generalization of the Kramers results for the rate coefficients in the asymptotic limit of large barriers [3], while the numerical analysis of FP equations usually supplies the relaxation times in the range of intermediate barriers.

A general method for the derivation of the master equation (2.8) has been proposed in Ref. [9] on the basis of the spectral decomposition of the FP operator

$$\Gamma\phi_j(x)p^{\mathrm{eq}}(x)=\lambda_j\phi_j(x)p^{\mathrm{eq}}(x) \tag{2.10}$$

for $j=0,1,2,\ldots$, with the eigenvalues ordered in magnitude $\lambda_j\leqslant\lambda_{j+1}$ [29] and the eigenfunctions normalized as

$$\langle\phi_j|p^{\mathrm{eq}}|\phi_{j'}\rangle\equiv\int_{\Omega}\phi_j(x)p^{\mathrm{eq}}(x)\phi_{j'}(x)dx=\delta_{jj'}. \tag{2.11}$$

Nonequilibrium distributions can be decomposed as

$$p(x,t)=\sum_j\ \phi_j(x)p^{\mathrm{eq}}(x)e^{-\lambda_j t}\langle\phi_j|p(\cdot,0)\rangle. \tag{2.12}$$

Each eigenfunction $\phi_j(x)$ represents an independent dynamical mode of the FP model with an intrinsic relaxation time $1/\lambda_j$. Let us consider a mean-field potential with M minima at positions $\hat{x}_\alpha$ for $\alpha=1,2,\ldots,M$. Then, in order to fulfill the condition (2.9), there should be a sharp separation on the time scale between the first M modes having a kinetic character and the remaining ones describing the local equilibration processes, that is,

$$\lambda_{M-1}\ll\lambda_M, \tag{2.13}$$

having chosen $1/\tau_{\mathrm{leq}}=\lambda_M$, into correspondence with the upper bound to the relaxation times for the local equilibration modes. Under this condition, only the kinetic modes survive in the decomposition (2.12) for times $t\gg\tau_{\mathrm{leq}}$. Thus the site populations $P_\alpha(t)$ should be related to the weight of the kinetic modes in the distribution $p(x,t)$. For a precise identification, a site-localizing function $G_\alpha(x)$ is introduced for the stable state associated with the potential minimum at $\hat{x}_\alpha$ such that

$$P_\alpha(t)=\int_{\Omega}G_\alpha(x)p(x,t)dx. \tag{2.14}$$

In order to recover from Eq. (2.14) site populations devoid of components along the local equilibration modes, the site-localizing functions must belong to the kinetic subspace defined as linear combinations of eigenfunctions $\phi_j(x)$ for $j=0,1,\ldots,M-1$. The unknown coefficients of the required linear combinations can be derived by using the condition

$$G_\alpha(\hat{x}_\beta)=\delta_{\alpha,\beta}. \tag{2.15}$$

It is justified on a phenomenological ground since the site populations of Eq. (2.14) should describe the probability density integrated in the neighborhood of the stable states. Therefore, as stated by Eq. (2.15), $G_\alpha(x)$ must be unitary for $x=\hat{x}_\alpha$, while it must vanish in correspondence with the other stable states. Notice that the constraint (2.15) leads to the relation between FP eigenfunctions and the site-localizing functions

$$\phi_j(x)=\sum_\alpha\ G_\alpha(x)\phi_j(\hat{x}_\alpha) \tag{2.16}$$

for $j=0,1,2,\ldots,M-1$ [9]. In particular, for the stationary mode $\phi_0(x)=1$ one obtains

$$1=\sum_\alpha\ G_\alpha(x), \tag{2.17}$$

that is, the decomposition of unity by the ensemble of site-localizing functions. In conclusion, a precise correspondence is established according to Eq. (2.14) between time-dependent solutions of FP equations and the site populations whose time evolution can be derived exactly in the form of the kinetic equation (2.8) with the rate coefficients specified in terms of matrix elements of the time evolution operator [9]. Such a method can be considered as an example of the general procedures leading to a ''coarse theory'' (i.e., the kinetic equations) from a ''fine theory'' (i.e., the Fokker-Planck equation) [30].

The numerical calculations done for several one- and two-dimensional models [9] have shown that site-localizing functions have a steplike behavior: they are nearly constant, with values close to zero or unity, in large regions separated by

narrow layers with a steep change between the two extrema. The gradient of site-localizing functions is concentrated in such layers (and for this reason they might be called gradient layers), which bring the main contribution to the kinetic coefficients. In fact, the transition rates can be calculated as

$$P_\alpha^{\rm eq} w(\alpha\to\beta) = -\langle G_\alpha|\Gamma|p^{\rm eq}G_\beta\rangle = -\int_\Omega p^{\rm eq}(x)\,\frac{\partial G_\alpha(x)}{\partial x^i}\,D^{ij}(x)\,\frac{\partial G_\beta(x)}{\partial x^j}\,dx \tag{2.18}$$

under the approximation of neglecting the superposition integral between different site-localizing functions [9]. In Eq. (2.18), $P_\alpha^{\rm eq}$ denotes the equilibrium population of the site α to be calculated by inserting the equilibrium distribution $p^{\rm eq}(x)$ into Eq. (2.14).

This general method can be applied in two different ways. As done in Ref. [9], one can calculate exactly the site-localizing functions from the numerical eigenfunctions of the FP operator in order to establish a direct correspondence between the FP model and the kinetic equations. Otherwise, by means of Eq. (2.18), estimates of the rate coefficients can be obtained from approximate site-localizing functions. As done in the past for similar problems [5,13,22,23,31–34], one can employ the following equation for deriving the site-localizing functions:

$$\Gamma p^{\rm eq}(x) G_\alpha(x) = 0. \tag{2.19}$$

In fact, the exponential vanishing of the kinetic eigenvalues with increasing barrier heights justifies the approximation $\Gamma\phi_j(x)p^{\rm eq}(x)\simeq 0$ for the kinetic modes. The same approximation can be applied also to site-localizing functions as long as they are given as linear combination of the kinetic FP modes. However, Eq. (2.19) has only one exact solution, namely, the stationary mode $\phi_0(x)=1$. Thus nontrivial solutions of Eq. (2.19) can be derived only as local approximations, in particular, within the gradient layer in order to recover good estimates of transition rates according to Eq. (2.18). Analytical approximations of site-localizing functions have been obtained for one-dimensional models or near the saddle points in multidimensional problems [33]. In the following sections it will be shown how Eq. (2.19) can be solved for the calculation of site-localizing functions in the entire gradient layer.

An alternative method is based on the following relation for the decay rate w_α of the stable state α:

$$w_\alpha \equiv \sum_{\beta\neq\alpha} w(\omega\to\beta) = \int \frac{p^{\rm eq}(x)}{P_\alpha^{\rm eq}}\,\frac{\partial G_\alpha}{\partial x^i}\,D^{ij}(x)\,\frac{\partial G_\alpha}{\partial x^j}\,dx, \tag{2.20}$$

which derives from Eq. (2.18), the summation on G_β for $\beta\neq\alpha$ being eliminated according to Eq. (2.17). Therefore, each decay rate w_α can be considered as a functional of the corresponding site localizing function $G_\alpha(x)$. Let us consider the projection operator $\mathcal{P}$ onto an arbitrary M-dimensional subspace and a given self-adjoint operator $A=A^\dagger$ with eigenvalues bounded from below and ordered as $\lambda_j\leqslant\lambda_{j+1}$ for $j=0,1,2,\ldots$. One can demonstrate the relation for the trace operation

$$\mathrm{Tr}\{\mathcal{P}A\} \geqslant \sum_{j=0}^{M-1}\lambda_j, \tag{2.21}$$

with the equality holding if $\mathcal{P}$ projects onto the subspace determined by the first M eigenfunctions of A [35]. It can be applied to FP-Smoluchowski problems by identifying A with the symmetrized evolution operator $\tilde{\Gamma}\equiv(p^{\rm eq})^{-1/2}\Gamma(p^{\rm eq})^{1/2}$, which is self-adjoint. In this case the minimum of $\mathrm{Tr}\{\mathcal{P}\tilde{\Gamma}\}$ determines the subspace of the kinetic modes. By using the projection operator in the subspace of site-localizing functions under the condition that the superposition integrals are negligible for $\alpha\neq\beta$ [36],

$$\mathcal{P} = \sum_\alpha |(p^{\rm eq})^{1/2}G_\alpha\rangle\,\frac{1}{P_\alpha^{\rm eq}}\,\langle G_\alpha(p^{\rm eq})^{1/2}|, \tag{2.22}$$

one derives that

$$\mathrm{Tr}\{\mathcal{P}\tilde{\Gamma}\} = \sum_\alpha \langle G_\alpha|\Gamma p^{\rm eq}|G_\alpha\rangle/P_\alpha^{\rm eq} = \sum_\alpha w_\alpha \geqslant \sum_{j=0}^{M-1}\lambda_j, \tag{2.23}$$

i.e., the condition of the minimum for the sum of decay rates when the site-localizing functions span the kinetic subspace. This is precisely the optimization condition to be imposed on the site-localizing functions. On the other hand, each decay rate is a functional of the corresponding site-localizing function only according to Eq. (2.20). In conclusion, each decay rate w_α has to minimized with respect to $G_\alpha(x)$ and this provides an independent variational criterion for the optimization of a site-localizing function. One can easily show the equivalence to the variational method based on the Rayleigh quotient [23] in the case of two sites with a symmetric configuration. The implementation of this variational method will be considered in Sec. V.

III. DETERMINISTIC APPROXIMATION IN THE COVARIANT FORM

By means of a change of variables $x\mapsto y(x)\equiv(y^1,y^2,\ldots,y^N)$, alternative probabilistic representations can be generated for the same physical problem. It is then important to deal with approximation schemes that are invariant with respect to the FP representation; otherwise the results would depend on the subjective choice of the coordinate system. This calls for a covariant form of the FP equation [16–19].

In order to address the problem in all generality, let us analyze the effects of the change of variables $x\mapsto y(x)$ in the calculation of the time-dependent average (2.7),

$$\overline{f(t)} = \int_{y(\Omega)} f(y)\,|\det(\partial x^k/\partial y^{k'})|\,[p(x,t)]_{x=x(y)}\,dy, \tag{3.1}$$

where $f(y)=[f(x)]_{x=x(y)}$ and $\det(\partial x^k/\partial y^{k'})$ is the determinant of the Jacobian matrix for the transformation. In order to preserve the formal structure of the FP model, the probability density in the new representation should be defined as

$$p(y,t) \equiv |\det(\partial x^k/\partial y^{k'})|\,[p(x,t)]_{x=x(y)}. \tag{3.2}$$

Correspondingly, the FP evolution operator with the same form as Eq. (2.6) holds also in the y representation, with the new equilibrium distribution defined according to Eq. (3.2) and the transformation of the diffusion tensor [37]

$$D^{ij}(y)=\left[\frac{\partial y^i}{\partial x^{i'}}\frac{\partial y^j}{\partial x^{j'}}D^{i'j'}(x)\right]_{x=x(y)}. \tag{3.3}$$

According to Eq. (3.2), the probability density $p(x,t)$ does not transform as a scalar function but as a measure. Probability functions behaving like true scalar fields provide a more convenient framework for a theory of transition rates that should be invariant with respect to changes of the coordinate representation. This can be achieved by using the metric tensor $g_{ij}(x)$ as defined by the differential geometry on Riemann manifolds [39], which transforms as

$$g_{ij}(y)=\left[\frac{\partial x^{i'}}{\partial y^i}\frac{\partial x^{j'}}{\partial y^j}g_{i'j'}(x)\right]_{x=x(y)}. \tag{3.4}$$

By means of the determinant $g(x)$ of the metric tensor

$$g(x)\equiv\det[g_{ij}(x)]=g(y)/[\det(\partial x^k/\partial y^{k'})]^2, \tag{3.5}$$

one can define a new probability distribution

$$p_s(x;t)\equiv p(x;t)/\sqrt{g(x)}, \tag{3.6}$$

which behaves like a true scalar field (and for this reason it will be called as scalar distribution in the following), as one immediately derives from Eq. (3.2),

$$p_s(y,t)=[p_s(x,t)]_{x=x(y)}. \tag{3.7}$$

In particular, a new mean-field potential $V_s(x)$, which also behaves like a scalar field, can be recovered from the scalar equilibrium distribution $p_s^{\mathrm{eq}}(x)$,

$$p_s^{\mathrm{eq}}(x)\equiv\frac{p^{\mathrm{eq}}(x)}{\sqrt{g(x)}}\propto\exp\{-V_s(x)/k_BT\},$$
$$V_s(x)=V(x)+\frac{k_BT}{2}\ln g(x). \tag{3.8}$$

In this way the change of variables in the integral for the calculation of the averages

$$\overline{f(t)}=\int_\Omega dx\sqrt{g(x)}f(x)p_s(x,t) \tag{3.9}$$

does not require the Jacobian of the transformation because of the factor $\sqrt{g}$.

The time evolution equation for the scalar distribution is readily derived from Eqs. (2.5) and (2.6):

$$\partial p_s(x,t)/\partial t=-\Gamma_s p_s(x,t), \tag{3.10}$$

with the evolution operator Γ_s in the covariant form

$$\Gamma_s=-\frac{1}{\sqrt{g(x)}}\frac{\partial}{\partial x^i}\sqrt{g(x)}D^{ij}(x)p_s^{\mathrm{eq}}(x)\frac{\partial}{\partial x^i}p_s^{\mathrm{eq}}(x)^{-1}. \tag{3.11}$$

The decomposition of the time evolution operator

$$\Gamma_s=\Gamma_{\mathrm{LB}}+\Gamma_d \tag{3.12}$$

is obtained by including in Γ_{LB} the terms independent of the equilibrium distribution $p_s^{\mathrm{eq}}(x)$,

$$\Gamma_{\mathrm{LB}}=-\frac{1}{\sqrt{g(x)}}\frac{\partial}{\partial x^i}\sqrt{g(x)}D^{ij}(x)\frac{\partial}{\partial x^j}. \tag{3.13}$$

The operator Γ_{LB}, which determines the probability evolution in the presence of a homogeneous equilibrium distribution $p_s(x)=$const, is the standard Laplace-Beltrami operator for second-order differential forms [39] when the metric tensor and the diffusion tensor are identical. The second term in Eq. (3.12),

$$\Gamma_d\equiv\frac{1}{\sqrt{g(x)}}\frac{\partial}{\partial x^i}\sqrt{g(x)}b^i(x), \tag{3.14}$$

which includes the drift field

$$b^i(x)\equiv D^{ij}(x)p_s(x)^{-1}\frac{\partial p_s(x)}{\partial x^j}=-\frac{D^{ij}(x)}{k_BT}\frac{\partial V_s(x)}{\partial x^j}, \tag{3.15}$$

is a first-order differential operator like the Liouville operator of classical mechanics and therefore can be associated with the deterministic dynamics of the system. If the contribution of the Laplace-Beltrami operator is neglected, deterministic solutions for the nonequilibrium probability are derived in the form

$$p_s(x;t)=\delta(x-\tilde{x}(t)), \tag{3.16}$$

where $\delta(x-x')$ is the Dirac delta distribution associated with integrals like Eq. (3.9):

$$f(x')=\int_\Omega\sqrt{g(x)}\delta(x-x')f(x)dx, \tag{3.17}$$

while $\tilde{x}(t)$ is a trajectory in the phase space Ω to be calculated from the drift

$$d\tilde{x}^{\,i}(t)/dt=b^i(\tilde{x}(t)). \tag{3.18}$$

A simple picture based on the flow in the phase space is recovered from Γ_d, but this alone cannot give rise to acceptable solutions of the FP equation because of the absence of relaxation to the equilibrium according to Eq. (2.4). On the other hand, the deterministic solutions such as Eq. (3.16) might be good approximations in the presence of large drifts when the contributions of the Laplace-Beltrami operator could be neglected.

It should be emphasized that trajectories (3.18) are invariant with respect to change of variables because of the covariant forms of the drift (3.15) and of the deterministic operator Γ_d of Eq. (3.14) as well. This invariance relies on the

existence of a metric tensor $g_{ij}(x)$ (the boundary layer expansion also requires a metric tensor for the definition of an invariant distance; see Sec. IV). For a given physical problem, one might choose it among the observables having the correct tensorial properties. On the other hand, only the diffusion tensor is available for this purpose in the formal analysis of the FP equation of Smoluchowski type. Therefore, as done in past analyses of the covariant FP equation [16–19], we shall first use the metric tensor

$$g_{ij}(x)=D_{ij}(x), \tag{3.19}$$

where D_{ij} denotes the inverse of matrix D^{ij}. Qualitatively, this is equivalent to scaling the distances according to the square root of the mean-squared displacements in the absence of the drift.

In order to show some typical trajectories deriving from Eq. (3.18), let us introduce the two-dimensional model that will be used as the test case throughout the paper. It is characterized by the bistable potential

$$V(x^1,x^2)/k_BT=\Delta[(x^1/x_0)^2-1]^2+2\Delta[(x^2/x_0)\cot(\alpha/2)]^2, \tag{3.20}$$

where $(\pm x_0,0)$ are the locations of the minima, Δ is the barrier height in k_BT units of the saddle point at the origin, and α is the aperture angle of the saddle point (i.e., the angle between the two branches of the equipotential contour passing through the saddle point for $x^1\geqslant 0$) and a constant but anisotropic diffusion tensor

$$D^{ij}=D_s d_s^i d_s^j+D_f d_f^i d_f^j, \tag{3.21}$$

where d_s and d_f are the principal directions ($d_k^i d_{k'}^i=\delta_{k,k'}$) of its slow D_s and fast D_f components ($D_s\leqslant D_f$). The angle θ between the fast component direction d_f with respect to the $x^2=0$ axis is used to specify its orientation. All the reported results, except those of Fig. 7, are derived with the set of parameters

$$\Delta=5, \quad \alpha=90°, \quad \theta=35°, \tag{3.22}$$

which ensures the validity of the time-scale separation (2.13) for the kinetic modes, λ_2/λ_1 being in this case greater than 30 independently of the anisotropy ratio D_f/D_s [40]. Since $\theta<\alpha/2$, the fast diffusion component D_f mainly drives the saddle-point crossing, and with large values of the diffusion anisotropy D_f/D_s one can investigate the kinetic control by the the slow diffusion component far away from the saddle point [13]. In the analysis of the effects of diffusion anisotropy, another model has often been employed characterized by a diagonal diffusion matrix, but with a potential function that includes also a bilinear coupling between the two coordinates [12,20,24]. However, one can easily demonstrate the equivalence of the two models by performing the linear change of variables leading to a separated potential like that of Eq. (3.20). In Fig. 1 a set of equipotential lines of Eq. (3.20) are represented, while some trajectories of Eq. (3.18) are drawn in Fig. 2 for $D_f/D_s=10$. The deterministic flow has three stationary points: the stable states of the system, which correspond to the potential minima $\hat{x}_1=(-x_0,0)$ and $\hat{x}_2=(x_0,0)$, and the saddle point at the origin, which is an unstable point of hyperbolic type. All the trajectories converge to the stable states, except those reaching the saddle point and defining the deterministic separatrix of the problem.

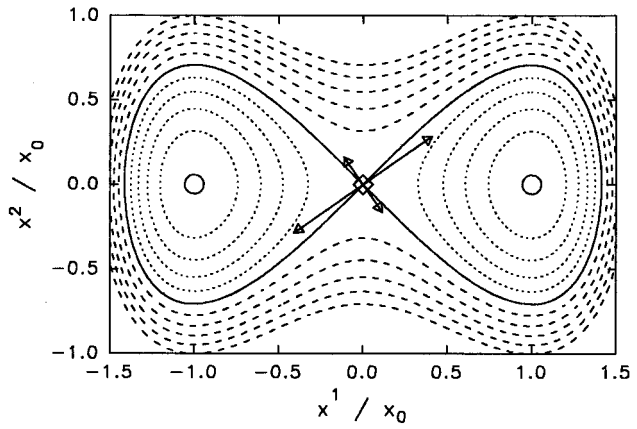

FIG. 1. Equipotential contours from Eq. (3.20) with parameters of Eq. (3.22) for $V/k_BT=1$–4 (dotted lines), for $V/k_BT=5$ (continuous line), and for $V/k_BT=6$–10 (dashed lines). The potential minima are located at the circles, while the diamond identifies the saddle point. The long (short) arrows are directed along the fast (slow) principal direction of the diffusion matrix (3.21).

The deterministic approximation for the site-localizing functions is derived by neglecting the Laplace-Beltrami contribution Γ_{LB} in Eq. (2.18),

$$b^i(x)\,\frac{\partial}{\partial x^i}\,G_\alpha(x)=0, \tag{3.23}$$

that is, by isolating the leading term of Eq. (2.19) in the low-temperature limit $T\to 0$ for fixed $D^{ij}(x)$ and $V_s(x)$. Then each site-localizing function should be constant along a trajectory derived from Eq. (3.18). Since all the trajectories (without considering those of the separatrix) converge to stable states, a site-localizing function can assume only a finite set of values. The standard partition [15] of the phase space is done by including into the same domain of attraction Ω_α all the points that belong to trajectories converging to the stable state $\hat{x}_\alpha$. The separatrix $\partial\Omega_{\alpha\beta}$ between two adjacent

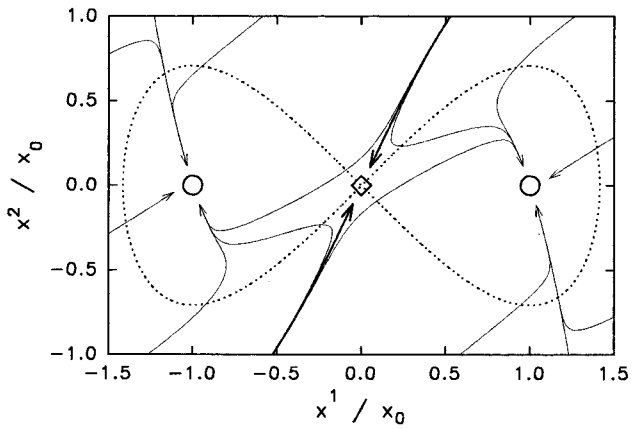

FIG. 2. Set of trajectories (3.18) for $D_f/D_s=10$, with the heavy lines identifying those of the separatrix. The dotted line is the equipotential contour through the saddle point.

domains of attraction Ω_α and Ω_β is associated with the trajectories ending up in unstable points of hyperbolic type (see Fig. 2). Correspondingly, the site-localizing functions in the deterministic approximation are constant within each domain of attraction and, by taking into account Eq. (2.15), one derives the explicit form

$$G_\alpha(x)=\begin{cases}1 & \text{if } x\in\Omega_\alpha\\ 0 & \text{otherwise,}\end{cases} \tag{3.24}$$

that is, a function that is discontinuous along the boundary of the domain Ω_α. Because of the discontinuity, such a simple approximation cannot be employed to derive the transition rate according to Eq. (2.18). The calculation of transition rates requires a continuous form of site-localizing functions with smooth changes from the extrema 0 and 1 when passing through the separatrix (see Sec. IV). On the contrary, the calculation of equilibrium site populations according to Eq. (2.14) is possible since the derivatives of site-localizing functions are not required. Because of the step function behavior (3.24), this is equivalent to the integration of $p_s^{\rm eq}(x)$ within a domain of attraction

$$\begin{aligned}P_\alpha^{\rm eq}&=\int_\Omega \sqrt{g(x)}G_\alpha(x)p_s^{\rm eq}(x)dx\\ &=(1/Z)\int_{\Omega_\alpha}\sqrt{g(x)}\exp\{-V_s(x)/k_BT\}dx,\end{aligned} \tag{3.25}$$

where Z is the normalization

$$Z\equiv\int_\Omega\sqrt{g(x)}\exp\{-V_s(x)/k_BT\}dx. \tag{3.26}$$

Sometimes it is convenient to use the free energy F_α to account for the equilibrium population of site α,

$$P_\alpha^{\rm eq}=\exp(-F_\alpha/k_BT)/Z. \tag{3.27}$$

An analytical estimate of the free energy F_α,

$$F_\alpha\simeq V(\hat{x}_\alpha)+\frac{k_BT}{2}\ln\left|\det\left\{\left[\frac{\partial^2V_s(x)}{\partial x^i\partial x^j}\right]_{x=\hat{x}_\alpha}\Big/2\pi k_BT\right\}\right|, \tag{3.28}$$

is derived by applying the Laplace method [41] to the integral (3.25) after a parabolic expansion of the potential about the stable state

$$V_s(x)\simeq V_s(\hat{x}_\alpha)+\frac{1}{2}\left[\frac{\partial^2V_s(x)}{\partial x^i\partial x^j}\right]_{x=\hat{x}_\alpha}(x^i-\hat{x}_\alpha^i)(x^j-\hat{x}_\alpha^j). \tag{3.29}$$

In Sec. II we identified the stable states with the minima of the mean-field potential $V(x)$ for the equilibrium distribution (2.2). This is certainly legitimate in the presence of a constant diffusion tensor, but in the general case the correct choice is necessarily provided by the stationary points of Eq. (3.18) since only they are truly invariant with respect to the change of variables. Therefore, the minima of the scalar potential $V_s(x)$ [Eq. (3.8)] should be considered as the true stable states $\hat{x}_\alpha$, thus taking into account the possible effects of the coordinate dependence of the diffusion tensor.

IV. LAYER EXPANSION ABOUT THE DETERMINISTIC SEPARATRIX

The method of the boundary layer expansion has been developed by Matkowsky and Schuss [14,15] originally for the calculation of the mean exit time from a given domain. The same method can be employed to derive a smooth form of site-localizing functions across the deterministic separatrix, thus overcoming the discontinuity of approximation (3.24). In this section the bistable problem in two dimensions will be examined in all generality, and the comparison with the exact numerical results will be done for the model system specified by Eqs. (3.20) and (3.21). The invariance of the layer expansion will be ensured by identifying the metric tensor with the diffusion tensor according to Eq. (3.19). We shall continue to use explicitly the notation for the metric tensor in order to introduce a geometrical description of the layer about the separatrix, which can be employed also with alternative choices of $g_{ij}(x)$ (see Sec. V).

The layer expansion takes a simple form by using as coordinates the displacement l along the separatrix and the distance ρ from the separatrix. Let us specify the coordinates of the separatrix $\partial\Omega_{\alpha\beta}$ in a parametrized form

$$x(\lambda)=(x^1(\lambda),x^2(\lambda)), \tag{4.1}$$

with a real parameter λ. Because of the Riemann structure attached to the manifold Ω, one can employ the invariant distance

$$l(\lambda)=\int_0^\lambda d\lambda'\left[\frac{dx^i(\lambda')}{d\lambda'}\,g_{ij}(x(\lambda'))\,\frac{dx^j(\lambda')}{d\lambda'}\right]^{1/2} \tag{4.2}$$

instead of the arbitrary parameter λ to denote the points at the separatrix. The unit vector locally tangent to the separatrix is then defined as

$$T^i(l)\equiv dx^i(l)/dl, \tag{4.3}$$

with the normalization derived from Eq. (4.2),

$$T^i(l)g_{ij}(x(l))T^j(l)=1. \tag{4.4}$$

As long as the separatrix is a particular solution of the deterministic equation (3.18), the drift at the separatrix must be parallel to $T^i(l)$, that is,

$$b^i(x(l))=B(l)T^i(l), \tag{4.5}$$

with the scalar function $B(l)$ determining the magnitude of the drift. The metric tensor allows the definition of the unit vector $N^i(l)$ orthogonal to separatrix at a given position l,

$$\begin{aligned}&N^i(l)g_{ij}(x(l))T^j(l)=0,\\ &N^i(l)g_{ij}(x(l))N^j(l)=1.\end{aligned} \tag{4.6}$$

Then the displacement along $N^i(l)$ supplies the natural choice for the distance ρ from the separatrix. In order to deal with an invariant distance ρ (which is, independent of the starting coordinate representation x), one should introduce

the geodetic line $x^i(\mu)$ orthogonal to the separatrix at the position l as solution of the equation

$$\frac{d^2x^i(\mu)}{d\mu^2}+\Gamma^i_{jk}(x(\mu))\,\frac{dx^j(\mu)}{d\mu}\,\frac{dx^k(\mu)}{d\mu}=0, \tag{4.7}$$

with $[dx^i(\mu)/d\mu]_{\mu=0}\propto N^i(l)$ as the boundary condition and Γ^i_{jk} the standard Christoffel symbol derived from the metric tensor [39]. Notice that the geodetics are straight lines if the metric tensor is constant since $\Gamma^i_{jk}=0$ in such a case. The invariant distance ρ from the separatrix is calculated by applying to these geodetics the definition of Riemann arclength measure like in Eq. (4.2).

Distances ρ and l supply the natural representation for performing the layer expansion, and in the following we shall employ the FP operator expressed in these coordinates:

$$y\equiv(y^1,y^2)\equiv(\rho,l). \tag{4.8}$$

Correspondingly, both the diffusion tensor and the drift field assume a very simple form at the separatrix. By taking into account that

$$\left(\frac{\partial x^i(y)}{\partial y^1}\right)_{\rho=0}=N^i(l),\quad \left(\frac{\partial x^i(y)}{\partial y^2}\right)_{\rho=0}=T^i(l), \tag{4.9}$$

a unitary metric tensor (and therefore the diffusion tensor also) is obtained from Eq. (3.4) for $\rho=0$,

$$[g_{ij}(y)]_{\rho=0}=\delta_{ij},\quad [D^{ij}(y)]_{\rho=0}=\delta^{ij}. \tag{4.10}$$

From Eq. (4.5) one derives the drift in the y representation

$$[b^1(y)]_{\rho=0}=0,\quad [b^2(y)]_{\rho=0}=B(l), \tag{4.11}$$

with, according to Eq. (3.15), $B(l)$ determined by the gradient of the scalar potential $V_s(y)$ along the separatrix

$$B(l)=-\frac{1}{k_BT}\left(\frac{\partial V_s(y)}{\partial l}\right)_{\rho=0}. \tag{4.12}$$

The dependence of $b^1(y)$ on the distance from the separatrix is specified for small ρ as

$$b^1(y)=\omega(l)\rho+O(\rho^2),\quad \omega(l)\equiv\left[\frac{\partial b^1(y)}{\partial\rho}\right]_{\rho=0}, \tag{4.13}$$

with $1/\omega(l)$ determining the time scale for the motion orthogonal to the separatrix. It should be stressed out that coordinates y employed in the layer expansion are independent of the starting representation x. Because of the invariant character of the measure employed in their definition, the same values of y would be recovered for a given point of Ω also by using a different set of coordinates $x'=x'(x)$ for the FP equation. This feature ensures the invariance of the layer expansion with respect to the different but equivalent representations of the same stochastic problem.

Let us now analyze Eq. (2.19) for the calculation of site-localizing functions in the y representation

$$\left[\frac{1}{\sqrt{g(y)}}\,\frac{\partial}{\partial y^i}\,\sqrt{g(y)}D^{ij}(y)\,\frac{\partial}{\partial y^j}+b^i(y)\,\frac{\partial}{\partial y^i}\right]G_\alpha(y)=0 \tag{4.14}$$

by performing a low-temperature expansion, that is, by scaling the temperature as $T\to\epsilon T$ for fixed $D^{ij}(y)$ and $V_s(y)$. Because of the relation (3.15), this is equivalent to scaling in Eq. (4.14) the drift as $b^i(y)\to b^i(y)/\epsilon$ by keeping the diffusion tensor fixed. The layer expansion is implemented by invoking the working hypothesis that in the limit $\epsilon\to0$ the site-localizing function G_α is invariant after an $\sqrt{\epsilon}$ scaling of the distance from the separatrix, that is,

$$G_\alpha(\rho,l)=\hat G_\alpha(\rho/\sqrt{\epsilon},l)+O(\sqrt{\epsilon}), \tag{4.15}$$

with $\hat G_\alpha(\cdot,\cdot)$ independent of ϵ. The leading term of Eq. (4.14) is of first order with respect to $1/\epsilon$ and can be isolated by considering the limit

$$\lim_{\epsilon\to0}\left[\frac{\epsilon}{\sqrt{g(y)}}\,\frac{\partial}{\partial y^i}\,\sqrt{g(y)}D^{ij}(y)\,\frac{\partial}{\partial y^j}+b^i(y)\,\frac{\partial}{\partial y^i}\right]\hat G_\alpha(\rho/\sqrt{\epsilon},l)=0. \tag{4.16}$$

After the substitution $\rho\to\rho\sqrt{\epsilon}$ the final equation for the asymptotic site-localizing functions is derived in the form

$$\left[\frac{\partial^2}{\partial\rho^2}+B(l)\,\frac{\partial}{\partial l}+\omega(l)\rho\,\frac{\partial}{\partial\rho}\right]G_\alpha(\rho,l)=0, \tag{4.17}$$

where $\epsilon=1$ has been restored. The third term of Eq. (4.17) derives from the contribution

$$\lim_{\epsilon\to0}b^1(\rho,l)\,\frac{\partial\hat G_\alpha(\rho/\sqrt{\epsilon},l)}{\partial\rho}=\frac{\partial\hat G_\alpha(\rho',l)}{\partial\rho'}\lim_{\epsilon\to0}\frac{b^1(\rho'\sqrt{\epsilon},l)}{\sqrt{\epsilon}}=\rho'\,\frac{\partial\hat G_\alpha(\rho',l)}{\partial\rho'}\lim_{\rho\to0}\frac{b^1(\rho,l)}{\rho}, \tag{4.18}$$

with $\rho'=\rho/\sqrt{\epsilon}$ and the last limit being the derivative of the orthogonal component of the drift since $b^1(0,l)=0$.

The trial function for the solution of Eq. (4.17) is provided by the error-function profile with respect to the distance ρ

$$G_\alpha(\rho,l)=\frac{1}{2}\pm\frac{1}{2}\,\mathrm{erf}\{\rho/2\sigma(l)\}=\frac{1}{2}\pm\frac{1}{\sqrt{\pi}}\int_0^{\rho/2\sigma(l)}e^{-t^2}dt, \tag{4.19}$$

with the width $\sigma(l)$ dependent on the position l along the separatrix. The proper choice of the sign in Eq. (4.19) ensures an asymptotic behavior for $\rho\to\pm\infty$ in agreement with Eq. (2.15). After substitution of the trial function into Eq. (4.17), the following differential equation is recovered for the width:

$$B(l)\,\frac{d\sigma(l)^2}{dl}=2\,\omega(l)\sigma(l)^2-1. \tag{4.20}$$

Near the saddle point l_S of $V_s(y)$, where $B(l_S)=0$, a linear expansion $B(l)\propto(l-l_S)$ holds. Therefore, the only finite solution of Eq. (4.20) at l_S is given by the equation

$$1/\sigma(l_S)^2=2\omega(l_S), \tag{4.21}$$

which supplies the boundary condition for solving Eq. (4.20) [42]. Even if the analytical solution can be written explicitly [14,15]

$$\sigma(l)^2=-\int_{l_S}^{l}\frac{dl'}{B(l')}\exp\left(\int_{l'}^{l}dl''2\omega(l'')/B(l'')\right), \tag{4.22}$$

the numerical solution of Eq. (4.20) with a finite-difference scheme is more convenient when functions $\omega(l)$ and $B(l)$ have a complicated dependence on l. On the other hand, Eq. (4.22) allows one to demonstrate that $\sigma(l)^2$ is always a positive quantity.

Once the profile of site-localizing functions has been determined, one can calculate the transition rate $w(\alpha\rightarrow\beta)$ according to Eq. (2.18). By using the y representation and taking into account that according to Eq. (2.17) $G_\beta(y)=1-G_\alpha(y)$, the following equation is derived:

$$w(\omega\rightarrow\beta)=\int dy\sqrt{g(y)}\exp\{[F_\alpha-V_s(y)]/k_BT - \rho^2/2\sigma(l)^2\}f_i(y)D^{ij}(y)f_j(y)/4\pi\sigma(l)^2, \tag{4.23}$$

where

$$f_1=1,\quad f_2(y)=-\frac{\rho}{\sigma(l)}\frac{d\sigma(l)}{dl} \tag{4.24}$$

and the normalization Z [Eq. (3.26)] of the equilibrium distribution is eliminated by means of the free energy F_α of Eq. (3.27).

In agreement with the hypothesis of the layer expansion, one can simplify the integration in Eq. (4.23) by (i) evaluating the diffusion tensor at the separatrix $D^{ij}(y)\simeq\delta^{ij}$ [and, correspondingly, $g(y)\simeq1$] and (ii) performing a parabolic expansion of the scalar potential with respect to the distance ρ,

$$V_s(y)\simeq[V_s(y)]_{\rho=0}+\frac{\rho^2}{2}\left[\frac{\partial^2V_s(y)}{\partial\rho^2}\right]_{\rho=0} =V_s(l)+\frac{\rho^2}{2}V_s^{(2)}(l), \tag{4.25}$$

with $V_s(l)=V_s(x(l))$. Notice that the first-order term is missing because $[b^1(y)]_{\rho=0}=0$. In this way the transition rate can be derived by means of an integration along the separatrix

$$w(\alpha\rightarrow\beta)=\int dl\ \exp\{[F_\alpha-V_s(l)]/k_BT\}\frac{\eta(l)}{\sqrt{8\pi}\sigma(l)^2} \times\{1+[\eta(l)\dot\sigma(l)/\sigma(l)]^2\}, \tag{4.26}$$

where $\dot\sigma(l)\equiv d\sigma(l)/dl$ and

$$1/\eta(l)\equiv\sqrt{1/\sigma(l)^2+V_s^{(2)}(l)/k_BT}. \tag{4.27}$$

In the presence of a sharply defined minimum of $V_s(l)$ corresponding to the saddle point at l_S, one can perform an analytical integration on l after the parabolic expansion of $V_s(l)$, thus recovering the result of the Kramers-Langer analysis of the flux at the saddle point

$$w(\omega\rightarrow\beta)=\frac{\omega(l_S)}{2\pi}\exp\{-[F_S-F_\alpha]/k_BT\}, \tag{4.28}$$

where the free energy F_S of the saddle point is given by the same relation (3.28), but with $x(l_S)$ substituting $\hat x_\alpha$. If the diffusion tensor is independent of the coordinates x, in which case $V_s(x)$ and $V(x)$ differ by an additive constant according to Eq. (3.8), Eq. (4.28) is equivalent to the result of the standard Kramers-Langer analysis of the flux over the saddle point [3]. However, some differences could emerge with diffusion tensors explicitly dependent on the coordinates because of the different functional form of the mean-field potential $V(x)$ and its scalar counterpart $V_s(x)$ [Eq. (3.8)]. In this case, the advantage of using a covariant formalism should be evident because it leads to an approximation (4.28) that is independent of the coordinate representation.

The full integration in Eq. (4.26) along the separatrix is required when the transition process is driven by the system dynamics far away from the saddle point. This is the case of our model system with the parameters of Eq. (3.22). In order to perform a comparison, standard numerical methods (see Appendix B of Ref. [9]) have been used to calculate both the exact transition rate $w=w(1\rightarrow2)=w(2\rightarrow1)$ and the exact site-localizing functions. A simple discretization of the integral in Eq. (4.26) allows the calculation of the approximate rate when a finite-difference scheme is employed for the solution of Eq. (4.20). Notice that the explicit relation

$$V^{(2)}(l)=N^i(l)N^j(l)\left(\frac{\partial^2V(x)}{\partial x^i\partial x^j}\right)_{\rho=0}=-\omega(l) \tag{4.29}$$

can be used for the frequency factor when the metric tensor is constant.

In Fig. 3 the numerical results from Eq. (4.26) and the approximations (4.28) for the transition rate are reported as a function of the diffusion anisotropy D_f/D_s. For large diffusion anisotropies, the transition rate is overestimated by the Kramers-Langer result, which takes into account only the crossing near the saddle point. This is the situation where the layer expansion should be useful as long as it describes the diffusion process far away from the saddle point. However, as the data of Fig. 3 clearly show, it improves only partially the Kramers-Langer result, and large deviations from the exact rates are found for increasing ratios D_f/D_s.

In order to understand this finding, the exact site-localizing functions obtained numerically should be analyzed. As matter of fact, they have an error-function profile in agreement with Eq. (4.19) (see Fig. 4), but about a line different from the deterministic separatrix. This features clearly emerges by comparing the deterministic separatrix with the so-called stochastic separatrix [43], defined as the points where the exact site-localizing functions take the value $\frac{1}{2}$, i.e., halfway between their extrema [44],

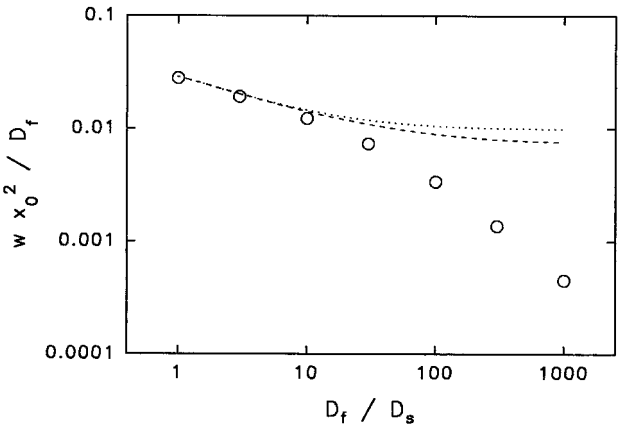

FIG. 3. Scaled transition rate as a function of the anisotropy ratio D_f/D_s with parameters of Eq. (3.22). Circles, exact numerical results; dashed line, Eq. (4.26) from the layer expansion about the deterministic separatrix; dotted line, Eq. (4.28) from the Kramers-Langer theory.

$$G_\alpha(x(\lambda))=\tfrac{1}{2}. \tag{4.30}$$

Of course, the two separatrices should be identical if the exact site-localizing functions is reproduced by the approximation (4.19). In Fig. 5 both deterministic and stochastic separatrices are represented in two cases. In the low range of diffusion anisotropy when the layer expansion reproduces correctly the transition rate, the two separatrices are close enough. On the contrary, for large values of D_f/D_s when significant deviations are found by calculating the rate according to Eq. (4.26), the deterministic and stochastic separatrices are quite different. Then one can draw the conclusion that the use of the deterministic separatrix is mainly responsible for the errors in the calculations of the transition rate. This calls for a generalization of the layer expansion, which should be able to determine also the correct stochastic separatrix, instead of using the deterministic separatrix from the very beginning (see Sec. V).

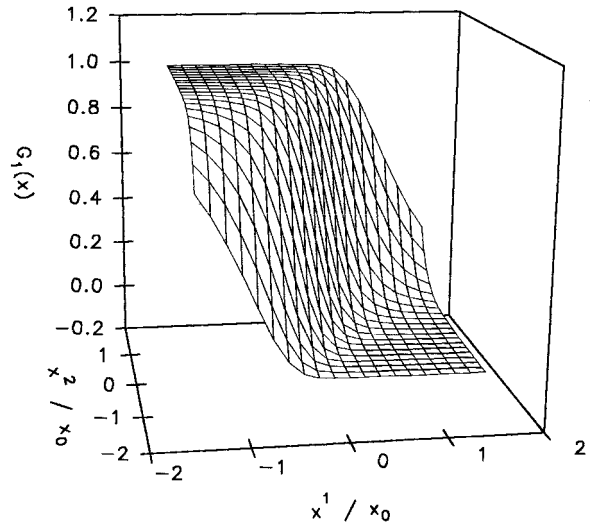

FIG. 4. Profile of the site localizing function $G_1(x)$ obtained numerically for $D_f/D_s=10$.

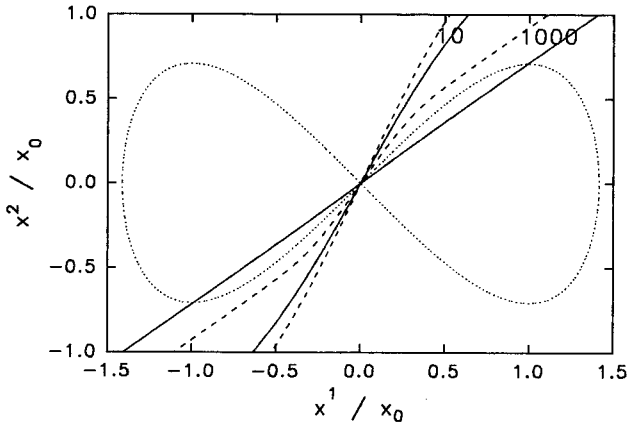

FIG. 5. Deterministic separatrices (dashed line) and stochastic separatrices (continuous line) for two values of diffusion anisotropy D_f/D_s as reported in the figure.

One should notice that the stochastic separatrix in Fig. 5 for $D_f/D_s=1000$ is almost a straight line parallel to the fast principal direction d_f of the diffusion tensor (3.21) (see Fig. 1 for the direction of d_f). Also the deterministic separatrix far from the origin tends to be parallel to d_f because in the solution of Eq. (3.18) the direction of fast diffusion is favored by a strong diffusion anisotropy. This behavior, however, cannot be realized near the origin because it would correspond to a deterministic trajectory climbing the potential energy.

Another drawback of the previous treatment derives from the small width of the layer where approximation (4.19) can be employed. As matter of fact, the coordinates $y=(\rho,l)$ are not suited to represent the FP equation in the entire domain Ω of the stochastic variables since the geodetic lines orthogonal to the separatrix in two different points l and l' might cross, say, in corresponding with distances ρ and ρ', respectively. Then the crossing point would be labeled by two sets of coordinates (l,ρ) and (l',ρ'). In order to deal with a uniquely defined representation y, only the points of Ω devoid of crossings among the orthogonal geodetics should be considered. The y representation can be adopted only on a strip of Ω defined by the condition

$$|\rho|<\rho_c(l), \tag{4.31}$$

where $\rho_c(l)$ is the shortest of all the crossing distances along the geodetic orthogonal to the separatrix at l. With large diffusion anisotropies, some crossing points are very close to the deterministic separatrix, as shown in Fig. 6, where line connecting the points at $\rho_c(l)$ is displayed for $D_f/D_s=1000$. Correspondingly, the layer expansion has to be confined to a very narrow strip much smaller than the width σ of the site-localizing functions. The ultimate reason for such a behavior is the identification (3.19) of the metric tensor with the diffusion tensor. In fact, the presence of a large diffusion anisotropy generates an orthogonal direction $N^i(l)$ almost parallel to $T^i(l)$ when there are small differences between the orientations of the separatrix and of the fast diffusion direction d_f. Such a distortion of the orthogonal direction lowers the crossing distances $\rho_c(l)$. Also this drawback should be

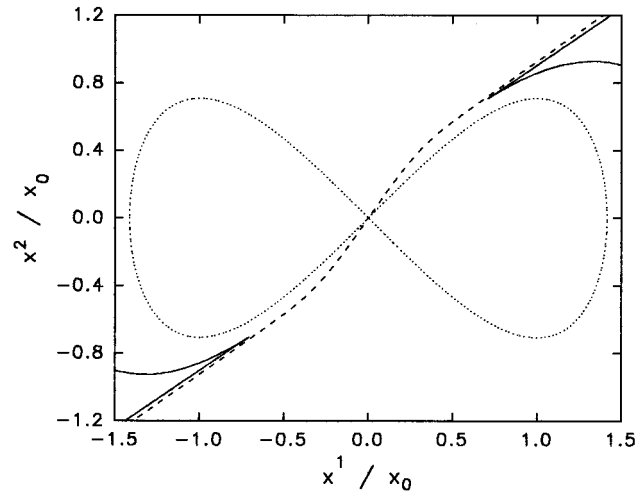

FIG. 6. Deterministic separatrix (dashed line) and the points at the crossing distance $\rho_c(l)$ (continuous line) for $D_f/D_s=10$.

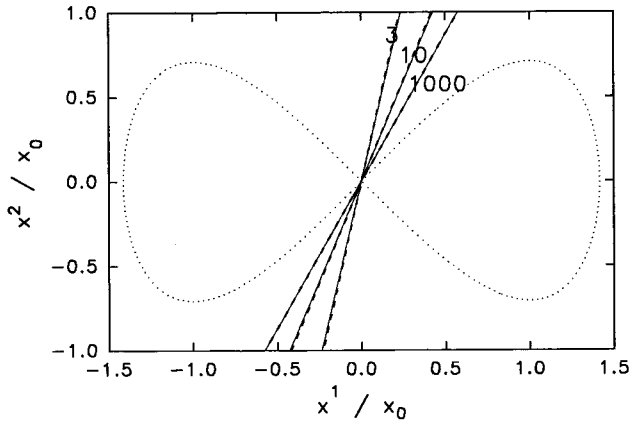

FIG. 7. Deterministic separatrices (dashed line) and stochastic separatrices (continuous line) for parameters of Eq. (4.34) and diffusion anisotropies D_f/D_s reported in the figure.

eliminated in a more accurate description of site-localizing functions by means of a different choice of the metric tensor.

Even if the crossing distance ρ_c can be enlarged by changing the metric tensor, in any case finite values are expected. Therefore, approximation (4.19) cannot be used in the entire phase space and some sort of cutoff should be introduced, while still preserving the continuity of site-localizing function in order to recover finite transition rates from Eq. (2.18). After introducing a parameter ξ satisfying the upper bound

$$\xi \leqslant \inf_l[\rho_c(l)/\sigma(l)], \tag{4.32}$$

the trial function (4.19) can be modified as

$$G_\alpha = \begin{cases} \dfrac{1}{2} \pm \dfrac{1}{2} \dfrac{\operatorname{erf}\{\rho/2\sigma(l)\}}{\operatorname{erf}(\xi/2)} & \text{if } |\rho/\sigma(l)| < \xi \\ 0 \text{ or } 1 & \text{otherwise.} \end{cases} \tag{4.33}$$

This function is still a solution of Eq. (4.17) within the layer $|\rho| < \xi\sigma(l)$; it is continuous and has the correct asymptotic behavior for $\rho \to \pm\infty$. One can use such a modified site-localizing function in the calculation of the transition rate, and the same result (4.26) is obtained besides the correction factor $1/\operatorname{erf}(\xi/2)^2$. This, however, does not improve at all the agreement with the exact kinetic rates shown in Fig. 3 since the correction factor is always greater than unity.

The comparison in Figs. 3 and 5 between the layer expansion method and the exact numerical results is done for the parameters of Eq. (3.22) of the model. However, the same general behavior is found in other cases as long as $|\theta| < \alpha/2$ (and provided the separability between kinetic modes and local equilibration modes [40]), that is, when the fast diffusion direction is within the aperture of the saddle point. In particular, the stochastic separatrix tends always to be aligned along the fast diffusion direction d_f for large diffusion anisotropies, this direction not being allowed for the deterministic separatrix near the saddle point. A completely different behavior is found when $|\theta| > \alpha/2$ for fast diffusion directions outside the aperture of the saddle point. In fact, satisfactory results are recovered from the layer expansion about the deterministic separatrix that is very close to the stochastic separatrix obtained numerically, as shown in Fig. 7 for the angle θ changed to 60°, i.e., for parameters

$$\Delta = 5, \quad \alpha = 90°, \quad \theta = 60°. \tag{4.34}$$

Also in this case the stochastic separatrix tends to be aligned with d_f for D_f/D_s large enough, but also the deterministic separatrix follows the same direction since for $|\theta| > \alpha/2$ this does not violate the requirement of deterministic motion with decreasing potential energy. Notice that, given the profiles of the separatrix, only the points near the saddle point contribute significantly to the rate in the integral (4.26) and therefore approximation (4.28) holds.

V. VARIATIONAL LAYER EXPANSION

Let us return to our model system with the parameters of Eq. (3.22). The general message arising from the comparison with the numerical results is that the error-function profile is reproduced by the exact site-localizing functions, but in locations far from the deterministic separatrix when the diffusion anisotropy is large enough. This suggests that the functional form (4.33) can be used to approximate the site-localizing functions provided that the separatrix is considered as a parametric function to be optimized in the variational framework based on the minimum of the decay rate (2.20). In order to avoid distortions of the directions for the distance ρ towards the separatrix, also the metric tensor should be optimized. Drozdov and Talkner too have proposed a variational method relying on the error-function profile of the characteristic functions, but assuming a straight stochastic separatrix [24].

The general method will be presented for problems in two dimensions without limitations on the number of stable states, while the bistable model specified by Eqs. (3.20)–(3.22) will be considered for its application. The site-localizing function of Eq. (4.33) will be employed in the variational calculations with the following parametric functions: the stochastic separatrix, the width of the site-localizing function, and the metric tensor along the separatrix

$$x(\lambda),\ \sigma(\lambda),\ g_{ij}(\lambda) \equiv g_{ij}(x(\lambda)). \tag{5.1}$$

Specific parametric functions should be considered for each site-localizing function $G_\alpha(x)$, but, in order to avoid a too cumbersome notation, the explicit reference to the state α is left implicit in functions of Eq. (5.1).

For the sake of convenience, we choose the parameter λ such that the tangent vector of $x(\lambda)$ has a unitary norm in the Cartesian metric δ_{ij} (i.e., $\delta_{ij}a^i a^j = a_i a^i \equiv a^i a^i$ for all a),

$$t(\lambda) \equiv dx(\lambda)/d\lambda, \quad t^i(\lambda) t^i(\lambda) = 1. \tag{5.2}$$

Also the unitary vector $t_\perp(\lambda)$, locally orthogonal to the separatrix in the Cartesian metric

$$t_\perp^i(\lambda) t_\perp^i(\lambda) = 1, \quad t_\perp^i(\lambda) t^i(\lambda) = 0, \tag{5.3}$$

will be employed when necessary.

Once the metric tensor $g_{ij}(\lambda)$ is given, the invariant distance l, the covariant tangent vector $T^i(l)$, and the covariant orthogonal vector $N^i(l)$ along the stochastic separatrix are derived like in Sec. IV, with the set of complementary vectors

$$N_i(l) \equiv g_{ij}(l) N^j(l), \quad T_i(l) \equiv g_{ij}(l) T^j(l), \tag{5.4}$$

where $g_{ij}(l) \equiv [g_{ij}(\lambda)]_{\lambda=\lambda(l)}$. For a given l, the distance ρ from the separatrix will be measured along the straight line at $N^i(l)$,

$$x^i(\rho, l) = x^i(l) + \rho N^i(l). \tag{5.5}$$

Also the coordinate dependence of the metric tensor should be fully specified and we will employ the following form requiring only the knowledge of the metric tensor at the stochastic separatrix:

$$[g_{ij}(x)]_{x=x(\rho,l)} = g_{ij}(l). \tag{5.6}$$

This is equivalent to assuming a unitary metric tensor in the $y=(\rho, l)$ representation:

$$g_{ij}(y) \equiv \frac{\partial x^{i'}}{\partial y^i}\, g_{i'j'}(x)\, \frac{\partial x^{j'}}{\partial y^j} = \delta_{ij}, \tag{5.7}$$

the derivatives of the coordinates being calculated according to Eq. (5.5). Correspondingly, the parameter ρ can be identified with the distance along the orthogonal geodetics given as straight lines in the direction of $N^i(l)$. Equation (5.6), which is equivalent to a constant continuation of the metric tensor along the orthogonal directions, is the first approximation invoked in our variational procedure. Of course, also the dependence of the metric tensor on the orthogonal distance should be explicitly parametrized in a more complete treatment, but this would require an overly complicated variational procedure. The shortcoming of such an approximation is that the metric tensor (5.6) relies on the choice of coordinates x. If the same definition (5.5) is applied after a nonlinear change of variables $x \mapsto x'(x)$, a different set of orthogonal geodetics would be derived and therefore also a different metric tensor. On the other hand, the layer expansion is meaningful in the coordinate representations such that the change of site-localizing function is confined to a narrow layer about the stochastic separatrix. Correspondingly, smooth changes of variables do not modify substantially approximation (5.6) as long as small distances ρ only need to be considered. It should also be evident that any linear change of variables with constant coefficients preserves the structure of the metric tensor (5.6).

Also in the variational treatment, the layer expansion is conveniently done by employing coordinates $y \equiv (\rho, l)$. The decay rate (2.20) can be written in the following form after insertion of the site-localizing function (4.33):

$$w_\alpha = \frac{1}{4\pi\ \mathrm{erf}^2(\xi/2)} \int e^{[F_\alpha - V_s(y)]/k_B T - \rho^2/2\sigma(l)^2} f_i(y) \times D^{ij}(y) f_j(y)/\sigma(l)^2 dy \tag{5.8}$$

with functions $f_i(y)$ given by Eq. (4.24) and parameter ξ satisfying the constraint Eq. (4.32). As the second approximation, the diffusion matrix is calculated at the separatrix

$$D^{ij}(y) \equiv D^{ij}(\rho, l) \simeq D^{ij}(0, l). \tag{5.9}$$

This is justified by the hypothesis that only the integration in a narrow layer about the separatrix is required because of the steep change of site-localizing functions. Notice that the required diffusion elements can be derived from the diffusion tensor in the x representation according to the relations

$$\begin{aligned} D^{11}(0,l) &= N_i(l) [D^{ij}(x)]_{x=x(l)} N_j(l), \\ D^{12}(0,l) &= N_i(l) [D^{ij}(x)]_{x=x(l)} T_j(l), \\ D^{22}(0,l) &= T_i(l) [D^{ij}(x)]_{x=x(l)} T_j(l). \end{aligned} \tag{5.10}$$

The same hypothesis allows the parabolic expansion of the potential V_s with respect to the distance ρ,

$$\begin{aligned} V_s(y) &\simeq [V_s(y)]_{\rho=0} + \rho \left[\frac{\partial V_s(y)}{\partial \rho}\right]_{\rho=0} + \frac{\rho^2}{2} \left[\frac{\partial^2 V_s(y)}{\partial \rho^2}\right]_{\rho=0} \\ &= V_s(l) + \rho V_s^{(1)}(l) + \frac{\rho^2}{2} V_s^{(2)}(l); \end{aligned} \tag{5.11}$$

this is the third approximation invoked in our variational treatment. Notice that because of Eq. (5.6), the metric tensor depends only on the coordinate l and therefore the derivative of g in the scalar potential (3.8) is not required in the calculation of the expansion coefficients:

$$\begin{aligned} V_s^{(1)}(l) &= N^i(l) \left[\frac{\partial V(x)}{\partial x^i}\right]_{x=x(l)}, \\ V_s^{(2)}(l) &= N^i(l) N^j(l) \left[\frac{\partial^2 V(x)}{\partial x^i \partial x^j}\right]_{x=x(l)}. \end{aligned} \tag{5.12}$$

A major difference with respect to the layer expansion about the deterministic separatrix is the presence of the first-order term in the expansion (5.11). Correspondingly, the exponential function in Eq. (5.8) generates a Gaussian profile centered at a finite distance $\Delta\rho$ from the separatrix,

$$\Delta\rho(l) = -\eta(l)^2 V_s^{(1)}(l)/k_B T, \tag{5.13}$$

where $\eta(l)$ is given by Eq. (4.27). With all these ingredients, the analytical integration on the distance ρ can be performed

in Eq. (5.8), thus obtaining the following relation requiring only functions parametrically dependent on the coordinate l along the separatrix:

$$w_\alpha = \frac{1}{\sqrt{8\pi}\ \mathrm{erf}^2(\xi/2)} \int e^{[F_\alpha - V_s(l)]/k_B T + \Delta\rho^2/2\eta^2} \frac{\eta}{\sigma^2} \times \{h_i(l) D^{ij}(0,l) h_j(l) + D^{22}(0,l)[\eta\dot{\sigma}(l)/\sigma]^2\} dl, \tag{5.14}$$

with

$$h_1 = 1, \quad h_2(l) = -\Delta\rho(l)\dot{\sigma}(l)/\sigma(l). \tag{5.15}$$

Equation (5.14) is the starting point for the numerical implementation of the variational method.

First, however, one should eliminate the degeneracy of the decay rate with respect to the variational parameters, i.e., the possibility that different sets of parametric functions (5.1) determine the same decay rate w_α. Only without degeneracy can one find unambiguously the optimal variational parameters from the minimization of the decay rate. The identification of the degeneracies might be difficult because of the inclusion of the metric tensor g_{ij} among the variational parameters. A much simpler analysis can be done by using the tangent vector T^i and the orthogonal vector N^i. On the other hand, a direct relation exists between these vectors and the metric tensor. In fact, the metric tensor can be easily derived once these vectors are fixed and, in particular, the determinant of the metric tensor can be calculated as

$$1/g(l) = [T^i(l) t^i(l)]^2 [N^j(l) t^j_\perp(l)]^2. \tag{5.16}$$

Therefore, in the following analysis, reference is made to vectors T^i and N^i with the metric tensor calculated accordingly.

One type of degeneracy is associated with the change in magnitude of the tangent vector. Let us consider a set of parametric functions distinct from those of Eq. (5.1) because of a new metric tensor $\overline{g_{ij}}(\lambda)$ defined by a scaling of the tangent vector

$$\overline{T^i} = \mu T^i, \tag{5.17}$$

with $\mu = \mu(\lambda)$, without modifications of the orthogonal vector N^i and, according to Eq. (5.5), also of the distance ρ. Correspondingly, the same parameters $V_s^{(1)}$, $V_s^{(2)}$, η, and $\Delta\rho$ are recovered at a given point of the separatrix while, because of Eq. (4.2), the new distance l along the separatrix should obey the relation

$$d\overline{l} = dl/\mu. \tag{5.18}$$

One can easily show that the quantity in the curly brackets in Eq. (5.14) does not require modifications. On the contrary, the exponential function does change because the scalar potential (3.8) includes also a contribution of the metric tensor

$$\exp\{-\overline{V_s}(\lambda)/k_B T\} = \frac{e^{-V(x(\lambda))/k_B T}}{\sqrt{\overline{g(\lambda)}}} = \mu \frac{e^{-V(x(\lambda))/k_B T}}{\sqrt{g(\lambda)}} = \mu e^{-V_s(\lambda)/k_B T}, \tag{5.19}$$

where Eq. (5.16) has been employed for transforming the determinant of the metric tensor. Given all these correspondences, the same decay rate w_α is obtained by performing into Eq. (5.14) the change of the variable of Eq. (5.18), thus demonstrating the equivalence of the two sets of variational parameters. Such a degeneracy is conveniently removed by using tangent vectors with an unitary norm in the Cartesian metric

$$T^i = t^i, \tag{5.20}$$

which, because of Eq. (5.2), ensures the identity $l = \lambda$.

A second type of degeneracy can be associated with the scaling of the orthogonal vector

$$\overline{N^i} = \mu N^i, \tag{5.21}$$

with $\mu = \mu(\lambda)$, without modifications of the tangent vector T^i. According to Eq. (5.5), the distance from the separatrix scales as $\overline{\rho} = \rho/\mu$. Correspondingly, a new metric tensor $\overline{g_{ij}}(\lambda)$ is derived and one can calculate the decay rate with the same separatrix of Eq. (5.1); the width changed as $\overline{\sigma} = \sigma/\mu$ in order to preserve the shape of the site-localizing function. Like in the previous analysis, one can demonstrate that the decay rate does not change. Also this degeneracy should be eliminated, and this is conveniently done by fixing the magnitude of the complementary vector N_i,

$$N_i = t^i_\perp . \tag{5.22}$$

In this way the parameter ρ, as well as the associated width σ, assumes the meaning of distance from the separatrix in the Cartesian metric, as derived from Eq. (5.5) for small enough ρ,

$$\rho = t^i_\perp [x^i - x^i(l)]. \tag{5.23}$$

Correspondingly, a unitary determinant $g(l) = 1$ of the metric tensor is obtained from Eq. (5.16) since $t^j_\perp N^j = N_j N^j = 1$, and the scalar potential along the separatrix can be identified with the original mean-field potential

$$V_s(l) \equiv V_s(x(l)) = V(x(l)) \equiv V(l). \tag{5.24}$$

Having fixed the magnitudes of both vectors $T^i(l)$ and $N_i(l)$, one variable only is left for parametrizing the metric tensor. The most convenient one is the angle ψ between vectors $N^i(l)$ and $t^i_\perp(l)$. In conclusion, the parametric functions required in the variational calculation is reduced to the set

$$x(\lambda), \quad \sigma(\lambda), \quad \psi(\lambda), \tag{5.25}$$

to be used in the calculation of the decay rate according to Eq. (5.14) with $\lambda = l$ as the integration variable.

In order to test the previous variational procedure, the bistable problem specified by Eqs. (3.20)–(3.22) has been considered. In this case one can exploit the symmetry of the FP problem by imposing the conditions $x(-\lambda) = -x(\lambda)$, $\sigma(-\lambda) = \sigma(\lambda)$, and $\psi(-\lambda) = \psi(\lambda)$, where the saddle point has been taken as the origin for the parameter λ. Correspondingly, the integral (5.14) for the transition rate can be confined to the positive values of $l = \lambda$. Moreover, for computa-

tional purposes the stochastic separatrix $x(\lambda)$ is conveniently parametrized in terms of the orientation $\theta(\lambda)$ of its tangent vector $t^i(\lambda)$.

In order to deal with a finite-dimensional variational problem, the discretization of the integral (5.14) is performed with equally spaced mesh points λ_i. In this way the rate $w=w_1=w_2$ becomes a function of the variational parameters $\theta(\lambda_i)$, $\sigma(\lambda_i)$, and $\psi(\lambda_i)$ at the mesh points and standard algorithms [45] can be employed for the minimization of w with respect to these parameters. The contribution of each integration point is weighted by the Boltzmann factor $\exp\{-V(\lambda_i)/k_BT\}$. The points very far from the origin bring a negligible contribution to the rate given the magnitude of the potential function (see Fig. 1). Correspondingly, it is impossible to optimize the variational parameters in such locations because of the insensitivity of the rate. Therefore, the optimization procedure has to be confined to a finite portion of the separatrix. In our calculations we have discretized the separatrix up to $V(\lambda)/k_BT=12$, i.e., seven k_BT units over the saddle point. Thirty mesh points for $\lambda\geqslant 0$ have been used to achieve an accuracy better than 1% for the integral (5.14). In order to find the minimum of the rate within reasonable computation times, only five independent values (and equally spaced with respect to λ) for each parametric function $\theta(\lambda)$, $\sigma(\lambda)$, and $\psi(\lambda)$ were considered by using a linear interpolation scheme to get their intermediate values. This procedure is justified by the rather smooth dependence on λ of these parameters in the considered examples ($0.1\leqslant\sigma/x_0\leqslant 0.2$, ψ and θ varying in a range of 20°).

Finally, a discussion is in order about the choice of the parameter ξ of the trial function (4.33). We have employed the operative definition of ξ,

$$\xi=c_\xi \inf_\lambda[\rho_c(\lambda)/\sigma(\lambda)], \tag{5.26}$$

which agrees with the constraint (4.32) when the numerical coefficient c_ξ is in the range $0<c_\xi\leqslant 1$. For a given coefficient c_ξ, ξ depends on the variational parameters of Eq. (5.25). In particular, small values of ξ are recovered from strongly bent separatrices characterized by short crossing distances $\rho_c(\lambda)$. Then comparably large rates w are derived from Eq. (5.14) because of the error function term $\mathrm{erf}(\xi/2)^2$ in the denominator. In other words, the finite size of the layer, which is taken into account by the parameter ξ, gives rise in the minimization procedure to a penalty for highly bent profiles of the separatrix. Only in this case can the variational procedure be implemented. In fact, the minimization of the rate becomes ill defined numerically when the layer with an infinite width is considered by assuming $\xi=\infty$, because of the instabilities arising from deformations of the separatrix towards serpentine profiles and self-crossings. In most cases, the results of the variational procedure are weakly dependent on the choice of the numerical parameter c_ξ, which can be well fixed at the upper bound $c_\xi=1$. Some difficulties, however, arise in the low range of diffusion anisotropies $D_f\sim D_s$ because of the presence of secondary minima for the rate, which can be eliminated by using smaller values of c_ξ. For this reason we have employed the coefficient $c_\xi=0.1$ in all the calculations

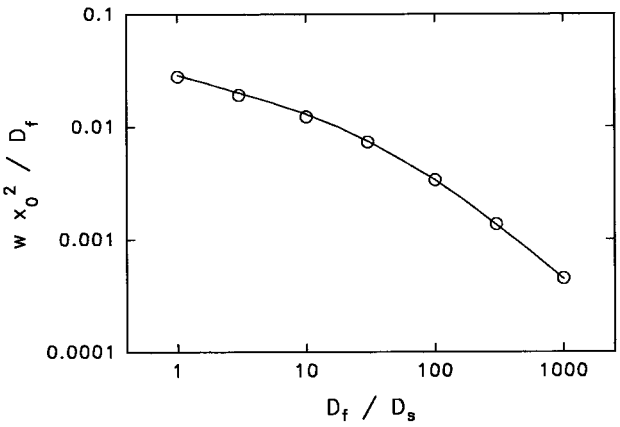

FIG. 8. Scaled transition rate as a function of the anisotropy ratio D_f/D_s with parameters of Eq. (3.22). Circles, exact numerical results; continuous line, variational layer expansion.

for the results reported in Figs. 8–10. Correspondingly, the factor $1/\mathrm{erf}^2(\xi/2)$ contributes at most by 2% to the optimized transitions rates.

A significant improvement with respect to the layer expansion about the deterministic separatrix is achieved by using the variational layer expansion. Satisfactory agreement is found in the comparison with the exact transition rates, as shown in Fig. 8, the largest deviation being about 5%. The evident improvement with respect to the method of Sec. IV (cf. Fig. 3) derives in the first instance from a correct identification of the separatrix. In Fig. 9 the comparison is done between the exact stochastic separatrix and the variational separatrix deriving from the optimization procedure for a set of diffusion anisotropies. In particular for large ratios D_f/D_s, the variational and the stochastic separatrices are nearly coincident. The progress with respect to the deterministic separatrix should appear evident by inspection of Fig. 5. Also the width of the site-localizing functions can be compared with the exact numerical solutions by examining their gradients. The derivative of G_α [Eq. (4.19)] with respect to the distance ρ at the separatrix is given as

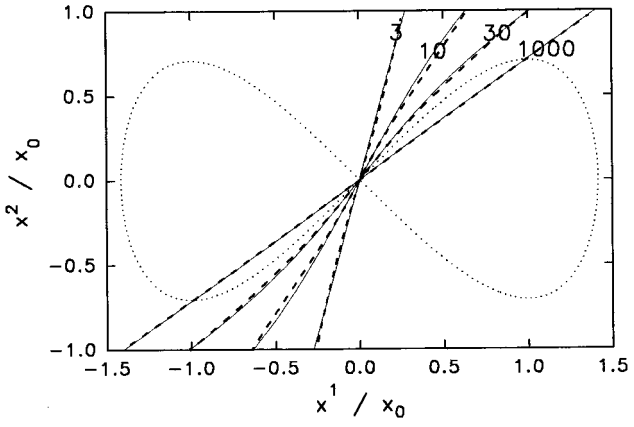

FIG. 9. Variational separatrices (dashed line) and stochastic separatrices (continuous line) for parameters of Eq. (3.22) and diffusion anisotropies D_f/D_s reported in the figure.

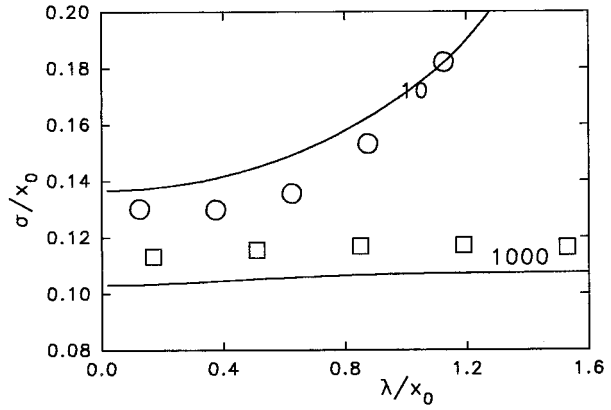

FIG. 10. Dependence of the width σ of the site-localizing function on the displacement λ along the separatrix, for parameters of Eq. (3.22) and two values of diffusion anisotropies, which are reported in the figure. Continuous lines, numerical exact results; circles and squares, variational layer expansion.

$$\left(\frac{\partial G_\alpha}{\partial \rho}\right)_{\rho=0} = \pm \frac{1}{2\sqrt{\pi}\sigma} . \tag{5.27}$$

According to Eq. (5.23), ρ can be identified with the Cartesian distance from the separatrix. Then the width σ can be derived from the the gradient of G_α at the separatrix, which can be easily computed from the exact site-localizing functions

$$1/\sigma = 2\sqrt{\pi}\left| t_\perp^i \frac{\partial G_\alpha}{\partial x^i}\right|_{\rho=0} . \tag{5.28}$$

In Fig. 10 both the exact numerical widths and those resulting from the variational procedure are reported in two cases as a functions of the position along the separatrix. Fair agreement is found with deviations of about 10%. On the contrary, no simple way exists for identifying the angle ψ from the exact site-localizing functions. On the other hand, the decay rate is weakly-dependent on this parameter, and substantially the same results reported in Figs. 8–10 would be recovered by constraining such an angle, say, $\psi=0$. One can understand it by considering the limit case of a straight separatrix with a fixed width σ. In this situation, the site-localizing function (4.19) does not depend on the angle ψ since Eq. (5.23) determines the distance from the separatrix independently of the direction of N^i. In actual cases, the separatrix is at least slightly bent, σ depends on the position, and the optimization procedure depends, even if weakly, on $\psi(\lambda)$.

It should be mentioned that the variational optimization is usually much faster than the exact numerical solution of the stochastic problem, but its efficiency strongly depends on the initial guess for the parameters of Eq. (5.25). On the other hand, the variational layer expansion supplies only approximate results and therefore it cannot replace the numerical solution of the FP problem if exact results are required.

As a by-product of layer expansion method, one can explain why the stochastic separatrix tends to be aligned with the direction d_f of fast diffusion motion for large diffusion anisotropies. Let us consider Eq. (5.14) under the simplifying condition of a constant width σ, such that the term in curly brackets can be replaced by

$$D^{11}(0,l) = t_\perp^i t_\perp^j [D^{ij}(x)]_{x=x(l)} \tag{5.29}$$

according to Eqs. (5.10) and (5.22). Its minimum is attained when $t_\perp^i = d_s^i$, i.e., $t^i = d_f^i$. It is precisely this term that dominates in the optimization of the decay rate when $D_f/D_s \to \infty$, so that the direction t^i of the separatrix is forced to be parallel to the fast diffusion direction.

VI. CONCLUSION

Our objective was the determination of the shape of the characteristic functions describing the kinetic processes (i.e., the site-localizing functions) in two-dimensional Fokker-Planck equations of Smoluchowski type. The starting point was provided by the boundary layer expansion developed by Matkowsky and Schuss [14,15]. This method has been implemented to the calculation of site-localizing functions by using the covariant form of the FP equation with the metric tensor chosen according to the diffusion matrix [16–19]. In this way we have derived a simple form of the site-localizing functions with an error-function shape about the deterministic separatrix imposed by the drift. Correspondingly, the transition rate was easily calculated by integration of a proper kernel along the separatrix. However, the comparison with the exact numerical results for a model bistable system evidenced large deviations on the transition rates when the diffusion matrix is highly anisotropic. The underlying reason is that the deterministic separatrix might be quite different from the stochastic separatrix of the exact numerical solutions. Therefore we have generalized the procedure by considering the separatrix as a parametric function to be optimized in the variational calculation of the transition rate. The variational method allowed also the inclusion of the metric tensor among the variational parameters. By taking into account the invariance properties of the transition rate, the variational problem is confined to the following parametric functions: the separatrix, the width of the site-localizing function, and the direction for the displacement from the separatrix. A finite-dimensional problem has been generated and solved by discretizing the integral for the transition rate. Satisfactory agreement is found for both the separatrix and the transition rate in comparison with the exact numerical solutions, thus demonstrating the capability of the variational procedure to reproduce the site-localizing functions.

ACKNOWLEDGMENTS

We thank the authors of Refs. [12, 24] for making available their work before publication. This work has been supported by the Italian Ministry for Universities and Scientific and Technological Research. G.J.M. acknowledges the support by the National Research Council (CNR), through its Centro Studi sugli Stati Molecolari, and the Committee for Information Science and Technology.

[1] H. A. Kramers, Physica **7**, 284 (1940).

[2] J. S. Langer, Ann. Phys. (N.Y.) **54**, 258 (1969).

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

[4] V. I. Melnikov, Phys. Rep. **209**, 1 (1991).

[5] A. Polimeno, P. L. Nordio, and G. Moro, Chem. Phys. Lett. **144**, 357 (1988); G. J. Moro and A. Polimeno, *ibid*. **189**, 133 (1992).

[6] R. Ferrando, R. Spadacini, and G. E. Tommei, Phys. Rev. E **48**, 2437 (1993).

[7] E. Pollack, J. Bader, B. J. Berne, and P. Talkner, Phys. Rev. Lett. **70**, 3299 (1993).

[8] A. M. Berezhkovskii, L. M. Berezhkovskii, and V. Yu. Zitserman, Chem. Phys. **130**, 55 (1989); A. M. Berezhkovskii and V. Yu. Zitserman, Physica A **166**, 585 (1990).

[9] G. J. Moro, J. Chem. Phys. **103**, 7514 (1995).

[10] E. Helfand, J. Chem. Phys. **54**, 4651 (1971); E. Helfand, Science **226**, 647 (1984).

[11] G. J. Moro, J. Phys. Chem. **100**, 16 419 (1996).

[12] A. M. Berezhkovskii, V. Yu. Zitserman, and A. Polimeno, J. Chem. Phys. **105**, 6342 (1996).

[13] G. J. Moro, Chem. Phys. **159**, 412 (1992).

[14] B. J. Matkowsky and Z. Schuss, SIAM J. Appl. Math. **33**, 365 (1977); **36**, 604 (1979); **40**, 242 (1981); **42**, 822 (1982).

[15] Z. Schuss, *Theory and Applications of Stochastic Differential Equations* (Wiley, New York, 1980).

[16] R. Graham, Z. Phys. B **26**, 387 (1977).

[17] H. Risken, *The Fokker-Planck Equation* (Springer, Berlin, 1989).

[18] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes* (North-Holland, Amsterdam, 1981).

[19] R. Bausch, R. Schmitz, and L. A. Turski, Phys. Rev. Lett. **73**, 2382 (1994).

[20] G. J. Moro, P. L. Nordio, and A. Polimeno, Chem. Phys. Lett. **182**, 575 (1991); A. Polimeno and P. L. Nordio, *ibid*. **192**, 509 (1992); Mol. Phys. **75**, 1203 (1992).

[21] D. Ryter, J. Stat. Phys. **49**, 751 (1987).

[22] M. M. Klozek, B. D. Matkowsky, and Z. Schuss, Ber. Bunsenges. Phys. Chem. **95**, 331 (1991).

[23] P. Talkner, Chem. Phys. **180**, 199 (1994).

[24] A. N. Drozdov and P. Talkner, J. Chem. Phys. **105**, 4117 (1996); Phys. Rev. E **54**, 6160 (1996).

[25] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer, Berlin, 1983).

[26] More general FP equations may include sink terms or may have absorbing boundary conditions. In such a case the norm conservation Eq. (2.3) would not be obeyed and the FP equation would not have a stationary solution.

[27] B. J. Berne, in *Physical Chemistry, An Advanced Treatise*, edited by H. Eyring, O. Henderson, and W. Jost (Academic, New York, 1971), Vol. 8B, p. 540.

[28] K. J. Laidler, *Chemical Kinetics* (Harper and Row, New York, 1976).

[29] Eigenfunctions reduced to their real form have been employed throughout. Notice that Eq. (2.10) has a trivial solution $\lambda_0=0$ and $\phi_0(x)=1$, in correspondence with the equilibrium distribution.

[30] R. G. Muncaster, Arch. Rat. Mech. Anal. **84**, 353 (1983).

[31] D. Ryter, Physica A **142**, 103 (1987).

[32] R. S. Larson and M. D. Kostin, J. Chem. Phys. **69**, 4821 (1978); **72**, 1392 (1980).

[33] G. Moro and P. L. Nordio, Mol. Phys. **57**, 947 (1986); G. J. Moro, J. Chem. Phys. **97**, 5749 (1992).

[34] Notice that Eq. (2.19) is used also for the definition of splitting probabilities [22,23,31]. In this case, however, the stochastic problem is modified by including suitable boundary conditions near the stable states, in order to impose the condition Eq. (2.15).

[35] One can easily verify that for $M=1$ Eq. (2.21) is equivalent to the Ritz theorem, which is the basis of many variational methods of quantum mechanics; see, for example, C. Cohen-Tannoudji, B. Diu, and F. Laloe, *Quantum Mechanics* (Wiley, New York, 1977), Vol. II. Equation (2.21) is a generalization of the Ritz theorem, which can be demonstrated by applying the same methods. First one has to derive the stationarity condition of $\mathrm{Tr}\{\mathcal{P}A\}$ with respect to the changes of subspace ϵ determined by $\mathcal{P}$. This is conveniently done by considering an orthonormal basis $f_1, f_2, \ldots, f_M$ for ϵ and the perturbations $f_i \to f_i + \delta f_i$ with δf_i orthogonal to ϵ. One derives that the linear variations of $\mathrm{Tr}\{\mathcal{P}A\}$ with respect to δf_i vanish only if ϵ is a closed subspace for A, that is, if ϵ is spanned by eigenfunctions of A, in which case $\mathrm{Tr}\{\mathcal{P}A\}$ is given by the sum of the corresponding eigenvalues. Therefore, the minimum of $\mathrm{Tr}\{\mathcal{P}A\}$ is attained for ϵ spanned by the first M eigenfunctions of A, and such a minimum is not degenerate if $\lambda_{M-1} < \lambda_M$.

[36] Such a condition is not essential for deriving the relation $\mathrm{Tr}\{\mathcal{P}\tilde{\Gamma}\} = \Sigma_\alpha w_\alpha$. By using the formalism of Ref. [9], which explicitly considers the superposition matrix, one can demonstrate it in all generality.

[37] There is a simple way to derive the transformation law (3.3) for the diffusion tensor by (i) calculating the time derivative of Eq. (2.7) according to the FP equation (2.5), (ii) integrating by parts with respect to the first derivative in operator Γ of Eq. (2.6), (iii) performing the change of variables $x \mapsto y$, and finally (iv) integrating by parts back to repristinate the second-order differential form for Γ (this procedure relies on the vanishing of the fluxes at the boundaries). An alternative, and more straightforward method utilizes the Piola identity [38] $\Sigma_i (\partial/\partial x^i) \det(\partial y^k/\partial x^{k'})(\partial x^i/\partial y^j) = 0$. By applying such an identity to the time derivative of probability density $p(y,t)$ Eq. (3.2), the structure of Eq. (2.6) is recovered for the evolution operator if the diffusion matrix is transformed according to Eq. (3.3).

[38] J. E. Marsden and T. J. R. Hughes, *Mathematical Foundations of Elasticity* (Prentice-Hall, Englewood Cliffs, NJ, 1983), p. 6.

[39] B. A. Dubrovin, A. T. Fomenko, and S. P. Novikov, *Modern Geometry—Methods and Applications* (Springer, Berlin, 1985), Vol. 1.

[40] It should be emphasized that for a fixed barrier height, the validity of Eq. (2.13) with large diffusion anisotropies strongly depends on the angle θ and, in particular, it fails when θ is too low [8,13,23]. In our case $\lambda_2/\lambda_1 < 5$ for $|\theta| < 15°$ when $\Delta = 5$ and $\alpha = 90°$.

[41] N. G. de Bruijn, *Asymptotic Methods in Analysis* (Dover, New York, 1981).

[42] The case of a deterministic separatrix with several saddle points has been examined by B. J. Matkovsky, Z. Schuss, and C. Tiers, SIAM J. Appl. Math. **43**, 673 (1983).

[43] The definition of stochastic separatrix of Refs. [21–23] utilizes the stationary solution of the FP equation with fixed boundary conditions at the minima in order to enforce the constraints Eq. (2.15). *Ad hoc* boundary conditions are not required by using

site-localizing functions calculated as linear combinations of FP kinetic eigenfunctions with the same constraints Eq. (2.15).

[44] In bistable problems, the same stochastic separatrix is derived from Eq. (4.30) independently of the site-localizing function since $G_1(x)+G_2(x)=1$ according to Eq. (2.17). This cannot be extended to problems with several stable states where a distinct stochastic separatrix $x_\alpha(\lambda)$ is obtained from Eq. (4.30) applied to each site-localizing function $G_\alpha(x)$. The phase space enclosed by $x_\alpha(\lambda)$ can be assigned to the (stochastic) domain of attraction Ω_α, but taking into account that these domains may superimpose and in general they do not cover the entire phase space [9].

[45] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes* (Cambridge University Press, Cambridge, 1989).