The Annals of Probability
1987, Vol. 15, No. 4, 1288–1305

METASTABILITY FOR A CLASS OF DYNAMICAL SYSTEMS SUBJECT TO SMALL RANDOM PERTURBATIONS

BY ANTONIO GALVES,[1] ENZO OLIVIERI AND MARIA EULALIA VARES[2]

Universidade de São Paulo, Università di Roma, La Sapienza and Instituto de Matemática Pura e Aplicada

We consider dynamical systems in $\mathbb{R}^d$ driven by a vector field $b(x) = -\nabla a(x)$, where a is a double-well potential with some smoothness conditions. We show that these dynamical systems when subjected to a small random disturbance exhibit metastable behavior in the sense defined in [2]. More precisely, we prove that the process of moving averages along a path of such a system converges in law when properly normalized to a jump Markov process. The main tool for our analysis is the theory of Freidlin and Wentzell [7].

0. Introduction. One of the challenging problems in the study of thermodynamical systems is to understand and to model in a rigorous way the phenomenon of metastability. The authors introduced in [2] a new (pathwise) approach to describe metastability in stochastic processes. The basic difference from other approaches (e.g., [8])—mostly based on the "evolution of ensembles"—is that it focuses on the statistics of each typical path.

The situation under consideration can be outlined as follows: A stochastic process with a unique stationary probability measure, which, for suitable initial conditions, behaves for a very long time as if it were described by another "stationary" measure (metastable state), performing, at the end, an abrupt transition to the correct equilibrium. In order to detect this behavior, it is suggested in [2] to look at time averages along typical trajectories; we should see: apparent stability–sharp transition–stability. To make this precise we consider a parametrized family of similar systems, and the statement will be made sharp if, by properly rescaling, the (measure-valued) processes of suitable time averages converge in law to a (measure-valued) Markov jump process. The Markov property accounts, in some sense, for the unpredictable character of the transition.

This behavior contrasts with the hydrodynamical one, whose main feature is a smooth evolution (in a certain space and time scale) through a continuum of "equilibrium states." We refer to [5] for a survey on this subject and to [2] for more general comments on the present approach to metastability.

In [2], the reader will also find two examples of stochastic processes showing metastable behavior: The first is a simple one-dimensional model, the so-called

Received November 1984; revised September 1986.

[1] Partially supported by CNPq grants 301301/79 and 20.0487/84.
[2] Partially supported by CNPq grant 401299/84.
AMS 1980 *subject classifications*. 60H10, 60J05.
Key words and phrases. Metastability, large deviations, dynamical systems.

Curie–Weiss model, and the other is a system of interacting particles on $\mathbb{Z}$, the basic contact process of Harris. This last example presents particularly interesting features, giving rise to some open questions on how the transition happens (i.e., spatial patterns related to the transition). Some extensions of [2] concerning the contact process have been obtained by Schonmann [12].

Here we prove that a certain class of dynamical systems in $\mathbb{R}^d$, when subjected to a small random disturbance (to be specified later), exhibits metastable behavior in the sense just described. (The parameter will be the "strength" of the perturbation.) The class includes systems driven by a vector field $b(x) = -\nabla a(x)$, where $a(\cdot)$ is a double-well potential (with some smoothness conditions). The assumption of deriving from a potential can be somewhat relaxed; due to the notion of "quasipotential" developed by Freidlin and Wentzell in their basic work [13], it is possible to extend the results to some nongradient cases, provided some restrictions are added.

This class of systems constitutes the simplest nontrivial example of processes in $\mathbb{R}^d$ exhibiting metastable behavior. In this case the results are quite intuitive and appealing (especially after [13, 14]). Moreover, the proofs involve interesting problems. For instance, the study of the law of the exit time from a domain containing both stable and unstable critical points is not covered by known results in this field [3, 4] and it is interesting by itself.

This material will be organized as follows. Section 1 describes the model and states the main results, besides establishing the basic notation. Section 2 provides the basic ingredients for the proofs, and contains the proof of Theorem 1 on the "exponentiality" of the jump time. In Section 3 we study the stability of time averages and prove Theorem 2.

1. Description of the model. Statement of the results. Throughout this article we shall be working with a family of diffusion processes, which is obtained from a dynamical system $X_0^x(t)$ on $\mathbb{R}^d$,

$$
\begin{aligned}
dX_0^x(t) &= b\big(X_0^x(t)\big)\,dt, \qquad t > 0, \\
X_0^x(0) &= x,
\end{aligned}
\tag{1.1}
$$

through an additive random noise. More specifically, we consider the family of diffusion processes $X_\varepsilon^x(\cdot)$, given by the stochastic Itô equation

$$
X_\varepsilon^x(t) = x + \int_0^t b\big(X_\varepsilon^x(s)\big)\,ds + \varepsilon B_t, \qquad t \ge 0,
\tag{1.2}
$$

where $(B_t)_{t \ge 0}$ is a standard d-dimensional Brownian motion on some complete probability space $(\Omega, \mathscr{F}, P)$; x denotes the initial position and ε is a positive real number.

As mentioned in the introduction we shall be particularly interested in the case of a gradient vector field which derives from a double-well potential. More precisely, let us assume

$$
b(x) = -\nabla a(x), \qquad x \in \mathbb{R}^d,
\tag{1.3}
$$

where the following conditions are satisfied:

(i) $a\colon \mathbb{R}^d \to \mathbb{R}$ is a C^2 function;
(ii) $a(x) \to +\infty$, as $|x| \to +\infty$;
(iii) $a(\cdot)$ has exactly three critical points, denoted by p, q and r, with $\det((\partial^2 a/\partial x_i\, \partial x_j)(x)) \neq 0$, for $x = p, q, r$, with p and q being points of stable equilibrium for (1.1), and r being a saddle point;
(iv) $a(q) < a(p) < a(r)$.

Let us also assume

(v) $|b(x)|^2 \leq K(1 + |x|^2)$, $|b(x) - b(y)|^2 \leq K|x - y|^2$, for all $x, y \in \mathbb{R}^d$, where K is some positive constant.

(See Figure 1.)

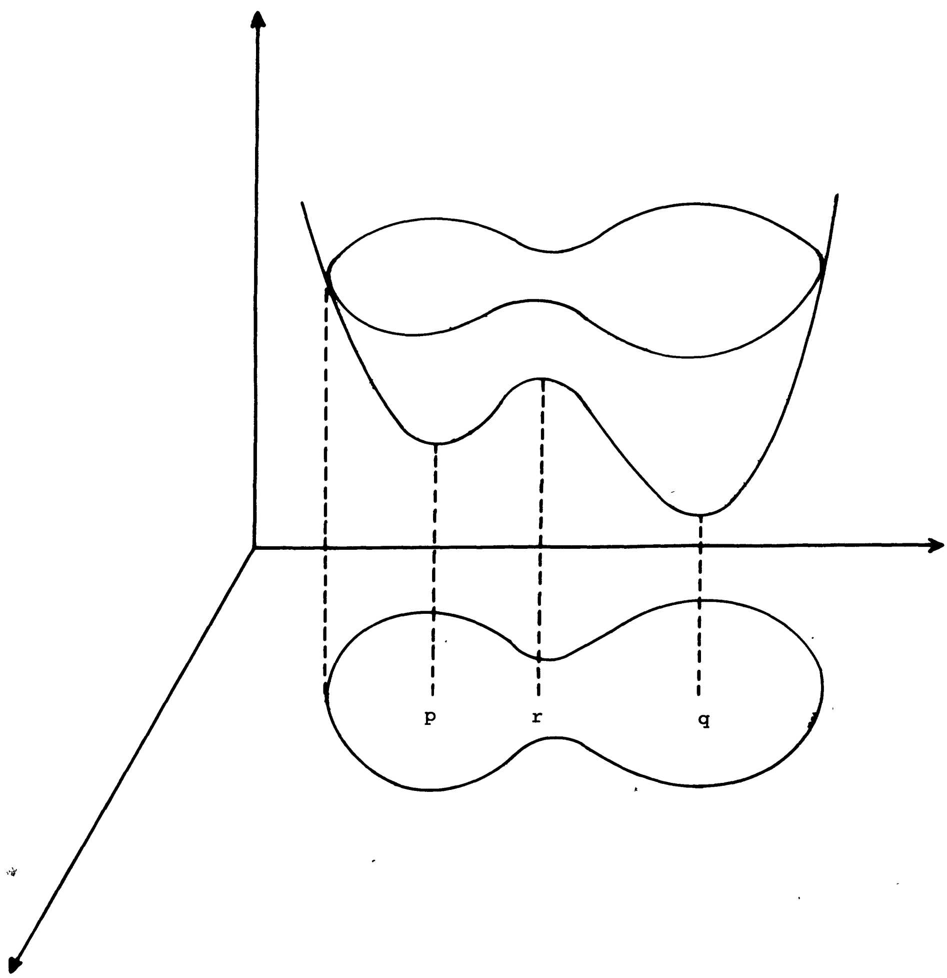

FIG. 1.

REMARKS.

(a) Condition (v) guarantees strong uniqueness of the solution for (1.2). It can be relaxed if we treat (1.2) in the weak sense, with uniqueness in law (via martingale problems). For us this will be irrelevant since the other conditions will allow us to restrict the study to $X^x(\cdot)$ before it leaves a certain compact set. In this case (i) automatically guarantees strong uniqueness of the stopped process.

(b) From (ii) and (iii) we know also that $A \stackrel{\text{def}}{=} ((\partial^2 a/\partial x_i\,\partial x_j)(r))$ has only one negative eigenvalue with a one-dimensional eigenspace. Thus, W_r^u is a one-dimensional manifold (cf. [10]). [Notation: as usual, we let W_r^u (W_r^s) denote the unstable (stable, respectively) manifold at r, for the system given by (1.1).]

Some notation.

1. $B_c(x)$ denote the closed ball centered at x, $x \in \mathbb{R}^d$, with radius $c > 0$ (for the Euclidean norm).
2. Let D_p (and D_q, respectively) denote the basin of attraction of p (and q, respectively) for the deterministic system $X_0^x(\cdot)$.
3. $(\mathscr{F}_t)_{t\geq 0}$ denotes the natural filtration on $(\Omega, \mathscr{F})$ associated to the given Brownian motion, i.e., $\mathscr{F}_t$ is the σ-field generated by $(B_s\colon s \leq t)$.
4. $C_b(\mathbb{R}^d)$ denotes the space of bounded continuous real functions on $\mathbb{R}^d$, with the supremum norm.
5. $\mathscr{M}_1$ denotes the space of probability measures on $\mathscr{B}(\mathbb{R}^d)$. We consider $\mathscr{M}_1$ with the weak*-topology induced on it as a subset of the dual of $C_b(\mathbb{R}^d)$.
6. $D([0, +\infty), \mathscr{M}_1)$ denotes the set of functions from $[0, +\infty)$ to $\mathscr{M}_1$, which are right continuous with left limits. Since $\mathscr{M}_1$ is a separable metric space we have no problem to define the Skorokhod topology on $D([0, +\infty), \mathscr{M}_1)$ (see [1, 11]).
7. Sometimes we shall write simply X_ε instead of X_ε^x, and then use P_x to denote the law of the process when the initial condition is x. Similarly, for T_ε^x and other objects defined below.
8. δ_x denotes the Dirac measure at x.
9. $\mathscr{C}([0, T], \mathbb{R}^d)$ denotes the space of continuous functions $f\colon [0, T] \to \mathbb{R}^d$; for any f, g belonging to $\mathscr{C}([0, T], \mathbb{R}^d)$,

$$\rho_T(f, g) = \sup_{0\leq t\leq T} |f(t) - g(t)|.$$

DEFINITION. Let us fix $c > 0$ so that $B_c(p) \subseteq D_p$ and $B_c(q) \subseteq D_q$, and let us define the $\mathscr{F}_t$-stopping times:

$$(1.4)\qquad \begin{aligned} T_\varepsilon^x &= \inf\{t > 0\colon X_\varepsilon^x(t) \in B_c(q)\},\\ \tilde{T}_\varepsilon^x &= \inf\{t > T_\varepsilon^x\colon X_\varepsilon^x(t) \in B_c(p)\}; \end{aligned}$$

the infimum being defined as $+\infty$ when the set is empty.

We may now state our main results.

THEOREM 1. *For $\varepsilon > 0$ let β_ε be defined through the relation $P_p(T_\varepsilon > \beta_\varepsilon) = e^{-1}$. Then, for all $x \in D_p$,*

$$P_x(T_\varepsilon/\beta_\varepsilon > t) \underset{\varepsilon \downarrow 0}{\rightarrow} e^{-t}, \quad \text{for all } t \geq 0. \tag{1.5}$$

THEOREM 2. *Let β_ε be as in Theorem 1. It is possible to find $R_\varepsilon > 0$ with $R_\varepsilon \to +\infty$ and $R_\varepsilon/\beta_\varepsilon \to 0$ as $\varepsilon \downarrow 0$, so that if we define the $\mathscr{M}_1$-valued processes $(\nu_t^\varepsilon)_{t \geq 0}$ via*

$$\nu_t^\varepsilon(f) = \frac{1}{R_\varepsilon} \int_{t\beta_\varepsilon}^{t\beta_\varepsilon + R_\varepsilon} f(X_\varepsilon(s))\, ds, \qquad f \in C_b(\mathbb{R}^d), \tag{1.6}$$

then, for each $x \in D_p$,

$$P_x\Big(\sup_{0 \leq s < (T_\varepsilon - 3R_\varepsilon)/\beta_\varepsilon} |\nu_s^\varepsilon(f) - f(p)| > \delta \Big) \to 0, \tag{1.7}$$

$$P_x\Big(\sup_{T_\varepsilon/\beta_\varepsilon \leq s < (\tilde{T}_\varepsilon - 3R_\varepsilon)/\beta_\varepsilon} |\nu_s^\varepsilon(f) - f(q)| > \delta \Big) \to 0, \tag{1.8}$$

as $\varepsilon \downarrow 0$ for each $\delta > 0$ and each $f \in C_b(\mathbb{R}^d)$. Finally, let

$$\begin{aligned} \tilde{\nu}_t^\varepsilon &= \nu_t^\varepsilon, \quad \text{if } t \notin \big[(T_\varepsilon - 3R_\varepsilon)/\beta_\varepsilon, T_\varepsilon/\beta_\varepsilon\big] \\ &= \nu^\varepsilon_{(T_\varepsilon - 3R_\varepsilon)/\beta_\varepsilon}, \quad \text{otherwise.} \end{aligned}$$

Then, for each $x \in D_p$, $(\tilde{\nu}_t^\varepsilon)_{t \geq 0}$ converges in law on $D([0, +\infty), \mathscr{M}_1)$ to a Markov jump process $(\nu_t)_{t \geq 0}$ given by

$$\begin{aligned} \nu_t &= \delta_p, \quad \text{if } t < T \\ &= \delta_q, \quad \text{if } t \geq T, \end{aligned}$$

where T is an exponential random variable with mean one.

REMARK. From Theorem 4.2, Chapter 4 of [7] it is known that $\varepsilon^2 \ln \beta_\varepsilon \to 2(a(r) - a(p))$ and our proof will show that given $0 < \alpha < \tilde{\alpha} < 2(a(r) - a(p))$, we can choose for R_ε any point in the interval $[\exp(\alpha/\varepsilon^2), \exp(\tilde{\alpha}/\varepsilon^2)]$. In particular, $R_\varepsilon = \beta_\varepsilon^\gamma$ with $0 < \gamma < 1$ is a possible choice for R_ε in Theorem 2.

SOME COMMENTS ON THE METHODOLOGY. As previously mentioned, we make strong use of the large deviation theory developed by Freidlin and Wentzell [7, 13, 14] for small random perturbations. Their results are by now well known, and for the class of diffusion processes we are studying, their very basic theorem gives approximations of large deviation probabilities for $(X_\varepsilon(t): 0 \leq t \leq T)$; these are described by the "action functional"

$$\mathscr{S}_T(\phi) = \frac{1}{2} \int_0^T |\dot{\phi}(t) - b(\phi(t))|^2\, dt$$

and "scaling" ε^{-2}, for each fixed $T < +\infty$, as defined in Chapter 3 of [7]. Applying the basic large deviation theory on finite intervals, Freidlin and Wentzell could study certain typical long-time behavior of these processes, for ε

going to zero. Examples of such situations are the problems of exit from a domain which, for the deterministic system, is completely attracted to a stable equilibrium point.

In our proof we use results from Chapter 4 of [7], such as Theorem 2.1 on the exit distribution and Theorems 4.1 and 4.2 on the asymptotic behavior of the mean exit time. These theorems also provide estimates of the scaling factor β_ε. In fact, we need a slightly stronger version of these results since we handle domains containing both a stable fixed point and a saddle point or only this last one. It is easy to check that the proofs in Chapter 4 of [7] actually cover the stronger results we need. In any case they are covered by the much more general results of Chapter 6 of [7].

2. Basic ingredients. Proof of Theorem 1. If C is a closed or open set we shall use $\tau_\varepsilon^x(C)$ to denote the first hitting time of C, for the process given by (1.2), starting at x,

$$\begin{aligned}\tau_\varepsilon^x(C) &= \inf\{t > 0: X_\varepsilon^x(t) \in C\}\\ &= +\infty, \quad \text{if the above set is empty.}\end{aligned}$$

The x, and eventually the ε, will be omitted when no confusion is possible.

For the questions we are concerned with, we may as well restrict our attention to a convenient bounded domain in $\mathbb{R}^d$. The hypothesis on the potential $a(\cdot)$, with a vector field $b(\cdot)$ given by (1.3), allows us to take a bounded domain G with a sufficiently smooth boundary (C^2 suffices) so that

(a) $B_c(p) \cup B_c(q) \subseteq G$, $r \in G$;

(b) $\langle b(x), n(x)\rangle < 0$ for all $x \in \partial G$, where $n(x)$ is the unit exterior normal vector to ∂G at x [in particular, $G \cup \partial G$ is invariant under the flow (1.1), as time evolves positively];

(c) $a(r) < \min_{y \in \partial G} a(y)$.

[For instance, we may take $G = \{x: a(x) < \bar{a}\}$ with $a(r) < \bar{a}$.]

W_r^s is a $(d-1)$-dimensional manifold, being a separatrix for the basins of attraction of p and q. From (b) we see that $W_r^s \cap G$ divides G into two subdomains, $G \cap D_p$ and $G \cap D_q$.

We are interested in the escape of the processes $X_\varepsilon(\cdot)$ from $G \cap D_p$ to $G \cap D_q$. In order to study exit problems it is convenient to have no critical point at the boundary of the domain [4, 7]. Therefore, we single out a "small" subdomain G_r containing r. The picture to keep in mind is that of Figure 2 (drawn in the 2-dimensional case). Lemma 1 will help us to define precisely G_r.

LEMMA 1. *Given $\alpha > 0$ we may take $\eta > 0$ so that if $C \overset{\text{def}}{=} \{y \in \overline{G}$: distance$(y, W_r^s \cap G) < \eta\}$, then*

$$\sup_{x \in C} E_x\big(\tau_\varepsilon(\mathbb{R}^d \setminus C)\big) < e^{\alpha/\varepsilon^2},$$

for ε sufficiently small.

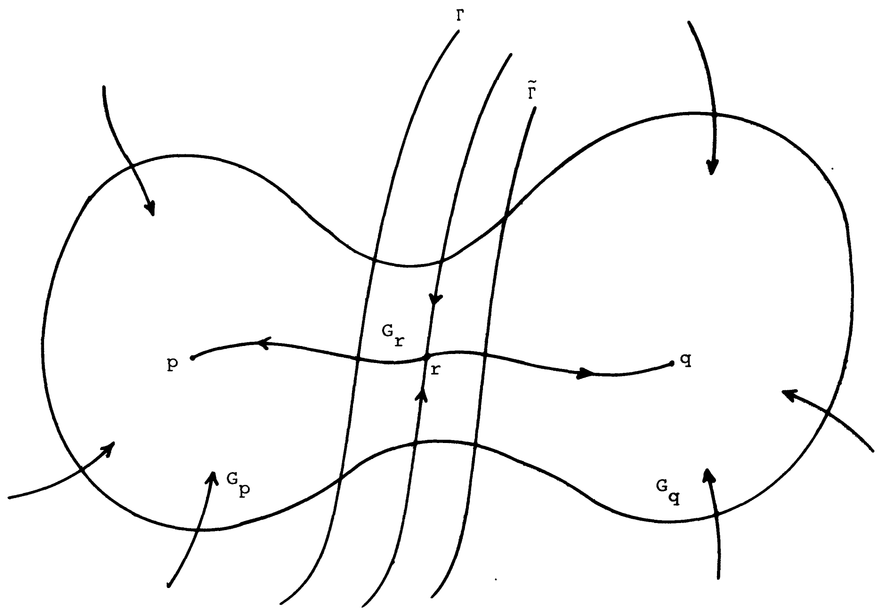

FIG. 2.

PROOF OF LEMMA 1. This is in fact a quite simple remark (analogous to Lemma 1.7, Chapter 6 of [7]). Denoting by $S_T(\cdot)$ the action functional of Freidlin and Wentzell, i.e.,

$$S_T(\phi) = \tfrac{1}{2}\int_0^T |\dot{\phi}(t) - b(\phi(t))|^2\,dt, \quad \text{if } \phi \text{ is absolutely continuous}$$

$$= +\infty, \quad \text{if not,}$$

the continuity of $b(\cdot)$ guarantees the existence of $L < +\infty$ such that for all $x, y \in \overline{G}$ one can find $\phi(\cdot)$ with $\phi(0) = x$, $\phi(T) = y$, where $T = |x - y|$ and $S_T(\phi) \le L|x - y|$.

Let us take $0 < \tilde{\eta} < \alpha/6L$ and $\tilde{T}_0 < +\infty$ so that, for each $x \in W_r^s \cap G$, $X_0^x(\cdot)$ reaches $B_{\tilde{\eta}}(r)$ during $(0, \tilde{T}_0]$. Let us fix z such that distance $(z, W_r^s \cap G) = |z - r| = \tilde{\eta}$ and let $\eta = \tilde{\eta}/2$. If $C = \{y \in \overline{G}\colon \text{distance}(y, W_r^s \cap G) < \eta\}$, then for each $x \in \overline{C}$ we can find ϕ_x with $\phi_x(0) = x$, $\phi_x(T) = z$ for some $T \le T_0 = \tilde{T}_0 + 4\eta$ and $S_T(\phi_x) \le 3L\tilde{\eta} < \alpha/2$. Thus, $\rho_{T_0}(X_\varepsilon^x(\cdot), \phi_x) < \eta$ implies $\tau_\varepsilon^x(\mathbb{R}^d \setminus C) < T_0$, and from the basic properties of the action functional [condition (II) of the definition, Chapter 3 of [7]] we have the existence of $\varepsilon_0 > 0$ such that for $\varepsilon < \varepsilon_0$,

$$\inf_{x \in \overline{C}} P_x\big(\tau_\varepsilon(\mathbb{R}^d \setminus C) < T_0\big) \ge \inf_{x \in \overline{C}} P_x\Big(\sup_{t \le T_0} |X_\varepsilon(t) - \phi_x(t)| < \eta\Big)$$

$$\ge e^{-3\alpha/4\varepsilon^2}.$$

Using the Markov property we get for $\varepsilon < \varepsilon_0$,

$$\sup_{x \in \overline{C}} P_x\big(\tau_\varepsilon(\mathbb{R}^d \setminus C) > nT_0\big) \le \big(1 - e^{-3\alpha/4\varepsilon^2}\big)^n,$$

for all $n \ge 1$, concluding the proof. □

We now fix some α_0 with $0 < \alpha_0 < 2(a(r) - a(p))$ and take η_0 according to Lemma 1. With this choice we decompose G by taking $(d-1)$-dimensional manifolds of class C^2, Γ and $\tilde{\Gamma}$, so that $\Gamma \subseteq G \cap D_p$, Γ divides G into exactly two subdomains G_p and $\tilde{\Lambda}$ with $B_c(p) \subseteq G_p \subseteq G \cap D_p$; analogously, $\tilde{\Gamma} \subseteq G \cap D_q$ and $\tilde{\Gamma}$ divides G into exactly two subdomains G_q and Λ with $B_c(q) \subseteq G_q \subseteq G \cap D_q$. We still ask that for each $x \in \Gamma \cup \tilde{\Gamma}$ its distance to $W_r^s \cap G$ is less than η_0, as fixed previously. Then we define $G_r = \Lambda \cap \tilde{\Lambda}$ and we have (see Figure 2):

(i) ∂G_r contains Γ, $\tilde{\Gamma}$ and two connected components of $\partial G \cap \overline{G}_r$.

(ii) G is the disjoint union of G_q, G_p, G_r, Γ and $\tilde{\gamma}$.

(iii) Since Γ and $\overline{\Gamma}$ are relatively compact, we may take $M < +\infty$ so that if $x \in \Gamma(\overline{\Gamma})$, then $X_0^x(\cdot)$ reaches $B_{c/2}(p)[B_{c/2}(q)$, respectively] during $(0, M)$.

(iv) For ε sufficiently small

$$\sup_{x \in G_r} E_x\big(\tau_\varepsilon(R^d \setminus G_r)\big) < e^{\alpha_0/\varepsilon^2}, \tag{2.1}$$

which is exponentially smaller than β_ε as $\varepsilon \downarrow 0$. (Theorem 4.2, Chapter 6 of [7] gives $\forall\, \delta > 0$: $\beta_\varepsilon > \exp\{2[a(r) - a(p) - \delta]/\varepsilon^2\}$ for ε sufficiently small.)

Also, according to our definition $\Lambda = G_p \cup \Gamma \cup G_r$. Setting $S_\varepsilon^x = \tau_\varepsilon^x(\partial\Lambda)$, we obviously have $S_\varepsilon^x \le T_\varepsilon^x$, for $x \in \Lambda$. Moreover, as we shall see, the difference between them is not important as $\varepsilon \downarrow 0$.

Lemma 2. *There exists a finite and positive constant M, such that*

$$P_x(T_\varepsilon < S_\varepsilon + M) \to 1, \tag{2.2}$$

as $\varepsilon \downarrow 0$ for each $x \in \Lambda$. Moreover, this convergence is uniform for $x \in B_c(p)$.

Proof. Let $M < +\infty$ be such that, for any $x \in \tilde{\Gamma}$, $X_0^x(\cdot)$ reaches $B_{c/2}(q)$ during $[0, M]$. From Theorem 1.2, Chapter 2 of [7], we know that

$$\sup_{x \in \tilde{\Gamma}} P_x\Big(\sup_{0 \le t \le M} |X_\varepsilon(t) - X_0(t)| > c/2\Big) \to 0, \quad \text{as } \varepsilon \to 0. \tag{2.3}$$

On the other hand, Theorem 5.2, Chapter 6 of [7] implies that for each $x \in \Lambda$,

$$P_x\big(X_\varepsilon(S_\varepsilon) \in \tilde{\Gamma}\big) \to 1, \quad \text{as } \varepsilon \to 0, \tag{2.4}$$

and moreover this happens uniformly for $x \in B_c(p)$, as can be easily seen from the proof. Lemma 2 follows from (2.3), (2.4) and the strong Markov property. □

Corollary. *Let γ_ε be defined by $P_p(S_\varepsilon > \gamma_\varepsilon) = e^{-1}$. Assuming that $P_p(S_\varepsilon > t\gamma_\varepsilon) \to e^{-t}$ as $\varepsilon \downarrow 0$, then, for each $x \in G \cap D_p$,* (1.5) *is equivalent to*

$$P_x(S_\varepsilon > t\gamma_\varepsilon) \to e^{-t}, \quad \text{as } \varepsilon \downarrow 0. \tag{2.5}$$

PROOF. Having proved that $P_p(S_\varepsilon > t\gamma_\varepsilon) \to e^{-t}$, Lemma 2 implies that $\gamma_\varepsilon/\beta_\varepsilon \to 1$ and the rest is immediate, again from Lemma 2. □

Let us outline the basic idea of the proof of Theorem 1. To make things simpler let us take $x = p$. If $S_\varepsilon/\gamma_\varepsilon$ converges in law to an exponential random variable, one hopes that $P_p(S_\varepsilon > (t+s)\gamma_\varepsilon)$ should behave like

$$P_p(S_\varepsilon > t\gamma_\varepsilon)P_p(S_\varepsilon > s\gamma_\varepsilon).$$

To prove this the most naive idea is simply to calculate $P_p(S_\varepsilon > (t+s)\gamma_\varepsilon)$ by conditioning on $[S_\varepsilon > t\gamma_\varepsilon]$ and on the position $X_\varepsilon(t\gamma_\varepsilon)$; the loss of memory should follow from the following two points and the strong Markov property:

(a) For y close to p [say $y \in \overline{B}_c(p)$], $P_y(S_\varepsilon > s\gamma_\varepsilon)$ should be close to

$$P_p(S_\varepsilon > s\gamma_\varepsilon).$$

(b) Starting at $y \in \Lambda \setminus \overline{B}_c(p)$ and conditioned on $[S_\varepsilon > s\gamma_\varepsilon]$ then, with overwhelming probability, $X_\varepsilon^y(\cdot)$ visits $B_c(p)$ before an interval of time $(0, \eta_\varepsilon]$ has elapsed, where η_ε is negligible compared to $s\gamma_\varepsilon$ as $\varepsilon \downarrow 0$.

These two points are made precise by the next two lemmas.

LEMMA 3. *We can find $\eta_\varepsilon > 0$ such that $\eta_\varepsilon/\gamma_\varepsilon \to 0$ as $\varepsilon \to 0$, and for all $u > 0$,*

$$\sup_{y \in B_c(p)} P_y(S_\varepsilon > u) \le P_p(S_\varepsilon > u - \eta_\varepsilon) + o(1), \tag{2.6a}$$

$$\inf_{y \in B_c(p)} P_y(S_\varepsilon > u) \ge P_p(S_\varepsilon > u + \eta_\varepsilon) - o(1), \tag{2.6b}$$

where $o(1)$ refers to a function of ε which goes to zero as ε tends to zero and does not depend on u for $u \ge 0$.

PROOF. Let us fix a compact K with smooth boundary so that $B_c(p) \subseteq \mathring{K} \subseteq K \subseteq G_p$. If we set

$$\bar{\alpha} = \inf_{y \in \partial K} [a(y) - a(p)], \tag{2.7}$$

then $\bar{\alpha} < a(r) - a(p)$, and so letting $\eta_\varepsilon = \exp(2\tilde{\alpha}/\varepsilon^2)$ for some $\bar{\alpha} < \tilde{\alpha} < a(r) - a(p)$, Theorem 4.2, Chapter 4 of [7] gives

$$\text{(i) } \eta_\varepsilon/\gamma_\varepsilon \to 0, \qquad \text{(ii) } \sup_{y \in B_c(p)} P_y(\tau_\varepsilon(K^c) > \eta_\varepsilon) \to 0. \tag{2.8}$$

For $y \in B_c(p)$, let $\mu_y^\varepsilon(\cdot)$ be the hitting distribution on ∂K for $X_\varepsilon^y(\cdot)$, i.e., $\mu_y^\varepsilon(A) = P_y(X_\varepsilon(\tau_\varepsilon(\partial K)) \in A)$ for A a Borel subset of K. By Theorem 1 of [3]

$$\sup_{y \in B_c(p)} \|\mu_y^\varepsilon - \mu_p^\varepsilon\| \to 0, \quad \text{as } \varepsilon \downarrow 0, \tag{2.9}$$

where $\|\cdot\|$ is the total variation norm.

Now setting $\tau_\varepsilon = \tau_\varepsilon(\partial K)$, the strong Markov property yields: For all $x, y \in B_c(p)$ and all $u, v > 0$,

$$\begin{aligned}
P_x(S_\varepsilon > u) &\geq P_x(S_\varepsilon > u + \tau_\varepsilon) = \int \mu_x^\varepsilon(dz) P_z(S_\varepsilon > u) \\
&\geq \int \mu_y^\varepsilon(dz) P_z(S_\varepsilon > u) - \|\mu_x^\varepsilon - \mu_y^\varepsilon\| \\
&= P_y(S_\varepsilon > u + \tau_\varepsilon) - \|\mu_x^\varepsilon - \mu_y^\varepsilon\| \\
&\geq P_y(S_\varepsilon > u + v) - P_y(\tau_\varepsilon > v) - \|\mu_x^\varepsilon - \mu_y^\varepsilon\|
\end{aligned} \tag{2.10}$$

and the lemma follows from (2.8)–(2.10). □

LEMMA 4. *Let $f_\varepsilon(t) = P_p(S_\varepsilon > \gamma_\varepsilon t)$ for $t > 0$. There exist positive numbers δ_ε with $\delta_\varepsilon \to 0$ as $\varepsilon \to 0$, such that for any s, t positive*

$$f_\varepsilon(s + \delta_\varepsilon) f_\varepsilon(t + \delta_\varepsilon) - o_t(1) \leq f_\varepsilon(t + s) \leq f_\varepsilon(s) f_\varepsilon(t - \delta_\varepsilon) + o_t(1), \tag{2.11}$$

where $o_t(1)$ is a function of ε and t which tends to zero as $\varepsilon \to 0$, this convergence being uniform for $t \geq t_0$, given any $t_0 > 0$.

PROOF. For each x in the basin of attraction of p, let $T(x)$ be the time needed for the orbit of (1.1) starting at x to arrive in $B_{c/2}(p)$. Then $T(x) < +\infty$ and by the upper semicontinuity $\sup_{x \in G_p} T(x) = M_0 < +\infty$. Then we easily get (as in Lemma 2)

$$\sup_{x \in \overline{G}_p} P_x(\tau_\varepsilon(B_c(p)) > M_0) \to 0, \quad \text{as } \varepsilon \to 0. \tag{2.12}$$

On the other hand, if we take $\eta'_\varepsilon = \exp(\alpha'/\varepsilon^2)$ where $\alpha_0 < \alpha' < 2(a(r) - a(p))$, then (2.1) implies that

$$\sup_{y \in \overline{G}_r} E_y(\tau_\varepsilon(G_r^c))/\eta'_\varepsilon \to 0, \quad \text{as } \varepsilon \to 0. \tag{2.13}$$

Moreover, as noticed before, Theorem 4.2, Chapter 4 of [7] gives that $\eta'_\varepsilon/\gamma_\varepsilon \to 0$ as $\varepsilon \downarrow 0$.

Decomposing the event $[S_\varepsilon > \eta'_\varepsilon, \tau_\varepsilon(B_c(p)) > \eta'_\varepsilon]$ according to the two possibilities of remaining in G_r during $[0, \eta'_\varepsilon/2]$ or reaching $\overline{G}_p$ during $[0, \eta'_\varepsilon/2]$, we have

$$\begin{aligned}
&\sup_{y \in \overline{G}_r} P_y(S_\varepsilon > \eta'_\varepsilon, \tau_\varepsilon(B_c(p)) > \eta'_\varepsilon) \\
&\quad \leq \sup_{y \in \overline{G}_r} P_y(\tau_\varepsilon(G_r^c) > \eta'_\varepsilon/2) + \sup_{y \in \overline{G}_p} P_y(\tau_\varepsilon(B_c(p)) > \eta'_\varepsilon/2),
\end{aligned} \tag{2.14}$$

where, in the second term on the r.h.s., we used the strong Markov property. From (2.12) and (2.13), both terms on the r.h.s. of (2.14) tend to zero. Thus, if we define for all $s \geq 0$,

$$R^s = \inf\{u > s\gamma_\varepsilon \colon X_\varepsilon(u) \in B_c(p)\},$$

then (2.12), (2.14) and the Markov property (at $s\gamma_\varepsilon$) yield

$$\sup_{y \in \Lambda} P_y(S_\varepsilon > s\gamma_\varepsilon + \eta'_\varepsilon, R^s > s\gamma_\varepsilon + \eta'_\varepsilon) \to 0, \quad \text{as } \varepsilon \to 0, \tag{2.15}$$

the limit being uniform on s. In particular, for any $s, t > 0$,

$$P_p(S_\varepsilon > (t+s)\gamma_\varepsilon,\ R^s > s\gamma_\varepsilon + \eta'_\varepsilon) \underset{\varepsilon\downarrow 0}{\to} 0,$$

uniformly on $(s, t) \in (0, +\infty) \times [t_0, +\infty)$ for any given $t_0 > 0$.

On the other hand,

$$P_p(S_\varepsilon > (t+s)\gamma_\varepsilon,\ R^s \le s\gamma_\varepsilon + \eta'_\varepsilon)$$
$$= \int_{(s\gamma_\varepsilon,\, s\gamma_\varepsilon + \eta'_\varepsilon]} P_p(R^s \in du,\ X_\varepsilon(R^s) \in dy,\ S_\varepsilon > u) P_y(S_\varepsilon > (t+s)\gamma_\varepsilon - u) \tag{2.16}$$

and the r.h.s. of (2.16) is bounded above by

$$P_p(S_\varepsilon > s\gamma_\varepsilon) \sup_{y \in B_c(p)} P_y(S_\varepsilon > t\gamma_\varepsilon - \eta'_\varepsilon), \tag{2.17}$$

and bounded below by

$$P_p(S_\varepsilon > s\gamma_\varepsilon + \eta'_\varepsilon,\ R^s \le s\gamma_\varepsilon + \eta'_\varepsilon) \inf_{y \in B_c(p)} P_y(S_\varepsilon > t\gamma_\varepsilon). \tag{2.18}$$

The lemma follows immediately from Lemma 3 and (2.15)–(2.18). □

Proof of Theorem 1. First, let us consider the case $x = p$. By the Corollary to Lemma 2 it suffices to prove that under P_p the random variables $S_\varepsilon/\gamma_\varepsilon$ converge in law to an exponential random variable with mean one. The tightness on $(0, +\infty)$ follows easily from Lemma 4: In fact, applying (2.11) with $t = 2$ and inductively on $s = 1, 2, \dots$, we see that for any $\delta > 0$ there exist ε_0, n_0 with $f_\varepsilon(d_0) < \delta$ and $f_\varepsilon(1/n_0) \ge 1 - \delta$ whenever $\varepsilon \le \varepsilon_0$. Now if μ is a probability measure on the Borel sets of $(0, +\infty)$ appearing as a weak*-limit point of the law of $S_\varepsilon/\gamma_\varepsilon$ and if we set $f(t) = \mu(t, +\infty)$, Lemma 4 implies that

$$f(t+s) = f(t)f(s), \tag{2.19}$$

for all $s, t > 0$ such that s, t and $s + t$ are continuity points of $f(\cdot)$. The right continuity of f implies that (2.19) holds for all $s, t \ge 0$. Thus, $f(t) = e^{-kt}$ for some $k \ge 0$. Since $f_\varepsilon(1) = e^{-1}$ for all $\varepsilon > 0$ one can easily see that in fact $k = 1$, proving the result in the case $x = p$.

The continuity of the limit implies that, in fact, $f_\varepsilon(t) \to e^{-t}$ uniformly in t. Now, from Lemma 3, we have

$$\sup_{x \in B_c(p)} P_x(S_\varepsilon > t\gamma_\varepsilon) \le f_\varepsilon(t - \eta_\varepsilon/\gamma_\varepsilon) + o(1),$$

$$\inf_{x \in B_c(p)} P_x(S_\varepsilon > t\gamma_\varepsilon) \ge f_\varepsilon(t + \eta_\varepsilon/\gamma_\varepsilon) - o(1),$$

where $\eta_\varepsilon/\gamma_\varepsilon \to 0$. Therefore for each $t \ge 0$, $P_x(S_\varepsilon > t\gamma_\varepsilon) \to e^{-t}$ uniformly on $x \in B_c(p)$, as $\varepsilon \to 0$.

When $x = p$ we know that $\gamma_\varepsilon/\beta_\varepsilon \to 1$. We use now Lemma 2 to conclude that for each $t \ge 0$,

$$\sup_{x \in B_c(p)} |P_x(T_\varepsilon > t\beta_\varepsilon) - e^{-t}| \to 0, \quad \text{as } \varepsilon \downarrow 0. \tag{2.20}$$

Now, given x in the basin of attraction of p, let $T(x) < +\infty$ be the time necessary for $X_0^x(\cdot)$ to reach $B_{c/2}(p)$. Since

$$P_x\Big(\sup_{t\le T(x)} |X_\varepsilon(t) - X_0(t)| > c/2\Big) \underset{\varepsilon\downarrow 0}{\to} 0,$$

we get

$$P_x(\tau_\varepsilon(B_c(p)) > T(x)) \underset{\varepsilon\downarrow 0}{\to} 0. \tag{2.21}$$

Thus

$$P_x(T_\varepsilon > t\beta_\varepsilon) - P_x(T_\varepsilon > t\beta_\varepsilon, \tau_\varepsilon(B_c(p)) \le T(x)) \to 0. \tag{2.22}$$

Using the strong Markov property at $\tau_\varepsilon(B_c(p))$ for the last term of the l.h.s. of (2.22), we conclude from (2.20) [and again (2.21)] that $P_x(T_\varepsilon > t\beta_\varepsilon) \to e^{-t}$. □

3. Stability of time averages. Proof of Theorem 2. Let us first state two lemmas: The first one (Lemma 5) is a particular case of a result of Freidlin and Wentzell and the second one gives an estimate needed in the proof of both (1.7) and (1.8).

LEMMA 5. (1) *For any $\alpha > 0$ we can find $\varepsilon_0 > 0$ such that for any $\varepsilon \le \varepsilon_0$,*

$$\sup_{x\in G_r} E_x(\tau_\varepsilon(G_r^c)) \le \exp(\alpha/\varepsilon^2). \tag{3.1}$$

(2) *For any $\theta > 0$ and any $\alpha > 0$ if we set $t_\varepsilon = \exp(\alpha/\varepsilon^2)$, then*

$$\sup_{x\in\Lambda} P_x\Big(\tau_\varepsilon(B_\theta(p)) > \sqrt{t_\varepsilon}, S_\varepsilon > t_\varepsilon\Big) \to 0, \quad \text{as } \varepsilon \to 0. \tag{3.2}$$

PROOF. Part (1) is explicitly written up just for clarity, since in fact it is an immediate consequence of Theorem 5.3, Chapter 6 of [7]. Indeed, using their definitions (page 193 of [7]) it is easy to see that, in our case, taking $D = G_r$,

$$W_D(x, y) = \min\{V_D(x, y), V_D(x, r) + V_D(r, y)\},$$

for all $x \in D$, $y \in \partial D$ and so $\min_{y\in\partial D} W_D(x, y) = 0$.

To obtain part (2), since $t_\varepsilon \to +\infty$, it suffices to prove that

$$\sup_{x\in\overline{G}_r} P_x\Big(\tau_\varepsilon(\overline{G}_p) > \sqrt{t_\varepsilon/2}, S_\varepsilon > t_\varepsilon\Big) \to 0.$$

But, for ε small enough, $t_\varepsilon > \sqrt{t_\varepsilon/2}$. Therefore,

$$P_x\Big(\tau_\varepsilon(\overline{G}_p) > \sqrt{t_\varepsilon/2}, S_\varepsilon > t_\varepsilon\Big) \le P_x\Big(\tau_\varepsilon(G_r^c) > \sqrt{t_\varepsilon/2}\Big)$$

and the result follows from part (1). □

LEMMA 6. *Let us call $I_\theta(\cdot)$ the indicator function of $B_\theta(p)$ and take $R'_\varepsilon = \exp(\alpha/\varepsilon^2)$ with $0 < \alpha < 2[a(r) - a(p)]$. Then for any $\theta > 0$, $\delta > 0$, there exists*

$\varepsilon_0 = \varepsilon_0(\theta, \delta) > 0$ *such that*

$$\sup_{x \in \Lambda} P_x\left\{\left|\frac{1}{R}\int_0^R I_\theta(X_\varepsilon(s))\,ds - 1\right| > \delta,\ S_\varepsilon > R\right\} \le \exp(-c_\delta R/R'_\varepsilon), \tag{3.3}$$

for $R > 2R'_\varepsilon$ *and* $\varepsilon \le \varepsilon_0$, *where* c_δ *is a positive constant.*

PROOF. First we fix α as in the statement of the lemma. Now, given $\theta > 0$, we choose θ_0 with $0 < \theta_0 \le \theta$ such that $0 < \min_{z:\,|z-p|=\theta_0}(a(z) - a(p)) < \alpha$. If we define t_ε as

$$t_\varepsilon = \exp\left\{\varepsilon^{-2} \min_{z:\,|z-p|=\theta}(a(z) - a(p))\right\}, \tag{3.4}$$

then $t_\varepsilon \le R'_\varepsilon$ and $t_\varepsilon/R'_\varepsilon \to 0$. Let us consider the decomposition $(0, R] = \bigcup_{i=1}^{N_\varepsilon}((i-1)t_\varepsilon, it_\varepsilon] \cup (N_\varepsilon t_\varepsilon, R]$, where $N_\varepsilon = [R/t_\varepsilon]$. If $t_\varepsilon/R < \delta/2$, then

$$\begin{aligned} &\left\{\left|\frac{1}{R}\int_0^R I_{\theta_0}(X_\varepsilon(s))\,ds - 1\right| > \delta,\ S_\varepsilon > R\right\} \\ &\quad \subseteq \left\{\left|\frac{1}{N_\varepsilon t_\varepsilon}\int_0^{N_\varepsilon t_\varepsilon} I_{\theta_0}(X_\varepsilon(s))\,ds - 1\right| > \delta/2,\ S_\varepsilon > N_\varepsilon t_\varepsilon\right\}. \end{aligned} \tag{3.5}$$

We may take K_1, K_2 compact neighborhoods of p so that

(i) $p \in \mathring{K}_1 \subseteq K_1 \subseteq \mathring{K}_2 \subseteq B_{\theta_0}(p)$;

(ii) K_1, K_2 have smooth boundaries and $\langle b(x), n_i(x)\rangle < 0$ for all $x \in \partial K_i$, where $n_i(x)$ denotes the unitary exterior normal vector to ∂K_i at x; and

$$\text{(iii)} \quad \min_{z:\,|z-p|=\theta}(a(z) - a(p)) < 2\min_{z \in \partial K_2}(a(z) - a(p)). \tag{3.6}$$

Let us define the random variables Y_ε^i, $i = 1, \dots, N_\varepsilon$, as follows:

$Y_\varepsilon^i = 0$, if $X_\varepsilon(\cdot)$ visits K_1 during $((i-1)t_\varepsilon, (i-1)t_\varepsilon + \sqrt{t_\varepsilon}\,]$ and spends the rest of time interval $((i-1)t_\varepsilon, it_\varepsilon\,]$ in K_2

$= 1$, otherwise.

Then

$$\begin{aligned} p(\varepsilon) &\overset{\text{def}}{=} \sup_{x\in\Lambda} P_x\left(Y_\varepsilon^1 = 1,\ S_\varepsilon > t_\varepsilon\right) \\ &\le \sup_{x\in\Lambda} P_x\left(\tau_\varepsilon(K_1) > \sqrt{t_\varepsilon},\ S_\varepsilon > t_\varepsilon\right) + \sup_{x\in K_1} P_x(\tau_\varepsilon(K_2^c) < t_\varepsilon). \end{aligned} \tag{3.7}$$

By part (2) of Lemma 5, we see that the first term on the r.h.s. of (3.7) goes to zero as $\varepsilon \to 0$. Moreover, using Theorem 4.2, Chapter 4 of [7], (3.4) implies that the second term also goes to zero.

Now, for $t_\varepsilon/R < \delta/2$ (which happens if $R > R'_\varepsilon$ and ε is small enough), we can write, using (3.5), that for all $x \in \Lambda$,

$$\begin{aligned} &P_x\left\{\left|\frac{1}{R}\int_0^R I_{\theta_0}(X(s))\,ds - 1\right| > \delta,\, S_\varepsilon > R\right\} \\ &\quad \le P_x\left\{\frac{1}{N_\varepsilon}\sum_{i=1}^{N_\varepsilon} Y_\varepsilon^i > \frac{\delta}{2} - \frac{1}{\sqrt{t_\varepsilon}},\, S_\varepsilon > N_\varepsilon t_\varepsilon\right\} \\ &\quad \le P_x\left\{\sum_{k=1}^{N_\varepsilon} Y_\varepsilon^i > \frac{\delta}{4}N_\varepsilon,\, S_\varepsilon > N_\varepsilon t\right\}, \end{aligned} \tag{3.8}$$

if ε is sufficiently small (so that $\sqrt{t_\varepsilon} > 4/\delta$).

Now we can estimate the r.h.s. of (3.8) as follows: For all $x \in \Lambda$,

$$\begin{aligned} &P_x\left(\sum_{i=1}^{N_\varepsilon} Y_\varepsilon > N_\varepsilon\delta/4,\, S_\varepsilon > N_\varepsilon t_\varepsilon\right) \\ &\quad \le \exp(-N_\varepsilon\delta/4)E_x\left\{1_{[S_\varepsilon > N_\varepsilon t_\varepsilon]}\exp\left(\sum_{i=1}^{N_\varepsilon} Y_\varepsilon^i\right)\right\} \\ &\quad \le \exp(-N_\varepsilon\delta/4)\left\{\sup_{y\in\Lambda} E_y 1_{[S_\varepsilon > t_\varepsilon]}\exp\left(Y_\varepsilon^1\right)\right\}^{N_\varepsilon}, \end{aligned} \tag{3.9}$$

where the last inequality is a consequence of the Markov property. Now

$$\begin{aligned} \sup_{y\in\Lambda} E_y\left[1_{(S_\varepsilon > t_\varepsilon)}\exp\left(Y_\varepsilon^1\right)\right] &\le \sup_{y\in\Lambda}\left[eP\left(Y_\varepsilon^1 = 1,\, S_\varepsilon > t_\varepsilon\right) + P\left(Y_\varepsilon^1 = 0,\, S_\varepsilon > t_\varepsilon\right)\right] \\ &= \sup_{y\in\Lambda}\left[P_y(S_\varepsilon > t_\varepsilon) + (e-1)p(\varepsilon)\right] \\ &\le \exp[(e-1)p(\varepsilon)], \end{aligned} \tag{3.10}$$

where $p(\varepsilon)$ is defined in (3.7).

From (3.8)–(3.10) we get

$$\begin{aligned} &\sup_{x\in\Lambda} P_x\left\{\left|\frac{1}{R}\int_0^R I_\theta(X_\varepsilon(s))\,ds - 1\right| > \delta,\, S_\varepsilon > R\right\} \\ &\quad \le \exp[-N_\varepsilon\delta/4 + (e-1)p(\varepsilon)] \\ &\quad \le \exp[-N_\varepsilon\delta/8] \le \exp[-c_8 r/R'_\varepsilon], \end{aligned}$$

with $c_\delta = \delta/16$ for ε sufficiently small and $R \ge 2R'_\varepsilon$. □

THEOREM 3. *There exist $R_\varepsilon \uparrow +\infty$ with $R_\varepsilon/\gamma_\varepsilon \to 0$ such that for all $\theta > 0$, $\delta > 0$,*

$$\sup_{X\in B_c(p)} P_x\left\{\sup_{0<s<S_\varepsilon - 2R_\varepsilon}\left|\frac{1}{R_\varepsilon}\int_s^{s+R_\varepsilon} I_\theta(X_\varepsilon(u))\,du - 1\right| > \delta\right\} \to 0, \tag{3.11}$$

as $\varepsilon \to 0$. *Also, in these conditions we may replace* S_ε *by* $T_\varepsilon - M$ *in* (3.11), *for a constant* M, *as given in Lemma* 2.

PROOF. Obviously, it is enough to prove (3.11). With $R_\varepsilon > 0$ to be chosen later, for each θ, δ positive and each l nonnegative integer we set

$$A_l^\varepsilon(\theta, \delta) = \left\{ \left| \frac{1}{R_\varepsilon} \int_{lR_\varepsilon}^{(l+1)R_\varepsilon} I_\theta(X_\varepsilon(s))\, ds - 1 \right| \le \delta \right\}.$$

Then, (3.11) will follow if we prove that, for suitable R_ε, with $R_\varepsilon/\gamma_\varepsilon \to 0$, we have

$$\inf_{x \in B_c(p)} P_x\left(\bigcap_{0 \le l < l_\varepsilon} A_l^\varepsilon(\theta, \delta),\ l_\varepsilon \ge 1 \right) \to 1,$$

where

$$l_\varepsilon = \inf\{ l \in \mathbb{N}: S_\varepsilon \le (l+1) R_\varepsilon \}.$$

To simplify the notation let us write A_l^ε for $A_l^\varepsilon(\theta, \delta)$. Then we have

$$\begin{aligned} P_x\left(\bigcap_{0 \le l < l_\varepsilon} A_l^\varepsilon,\ l_\varepsilon \ge 1 \right) &= \sum_{L=1}^{+\infty} P_x\left(\bigcap_{0 \le l < L} A_l^\varepsilon,\ l_\varepsilon = L \right) \\ &= P_x(l_\varepsilon \ge 1) - \sum_{L=1}^{\infty} P_x\left(\bigcup_{0 \le l < L} (\Omega \setminus A_l^\varepsilon),\ l_\varepsilon = L \right) \\ &\ge P_x(l_\varepsilon \ge 1) - \sum_{L=1}^{K_\varepsilon} P_x\left(\bigcup_{0 \le l < L} (\Omega \setminus A_l^\varepsilon),\ l_\varepsilon = L \right) \\ &\quad - P_x(l_\varepsilon > K_\varepsilon). \end{aligned}$$

Now, to conclude the proof of (3.11), we want to find R_ε and K_ε such that

$$(3.12)\qquad \begin{aligned} &(1) \qquad \inf_{x \in B_c(p)} P_x(l_\varepsilon \ge 1) \to 1, \\ &(2) \qquad \sup_{x \in B_c(p)} P_x(l_\varepsilon > K_\varepsilon) \to 0, \\ &(3) \qquad \sup_{x \in B_c(p)} \sum_{L=1}^{K_\varepsilon} P_x\left[\bigcup_{0 \le l < L} (\Omega \setminus A_l^\varepsilon),\ l_\varepsilon = L \right] \to 0, \quad \text{when } \varepsilon \to 0. \end{aligned}$$

Notice that from the definition of l_ε and from Lemma 3 there exist $\eta_\varepsilon > 0$ with $\eta_\varepsilon/\gamma_\varepsilon \to 0$, when $\varepsilon \to 0$, such that

$$\begin{aligned} \inf_{x \in B_c(p)} P_x(l_\varepsilon \ge 1) &= \inf_{x \in B_c(p)} P_x(S_\varepsilon > R_\varepsilon) \\ &\ge P_p(S_\varepsilon > R_\varepsilon + \eta_\varepsilon) - o(1) \\ &= f_\varepsilon((R_\varepsilon + \eta_\varepsilon)/\gamma_\varepsilon) - o(1), \end{aligned}$$

which, by Theorem 1, goes to 1, provided we have

CONDITION C1. $R_\varepsilon/\gamma_\varepsilon \to 0$ as $\varepsilon \downarrow 0$.

From Lemma 3 we still have

$$\begin{aligned}\sup_{x \in B_c(p)} P_x(l_\varepsilon > K_\varepsilon) &\le \sup_{x \in B_c(p)} P_x(S_\varepsilon > (K_\varepsilon + 1)R_\varepsilon)\\ &\le P_p(S_\varepsilon > K_\varepsilon R_\varepsilon - \eta_\varepsilon) + o(1)\\ &= f_\varepsilon((K_\varepsilon R_\varepsilon - \eta_\varepsilon)/\gamma_\varepsilon) + o(1)\end{aligned}$$

and from Theorem 1 this goes to zero if

CONDITION C2. $K_\varepsilon R_\varepsilon/\gamma_\varepsilon \to +\infty$ as $\varepsilon \downarrow 0$.

Finally, we estimate (3) in (3.12): For each $x \in B_c(p)$,

$$\begin{aligned}&\sum_{L=1}^{K_\varepsilon} P_x\Big(\bigcup_{0 \le l < L} (\Omega \setminus A_l^\varepsilon), l_\varepsilon = L\Big)\\ &\quad \le \sum_{L=1}^{K_\varepsilon} P_x \bigcup_{0 \le l < L} (\Omega \setminus A_l^\varepsilon, l_\varepsilon > l)\\ &\quad \le \sum_{L=1}^{K_\varepsilon} \sum_{0 \le l < L} P_x(\Omega \setminus A_l^\varepsilon, l_\varepsilon > l)\\ &\quad \le K_\varepsilon^2 \sup_{0 \le l < K_\varepsilon} P_x(\Omega \setminus A_l^\varepsilon, l_\varepsilon > l).\end{aligned}$$

From Lemma 6 if we set $R'_\varepsilon = \exp(\eta/\varepsilon^2)$ with $0 < \eta < 2(a(r) - a(p))$ and let $R_\varepsilon \ge 2R'_\varepsilon$ this is uniformly bounded for $x \in \Lambda$ by

$$K_\varepsilon^2 \exp(-c_\delta R_\varepsilon/R'_\varepsilon),$$

provided ε is sufficiently small. Thus, (3) in (3.12) will be true if K_ε and R_ε satisfy $R_\varepsilon \ge \exp(\eta/\varepsilon^2)$ for some $\alpha > 0$, and

CONDITION C3. $K_\varepsilon^2 \exp(-\delta R_\varepsilon/R'_\varepsilon) \to 0$ as $\varepsilon \downarrow 0$ for any $\delta > 0$.

Now it is easy to find $K_\varepsilon, R_\varepsilon$ satisfying Conditions C1–C3. As already mentioned, Theorem 4.2, Chapter 4 of [7] tells us that given $\alpha < 2(a(r) - a(p)) < \tilde{\alpha}$, then

$$\exp(\alpha/\varepsilon^2) \le \gamma_\varepsilon \le \exp(\tilde{\alpha}/\varepsilon^2),$$

for ε sufficiently small. Thus, taking, for instance, $R_\varepsilon = \exp(\alpha/\varepsilon^2)$ with fixed $0 < \alpha < 2(a(r) - a(p))$, then C1 is already satisfied. To satisfy C2 and C3 we may take, e.g., $R'_\varepsilon = R_\varepsilon^{1/2}$ and $K_\varepsilon = \exp(\tilde{C}/\varepsilon^2)$, with $\tilde{C} > 2(a(r) - a(p)) - \alpha$. This concludes the proof of Theorem 3. □

PROOF OF THEOREM 2. To fix ideas let us take $R_\varepsilon = \exp(\alpha/\varepsilon^2)$, where $0 < \alpha < 2(a(r) - a(p))$. From Theorem 3, it follows immediately that for $\delta > 0$, and for each $f \in C_b(\mathbb{R}^d)$,

$$\sup_{x \in B_c(p)} P_x\left\{ \sup_{0 \le s < T_\varepsilon - 3R_\varepsilon} \left| \frac{1}{R_\varepsilon} \int_s^{s+R_\varepsilon} f(X_\varepsilon(u))\,du - f(p) \right| > \delta \right\} \to 0. \tag{3.13}$$

Now given x in the basin of attraction of p we can take $T(x) < +\infty$ to be the hitting time of $B_{c/2}(p)$ for the orbit $X_0^x(\cdot)$. Since $P_x(\tau_\varepsilon(B_c(p)) \le T(x))$ goes to one [cf. (2.21)] it is very easy to conclude from (3.13) and the strong Markov property, that (1.7) holds for such x.

Since $a(r) - a(p) < a(r) - a(q)$, the above choice of R_ε also works if we reverse the roles of p and q in Theorem 3. This fact and the use of strong Markov property at $\tilde{T}_\varepsilon$ allows us to conclude the validity of (1.8).

Having already proved (1.7), (1.8) and Theorem 1 the final conclusion of Theorem 2 is clear. Let us recall that: (1) $\mathscr{M}_1$ is a Polish space for the w^*-topology and the Skorokhod topology on $D([0, +\infty), \mathscr{M}_1)$ can be defined as in [1, 6, 9, 11]. (2) We can find a sequence $g_1, g_2, \dots$ in $C_b(\mathbb{R}^d)$, with $\|g_k\|_\infty = 1$ for all k and such that $\mu_n \to_{w^*} \mu$ in $\mathscr{M}_1$ if and only if $\int g_k\,d\mu_n \to \int g_k\,d\mu$ as $n \to +\infty$ for each k. Thus, if we set

$$\rho(\mu, \eta) = \sum_k \frac{1}{2^k} \left| \int g_k\,d\mu - \int g_k\,d\eta \right|, \tag{3.14}$$

then $\rho(\cdot)$ is a metric generating the w^*-topology on $\mathscr{M}_1$. From (1.7), (1.8) and (3.14) we have: For each $\delta > 0$ and each x in the basin of attraction of p,

$$P_x\left(\sup_{0 \le s < T_\varepsilon/\beta_\varepsilon - 3R_\varepsilon/\beta_\varepsilon} \rho\left(\nu_s^\varepsilon, \delta_p\right) > \delta \right) \to 0$$

and

$$P_x\left(\sup_{T_\varepsilon/\beta_\varepsilon \le s < (\tilde{T}_\varepsilon - 3R)/\beta_\varepsilon} \rho\left(\nu_s^\varepsilon, \delta_q\right) > \delta \right) \to 0, \quad \text{as } \varepsilon \to 0.$$

Moreover, if we define $\tilde{\beta}_\varepsilon$ via $P_q(\tilde{T}_\varepsilon > \tilde{\beta}_\varepsilon) = e^{-1}$, reversing the roles of p and q in Theorem 1, it follows that for each $t > 0$,

$$P_x\left(\tilde{T}_\varepsilon > t\tilde{\beta}_\varepsilon\right) \to e^{-t} \quad \text{as } \varepsilon \to 0,$$

uniformly for $x \in B_c(q)$. Moreover, Theorem 4.2, Chapter 4 of [7] implies that for any $0 < c < 2(a(p) - a(q))$ we must have $\tilde{\beta}_\varepsilon/\beta_\varepsilon \ge \exp(c/\varepsilon^2)$, provided ε is sufficiently small. In particular, $\tilde{\beta}_\varepsilon/\beta_\varepsilon \to +\infty$ and we get that for all $t > 0$,

$$\inf_{x \in B_c(q)} P_x\left(\tilde{T}_\varepsilon/\beta_\varepsilon > t\right) \to 1, \quad \text{as } \varepsilon \to 0. \tag{3.15}$$

Using (1.7), (1.8), Theorem 1 and (3.15) we conclude the proof of Theorem 2. □

Final remark. The final statement in Theorem 2 is written in terms of $\tilde{\nu}_t^\varepsilon$, which is slightly different than ν_t^ε, i.e., they differ only for t in an interval of length at most $3R_\varepsilon/\beta_\varepsilon$ (which goes to zero). Conceptually this does not change anything, and the important facts are (1.7), (1.8) and the exponentiality of $T_\varepsilon/\beta_\varepsilon$ as $\varepsilon \downarrow 0$ for $x \in D_p$, together with the fact that $\tilde{\beta}_\varepsilon/\beta_\varepsilon \to +\infty$, which means that with the scaling $t\beta_\varepsilon$ we do not see the return to $B_c(p)$.

Acknowledgments. E.O. wishes to thank the IME-USP (São Paulo) and the IMPA (Rio de Janerio). A.G. and M.E.V. thank the IHES (Bures-sur-Yvette). Very useful discussions with Rémi Langevin and Edson Vargas are also acknowledged.

NOTE. After this work was concluded we learned that a result very similar to our Theorem 1 had been obtained by P. Videlaine-Maggi in her thesis, "Petites perturbations aléatoires de systèmes dynamiques" (Univ. Paris 7, June 1984).

REFERENCES

[1] BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

[2] CASSANDRO, M., GALVES, A., OLIVIERI, E. and VARES, M. E. (1984). Metastable behavior of stochastic dynamics: A pathwise approach. *J. Statist. Phys.* **35** 603–634.

[3] DAY, M. (1982). Exponential levelling of stochastically perturbed dynamical systems. *SIAM J. Math. Anal.* **13** 532–540.

[4] DAY, M. (1983). On the exponential exit law in the small parameter exit problem. *Stochastics* **8** 297–323.

[5] DE MASI, A., IANIRO, N., PELLEGRINOTTI, A. and PRESUTTI, E. (1984). A survey of the hydrodynamical behavior of many-particle systems. In *Studies in Statistical Mechanics* **11** (J. L. Lebowitz and E. W. Montroll, eds.) 124–294. North-Holland, Amsterdam.

[6] FLEMING, W. and VIOT, M. (1979). Some measure-valued Markov processes in population genetics. *Indiana Univ. Math. J.* **28** 817–843.

[7] FREIDLIN, M. I. and WENTZELL, A. D. (1984). *Random Perturbations of Dynamical Systems*. Springer, New York.

[8] LEBOWITZ, J. L. and PENROSE, O. (1979). Towards a rigorous molecular theory of metastability. In *Fluctuation Phenomena*; *Studies in Statistical Mechanics*. (J. L. Lebowitz and E. W. Montroll, eds.). North-Holland, Amsterdam.

[9] LINDWALL, T. (1973). Weak convergence of probability measures and random functions in the function space $D[0, +\infty)$. *J. Appl. Probab.* **10** 109–121.

[10] PALIS, J. and MELO, W. (1980). *Geometric Theory of Dynamical Systems*. Springer, New York.

[11] PARTHASARATHY, K. R. (1967). *Probability Measures on Metric Spaces*. Academic, New York.

[12] SCHONMANN, R. H. (1985). Metastability for the contact process. *J. Statist. Phys.* **41** 445–464.

[13] WENTZELL, A. D. and FREIDLIN, M. I. (1970). On small random perturbations of dynamical systems. *Russian Math. Surveys* **25** 1–55.

[14] WENTZELL, A. D. and FREIDLIN, M. I. (1972). Some problems concerning stability under small random perturbations. *Theory Probab. Appl.* **17** 269–283.

ANTONIO GALVES
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
UNIVERSIDADE DE SÃO PAULO
CAIXA POSTAL 20570 (AG. IGUATEMI)
SÃO PAULO
BRASIL

ENZO OLIVIERI
DIPARTAMENTO MATEMÁTICO–C.N.R.-G.N.F.M.
UNIVERSITÀ DI ROMA, LA SAPIENZA
00185-ROMA
ITÁLIA

MARIA EULALIA VARES
INSTITUTO DE MATEMÁTICA PURA E APLICADA
ESTRADA D. CASTORINA
110-JARDIM BOTÂNICO, 22460
RIO DE JANEIRO
RJ-BRASIL