The Annals of Probability
2017, Vol. 45, No. 5, 3202–3222
DOI: 10.1214/16-AOP1134

ON THE BEHAVIOR OF DIFFUSION PROCESSES WITH TRAPS

BY M. FREIDLIN[*,1], L. KORALOV[*,2] AND A. WENTZELL[†]

University of Maryland[*] and Tulane University[†]

We consider processes that coincide with a given diffusion process outside a finite collection of domains. In each of the domains, there is, additionally, a large drift directed towards the interior of the domain. We describe the limiting behavior of the processes as the magnitude of the drift tends to infinity, and thus the domains become trapping, with the time to exit the domains being exponentially large. In particular, in exponential time scales, metastable distributions between the trapping regions are considered.

1. Introduction. Let v be an infinitely differential vector field on the d-dimensional torus $\mathbb{T}^d$. (General manifolds, whether compact or not, can also be considered, but we will stick with the torus for the sake of simplicity of later notation.) Consider the process $\tilde{X}_t^{x,\varepsilon}$ defined via

$$d\tilde{X}_t^{x,\varepsilon} = v(\tilde{X}_t^{x,\varepsilon})\,dt + \sqrt{\varepsilon}\,d\tilde{W}_t, \qquad \tilde{X}_0^{x,\varepsilon} = x, \tag{1}$$

where ε is a small parameter and $\tilde{W}$ is a d-dimensional Wiener process. We are interested in the large-time behavior of this process when $\varepsilon \downarrow 0$.

It will be more convenient to "speed up" the time by the factor ε, thus considering the process $X_t^{x,\varepsilon}$ defined via

$$dX_t^{x,\varepsilon} = \frac{1}{\varepsilon} v(X_t^{x,\varepsilon})\,dt + dW_t, \qquad X_0^{x,\varepsilon} = x,$$

where ε is a small parameter and W is a d-dimensional Wiener process. Assume that the collection of asymptotically stable limit sets of the unperturbed flow consists of n equilibrium points $O_1, \dots, O_n$. Let $D_1, \dots, D_n$ denote the sets that are attracted to $O_1, \dots, O_n$, respectively. Intuitively, when ε is small and t is large, $X_t^{x,\varepsilon}$ is located near one of these points with overwhelming probability. The transitions between small neighborhoods of the equilibriums are governed by the matrix

$$V_{ij} = \frac{1}{2} \inf\left\{ \int_0^T |\dot{\varphi}_s - v(\varphi_s)|^2\,ds, \varphi_0 = O_i, \varphi_T = O_j \right\}, \qquad i, j \in \{1, \dots, n\}, \tag{2}$$

Received October 2015; revised April 2016.
[1]Supported by NSF Grant DMS-14-11866.
[2]Supported by NSF Grant DMS-13-09084.
MSC2010 subject classifications. 60F10, 35J25, 47D07, 60J60.
Key words and phrases. Metastable distributions, large deviations, exit problem, nonstandard boundary problem.

where the infimum is taken over all $T \geq 0$ and all absolutely continuous functions $\varphi : [0, T] \to \mathbb{T}^d$. Loosely speaking, for each $i = 1, \dots, n$, $x \in D_i$, and $\lambda > 0$ (except a finite subset of λ's, as discussed below), there is an index $k = k(x, \lambda) \in \{1, \dots, n\}$ such that

$$\text{(3)} \qquad \operatorname{dist}\big(X^{x,\varepsilon}_{\exp(\lambda/\varepsilon)}, O_k\big) \to 0 \qquad \text{in probability as } \varepsilon \downarrow 0.$$

The equilibrium O_k is called the metastable state for the process $X_t^{x,\varepsilon}$ corresponding to the initial point x and time scale $\exp(\lambda/\varepsilon)$. The state O_k can be determined by comparing λ with certain linear expressions involving the numbers V_{ij} (see Chapter 6 of [4]). Typically, there is a finite set Λ of values of λ where transitions from one metastable state to another happen, that is, the notion of a metastable state is defined for $\lambda \in (0, \infty) \setminus \Lambda$.

On the other hand, for certain geometries of the unperturbed flow, it may happen that $V_{ij_1} = V_{ij_2}$ for some $j_1 \neq j_2$ or, more generally, the sums of two distinct collections of V_{ij}'s may be equal. The analysis of the asymptotics of $X_t^{x,\varepsilon}$ is then more intricate. The notion of rough symmetry, of which the simplest examples are due to geometric symmetries of the flow, was introduced and to some extent analyzed in [3]. In the presence of rough symmetry, metastable states may need to be replaced by metastable distributions between the asymptotically stable attractors.

In the current paper, we analyze an interesting situation where V_{ij} do not depend on j. Consequently, $X_t^{x,\varepsilon}$ is distributed between the neighborhoods of several equilibriums at exponential time scales. Now this phenomenon is due not to geometric symmetries but to the vanishing of the vector field in certain regions of the state space. Namely, let $D_1, \dots, D_n$ be open connected domains with infinitely differentiable boundaries ∂D_k, $k = 1, \dots, n$, on the d-dimensional torus $\mathbb{T}^d$. The closures $\overline{D}_k$ are assumed to be disjoint. Let v be a vector field on $\mathbb{T}^d$ that is equal to zero on $\mathbb{T}^d \setminus \bigcup_{k=1}^n \overline{D}_k$. It is assumed to be infinitely differentiable in the sense that there is an infinitely differentiable field on $\mathbb{T}^d$ that agrees with v on $\bigcup_{k=1}^n \overline{D}_k$. Assume, for the moment, that all the points of $\overline{D}_k$ are attracted to an equilibrium $O_k \in D_k$, $k = 1, \dots, n$ (see Figure 1).

The quantities V_{ij} are easily seen not to depend on j since

$$\begin{aligned} &\inf\left\{\int_0^T |\dot{\varphi}_s - v(\varphi_s)|^2\, ds, \varphi_0 \in \partial D_i, \varphi_T \in \partial D_j\right\} \\ &\quad = \inf\left\{\int_0^T |\dot{\varphi}_s|^2\, ds, \varphi_0 \in \partial D_i, \varphi_T \in \partial D_j\right\} \\ &\quad = 0, \end{aligned}$$

where the infimum is taken over all $T > 0$ and $\varphi \in C^1([0, T], \mathbb{T}^d)$. There are two issues involved in understanding the transitions of the process $X_t^{x,\varepsilon}$ between the neighborhoods of different equilibriums. The first issue is to describe how the process that starts near O_k exits the domain D_k. Fortunately, this questions has been

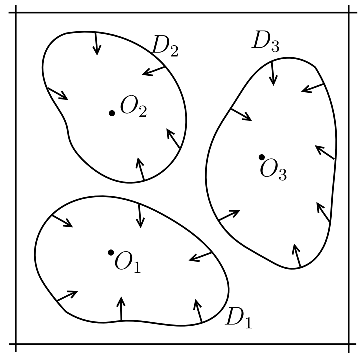

FIG. 1. *Torus with multiple trapping regions.*

well studied. One needs to look at the quasi-potential:

$$V_k(x) = \frac{1}{2}\inf\left\{\int_0^T |\dot{\varphi}_s - v(\varphi_s)|^2\,ds, \varphi_0 = O_k, \varphi_T = x\right\}, \qquad x \in \partial D_k, \tag{4}$$

where the infimum is taken over all $T \geq 0$ and all $\varphi \in C^1([0, T], \overline{D}_k)$. If the minimum of $V_k(x), x \in \partial D_k$, is achieved at a single point x_k, then the process starting near O_k exits D_k in a small neighborhood of x_k with probability that tends to one as $\varepsilon \downarrow 0$ (see Chapter 4 of [4]). The time to exit is of order $\exp(V_k(x_k)/\varepsilon)$. Even if the minimum is not achieved at a single point, there are cases when the exit from the domain is well understood (e.g., in the simplest example when v is spherically symmetric and D_k is a ball centered at O_k, also see [1] and references there). We will simply assume that for each compact $K \subset D_k$, the exit time (appropriately rescaled) and the exit location have limiting distributions that do not depend on the starting point within K, as in the case of a single minimum for the quasi-potential and in the symmetric case mentioned above.

The second issue concerns the transitions between the domains D_k, that is, the behavior of the process $X_t^{x,\varepsilon}$ on $\mathbb{T}^d \setminus \bigcup_{k=1}^n D_k$. In order to get a meaningful limiting object, we introduce a new process $Y_t^{x,\varepsilon}$ by running the clock only when $X_t^{x,\varepsilon} \in \mathbb{T}^d \setminus \bigcup_{k=1}^n D_k$. While this process obviously coincides with a Wiener process away from boundary $\bigcup_{k=1}^n \partial D_k$, the regions D_k still play an important role by trapping the process when it reaches the boundary and then re-distributing it on ∂D_k. The description of the limiting behavior of $Y_t^{x,\varepsilon}$ is the main result of this paper.

In Section 2, we describe the limiting process. It belongs to a peculiar class of processes that, it seems, have not been discussed in the literature previously. In Section 3, we prove the convergence of $Y_t^{x,\varepsilon}$ to the limit. In Section 4, we describe the asymptotics of the original process $X_t^{x,\varepsilon}$ at exponential time scales and make several additional remarks.

One of the motivations for this paper is the study of cluster formation for particles moving in a short-range potential (see [5], [7] and references therein). The unperturbed motion of particles is governed by the equation

$$\dot{X}_t^x = v(X_t^x) := -\nabla A(X_t^x), \qquad X_0^x = x \in \mathbb{T}^d,$$

where $A(x) = \sum_{k=1}^n A_k(x)$. Here, A_k, $k = 1, \dots, n$, are continuous, supported in D_k, and each has a unique local maximum O_k in D_k. By saying that the potential is short-range, we mean that the supports of the functions A_k do not overlap (see Section 4 for a more general discussion). Since the vector field is potential in this example, we simply have $V_k = 2A_k(O_k)$ for the quasi-potentials V_k defined in (4). Therefore, the perturbed process $\tilde{X}_t^{x,\varepsilon}$, defined in (1), exits D_k in time that is logarithmically equivalent to $\exp(2A_k(O_k)/\varepsilon)$ ([4], Chapter 4). The analysis of the limiting behavior of $Y_t^{x,\varepsilon}$ applies in this case. As we discussed, this allows one to describe the distribution of $\tilde{X}_{T(\varepsilon)}^{x,\varepsilon}$ (and its time-changed version $X_t^{x,\varepsilon}$) for $T(\varepsilon) \sim \exp(\lambda/\varepsilon)$, $\lambda > 0$.

2. Description of the limiting process. In this section, we define the family of processes X_t^x, which later will be proved to be the limiting processes for $Y_t^{x,\varepsilon}$ as $\varepsilon \downarrow 0$. Let $D_1, \dots, D_n \subset \mathbb{T}^d$ be open connected domains with infinitely differentiable boundaries ∂D_k, $k = 1, \dots, n$. The closures $\overline{D}_k$ are assumed to be disjoint. Let $U = \mathbb{T}^d \setminus \bigcup_{k=1}^n \overline{D}_k$. The closure of this domain will be denoted by $\overline{U}$. Let U' be the metric space obtained from $\overline{U}$ by identifying all points of ∂D_k, turning every ∂D_k, $k = 1, \dots, n$, into one point d_k.

The family of processes X_t^x, $x \in U'$, will be defined in terms of its generator. Since we expect X_t^x to coincide with a Wiener process inside U, the generator coincides with $\frac{1}{2}\Delta$ on a certain class of functions. The domain, however, should be restricted by certain boundary conditions to account for nontrivial behavior of X_t^x on the boundary of U. We will use the Hille–Yosida theorem stated here in the form that is convenient for considering closures of linear operators (see [8]).

THEOREM 2.1. *Let K be a compact space, $C(K)$ be the space of continuous functions on it. The space $C(K)$ is endowed with the supremum norm. Suppose that a linear operator A on $C(K)$ has the following properties*:

(a) *The domain $\mathcal{D}(A)$ is dense in $C(K)$.*

(b) *The constant function $\mathbf{1}$ belongs to $\mathcal{D}(A)$ and $A\mathbf{1} = 0$.*

(c) *The maximum principle*: *If S is the set of points where a function $f \in \mathcal{D}(A)$ reaches its maximum, then $Af(x) \le 0$ for at least one point $x \in S$.*

(d) *For a dense set $\Psi \subseteq C(K)$, for every $\psi \in \Psi$, and every $\lambda > 0$, there exists a solution $f \in \mathcal{D}(A)$ of the equation $\lambda f - Af = \psi$.*

Then the operator A is closable and its closure $\overline{A}$ is the infinitesimal generator of a unique semi-group of positivity-preserving operators T_t, $t \ge 0$, on $C(K)$ with $T_t\mathbf{1} = \mathbf{1}$, $\|T_t\| \le 1$.

Suppose that we are given positive finite measures $\nu_1, \dots, \nu_n$ concentrated on $\partial D_1, \dots, \partial D_n$, respectively. The Hille–Yosida theorem will be applied to the space $K = U'$. Let us define the linear operator A in $C(U')$. First, we define its domain. It consists of all functions $f \in C(U')$ that satisfy the following conditions:

(1) f is twice continuously differentiable in U.

(2) The limits of all the first- and second-order derivatives of f exist at all the points of the boundary $\partial U = \bigcup_{k=1}^n \partial D_k$.

(3) There are constants $g_1, \dots, g_n$ such that

$$\lim_{y \in U, \operatorname{dist}(y, \partial D_k) \downarrow 0} \Delta f(y) = g_k, \qquad k = 1, \dots, n.$$

(4) For each $k = 1, \dots, n$,

$$\int_{\partial D_k} \langle \nabla f(x), n(x) \rangle \nu_k(dx) = 0, \tag{5}$$

where $n(x)$ is the unit exterior normal at $x \in \partial D_k$ (with respect to U).

For $f \in \mathcal{D}(A)$ and $x \in U'$, we define

$$Af = \begin{cases} \dfrac{1}{2} \Delta f(x), & \text{if } x \in U, \\ \dfrac{1}{2} g_k, & \text{if } x = d_k, k = 1, \dots, n. \end{cases}$$

Let us check that the conditions of the Hille–Yosida theorem are satisfied.

(a) Consider the set G of functions g that are infinitely differentiable and have the following property: for each $k = 1, \dots, n$ there is a set V_k open in U' such that $\partial D_k \subset V_k$ and g is constant on V_k. It is clear that $G \subset \mathcal{D}(A)$ and G is dense in $C(U')$.

(b) Clearly, $\mathbf{1} \in \mathcal{D}(A)$ and $A\mathbf{1} = 0$.

(c) If f has a maximum at $x \in U$, it is clear that $\Delta f(x) \le 0$. Now suppose that f has a maximum at d_k. We can view f as an element of $C^2(\overline{U})$ that is constant on each component of the boundary, in particular on ∂D_k. Note that $\langle \nabla f(x), n(x) \rangle$ is identically zero on ∂D_k, since otherwise it would be negative at some points due to (5), which would contradict the fact that f reaches its maximum on ∂D_k. Then the second derivative of f in the direction of n is nonpositive at all points $x \in \partial D_k$. Since f is constant on ∂D_k, its second derivative in any direction tangential to the boundary is equal to zero. Therefore, $\Delta f(x) \le 0$ for $x \in \partial D_k$, that is, $Af(d_k) \le 0$, as required.

(d) Let Ψ be the set of functions $\psi \in C(U')$ that have limits of all the first-order derivatives as $y \to x$, $y \in U$, at all points $x \in \partial U$. It is clear that Ψ is dense in $C(U')$. Let $\widetilde{f} \in C^2(\overline{U})$ be the solution of the equation $\lambda \widetilde{f} - \frac{1}{2} \Delta \widetilde{f} = \psi$ in U, $\widetilde{f} = 0$ on ∂U. Let $h_k \in C^2(\overline{U})$ be the solution of the equation

$$\lambda h_k(x) - \frac{1}{2} \Delta h_k(x) = 0, \qquad x \in U,$$

$$h_k(x) = 1, \qquad x \in \partial D_k; \qquad h_k(x) = 0, \qquad x \in \partial U \setminus \partial D_k.$$

Let us look for the solution $f \in \mathcal{D}(A)$ of $\lambda f - Af = \psi$ in the form $f = \widetilde{f} + \sum_{k=1}^n c_k h_k$. We get n linear equations for $c_1, \dots, c_n$. The solution is unique because of the maximum principle. Therefore, the determinant of the system is nonzero, and the solution exists for all the right-hand sides.

Let $\overline{A}$ be the closure of A. Let T_t, $t \geq 0$, be the corresponding semi-group on $C(U')$, the existence of which is guaranteed by the Hille–Yosida theorem. By the Riesz–Markov–Kakutani representation theorem, for $x \in U'$ there is a measure $P(t, x, dy)$ on $(U', \mathcal{B}(U'))$ such that

$$(T_t f)(x) = \int_{U'} f(y) P(t, x, dy), \qquad f \in C(U').$$

It is a probability measure since $T_t \mathbf{1} = \mathbf{1}$. Moreover, it can be easily verified that $P(t, x, B)$ is a Markov transition function. Let X_t^x, $x \in U'$, be the corresponding Markov family. In order to show that a modification with continuous trajectories exists, it is enough to check that $\lim_{t \downarrow 0} P(t, x, B)/t = 0$ for each closed set B that does not contain x (Theorem I.5 of [6], see also [2]). Let $f \in \mathcal{D}(A)$ be a nonnegative function that is equal to one on B and whose support does not contain x. Then

$$\lim_{t \downarrow 0} \frac{P(t, x, B)}{t} \leq \lim_{t \downarrow 0} \frac{(T_t f)(x) - f(x)}{t} = Af(x) = 0,$$

as required. Thus, X_t^x can be assumed to have continuous trajectories.

3. Convergence of the trace of the process. Let $\pi : \mathbb{T}^d \to U'$ be the mapping defined by $\pi(x) = x$ for $x \in U$ and $\pi(x) = d_k$ for $x \in \overline{D}_k$. In this section, we prove the convergence of $Y_t^{x,\varepsilon}$, obtained from $X_t^{x,\varepsilon}$ by running the clock only when the process is in $\overline{U}$, to the limiting process $X_t^{\pi(x)}$.

Let v be a vector field that is smooth in $\bigcup_{k=1}^n \overline{D}_k$ (i.e., it admits a smooth continuation from $\bigcup_{k=1}^n \overline{D}_k$ to the whole space) and is equal to zero outside $\bigcup_{k=1}^n \overline{D}_k$.

Let

$$a(x) = \langle v(x), n(x) \rangle, \qquad x \in \partial U,$$

where n is the unit exterior normal to the boundary (with respect to U). We will assume that $a(x) > 0$ for all $x \in \partial U$. Recall that the process $X_t^{x,\varepsilon}$ is defined via

$$dX_t^{x,\varepsilon} = \frac{1}{\varepsilon} v(X_t^{x,\varepsilon})\, dt + dW_t, \qquad X_0^{x,\varepsilon} = x.$$

For $B \subset \mathbb{T}^d$, let $\tau^{x,\varepsilon}(B) = \inf\{t \geq 0 : X_t^{x,\varepsilon} \in B\}$. Let $\mu_k^{x,\varepsilon}$ be the measure on ∂D_k induced by $X_{\tau^{x,\varepsilon}(\partial D_k)}^{x,\varepsilon}$. We will assume that there are measures μ_k, $k = 1, \dots, n$, such that for each compact set $K \subset D_k$ and each continuous function φ on ∂D_k we have

$$\lim_{\varepsilon \downarrow 0} \int_{\partial D_k} \varphi \, d\mu_k^{x,\varepsilon} = \int_{\partial D_k} \varphi \, d\mu_k \tag{6}$$

uniformly in $x \in K$. Define the measures ν_k via

$$\nu_k(dx) = \big(2a(x)\big)^{-1}\mu_k(dx), \qquad x \in \partial D_k. \tag{7}$$

Let X_t^x be the Markov family of continuous U'-valued processes defined above, corresponding to the measures ν_k, $k = 1, \dots, n$. As discussed in the Introduction, if the minimum of the quasi-potential $V_k(x)$ defined in (4) is achieved at a single point x_k, then (6) is satisfied with μ_k being the delta-measure at x_k. Then ν_k is a constant multiple of the delta-measure, and the integral condition (5) becomes simply $\langle \nabla f(x_k), n(x_k)\rangle = 0$.

Define

$$s(t) = \inf\big(s : \lambda\big(u : u \le s, X_u^{x,\varepsilon} \in \overline{U}\big) > t\big),$$

where λ is the Lebesgue measure on the real line, and let

$$Y_t^{x,\varepsilon} = X_{s(t)}^{x,\varepsilon}. \tag{8}$$

Thus, $Y_t^{x,\varepsilon}$ is a right-continuous process with values in $\overline{U}$, which also can be viewed as a continuous U'-valued process. It can be obtained from $X_t^{x,\varepsilon}$ by running the clock only when $X_t^{x,\varepsilon}$ is in $\overline{U}$. The main result of this section is the following.

THEOREM 3.1. *For each $x \in \mathbb{T}^d$, the measures on $C([0,\infty), U')$ induced by the processes $Y_t^{x,\varepsilon}$ converge weakly, as $\varepsilon \downarrow 0$, to the measure induced by $X_t^{\pi(x)}$.*

The key ingredient in the proof of this theorem is the following proposition.

PROPOSITION 3.2. *Suppose that $f \in \mathcal{D}(A)$. Then*

$$\lim_{\varepsilon\downarrow 0} \mathrm{E}\bigg(f\big(Y_t^{x,\varepsilon}\big) - f(x) - \frac{1}{2}\int_0^t \Delta f\big(Y_u^{x,\varepsilon}\big)\,du\bigg) = 0 \tag{9}$$

for each $t \ge 0$, uniformly in $x \in \overline{U}$.

PROOF. For the sake of notational simplicity, we will assume that there is just one domain where the vector field v is nonzero. This does not lead to any loss of generality as the proof in the case of multiple domains is similar. We will denote the domain by D and will drop the subscript k from the notation everywhere. For example, (5) now takes the form

$$\int_{\partial D} \langle \nabla f(x), n(x)\rangle\, d\nu(x) = 0. \tag{10}$$

Let $S_r = \{x \in \overline{U} : \mathrm{dist}(x, \partial D) = r\}$ for $r \ge 0$, $S_r = \{x \in D : \mathrm{dist}(x, \partial D) = -r\}$ for $r < 0$. These are smooth surfaces if r is sufficiently small. Let $\Gamma_r = \{x \in \overline{U} : \mathrm{dist}(x, \partial D) \le r\}$ for $r \ge 0$.

Let $\sigma_0^{x,\varepsilon}=0$, $\tau_1^{x,\varepsilon}=\tau^{x,\varepsilon}(\partial D)$, $\sigma_n^{x,\varepsilon}=\inf(t\ge\tau_n^{x,\varepsilon}:X_t^{x,\varepsilon}\in S_{\sqrt{\varepsilon}})$, $n\ge 1$, while $\tau_n^{x,\varepsilon}=\inf(t\ge\sigma_{n-1}^{x,\varepsilon}:X_t^{x,\varepsilon}\in\partial D)$, $n\ge 2$. Then

$$\begin{aligned}
&\mathrm{E}\biggl(f(Y_t^{x,\varepsilon})-f(x)-\frac{1}{2}\int_0^t\Delta f(Y_u^{x,\varepsilon})\,du\biggr)\\
&\quad=\mathrm{E}\sum_{n=1}^{\infty}\biggl(f(X^{x,\varepsilon}_{\tau_n^{x,\varepsilon}\wedge s(t)})-f(X^{x,\varepsilon}_{\sigma_{n-1}^{x,\varepsilon}\wedge s(t)})-\frac{1}{2}\int_{\sigma_{n-1}^{x,\varepsilon}\wedge s(t)}^{\tau_n^{x,\varepsilon}\wedge s(t)}\Delta f(X_u^{x,\varepsilon})\,du\biggr)\qquad(11)\\
&\qquad+\mathrm{E}\sum_{n=1}^{\infty}\biggl(f(X^{x,\varepsilon}_{\sigma_n^{x,\varepsilon}\wedge s(t)})-f(X^{x,\varepsilon}_{\tau_n^{x,\varepsilon}\wedge s(t)})-\frac{1}{2}\int_{\tau_n^{x,\varepsilon}\wedge s(t)}^{\sigma_n^{x,\varepsilon}\wedge s(t)}\Delta f(X_u^{x,\varepsilon})\,du\biggr),
\end{aligned}$$

where we put $\Delta f\equiv 0$ on D. The first expectation on the right-hand side is equal to zero since $X_t^{x,\varepsilon}$ is a Wiener process on $\overline{U}$. Our goal is to show that the second expectation tends to zero. We separate the proof into several steps, most of which concern the stopping times $\sigma_n^{x,\varepsilon}$ and the behavior of the process near ∂D.

(1) In order to deal with the second expectation on the right-hand side of (11), we need to control the number of terms in the sum.

LEMMA 3.3. *There is $c=c(t)>0$ such that*

$$\mathrm{P}(\sigma_n^{x,\varepsilon}\le s(t))\le\exp(-cn\sqrt{\varepsilon}),\qquad x\in\overline{U},n\ge 2.\qquad(12)$$

PROOF. Since ∂D is smooth, there is $r>0$ such that the ball of radius r tangent to ∂D at x lies entirely in $\overline{U}$. Let η^ε be the time it takes a Wiener process starting inside a ball of radius r at a distance $\sqrt{\varepsilon}$ from the boundary to reach the boundary. It is easy to see that there is $c=c(t)>0$ such that

$$\mathrm{P}(\eta^\varepsilon\le t)\le\exp(-c\sqrt{\varepsilon}).$$

Therefore, if η_k^ε, $k\ge 1$, is a sequence of such independent random variables, then

$$\mathrm{P}(\eta_1^\varepsilon+\cdots+\eta_n^\varepsilon\le t)\le\exp(-cn\sqrt{\varepsilon}),\qquad n\ge 1.\qquad(13)$$

By our construction, $\mathrm{P}(\tau_n^{x,\varepsilon}-\sigma_{n-1}^{x,\varepsilon}>z|X^{x,\varepsilon}_{\sigma_{n-1}^{x,\varepsilon}})\ge\mathrm{P}(\eta^\varepsilon>z)$ for each $n\ge 2$ and $z>0$. Therefore, estimate (12), with $\tau_{n+1}^{x,\varepsilon}$ instead of $\sigma_n^{x,\varepsilon}$, follows from (13) and the strong Markov property. Thus, the original formula (12) also holds, with a different constant c. □

(2) Next, we consider auxiliary processes that will appear later, when we analyze the behavior of $X_t^{x,\varepsilon}$ in the vicinity of ∂D. First, consider a one-dimensional process Z_t^z that satisfies

$$dZ_t^z=dB_t-\mathrm{v}\chi_{(-\infty,0)}(Z_t^z)\,dt,\qquad Z_0^z=z,$$

where $\mathrm{v} > 0$ and B_t is a one-dimensional Wiener process. Also consider its perturbation defined via

$$d\widetilde{Z}_t^z = dB_t - \mathrm{v}\chi_{(-\infty,0)}(\widetilde{Z}_t^z)\,dt + A_t^z\,dt, \qquad \widetilde{Z}_0^z = z, \tag{14}$$

where A_t^z is a continuous adapted process satisfying $|A_t^z| \le \mathrm{v}/2$. It is easy to see that for each $\eta > 0$ there is a $z_0 < 0$ such that

$$\mathrm{P}\Big(\sup_{t\ge 0} \widetilde{Z}_t^{z_0} \ge 0\Big) \le \eta, \tag{15}$$

while for each $z_0 < 0$ there is a t_0 such that for $z \in [z_0, 1]$,

$$\mathrm{P}\big(\widetilde{Z}_t^z \text{ reaches } \{z_0\} \cup \{1\} \text{ before time } t_0\big) \ge 1 - \eta. \tag{16}$$

A direct calculation shows that

$$\mathrm{P}\Big(\sup_{t\ge 0} Z_t^0 \ge 1\Big) = (1 + 2\mathrm{v})^{-1}.$$

Fix such z_0 that (15) holds and

$$\mathrm{P}\big(Z_t^0 \text{ reaches 1 before reaching } z_0\big) \ge (1 + 2\mathrm{v})^{-1} - \eta. \tag{17}$$

By the Girsanov formula, the use of which is justified by (16) with $\widetilde{Z}_t^z$ replaced by Z_t^0, there is $\varkappa > 0$ such that

$$\begin{aligned} &\mathrm{P}\big(\widetilde{Z}_t^0 \text{ reaches 1 before reaching } z_0\big) \\ &\qquad \in \big[(1 + 2\mathrm{v})^{-1} - 3\eta, (1 + 2\mathrm{v})^{-1} + 3\eta\big], \end{aligned} \tag{18}$$

provided that $|A_t^0| \le \varkappa$. Below we will encounter a related ε-dependent process. Namely, suppose that $(\widetilde{Z}_t^{z,\varepsilon}, \widehat{Z}_t^{z,\varepsilon}) \in \mathbb{R} \times \mathbb{R}^{d-1}$ satisfy

$$d\widetilde{Z}_t^{z,\varepsilon} = d\widetilde{B}_t - \frac{\mathrm{v}}{\varepsilon}\chi_{(-\infty,0)}(\widetilde{Z}_t^{z,\varepsilon})\,dt + \frac{\widetilde{A}_t^z}{\varepsilon}\,dt, \qquad \widetilde{Z}_0^{z,\varepsilon} = \widetilde{z}, \tag{19}$$

$$d\widehat{Z}_t^{z,\varepsilon} = \sigma(\widetilde{Z}_t^{z,\varepsilon}, \widehat{Z}_t^{z,\varepsilon})\,d\widehat{B}_t + \frac{\widehat{A}_t^z}{\varepsilon}\,dt, \qquad \widehat{Z}_0^{z,\varepsilon} = \widehat{z}, \tag{20}$$

where $\widetilde{B}_t$ is a one-dimensional Wiener process, $\widehat{B}_t$ is a d-dimensional Wiener process, possibly correlated with $\widetilde{B}_t$, σ is a $(d-1) \times d$ matrix, and $z = (\widetilde{z}, \widehat{z}) \in \mathbb{R}^d$. We will assume that there is $C > 0$ such that

$$|\widetilde{A}_t^z| \le \mathrm{v}/2, \qquad |\sigma|, |\widehat{A}_t^z| \le C. \tag{21}$$

For $A \subset \mathbb{R}$, let $\widetilde{\tau}^{z,\varepsilon}(A) = \inf\{t \ge 0 : \widetilde{Z}_t^{z,\varepsilon} \in A\}$.

LEMMA 3.4. *There are* $\varepsilon_0 > 0$ *and* $L > 0$ *such that*

$$\mathrm{E}\big(\lambda\big(t : \widetilde{Z}_t^{0,\varepsilon} \in [0, \varepsilon], t \le \widetilde{\tau}^{0,\varepsilon}(\{-\sqrt{\varepsilon}\} \cup \{\varepsilon\})\big)\big) \le L\varepsilon^2, \tag{22}$$

provided that $\varepsilon \le \varepsilon_0$. *If, additionally,*

$$|\widetilde{A}_t^0| \le \varkappa(\varepsilon) \qquad \text{whenever } |\widetilde{Z}_t^{0,\varepsilon}| + |\widehat{Z}_t^{0,\varepsilon}| \le \sqrt{\varepsilon}, \text{ with } \varkappa(\varepsilon) \downarrow 0 \text{ as } \varepsilon \downarrow 0, \tag{23}$$

then

$$\lim_{\varepsilon \downarrow 0} \mathrm{P}\big(\widetilde{\tau}^{0,\varepsilon}(\{\varepsilon\}) < \widetilde{\tau}^{0,\varepsilon}(\{-\sqrt{\varepsilon}\})\big) = (1 + 2\mathrm{v})^{-1}. \tag{24}$$

PROOF. Let $\eta \in (0, 1)$. Find $z_0 < 0$ such that (17) holds and

$$\mathrm{P}\big(\widetilde{\tau}^{z_0\varepsilon,\varepsilon}(\{0\}) < \infty\big) \le \eta, \tag{25}$$

which is possible by (15) and the scaling-invariance of a Wiener process. By (16), we can find t_0 such that for $z \in [z_0, 1]$,

$$\mathrm{P}\big(\widetilde{\tau}^{z\varepsilon,\varepsilon}(\{z_0\varepsilon\} \cup \{\varepsilon\}) \le t_0\varepsilon^2\big) \ge 1 - \eta. \tag{26}$$

The combination of these two inequalities and the strong Markov property (the use of which is allowed since a time-shift of the process $\widetilde{A}_t^z$ is also bounded by $\mathrm{v}/2$) imply (22). By (19), (20), (23) and (26),

$$\begin{aligned}
&\mathrm{P}\Big(\sup_{t \le \widetilde{\tau}^{0,\varepsilon}(\{z_0\varepsilon\}\cup\{\varepsilon\})} |\widetilde{A}_t^0| \le \varkappa(\varepsilon)\Big) \\
&\qquad \ge \mathrm{P}\Big(\sup_{t \le \widetilde{\tau}^{0,\varepsilon}(\{z_0\varepsilon\}\cup\{\varepsilon\})} (|\widetilde{Z}_t^{0,\varepsilon}| + |\widehat{Z}_t^{0,\varepsilon}|) \le \sqrt{\varepsilon}\Big) \\
&\qquad \ge 1 - 2\eta
\end{aligned}$$

for all sufficiently small ε. Therefore, by (18), which can be applied after rescaling,

$$\mathrm{P}\big(\widetilde{\tau}^{0,\varepsilon}(\{\varepsilon\}) < \widetilde{\tau}^{0,\varepsilon}(\{z_0\varepsilon\})\big) \in \big[(1 + 2\mathrm{v})^{-1} - 5\eta, (1 + 2\mathrm{v})^{-1} + 5\eta\big].$$

Together with (25), this implies that

$$\mathrm{P}\big(\widetilde{\tau}^{0,\varepsilon}(\{\varepsilon\}) < \widetilde{\tau}^{0,\varepsilon}(\{-\sqrt{\varepsilon}\})\big) \in \big[(1 + 2\mathrm{v})^{-1} - 6\eta, (1 + 2\mathrm{v})^{-1} + 6\eta\big].$$

Since η was arbitrary, this implies (24). □

From the presence of a strong drift to the left in (19), it easily follows that there is $L > 0$ such that

$$\sup_{z \in [-\sqrt{\varepsilon}, \varepsilon]} \mathrm{E}\widetilde{\tau}^{z,\varepsilon}\big(\{-\sqrt{\varepsilon}\} \cup \{\varepsilon\}\big) \le L\varepsilon^{\frac{3}{2}}. \tag{27}$$

From (19), (20) and (27), it follows that

$$\lim_{\varepsilon \downarrow 0} \mathrm{P}\Big(\sup_{t \le \widetilde{\tau}^{0,\varepsilon}(\{-\sqrt{\varepsilon}\}\cup\{\varepsilon\})} (|\widetilde{Z}_t^{0,\varepsilon}| + |\widehat{Z}_t^{0,\varepsilon}|) > \varepsilon^{\frac{1}{3}}\Big) = 0. \tag{28}$$

(3) Now we will apply Lemma 3.4 to study the behavior of the process $X_t^{x,\varepsilon}$ in the vicinity of ∂D. Namely, we will prove the following lemma.

LEMMA 3.5. *For each sufficiently small $\delta > 0$, there are $\varepsilon_0 > 0$ and $L > 0$ such that*

$$\mathrm{E}\big(\lambda\big(t : X_t^{x,\varepsilon} \in \overline{U}, t \le \min\big(\tau^{x,\varepsilon}(S_{-\delta}), \tau^{x,\varepsilon}(S_\varepsilon)\big)\big)\big) \le L\varepsilon^2, \qquad x \in \partial D, \tag{29}$$

provided $\varepsilon \le \varepsilon_0$. Moreover,

$$\lim_{\varepsilon \downarrow 0} \mathrm{P}\big(\tau^{x,\varepsilon}(S_\varepsilon) < \tau^{x,\varepsilon}(S_{-\delta})\big) = \big(1 + 2a(x)\big)^{-1} \qquad \textit{uniformly in } x \in \partial D. \tag{30}$$

PROOF. First, observe that

$$\lim_{\varepsilon \downarrow 0} \varepsilon^{-2} \mathrm{P}\big(\tau^{x,\varepsilon}(S_\varepsilon) < \tau^{x,\varepsilon}(S_{-\delta})\big) = 0 \qquad \text{uniformly in } x \in S_{-\sqrt{\varepsilon}}, \tag{31}$$

due to the presence of the strong drift inside D. Also, there is $L > 0$ such that

$$\mathrm{E}\min\big(\tau^{x,\varepsilon}(\partial D), \tau^{x,\varepsilon}(S_\varepsilon)\big) \le L\varepsilon^2, \qquad x \in \Gamma_\varepsilon, \tag{32}$$

for all sufficiently small ε, since $X_t^{x,\varepsilon}$ is a Wiener process on U.

Let us describe a change of coordinates in a neighborhood of a point $x \in \partial D$. Let $V_\varepsilon^r = [-\sqrt{\varepsilon}, \varepsilon] \times B_r \subset \mathbb{R}^d$, where $B_r \subset \mathbb{R}^{d-1}$ is the closed ball of radius r centered at the origin. Let m_x be an isometric mapping of B_r to the $(d-1)$-dimensional ball of radius r centered at x in the tangent plane to ∂D at x. For $y \in B_r$, we take the straight line passing through $m_x(y)$ and the point on ∂D closest to $m_x(y)$ [this line is perpendicular to ∂D if $m_x(y) \notin \partial D$; we define it as the perpendicular if $m_x(y) \in \partial D$]. For $z \in [-\sqrt{\varepsilon}, \varepsilon]$ and $y \in B_r$, define $\varphi_x(z, y)$ as the point on the perpendicular that belongs to S_z. If r and ε are sufficiently small, then φ_x is a diffeomorphism from V_ε^r to a domain $U_\varepsilon^r(x)$ for each x (see Figure 2).

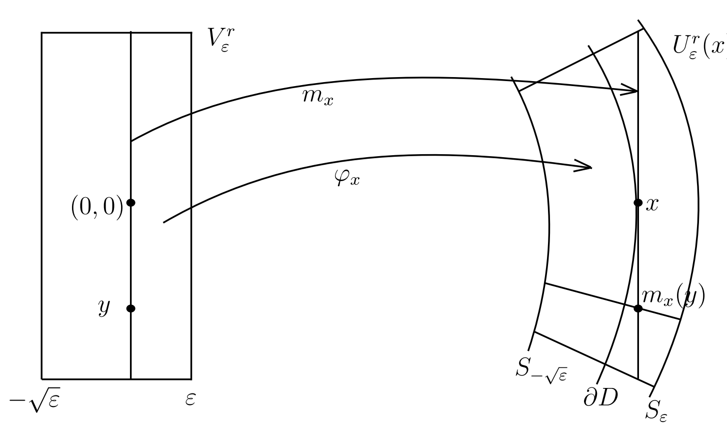

FIG. 2. *The diffeomorphism φ_x.*

For $z \in V_\varepsilon^r$, let $\overline{X}_t^{z,\varepsilon} = \varphi^{-1}(X_t^{\varphi(z),\varepsilon})$ be the process written in the new coordinates (stopped when it reaches the boundary of V_ε^r). It satisfies

$$d\overline{X}_t^{z,\varepsilon} = \frac{1}{\varepsilon}\overline{v}(\overline{X}_t^{z,\varepsilon})\chi_{[-\sqrt{\varepsilon},0]\times\mathbb{R}^{d-1}}(\overline{X}_t^{z,\varepsilon})\,dt + \overline{\beta}(\overline{X}_t^{z,\varepsilon})\,dt + \overline{\sigma}(\overline{X}_t^{z,\varepsilon})\,d\overline{W}_t, \qquad \overline{X}_0^{z,\varepsilon} = z.$$

The coefficients $\overline{v}, \overline{\beta}, \overline{\sigma}$ are bounded in $C^1(V_\varepsilon^r)$ (the change of coordinates and, therefore, the coefficients depend on x, but the bound is uniform in x) and satisfy: $\overline{v}_1(0) = -a(x)$, $\overline{\sigma}(0)$ is an orthogonal matrix. Let

$$\alpha(z) = (\sigma_{11}^2(z) + \cdots + \sigma_{1d}^2(z))^{-\frac{1}{2}}.$$

Note that this is a smooth function such that $\alpha(0) = 1$. The process

$$d\widetilde{X}_t^{z,\varepsilon} = \frac{1}{\varepsilon}\alpha^2\overline{v}\chi_{[-\sqrt{\varepsilon},0]\times\mathbb{R}^{d-1}}(\widetilde{X}_t^{z,\varepsilon})\,dt + \alpha^2\overline{\beta}(\widetilde{X}_t^{z,\varepsilon})\,dt + \alpha\overline{\sigma}(\widetilde{X}_t^{z,\varepsilon})\,d\overline{W}_t, \qquad \widetilde{X}_0^{z,\varepsilon} = z,$$

is different from $\overline{X}_t^{z,\varepsilon}$ by a random change of time. The coefficients of $\widetilde{X}_t^{z,\varepsilon}$ can be extended from $V_\varepsilon^r = [-\sqrt{\varepsilon}, \varepsilon] \times B_r \subset \mathbb{R}^d$ to $[-\sqrt{\varepsilon}, \varepsilon] \times \mathbb{R}^{d-1}$ by requiring that they do not vary in the radial direction outside B_r. This way the process $\widetilde{X}_t^{z,\varepsilon}$ can be defined until the time it reaches the boundary of $[-\sqrt{\varepsilon}, \varepsilon] \times \mathbb{R}^{d-1}$. Let $\widetilde{Z}_t^{z,\varepsilon}$ be the first coordinate of $\widetilde{X}_t^{z,\varepsilon}$ and $\widehat{Z}_t^{z,\varepsilon}$ be the vector consisting of the remaining $d-1$ coordinates of $\widetilde{X}_t^{z,\varepsilon}$. Note that the process $(\widetilde{Z}_t^{z,\varepsilon}, \widehat{Z}_t^{z,\varepsilon})$ can be written in the form (19)–(20) with the coefficients satisfying (21), (23), provided that r and ε are chosen to be sufficiently small, independently of x.

By (28),

$$\lim_{\varepsilon\downarrow 0} \mathrm{P}(\min(\tau^{x,\varepsilon}(S_{-\sqrt{\varepsilon}}), \tau^{x,\varepsilon}(S_\varepsilon)) > \tau^{x,\varepsilon}(\partial U_\varepsilon^r(x))) = 0, \tag{33}$$

and it is not difficult to see that the limit is uniform in $x \in \partial D$. Since α is bounded from above and below, from Lemma 3.4 [formulas (22) and (24)] it follows that there are $\varepsilon_0 > 0$ and $L > 0$ such that

$$\mathrm{E}(\lambda(t : X_t^{x,\varepsilon} \in \overline{U}, t \le \tau^{x,\varepsilon}(\partial U_\varepsilon^r(x)))) \le L\varepsilon^2, \qquad x \in \partial D, \tag{34}$$

provided that $\varepsilon \le \varepsilon_0$. Moreover,

$$\lim_{\varepsilon\downarrow 0} \mathrm{P}(\tau^{x,\varepsilon}(S_\varepsilon) < \tau^{x,\varepsilon}(S_{-\sqrt{\varepsilon}})) = (1 + 2a(x))^{-1}, \qquad x \in \partial D. \tag{35}$$

The convergence is uniform in x since the dependence of the process $(\widetilde{Z}_t^{z,\varepsilon}, \widehat{Z}_t^{z,\varepsilon})$ on x manifests itself through the value of $\mathrm{v} = a(x)$ and through the values of C, $\varkappa(\varepsilon)$ in (21), (23) in a way that does not affect the applicability of (22) and (24).

The strong Markov property of the process $X_t^{x,\varepsilon}$, together with (31), (32) and (33), allows us to obtain (29) from (34) and (30) from (35). $\square$

(4) Let us get a bound on $\mathrm{E}\lambda(u : u \le \sigma_1^{x,\varepsilon}, X_u^{x,\varepsilon} \in \overline{U})$, $x \in \partial D$. Since $X_t^{x,\varepsilon}$ is a Wiener process on U, there are $\varepsilon_0 > 0$ and $L > 0$ such that

$$\mathrm{E}\min(\tau^{x,\varepsilon}(\partial D), \tau^{x,\varepsilon}(S_{\sqrt{\varepsilon}})) \le L\varepsilon^{\frac{3}{2}}, \qquad x \in S_\varepsilon, \tag{36}$$

provided that $\varepsilon \le \varepsilon_0$, while

$$\lim_{\varepsilon \downarrow 0}(\varepsilon^{-\frac{1}{2}}\mathrm{P}(\tau^{x,\varepsilon}(S_{\sqrt{\varepsilon}}) < \tau^{x,\varepsilon}(\partial D))) = 1 \qquad \text{uniformly in } x \in S_\varepsilon. \tag{37}$$

Let δ be sufficiently small for (29) and (30) to hold. Formulas (36), (37), together with (29), (30), and the strong Markov property of the process, imply that there are $\varepsilon_0 > 0$ and $L > 0$ such that for all sufficiently small ε,

$$\mathrm{E}\lambda(u : u \le \sigma_1^{x,\varepsilon}, X_u^{x,\varepsilon} \in \overline{U}) \le L\varepsilon, \qquad x \in \partial D. \tag{38}$$

Let $\xi_n^{x,\varepsilon} = \lambda(u : \tau_n^{x,\varepsilon} \le u \le \sigma_n^{x,\varepsilon}, X_u^{x,\varepsilon} \in \overline{U})$. Formulas (38) and (12), together with the strong Markov property of the process, imply that there is $c = c(t)$ such that

$$\mathrm{E}\sum_{n=0}^{\infty} \xi_{n+1}^{x,\varepsilon} \chi_{\{\sigma_n^{x,\varepsilon} \le s(t)\}} \le c\sqrt{\varepsilon}, \qquad x \in \overline{U}. \tag{39}$$

(5) The next chain of arguments will relate the stopping times to the coefficient $1/2a(x)$, which appears in the definition of the measure ν_k in (7). We claim that for sufficiently small δ there are $\varepsilon_0 > 0$ and $L > 0$ such that

$$\mathrm{E}(|X^{x,\varepsilon}_{\tau^{x,\varepsilon}(S_\varepsilon)} - x|^2, \tau^{x,\varepsilon}(S_\varepsilon) < \tau^{x,\varepsilon}(S_{-\delta})) \le L\varepsilon^2, \qquad x \in \partial D, \tag{40}$$

provided that $\varepsilon \le \varepsilon_0$. Let us sketch a proof of this statement. First, by observing the process in the ε-neighborhood of ∂D, it is easy to see that

$$\mathrm{E}|X^{x,\varepsilon}_{\tau^{x,\varepsilon}(S_\varepsilon) \wedge \tau^{x,\varepsilon}(S_{-\varepsilon})} - x|^2 \le L\varepsilon^2, \qquad x \in \partial D. \tag{41}$$

From the the presence of the strong drift inside D, it follows that there is $c > 0$ such that

$$\mathrm{P}(\tau^{x,\varepsilon}(\partial D) > t\varepsilon^2, \tau^{x,\varepsilon}(\partial D) < \tau^{x,\varepsilon}(S_{-\delta})) \le e^{-ct}, \qquad x \in \partial S_{-\varepsilon}, t \ge 0, \tag{42}$$

and consequently,

$$\mathrm{E}(|X^{x,\varepsilon}_{\tau^{x,\varepsilon}(\partial D)} - x|^2, \tau^{x,\varepsilon}(\partial D) < \tau^{x,\varepsilon}(S_{-\delta})) \le L\varepsilon^2, \qquad x \in \partial S_{-\varepsilon}. \tag{43}$$

Also note that there is $c > 0$ such that

$$\mathrm{P}(\tau^{x,\varepsilon}(\partial D) < \tau^{x,\varepsilon}(S_{-\delta})) \le 1 - c, \qquad x \in \partial S_{-\varepsilon}. \tag{44}$$

By considering consecutive visits of the process to $S_{-\varepsilon}$ and ∂D, employing (41), (43), (44) and using the strong Markov property, we obtain (40).

Take the compact set $K \subset D$ such that $\partial K = S_{-\delta}$, where δ is sufficiently small for (29), (30) and (40) to hold. Observe that, since $X_t^{x,\varepsilon}$ is a Wiener process in U, by (36),

$$\begin{aligned} \mathrm{E}\big|X^{x,\varepsilon}_{\tau^{x,\varepsilon}(\partial D)\wedge\tau^{x,\varepsilon}(S_{\sqrt{\varepsilon}})} - x\big|^2 &= d\cdot \mathrm{E}\big(\tau^{x,\varepsilon}(\partial D)\wedge \tau^{x,\varepsilon}(S_{\sqrt{\varepsilon}})\big) \\ &\le L\varepsilon^{\frac{3}{2}}, \qquad x\in S_\varepsilon. \end{aligned} \tag{45}$$

Therefore, by (40), (30) and (37), it follows from the strong Markov property that

$$\begin{aligned} \lim_{\varepsilon\downarrow 0}\big(\varepsilon^{-\frac{1}{2}}\mathrm{P}(\sigma_1^{x,\varepsilon} < \tau^{x,\varepsilon}(K))\big) &= \sum_{n=1}^{\infty}\left(\frac{1}{1+2a(x)}\right)^n \\ &= \frac{1}{2a(x)} \qquad \text{uniformly in } x\in\partial D. \end{aligned} \tag{46}$$

(6) Next, we obtain an estimate on the displacement of the process before the time $\sigma_1^{x,\varepsilon}$. Combining (30), (37), (40) and (45), and using the strong Markov property, we obtain that there are $\varepsilon_0 > 0$ and $L > 0$ such that

$$\mathrm{E}\big(\big|X^{x,\varepsilon}_{\sigma_1^{x,\varepsilon}} - x\big|^2, \sigma_1^{x,\varepsilon} < \tau^{x,\varepsilon}(K)\big) \le L\varepsilon^{\frac{3}{2}}, \qquad x\in\partial D, \tag{47}$$

provided that $\varepsilon \le \varepsilon_0$. Therefore,

$$\mathrm{E}\big(\big|X^{x,\varepsilon}_{\sigma_1^{x,\varepsilon}} - x\big|, \sigma_1^{x,\varepsilon} < \tau^{x,\varepsilon}(K)\big) \le L\varepsilon, \qquad x\in\partial D, \tag{48}$$

provided that $\varepsilon \le \varepsilon_0$, with a different constant L.

(7) Recall that the second sum on the right-hand side of (11) contains the terms $f(X^{x,\varepsilon}_{\sigma_n^{x,\varepsilon}\wedge s(t)}) - f(X^{x,\varepsilon}_{\tau_n^{x,\varepsilon}\wedge s(t)})$. The next lemma will help us bound the expectations of such expressions. Note that it is precisely in the proof of this lemma where condition (6) on the distribution of the exit location and condition (5) [in the form (10)] on the function f are used.

LEMMA 3.6. *Let $\overline{f}$ be the value of f on ∂D. Then*

$$\lim_{\varepsilon\downarrow 0}\big(\varepsilon^{-\frac{1}{2}}\mathrm{E}\big(f\big(X^{x,\varepsilon}_{\sigma_1^{x,\varepsilon}}\big) - \overline{f}\big)\big) = 0 \qquad \textit{uniformly in } x\in\partial D. \tag{49}$$

PROOF. Introduce the following two sequences of stopping times: $\overline{\tau}_1^{x,\varepsilon} = \tau^{x,\varepsilon}(\partial D)$, $\overline{\sigma}_n^{x,\varepsilon} = \inf(t \ge \overline{\tau}_n^{x,\varepsilon} : X_t^{x,\varepsilon} \in K)$, $n \ge 1$, while $\overline{\tau}_n^{x,\varepsilon} = \inf(t \ge \overline{\sigma}_{n-1}^{x,\varepsilon} : X_t^{x,\varepsilon} \in \partial D)$, $n \ge 2$. Then

$$\begin{aligned} &\mathrm{E}f\big(X^{x,\varepsilon}_{\sigma_1^{x,\varepsilon}}\big) - \overline{f} \\ &\qquad = \sum_{n=1}^{\infty}\mathrm{E}\big(\big(f\big(X^{x,\varepsilon}_{\sigma_1^{x,\varepsilon}}\big) - \overline{f}\big), \overline{\tau}_n^{x,\varepsilon} < \sigma_1^{x,\varepsilon} < \overline{\tau}_{n+1}^{x,\varepsilon}\big) \end{aligned}$$

$$= \mathrm{E}\big(\big(f(X^{x,\varepsilon}_{\sigma_1^{x,\varepsilon}}) - \overline{f}\big), \sigma_1^{x,\varepsilon} < \overline{\tau}_2^{x,\varepsilon}\big)$$

$$+ \sum_{n=2}^{\infty} \mathrm{E}\big(\big(f(X^{x,\varepsilon}_{\sigma_1^{x,\varepsilon}}) - \overline{f}\big), \sigma_1^{x,\varepsilon} < \overline{\tau}_{n+1}^{x,\varepsilon} | \sigma_1^{x,\varepsilon} > \overline{\sigma}_{n-1}^{x,\varepsilon}\big) \mathrm{P}\big(\sigma_1^{x,\varepsilon} > \overline{\sigma}_{n-1}^{x,\varepsilon}\big).$$

By (46) and the strong Markov property of the process, there is $c > 0$ such that

$$\mathrm{P}\big(\sigma_1^{x,\varepsilon} > \overline{\sigma}_{n-1}^{x,\varepsilon}\big) \le (1 - c\sqrt{\varepsilon})^{n-1}.$$

Also observe that

$$\big|\mathrm{E}\big(\big(f(X^{x,\varepsilon}_{\sigma_1^{x,\varepsilon}}) - \overline{f}\big), \sigma_1^{x,\varepsilon} < \overline{\tau}_{n+1}^{x,\varepsilon} | \sigma_1^{x,\varepsilon} > \overline{\sigma}_{n-1}^{x,\varepsilon}\big)\big|$$

$$\le \sup_{x \in K} \big|\mathrm{E}\big(\big(f(X^{x,\varepsilon}_{\sigma_1^{x,\varepsilon}}) - \overline{f}\big), \sigma_1^{x,\varepsilon} < \overline{\tau}_2^{x,\varepsilon}\big)\big|$$

$$\le \sup_{x \in K} \big|\mathrm{E}\big(\big\langle \nabla f(X^{x,\varepsilon}_{\overline{\tau}_1^{x,\varepsilon}}), X^{x,\varepsilon}_{\sigma_1^{x,\varepsilon}} - X^{x,\varepsilon}_{\overline{\tau}_1^{x,\varepsilon}}\big\rangle, \sigma_1^{x,\varepsilon} < \overline{\tau}_2^{x,\varepsilon}\big)\big|$$

$$+ C \sup_{x \in \partial D} \mathrm{E}\big(\big|X^{x,\varepsilon}_{\sigma_1^{x,\varepsilon}} - x\big|^2, \sigma_1^{x,\varepsilon} < \overline{\tau}_2^{x,\varepsilon}\big),$$

where C depends on the $C^2(\overline{U})$-norm of f. The second term on the right-hand side is bounded by $C\varepsilon^{\frac{3}{2}}$, with a different constant C, using (47). In order to estimate the first term, we notice that if $x \in \partial D$ and $y \in S_{\sqrt{\varepsilon}}$, then, since $\nabla f(x)$ is orthogonal to ∂D,

$$\big|\langle \nabla f(x), y - x\rangle + \sqrt{\varepsilon}\langle \nabla f(x), n(x)\rangle\big| \le c|x - y|^2$$

for some $c > 0$, where $n(x)$ is the unit inward normal to the boundary at x (with respect to D). Therefore, for $x \in K$,

$$\big|\mathrm{E}\big(\big\langle \nabla f(X^{x,\varepsilon}_{\overline{\tau}_1^{x,\varepsilon}}), X^{x,\varepsilon}_{\sigma_1^{x,\varepsilon}} - X^{x,\varepsilon}_{\overline{\tau}_1^{x,\varepsilon}}\big\rangle, \sigma_1^{x,\varepsilon} < \overline{\tau}_2^{x,\varepsilon}\big)$$

$$+ \sqrt{\varepsilon}\mathrm{E}\big(\big\langle \nabla f(X^{x,\varepsilon}_{\overline{\tau}_1^{x,\varepsilon}}), n\big(X^{x,\varepsilon}_{\overline{\tau}_1^{x,\varepsilon}}\big)\big\rangle, \sigma_1^{x,\varepsilon} < \overline{\tau}_2^{x,\varepsilon}\big)\big|$$

$$\le C\mathrm{E}\big(\big|X^{x,\varepsilon}_{\sigma_1^{x,\varepsilon}} - X^{x,\varepsilon}_{\overline{\tau}_1^{x,\varepsilon}}\big|^2, \sigma_1^{x,\varepsilon} < \overline{\tau}_2^{x,\varepsilon}\big),$$

where C depends on the $C^1(\overline{U})$-norm of f. The right-hand side is bounded by $C\varepsilon^{\frac{3}{2}}$, with a different constant C, using (47) and the strong Markov property. Observe that

$$\lim_{\varepsilon \downarrow 0} \sup_{x \in K} \big|\varepsilon^{-\frac{1}{2}}\mathrm{E}\big(\big\langle \nabla f(X^{x,\varepsilon}_{\overline{\tau}_1^{x,\varepsilon}}), n\big(X^{x,\varepsilon}_{\overline{\tau}_1^{x,\varepsilon}}\big)\big\rangle, \sigma_1^{x,\varepsilon} < \overline{\tau}_2^{x,\varepsilon}\big)\big| = 0$$

by (6), (7), (10), (46) and the strong Markov property of the process. By (48), we also have

$$\lim_{\varepsilon \downarrow 0} \sup_{x \in \partial D} \big|\varepsilon^{-\frac{1}{2}}\mathrm{E}\big(\big(f(X^{x,\varepsilon}_{\sigma_1^{x,\varepsilon}}) - \overline{f}\big), \sigma_1^{x,\varepsilon} < \overline{\tau}_2^{x,\varepsilon}\big)\big| = 0.$$

Combining the estimates above, we obtain (49). □

(8) Finally, let us gather all the ingredients and complete the proof of the proposition. Let us examine the second term in the right-hand side of (11). From (39) and the boundedness of Δf, it follows that

$$\lim_{\varepsilon\downarrow 0} \mathrm{E} \sum_{n=1}^{\infty} \int_{\tau_n^{x,\varepsilon}\wedge s(t)}^{\sigma_n^{x,\varepsilon}\wedge s(t)} \Delta f\big(X_u^{x,\varepsilon}\big)\,du = 0.$$

It remains to show that

$$\lim_{\varepsilon\downarrow 0} \mathrm{E} \sum_{n=1}^{\infty} \big(f\big(X^{x,\varepsilon}_{\sigma_n^{x,\varepsilon}\wedge s(t)}\big) - f\big(X^{x,\varepsilon}_{\tau_n^{x,\varepsilon}\wedge s(t)}\big)\big) = 0. \tag{50}$$

Introduce the stopping time

$$s'(t) = \begin{cases} \sigma_n^{x,\varepsilon}, & \text{if } \tau_n^{x,\varepsilon} < s(t) \le \sigma_n^{x,\varepsilon}, \\ s(t), & \text{otherwise.} \end{cases}$$

Now (50) will follow if we show that

$$\lim_{\varepsilon\downarrow 0} \mathrm{E} \sum_{n=1}^{\infty} \big(f\big(X^{x,\varepsilon}_{\sigma_n^{x,\varepsilon}\wedge s'(t)}\big) - \overline{f}\big) = 0, \tag{51}$$

since the difference between (51) and (50) is estimated from above by $2\sup_{x\in S_{\sqrt{\varepsilon}}} |f(x) - \overline{f}|$, which goes to zero as $\varepsilon \downarrow 0$. Let $N^{x,\varepsilon} = \max(n : \sigma_n^{x,\varepsilon} \le s'(t))$. By the strong Markov property,

$$\sup_{x\in\overline{U}} \mathrm{E} \sum_{n=1}^{\infty} \big(f\big(X^{x,\varepsilon}_{\sigma_n^{x,\varepsilon}\wedge s'(t)}\big) - \overline{f}\big) \le \sup_{x\in\overline{U}} \mathrm{E} N^{x,\varepsilon} \sup_{x\in\partial D} \mathrm{E}\big(f\big(X^{x,\varepsilon}_{\sigma_1^{x,\varepsilon}}\big) - \overline{f}\big).$$

The right-hand side tends to zero by (12) and (49). This concludes the proof of Proposition 3.2. □

PROOF OF THEOREM 3.1. Recall that Ψ is the set of functions $\psi \in C(U')$ that have limits of all the first-order derivatives as $y \to x$, $y \in U$, at all points $x \in \partial U$. This is a measure-defining class of functions on U', that is, if μ_1 and μ_2 satisfy $\int_{U'} \psi\, d\mu_1 = \int_{U'} \psi\, d\mu_2$ for every $\psi \in \Psi$, then $\mu_1 = \mu_2$. As shown in Section 2, for every $\psi \in \Psi$ and every $\lambda > 0$, there is $f \in \mathcal{D}(A)$ that satisfies $\lambda f - Af = \psi$. We have demonstrated that (9) holds for $f \in \mathcal{D}(A)$. The extension of (9) from $x \in \overline{U}$ to $x \in \mathbb{T}^d$ is trivial. By Lemma 3.1 in Chapter 8 of [4], this is sufficient to guarantee the convergence if, in addition, the family $\{Y_t^{x,\varepsilon}\}$, $\varepsilon > 0$, $x \in \mathbb{T}^d$, is tight. The tightness, however, is clear since the processes coincide with a Wiener process inside U, while all the points of ∂U are identified. □

4. Applications, generalizations and remarks.

4.1. *The behavior of the process at exponential time scales.* Let us now discuss the behavior, as $\varepsilon \downarrow 0$, of the original process $X_t^{x,\varepsilon}$ (rather than its trace $Y_t^{x,\varepsilon}$

on $\overline{U}$). If the process starts in a small neighborhood of O_k, then it takes time of order $\exp(V_k/\varepsilon)$ (in the sense of logarithmic equivalence) for it to reach ∂D_k, where $V_k = \inf_{x\in\partial D_k} V_k(x)$ and $V_k(x)$ is the quasi-potential defined in (4). Thus, it is reasonable to study the behavior of $X_t^{x,\varepsilon}$ at exponential time scales, that is, at times of order $\exp(\lambda/\varepsilon)$ with fixed λ.

The transitions between small neighborhoods of the equilibriums are governed by the matrix V_{ij} defined in (2). In our case, $V_{ij} = V_i$ for all i, j, as explained in the Introduction. Because of this "rough symmetry" ([3]), the notion of a metastable state [see (3)] should be replaced by that of a metastable distribution between the equilibriums.

To describe the metastable distribution for a given initial point $x \in \mathbb{T}^d$ and time scale $\exp(\lambda/\varepsilon)$ with $\lambda > 0$, assume that $V_1 < V_2 < \cdots < V_n$, and put $V_0 = 0$, $V_{n+1} = \infty$. We introduce the following nonstandard boundary problem, which will be referred to as the (k, j)-problem on U. Namely, for each $1 \le k \le n$ and $k \le j \le n$, let $u_{k,j}$ solve the problem

$$
\begin{aligned}
&\Delta u_{k,j}(x) = 0, \qquad x \in U;\\
&u_{k,j}(x) = c^i_{k,j}, \qquad x \in \partial D_i \quad \text{and} \quad \int_{\partial D_i} \langle \nabla u_{k,j}(x), n(x)\rangle \nu_i(dx) = 0\\
&\hspace{20em} \text{for } 1 \le i < k;\\
&u_{k,j}(x) = 0, \qquad x \in \partial D_i \qquad \text{for } i \ge k, i \ne j;\\
&u_{k,j}(x) = 1, \qquad x \in \partial D_j.
\end{aligned}
$$

The constants $c^i_{k,j}$ are not prescribed, that is, solving the (k, j)-problem includes finding the boundary values of the function on ∂D_i, $i < k$. As is shown in Section 2, the solution exists and is unique in $C^2(\overline{U})$.

THEOREM 4.1. *Assume that $V_{k-1} < \lambda < V_k$ for some $1 \le k \le n$. Suppose that $T(\varepsilon)$ is such that $\lim_{\varepsilon\downarrow 0}(\varepsilon \ln T(\varepsilon)) = \lambda$. Let $\mathcal{E}_i \subseteq D_i$, $i = 1, \dots, n$, be arbitrary neighborhoods of O_i, $i = 1, \dots, n$. Then*

$$\lim_{\varepsilon\downarrow 0} \mathrm{P}\big(X^{x,\varepsilon}_{T(\varepsilon)} \in \mathcal{E}_j\big) = u_{kj}(x), \qquad x \in U, j \ge k; \tag{52}$$

$$\lim_{\varepsilon\downarrow 0} \mathrm{P}\big(X^{x,\varepsilon}_{T(\varepsilon)} \in \mathcal{E}_j\big) = c^i_{k,j}, \qquad x \in \overline{D}_i, i < k, j \ge k; \tag{53}$$

$$\lim_{\varepsilon\downarrow 0} \mathrm{P}\big(X^{x,\varepsilon}_{T(\varepsilon)} \in \mathcal{E}_j\big) = 1, \qquad x \in \overline{D}_j, j \ge k, \tag{54}$$

where $u_{k,j}$ is the solution of the (k, j)-problem and $c^i_{k,j}$ is the value of $u_{k,j}$ on ∂D_i. If $\lambda \ge V_n$, then

$$\lim_{\varepsilon\downarrow 0} \mathrm{P}\big(X^{x,\varepsilon}_{T(\varepsilon)} \in \mathcal{E}_n\big) = 1, \qquad x \in \mathbb{T}^d. \tag{55}$$

Observe that $\sum_{j=k}^{n} u_{kj}(x) = 1$ for $x \in U$, while $\sum_{j=k}^{n} c_{k,j}^{i} = 1$ for $x \in \overline{D}_i$ if $i < k$. Thus, with probability close to one, $X_{T(\varepsilon)}^{x,\varepsilon}$ is located in a small neighborhood of one of the equilibrium points.

SKETCH OF THE PROOF. Without loss of generality, we can assume that $\overline{\mathcal{E}}_i \subset D_i$, $i = 1, \dots, n$. First, consider the case when $V_{k-1} < \lambda < V_k$ for some $1 \le k \le n$. Observe that $\tau^{x,\varepsilon}(\overline{\mathcal{E}}_j)/T(\varepsilon) \to 0$ in probability as $\varepsilon \downarrow 0$ since the unperturbed system starting at $x \in \overline{D}_j$ is attracted to O_j. If $x \in \overline{\mathcal{E}}_j$, $j \ge k$, then $\lim_{\varepsilon \downarrow 0} \mathrm{P}(X_{T'(\varepsilon)}^{x,\varepsilon} \in \mathcal{E}_j) = 1$ uniformly in $x \in \overline{\mathcal{E}}_j$ and $T'(\varepsilon) \in [T(\varepsilon)/2, T(\varepsilon)]$. Combining these two observations and using the strong Markov property of the process, we obtain (54).

Now assume that $x \in \bigcup_{i<k} \overline{D}_i \cup U$. In this case, it is easy to see ([4], Chapter 4) that the time it takes $X_t^{x,\varepsilon}$ to reach $\bigcup_{j \ge k} \partial D_j$ is significantly smaller than $T(\varepsilon)$, that is, $\tau^{x,\varepsilon}(\bigcup_{j \ge k} \partial D_j)/T(\varepsilon) \to 0$ in probability as $\varepsilon \downarrow 0$. Therefore, with probability close to one, the process will reach ∂D_j for some $j \ge k$ before time $T(\varepsilon)/2$. By (54) (applied to $T'(\varepsilon) \in [T(\varepsilon)/2, T(\varepsilon)]$ instead of $T(\varepsilon)$) and the strong Markov property, $X_{T(\varepsilon)}^{x,\varepsilon} \in \bigcup_{j \ge k} \mathcal{E}_j$ with probability close to one. The choice of $\mathcal{E}_j$ is determined by the behavior of the process $Y_t^{x,\varepsilon}$ as $\varepsilon \downarrow 0$. In fact, the solution of the (k, j)-problem gives the limiting probability of $Y_t^{x,\varepsilon}$ hitting ∂D_j prior to hitting ∂D_i with $i \ge k$, $i \ne j$. This justifies (52) and (53).

Finally, if $\lambda \ge V_n$, then at time $T(\varepsilon)$ the process will be located in an arbitrarily small neighborhood of O_n with probability close to one ([4], Chapter 6), that is, (55) holds. □

4.2. *Trapping regions with multiple equilibriums*. Up to now, we assumed that there was just one attractor (asymptotically stable equilibrium) inside each trapping region. Let us now consider an example where this is not the case. For simplicity, assume that there is one trapping region D containing two equilibriums O_1 and O_2 and one saddle point S. The structure of the vector v field on $\overline{D}$ is assumed to be as shown in Figure 3. As before, the vector field is equal to zero in $U = \mathbb{T}^2 \setminus \overline{D}$.

Let $D_k \subset D$ and $\gamma_k \subset \partial D$, $k = 1, 2$, be the sets of points that are carried to an arbitrarily small neighborhood of O_k by the deterministic flow $\dot{x}(t) = v(x(t))$. Let $A, B \in \partial D$ be the points that separate γ_1 from γ_2. Let γ be the curve that connects A with B and consists of two flow lines and the saddle point (see Figure 3). The asymptotic behavior of the process $X_t^{x,\varepsilon}$ (in exponential time scales) and of the trace $Y_t^{x,\varepsilon}$ is determined by the numbers V_{ij} defined in (2) and by the values $V_k = \inf_{x \in \partial D_k} V_k(x)$, where the quasi-potentials $V_k(x)$ are defined in (4).

Consider the case when $V_1 < V_{12}$, $V_2 < V_{21}$, and the infimum in the definition of V_k is achieved at a unique point $x_k \in \gamma_k$, $k = 1, 2$. Then, with probability that tends to one as $\varepsilon \downarrow 0$, the process exits D in an arbitrarily small neighborhood of x_k, provided that it starts in a small neighborhood of O_k, $k = 1, 2$ (see Section 6.5 of [4]).

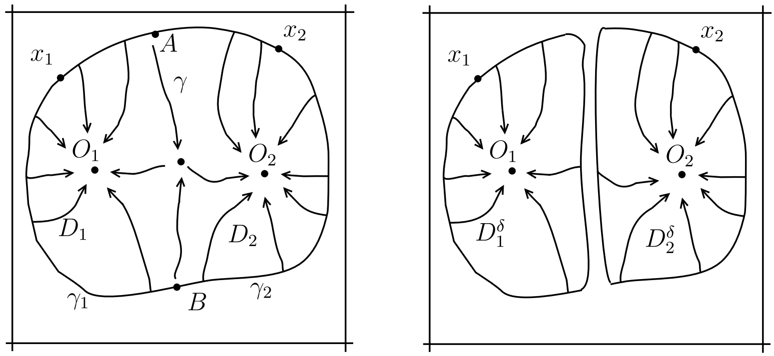

FIG. 3. *The flow lines of the original and the modified vector fields.*

For $\delta > 0$, one can consider the following auxiliary system. Let V_δ be the δ-neighborhood of γ. Let $D_k^\delta \subset D_k$, $k = 1, 2$, be domains with smooth boundaries such that $D_k^\delta \setminus V_\delta = D_k \setminus V_\delta$, yet $\overline{D_k^\delta} \cap \gamma = \varnothing$. Moreover, we can modify the vector field v (i.e., replace it by a new vector field v^δ) in such a way that $v^\delta(x) = v(x)$ for $x \notin V_\delta$, while the field v^δ satisfies the assumptions with respect to the domains $\overline{D_1^\delta}$ and $\overline{D_2^\delta}$ that were imposed on v in Section 3, that is, it is equal to zero outside $\overline{D_1^\delta} \cup \overline{D_2^\delta}$, is directed inside each of the domains on the boundary, and all the points of $\overline{D_k^\delta}$ are attracted to O_k.

The analysis of Section 3 applies to the process $X_t^{\delta,x,\varepsilon}$ defined via

$$dX_t^{\delta,x,\varepsilon} = \frac{1}{\varepsilon} v^\delta\big(X_t^{\delta,x,\varepsilon}\big)\,dt + dW_t, \qquad X_0^{\delta,x,\varepsilon} = x \in U.$$

Let $X_t^{\delta,x}$ denote the limit of the trace process in $\mathbb{T}^2 \setminus (D_1^\delta \cup D_2^\delta)$. Since $x_1, x_2 \notin V_\delta$ for all sufficiently small δ, it is not difficult to show that the probability that $X_t^{\delta,x}$ enters V_δ prior to time T tends to zero for each finite T. Moreover, $X_t^{\delta,x}$ has a limit in probability as $\delta \downarrow 0$. This limit will be denoted by X_t^x. This process is the limit, as $\varepsilon \downarrow 0$, of the trace $Y_t^{x,\varepsilon}$ of the original process $X_t^{x,\varepsilon}$. A direct construction of the process X_t^x (in terms of the generator rather that via approximating processes) seems to be technically complicated and is not presented here.

Another case where the limit of the trace process can be easily described is when $V_1 < V_{12}$ if we assume that the infimum in the definition of V_1 is achieved at a unique point $x_1 \in \gamma_1$, while the infimum in the definition of V_2 is achieved at a unique point $x_2 \in \gamma \setminus \partial D$ (which implies that $V_2 = V_{21}$). Thus, with probability that tends to one as $\varepsilon \downarrow 0$, the process exits D in an arbitrarily small neighborhood of x_1, irrespective of whether it starts in D_1 or D_2. The results of Section 3 then apply, with the limit of the exit measure μ being the point mass concentrated at x_1.

A more general situation of several equilibriums within D with various relations on the quantities V_{ij} and V_k can be analyzed using the construction above based on removing the δ-neighborhoods of the boundaries of D_k and the results of Sections 6.5–6.6 of [4] on the hierarchies of cycles.

4.3. *Other generalizations.* If the process $X_t^{x,\varepsilon}$ is governed by a more general elliptic operator,

$$L^\varepsilon = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i(x)\frac{\partial}{\partial x_i} + \frac{1}{\varepsilon}v(x)\nabla, \tag{56}$$

then the results and the proofs are similar. The definition of the numbers V_{ij} and of the quasi-potential $V_k(x)$ should now be based on the action functional corresponding to the operator L^ε. The definition (7) of the measures ν_k needs to be modified to account for the variable diffusion coefficients of the process $X_t^{x,\varepsilon}$. However, if the infimum of $V_k(x)$, $x \in \partial D_k$, is achieved at a single point x_k, then ν_k is still the δ-measure concentrated at x_k.

The assumptions on the vector field v that we made in Section 3 do not specify that D necessarily contains a single equilibrium point. They may hold, for example, if D contains a single limit cycle instead. The case of several limit cycles is technically not different from the case of several equilibrium points that we discussed above.

The results also apply to processes on general smooth manifolds, not only on a torus.

4.4. *More on the limiting process.* The nonstandard boundary problem introduced in Section 2 and the corresponding Markov process with jumps at the boundary arise in other situations, not just in the large deviation case. Consider, for example, a vector field v with closed flow lines that is equal to zero outside of $\overline{D}$ such that ∂D serves as one of the flow lines (see Figure 4).

We expect that the trace in $\mathbb{T}^2 \setminus D$ of the process $X_t^{x,\varepsilon}$ with generator (56) converges, as $\varepsilon \downarrow 0$, to the process described in Section 2. The measure ν on ∂D

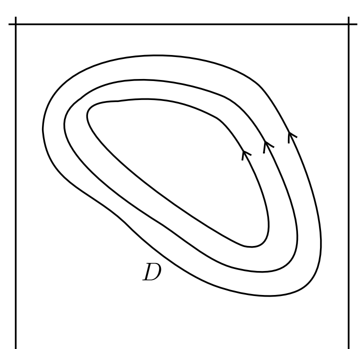

FIG. 4. *The flow lines inside D.*

will be defined by the values of v in an arbitrarily small neighborhood of ∂D and by the diffusion coefficients a_{ij} on ∂D and can be calculated explicitly.

REFERENCES

[1] DAY, M. V. (1999). Mathematical approaches to the problem of noise-induced exit. In *Stochastic Analysis*, *Control*, *Optimization and Applications* (W. McEneauey, G. Yin and Q. Zhang, eds.). 269–287. Birkhäuser, Boston, MA. MR1702965

[2] DYNKIN, E. B. (1965). *Markov Processes*. Springer, Berlin. MR0193671

[3] FREIDLIN, M. (2014). On stochastic perturbations of dynamical systems with a "rough" symmetry. Hierarchy of Markov chains. *J. Stat. Phys.* **157** 1031–1045. MR3277758

[4] FREIDLIN, M. I. and WENTZELL, A. D. (1998). *Random Perturbations of Dynamical Systems*, 2nd ed. Springer, New York. MR1652127

[5] HOLMES-CERFON, M., GORTLER, S. J. and BRENNER, M. P. (2013). A geometrical approach to computing free-energy landscapes from short-ranged potentials. *Proc. Natl. Acad. Sci. USA* **110** E5–E14.

[6] MANDL, P. (1968). *Analytical Treatment of One-Dimensional Markov Processes*. Springer, New York. MR0247667

[7] MENG, G., ARKUS, N., BRENNER, M. P. and MANOHARAN, V. (2010). The free-energy landscape of clusters of attractive hard spheres. *Science* **327** 560–563.

[8] VENTCELL, A. D. (1959). On lateral conditions for multidimensional diffusion processes. *Teor. Veroyatn. Primen.* **4** 172–185. MR0106515

M. FREIDLIN
L. KORALOV
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND 20742
USA
E-MAIL: mif@math.umd.edu
koralov@math.umd.edu

A. WENTZELL
DEPARTMENT OF MATHEMATICS
TULANE UNIVERSITY
NEW ORLEANS, LOUISIANA 70118
USA
E-MAIL: wentzell@math.tulane.edu