Journal of Statistical Physics, Vol. 93, Nos. 1/2, 1998

Couplings and Asymptotic Exponentiality of Exit Times

S. Brassesco,[1] E. Olivieri,[2] and M. E. Vares[3]

Received July 8, 1998

The goal of this note is simply to call attention to the resulting simplification in the proof of asymptotic exponentiality of exit times in the Freidlin–Wentzell regime (as proved by F. Martinelli *et al.*) by using the coupling proposed by T. Lindvall and C. Rogers.

KEY WORDS: Exit times; exponentiality; metastability; couplings.

1. INTRODUCTION

In this note we examine a classical problem in the framework of the theory of small random perturbations of dynamical systems: the first exit from a suitable domain G.

In particular, for a class of Itô equations, we address the question of the asymptotic exponentiality, in the limit of small noise, of the suitably normalized first exit time from G. We are interested in the general case of G containing many attractors of the unperturbed system. This problem is, on one side, interesting in itself; it amounts to considerably strengthen the classical Freidlin–Wentzell results on the asymptotics of the first exit time from a domain G. It is, on the other side, also related to the so-called metastable behavior of the particular stochastic dynamics described by our Itô equations, in the framework of the pathwise approach to metastability introduced in [CGOV].

From a probabilistic point of view the asymptotic exponentiality (or asymptotic unpredictability) of the exit time is related to a particular exit

[1] Instituto Venezolano de Investigaciones Científicas, Caracas, Venezuela.
[2] II Università di Roma "Tor Vergata," Rome, Italy.
[3] IMPA, Rio de Janeiro, Brazil.

mechanism: the repetition of a large number of almost independent trials. Among the various different ways the large deviation theory is able to select a particularly efficient one. So the heuristic explanation of the asymptotic exponentiality is based on a long sequence of recurrences inside G together with a loss of memory and eventually a successful exit attempt. In [GOV], very sophisticated analytical results due to Day (cf. [D]) were used to extend to a "tunneling" problem the previous results relative to the case of a domain G completely attracted by a unique asymptotically stable point. In [MOS], for a general class of domains G, the analytical methods of Day were replaced by probabilistic arguments based on contraction properties of the stochastic map (depending on the noise) which associates to the initial datum of our stochastic equation the solution at a given time T. Again the ingredient of loss of memory, necessary for the asymptotic exponentiality is based on delicate and highly non-trivial arguments developed in [MS].

In the present note, in the general [MOS] context, we give another proof of the asymptotic exponentiality by using a simple and beautiful coupling argument due to Lindvall and Rogers (cf. [LR]). The goal is to stress the resulting simplicity.

Coupling methods have been also successfully used to show asymptotic exponentiality for an infinite dimensional case, as the stochastically perturbed non-linear heat equation, also considered in [MOS]. Using a coupling introduced by Mueller in [M], Brassesco (cf. [B]) was able to treat escape times which were not treatable with the techniques considered in [MOS].

2. THE RESULT

Let $X_t^{x,\varepsilon}$ be the Markov process obtained as the unique solution of the following Itô equation:

$$\begin{aligned} dX_t^{x,\varepsilon} &= b(X_t^{x,\varepsilon})\,dt + \varepsilon\,dW_t \\ X_0^{x,\varepsilon} &= x \end{aligned} \tag{1}$$

where (W_t) is a standard d-dimensional Brownian motion, $x \in \mathbb{R}^d$, $\varepsilon > 0$, and the vector field b is assumed to be globally Lipschitz. Let us in fact, and to simplify, assume b to be of class C^1 with bounded gradient. In particular, as it is well known, this implies strong uniqueness of the solution of (1), for any given Brownian motion (W_t), as well as the strong Markov property for $(X_t^{x,\varepsilon})_t$. Of course, more general assumptions on the field b can be taken, and an extension to varying diffusion coefficients is also possible, cf. Remark 4.

Notation. Though everything is done on any probability space $(\Omega, \mathscr{A}, P)$ where (W_t) is defined, through a pathwise (and continuous) transformation, sometimes it is more convenient to relax the notation, eliminating the superscript x on $X^{x,\varepsilon}$ and using P_x to denote the probability under the condition $X_0^\varepsilon = x$.

Our goal is to discuss the asymptotic behavior, as $\varepsilon \to 0$, of the first exit time $\tau^\varepsilon =^{\text{DEF}} \inf\{t > 0; X_t^\varepsilon \notin G\}$, when $X_0^\varepsilon = x \in G$, and where G is a bounded domain verifying certain conditions. A possible set of assumptions would be, similarly to [MS] and [MOS]:

(a1) G is a bounded domain in $\mathbb{R}^d$, with a smooth boundary ∂G, taken as of class C^2. If $I_{0,T}(\varphi)$ denotes the rate functional

$$I_{0,T}(\varphi) = \begin{cases} \frac{1}{2} \int_0^T |\dot{\varphi}_t - b(\varphi_t)|^2 \, dt & \text{if } \varphi \text{ is absolutely continuous} \\ +\infty & \text{otherwise} \end{cases} \tag{2}$$

defined on the space $C([0, T], \mathbb{R}^d)$ and corresponding to the large deviation principle associated to the family of laws of $(X^{x,\varepsilon})_{0 \leqslant t \leqslant T}$ on this space, $V(x, y)$ is the associated quasi potential of Freidlin and Wentzell:

$$V(x, y) = \inf_{\substack{\varphi : \varphi(0) = x, \varphi(T) = y \\ T > 0}} \{I_{0,T}(\varphi)\} \tag{3}$$

and one considers the equivalence relation

$$x \sim y \quad \text{iff} \quad V(x, y) = V(y, x) = 0 \tag{4}$$

then one assumes:

(a2) There are finitely many compact sets $K_1, \ldots, K_m$, equivalence classes for $\sim$, and such that:

(i) each ω-limit set of the deterministic system given by $\dot{x}(t) = b(x(t))$ is contained in some K_i.

(ii) The stable classes are $K_1, \ldots, K_\ell$ ($\ell < m$) and each of them consists of a fixed point of the deterministic system. These are denoted by x_i, $i = 1, \ldots, \ell$ and we assume that $\{x_1, \ldots, x_\ell\} \cap \partial G = \varnothing$. Here the notion of a "stable" class is that coming from Freidlin and Wentzell theory:

Definition 1. An equivalence class K is said to be stable if $V(x, y) > 0$ for all $x \in K$, all $y \notin K$.

We know that $V(x, y)$ is constant for all $x \in K_i$, all $y \in K_j$.

Let $V_{i,j}$ denote this constant, so that K_i is stable iff

$$\inf_{j \neq i} V_{i,j} > 0$$

Let $1 \leqslant k \leqslant \ell$ be such that $\{x_1,..., x_\ell\} \cap G = \{x_1,..., x_k\}$ and let $\delta > 0$ such that $B_\delta(x_i)$ is the closed euclidean ball with center x_i and radius σ, is contained in the basin of attraction of x_i as well as in G, for $i = 1, ..., k$. Moreover, let us assume that σ is such that all positive orbits of the deterministic system starting in $B(x_i)$ do not leave G. Let

$$D_i = B_\delta(x_i) \qquad D = \bigcup_{i=1}^{k} D_i$$

We assume further

(a3) Among $x_1,..., x_k$ at least one of them is a hyperbolic fixed point, i.e., there exists $i_0 \in \{1,..., k\}$ such all the eigenvalues of the Jacobian matrix $(\partial b_r/\partial x^s)_{r,s}|_{x=x_{i_0}}$ have negative real part.

The last assumption concerns the "cycle" property:

(a4) Let $V = \max_{i,j \leqslant k} V_{i,j}$ and $V_G = \min_{1 \leqslant i \leqslant k} \min_{y \in \partial G} V(x_i, y)$. We assume that $V_G > V$.

We may now state

Theorem 1. Under above assumptions, and if we define β_ε through the relation

$$\sup_{x \in D} P_x(\tau^\varepsilon > \beta_\varepsilon) = e^{-1} \tag{5}$$

then:

(i) $\lim_{\varepsilon \to 0} P_x(\tau^\varepsilon > t\beta_\varepsilon) = e^{-t}$, for each $x \in D$, each $t > 0$

(ii) $\lim_{\varepsilon \to 0} E_x(\tau^\varepsilon)/\beta_\varepsilon = 1$, $\forall x \in D$.

Remark 1. If G is confining, i.e., $\langle b(x), n(x) \rangle < 0$ for each $x \in \partial G$, where $n(x)$ indicates the outward unit normal vector to ∂G, at the point x, then we may take any $x \in G$ in (i) and (ii) of Theorem 1. (G is open)

Remark 2. Contrarily to what happens in the case of a domain contained in the basin of attraction of a single fixed point or a periodic orbit, we do not always have asymptotic equivalence (even logarithmically) between a quantile of the distribution of τ^ε under P_x $(x \in D)$ and $E_x \tau^\varepsilon$. For a counterexample see, e.g., [FW], p. 197.

Nevertheless, if β_ε is defined through Eq. (5), as observed in [MOS], the bound

$$\sup_{x \in D} \frac{E_x \tau^\varepsilon}{\beta_\varepsilon} \leqslant C < +\infty \tag{6}$$

for some finite constant C, holds independently of (i) of Theorem 1.

Moreover, from Eq. (6) and the known results of Freidlin and Wentzell on the asymptotic behaviour of $\varepsilon^2 \log E_x \tau_\varepsilon$, we get

$$\varliminf_{\varepsilon \to 0} \varepsilon^2 \log \beta_\varepsilon \geqslant V_G \tag{7a}$$

On the other side, and this is the reason for the name "cycle," if $x \in D_i$ one has for any $h > 0$

$$\lim_{\varepsilon \to 0} P_x(\tau_\varepsilon(D_j) \leqslant e^{(V+h)/\varepsilon^2}) = 1 \tag{7b}$$

using the notation $\tau_\varepsilon(A)$ to denote the first hitting time of the set A.

For convenience of the reader let us recall the verification of Eq. (6), as in [MOS]:

$$\begin{aligned} \frac{E_x \tau^\varepsilon}{\beta_\varepsilon} &= \frac{1}{\beta_\varepsilon} \int_0^{+\infty} P_x(\tau^\varepsilon > t)\, dt \\ &= \int_0^{+\infty} P_x(\tau^\varepsilon > t\beta_\varepsilon)\, dt \\ &\leqslant \int_0^{+\infty} g_\varepsilon(t)\, dt \end{aligned} \tag{8}$$

where $g_\varepsilon(t) =^{\text{DEF}} \sup_{x \in G} P_x(\tau^\varepsilon > t\beta_\varepsilon)$. But the Markov property implies that

$$g_\varepsilon(t+s) \leqslant g_\varepsilon(t)\, g_\varepsilon(s)$$

so that

$$g_\varepsilon(2k) \leqslant (g_\varepsilon(2))^k$$

As in [MOS] we can see that $g_\varepsilon(2) \leqslant r < 1$ for ε small, and so we get (6).

In fact,

$$g_\varepsilon(2) \leqslant \sup_{x \in G} P_x(\tau_\varepsilon(D \cup G^c) > \beta_\varepsilon) + \sup_{x \in D} P_x(\tau^\varepsilon > \beta_\varepsilon) \tag{9}$$

the second term on the r.h.s. of Eq. (9) is e^{-1}, and using Freidlin and Wentzell estimates we see that the first term goes to zero, so that we get the claimed upper bound.

Remark 3. The argument just described allows also to make use of the Dominated Convergence Theorem and, from Eq. (8), to get (ii) of Theorem 1, once part (i) is proved.

Moreover, and as in [GOV], for the proof of part (i) in Theorem 1 in the case $x = x_{i_0}$, it suffices to prove the following:

Lemma 1. Under the assumptions of Theorem 1, with $\beta_\varepsilon > 0$ given by Eq. (5) and letting

$$f_\varepsilon(t) = P_{x_{i_0}}(\tau^\varepsilon > t\beta_\varepsilon)$$

for $t > 0$, $\varepsilon > 0$, then there exist positive numbers δ_ε which tend to zero as $\varepsilon \to 0$, and such that for each $s, t > 0$:

$$f_\varepsilon(s + \delta_\varepsilon)\, f_\varepsilon(t + \delta_\varepsilon) - o_t(1) \leqslant f_\varepsilon(t + s) \leqslant f_\varepsilon(s)\, f_\varepsilon(t - \delta_\varepsilon) + o_t(1) \qquad (10)$$

where $o_t(1)$ is a function of t and ε, which tends to zero as $\varepsilon \to 0$, uniformly on $t \geqslant t_0$, for any given $t_0 > 0$.

The Proof of Lemma 1, as presented below, is similar to that of Lemma 4 in [GOV] with assumption (a4) and the Freidlin and Wentzell theory being used to control the time needed to arrive to a suitably small neighborhood of x_{i_0}, and using the coupling method proposed by Lindvall and Rogers (cf. [LR], Sections 2 and 3) to ensure the loss of memory.

For this, let B denote the Jacobian matrix of b at x_{i_0}. It is well known that, under assumption (a3) there exists a unique symmetric positive definite matrix L such that

$$B^T L + LB = -\mathbb{I} \qquad (11)$$

(We are using ^{T} for transposition, and $\mathbb{I}$ to denote the identity $d \times d$ matrix.)

Define next the norm ρ in $\mathbb{R}^d$, by

$$\rho(x) = (\langle x, Lx \rangle)^{1/2} \qquad (12)$$

Recall that $b(\cdot)$ is assumed of class C^1, which implies, using (11) and standard facts, that,

$$\begin{aligned}\langle b(x) - b(x'), L(x - x') \rangle &= \langle A(x, x')(x - x'), L(x - x') \rangle \\ &\leqslant -\tfrac{1}{4}\, |x - x'|^2 \leqslant 0 \end{aligned} \qquad (13)$$

if $x, x' \in \mathscr{B}_{\delta_1}(x_{i_0})$, for $0 < \delta_1 < \delta$ small enough, and where $A(x, x')$ is a $d \times d$ matrix with entries $\partial_s b_r(y_{r,s})$ for suitable $y_{r,s}$ with $|y_{r,s} - x_{i_0}| \leqslant C\delta_1$, C being a positive constant (depending only on the dimension).

The coupling proposed in [LR] may thus be used to replace the analytical results of [Day] used in [GOV], or the exponential joining proposed by [MS], and used in [MOS], and allows us to write the following

Lemma 2. There exists $s_0 > 0$ so that if $x \in B_{\delta_0}(x_{i_0})$ then

$$P_x(\tau^\varepsilon > t\beta_\varepsilon) - P_{x_{i_0}}(\tau^\varepsilon > t\beta_\varepsilon) \to 0 \qquad \text{as} \quad \varepsilon \to 0 \tag{14}$$

uniformly on $t \geqslant t_0$, for any given $t_0 > 0$.

Proof. Let $\sigma > 0$ be chosen so that (13) holds and let us now choose $\sigma_0 \in (0, \sigma_1)$ so that all positive orbits of the deterministic system issued from some point in $B(x_{i_0})$ converge to x_{i_0} without leaving $B\sigma_1(x_{i_0})$. For the proof of (14) it suffices to present a coupling of the two processes $X_t^{x,\varepsilon}$ and $X_t^{x_{i_0},\varepsilon}$ in such a way that with probability tending to one they will meet before leaving $B_{2s_1}(x_{t_0})$, and this in time of order shorter than β_ε.

In order to do so, we consider the coupling proposed by Lindvall and Rogers (Sections 2 and 3 of [LR]), which is particularly simple in the case of constant diffusion coefficient (Example 5 of [LR]). The processes $X_t^{x,\varepsilon}$ and $X_t^{x_{i_0},\varepsilon}$ are constructed using the same noise, as follows: take $X_t^{x,\varepsilon}$ and $X_t^{x_{i_0},\varepsilon}$ as solutions of the Itô equations

$$\begin{aligned} dX_t^{x,\varepsilon} &= b(X_t^{x,\varepsilon})\,dt + \varepsilon\,dW_t; \qquad X_0^{x,\varepsilon} = x \\ dX_t^{x_{i_0},\varepsilon} &= b(X_t^{x_{i_0},\varepsilon})\,dt + \varepsilon H(X_t^{x,\varepsilon}, X_t^{x_{i_0},\varepsilon})\,dW_t; \qquad X_0^{x_{i_0},\varepsilon} = x_{i_0} \end{aligned} \tag{15}$$

where W_t is a standard d-dimensional Brownian motion and $H(x, y)$ is the $d \times d$ orthogonal matrix with determinant -1 given by

$$H(x, y) = \mathbb{I} - 2\frac{(x-y)}{|x-y|}\left[\frac{(x-y)}{|x-y|}\right]^T \tag{16}$$

The geometric idea behind this construction is clear: Consider $x \neq y \in \mathbb{R}^d$. From (16), we have

$$H(x, y)\left(\frac{(x-y)}{|x-y|}\right) = -\left(\frac{x-y}{|x-y|}\right) \tag{17}$$

and, acting on vectors that belong to the plane orthogonal to $x-y$, $H(x, y)$ is just the identity. Thus, $H(x, y)$ is simply the specular reflexion through the plane (by the origin) orthogonal to the vector $x-y$, and has determinant -1. Then, given $z \in \mathbb{R}^d$, $z = x + b$, consider $z' = H(x, y)\, b + y$. Then, z' is the reflexion of z by the plane orthogonal to $x-y$, that passes by the middle point between x and y. In particular, if Z_t is a d-dimensional Brownian motion starting at x, then Z'_t as obtained by the above described reflexion (for each point in the path), is a d-dimensional Brownian motion starting at y.

Then, the processes $X_t^{x,\varepsilon}$ and $X_t^{x_{i_0},\varepsilon}$ are both solutions of our original Itô equation, and if one considers the function $g: \mathbb{R}^{2d} \to \mathbb{R}$, $g(x, y) = \rho(x-y)$ (for ρ defined in (12)), then, Itô's formula (which is valid as long as $g(x-y) > 0$), yields for the one-dimensional process Y_t^ε, given by

$$
\begin{aligned}
&Y_t^\varepsilon = g(X_t^{x,\varepsilon}, X_t^{x_{i_0},\varepsilon}):\\
&dY_t^\varepsilon = \left\langle b(X_t^{x,\varepsilon}) - b(X_t^{x_{i_0},\varepsilon}), \frac{L(X_t^{x,\varepsilon} - X_t^{x_{i_0},\varepsilon})}{\rho(X_t^{x,\varepsilon} - X_t^{x_{i_0},\varepsilon})} \right\rangle dt + \frac{\varepsilon^2}{2} A_t^\varepsilon\, dt\\
&\qquad + \varepsilon \left\langle [\mathbb{I} - H(X_t^{x,\varepsilon}, X_t^{x_{i_0},\varepsilon})] \frac{L(X_t^{x,\varepsilon} - X_t^{x_{i_0},\varepsilon})}{\rho(X_t^{x,\varepsilon} - X_t^{x_{i_0},\varepsilon})}, dW_t \right\rangle\\
&\qquad Y_0^\varepsilon = \rho(x - x_{i_0})
\end{aligned}
\tag{18}
$$

The process A_t^ε above is that coming from the second derivative of g, $D^2 g$ in Itô's formula, and it is given by

$$
\begin{aligned}
A_t^\varepsilon &= tr(D^2 g(X_t^{x,\varepsilon}, X_t^{x_{i_0},\varepsilon}))\\
&\quad - \Bigg[tr(D^2 g(X_t^{x,\varepsilon}, X_t^{x_{i_0},\varepsilon})) + \frac{4}{\rho(X_t^{x,\varepsilon} - X_t^{x_{i_0},\varepsilon})\, |X_t^{x,\varepsilon} - X_t^{x_{i_0},\varepsilon}|^2}\\
&\quad \times \Bigg(\frac{\langle X_t^{x,\varepsilon} - X_t^{x_{i_0},\varepsilon}, L(X_t^{x,\varepsilon} - X_t^{x_{i_0},\varepsilon}) \rangle^2}{\rho^2(X_t^{x,\varepsilon} - X_t^{x_{i_0},\varepsilon})}\\
&\quad - \langle X_t^{x,\varepsilon} - X_t^{x_{i_0},\varepsilon}, L(X_t^{x,\varepsilon} - X_t^{x_{i_0},\varepsilon}) \rangle \Bigg) \Bigg]
\end{aligned}
$$

Since A_t^ε results to be zero in our case, it follows from (13) that the drift part in (18) is less or equal than zero, as long as $X_t^{x,\varepsilon}$ and $X_t^{x_{i_0},\varepsilon}$ remain in $\mathscr{B}_{\delta_1}(x_{i_0})$. From (16), the martingale term on the r.h.s. of Eq. (18) is

$$dM_t^\varepsilon = \varepsilon \left\langle [\mathbb{I} - H(X_t^{x,\varepsilon}, X_t^{x_{i_0},\varepsilon})] \frac{L(X_t^{x,\varepsilon} - X_t^{x_{i_0},\varepsilon})}{\rho(X_t^{x,\varepsilon} - X_t^{x_{i_0},\varepsilon})}, dW_t \right\rangle$$

$$= 2\varepsilon \left\langle \frac{\rho(X_t^{x,\varepsilon} - X_t^{x_{i_0},\varepsilon})}{|X_t^{x,\varepsilon} - X_t^{x_{i_0},\varepsilon}|} \frac{(X_t^{x,\varepsilon} - X_t^{x_{i_0},\varepsilon})}{|X_t^{x,\varepsilon} - X_t^{x_{i_0},\varepsilon}|}, dW_t \right\rangle$$

which implies that the process Y_t^ε satisfies

$$Y_t^\varepsilon = Y_0 + \int_0^t \left\langle b(X_s^{x,\varepsilon}) - b(X_s^{x_{i_0},\varepsilon}), \frac{L(X_s^{x,\varepsilon} - X_s^{x_{i_0},\varepsilon})}{\rho(X_s^{x,\varepsilon} - X_s^{x_{i_0},\varepsilon})} \right\rangle ds$$

$$+ 2\varepsilon \int_0^t \frac{\rho(X_s^{x,\varepsilon} - X_s^{x_{i_0},\varepsilon})}{|X_s^{x,\varepsilon} - X_s^{x_{i_0},\varepsilon}|} dB_s$$

for B_t a standard one dimensional Brownian motion. Next, let S^ε be the coupling time, $T^\varepsilon(y)$ the exit time from $B_{2s_1}(x_{i_0})$ of the solution $X_t^{y,\varepsilon}$ and $\tilde{S}^\varepsilon$ the time it takes for $Y_0^\varepsilon + 2\varepsilon \int_0^t (\rho(X_s^{x,\varepsilon} - X_s^{x_{i_0},\varepsilon})/|X_s^{x,\varepsilon} - X_s^{x_{i_0},\varepsilon}|)\, dB_s$ to hit zero:

$$S^\varepsilon = \inf\{t \geqslant 0 : |Y_t^\varepsilon| = 0\}$$

$$T^\varepsilon(y) = \inf\{t \geqslant 0 : X_t^{y,\varepsilon} \notin B_{2s_1}(x_{i_0})\}$$

$$\tilde{S}^\varepsilon = \inf\left\{t \geqslant 0 : Y_0^\varepsilon + 2\varepsilon \int_0^t \frac{\rho(X_s^{x,\varepsilon} - X_s^{x_{i_0},\varepsilon})}{|X_s^{x,\varepsilon} - X_s^{x_{i_0},\varepsilon}|} dB_s = 0\right\}$$

Now, from (13) and the above remarks,

$$\begin{aligned} P(S^\varepsilon < \varepsilon^{-3}) &\geqslant P(S^\varepsilon < \varepsilon^{-3}, T^\varepsilon(x) \wedge T^\varepsilon(x_{i_0}) > \varepsilon^{-4}) \\ &\geqslant P(\tilde{S}^\varepsilon < \varepsilon^{-3}, T^\varepsilon(x) \wedge T^\varepsilon(x_{i_0}) > \varepsilon^{-4}) \\ &\geqslant P(\tilde{S}^\varepsilon < \varepsilon^{-3}) - P(T^\varepsilon(x) \wedge T^\varepsilon(x_{i_0}) \leqslant \varepsilon^{-4}) \end{aligned} \tag{19}$$

where we denoted by $t \wedge s$ the minimum between t and s. But, from the Freidlin and Wentzell theory we know that there exists $a, c > 0$ so that

$$P(T^\varepsilon(x) \wedge T^\varepsilon(x_{i_0}) \leqslant \varepsilon^{-4}) \leqslant 2 \sup_{y \in B_{\delta_0}(x_{i_0})} P(T^\varepsilon(y) \leqslant \exp\{a/\varepsilon^2\}) \leqslant \exp\{c/\varepsilon^2\} \tag{20}$$

For the other term, we have that the continuous martingale $M_t^\varepsilon = 2\varepsilon \int_0^t (\rho(X_s^{x,\varepsilon} - X_s^{x_{i_0},\varepsilon})/|X_s^{x,\varepsilon} - X_s^{x_{i_0},\varepsilon}|)\, dB_s$ has compensator $\langle M^\varepsilon \rangle_t = 4\varepsilon^2 \int_0^t (\rho^2(X_s^{x,\varepsilon} - X_s^{x_{i_0},\varepsilon})/|X_s^{x,\varepsilon} - X_s^{x_{i_0},\varepsilon}|^2)\, ds \geqslant \varepsilon^2 k_1 t$ for some constant k_1, by the equivalence of the norms in $\mathbb{R}^d$. Thus we may use a classical result on time-change for martingales (see [KS], p. 174) which allows us to write

$M_t^\varepsilon = \tilde{B}_{\langle M^\varepsilon \rangle_t}$, where $\tilde{B}$ is a standard one dimensional Brownian motion ($\tilde{B}_t = M^\varepsilon_{C_t^\varepsilon}$ where the process C_t^ε is the inverse of the compensator, cf. Theorem 4.6, Ch. 3 in [KS]). Thus

$$P(\tilde{S}^\varepsilon < \varepsilon^{-3}) = P(\inf_{s<\varepsilon^{-3}} M_s^\varepsilon \leqslant -Y_0^\varepsilon) \geqslant P(\inf_{s\leqslant k_1\varepsilon^{-1}} \tilde{B}_s \leqslant -Y_0^\varepsilon)$$
$$= 1 - \int \mathbf{1}_{\{|x| \leqslant Y_0^\varepsilon \varepsilon^{1/2}/k_1^{1/2}\}} \frac{e^{-x^2/2}}{\sqrt{2\pi}} \geqslant 1 - \frac{k_2 \delta_0 \varepsilon^{1/2}}{k_1^{1/2}} \tag{21}$$

where k_2 is a positive constant so that $\rho(v) \leqslant k_2 |v|$, for all $v \in \mathbb{R}^d$. From (19), (20) and (21), it follows that $P(S^\varepsilon < \varepsilon^{-3}) \to 1$ as $\varepsilon \to 0$, which implies Lemma 2 from (7a).

Proof of Lemma 1. As in [GOV], the point is to show the existence of $\eta_\varepsilon > 0$ such that $\eta_\varepsilon/\beta_\varepsilon \to 0$ and such that

$$\lim_{\varepsilon\to 0} \sup_{x\in G} P_x(\tau^\varepsilon > \eta_\varepsilon, \tau_\varepsilon(B_{\delta_0}(x_{i_0})) > \eta_\varepsilon) = 0 \tag{22}$$

To verify (22) let us take

$$V < \alpha < V_G$$

and let $\eta_\varepsilon = e^{\alpha/\varepsilon^2}$.

Since $\alpha > 0$, $G \backslash D$ is bounded, and all stable classes in $G \cup \partial G$ are contained in D, from Freidlin and Wentzell theory we know that

$$\lim_{\varepsilon\to 0} \sup_{x\in G\backslash D} P_x(\tau_\varepsilon(G^c \cup D) > \eta_\varepsilon) = 0$$

On the other side, by assumption (a4) and since $V < \alpha$. Freidlin and Wentzell theory implies that

$$\sup_{x\in D} P_x(\tau_\varepsilon(B_{\delta_0}(x_{i_0})) > \eta_\varepsilon/2) \to 0$$

Thus we get:

$$\sup_{x\in G} P_x(\tau_\varepsilon > \eta_\varepsilon, \tau_\varepsilon(B_{\delta_0}(x_{i_0})) > \eta_\varepsilon)$$
$$\leqslant \sup_{x\in G} P_x(\tau_\varepsilon > \eta_\varepsilon, \tau_\varepsilon(D) > \eta_\varepsilon/2) + \sup_{x\in D} P_x(\tau_\varepsilon(B_{\delta_0}(x_{i_0})) > \eta_\varepsilon/2)$$

which both tend to zero, yielding Eq. (22).

To complete the proof of Lemma 1, we proceed as in [GOV]: Let $s > 0$ and

$$R^s = \inf\{u > s\beta_\varepsilon : X_u^\varepsilon \in B_{\delta_0}(x_{i_0})\}$$

then

$$\sup_{x\in G} P_x(\tau^\varepsilon > s\beta_\varepsilon + \eta_\varepsilon,\ R^s > s\beta_\varepsilon + \eta_\varepsilon) \leqslant \sup_{x\in G} P_x(\tau^\varepsilon > \eta_\varepsilon,\ \tau_\varepsilon(B_{\delta_0}(x_{i_0})) > \eta_\varepsilon)$$

which tends to zero as $\varepsilon \to 0$.

Due to Eq. (6) (cf. Remark 2) and the choice of α, $\eta_\varepsilon/\beta_\varepsilon \to 0$, and we have that uniformly on $(s, t) \in [0, +\infty) \times [t_0, +\infty)$ for any given $t_0 > 0$:

$$\sup_{x\in G} P_x(\tau^\varepsilon > (s+t)\,\beta_\varepsilon,\ R^s \geqslant s\beta_\varepsilon + \eta_\varepsilon) \to 0 \tag{23}$$

But, as in Eq. (2.16) of [GOV]:

$$\begin{aligned} &P_{x_{i_0}}(\tau^\varepsilon > (s+t)\,\beta_\varepsilon,\ R^s \leqslant s\beta_\varepsilon + \eta_\varepsilon) \\ &\qquad \leqslant P_{x_{i_0}}(\tau^\varepsilon > s\beta_\varepsilon) \sup_{y\in B_{\delta_0}(x_{i_0})} P_y(\tau^\varepsilon > t\beta_\varepsilon - \eta_\varepsilon) \end{aligned} \tag{24}$$

and

$$\begin{aligned} &P_{x_{i_0}}(\tau^\varepsilon > (s+t)\,\beta_\varepsilon,\ R^s \leqslant s\beta_\varepsilon + \eta_\varepsilon) \\ &\qquad \geqslant P_{x_{i_0}}(\tau^\varepsilon > s\beta_\varepsilon + \eta_\varepsilon,\ R^s \leqslant s\beta_\varepsilon + \eta_\varepsilon) \inf_{y\in B_{\delta_0}(x_{i_0})} P_y(\tau^\varepsilon > t\beta_\varepsilon) \end{aligned} \tag{25}$$

so that Lemma 1 follows easily from Eqs. (23)–(25), and Lemma 2.

Proof of Theorem 1. As already noticed it suffices us to prove part (i).

Also if $x = x_{i_0}$, (i) follows at once from Lemma 1. Using Lemma 2 we extend to any $x \in B_{\delta_0}(x_{i_0})$. To conclude we need to recall, as in Eq. (7b) that if $V < \alpha < V_G$ and $\eta_\varepsilon = e^{\alpha/\varepsilon^2}$ then

$$P_x(\tau_\varepsilon(B_{\delta_0}(x_{i_0})) \leqslant \eta_\varepsilon) \to 1.$$

Using the strong Markov property at $\tau_\varepsilon(B_{\delta_0}(x_{i_0}))$ we then conclude the proof as before.

Remark 4. The assumption of constant diffusion coefficient and $b(\cdot)$ satisfying Eq. (13) makes the coupling time of the two processes $X^{x,\varepsilon}$, $X^{x_{i_0},\varepsilon}$—if we use the coupling designed in [LR]—particularly easy to evaluate, and directly comparable with $\tilde{S}$ where $\tilde{S}$ is the time for a one dimensional Brownian motion starting at some point $r = |x-y|/2\varepsilon$ to reach the origin.

On the other hand if $\sigma(\cdot)$ is not constant, one needs to examine condition (23) of [LR] to verify if coupling occurs. Since we not only want to see the finiteness of the coupling time, but also its ε-dependence, we need to make a further comparison, and we do not enter this.

Remark 5. Calculations become simpler in the particular case satisfying condition $\langle b(x)-b(x'), x-x'\rangle \leqslant 0$ for x, x' in some neighborhood of the fixed point x_{i_0}. This would be the case, e.g., in the situation of $b=-\nabla a$ and x_{i_0} being a hyperbolic local minimum. Without assuming hyperbolicity it will also work if the potential a is convex in a neighborhood of its local minimum x_{i_0}. Then, of course we may work with L as the identity and the martingale M^ε becomes a Brownian motion. (This example covers the cases treated in [GOV]. Cf. also Example 5 in [LR].)

ACKNOWLEDGMENT

The work of M.E.V. was partially supported by CNP_q FINEP (Pronex project).

REFERENCES

[B] S. Brassesco, Unpredictability of an exit time, *Stoch. Proc. and Their Applic.* **63**:55–65 (1996).

[CGOV] M. Cassandro, A. Galves, E. Olivieri, and M. E. Vares, Metastable behavior of stochastic dynamics: A pathwise approach, *J. Stat. Phys.* **35**(5/6):603–634 (1984).

[D] M. Day, On the exponential exit law in the small parameter exit problem, *Stochastics* **8**:297 (1989).

[FW] M. Freidlin and A. P. Wentzell, *Random Perturbations of Dynamical Systems* (Springer-Verlag, 1984).

[GOV] A. Galves, E. Olivieri, and M. E. Vares, Metastability for a class of dynamical systems subject to small random perturbations, *Ann. Prob.* **15**:1288–1305 (1987).

[KS] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*, 2nd ed. (Springer-Verlag, 1991).

[LR] T. Lindvall and L. C. G. Rogers, Couplings of multidimensional diffusions by reflexion, *Ann. Prob.* **14**:860–872 (1986).

[M] C. Mueller, Coupling and invariant measures for the heat equation with noise, *Ann. Prob.* **21**:2189–2198 (1993).

[MOS] F. Martinelli, E. Olivieri, and E. Scoppola, Small random perturbations of finite and infinite dimensional dynamical systems: unpredictability of exit times, *J. Stat. Phys.* **55**:478–503 (1989).

[MS] F. Martinelli and E. Scoppola, Small random perturbations of dynamical systems: exponential loss of memory of the initial condition, *CMP* **120**:25–69 (1988).

Communicated by P. Picco