The Annals of Probability
1990, Vol. 18, No. 4, 1676–1692

A DISCRETE-TIME VERSION OF THE WENTZELL–FREIDLIN THEORY[1]

BY YURI KIFER

The Hebrew University

Dedicated to E. B. Dynkin on his 65th birthday

We present a version of the Wentzell–Freidlin theory for Markov chains which includes random perturbations not only of deterministic motions but also of Markov chains. Some results for the continuous-time case are obtained as corollaries. In particular, by this method one can treat random perturbations of degenerate diffusions even when the large deviations principle fails.

1. Introduction. Let X_n^ε, $\varepsilon > 0$, $n = 0, 1, \dots,$ be a family of Markov chains on a compact metric space M with transition probabilities $P^\varepsilon(x, \cdot)$, $x \in M$, which are Borel measures Borel measurably depending on x and such that for any open set $U \subset M$ uniformly in $x \in M$,

$$\lim_{\varepsilon \to 0} \varepsilon \log P^\varepsilon(x, U) = -\inf_{y \in U} \rho(x, y), \tag{1.1}$$

where $\rho(x, y) \geq 0$ is a continuous function on $M \times M$. Wentzell and Freidlin [11] considered diffusion processes X_t^ε generated by operators of the form $L^\varepsilon = \varepsilon L + b$, where L is a nondegenerate elliptic differential operator of the second order and b is a vector field, i.e., a differential operator of the first order. They studied the asymptotic behavior as $\varepsilon \to 0$ of invariant measures of processes X_t^ε, of the distribution of exit points of X_t^ε from a bounded domain and of the principal eigenvalue of the operator L^ε by estimating the probabilities for processes X_t^ε to stay in tube neighborhoods of different curves. We shall present here a discrete-time version of their results which works both for diffusion-type random perturbations and for perturbations by means of processes with jumps considered in Section 2 of Chapter 5 in [5]. Since the transition probabilities $P^\varepsilon(t, x, \cdot)$ of X_t^ε satisfy in these cases some kind of (1.1), it turns out that their results can be derived from ours by considering X_t^ε only at moments $t = k\Delta$, $k = 0, 1, 2, \dots,$ for some $\Delta > 0$. We remark that there now exist viscosity solutions methods (see [3]) which, studying a nonlinear equation for $\varepsilon \log P^\varepsilon(t, x, U)$, enable one to obtain directly limits of the sort (1.1) without employing probabilistic large deviations estimates from [11]. Via a more careful analysis one can relax the compactness and the continuity assumptions on M and ρ, respectively.

Received April 1988; revised November 1988.
[1]Partly supported by the U.S. Army Research Office through the Mathematical Sciences Institute at Cornell University. This work was done while the author visited the Department of Mathematics at Cornell University.
AMS 1980 *subject classifications*. Primary 60F10; secondary 60J05, 60J60.
Key words and phrases. Random perturbations, large deviations, attractors.

The general setup via (1.1) does not even presume that the Markov chains X_n^ε are perturbations of something else and, in fact, the study of the asymptotic behavior as $\varepsilon \to 0$ goes on without this additional precondition. On the other hand, if the probability measures $P^\varepsilon(x, \cdot\,)$ converge in some sense as $\varepsilon \to 0$ to probability measures $P^0(x, \cdot\,)$ yielding a Markov chain X_n^0 we may view X_n^ε as perturbations of X_n^0 which generalizes models of random perturbations of deterministic transformations (see [8]). In the continuous-time case this corresponds to random perturbations of degenerate diffusions studied in Section 4.4 of [4] and in [1]. In this case X_t^ε is a diffusion generated by an operator L^ε of the form $L^\varepsilon = \varepsilon L + L_0$, where L is the same as before but L_0 now is a second-order elliptic operator whose matrix of coefficients in second derivatives may degenerate. It is not difficult to see that a version of (1.1) follows from large deviations estimates established in [4] and [2] under certain conditions on coefficients of L^ε. On the other hand, a counterexample in [2] shows that in this case large deviations estimates may fail though a kind of relation (1.1) is still valid. This is due to the fact that the probabilities $P^\varepsilon(t, x, U) = P\{X_t^\varepsilon \in U | X_0^\varepsilon = x\}$ being solutions of the equation $\partial P^\varepsilon/\partial t = L^\varepsilon P^\varepsilon$ (L^ε acts in x) behave more regularly than probabilities that the paths of X_t^ε belong to a subset of a functional space.

We have in mind also the following model considered in [1]. Suppose that $b_1, \dots, b_k$ are vector fields given on a manifold M. Next one considers a process X_t^ε governed by equations of the form

$$dX_t^\varepsilon = b_{Y(t)}(X_t^\varepsilon)\,dt + \varepsilon\tilde{b}(X_t^\varepsilon)\,dt + \varepsilon^{1/2}\sigma(X_t^\varepsilon)\,dw(t), \tag{1.2}$$

where $Y(t)$ is a time-homogeneous Markov chain with the states $\{1, \dots, k\}$ independent of the Wiener process $w(t)$, i.e., $P\{Y(t + \Delta t) = j | Y(t) = i\} = p_{ij}\,\Delta t + O(\Delta t), i \neq j$. Then the pair $(X_t^\varepsilon, Y(t))$ is a Markov process and it follows from [1] that transition probabilities $P^\varepsilon(t, (x, i), U \times \{j\}) = P\{X_t^\varepsilon \in U$ and $Y(t) = j | X_0^\varepsilon = x$ and $Y(0) = i\}$ satisfy

$$\lim_{\varepsilon \to 0} \varepsilon \log P^\varepsilon(t, (x, i), U \times \{j\}) = -\inf_{y \in U} B_t^{i,j}(x, y), \tag{1.3}$$

where

$$B_t^{i,j}(x, y) = \inf_{\varphi_0 = x,\ \varphi_t = y} \int_0^t \inf_{\gamma \in \Gamma_{0,t}(i,j)} \left\| \sigma^{-1}(\varphi_s)\big(\dot{\varphi}_s - f_{\gamma(s)}(\varphi_s)\big) \right\|^2 ds,$$

$\Gamma_{0,t}(i, j)$ is the space of possible paths of the Markov chain $Y(t)$ starting at i at time 0 and ending at j at time t, and $\{\varphi_s,\ 0 \le s \le t\}$ are absolutely continuous curves so that $\dot{\varphi}_s = d\varphi_s/ds$ are defined. If $Y(t)$ cannot pass from i to j with positive probability, i.e., $\Gamma_{0,t}(i, j)$ is empty for all $t > 0$ then the limit (1.3) equals $-\infty$, and so we must put $B_t^{i,j}(x, y) = \infty$ for any $x, y \in M$ and $t > 0$. Nevertheless, our methods will go through since $B_1^{i,j}(x, y)$ can be viewed formally as continuous on $\{1, \dots, k\} \times \{1, \dots, k\} \times M \times M$ because it is truly continuous on $M \times M$ for all i and j such that $\Gamma_{0,1}(i, j) \neq \varnothing$ and $B_1^{i,j}(x, y) = \infty$ for all $x, y \in M$ if $\Gamma_{0,1}(i, j) = \varnothing$. Anyway all forbidden passages can be disregarded, and so these infinite values will not appear in estimates.

As a genuine discrete-time example we shall mention the following model of perturbations of random transformations. Let μ be a probability measure on the space of continuous maps of M into itself. Put

$$P^{\varepsilon}(x,U) = \int Q^{\varepsilon}_{fx}(U)\, d\mu(f), \tag{1.4}$$

where a family of probability measures Q^{ε}_z satisfy uniformly in $z \in M$,

$$\lim_{\varepsilon\to 0} \varepsilon \log Q^{\varepsilon}_z(U) = -\inf_{y\in U} r(z,y), \tag{1.5}$$

for any open U, where $r(z,y) \geq 0$ is a continuous function. Then (1.1) holds true with

$$\rho(x,y) = \inf_{f\in \operatorname{supp}\mu} r(fx,y). \tag{1.6}$$

The meaning is that first we apply a random map with the distribution μ and then we perturb it independently by applying, say, a diffusion for the time ε. In the last case $r(z,y)$ will be equal to the square of the distance (corresponding to the diffusion matrix) between z and y. If $\operatorname{supp}\mu$ is just one map we obtain models of random perturbations of dynamical systems considered in [7] and [8]. The distribution μ above may also depend on ε and then, in general, we shall not have a perturbation of some limiting Markov chain corresponding to $\varepsilon = 0$ but still our results will remain applicable.

This paper has the following structure. In the next section we introduce an equivalence relation corresponding to the function $\rho(x,y)$, study the behavior of the unperturbed Markov chain X^0_n (if it can be defined) and derive a version of the Wentzell–Freidlin lower and upper bounds for probabilities to stay in tube neighborhoods. In the subsequent two sections we obtain corresponding results about the asymptotical behavior as $\varepsilon \to 0$ of invariant measures of X^{ε}_n, of the exit distribution and the mean exit time of X^{ε}_n from an open set and of the biggest eigenvalue of the transition operator of X^{ε}_n corresponding to an open set.

2. Preliminaries. Let A_N be a function on the N-fold product $M^N = M \times \cdots \times M$ defined for $\xi = (\xi_0, \dots, \xi_{N-1}) \in M^N$, $\xi_i \in M$, $i = 0, \dots, N-1$, by the formula

$$A_N(\xi) = \sum_{i=0}^{N-2} \rho(\xi_i, \xi_{i+1}) \quad \text{for } N > 1 \text{ and } A_1 \equiv 0. \tag{2.1}$$

For any pair of points $x, y \in M$ put

$$B(x,y) = \inf\{A_n(\xi) \colon n \geq 1,\ \xi = (\xi_0, \dots, \xi_{n-1}),\ \xi_0 = x,\ \xi_{n-1} = y\}. \tag{2.2}$$

The function B induces a preorder writing $y \succ_\rho x$ if $B(x,y) = 0$. This yields a ρ-equivalence relation if we write $x \sim_\rho y$ provided $x \succ_\rho y$ and $y \succ_\rho x$. A ρ-equivalence class containing $x \in M$ will be denoted by $[x]_\rho$. It will be called a basic ρ-equivalence class if either $\rho(x, x) = 0$ or $[x]_\rho$ contains more than one point.

We have the following easy fact proved in [8], pages 58 and 59.

LEMMA 2.1. *The function $B(x, y)$ is continuous in both variables, and so ρ-equivalence classes are closed sets.*

Next, we introduce a partial order among ρ-equivalence classes saying $[y]_\rho \succ_\rho [x]_\rho$ if $y \succ_\rho x$. Any maximal in this partial-order ρ-equivalence class will be called a ρ-attractor. This definition will be justified by Proposition 2.1 and Corollary 2.1 below. Since M is compact then for each $x \in M$ there exists $\varepsilon_i(x) \to 0$ such that

$$P^{\varepsilon_i(x)}(x, \cdot) \to \tilde{P}(x, \cdot) \quad \text{weakly as } i \to \infty. \tag{2.3}$$

Then, clearly, for any open set U,

$$\liminf_{\varepsilon \to 0} P^\varepsilon(x, U) \geq \tilde{P}(x, U). \tag{2.4}$$

If $U \cap \operatorname{supp} \tilde{P}(x, \cdot) \neq \varnothing$ then $\tilde{P}(x, U) > 0$, and so (1.1) together with (2.4) imply $\inf_{y \in U} \rho(x, y) = 0$. By the continuity of ρ it follows that

$$\rho(x, y) = 0 \quad \text{if } y \in \operatorname{supp} \tilde{P}(x, \cdot), \tag{2.5}$$

in particular,

$$y \succ_\rho x \quad \text{if } y \in \operatorname{supp} \tilde{P}(x, \cdot). \tag{2.6}$$

From this we conclude that for each x there exists y with $\rho(x, y) = 0$ and any ρ-attractor $[x]_\rho$ is a basic equivalence class such that if $y \succ_\rho x$ then $y \in [x]_\rho$. The existence of ρ-attractors follows from the Zorn lemma.

LEMMA 2.2. *Let $z_0, z_1, \ldots$ be an infinite sequence of points from M such that $\rho(z_k, z_{k+1}) = 0$ for all $k = 0, 1, \ldots$. Then all limit points of the sequence $z_0, z_1, \ldots$ belong to one basic equivalence class. In particular, $\bigcup_k \{z_k\}$ has a nonempty intersection with one of basic equivalence classes.*

PROOF. If the whole sequence converges to a point z then passing to the limit in $\rho(z_k, z_{k+1}) = 0$ we get $\rho(z, z) = 0$ and so $[z]_\rho$ is a basic equivalence class. Suppose now that $z_{k_i} \to z^{(1)}$ and $z_{l_i} \to z^{(2)}$ as $i \to \infty$ for some $z^{(1)} \neq z^{(2)}$. We can choose these subsequences so that $k_{i+1} > l_i > k_i$. Then $B(z_{k_i}, z_{l_i}) = 0$ and $B(z_{l_i}, z_{k_{i+1}}) = 0$. Since B is continuous then letting here $i \to \infty$ we obtain $B(z^{(1)}, z^{(2)}) = B(z^{(2)}, z^{(1)}) = 0$ and so $z^{(1)}, z^{(2)}$ belong to a basic equivalence class. □

PROPOSITION 2.1. *Let $[x]_\rho$ be a ρ-attractor having an open neighborhood $G \supset [x]_\rho$ disjoint from other basic ρ-equivalence classes except for $[x]_\rho$. Then there exists an open set $U \supset [x]_\rho$ such that for any open set $V \supset [x]_\rho$ one can find an integer $n(V) > 0$ so that for any $n \geq n(V)$ and each finite sequence $\xi = (\xi_0, \ldots, \xi_{n-1})$ satisfying $\xi_0 \in U$ and $A_n(\xi) = 0$ one has $\xi_{n-1} \in V$.*

Moreover, if $z_0, z_1, \dots$ *is an infinite sequence of points from* M *such that* $z_0 \in U$ *and* $\rho(z_k, z_{k+1}) = 0$ *for all* $k = 0, 1, \dots$ *then* $\operatorname{dist}(z_k, [x]_\rho) \to 0$ *as* $k \to \infty$.

PROOF. Remark that

$$\{y: \rho(x,y) = 0\} \subset \{y: y \succ_\rho x\} \subset [x]_\rho. \tag{2.7}$$

Put $D_\delta = \{y: B(x,y) < \delta\}$ which is an open set for each $\delta > 0$ since B is a continuous function. We claim that there exists $\delta_0 > 0$ such that $\overline{D}_{\delta_0} \subset G$, and so $\overline{D}_\delta \subset G$ for all $\delta \le \delta_0$. Indeed, if it were not true then one could choose a sequence of numbers $\delta_n \downarrow 0$ and a collection of sequences $\xi^{(n)} = (\xi_0^{(n)}, \dots, \xi_{k_n-1}^{(n)})$ with $A_{k_n}(\xi^{(n)}) \le \delta_n$ which start at points $y_n \in \xi_0^{(n)} \in [x]_\rho$ and end at points $z_n = \xi_{k_n-1}^{(n)} \notin G$. Then there would exist a subsequence n_i such that $y_{n_i} \to y \in [x]_\rho$ and $z_{n_i} \to z \notin G$, and so $B(y,z) = 0$. Hence $z \succ y \in [x]_\rho$ and by (2.7), $z \in [x]_\rho$, which is a contradiction. Thus $D_{\delta_0} \subset G$ for some $\delta_0 > 0$. Since $B(x,w) \le B(x,y) + B(y,w)$ then $y \in D_\delta$ and $B(y,w) = 0$ imply $w \in D_\delta$. In particular, if $\xi_0 \in D_\delta$ and $\xi = (\xi_0, \dots, \xi_{n-1})$ satisfies $A_n(\xi) = 0$ then $\xi_i \in D_\delta$ for all $i = 0, 1, \dots, n-1$. Now put $U = D_{\delta_0}$. Take an arbitrary open set $V \supset [x]_\rho$, $V \subset D_{\delta_0}$. We claim that there exists an integer $n(V) > 0$ such that any sequence $\xi = (\xi_0, \dots, \xi_{n-1})$ satisfying $n \ge n(V)$, $\xi_0 \in D_{\delta_0}$ and $A_n(\xi) = 0$ must have $\xi_{n-1} \in V$. Indeed, since $\bigcap_{\delta > 0} D_\delta = [x]_\rho$ we can choose $\delta(V) > 0$ such that $D_{\delta(V)} \subset V$. We shall even show that $\xi_{n-1} \in D_{\delta(V)}$ if $n \ge n(V)$ and $n(V)$ is large enough. If we were not able to choose such $n(V)$ this would mean that there exist sequences $\xi^{(n)} = (\xi_0^{(n)}, \dots, \xi_{k_n-1}^{(n)})$ with $k_n \to \infty$ as $n \to \infty$, $A_{k_n}(\xi^{(n)}) = 0$ and $\xi_i^{(n)} \in D_{\delta_0} \setminus D_{\delta(V)}$ for all $i = 0, 1, \dots, k_n - 1$. Choosing first a subsequence $n_i \to \infty$ such that $\xi_0^{(n_i)} \to z_0$, from this subsequence choosing another subsequence n_{i_j} such that $\xi_1^{(n_{i_j})} \to z_1$, etc., we will end up with an infinite sequence of points $z_k \in D_{\delta_0} \setminus D_{\delta(V)}$ satisfying $\rho(z_k, z_{k+1}) = 0$ for all $k = 0, 1, \dots$, which is impossible in view of Lemma 2.2. The last assertion of Proposition 2.1 follows, as well. □

COROLLARY 2.1. *Let* $[x]_\rho$ *and* $G \supset [x]_\rho$ *be the same as in Proposition* 2.1. *Suppose that for any* $x \in M$,

$$P^\varepsilon(x, \cdot) \to P^0(x, \cdot) \quad \textit{weakly as } \varepsilon \to 0. \tag{2.8}$$

Then there exists an open set $U \supset [x]_\rho$ *such that for any open set* $V \supset [x]_\rho$ *one can find an integer* $n(V) > 0$ *so that for any* $y \in U$ *and* $n \ge n(V)$,

$$P^0(n, y, V) = 1, \tag{2.9}$$

where $P^0(n, y, \cdot)$ *is the n-step transition probability of a Markov chain* X_n^0 *whose one-step transition probabilities are* $P^0(z, \cdot)$. *In particular, if* $X_0^0 \in U$ *then with the probability* 1,

$$\operatorname{dist}(X_n^0, [x]_\rho) \to 0 \quad \textit{as } n \to \infty. \tag{2.10}$$

PROOF. Since by (2.5), $\rho(y, z) = 0$ whenever $z \in \operatorname{supp} P^0(y, \cdot)$ then the result follows immediately from Proposition 2.1 by the Chapman–Kolmogorov formula. □

Next, we shall estimate the exit time from a neighborhood of a ρ-attractor.

LEMMA 2.3. *Let $K = [x]_\rho$ be a ρ-attractor satisfying conditions of Proposition 2.1. Then for any open set $V \supset K$ there exist numbers $r, \beta, \varepsilon_0 > 0$ such that for all $N = 1, 2, \dots$ one has*

$$P_x^\varepsilon\{\tau_{M\setminus V} < N\} < N^2 e^{-\beta/\varepsilon}, \tag{2.11}$$

provided $x \in U_r(K) = \{y\colon \operatorname{dist}(y, K) < r\}$, $0 < \varepsilon < \varepsilon_0$, where

$$\tau_W = \inf\{n\colon X_n^\varepsilon \in W\}.$$

In particular,

$$E_x^\varepsilon \tau_{M\setminus V} > \tfrac{1}{4} e^{\beta/2\varepsilon}. \tag{2.12}$$

PROOF. We shall call a δ-chain any finite sequence of points $\{z_l,\ l = 0, \dots, k\}$ such that $z_{l+1} \in W_\delta(z_l) = \{v\colon \operatorname{dist}(v, W(z_l)) \le \delta\}$, where $W(z) = \{v\colon \rho(z, v) = 0\}$.

In the same way as on page 64 of [8] we see that $P_x^\varepsilon\{\tau_{M\setminus V} < N\}$ is bounded by the sum of multiple integrals along δ-chains starting at x and ending outside V plus the expression

$$\frac{N(N-1)}{2} \sup_{z\in V} P^\varepsilon(z, M\setminus W_\delta(z)).$$

We claim that if $r, \delta > 0$ are small enough then there exists no δ-chain starting inside $U_r(K)$ and ending outside V which means that the multiple integrals in question are 0. Indeed, for otherwise we would have sequences of numbers $r_n \to 0$ and $\delta_n \to 0$ as $n \to \infty$ and a sequence of δ_n-chains $\{z_l^{(n)},\ l = 0, \dots, k_n\}$ such that $z_0^{(n)} \in U_{r_n}(K)$, $z_l^{(n)} \in V$ for all $l = 0, \dots, k_n - 1$, and $z_{k_n}^{(n)} \in M \setminus V$. Then taking a subsequence n_i so that $z_{k_{n_i}} \to y_0 \in M \setminus V$ as $i \to \infty$, from this subsequence choosing another subsequence n_{i_j} so that $z_{k_{n_{ij}}}^{(n_{ij})} - 1 \to y_{-1}$ as $j \to \infty$, etc., we shall obtain in view of Lemma 2.2 a sequence of points $\dots, y_{-2}, y_{-1}, y_0$ such that $y_0 \in M \setminus V$, $\rho(y_l, y_{l+1}) = 0$ for all $l = -1, -2, -3, \dots$, and $\operatorname{dist}(y_l, K) \to 0$ as $l \to -\infty$. Then it will follow that $y_0 \succ x$, which is impossible since $[x]_\rho$ is a ρ-attractor.

Next, it remains to estimate $\sup_z P^\varepsilon(z, M \setminus W_\rho(z))$. Since, clearly

$$\inf_{z\in M} \inf_{v\in M\setminus W_\rho(z)} \rho(z, v) = \gamma(\delta) > 0, \tag{2.13}$$

then by (1.1),

$$\sup_z P^\varepsilon\big(z, M \setminus W_\rho(z)\big) \le e^{-\gamma(\delta)/2\varepsilon}, \tag{2.14}$$

provided $\varepsilon > 0$ is small enough. This yields (2.11). We obtain (2.12) noting that

$$E_x^\varepsilon \tau_{M\setminus V} \geq NP_x^\varepsilon\{\tau_{M\setminus V} > N\} \geq N(1 - N^2 e^{-\beta/\varepsilon})$$

for N of order $\frac{1}{3}e^{\beta/2\varepsilon}$. □

Next, one obtains a version of the Wentzell–Freidlin key lower and upper bounds of the probability for Markov chains X_n^ε to stay in a small tube near a fixed sequence of points as in Theorem 1.5.2 and Corollary 1.5.2 of [8].

We shall also need the following lemma.

LEMMA 2.4. *Let K be a compact subset of M which does not contain entirely any infinite sequence of points $z_0, z_1, z_2, \dots$ satisfying $\rho(z_k, z_{k+1}) = 0$ for all $k = 0, 1, \dots$. Then there exist numbers $a = a(K) > 0$ and $N = N(K) > 0$ such that*:

(i) *For any sequence $\xi = (\xi_0, \dots, \xi_{n-1})$ with $n > N$ and $\xi_i \in K_i$, $i = 0, \dots, n-1$, one has $A_n(\xi) > (n - N)a$.*

(ii) *There exists $\varepsilon_0 > 0$ such that for any $n > N$,*

$$P_x^\varepsilon\{\tau_{M\setminus K} > n\} \leq e^{-[(n-N)/\varepsilon]a}, \tag{2.15}$$

provided $x \in K$ and $0 < \varepsilon < \varepsilon_0$, where $\tau_V = \inf\{m > 0: X_m^\varepsilon \in V\}$.

PROOF. We claim that there exists an integer $N_1 > 0$ such that any sequence $\xi = (\xi_0, \dots, \xi_{n-1})$ with $A_n(\xi) = 0$ and $\xi_i \in K$ for all $i = 0, \dots, n-1$ must contain less than N_1 points. Indeed, for otherwise we would have an infinite collection of sequences $\xi^{(l)} = (\xi_0^{(l)}, \dots, \xi_{k_l-1}^{(l)})$ with $k_l \to \infty$ as $l \to \infty$, $A_{k_l}(\xi^{(l)}) = 0$ and $\xi_i^{(l)} \in K$ for all $i = 0, \dots, k_l - 1$. Since K is compact we could choose then similarly to the end of the proof of Proposition 2.1 an infinite sequence of points $z_0, z_1, \dots$ from K satisfying $\rho(z_k z_{k+1}) = 0$ for all $k = 0, 1, \dots$, which contradicts the assumption on K. The rest of the proof is the same as on pages 73 and 74 of [8], where one has to replace orbits of a map F by sequences of points $\{z_k\}$ satisfying $\rho(z_k, z_{k+1}) = 0$. □

3. Invariant measures. In this section we shall study the asymptotic behavior as $\varepsilon \to 0$ of invariant measures of the Markov chains X_n^ε, i.e., of the probability measures μ^ε on M satisfying

$$\mu^\varepsilon(\Gamma) = \int_M d\mu^\varepsilon(x) P^\varepsilon(x, \Gamma), \tag{3.1}$$

for any Borel set $\Gamma \subset M$. We shall employ the following well-known result (see [10], Proposition 5, and [8], pages 70 and 71).

COROLLARY 3.1. *Let X_n be a Markov chain in a measurable space $(M, \mathscr{B})$ with transition probabilities $P(x, \Gamma)$ having an invariant probability measure*

μ. Let $V \subset M$ be a measurable set such that

$$\sup_{x\in M} E_x \tau_V < \infty, \tag{3.2}$$

where $\tau_V = \inf\{n > 0\colon X_n \in V\}$. Then $\mu(V) > 0$ and we can define another Markov chain ${}^V X_n$ (called the induced Markov chain) on V by its transition probabilities ${}^V P(x, \Gamma)$, $x \in V$, having the form

$${}^V P(x, \Gamma) = P_x\{X_{\tau_V} \in \Gamma\}, \tag{3.3}$$

where Γ is a measurable subset of V and $P_x\{\ \}$ denotes the probability for the Markov chain X_n starting at x. Then the restriction μ_V of $(\mu(V))^{-1}\mu$ to V is the probability invariant measure of the Markov chain ${}^V X_n$ and for any measurable set $G \subset M$,

$$\begin{aligned} \mu(G) &= \mu(V)\int_V d\mu_V(x) E_x \sum_{k=0}^{\tau_V - 1} \chi_G(X_k) \\ &= \int_V d\mu(x) E_x \sum_{k=0}^{\tau_V - 1} \chi_G(X_k), \end{aligned} \tag{3.4}$$

which gives the representation of μ via μ_V, where χ_G denotes the indicator of a set G.

Remark 3.1. The existence of an invariant measure for X_n will follow if, for instance, M is compact and the measures $P(x, \cdot\,)$ depend continuously on x in the weak topology or if these measures have positive densities with respect to a fixed measure.

Next, we proceed similarly to the original paper of Wentzell and Freidlin [11]. The arguments below will rely on the following assumption.

Assumption 3.1. There exists only a finite number of basic ρ-equivalence classes $K_1, \dots, K_\nu$.

By Lemma 2.1 $K_1, \dots, K_n$ are compact. Let V_i be open sets such that

$$K_i \subset V_i \subset U_r(K_i) = \{y\colon \operatorname{dist}(y, K_i) < r\}. \tag{3.5}$$

We shall always take $r > 0$ above to be small enough so that V_i, $i = 1, \dots, \nu$, will be disjoint. Denote $V = \bigcup_{1 \le i \le \nu} V_i$ and consider the Markov chain ${}^V X_n^\varepsilon$ introduced in the same way as in Proposition 3.1 by means of transition probabilities ${}^V P^\varepsilon(x, \Gamma) = P_x^\varepsilon\{X_{\tau_V}^\varepsilon \in \Gamma\}$, where $\tau_V = \inf\{n > 0\colon X_n^\varepsilon \in V\}$ and Γ is a Borel subset of V. In view of Lemma 2.4(ii) it is clear that (3.2) will then be satisfied and so Proposition 3.1 is applicable. Since K_i and K_j are equivalence classes the value $B(x, y)$ defined by (2.2) remains the same for all $x \in K_i$ and $y \in K_j$, and it will be denoted by B_{ij}. Clearly, if $i \ne j$ then at least one of the numbers B_{ij} and B_{ji} is positive. It is clear from the definition that K_i is a ρ-attractor if and only if $B_{ij} > 0$ for any $j \ne i$.

Next, one obtains key bounds for the transition probabilities of the Markov chain ${}^V X_n^\varepsilon$ when ε is small in the form

(3.6) $$\exp(-(B_{kl}+\beta)/\varepsilon) < {}^V P^\varepsilon(N,x,V_l) < \exp((-B_{kl}+\beta)/\varepsilon),$$

provided $x \in V_k$, $0 < \varepsilon < \varepsilon_0$ *and* $1 \le k, l \le \nu$.

The proof of these bounds in our case repeats verbatim the proof of Lemma 1.5.4 on pages 75–80 of [8] for the case of random perturbations of a map F. The only change one has to do is to replace orbits of the map F appearing on pages 77 and 79 by sequences of points $\{z_k\}$ such that $\rho(z_k, z_{k+1}) = 0$ for all k.

Let L be a finite set, whose elements will be denoted by the letters i, j, k, m, n, etc. Given $i \in L$, a graph consisting of arrows $m \to n$ ($m \ne i$, $m, n \in L$, $n \ne m$) is called an i-graph if it satisfies the following conditions: Every point $m \ne i$ is the origin of exactly one arrow, and the graph has no cycles.

Let $L = \{1, \dots, \nu\}$, $i \in L$,

$$B(i) = \min_{g \in G(i)} \sum_{(m \to n) \in g} B_{mn}$$

and

$$L_{\min} = \left\{ i \in L \colon B(i) = \min_{j \in L} B(j) \right\}.$$

Now we can formulate the main result of this section.

THEOREM 3.1. *If* $i \in L_{\min}$ *then* K_i *is a* ρ*-attractor. Let* $\Gamma \subset M$ *be a closed set disjoint with* $\bigcup_{i \in L_{\min}} K_i$. *Then any invariant probability measures* μ^ε *of the Markov chain* X_n^ε *satisfy*

(3.7) $$\lim_{\varepsilon \to 0} \mu^\varepsilon(\Gamma) = 0,$$

and so any weak limit of measures μ^ε *as* $\varepsilon \to 0$ *has support in* $\bigcup_{i \in L_{\min}} K_i$.

PROOF. After preparations of this and the previous sections the proof of this theorem proceeds verbatim as the proof of Theorem 1.5.4 on pages 83 and 84 in [8] from showing that any K_i, $i \in L_{\min}$, is a ρ-attractor until formula (1.5.51) which asserts that the invariant measure $\mu_V^\varepsilon = (\mu^\varepsilon(V))^{-1}\mu^\varepsilon$ of the Markov chain ${}^V X_n^\varepsilon$ satisfies

(3.8) $$(\mu^\varepsilon(V))^{-1} \mu^\varepsilon\left(\bigcup_{j \notin L_{\min}} V_j \right) < e^{-\gamma/e},$$

for some $\gamma > 0$ and $\varepsilon > 0$ small enough, and so

(3.9) $$(\mu^\varepsilon(V))^{-1} \mu^\varepsilon\left(\bigcup_{i \in L_{\min}} V_i \right) \to 1 \quad \text{as } \varepsilon \to 0.$$

It remains to show that

(3.10) $$\mu^\varepsilon(M \setminus V) \to 0 \quad \text{as } \varepsilon \to 0.$$

Since any K_i, $i \in L_{\min}$, is a ρ-attractor then by Lemma 2.3 we can choose $r, \beta > 0$ so that

(3.11) $$P_x^\varepsilon\{\tau_{M \setminus V_i} < N\} < N^2 e^{-\beta/\varepsilon},$$

for any $x \in U_r(K_i)$, $i \in L_{\min}$, all $\varepsilon > 0$ small enough, and each $N = 1, 2, \dots$. Denote $\tilde{V}_i = U_r(K_i)$, $i = 1, \dots, \nu$, and $\tilde{V} = \bigcup_{1 \le i \le \nu} \tilde{V}_i$. By Lemma 2.4(ii) there exist $\tilde{N} = N(M \setminus \tilde{V}) + 1$ and $a > 0$ such that

$$P_x^\varepsilon\{\tau_{\tilde{V}} > n\} \le e^{-[(n-\tilde{N})/\varepsilon]a}, \tag{3.12}$$

for any $x \in M$ and $n > \tilde{N}$. Finally, by (3.4), (3.8), (3.11) and (3.12) for $\varepsilon > 0$ small enough

$$\begin{aligned} \mu^\varepsilon(M \setminus V) &= \int_{\tilde{V}} d\mu^\varepsilon(x) E_x^\varepsilon \sum_{k=0}^{\tau_{\tilde{V}}-1} \chi_{M \setminus V}(X_k^\varepsilon) \\ &\le \sum_{i \in L_{\min}} \int_{V_i} d\mu^\varepsilon(x) E_x^\varepsilon \sum_{k=0}^{\tilde{N}+1} \chi_{M \setminus V}(X_k^\varepsilon) + \tfrac{1}{2} e^{-a/3} + e^{-\gamma/\varepsilon}(\tilde{N} + 2) \\ &\le \nu(\tilde{N} + 1)^3 e^{-\beta/\varepsilon} + \tfrac{1}{2} e^{-a/\varepsilon} + (\tilde{N} + 2) e^{-\gamma/\varepsilon}, \end{aligned} \tag{3.13}$$

proving (3.10). A more careful analysis enables one to get more precise estimates of $\mu^\varepsilon(M \setminus \bigcup_{i \in L_{\min}} V_i)$ the same as in Theorem 4.1 on page 186 of [5]. □

We obtained Theorem 3.1 without assuming that the Markov chains X_n^ε are perturbations of some other Markov chain X_n^0, but if it is the case then under the condition below all weak limits of μ^ε as $\varepsilon \to \infty$ turn out to be invariant measures of X_n^0 and so Theorem 3.1 describes support of such measures.

PROPOSITION 3.2. *Suppose that for any continuous function f on M,*

$$\limsup_{\varepsilon \to 0} \sup_x \left| \int_M P^\varepsilon(x, dy) f(y) - \int_M P^0(x, dy) f(y) \right| = 0, \tag{3.14}$$

where $P^0(x, \cdot)$, $x \in M$, is a family of probability measures on M continuously dependent on x in the weak topology of measures. Then any weak limit as $\varepsilon \to 0$ of invariant measures of Markov chains X_n^ε with transition probabilities $P^\varepsilon(x, \cdot)$ is an invariant measure of the Markov chain X_n^0 with transition probabilities $P^0(x, \cdot)$.

PROOF. Suppose that $\mu^{\varepsilon_i} \to_w \mu$ then for any continuous function f on M,

$$\begin{aligned} &\left| \int f(x)\, d\mu(x) - \iint f(y) P^0(x, dy)\, d\mu(x) \right| \\ &\quad \le \left| \int f d\mu - \int f d\mu^{\varepsilon_i} \right| \\ &\qquad + \int \left| \int f(y) P^{\varepsilon_i}(x, dy) - \int f(y) P^0(x, dy) \right| d\mu^{\varepsilon_i}(x) \\ &\qquad + \left| \iint f(y) P^0(x, dy)\, d\mu^{\varepsilon_i}(x) - \iint f(y) P^0(x, dy)\, d\mu(x) \right| \\ &\quad \to 0 \quad \text{as } \varepsilon_i \to 0, \end{aligned} \tag{3.15}$$

in view of (3.14) and the fact that $\int f(y)P^0(x, dy)$ is a continuous function in x. Thus

$$\int f(x)\, d\mu(x) = \iint f(y) P^0(x, dy)\, d\mu(x), \tag{3.16}$$

for any continuous function f, and so μ is an invariant measure of the Markov chain X_n^0. □

Next, we shall see how our discrete-time results imply the corresponding continuous-time results from [11]. Wentzell and Freidlin dealt with the asymptotic behavior of invariant measures of diffusion-type random perturbations. This model considered on a smooth Riemannian manifold M leads to a diffusion Markov process X_t^ε generated by operators $L^\varepsilon = \varepsilon L + b$, where L is an elliptic second-order differential operator and b is a vector field. This means that transition probabilities $P^\varepsilon(t, x, \Gamma)$ satisfy the parabolic equation $\partial P^\varepsilon/\partial t = L^\varepsilon P^\varepsilon$ with the initial condition $P^\varepsilon|_{t=0} = \chi_\Gamma$. The Markov processes X_t^ε are viewed as random perturbations of a flow F^t solving the ordinary differential equation

$$\frac{dF^t x}{dt} = b(F^t x), \qquad F^0 x = x.$$

We will not discuss here the specific features of such random perturbations since the only fact we will need is the following property of transition probabilities similar to (1.1):

$$\lim_{\varepsilon \to 0} \varepsilon \log P^\varepsilon(t, x, U) = -\inf_{y \in U} B_t(x, y), \tag{3.17}$$

for any $x \in M$ and an open set U, where

$$\begin{aligned} B_t(x,y) &= \inf_{\varphi_0 = x,\ \varphi_t = y} A_t(\varphi) \\ &= \inf_{\varphi_0 = x,\ \varphi_y = u} \int_0^t \|b(\varphi_s) - \dot{\varphi}_s)\|^2\, ds, \end{aligned} \tag{3.18}$$

where the infimum is taken over absolutely continuous curves φ_s, $0 \le s \le t$, on M starting at x and ending at y, $\dot{\varphi}_s = d\varphi_s/ds$ denotes the tangent (speed) vector to φ_s, and $\|\cdot\|$ denotes certain Riemannian norm in the tangent bundle constructed by means of diffusion coefficients of X_t^ε. The relation (3.17) follows from more general results which can be found in Chapter 4 of [11] and in Chapter 14 of [6] but can be proved now also directly by the PDE viscosity solutions methods.

If we apply our theory to $F = F^1$ and X_t^ε considered only for integer $t = 0, 1, 2, \dots$, then the results concerning invariant measures will remain valid for the continuous-time process X_t^ε since the invariant measures of X_t^ε will be, of course, invariant with respect to X_n^ε. The only fact needed to be checked is the coincidence of Assumption 3.1 with the corresponding assumption formulated by Wentzell and Freidlin for the continuous-time case and that the numbers B_{ij} will be the same both for the discrete- and continuous-time cases. In the continuous-time case one calls x and y equivalent (written $x \sim y$) if and only if $\inf_{t \geq 0} B_t(x, y) = \inf_{t \geq 0} B_t(y, x) = 0$. Our definition of the equivalence relation which we will denote here by $\sim_1$ corresponds to the case when the above infimum is taken only over integers; $x \sim_1 y$ if and only if $\inf_{\text{integer } n \geq 0} B_n(x, y) = \inf_{\text{integer } n \geq 0} B_n(y, x) = 0$. Denote the equivalence classes containing a point x and corresponding to $\sim$ and $\sim_1$ by $[x]$ and $[x]^{(1)}$, respectively. Then one has the following result proved on pages 88 and 89 of [8].

PROPOSITION 3.3. *For any* $x \in M$, $[x] = [x]^{(1)}$.

It remains to establish the following proposition.

PROPOSITION 3.4. *Let K be a basic equivalence class. Then for any pair of points $x, y \in M$ such that either $x \in K$ or $y \in K$ one has*

$$\inf_{\text{integer } n \geq 0} B_n(x, y) = \inf_{t \geq 0} B_t(x, y). \tag{3.19}$$

PROOF. First, it is obvious that the above expression does not depend on the choice of the point in K. Clearly,

$$\inf_{\text{integer } n \geq 0} B_n(x, y) \geq \inf_{t \geq 0} B_t(x, y) = \tilde{B}. \tag{3.20}$$

It is easy to see that there exist a sequence of numbers $t_n \to \infty$ and a sequence of piecewise smooth curves $\varphi_s^{(n)}$, $0 \leq s \leq t_n$, $\varphi_0^{(n)} = x$, $\varphi_{t_n}^{(n)} = y$ such that

$$A_{t_n}(\varphi^{(n)}) \to \tilde{B} \quad \text{as } t_n \to \infty. \tag{3.21}$$

Define new curves $\psi_s^{(n)} = \varphi_{st_n([t_n]+1)^{-1}}^{(n)}$, where $[\cdot]$ denotes the integral part. Then

$$\begin{aligned} B_{[t_n]+1}(x, y) &\leq \int_0^{[t_n]+1} \left\| b(\psi_s^{(n)}) - \dot{\psi}_s^{(n)} \right\|^2 ds \\ &= \int_0^{[t_n]+1} \left\| b\left(\varphi_{st_n([t_n]+1)^{-1}}^{(n)}\right) - \frac{t_n}{[t_n]+1} \dot{\varphi}_{st_n([t_n]+1)^{-1}}^{(n)} \right\|^2 ds \\ &= \frac{[t_n]+1}{t_n} \int_0^{t_n} \left\| b(\varphi_u^{(n)}) - \frac{t_n}{[t_n]+1} \dot{\varphi}_u^{(n)} \right\|^2 du \\ &\leq \frac{(1+\alpha)t_n}{[t_n]+1} A_{t_n}(\varphi^{(n)}) + \frac{(1+1/\alpha)}{t_n([t_n]+1)} \int_0^{t_n} \| b(\varphi_u^{(n)}) \|^2 du \\ &\leq (1+\alpha) A_{t_n}(\varphi^{(n)}) + \left(1 + \frac{1}{\alpha}\right) t_n^{-1} \sup_x \| b(x) \|^2, \end{aligned} \tag{3.22}$$

for any $\alpha > 0$, where we used the inequality $2(\xi, \zeta) \leq \alpha \|\xi\|^2 + (1/\alpha)\|\zeta\|^2$ for any pair of vectors ξ and ζ. Now letting $t_n \to \infty$ and noting that the left-hand side of (3.20) does not exceed the left-hand side of (3.22), we derive in view of (3.21) that

$$\inf_{\text{integer } n \geq 0} B_n(x, y) \leq (1 + \alpha)\tilde{B}.$$

Since $\alpha > 0$ is arbitrary this together with (3.23) yields (3.19). □

REMARK 3.2. The same arguments produce the continuous-time result from its discrete-time counterpart for the more general case (1.2) and (1.3) described in Introduction, as well, as for other action functionals of similar structure.

4. Exit problems. In this section M will be a compact subset of a locally compact metric space S such that M coincides with the closure of its interior int M.

Let X_n^ε, $\varepsilon > 0$, $n = 0, 1, \dots,$ be a family of Markov chains on S with Borel transition probabilities $P^\varepsilon(x, \cdot\,)$, $x \in S$, Borel measurable in x and such that for any open set $U \subset S$ uniformly in $x \in \text{int}\, M$,

$$\lim_{\varepsilon \to 0} \varepsilon \log P^\varepsilon(x, U) = -\inf_{y \in U} \rho(x, y), \tag{4.1}$$

where $\rho(x, y) \geq 0$ defined on $M \times S$ and for some open set $W \supset M$ with a compact closure $\overline{W}$ the function $\rho(x, y)$ is continuous on $M \times \overline{W}$ and $\rho(x, y) = \infty$ for $x \in M$ and $y \notin \overline{W}$. This last condition can be substituted by

$$\sup_{x \in M,\, y \in W} \rho(x, y) < \inf_{x \in M,\, y \notin \overline{W}} \rho(x, y). \tag{4.2}$$

In this section we will study the distribution of the exit points $X^\varepsilon_{\tau_{\overline{M^c}}}$ from int M, where $M^c = S \setminus M$, and the expectation of $\tau_{\overline{M^c}}$. The main results will be obtained under the condition

$$\inf_{x \in M,\, y \in W \setminus M} \rho(x, y) > 0, \tag{4.3}$$

which in the case of random perturbations of dynamical systems corresponds to perturbations of transformations whose orbits enter the set int M.

Let A_N be a function on the N-fold product $M^{N-1} \times S = M \times \cdots \times M \times S$ defined by formula (2.1) for any sequence $\xi = (\xi_0, \dots, \xi_{N-1})$ with $\xi_i \in M$ if $i = 0, \dots, N-2$ and $\xi_{N-1} \in S$. For any pair of points $x \in M$ and $y \in S$ we define $B(x, y)$ by (2.2), where the infimum is taken over all sequences $\xi = (\xi_0, \dots, \xi_{n-1})$ with $\xi_0 = x$, $\xi_{n-1} = y$ and $\xi_i \in M$ for all $i = 1, 2, \dots, n-2$. By the continuity of the function ρ the value of $B(x, y)$ will not change if this infimum is taken over sequences $\xi = (\xi_0, \dots, \xi_{n-1})$ with $\xi_0 = x$, $\xi_{n-1} = y$ and $\xi_i \in \text{int}\, M$ for $i = 1, \dots, n-2$. In the same way as in Section 2 the function $B(\cdot, \cdot)$ induces a preorder and a partial-order $\succ_\rho$ and a ρ-equivalence relation $\sim_\rho$. The definitions of ρ-equivalence classes, basic ρ-equivalence classes and ρ-attractors remain the same as in Section 2. The conditions (4.2) and (4.3) ensure that if $x \in M$ and $y \succ_\rho x$ then $y \in \text{int}\, M$. Moreover, by (4.3) and the continuity of ρ there exists $\delta > 0$ such that

$$\inf\{\rho(x,y): x \in M, y \notin M \setminus U_\delta(\partial M)\} \geq \delta, \tag{4.4}$$

where $\partial M = M \setminus \operatorname{int} M$ and $U_\delta(V) = \{z: \operatorname{dist}(z,V) < \delta\}$. Thus if $y \succ_\rho x \in M$ then $y \in M \setminus U_\delta(\partial M)$ and so all basic ρ-equivalence classes must be contained in $M \setminus U_\delta(\partial M)$. We shall work under the following assumption.

ASSUMPTION 4.1. There exists only a finite number of basic ρ-equivalence classes $K_1, \dots, K_\nu$ in M.

Since by Lemma 2.1 $K_1, \dots, K_\nu$ are compacts then they stay on positive distance from each other and from ∂M. Thus we can pick up disjoint open sets $V_i \subset \operatorname{int} M$ such that (3.5) holds true. We shall denote again by B_{ij} the value $B(x,y)$ which is the same for all $x \in K_i$ and $y \in K_j$, and introduce also the following notation: B_{iy} for $B(x,y)$ with $x \in K_i$, B_{xj} for $B(x,y)$ with $y \in K_j$,

$$B_{i\partial} = \inf_{y \in M^c} B_{iy} \quad \text{and} \quad B_{x\partial} = \inf_{y \in M^c} B(x,y). \tag{4.5}$$

In view of (4.2) both infimuma in (4.5) are attained at points of $\overline{W \setminus M}$.

We remark that under our conditions the exit time $\tau_{\overline{M}^c}$ from int M is finite with probability 1 and, moreover, its expectation is finite, as well. Indeed, if $L = \sup_{x \in M,\, y \in W} \rho(x,y)$ then by (4.1) if $\varepsilon > 0$ is small enough,

$$P^\varepsilon(x, W \setminus M) \geq e^{-2L/\varepsilon}, \tag{4.6}$$

for any $x \in M$. Thus, by the Markov property

$$P_x^\varepsilon\{\tau_{\overline{M}^c} > n\} \leq (1 - e^{-2L/\varepsilon})^n, \tag{4.7}$$

and so

$$E_x^\varepsilon \tau_{\overline{M}^c} \leq e^{2L/\varepsilon}. \tag{4.8}$$

Later we will obtain a more precise estimate of this expectation.

Denote by ∂_i the set of points $y \in M^c$ for which $B_{i\partial} = B_{iy}$. In view of the remark after (4.5), $\partial_i \subset \overline{W \setminus M}$ and it is a closed set. By Lemma 2.2 any infinite sequence of points $\mathscr{z} = (z_0, z_1, \dots)$, $z_k \in M$, $\rho(z_k, z_{k+1}) = 0$, $k = 0, 1, \dots,$ attracts to one of K_i whose index we denote by $i(\mathscr{z})$. For any $x \in M$ we denote by $I(x)$ the set of indices $i(\mathscr{z})$ for all $\mathscr{z}$ starting at x. Consider the set $G(\partial)$ of graphs with vertices in the set $L = \{1, 2, \dots, \nu, \partial\}$ consisting of exactly one arrow emanating from each vertex except for ∂ and having no cycles. Among such graphs we choose those at which the minimum

$$B = \min_{g \in G(\partial)} \sum_{(\alpha \to \beta) \in g} B_{\alpha\beta} \tag{4.9}$$

is attained. In each of them we consider the chain of arrows leading from i to ∂. Let $j \to \partial$ be the last arrow in this chain. The set of all these j in all chosen above graphs is denoted by $R(i)$.

THEOREM 4.1. *For each* $x \in \operatorname{int} M$ *and any open neighborhood* U *of the set* $\partial(x) = \bigcup_{i \in I(x)} \bigcup_{j \in R(i)} \partial_j$ *one has*

$$P_x^\varepsilon\{X_{\tau_{\overline{M}^c}}^\varepsilon \in U\} \to 1 \quad \text{as } \varepsilon \to 0. \tag{4.10}$$

Similarly to Theorem 9.1 of [11] and Section 5 of Chapter 6 in [5], the proof of this theorem relies upon the study of the induced Markov chain $\tilde{X}_n^\varepsilon$ on

$$\tilde{V} = \Big(\bigcup_{1 \le i \le \nu} V_i \Big) \cup (U_1 \cap \overline{M}^c) \cup (U \cap \overline{M}^c \setminus U_1) \cup (\overline{M}^c \setminus U), \tag{4.11}$$

which stops at the arrival to $\overline{M}^c$, where $U_1 \supset \partial(x)$ is an open set such that $\overline{U}_1 \subset U$. The one-step transition probabilities of $\tilde{X}_n^\varepsilon$ have the form

$$\tilde{P}^\varepsilon(x, \Gamma) = P_x^\varepsilon\{X_{\tau_{\tilde{V}}}^\varepsilon \in \Gamma\},$$

if $x \in V = \bigcup_i V_i$. for $x \in \tilde{V} \setminus V$ we put $\tilde{P}^\varepsilon(x, \{x\}) = 1$. For an appropriate N the N-step transition probabilities $\tilde{P}^\varepsilon(N, x, V_l)$ can be estimated by formula (3.6) if $x \in V_l$. Similarly, one can show that

$$\exp(-B_{k\partial} + \beta)/\varepsilon) < \tilde{P}^\varepsilon(N, x, U_1 \cap \overline{M}^c) < \exp((-B_{k\partial} + \beta)/\varepsilon) \tag{4.12}$$

and

$$\tilde{P}^\varepsilon(N, x, \overline{M}^c \setminus U) < \exp(-(B_{k\partial} + \gamma)/\varepsilon), \tag{4.13}$$

for $x \in V_k$, where $\beta > 0$ can be made much smaller than $\gamma > 0$ for appropriately chosen U_1 and V_i, $i = 1, \dots, \nu$. To derive Theorem 4.1 from estimates (3.6), (4.12) and (4.13), one needs certain results about Markov chains proved in [11], Lemma 7.3, and [5], Lemma 3.3 of Chapter 6.

After this result the remainder of the proof of Theorem 4.1 is easy and it proceeds in the same way as in Section 9 of [11] and in Section 5 of Chapter 6 in [5] with simplifications due to the discrete time. The details are left to the reader.

Let $G(x \nrightarrow \partial)$ denote the set of oriented graphs without cycles on the set $L = \{1, \dots, \nu, x, \partial\}$ consisting of ν arrows $\alpha \to \beta$ and not containing chains of arrows leading from x to ∂. Put

$$B(x) = \min_{g \in G(x \nrightarrow \partial)} \sum_{(\alpha \to \beta) \in G} B_{\alpha\beta}. \tag{4.14}$$

The following result can be proved in the same way as Theorem 5.3 in Chapter 6 of [5].

THEOREM 4.2. *Uniformly in* x *belonging to any compact subset of* $\operatorname{int} M$ *one has*

$$\lim_{\varepsilon \to 0} \varepsilon \log E_x^\varepsilon \tau_{\overline{M}^c} = B - B(x), \tag{4.15}$$

where B *is defined by* (4.9).

If the Markov chain X_n^ε is a diffusion process X_t^ε considered only at integer $t = 0, 1, 2, \dots$ as described in the end of Section 3 with (3.17) and (3.18) satisfied, then by Proposition 3.4 the functions $B(x, y)$ are the same whether the infimum of $B_t(x, y)$ is taken over nonnegative integers or nonnegative reals provided x or y belongs to a basic equivalence class. This implies that the corresponding numbers $B_{\alpha\beta}$, B and $B(x)$ defined above will be also the same for both cases yielding that the asymptotical behavior as $\varepsilon \to 0$ of the exit distribution and the mean exit time will be the same whether one considers X_t^ε for all $t \geq 0$ or only for integer $t \geq 0$.

Next, we shall discuss the eigenvalue problem. Suppose in addition to (4.1)–(4.3) and Assumption 4.1 that

$$P^\varepsilon(x, \operatorname{int} M) = 0 \quad \text{for any } x \notin \operatorname{int} M. \tag{4.16}$$

Then the operator P^ε acting on bounded Borel functions f on S by the formula

$$P_\varepsilon f(x) = \int_S f(y) P^\varepsilon(x, dy) \tag{4.17}$$

transforms the space $\mathscr{F}_0(M)$ of bounded Borel functions on S which are 0 outside of int M into itself. If $\|\cdot\|$ is the sup-norm on $\mathscr{F}_0(M)$ then the limit

$$\lambda^\varepsilon = \lim_{n \to \infty} \frac{1}{n} \log \|P_\varepsilon^n\| \tag{4.18}$$

exists by the standard subadditivity argument, $\lambda^\varepsilon \leq 0$, and e^{λ^ε} is the spectral radius of P_ε. It is easy to see that λ^ε can be obtained in the following way:

$$\lambda^\varepsilon = \lim_{n \to \infty} \frac{1}{n} \log\Big(\sup_{x \in \operatorname{int} M} P_x^\varepsilon\{\tau_{\overline{M}^c} > n\}\Big), \tag{4.19}$$

and so by (4.7), $\lambda^\varepsilon < 0$. If $P^\varepsilon(x, \partial M) = 0$ for all $x \in \operatorname{int} M$ and $P^\varepsilon(x, \cdot)$ depends continuously on x in the topology of weak convergence, then one can replace $\mathscr{F}_0(M)$ by the space $\mathscr{C}_0(M)$ of continuous functions which are 0 outside of int M. In this case the operator P_ε is completely continuous and e^{λ^ε} is the absolute value of its principal eigenvalue. If P_ε is taken from a semigroup generated by an elliptic operator L^ε then λ^ε itself is the principal eigenvalue of L^ε, i.e., its eigenvalue with the biggest real part.

Adapting methods of [10] and [9] to our discrete-time framework in the spirit of this paper one derives the following theorem.

THEOREM 4.3. *Suppose that* (4.1)–(4.3), (4.16) *and Assumption* 4.1 *hold. Then*

$$\lim_{\varepsilon \to 0} \varepsilon \log(-\lambda^\varepsilon) = -(B^{(1)} - B^{(2)}), \tag{4.20}$$

where

$$B^{(k)} = \min_{g \in G(k)} \sum_{(\alpha \to \beta) \in g} B_{\alpha\beta}, \qquad k = 1, 2,$$

and $G(k)$ is the set of oriented graphs without cycles with vectors in $L = \{1, \dots, \nu, \partial\}$ consisting of one chain of $(\nu - k + 1)$ arrows.

REMARK 4.1. Suppose that we replace (4.2), (4.3) and Assumption 4.1 by the condition that M does not contain any basic ρ-equivalence classes, and so by Lemmas 2.2 and 2.5 there exists N such that for each $x \in M$ one can find a sequence of points $z_0, z_1, \dots, z_n$ with $n \le N$ such that $z_0 = x$, $z_n \notin M$ and $\rho(z_k, z_{k+1}) = 0$ for all $k = 0, \dots, n-1$. In this case similarly to Theorem 7.1 of Chapter 6 in [5] one can show that $\lambda^\varepsilon \to -\infty$ as $\varepsilon \to 0$ and

$$\lim_{\varepsilon \to 0} \varepsilon \lambda^\varepsilon = -\lim_{N \to \infty} N^{-1} \min\{A_N(\xi) : \xi = (\xi_0, \dots, \xi_{N-1}) \in M^N\}. \tag{4.21}$$

REFERENCES

[1] BEZUIDENHOUT, C. (1987). A large deviations principle for small perturbations of random evolution equations. *Ann. Probab.* **15** 646–658.

[2] BEZUIDENHOUT, C. (1987). Singular perturbations of degenerate diffusions. *Ann. Probab.* **15** 1014–1043.

[3] FLEMING, W. H. and SOUGANIDIS, P. E. (1986). PDE-viscosity solution approach to some problems of large deviations. *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **13** 171–192.

[4] FREIDLIN, M. (1985). *Functional Integration and Partial Differential Equations.* Princeton Univ. Press.

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

[6] FRIEDMAN, A. (1986). *Stochastic Differential Equations and Applications.* Academic, New York.

[7] KIFER, Y. I. (1974). On small random perturbations of diffeomorphisms. *Uspehi Mat. Nauk* **29** 173–174. (In Russian.)

[8] KIFER, YU. (1988). *Random Perturbations of Dynamical Systems.* Birkhäuser, Boston.

[9] WENTZELL, A. D. (1972). On the asymptotic behavior of the greatest eigenvalue of a second-order elliptic differential operator with a small parameter in the high derivatives. *Soviet Math. Dokl.* **13** 13–17.

[10] WENTZELL, A. D. (1973). Formulae for eigenfunctions and eigenmeasures associated with a Markov process. *Theory Probab. Appl.* **18** 1–26.

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

INSTITUTE OF MATHEMATICS
THE HEBREW UNIVERSITY
JERUSALEM
ISRAEL