The Annals of Probability
1987, Vol. 15, No. 3, 1014–1043

SINGULAR PERTURBATIONS OF DEGENERATE DIFFUSIONS

BY CAROL BEZUIDENHOUT

University of Minnesota

Ventsel and Freidlin studied small random perturbations $dx^\varepsilon(t) = b(x^\varepsilon(t))\,dt + \sqrt{\varepsilon}\,\sigma(x^\varepsilon(t))\,d\beta(t)$ of the dynamical system $dx(t) = b(x(t))\,dt$. They proved a large deviations theorem for the family $\{x^\varepsilon(t)\}$, and used the large deviations estimates to prove (among other things) a result about exit from a bounded domain containing an equilibrium of the unperturbed system. Fleming rederived this exit theorem using techniques from stochastic control theory.

In this paper we study analogous questions for singular perturbations of degenerate diffusions. We consider the family of processes $dx^\varepsilon(t) = b(x^\varepsilon(t))\,dt + \tau(x^\varepsilon(t)) \circ dz(t) + \sqrt{\varepsilon}\,\sigma(x^\varepsilon(t))\,d\beta(t)$ for $\varepsilon \geq 0$, where the process $x^0(t)$ is degenerate. We show that a large deviations principle need not hold, but we derive bounds sufficient for obtaining some estimates on probabilities. We also adapt Fleming's approach to prove an exit theorem.

1. Introduction. Ventsel and Freidlin [25] begin with a dynamical system $dx_t^0 = b(x_t^0)\,dt$, and perturb it by adding a small random noise to obtain the stochastic differential equation $dx_t^\varepsilon = b(x_t^\varepsilon)\,dt + \sqrt{\varepsilon}\,\sigma(x_t^\varepsilon)\,d\beta(t)$. They show that for fixed $x \in \Re^d$, and fixed $T > 0$, the family of measures induced on the space $E(x) = C([0, T], \Re^D, x)$ by the processes $x^\varepsilon_\cdot|_{[0,T]}$ satisfies a large deviations principle. They use their large deviations estimates to prove, among other things, a result about exit from a bounded domain containing a stable equilibrium of the unperturbed system. Under the assumption that there is a unique point of "minimal energy" on the boundary of the domain, they show that the exit measures induced on the boundary of the domain by the processes x^ε converge weakly to a point mass concentrated at the point of minimal energy. Fleming [5] has shown how one can derive this exit result using techniques from the theory of optimal stochastic control. Expositions of the original Ventsel–Friedlin work appear in [9] and [10].

In this paper, we try to prove analogous results for singular perturbations of degenerate diffusions. Throughout, we shall assume that the dimension is $d \geq 2$. The case $d = 1$ was dealt with in another publication [3].

Consider the stochastic differential equation

$$dx_t^0 = b\left(x_t^0\right)dt + \tau\left(x_t^0\right) \circ dz(t), \tag{1.1}$$

where $\circ$ denotes the Stratonovich differential. (See [13], Chapter III, Section 1; [14].) For $\varepsilon > 0$, we introduce the following perturbation of (1.1):

$$dx_t^\varepsilon = b\left(x_t^\varepsilon\right)dt + \tau\left(x_t^\varepsilon\right) \circ dz(t) + \sqrt{\varepsilon}\,\sigma\left(x_t^\varepsilon\right)d\beta(t). \tag{1.2}$$

Received July 1985; revised June 1986.
AMS 1980 *subject classifications.* Primary 60F10; secondary 34F05, 35B25, 60H10, 60H25, 60J60.
Key words and phrases. Large deviations, singular perturbations, degenerate diffusions, exit problem, Ventsel–Freidlin theory, stochastic control theory.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Probability.* www.jstor.org

For fixed $x \in \Re^d$, and fixed $T > 0$, let P_x^ε denote the measure induced on the space $E(x) = C([0,T], \Re^d, x)$ by the solution of (1.2). For x and v in $\Re^d$, define

$$L(x,v) = \inf_w \tfrac{1}{2}\langle a^{-1}(x)(v - b(x) + \tau(x)w), v - b(x) + \tau(x)w\rangle, \tag{1.3}$$

where $a(x) = \sigma(x)\sigma^*(x)$. If $g \in E(x)$ and g is absolutely continuous with square integrable derivative g', we define

$$I(g) = \int_0^T L(g_t, g_t')\,dt. \tag{1.4}$$

Otherwise $I(g) = +\infty$.

Remark. If $\sigma(x)$ and $\tau(x)$ commute for every $x \in \Re^d$, then $L(x,v)$ has a simpler form, and hence $I(g)$ has a simpler form. In this case,

$$L(x,v) = \tfrac{1}{2}\left|P_{R^\perp \tau(x)}\left[\sigma^{-1}(x)(v - b(x))\right]\right|^2, \tag{1.5}$$

where $P_{R^\perp \tau(x)}$ denotes the orthogonal projection onto the orthogonal complement in $\Re^d$ of the range of the linear map $\tau(x)$.

This paper has two main parts. First, we investigate the problem of whether the measures $\{P_x^\varepsilon\}$ on $E(x) = C([0,T], \Re^d, x)$ obey a large deviations principle. It turns out that this is not true in general. This is demonstrated by means of a counterexample (Lemma 2.11). It is possible, however, to obtain, under minimal assumptions, the lower bound on the probabilities of open sets which would be required to establish a large deviations principle (Proposition 2.1). Two proofs are given. The first, using Girsanov's theorem, is similar to the original proof in [25]. The second uses Schilder's theorem [16] and ideas related to the "contraction principle" (see [24], Remark 1, page 5). We also prove an upper bound (Proposition 2.6) and use this, together with the lower bound to estimate some probabilities (Lemma 2.8). These results are described in more detail in Section 2, and proved in Section 4.

The second part of this paper is devoted to proving an exit result (Theorem 3.16). We have had to impose stringent conditions on the matrix $\tau(x)$ in order to do this (see Assumptions 3.2 and 3.3). This exit result has an interpretation in terms of singular perturbations of degenerate second order elliptic equations (Corollary 3.19). The lower bound (Proposition 2.1) mentioned in the last paragraph is used to prove the lower bound (5.2.2) needed in the exit result. The upper bound (5.2.3) required for the exit result does not, however, follow from the general upper bound result, Proposition 2.6 (an example is provided demonstrating this fact; see Section 5.5), and so an additional argument is needed. An adaptation of Fleming's technique is used. In a special case, another argument establishing the upper bound is given; see Section 5.7. The exit result is stated in Section 3, and proved in Sections 5 and 6.

Although the author was not aware of this while writing this paper, and all the results here were obtained entirely independently, similar problems had been studied earlier by Freidlin and Gartner [8], [11]. An exposition of these results has appeared in Freidlin's new book [7]. In [8], Freidlin and Gartner indicate how one can study the case where τ and σ are both constant. They advocate the use of the same approach as in the original work of Ventsel and Freidlin [25], namely, to prove the lower bound, use Girsanov's theorem, and to prove the upper bound, approximate x^ε by some sort of piecewise smooth path. I have not seen [11]. However, in his book, Freidlin [7] examines the case where $c = \tau\tau^*$ has the special form

$$\begin{bmatrix} C^{(11)} & 0 \\ 0 & 0 \end{bmatrix}. \tag{1.6}$$

He obtains upper and lower bounds, using techniques similar to the original ones in [25] (Girsanov's theorem for the lower bound and piecewise smooth approximations, which are quite complicated in this case, for the upper bound). The lower bound, Theorem IV 4.1 in [7], encodes some information about lower order terms, and thus gives finer information than our bound (2.1). However, it is proved under a more restrictive hypothesis [namely that c has the form (1.6)]. The upper bound, Theorem IV 4.2 in [7], is better than Proposition 2.6 here. Freidlin states a result, Theorem IV 5.1, which can be interpreted as a large deviations theorem for subsets A of $E(x)$ which depend only on those coordinates in which the limiting diffusion has no noise. The exit result (Theorem 3.16) given here could, in principle, be derived from the results in [7], subject to the restriction that they would have to be proved for diffusions on a manifold, as was done in the original work [25]. In [7], Freidlin studies only the case where the diffusions are on $\mathfrak{R}^d$. If the manifold is "nicely foliated" by the vector fields making up the noise (as is the case in Theorem 3.16), then, in local coordinates, c has the form (1.6). In fact, in Theorem IV 5.2, Freidlin gives an exit result similar in spirit to Theorem 3.16 in the special case where $d = 2$ and the diffusion lives on circles. (See Example 3.8.)

2. General results.

Assumptions on the coefficients. Suppose τ is a $d \times d$ matrix-valued function defined on $\mathfrak{R}^d$ so that for each $i, j = 1, \dots, d$, τ^{ij} is in $C_b^2(\mathfrak{R}^d)$, the space of continuous bounded real-valued functions defined on $\mathfrak{R}^d$ which have continuous bounded derivatives of first and second orders. For $x \in \mathfrak{R}^d$, let $\sigma(x)$ be a $d \times d$ matrix so that for each $i, j = 1, \dots, d$, $\sigma^{ij} \in C_b^2(\mathfrak{R}^d)$. For $x \in \mathfrak{R}^d$, let $c(x) = \tau(x)\tau^*(x)$, where $\tau^*(x)$ denotes the transpose of $\tau(x)$. We assume only that $c(x)$ is nonnegative definite. Let $a(x) = \sigma(x)\sigma^*(x)$, where $\sigma^*(x)$ denotes the transpose of $\sigma(x)$. Assume that $a(x)$ is uniformly positive definite on $\mathfrak{R}^d$. Suppose that b: $\mathfrak{R}^d \to \mathfrak{R}^d$ is bounded and uniformly Lipschitz continuous.

Let $(\Omega, \mathbf{F}, P)$ be a probability space, and $\{\mathbf{F}_t\}_{t \geq 0}$ an increasing family of sub-σ-algebras of $\mathbf{F}$. Let $z(t)$ be an $\mathbf{F}_t$-adapted d-dimensional Brownian motion. Consider the stochastic differential equation (1.1). Let $\beta(t)$ be an $\mathbf{F}_t$-adapted d-dimensional Brownian motion independent of the Brownian motion $z(t)$. For $\varepsilon > 0$, consider the perturbation (1.2) of (1.1).

Large deviations bounds and examples. Under the preceding hypotheses on the coefficients, we establish some large deviations bounds for the family of measures $\{P_x^\varepsilon\}$ induced by the processes $x^\varepsilon|_{[0,T]}$ on the space $E(x) = C([0,T], \Re^d, x)$ of $\Re^d$-valued continuous functions defined on the interval $[0,T]$, and taking the value x at time $t = 0$. Throughout, it is understood that the space $E(x)$ is endowed with the topology of uniform convergence on $[0,T]$. The lower bound is stated as Proposition 2.1, and the upper bound as Proposition 2.6. The upper bound pertains only to the special case $\sigma \equiv I$, the $d \times d$ identity matrix, although a similar result could be proved for σ in a wider class by using ideas from [1] to adapt the proof given in Section 4. The upper bound is not a priori the one which would be needed to establish a large deviations principle, but in some cases it agrees with the lower bound. We give an example in which the bounds agree and therefore can be used to show exponential decay of certain probabilities in the special case where σ and τ are both constant matrices; see Lemma 2.8. However, by presenting a counterexample, we shall show that a large deviations principle need not hold for the family of measures $\{P_x^\varepsilon\}$ with rate function given either by I or its lower semicontinuous regularization I^*; see Lemma 2.11. Statements of these results follow. Their proofs appear in Section 4.

PROPOSITION 2.1. *Let $I(\cdot)$ denote the rate function on $E(x)$ defined in* (1.4). *Then if A is an open subset of $E(x)$,*

$$\liminf_{\varepsilon \to 0} \varepsilon \log P_x^\varepsilon(A) \geq -\inf\{I(g) \colon g \in A\}. \tag{2.2}$$

For $A \subset E(x)$, define $\Pi(A) \subset E(0) = C([0,T], \Re^d, 0)$ by

$$\Pi(A) = \{f \in E(0) \colon \exists \varphi \in E(0) \ni \varphi' \in L^2([0,T], \Re^d) \,\&\, F(f,\varphi) \in A\}, \tag{2.3}$$

where for $f \in E(0)$ and $\varphi \in E(0)$, f and φ absolutely continuous, $F(f,\varphi) = g$ is defined to be the solution for $t \in [0,T]$ of

$$g(t) = x + \int_0^t \{b(g(s)) + \tau(g(s))\varphi'(s) + \sigma(g(s)) f'(s)\}\, ds. \tag{2.4}$$

Let I_0 be the rate function for ε (Brownian motion), i.e., for $f \in E(0)$,

$$I_0(f) = \tfrac{1}{2}\int_0^T |f'(t)|^2\, dt \tag{2.5}$$

if f is absolutely continuous with square integrable derivative f', and $I_0(f) = +\infty$ otherwise. Let A be a closed subset of $E(x)$. For $\delta > 0$, let $A^\delta = \{g \in C([0,T], \Re^d, x) \colon \operatorname{dist}(g, A) \leq \delta\}$. Let Cl stand for closure in the space $C([0,T], \Re^d, 0)$. Then:

PROPOSITION 2.6. *In* (1.2), *let $\sigma(x)$ be the $d \times d$ identity matrix. Then if $A \subset E(x)$ is closed,*

$$\limsup_{\varepsilon \to 0} \varepsilon \log P_x^\varepsilon(A) \leq -\inf\Big\{I_0(f) \colon f \in \bigcap_{\delta > 0} \mathrm{Cl}(\Pi(A^\delta))\Big\}. \tag{2.7}$$

An application of the large deviations bounds. We give an application of the estimates in Propositions 2.1 and 2.6 to the problem of calculating the probability that a certain process hits a set before a prescribed time.

LEMMA 2.8. *In* (1.2), *take* $\sigma(x)$ *to be the* $d \times d$ *identity matrix, and* $\tau(x)$ *to be a constant* $d \times d$ *matrix of rank strictly less than d. Let* $\Delta \subset \Re^d$ *be a fixed closed set which is equal to the closure of its interior. Define* $A = A(\Delta) \subset E(x) = C([0,T], \Re^d, x)$ *by*

$$A = \{g \in E(x) \colon g_t \in \Delta \text{ for some } t \in [0,T]\}. \tag{2.9}$$

Then

$$\lim_{\varepsilon \to 0} \varepsilon \log P_x^{\varepsilon}(A(\Delta)) = -\inf\{I(g) \colon g \in A(\Delta)\}. \tag{2.10}$$

A counterexample. We shall show that it is not true in general that the measures P_x^{ε} induced on $E(x)$ by the solution to (1.2) satisfy a large deviations principle with rate function I [as in (1.4)] or I^*, the lower semicontinuous regularization of I. In fact we shall construct examples of closed subsets A of $E(x)$ for which $P_x^{\varepsilon}(A)$ tends to zero arbitrarily slowly, but for which $\inf\{I^*(g) \colon g \in A\}$ is positive, so that, were a large deviations principle to hold, $P_x^{\varepsilon}(A)$ would have to decrease to zero exponentially fast.

LEMMA 2.11. *In* (1.2), *take* $d = 2$, $x = 0 \in \Re^2$, $b \equiv 0$, $T = 1$, $\sigma \equiv I_{2\times 2}$ *and* τ *constant with* $\tau_{11} = 1$ *and* $\tau_{ij} = 0$ *otherwise. Let* $H \colon [0,\infty) \to [0,\infty)$ *be any function so that* $H(\varepsilon)$ *decreases to zero as* $\varepsilon \to 0$. *Then there are sequences* $\{q_n\} \subset (0,\infty)$ *and* $\{a_n\} \subset [0,\infty)$ *so that* q_n *decreases to zero as* $n \to \infty$, $\{a_n\}$ *is increasing,* $a_0 = 0$ *and so that if*

$$V_n = \{f \in E(0) \colon \operatorname{dist}(f, C_1) \ge q_n\}, \tag{2.12}$$

where C_1 *is as in* (4.5) *with* $a = 1$ *and*

$$D_n = \{\varphi \in E(0) \colon a_n \le |\varphi(1)| \le a_{n+1}\}, \tag{2.13}$$

and if $A \subset E(x)$ *is the set*

$$A = \bigcup_{n \ge 1} D_n \times V_n, \tag{2.14}$$

then A is closed in $E(x)$, $\inf\{I^*(g) \colon g \in A\} \ge 1$, *but* $P_x^{\varepsilon}(A) \ge H(\varepsilon)$.

3. An exit problem. For fixed $x \in \Re^d$, and $\varepsilon \ge 0$ let P_x^{ε} denote the measure induced on $E(x) = C([0,\infty), \Re^d, x)$ by the solution of (1.2) starting at x at time $t = 0$. Let D be a bounded region in $\Re^d$ with smooth boundary ∂D. Let T_D denote the time of first exit from D. Suppose that for some $x \in D$, $P_x^0(T_D < \infty) = 0$. Since for $\varepsilon > 0$, the process x^{ε} is nondegenerate, a.s. the trajectories exit from the bounded region D in finite time. Let μ_x^{ε} be the measure induced on the boundary ∂D of D by x^{ε} at time $T_D(x^{\varepsilon})$. For $N \subset \partial D$,

$$\mu_x^{\varepsilon}(N) = P_x^{\varepsilon}(x(T_D) \in N). \tag{3.1}$$

In this section, we shall show that under certain restrictions on the coefficients σ and τ it is possible to prove an exit result analogous to the one proved by Ventsel and Freidlin. We shall show that for a certain class of processes, if the region D has the property that there is a unique point of "minimal energy" (in a sense to be made precise later) on the boundary of D, then the measures μ_x^ε converge weakly.

In order to prove a result of this kind, it is necessary to derive upper and lower bounds on probabilities. The lower bound is derived as a consequence of Proposition 2.1. The upper bound cannot be deduced from Proposition 2.6. This will be shown in Section 5.5 by means of an example. Therefore, an extra argument is required. The argument we present here uses techniques from optimal stochastic control. It is an adaptation of the technique used by Fleming in [5]. The method used here requires that we consider only processes of a very special type. A second proof of the upper bound is given in a special case in Section 5.7.

The main results are stated as Lemma 3.12, Theorem 3.16 and Corollary 3.19. Proofs are given in Sections 5 and 6. Assumptions 3.2, 3.3, 3.5 and 3.6 are needed throughout the proof. Assumption 3.15 and the assumption that there is a unique point of minimal energy are needed to state a nice theorem; see the statement of Theorem 3.16 and Corollary 3.19.

Assumptions.

ASSUMPTION 3.2. EXISTENCE OF THE "RADIAL FUNCTION" R. We shall assume that there is a function R: $\Re^d \to \Re^d$ such that:

(i) $R \in C^2(\Re^d \setminus \{0\})$.
(ii) $R(0) = 0$; $R(x) \geq 0$ for $x \in \Re^d$.
(iii) $\tau^*(x)\nabla R(x) = 0$ for $x \in \Re^d$.
(iv) If $0 < \delta < R < \infty$, then there exists a constant $\eta(\delta, R)$ so that $|\nabla R(x)| \geq \eta(\delta, R)$ if $\delta \leq R(x) \leq R$. There is a constant $M < \infty$ so that $|\nabla R(x)| \leq M$ whenever $R(x) \neq 0$.
(v) For $r \geq 0$, the set $\{x\colon R(x) \leq r\}$ is compact. If $R(x) = R(y) = r$, there is an absolutely continuous function φ on an interval $[0, T]$ with $\varphi(0) = x$, $\varphi(T) = y$ and $R(\varphi(t)) = r$ for $0 \leq t \leq T$.
(vi) For $r, s > 0$, define $D(r, s) = \max\{\operatorname{dist}(x, \{y\colon R(y) = s\})\colon R(x) = r\}$. Then if $0 < \delta < R < \infty$, there exists a constant $K(\delta, R)$ so that $D(r, s) \leq K(\delta, R)|r - s|$ if $\delta \leq R(x) \leq R$.

REMARK. Note that the assumptions that $R(0) = 0$ and that $\{R(x) \leq r\}$ is compact for every $r > 0$ imply that R assumes every nonnegative value.

ASSUMPTION 3.3. ASSUMPTIONS ON τ, σ. We assume that for every $x \in \Re^d$, $\sigma(x)$ and $\tau(x)$ are symmetric, and that $\sigma(x)$ and $\tau(x)$ commute. If $R(x) \neq 0$, assume that the rank of the matrix $\tau(x)$ is $d - 1$. Since $P_{R^\perp \tau(x)}$ and $\sigma^{-1}(x)$

commute, this ensures that for such x,

$$\begin{aligned}(3.4)\qquad |P_{R^{\perp}\tau(0)}\sigma^{-1}(x)y| &= |\nabla R(x)/|\nabla R(x)| \cdot \sigma^{-1}(x)y| \\ &= |\nabla R(x)\cdot y|\,|\sigma^{-1}(x)\nabla R(x)|/|\nabla R(x)|^2.\end{aligned}$$

ASSUMPTIONS ON b. We assume that $b(0) = 0$ and that the vector field b points inward along level sets of $R(\cdot)$, i.e., that

$$(3.5)\qquad \nabla R(x)\cdot b(x) < 0 \quad \text{whenever } R(x) \neq 0.$$

ASSUMPTION ON D. D will always denote a bounded open subset of $\Re^d$ with C^2 boundary ∂D. Assume in addition that

$$(3.6)\qquad 0 \in D, \qquad \min_{y\in\partial D} R(y) \equiv r(D) > 0.$$

This implies, in particular, that $\{x\colon R(x) < r(D)\} \subset D$.

REMARK 3.7. If $x^0(t)$ satisfies (1.1), then by the chain rule for Stratonovich differentials [14] and Assumption 3.2(iii), $dR(x^0(t)) = \nabla R(x^0(t))\cdot b(x^0(t))$, so that Assumption 3.5 implies that a.s. $R(x^0(t))$ is a nonincreasing function of t which is strictly decreasing as long as $R(x^0(t)) \neq 0$. Therefore, a.s. $R(x^0(t)) \to 0$ as $t \to \infty$. In particular, if $x \in D$ is such that $R(x) < r(D)$, then $P_x^0(T_D < \infty) = 0$.

Examples.

THE SPHERICAL CASE. One example where Assumption 3.2 holds is the case where $\tau(x) = (\tau^{ij}(x))$ has the form

$$(3.8)\qquad \tau^{ij}(x) = \delta_{ij} - x_i x_j/|x|^2.$$

In this case $R(x) = |x|$. Note that $\tau(x) = P_{\langle x\rangle^{\perp}}$, the orthogonal projection onto the orthogonal complement of the subspace of $\Re^d$ spanned by x; see [18] and [14].

THE ELLIPTICAL CASE. A slight generalization of the example just given is obtained by replacing spheres centered at the origin by ellipsoids: Let A be any constant symmetric positive definite $d\times d$ matrix, and take $\tau(x) = \tau_A(x) = AP_{\langle A^{-1}x\rangle^{\perp}}$. Here $R(x) = R_A(x) = |A^{-1}x|$.

NOTATION. The exit time from D of a member φ of $C([0,\infty), \Re^d)$ will be denoted by $T_D(\varphi)$. Thus $T_D(\varphi) = \inf\{t \geq 0\colon \varphi(t) \notin D\}$.

If Δ is a subset of D and N a subset of the boundary $\partial\Delta$ of Δ, then for $x \in D$, we define

$$(3.9)\qquad I_\Delta(x, N) = \inf_{T,g} I_{0,T}(g),$$

where the infimum is taken over all $T > 0$ and all absolutely continuous functions $g \in C([0, T], \mathfrak{R}^d, x)$ for which $g(T) \in N$ and $g(t) \in \mathrm{Cl}(\Delta)$ for $0 \leq t \leq T$.

Define K to be the following subset of ∂D:

$$K = \{y \in \partial D \colon I_D(0, \{y\}) = I_D(0, \partial D)\}. \tag{3.10}$$

We shall show below (5.1.3) that if R is the "radial function" of Assumption 3.2 and $r(D)$ is as in (3.6), then

$$K = \{y \in \partial D \colon R(y) = r(D)\}. \tag{3.11}$$

Statement of results.

LEMMA 3.12. *Assume that Assumptions* 3.2, 3.3 *and* 3.5, 3.6 *hold. If $N \subset \partial D$ is any neighbourhood in ∂D of K [defined in* (3.10)] *and if $x \in D$ is such that $R(x) < r(D)$, then*

$$\lim_{\varepsilon \to 0} P_x^\varepsilon(x(T_D) \in N) = 1.$$

Lemma 3.12 deals with those points in D for which $R(x) < r(D)$. In order to say what happens starting at the other points of D, we need to be sure that the unperturbed process x^0 does not hit irregular boundary points with positive probability. Thus it is necessary to make another assumption. To state the assumption we need some definitions.

NOTATION. If $G \subset \mathfrak{R}^d$ is any open set and $g \in C([0, \infty), \mathfrak{R}^d)$, we shall define

$$\begin{aligned} T_G(g) &\equiv \inf\{t \geq 0 \colon g(t) \notin G\}, \\ T_G'(g) &\equiv \inf\{t \geq 0 \colon g(t) \notin \overline{G}\}. \end{aligned} \tag{3.13}$$

DEFINITION 3.14. Let D and $r(D)$ be as in (3.6). Define D^* by

$$D^* = \{x \in D \colon R(x) > r(D)\}.$$

For $x \in D^*$, let μ_x^* denote the measure induced on ∂D^* by x_x^0.

For fixed $x \in \mathfrak{R}^d$, let P_x^0 be the measure induced on $C([0, \infty), \mathfrak{R}^d)$ by the solution of (1.1).

ASSUMPTION 3.15. We shall assume that for every $x \in D^*$,

$$P_x^0[T_{D^*}'(x(\cdot)) = T_{D^*}(x(\cdot))] = 1.$$

The statement of the main result follows. As in Ventsel and Freidlin's work, in order to state a nice result, we assume that there is a unique point of minimal energy on the boundary of D, i.e., that K [defined in (3.10)] consists of a single point.

THEOREM 3.16. *Suppose that Assumptions* 3.2, 3.3 *and* (3.5), (3.6) *and, in addition, Assumption* 3.15 *hold. Suppose also that $K = \{y_0\}$, where K is as in*

(3.10). *Let* μ_x^ε *and* μ_x^* *be as in* (3.1) *and Definition* 3.14, *respectively. Then for every* $x \in D$, *as* $\varepsilon \to 0$,

$$\mu_x^\varepsilon \Rightarrow \mu_x^*|_{\partial D} + \mu_x^*(\{y\colon R(y) = r(D)\})\delta_{y_0}.$$

Theorem 3.16 can be interpreted as a statement about convergence of solutions of certain boundary value problems. Let $\mathbf{L}$ be the differential generator of the process x^0, i.e.,

$$\mathbf{L} = b^{\sim}(x) \cdot \nabla + \tfrac{1}{2}\sum c_{ij}(x)\partial^2/\partial x_i\,\partial x_j, \tag{3.17}$$

where, for $x \in \Re^d$, $b^{\sim}(x) = b(x) + \frac{1}{2}\tau'\tau(x)$ and $\tau'\tau(x) \in \Re^d$ has ith component $(\partial/\partial x_k\, \tau^{ij}(x))(\tau^{kj}(x))$ (summing over repeated indices). (See [13, page 235].) Let $\mathbf{L}^\varepsilon$ be the differential generator of the perturbed process. So

$$\mathbf{L}^\varepsilon = \mathbf{L} + \varepsilon/2\sum a_{ij}(x)\partial^2/\partial x_i\,\partial x_j. \tag{3.18}$$

COROLLARY 3.19 (Analytical interpretation). *Let* D^* *be as in Definition* 3.14. *Suppose that Assumptions* 3.2, 3.3, *and* (3.5), (3.6) *and, in addition, Assumption* 3.15 *hold. Suppose also that* $K = \{y_0\}$, *where* K *is as in* (3.10). *Let* g *be a continuous function defined on* ∂D. *Define* $g^* \in C(\partial D^*, \Re)$ *by* $g^*|_{\partial D} = g|_{\partial D}$ *and* $g^*(y) = g(y_0)$ *if* $R(y) = r(D)$. *Then the problem* $\mathbf{L}u^* = 0$ *in* D^*; $u^*|_{\partial D^*} = g^*$ *has a solution. For* $\varepsilon > 0$ *define*

$$u^\varepsilon(x) = E_x[g(x^\varepsilon(T_D))]. \tag{3.20}$$

Then u^ε *solves the boundary value problem*: $\mathbf{L}^\varepsilon u^\varepsilon = 0$ *in* D; $u^\varepsilon|_{\partial D} = g$. *For every* $x \in D$,

$$u^\varepsilon(x) \to u^*(x) \quad \text{as } \varepsilon \to 0. \tag{3.21}$$

4. The rate function, proofs of large deviations bounds and examples.

The rate function. In the case where σ and τ are constant matrices, the function F in (2.4) is well defined for all f and φ in $E(0)$ and defines a continuous function from $E(0) \times E(0) \to E(x)$. A result analogous to the "contraction principle" (see [4]) suggests that the "correct" formula for the rate function is

$$\begin{aligned} I'(g) &= \inf\{I_0(f)\colon f \in E(0)\ \&\ \exists\varphi \in E(0) \ni \varphi' \\ &\qquad \in L^2([0,T],\Re^d)\ \&\ F(f,\varphi) = g\} \\ &= \inf_{\Psi \in L^2} \tfrac{1}{2}\int_0^T \left|\sigma^{-1}(g_t)[g_t' - b(g_t) - \tau(g_t)\Psi_t]\right|^2 dt. \end{aligned} \tag{4.1}$$

From the second expression, it is clear that $I'(g) \ge I(g)$, where $I(g)$ is as in (1.4). To see that these two quantities are actually equal, assume that $I(g) < \infty$. Then g' is in L^2. Define

$$L_n(x,v) = \inf_{|w| \le n} \tfrac{1}{2}\langle a^{-1}(x)(v - b(x) + \tau(x)w), v - b(x) + \tau(x)w\rangle.$$

Then $I(g) = \lim_{n\to\infty}\int_0^T L_n(g_t, g_t')\,dt = \lim_{n\to\infty} I_n(g)$ [by Lebesgue's dominated convergence theorem, because $L_n(g_t, g_t')$ tends pointwise to $L(g_t, g_t')$, and is bounded by $L_1(g_t, g_t')$, which is integrable because g' is in L^2]. By taking a measurable selection Ψ_t from the set $\{w\colon |w| \le n, \langle a^{-1}(g_t)(g_t' - b(g_t) + \tau(g_t)w), g_t' - b(g_t) + \tau(g_t)w\rangle = 2L_n(g_t, g_t')\}$, one sees that $I'(g) \le I_n(g)$. So $I'(g) \le I(g)$.

Two proofs of the lower bound of Proposition 2.1. We give two proofs of the lower bound stated as Proposition 2.1. The first uses the same technique as Ventsel and Freidlin used in their context, namely Girsanov's theorem. Freidlin and Gartner also use this method in [8]. In the second proof, one thinks of (1.2) as defining a random measure on $E(x)$. With this idea it is easy to deduce the lower bound from the corresponding lower bound for $\varepsilon\beta$ (Schilder's theorem [16]). This proof is given only in the special case $\sigma \equiv I$, but Azencott's ideas [1] can be used to deduce the result for more general σ.

REMARK. In both proofs, the coefficients can be allowed to depend on t as well as on x, as long as they are assumed to be jointly measurable in t and x, and sufficiently smooth in t.

PROOF 1: PROOF BY ABSOLUTELY CONTINUOUS SUBSTITUTION OF MEASURES. The equation (1.2) can be written in Itô form as

$$dx_t^\varepsilon = b^\sim(x_t^\varepsilon)\,dt + \tau(x_t^\varepsilon)\,dz_t + \sqrt{\varepsilon}\,\sigma(x_t^\varepsilon)\,d\beta_t, \tag{4.2}$$

where $b^\sim(x)$ was defined after (3.17).

In order to prove (2.2) for an arbitrary open subset of $E(x) = C([0, T], \Re^d, x)$, it is enough to show that if $g \in E(x)$ is twice continuously differentiable, and $\delta > 0$, then $\liminf_{\varepsilon \to 0} \varepsilon \log P_x(\|x^\varepsilon - g\| < \delta) \ge -I(g)$, where $\|\cdot\|$ denotes the supremum norm.

Let $\xi_t^\varepsilon = x_t^\varepsilon - g_t$ for $t \in [0, T]$. Then $\xi_0^\varepsilon = 0$, and $\xi_\cdot^\varepsilon$ satisfies the stochastic differential equation $d\xi_t^\varepsilon = [b^\sim(\xi_t^\varepsilon + g_t) - g_t']\,dt + \tau(\xi_t^\varepsilon + g_t)\,dz_t + \sqrt{\varepsilon}\,\sigma(\xi_t^\varepsilon + g_t)\,d\beta_t$. By Girsanov's theorem [12], the measure μ_{ξ^ε} induced on $E(0)$ by $\xi_\cdot^\varepsilon$ is absolutely continuous with respect to the measure μ_{ζ^ε} induced on this space by the solution ζ^ε starting at $0 \in \Re^d$ at time $t = 0$ of $d\zeta_t^\varepsilon = \frac{1}{2}\tau'\tau(\zeta_t^\varepsilon + g_t)\,dt + \tau(\zeta_t^\varepsilon + g_t)\,dz_t + \sqrt{\varepsilon}\,\sigma(\zeta_t^\varepsilon + g_t)\,d\beta_t$, and

$$d\mu_{\xi^\varepsilon}/d\mu_{\zeta^\varepsilon} = \exp\left[\int_0^T \Psi_t\,dz_t + \varepsilon^{-1/2}\Phi(t, \zeta_t^\varepsilon + g_t)\,d\beta_t - \tfrac{1}{2}|\Psi_t|^2\,dt - \tfrac{1}{2}\varepsilon^{-1}\left|\Phi(t, \zeta_t^\varepsilon + g_t)\right|^2 dt\right],$$

where $\Phi(t, x) = \sigma^{-1}(x)[b(x) - g_t' - \tau(x)\Psi_t]$ and $\Psi_t \in L^2$ is chosen so that $\frac{1}{2}\int_0^T |\Phi(t, g_t)|^2\,dt \le I(g) + \eta$, where η is some small preassigned number [see (4.1)].

One can choose $0 < \delta' < \delta$ so that on $\|\zeta^\varepsilon\| < \delta'$, $\frac{1}{2}\int_0^T |\Phi(t, \zeta_t^\varepsilon + g_t)|^2\,dt \le \frac{1}{2}\int_0^T |\Phi(t, g_t)|^2\,dt + \eta$. Then

$$\begin{aligned} P_x(\|x^\varepsilon - g\| < \delta) &= P_0(\|\xi^\varepsilon\| < \delta) \ge P_0(\|\xi^\varepsilon\| < \delta') \\ &= \int_{\|\zeta^\varepsilon\| < \delta'} [d\mu_{\xi^\varepsilon}/d\mu_{\zeta^\varepsilon}(\zeta^\varepsilon)]\, d\mu_{\zeta^\varepsilon} \\ &\ge \exp\{-1/\varepsilon[I(g) + 4\eta]\} P\bigg(\|\zeta^\varepsilon\| < \delta', \int_0^T \Psi_t\, dz_t > -\eta/\varepsilon, \\ &\qquad\qquad \int_0^T \Phi(t, \zeta_t^\varepsilon + g_t)\, d\beta_t > -\varepsilon^{-1/2}\eta\bigg). \end{aligned}$$

The result follows from the

Claim. If δ' and η are fixed, and $\varepsilon > 0$ is sufficiently small, then

$$P\bigg(\|\zeta^\varepsilon\| < \delta', \int_0^T \Psi_t\, dz_t > -\eta/\varepsilon, \int_0^T \Phi(t, \zeta_t^\varepsilon + g_t)\, d\beta_t > -\varepsilon^{-1/2}\eta\bigg) \ge \lambda > 0.$$

Proof of Claim. Define ζ_t to be the solution of the equation $d\zeta_t = \frac{1}{2}\tau'\tau(\zeta_t + g_t)\,dt + \tau(\zeta_t + g_t)\,dz_t$. By a standard argument using Gronwall's inequality, there is a constant K so that $E(\|\zeta^\varepsilon - \zeta\|^2) \le K\varepsilon$. So, by Chebyshev's inequality, $P(\|\zeta^\varepsilon\| \ge \delta') \le P(\|\zeta\| > \delta'/2) + 4K\varepsilon/(\delta')^2$. But $P(\|\zeta\| > \delta'/2) < 1$. [For example, by Stroock and Varadhan's support theorem, [21], $0 \in \text{Support}(z(\cdot))$.] Therefore, for sufficiently small $\varepsilon > 0$ (and fixed $\delta' > 0$), $P(\|\zeta^\varepsilon\| < \delta') \ge \rho_0 > 0$. Now

$$P\bigg(\int_0^T \Psi_t\, dz_t \le -\eta/\varepsilon\bigg) \le (\varepsilon^2/\eta^2)\int_0^T |\Psi(t)|^2\, dt \le \text{const.}\ \varepsilon^2/\eta^2.$$

Further,

$$\begin{aligned} &P\bigg(\int_0^T \Phi(t, \zeta_t^\varepsilon + g_t)\, d\beta_t \le -\varepsilon^{-1/2}\eta\bigg) \\ &\qquad \le (\varepsilon/\eta^2) E\bigg(\int_0^T |\Phi(t, \zeta_t^\varepsilon + g_t)|^2\, dt\bigg) \le \text{const.}\ \varepsilon/\eta^2. \end{aligned}$$

So for sufficiently small $\varepsilon > 0$,

$$P\bigg(\|\zeta^\varepsilon\| < \delta', \int_0^T \Psi_t\, dz_t > -\eta/\varepsilon, \int_0^T \Phi(t, \zeta_t^\varepsilon + g_t)\, d\beta_t > -\varepsilon^{-1/2}\eta\bigg) \ge \rho_0/2. \quad \square$$

Proof 2: Using Schilder's theorem. Assume $\sigma \equiv I$. For more general σ, one can use Azencott's ideas [1] to modify the proof given here. The solution to (1.2) can be realised on a product probability space $(\Omega \times \Omega', \mathbf{F} \times \mathbf{F}', P \times P')$, where $z(t)$ and $\beta(t)$ are Brownian motions on $(\Omega, \mathbf{F}, P)$ and $(\Omega', \mathbf{F}', P')$, respectively. For fixed $x \in \mathfrak{R}^d$, let x^ε be the solution starting at x of

$$(4.3)\quad dx_t^\varepsilon(\omega, \omega') = b(x_t^\varepsilon(\omega, \omega'))\, dt + \tau(x_t^\varepsilon(\omega, \omega')) \circ dz_t(\omega) + \sqrt{\varepsilon}\, d\beta_t(\omega').$$

Define $I(\cdot)$ on $E(x) = C([0, T], \mathfrak{R}^d, x)$ as in (1.4). Let P_x^ε be the measure induced on $E(x)$ by $x^\varepsilon_\cdot|_{[0,T]}$. We shall show that for every open set $A \subset C([0, T], \mathfrak{R}^d, x)$, the conclusion (2.2) of Proposition 2.1 holds.

For $f \in E(0)$, f absolutely continuous, let x_t^f be the solution starting at x at time $t = 0$ of

$$dx_t^f(\omega) = \left[b\left(x_t^f(\omega)\right) + f_t'\right] dt + \tau\left(x_t^f(\omega)\right) \circ dz_t(\omega). \tag{4.4}$$

Then $x^f(\cdot)|_{[0,T]}$ is an element of the space $L^2(\Omega, E(x), P)$ of random variables X defined on Ω, taking values in $E(x)$ and satisfying $|||X||| \equiv [E^P\|X\|_{0,T}^2]^{1/2} < \infty$. For fixed ω', $x^\varepsilon_\cdot|_{[0,T]}(\cdot, \omega')$ is also an element of $L^2(\Omega, E(x), P)$. By standard arguments using Gronwall's inequality, it follows that with

$$C_a = \{f \in E(0)\colon f \text{ is absolutely continuous and } I_0(f) \le a\} \tag{4.5}$$

[I_0 was defined in (2.5)], if $f \in C_a$, then

$$E^P\|x^\varepsilon - x^f\|_{0,T}^2 \le K(a)\left\| f - \sqrt{\varepsilon}\beta_\cdot \right\|_{0,T}^2, \tag{4.6}$$

where $K(a)$ is a constant which depends only on a, T, x, d and on the bounds on τ, b and $\tau'\tau$. So $|||x^\varepsilon - x^f||| \le K(a)\|f - \sqrt{\varepsilon}\beta_\cdot\|_{0,T}^2$.

Define a function $\mathfrak{I}$ on $L^2(\Omega, E(x), P)$ by $\mathfrak{I}(X) = \inf\{I_0(f)\colon X = x^f\}$, where the infimum is taken to be $+\infty$ if the set is empty. Let $A \subset L^2(\Omega, E(x), P)$ be open. If $X \in A$, $\mathfrak{I}(X) < a < \infty$, and $X = x^f$, with $f \in C_a$, then for some $\delta > 0$, $P'(x^\varepsilon \in A) \ge P'(|||x^\varepsilon - x^f|||^2 < \delta) \ge P'(\|f - \sqrt{\varepsilon}\beta_\cdot\|_{0,T}^2 < \delta/K(a))$. It now follows from Schilder's theorem [16, 19] and the fact that this holds for every $X \in A$ and every corresponding f that

$$\liminf_{\varepsilon \to 0} \varepsilon \log P'(x^\varepsilon \in A) \ge -\inf\{\mathfrak{I}(X)\colon X \in A\}.$$

Let $\mathbf{P}$ denote the product measure $P \times P'$. We really want estimates on probabilities of the form $\mathbf{P}(x^\varepsilon \in A)$, where $A \subset E(x)$ is open. Note that

$$\mathbf{P}(x^\varepsilon \in A) = E^{P'}[P(x^\varepsilon(\cdot, \omega') \in A)] = E^{P'}[\Phi(x^\varepsilon)],$$

where for $X \in L^2(\Omega, E(x), P)$, $\Phi(X) = P(X \in A)$.

We shall need the results of the following two easy lemmas.

LEMMA 4.7. *If $A \subset E(x)$ is open, and Φ: $L^2(\Omega, E(x), P) \to \mathfrak{R}$ is defined by $\Phi(X) = P(X \in A)$, then Φ is lower semicontinuous.*

PROOF. It must be shown that if $X_n \to X$ in $L^2(\Omega, E(x), P)$, then $\liminf_{n\to\infty}\Phi(X_n) \ge \Phi(X)$. Since $X_n \to X$ in $L^2(\Omega, E(x), P)$, the measures induced on $E(x)$ by X_n converge weakly to the measure induced by X. By a general result on weak convergence of measures on separable metric spaces [22, Theorem 1.1.1], the result follows. □

LEMMA 4.8. *Suppose that $\{P^\varepsilon\}$ is a family of probability measures on a Polish space E which satisfies a large deviations principle with rate function I. Suppose Φ: $E \to \mathfrak{R}$ is nonnegative and lower semicontinuous. Then*

$$\liminf_{\varepsilon \to 0} \varepsilon \log \int \Phi \, dP^\varepsilon \ge -\inf\{I(x)\colon \Phi(x) > 0\}.$$

PROOF. By lower semicontinuity, for every $\eta > 0$, the set $\{x \in E: \Phi(x) > \eta\}$ is open. Clearly $\int \Phi \, dP^\varepsilon \geq \eta P^\varepsilon(\{x \in E: \Phi(x) > \eta\})$. Since, by assumption, the family $\{P^\varepsilon\}$ obeys a large deviations principle with rate I, this implies that $\liminf \varepsilon \log \int \Phi \, dP^\varepsilon \geq -\inf\{I(x): \Phi(x) > \eta\}$. Letting $\eta \to 0$, we get the result. □

From Lemmas 4.7 and 4.8, it follows that if $A \subset E(x)$ is open, then

$$\liminf_{\varepsilon \to 0} \varepsilon \log \mathbf{P}(x^\varepsilon \in A) \geq -\inf\{\Im(X): P(X \in A) > 0\}. \tag{4.9}$$

For $A \subset E(x)$ define $\Pi(A) \subset E(0)$ as in (2.3). One can show that if $g \in E(x)$, and I is the rate function defined in (1.4), then $I(g) = \inf\{I_0(f): f \in \Pi(\{g\})\}$. From this, it follows easily that for $A \subset E(x)$,

$$\inf\{I(g): g \in A\} = \inf\{I_0(f): f \in \Pi(A)\}. \tag{4.10}$$

The assertion of Proposition 2.1 will follow from (4.9) once we have established

LEMMA 4.11. *Suppose $A \subset C([0, T], \Re^d, x)$ is open. Then*

$$\inf\{\Im(X): P(X \in A) > 0\} = \inf\{I_0(f): f \in \Pi(A)\}.$$

PROOF. If $X \in L^2(\Omega, E(x), P)$, $P(X \in A) > 0$ and $\Im(X) < \infty$, then for some $f \in E(0)$, $X = x^f$ and $\Im(X) = I_0(f)$. Since $P(X \in A) > 0$, $A \cap \operatorname{Support}(X)$ is nonempty. By the support theorem of Stroock and Varadhan [21], $\operatorname{Support}(x^f)$ is the closure in $E(x)$ of the set $\{g \in E(x): g = F(f, \varphi)$, some $\varphi\}$. Hence, since A is open, we can find φ so that $F(f, \varphi) \in A$. So $f \in \Pi(A)$, and hence $\inf\{I_0(h): h \in \Pi(A)\} \leq I_0(f) = \Im(X)$. This establishes that the right hand side does not exceed the left hand side. For the opposite inequality, suppose that $f \in \Pi(A)$. Then $F(f, \varphi) \in A$ for some φ. So by the support theorem [21], since A is open, $P(x^f \in A) > 0$, and hence $\inf\{\Im(X): P(X \in A) > 0\} \leq \Im(x^f) \leq I_0(f)$. The result follows. □

Proof of the upper bound of Proposition 2.6. As noted earlier, the upper bound of Proposition 2.6 is not the one needed to establish a large deviations principle for the family of measures $\{P_x^\varepsilon\}$. However, since it agrees with the lower bound established in Proposition 2.1 in some situations, it can be used to derive some results about exponential rates of convergence.

PROOF. Assume $\sigma \equiv I$. Let x^f be as in (4.4). By the support theorem of Stroock and Varadhan [21], the support of the measure induced on $E(x)$ by x^f is the closure in $E(x)$ of $S(f) \equiv \{g \in E(x): \exists \varphi \ni g = F(f, \varphi)\}$. See (2.4) for the definition of F. Suppose the quantity on the right hand side of (2.7) is $< -a$. Then for some $\delta > 0$, $\inf\{I_0(f): f \in \operatorname{Cl} \Pi(A^\delta)\} > a$. Fix $f \in C_a$ [defined in (4.5)]. Then $S(f) \cap A^\delta = \varnothing$. So if $h \in A$, $S(f) \cap B(h, \delta) = \varnothing$ and so $P(\|x^f - h\| < \delta) = 0$. If $f \in C_a$, $\|\varepsilon\beta - f\| < \eta$, and if x^ε satisfies (4.3), then $E^P(\|x^f - x^\varepsilon\|^2) < K(a)\eta^2$ by (4.6). So $P(\|x^f - x^\varepsilon\| \geq \delta) \leq K(a)\eta^2/\delta^2 < 1$ if η is small enough. So $x^\varepsilon \notin A$. Hence $\mathbf{P}(x^\varepsilon \in A) \leq P(d(\varepsilon\beta, C_a) \geq \eta)$, and so by Schilder's theorem [16], the left hand side of (2.7) is $\leq -a$. □

An application of the bounds.

PROOF OF LEMMA 2.8. In order to establish (2.10), it suffices, by Propositions 2.1 and 2.6, to show that

$$\text{(4.12)}\quad \inf\{I(g)\colon g \in A(\Delta)\} = \inf\Big\{I_0(f)\colon f \in \bigcap_{\delta>0} \mathrm{Cl}\Big[\Pi\big((A(\Delta))^{\delta}\big)\Big]\Big\}.$$

For notational convenience, write $B = \bigcap_{\delta>0}\mathrm{Cl}[\Pi(A^{\delta})]$. Since $\Pi(A) \subset B$, it follows from (4.10) that the right hand side of (4.12) does not exceed the left hand side.

To prove the opposite inequality, suppose that $f \in B$. Then for $n \ge 1$, we can find $f_n \in C([0,T], \Re^d, 0)$ so that $\|f - f_n\| \le 1/Kn$, where K is a constant depending on the Lipschitz constant for b, and φ_n absolutely continuous with φ_n' in $L^2([0,T], \Re^d)$ so that $g_n \equiv F(f_n, \varphi_n)$ is in $A^{1/n}$. Let $g_n^* = F(f, \varphi_n)$. By the usual argument using Gronwall's inequality, $\|g_n - g_n^*\| \le 1/n$. (Note that this bound depends on the fact that σ and τ are assumed to be constants. Otherwise we could conclude only that $\|g_n - g_n^*\| \le K_n\|f - f_n\|$, where K_n may depend on the L^2 norms of φ_n' and f_n'.) So $g_n^* \in A^{2/n}$. Let $t_n = \inf\{t \ge 0\colon g_n(t) \in \Delta^{2/n}\}$. Choose $x_n \in \Delta^{2/n}$ so that $|g_n(t_n) - x_n| = 2/n$.

Note that, since σ and τ commute ($\sigma \equiv I_{d\times d}$), $L(x, v)$ in the definition of $I(g)$ has the form (1.5).

Consider two cases:

CASE 1. $t_n < T - 1/n$ infinitely often. We may assume by taking a subsequence if necessary, that $t_n < T - 1/n$ for every n. Define $h_n \in A$ as

$$h_n(t) = \begin{cases} g_n^*(t), & 0 \le t \le t_n, \\ g_n^*(t_n) + n(t - t_n)(x_n - g_n^*(t_n)), & t_n \le t \le t_n + 1/n, \\ \text{the solution of } \begin{cases} g'(t) = b(g(t)), \\ g(t_n + 1/n) = x_n, \end{cases} & t_n + 1/n \le t \le T. \end{cases}$$

Then $\inf\{I(g)\colon g \in A\} \le I(h_n)$, which equals

$$\begin{aligned} &\tfrac{1}{2}\int_0^{t_n} \big|P_{R^{\perp}\tau}\big(g_n'^*(s) - b(g_n^*(s))\big)\big|^2\,ds \\ &\quad + \tfrac{1}{2}\int_{t_n}^{t_n+1/n} \big|P_{R^{\perp}\tau}\big(n(x_n - g_n^*(t_n)) - b(h_n(s))\big)\big|^2\,ds \\ &\qquad \le \tfrac{1}{2}\int_0^{T} |f'(s)|^2\,ds + 1/n(2+M)^2, \end{aligned}$$

where M is the L^{∞} norm of b. Letting $n \to \infty$, we deduce that $\inf\{I(g)\colon g \in A\} \le I_0(f)$.

CASE 2. $t_n \ge T - 1/n$ eventually. Assume this happens for every n. Here, we define $h_n \in A$ by

$$h_n(t) = \begin{cases} g_n^*(c_n t), & 0 \le t \le T - 1/n, \\ g_n^*(t_n) + n(t - T + 1/n)(x_n - g_n^*(t_n)), & T - 1/n \le t \le T, \end{cases}$$

where $c_n = t_n/(T - 1/n)$. Then

$$\inf\{I(g): g \in A\} \le I(h_n) \le \tfrac{1}{2}\int_0^{t_n} |P_{R^{\perp}\tau}[c_n f'(s) + (c_n - 1)b(g_n^*(s))]|^2 c_n^{-1}\, ds$$

$$+ 1/n[2 + M]^2.$$

Since f' is square integrable and b is bounded, one can use Lebesgue's dominated convergence theorem to conclude that, as $n \to \infty$, the expression on the right approaches

$$\tfrac{1}{2}\int_0^T |P_{R^{\perp}\tau}[f'(s)]|^2\, ds,$$

which is not greater than $I_0(f)$. The result follows. □

A counterexample.

PROOF OF LEMMA 2.11. Under the hypotheses of Lemma 2.11, x^{ε} has the same distribution as $\xi^{\varepsilon} = ((1 + \varepsilon)^{1/2} B_1(t), \varepsilon B_2(t))$, where $B_1(t)$ and $B_2(t)$ are independent one-dimensional Brownian motions. Define V_n and D_n as in (2.12) and (2.13), respectively, where $\{q_n\}$ will be chosen later and $\{a_n\}$ is such that $a_0 = 0$ and for $n \ge 1$,

$$(2\pi)^{-1/2} \int_{a_n}^{a_{n+1}} \exp\{-t^2/2\}\, dt = \tfrac{1}{2} p_n,$$

and $p_n \ge 0$, $\sum p_n = 1$ and $p_n > 0$ infinitely often. Define $A \subset E(x)$ as in (2.14). Then A is closed in $E(x)$. Using (4.9), it is easy to see that $\inf\{I^*(g): g \in A\} \ge 1$. However,

$$P_x^{\varepsilon}(A) = P(\xi^{\varepsilon} \in A) = 2\sum (2\pi)^{-1/2} \int_{a_{n+1}^{\varepsilon}}^{a_n^{\varepsilon}} \exp\{-t^2/2\}\, dt\, P(\varepsilon B_2 \in V_n)$$

$$\ge (1 + \varepsilon)^{-1/2} \sum p_n P(\varepsilon B_2 \in V_n).$$

In the integral, $a_n^{\varepsilon} = (1 + \varepsilon)^{-1/2} a_n$. By choosing q_n appropriately, one can ensure that $\sum p_n P(\varepsilon B_2 \in V_n)$ tends to zero as slowly as desired. This is a consequence of the following lemma, which is given here without proof. For a proof see [2] or [4].

LEMMA. *Let $\{p_n\}$ be a sequence of nonnegative numbers with $p_n > 0$ infinitely often and $\sum p_n = 1$. Let $H: [0, \infty) \to [0, \infty)$ be any function so that $H(\varepsilon)$ decreases to zero as $\varepsilon \to 0$. Then there exists a decreasing sequence of strictly positive numbers $\{q_n\}$ so that for sufficiently small ε, $\sum p_n P\{\mathrm{dist}(\varepsilon B_2, C(1)) \ge q_n\} \ge H(\varepsilon)$, where B_2 is a one-dimensional Brownian motion.* □

5. Proof of Lemma 3.12. The hard part of the proof of Theorem 3.16 is to prove Lemma 3.12. This lemma deals with those points x in D for which $R(x) < r(D)$, where R is the radial function of Assumption 3.2 and $r(D)$ was defined in (3.6). Afterwards, in Section 6, we shall take care of the points for which $R(x) \ge r(D)$. This amounts to taking care of some technicalities. This section is devoted to the proof of Lemma 3.12. To facilitate the reading of this section, a list of its contents follows.

5.1. *Definitions and preliminary results.* To outline the proof of Lemma 3.12, we need some definitions. We also give an expression, in (5.1.3), for $I_D(0, \partial D)$ [see (3.9) and (3.10)], from which it is clear that (3.11) holds.

For $x \in \Re^d$ such that $R(x) \neq 0$, define

$$B(x) = -\nabla R(x) \cdot b(x) \big| \sigma(x)^{-1} \nabla R(x) / |\nabla R(x)|^2 \big|^2. \tag{5.1.1}$$

For $r > 0$ in $\Re$, define

$$\beta(r) = \min\{B(x) \colon R(x) = r\}. \tag{5.1.2}$$

It will follow from Lemma 5.3.5 that if K, $r(D)$ and I are as defined in (3.10), (3.6) and (3.9), respectively, then

$$I_D(0, \partial D) = 2\int_0^{r(D)} \beta(s)\, ds. \tag{5.1.3}$$

(5.1.3) obviously implies (3.11).

5.2. *Outline of the proof of Lemma* 3.12. In order to establish Lemma 3.12, we shall prove upper and lower bounds. To be more precise, we shall establish the following facts.

Fix $\delta < r(D)$, where $r(D)$ was defined in (3.6), and let Δ be the set $\{x \in D\colon R(x) > \delta\}$. Let $T_\Delta(\varphi)$ denote the first exit time of $\varphi \in C([0,\infty), \Re^d)$ from Δ. If N is an open subset of ∂D, define

$$r(N) = \inf\{R(y) \colon y \in N\}. \tag{5.2.1}$$

Lower bound. If $N \subset \partial D$ is a neighbourhood of K [see (3.10)], then we shall show that

$$\liminf_{\varepsilon \to 0} \varepsilon \log P_x^\varepsilon\big(x(T_\Delta) \in N\big) \geq -2\int_{R(x)}^{r(N)} \beta(r)\, dr. \tag{5.2.2}$$

We shall show in Section 5.4 how this follows from Proposition 2.1.

Upper bound. If N is any open subset of ∂D,

$$\limsup_{\varepsilon \to 0} \varepsilon \log P_x^\varepsilon\big(x(T_\Delta) \in N\big) \leq -2\int_{R(x)}^{r(N)} \beta(r)\, dr. \tag{5.2.3}$$

This will be proved in Section 5.6 by an adaptation of Fleming's technique. A second proof, which does not use the machinery of stochastic control theory, will be given in Section 5.7 in the case where τ has the special form (3.8). An example is given in Section 5.5 to show that (5.2.3) does not follow from the upper bound in Proposition 2.6.

PROOF OF LEMMA 3.12 FROM (5.2.2) AND (5.2.3). The proof of this statement is exactly the same as that of the corresponding result in Fleming's treatment of the Ventsel–Freidlin exit result [5, page 339]. Choose $\delta > 0$ so that $2\delta < r(D)$. As before, let $\Delta = \{x \in D\colon R(x) > \delta\}$. Since $\partial D \setminus N$ is compact, and for fixed $y \in \partial D \setminus N$, $R(y) > r(D)$, there exists $\rho > 0$ so that $R(y) \geq r(D) + \rho$ for $y \in \partial D \setminus N$. It follows from (3.5) that the function β is strictly positive on $(0, \infty)$. Therefore, it follows from the estimates (5.2.2) and (5.2.3) that

$$\limsup_{\varepsilon \to 0} \sup_{z \in \Gamma} \frac{P_x^\varepsilon(x(T_\Delta) \in \partial D \setminus N)}{P_x^\varepsilon(x(T_\Delta) \in N)} = 0,$$

where $\Gamma = \{z\colon R(z) = 2\delta\}$. The result follows from a standard stopping time argument; see [5] and [23]. □

5.3. *Technical lemmas.*

LEMMA 5.3.1. *Let $B(x)$ be defined for $x \in \Re^d$ such that $R(x) \neq 0$ by* (5.1.1). *Then B is uniformly Lipschitz continuous on $\{x\colon \delta \leq R(x) \leq R\}$ whenever $\delta < R < \infty$.*

PROOF. This is an easy consequence of the assumptions on b and σ (Section 2) and Assumptions 3.2(i) and (iv). □

LEMMA 5.3.2. *Let $\beta(r)$ be defined for $r > 0$ by* (5.1.2). *Then β is uniformly Lipschitz continuous on $[\delta, R]$ whenever $\delta < R < \infty$.*

PROOF. This follows from Lemma 5.3.1 and Assumption 3.2(vi). □

LEMMA 5.3.3. *Suppose $\varphi \in C([0,T], \Re^d)$ is absolutely continuous. If $r(\varphi) = \max\{R(\varphi(t))\colon t \in [0,T]\}$ and I is as in* (1.4), *then*

$$I(\varphi) \geq 2\int_{R(\varphi(0))}^{r(\varphi)} \beta(r)\,dr. \tag{5.3.4}$$

PROOF. Fix $n \geq 1$. For $k = 1, \dots, n$, let $r_k = R(\varphi(0)) + (k/n)(r(\varphi) - R(\varphi(0)))$. Let t_k be the first time that $R(\varphi(t))$ equals r_k. Let s_k be the last time before t_k that $\varphi(t)$ equals r_{k-1}. Then

$$r_k - r_{k-1} = \int_{s_k}^{t_k} \nabla R(\varphi(t)) \cdot \varphi'(t)\,dt \leq \int_{s_k}^{t_k} \nabla R(\varphi(t)) \cdot \varphi'(x)\chi_S(t)\,dt,$$

where $S = \{t \in [0,T]\colon \nabla R(\varphi(t)) \cdot \varphi \ni (t) > 0\}$. Hence, from (3.4) and (1.5), we get that $I(\varphi)$ is at least

$$\begin{aligned}
&\sum_{k=1}^{n} \int_{s_k}^{t_k} L(\varphi(t), \varphi'(t))\, dt \\
&\quad \geq \tfrac{1}{2} \sum_{k=1}^{n} \int_{s_k}^{t_k} \Big\{ \nabla R(\varphi(t)) \cdot [\varphi'(t) - b(\varphi(t))] \\
&\qquad\qquad \times \left|\sigma^{-1}(\varphi(t)) \nabla R(\varphi(t))\right| \Big/ \left|\nabla R(\varphi(t))\right|^2 \Big\}^2 \chi_S(t)\, dt \\
&\quad \geq 2 \sum_{k=1}^{n} \int_{s_k}^{t_k} B(\varphi(t)) [\nabla R(\varphi(t)) \cdot \varphi'(t)] \chi_S(t)\, dt \\
&\quad \geq 2 \sum_{k=1}^{n} \min_{r_{k-1} \leq r \leq r_k} \beta(r)(r_k - r_{k-1}).
\end{aligned}$$

The continuity of β ensures that the last expression approaches the right hand side of (5.3.4) as $n \to \infty$. □

LEMMA 5.3.5. *Suppose that $R(x) < R(y)$. Then*

$$\inf_{T, \varphi} I(\varphi) = 2 \int_{R(x)}^{R(y)} \beta(r)\, dr, \tag{5.3.6}$$

where the infimum is taken over all $T > 0$ and all absolutely continuous φ defined on $[0, T]$ so that $\varphi(0) = x$ and $\varphi(T) = y$.

PROOF. That the left hand side is at least as big as the right side follows from Lemma 5.3.3. To prove the opposite inequality it suffices to construct a sequence of absolutely continuous functions $\{\varphi_n\}$ so that φ_n is defined on an interval $[0, T_n]$, $\varphi_n(0) = x$, $\varphi_n(T_n) = y$ and $I(\varphi_n)$ approaches the right side of (5.3.6) as $n \to \infty$. Fix $n \geq 1$. For $k = 1, \dots, n$ let $r_k = R(x) + k/n(R(y) - R(x))$. Choose x_k so that $R(x_k) = r_k$ and $B(x_k) = \beta(r_k)$. Assume for simplicity that $R(x) \neq 0$. A simple modification of the argument given below can be used if $R(x) = 0$. Let Ψ_0 be an absolutely continuous function defined on an interval $[0, t_0]$ with $t_0 \leq 1/n^2$ so that $R(\Psi(t)) = r_0$ for $0 \leq t \leq t_0$, $\Psi(0) = x$, $\Psi(t_0) = x_0$. [See Assumption 3.2(v)—the time parameter can be changed to make the interval of definition as short as desired.] Let γ_0 be a function satisfying the equation

$$\gamma'(t) = [-b(\gamma(t)) \cdot \nabla R(\gamma(t))] |\nabla R(\gamma(t))|^{-2} \nabla R(\gamma(t)) \tag{5.3.7}$$

with the initial condition $\gamma_0(0) = x_0$, and let $s_0 = \inf\{t > 0\colon R(\gamma_0(t)) = r_1\}$. Proceed by induction. For $k = 1, \dots, n-1$, let Ψ_k be an absolutely continuous function defined on an interval $[0, t_k]$ with $t_k \leq 1/n^2$ so that $R(\Psi(t)) = r_k$ for $0 \leq t \leq t_k$, $\Psi(0) = \gamma_{k-1}(s_{k-1})$, $\Psi(t_k) = x_k$. Let γ_k be a function satisfying (5.3.7) with initial condition $\gamma_k(0) = x_k$, and let $s_k = \inf\{t > 0\colon R(\gamma_k(t)) = r_{k+1}\}$. Finally, let Ψ_n be an absolutely continuous function defined on an interval $[0, t_n]$ with $t_n \leq 1/n^2$ so that $R(\Psi(t)) = r_n$ for $0 \leq t \leq t_n$, $\Psi(0) = \gamma_{n-1}(s_{n-1})$,

$\Psi(t_n) = y$. Define φ_n on $[0, T_n]$ where $T_n = \Sigma(t_k + s_k) + t_n$, where k ranges from 0 to $n-1$. We define φ_n by joining up the pieces previously defined: $\varphi_n(t)$ equals $\Psi_0(t)$ for $0 \le t \le t_0$, $\gamma_0(t - t_0)$ for $t_0 \le t \le t_0 + s_0$, $\Psi_1(t - t_0 - s_0)$ for $t_0 + s_0 \le t \le t_0 + s_0 + t_1$, etc.

Clearly $I(\varphi_n) = \Sigma I(\Psi_k) + \Sigma I(\gamma_k)$ where, in the first sum, k ranges from 0 to n, and in the second, from 1 to $n-1$. Since $R(\Psi_k(t)) \equiv r_k$ on $[0, t_k]$, $P_{R^\perp \tau(\Psi(t))} \Psi_k'(t) = 0$ for $0 \le t \le t_k$. Therefore, $I(\Psi_k) \le (M/n)^2$, where M is a bound on b. Using (5.3.7), the definition (1.4) of I and (3.4), it is easy to show that

$$I(\gamma_k) = 2\int_0^{s_k} B(\gamma_k(t))d/dt(R(\gamma_k(t)))\,dt. \tag{5.3.8}$$

Since $B(\gamma_k(0)) = B(x_k) = \beta(r_k) = \beta(R(\gamma_k(0)))$, and since B and R are Lipschitz continuous on $\{z\colon R(x) \le R(z) \le R(y)\}$, and b is Lipschitz on $[R(x), R(y)]$,

$$\begin{aligned} B(\gamma_k(t)) &\le \beta(R(\gamma_k(t))) + |\beta(R(\gamma_k(t))) - \beta(R(\gamma_k(0)))| \\ &\quad + |B(\gamma_k(0)) - B(\gamma_k(t))| \\ &\le \beta(R(\gamma_k(t))) + K|\gamma_k(t) - \gamma_k(0)|, \end{aligned} \tag{5.3.9}$$

where K is a constant. It follows from the equation satisfied by $\gamma_k(t)$ [(5.3.8)], the Assumption 3.2(iv) that $|\nabla R(z)|$ is bounded below on $\{z\colon R(x) \le R(z) \le R(y)\}$, and the Lipschitz continuity of R on this set, that $|\gamma_k(t) - \gamma_k(0)| \le K(r_{k+1} - r_k)$ for $0 \le t \le s_k$, where K is a (different) constant independent of n. So from (5.3.8) and (5.3.9) it follows that

$$\begin{aligned} \sum_{k=0}^{n-1} I(\gamma_k) &\le 2\sum_{k=0}^{n-1}\int_0^{s_k} \beta(R(\gamma_k(t)))d/dt(R(\gamma_k(t)))\,dt \\ &\quad + K\sum_{k=0}^{n-1}(r_{k+1} - r_k)\int_0^{s_k} d/dt(R(\gamma_k(t)))\,dt \\ &= 2\sum_{k=0}^{n-1}\int_{r_k}^{r_{k+1}} \beta(r)\,dr + K\sum_{k=0}^{n-1}(r_{k+1} - r_k)^2. \end{aligned}$$

K is a constant independent of n whose value may change from line to line. From this it follows easily that $\limsup_{n\to\infty} I(\varphi_n)$ does not exceed the right hand side of (5.3.6) and the result follows. □

5.4. *Proof of the lower bound* (5.2.2). From Proposition 2.1 it follows that it suffices to construct a sequence of functions $\{\varphi_n\}_{n\ge1}$ so that φ_n lies in the interior of the set $A(T_n) = \{\varphi \in C([0, T_n], \Re^d, x)\colon T_\Delta(\varphi) \le T_n;\ \varphi(T_\Delta(\varphi)) \in N\}$, and so that as $n \to \infty$, $-I(\varphi_n)$ approaches the quantity on the right hand side of (5.2.2).

Choose y_n in the interior of N so that $R(y_n) \to r(N)$ as $n \to \infty$. By Lemma 5.3.5 one can construct a function φ_n on an interval $[0, S_n]$ so that

$$\left| I_{0,Sn}(\varphi_n) - \int_{R(x)}^{R(y_n)} \beta(s)\,ds \right| \to 0 \quad \text{as } n \to \infty.$$

Extend the definition of φ_n to an interval $[0, T_n]$ with $T_n > S_n$ in such a way that

φ_n penetrates a positive distance into the complement of the closure of Δ during the interval $[S_n, T_n]$. Then any path sufficiently close to φ_n in the topology of uniform convergence on $[0, T_n]$ will also leave Δ during $[0, T_n]$ and will do so through N. So φ_n will lie in the interior of $A(T_n)$. By making $T_n - S_n$ and the distance travelled by φ_n during the interval $[S_n, T_n]$ small, one can ensure that $I_{S_n, T_n}(\varphi_n) \to 0$ as $n \to \infty$. □

5.5. *An example motivating the introduction of a new technique for proving the upper bound.* In this section we give an example showing that the upper bound (2.6) is not good enough to yield the upper estimate (5.2.3) which we need to prove the exit result, even in the "spherical case" (3.8). The details are given in [2].

In (3.8) with $\sigma \equiv I_{d\times d}$, take $d = 2$, and, for $x \in \Re^2$, define $b(x) = -x$. Then $x \cdot b(x) = -|x|^2 < 0$ for every $x \neq 0$. Let $A \subseteq E(0)$ be the set of functions

$$A = \{g\colon g(0) = 0, |g(t)| \geq 1 \text{ for some } t \in [0,1]\}.$$

We shall show that

$$\inf\{I_0(f)\colon f \in \mathrm{Cl}(\Pi(A))\} = 0, \tag{5.5.1}$$

which clearly implies that the expression on the right hand side of (2.7) equals zero, so that Proposition 2.6 says nothing. To show (5.5.1), it suffices to construct sequences $\{f_m\}$, $\{\varphi_m\}$ and $\{g_m\}$ in $E(0)$ so that $g_m \in A$ and $F(f_m, \varphi_m) = g_m$, and so that $\|f_m\| \to 0$. This implies that $0 \in \mathrm{Cl}(\Pi(A))$, which clearly implies (5.5.1).

Define, for $t \in [0,1]$,

$$g_m(t) = t(\cos mt, \sin mt),$$

$$f_m(t) = m^{-2}(\cos mt + m(1+t)\sin mt - 1, \sin mt - m(1+t)\cos mt + m),$$

$$\varphi_m(t) = (t\cos mt - m^{-1}\sin mt, t\sin mt + m^{-1}\cos mt - m^{-1}).$$

An elementary computation shows that these sequences have the desired properties.

5.6. *Proof of the upper bound using stochastic control.* This section is devoted to proving the estimate (5.2.3). The technique is a modification of Fleming's. Note that until after Remark 5.6.18, we do not make use of the special assumptions introduced in Section 3, except Assumption 3.3.

Let Φ be any function satisfying

$$\Phi \in C^2(\Re^d); \qquad \Phi(x) = 0, \quad x \in N; \qquad \Phi(x) > 0, \quad x \notin N. \tag{5.6.1}$$

For $x \in \Delta$, define

$$\begin{aligned} g^\varepsilon(x) &= E_x[\exp\{-\Phi(x^\varepsilon(T_\Delta))/\varepsilon\}], \\ J^\varepsilon(x) &= -\varepsilon\log(g^\varepsilon(x)). \end{aligned} \tag{5.6.2}$$

Then for $x \in \Delta$,

$$\limsup_{\varepsilon\to 0} \varepsilon \log P_x^\varepsilon(x(T_\Delta) \in N) \leq -\liminf_{\varepsilon\to 0} J^\varepsilon(x). \tag{5.6.3}$$

We shall take a sequence $\{\Phi^M\}$ of functions satisfying (5.6.1) so that

(5.6.4) $$\Phi^M(x) \to \infty \quad \text{as} \quad M \to \infty \quad \text{for } x \notin N.$$

Letting $J^{\varepsilon, M}$ denote the function obtained from Φ^M using the definition (5.6.2), we shall show

(5.6.5) $$\liminf_{M \to \infty} \liminf_{\varepsilon \to 0} J^{\varepsilon, M}(x) \ge 2 \int_{R(x)}^{r(N)} \beta(r)\, dr.$$

This, together with (5.6.3), will give the desired estimate (5.2.3).

It follows by standard results that g^ε solves the boundary value problem: $\mathbf{L}^\varepsilon g^\varepsilon = 0$ in Δ; $g^\varepsilon(x) = \exp\{-\Phi(x)/\varepsilon\}$ for $x \in \partial\Delta$. $\mathbf{L}^\varepsilon$ was defined in (3.18). Therefore, from the definition (5.6.2) of J^ε it follows that J^ε solves the nonlinear equation

$$\tfrac{1}{2}\sum \left(c_{ij}(x) + \varepsilon a_{ij}(x)\right)^{1/2} J_{ij}^\varepsilon(x) + J_x^\varepsilon(x) \cdot b^\sim(x)$$
$$-\varepsilon^{-1}/2 J_x^\varepsilon(x) \cdot \left(c(x) + \varepsilon a(x)\right) J_x^\varepsilon(x) = 0,$$

where $J_{ij}^\varepsilon(x) = \partial^2/\partial x_i\, \partial x_j (J^\varepsilon(x))$ and $J_x^\varepsilon(x) = \nabla J^\varepsilon(x)$, with boundary condition

(5.6.6) $$J^\varepsilon(x) = \Phi(x) \quad \text{for } x \in \partial\Delta.$$

For $x \in \Re^d$ and $p \in \Re^d$, define

(5.6.7) $$H^\varepsilon(x, p) = -\tfrac{1}{2} p \cdot \varepsilon^{-1}(c(x) + \varepsilon a(x))p + p \cdot b^\sim(x).$$

Then J^ε satisfies

(5.6.8) $$\tfrac{1}{2}\sum \left(c_{ij}(x) + \varepsilon a_{ij}(x)\right)^{1/2} J_{ij}^\varepsilon(x) + H^\varepsilon(x, J_x^\varepsilon(x)) = 0.$$

$H^\varepsilon(x, \cdot)$ is dual in the sense of duality for convex and concave functions to $L^{\sim\varepsilon}(x, \cdot)$, defined by

(5.6.9) $$L^{\sim\varepsilon}(x, v) = \tfrac{1}{2}(b^\sim(x) - v) \cdot \varepsilon\left(c(x) + \varepsilon a(x)\right)^{-1}(b^\sim(x) - v)$$

for $x \in \Re^d$ and $v \in \Re^d$. This means that

(5.6.10) $$H^\varepsilon(x, p) = \min_{v \in \Re^d} \{L^{\sim\varepsilon}(x, v) + p \cdot v\}.$$

The minimum in (5.6.10) is attained when

(5.6.11) $$v = b^\sim(x) - \varepsilon^{-1}(c(x) + \varepsilon a(x))p.$$

The equation (5.6.8) is the dynamic programming equation for the following stochastic control problem. In the notation of Section 1, let $\mathbf{v}(t)$ be a control which is $\mathbf{F}_t$-adapted and bounded. The corresponding state of the controlled system, $\eta(t)$, is the solution of

(5.6.12) $$d\eta(t) = \mathbf{v}(t)\, dt + \tau(\eta(t)) \circ dz(t) + \sqrt{\varepsilon}\, \sigma(\eta(t))\, d\beta(t).$$

The problem is to minimize

(5.6.13) $$\Im^\varepsilon(x, \mathbf{v}) = E_x\left\{ \int_0^{T_\Delta} L^\varepsilon[\eta(t), \mathbf{v}(t)]\, dt + \Phi(\eta(T_\Delta))\right\}$$

over all such controls v, for fixed initial state x, where for $x \in \Re^d$ and $v \in \Re^d$,

$$L^\varepsilon(x, v) = L^{\sim\varepsilon}\big(x, v + \tfrac{1}{2}\tau'\tau(x)\big). \tag{5.6.14}$$

So L^ε is defined analogously to (5.6.9) with $b^\sim(x)$ replaced by $b(x)$.

Lemma 5.6.15. *$J^\varepsilon(x)$ is the infimum over all controls* v *of* $\mathfrak{J}^\varepsilon(x, \mathbf{v})$. *Moreover, there is an optimal feedback control* $\mathbf{v}^\varepsilon(t) = V^\varepsilon(\eta^\varepsilon(t))$, *where* $\eta^\varepsilon(t)$ *satisfies the control equation* (5.6.12) *with* $\mathbf{v}(t)$ *replaced by* $\mathbf{v}^\varepsilon(t)$, *and* $V^\varepsilon(x)$ *is defined for* $x \in \Re^d$ *by*

$$V^\varepsilon(x) = b(x) - \varepsilon^{-1}\big(c(x) + \varepsilon a(x)\big)J_x^\varepsilon(x). \tag{5.6.16}$$

Proof. This follows from a standard argument using Itô's formula, the equation and boundary condition satisfied by J^ε (5.6.6), (5.6.8), and the duality between $L^{\sim\varepsilon}$ and H^ε; see [5] and [6]. □

We now give the connection between the functions L and L^ε defined in (1.5) and (5.6.14).

Lemma 5.6.17. *For fixed* $x \in \Re^d$ *and* $v \in \Re^d$, $L^\varepsilon(x, v) \downarrow L(x, v)$ *as* $\varepsilon \downarrow 0$.

Proof. This follows from the following fact:

Lemma. *Suppose τ and σ are constant symmetric $d \times d$ matrices which commute, and let $a = \sigma^2$ and $c = \tau^2$. Suppose that a is strictly positive definite. Then for any $w \in \Re^d$, as $\varepsilon \downarrow 0$,*

$$\varepsilon w \cdot (c + \varepsilon a)^{-1} w \downarrow |P_{R^\perp\tau}\sigma^{-1}w|^2.$$

Proof. This follows by an elementary argument, using the fact that if τ and σ commute, they can be simultaneously diagonalized. □

Remark 5.6.18. Note that so far in Section 5.6 we have not used any of the special assumptions introduced in Section 3, except Assumption 3.3. From this point on, however, we shall assume that all the assumptions imposed in that section hold.

Let $\mathbf{v}^\varepsilon(t)$ and $\eta^\varepsilon(t)$ be as in Lemma 5.6.15. Applying Itô's formula to the process $\eta^\varepsilon(t)$ and the function R in (3.2), one obtains

$$\begin{aligned} dR(\eta^\varepsilon(t)) = {} & \big\{\nabla R(\eta^\varepsilon(t)) \cdot \mathbf{v}^\varepsilon(t) + (\varepsilon/2)\textstyle\sum R_{ij}(\eta^\varepsilon(t))a_{ij}(\eta^\varepsilon(t))\big\}\, dt \\ & + \sqrt{\varepsilon}\, \nabla R(\eta^\varepsilon(t)) \cdot \sigma(\eta^\varepsilon(t))\, d\beta(t). \end{aligned} \tag{5.6.19}$$

Define a process $R^\varepsilon(t)$ by

$$R_\varepsilon(t) = R(\eta^\varepsilon(t)). \tag{5.6.20}$$

Define R_ε^* by

$$R_\varepsilon^*(t) = R_\varepsilon(0) + \int_0^t \nabla R(\eta^\varepsilon(s)) \cdot v^\varepsilon(s)\, ds. \tag{5.6.21}$$

In order to demonstrate (5.6.5) and hence (5.2.3), it suffices to show that for every $r < r(N)$, the left hand side of (5.6.5) is not less than the integral from $R(x)$ to r of $2\beta(r)$. Let r be an arbitrary number with $R(x) < r < r(N)$. Fix $n \geq 1$, and for $k = 1, \dots, n$, let $r_k = R(x) + (k/n)(r - R(x))$. Define random times T_k and S_k by

$$\begin{aligned} T_k &= \inf\{t \geq 0\colon R_\varepsilon^*(t) = r_k\}, \qquad k = 0, \dots, n, \\ S_k &= \sup\{t < T_k\colon R_\varepsilon^*(t) = r_{k-1}\}, \qquad k = 1, \dots, n. \end{aligned} \tag{5.6.22}$$

Then

$$r_k - r_{k-1} \leq \int_{S_k}^{T_k} d/dt(R_\varepsilon^*(t))\chi_B(t)\,dt, \tag{5.6.23}$$

where B is the random set of times $\{t\colon d/dt(R_\varepsilon^*(t)) > 0\}$. By Lemma 5.6.15, $J^\varepsilon(x) = \mathfrak{J}^\varepsilon(x, v^\varepsilon)$. Using this, the fact that for $x \in \mathfrak{R}^d$ and $v \in \mathfrak{R}^d$, $L^\varepsilon(x, v) \geq L(x, v)$ (see Lemma 5.6.17) and (3.4), we deduce from (5.6.23) that

$$\begin{aligned} J^\varepsilon(x) &\geq E_x\Bigg[\sum_{k=1}^n \int_{S_k \wedge T_\Delta}^{T_k \wedge T_\Delta} \tfrac{1}{2}\Big\{\nabla R(\eta^\varepsilon(t)) \cdot (v^\varepsilon(t) - b(\eta^\varepsilon(t))) \\ &\qquad \times |\sigma^{-1}(\eta^\varepsilon(t))\nabla R(\eta^\varepsilon(t))||\nabla R(\eta^\varepsilon(t))|^{-2}\Big\}^2 \chi_B(t)\,dt\Bigg] \\ &\geq E_x\Bigg[\sum_{k=1}^n \int_{S_k \wedge T_\Delta}^{T_k \wedge T_\Delta} 2B(\eta^\varepsilon(t))d/dt(R_\varepsilon^*(t))\chi_B(t)\,dt\Bigg] \\ &\geq 2\sum_{k=1}^{n-1} E_x\Big[\min_{x \in A_k} B(x)\Big](r_k - r_{k-1}), \end{aligned}$$

where A_k is the random subset of $\mathfrak{R}^d$ $A_k = \{\eta^\varepsilon(t)\colon t \in [S_k \wedge T_\Delta, T_k \wedge T_\Delta]\}$. If $\sup\{|R_\varepsilon(t) - R_\varepsilon^*(t)|\colon t \in [0, T_\Delta]\} \leq \rho$, and if $T_k \leq T_\Delta$, then $A_k \subseteq \{x \in \mathfrak{R}^d\colon r_{k-1} - \rho \leq R(x) \leq r_k + \rho\}$. Since $T_k \leq T_{n-1}$ for $k = 1, \dots, n-1$,

$$\begin{aligned} E_x\Big[\min_{x \in A_k} B(x)\Big] &\geq \min_{r_{k-1} - \rho \leq R(x) \leq r_k + \rho} B(x) \\ &\qquad \times P_x\big[T_{n-1} \leq T_\Delta; \sup\{|R_\varepsilon(t) - R_\varepsilon^*(t)|\colon t \in [0, T_\Delta]\} \leq \rho\big]. \end{aligned}$$

On the set where $\sup\{|R_\varepsilon(t) - R_\varepsilon^*(t)|\colon t \in [0, T_\Delta]\} \leq \rho$ and $R(\eta^\varepsilon(T_\Delta)) > r$, we have $R_\varepsilon^*(T_\Delta) \geq r - \rho = r_n - \rho$, and so if $\rho < r_n - r_{n-1}$, then $T_{n-1} < T_\Delta$ on this set. Therefore,

$$\begin{aligned} J^\varepsilon(x) &\geq 2S(n)P_x\big(R(\eta^\varepsilon(T_\Delta)) > r; \\ &\qquad \sup\{|R_\varepsilon(t) - R_\varepsilon^*(t)|\colon t \in [0, T_\Delta]\} \leq \rho\big), \end{aligned} \tag{5.6.24}$$

where

$$S(n) = \sum_{k=1}^{n-1} \min_{r_{k-1} - \rho \leq r \leq r_{k+\rho}} \beta(r)(r_k - r_{k-1}). \tag{5.6.25}$$

Using the definitions (5.6.20) and (5.6.21) of R_ε and R_ε^*, equation (5.6.19), and

Assumption 3.2(i) and (iv) and standard arguments using Doob's martingale theorem, we get that there is a constant K so that for any $T > 0$,

$$\begin{aligned} P_x(\sup\{|R_\varepsilon(t) - R_\varepsilon^*(t)|: t \in [0, T_\Delta]\} \geq \rho) \\ \leq K(1/T + \varepsilon/\rho)E_x(T_\Delta) + \varepsilon KT/\rho^2. \end{aligned} \tag{5.6.26}$$

We shall show in the following that there is a constant C independent of $\varepsilon > 0$ and of $M \geq 1$ so that

$$E_x(T_\Delta) \leq C[1 + J^{\varepsilon, M}(x)]. \tag{5.6.27}$$

To prove (5.6.5), we may assume that the left side of (5.6.5) is finite and that there are sequences $\{M(k)\}_{k\geq 1}$ and, for each $k \geq 1$ $\{\varepsilon(k, m)\}_{m\geq 1}$ so that the left side of (5.6.5) equals

$$\lim_{k\to\infty} \lim_{m\to\infty} J^{M(k), \varepsilon(k, m)}(x). \tag{5.6.28}$$

In particular, $J^{M(k), \varepsilon(k,m)}$ is bounded independently of k and of m. Restricting (M, ε) to lie in the set $\{(M(k), \varepsilon(k, m)): k \geq 1, m \geq 1\}$, we may assume, by (5.6.27), that $E_x(T_\Delta)$ is bounded independently of ε and M. This, together with (5.6.26) implies that for any $\lambda > 0$,

$$\begin{aligned} \liminf_{k\to\infty} \liminf_{m\to\infty} P_x\Big(R\big(\eta^{\varepsilon(k,m)}(T_\Delta)\big) > r; \\ \sup\big\{|R_{\varepsilon(k,m)}(t) - R_{\varepsilon(k,m)}{}^*(t)|: t \in [0, T_\Delta]\big\} \leq \rho\Big) \\ \geq (1 - \lambda) \liminf_{k\to\infty} \liminf_{m\to\infty} P_x\Big(R\big(\eta^{\varepsilon(k,m)}(T_\Delta)\big) > r\Big). \end{aligned}$$

This, together with (5.6.24) and (5.6.28) implies that the left side of (5.6.5) is not less than

$$2S(n)(1 - \lambda) \liminf_{k\to\infty} \liminf_{m\to\infty} P_x\Big(R\big(\eta^{\varepsilon(k,m)}(T_\Delta)\big) > r\Big).$$

The continuity of β implies that, as $n \to \infty$, $S(n)$ approaches the integral from $R(x)$ to r of $\beta(s)\,ds$. (5.6.5) will follow if we show that for every $r \in (R(x), r(N))$,

$$\liminf_{k\to\infty} \liminf_{m\to\infty} P_x\Big(R\big(\eta^{\varepsilon(k,m)}(T_\Delta)\big) > r\Big) = 1.$$

This follows from (5.6.4) because $\{J^{M(k), \varepsilon(k,m)}: k \geq 1, m \geq 1\}$ is bounded and because for any $\varepsilon > 0$ and $M \geq 1$,

$$J^{\varepsilon, M}(x) \geq \inf\{\Phi^M(y): R(y) \leq r, y \in \partial D\} P_x\big(R(\eta^\varepsilon(T_\Delta)) \leq r\big).$$

To complete the proof, we must show that the inequality (5.6.27) holds.

PROOF OF (5.6.27). From (5.6.20) and (5.6.19), it follows that

$$\begin{aligned} dR_\varepsilon(t) &= \nabla R(\eta^\varepsilon(t)) \cdot (v^\varepsilon(t) - b(\eta^\varepsilon(t)))\,dt \\ &\quad + \big[\nabla R(\eta^\varepsilon(t)) \cdot b(\eta^\varepsilon(t)) + (\varepsilon/2) \textstyle\sum R_{ij}(\eta^\varepsilon(t)) a_{ij}(\eta^\varepsilon(t))\big]\,dt \\ &\quad + \sqrt{\varepsilon}\,\nabla R(\eta(t)) \cdot \sigma(\eta(t))\,d\beta(t). \end{aligned}$$

Since D is bounded, and $\Delta = \{x \in D: \delta < R(x)\}$, it follows from (3.5) and

Assumption 3.2(i) and (iv) that if ε is sufficiently small and $t \le T_\Delta$, $\nabla R(\eta(t)) \cdot b(\eta^\varepsilon(t)) + (\varepsilon/2)\Sigma R_{ij}(\eta(t))a_{ij}(\eta(t)) \le -C < 0$, where C is a constant. Using this and Assumption 3.2, we get that there is a constant M so that if $R(D) = \sup\{R(y)\colon\ y \in D\}$ and $T_\Delta < T$,

$$\delta \le R(D) - CT + M\int_0^{T_\Delta \wedge T} L[\eta^\varepsilon(t), \mathbf{v}^\varepsilon(t)]\,dt$$
$$+\sqrt{\varepsilon}\left|\int_0^{T_\Delta \wedge T} \nabla R(\eta(t)) \cdot \sigma(\eta(t))\,d\beta(t)\right|.$$

If T is large enough to ensure that $R(D) - CT \le 0$, this implies that

$$P_x(T_\Delta < T) \le P_x\left[\int_0^{T_\Delta \wedge T} ML[\eta^\varepsilon(t), \mathbf{v}^\varepsilon(t)]\,dt > \delta/2\right]$$
$$+P_x\left[\left|\int_0^{T} \nabla R(\eta(t)) \cdot \sigma(\eta(t))\,d\beta(t)\right| > \delta/2\sqrt{\varepsilon}\right].$$

This implies by Chebyshev's inequality that if $\alpha \in (0,1)$ is fixed and if $R(D) - CT \le 0$, then there is a constant A so that for sufficiently small ε and every $x \in \Delta$,

$$P_x(T_\Delta < T) \le AE_x\left[\int_0^{T_\Delta \wedge T} L[\eta^\varepsilon(t), \mathbf{v}^\varepsilon(t)]\,dt\right] + \alpha. \tag{5.6.29}$$

From this, it follows by a straightforward induction argument using the strong Markov property that for every $n \ge 1$,

$$P_x(T_\Delta > nT) \le \alpha^{n-1}P_x(T_\Delta > T) + A\sum_{k=2}^{n} \alpha^{n-k}E_x\left[\int_{T_\Delta \wedge (k-1)T}^{T_\Delta \wedge kT} L[\eta^\varepsilon(t), \mathbf{v}^\varepsilon(t)]\,dt\right].$$

We use this to estimate $E_x(T_\Delta)$:

$$\begin{aligned} E_x(T_\Delta) &= \int_0^\infty P_x(T_\Delta > t)\,dt \\ &\le T\sum_{n=0}^{\infty} P_x(T_\Delta > nT) \\ &\le T + T(1-\alpha)^{-1}P_x(T_\Delta > T) \\ &\quad + AT\sum_{k=2}^{\infty}\sum_{n=k}^{\infty} \alpha^{n-k}E_x\left[\int_{T_\Delta \wedge (k-1)T}^{T_\Delta \wedge kT} L[\eta^\varepsilon(t), \mathbf{v}^\varepsilon(t)]\,dt\right]. \end{aligned}$$

Interchanging the order of summation and using (5.6.29) again to estimate $P_x(T_\Delta > T)$, we deduce from this that for $x \in \Delta$,

$$\begin{aligned} E_x(T_\Delta) &\le T(1-\alpha)^{-1}\left\{1 + A\sum_{k=2}^{\infty} E_x\left[\int_{T_\Delta \wedge (k-1)T}^{T_\Delta \wedge kT} L[\eta^\varepsilon(t), \mathbf{v}^\varepsilon(t)]\,dt\right]\right\} \\ &= T(1-\alpha)^{-1}\left\{1 + AE_x\left[\int_0^{T_\Delta} L[\eta^\varepsilon(t), \mathbf{v}^\varepsilon(t)]\,dt\right]\right\}. \end{aligned}$$

This clearly implies (5.6.27) since $\Phi \ge 0$ and $L^\varepsilon(\cdot,\cdot) \ge L(\cdot,\cdot)$. $\square$

5.7. *An easier proof of the upper bound in the spherical case.* In this section, we give an easier proof of the estimate (5.2.3) in the case where τ has the special form (3.8) and $\sigma \equiv I_d$. In this case, $R(x) = |x|$. Since we are interested only in what happens until the exit time T_Δ from Δ, we can redefine τ in $\{y: R(y) < \delta/4\}$ in such a way as to remove the singularity at the origin, without affecting the process up to time T_Δ: By Itô's formula, if $R^\varepsilon(t) = R(x^\varepsilon(t))$, then $dR^\varepsilon(t) = [(R^\varepsilon(t))^{-1}x^\varepsilon(t) \cdot b(x^\varepsilon(t)) + \varepsilon(d-1)(2R^\varepsilon(t))^{-1}]\,dt + \sqrt{\varepsilon}\,d\omega(t)$, where $\omega(t)$ is a one-dimensional Brownian motion. For $s \in \Re$, define $b_\varepsilon^*(s) = -\beta(s \wedge \delta/2) + \varepsilon(d-1)(2s \wedge \delta)^{-1}$. Compare $R^\varepsilon(t)$ with the solution of

$$dr^\varepsilon(t) = b_\varepsilon^*(r^\varepsilon(t))\,dt + \sqrt{\varepsilon}\,d\omega(t).$$

Since, if $|x| \geq \delta/2$, $|x|^{-1}x \cdot b(x) + \varepsilon(d-1)(2|x|)^{-1} \leq b_\varepsilon^*(|x|)$, it follows by the comparison theorem for one-dimensional diffusions [13, Chapter VI, Theorem 1.1] that for $t \leq T_{\delta/2} \equiv \inf\{t \geq 0: |x^\varepsilon(t)| \leq \delta/2\}$, $R^\varepsilon(t) \leq r^\varepsilon(t)$. Now $b_\varepsilon^*(s) \to b^*(r) \equiv -\beta(s \wedge \delta/2)$ uniformly for $s \in \Re$, and $b^*(\cdot)$ is bounded and uniformly Lipschitz on $\Re$. Therefore, by Ventsel and Freidlin's result [25], if A is a closed subset of $C([0,T], \Re, r)$, where $T > 0$ and $r \in \Re$ are fixed, then

$$\limsup_{\varepsilon \to 0} \varepsilon \log P_r\big(r^\varepsilon|_{[0,T]} \in A\big) \leq -\inf\{I_1(g): g \in A\}, \tag{5.7.1}$$

where if $g \in C([0,T], \Re, r)$ and g is absolutely continuous, then

$$I_1(g) = \tfrac{1}{2}\int_0^T [g'(t) - b^*(g(t))]^2\,dt \tag{5.7.2}$$

and otherwise $I_1(g) = +\infty$. If T_δ and $T_{r(N)}$ denote the hitting times of δ and $r(N)$, respectively,

$$\begin{aligned} P_x^\varepsilon(x(T_\Delta) \in N) &\leq P_{|x|}\big(T_\delta(R^\varepsilon) > T_{r(N)}(R^\varepsilon)\big) \\ &\leq P_{|x|}\big(T_\delta(r^\varepsilon) > T_{r(N)}(r^\varepsilon)\big) \\ &\leq P_{|x|}\big(T_{r(N)}(r^\varepsilon) \leq T\big) + P_{|x|}\big(T_{r(N)}(r^\varepsilon) \geq T;\ T_\delta(r^\varepsilon) \geq T\big) \\ &\leq P_{|x|}\big(r^\varepsilon|_{[0,T]} \in A_1(T)\big) + P_{|x|}\big(r^\varepsilon|_{[0,T]} \in A_2(T)\big), \end{aligned}$$

where $A_1(T)$ and $A_2(T) \subset C([0,T], \Re, |x|)$ are the sets $A_1(T) = \{g: \sup\{|g(t)|: 0 \leq t \leq T\} \geq r(N)\}$ and $A_2(T) = \{g: \delta \leq g(t) \leq r(N), 0 \leq t \leq T\}$. Therefore, the estimate (5.2.3) will follow from (5.7.1) once we have established the two estimates

$$\inf\{I_1(g): g \in A_1(T)\} \geq 2\int_{|x|}^{r(N)} \beta(s)\,ds, \quad \text{for every } T > 0, \tag{5.7.3}$$

$$\inf\{I_1(g): g \in A_1(T)\} \to +\infty, \quad \text{as } T \to +\infty. \tag{5.7.4}$$

PROOF OF (5.7.3). By Ventsel and Freidlin's argument [25, Lemma 4.1, page 30], $\inf\{I_1(g): g \in A_1(T)\}$ equals $\inf\{I_1(g): g(0) = |x|; g(T) = r(N), g'(t) = -b^*(g(t))\}$. For g in this set, substituting into (5.7.2), using the fact that $\frac{1}{2}[g'(t) - b^*(g(t))]^2 = 2\beta(g(t))g'(t)$ and making a change of variables, one sees that $I_1(g)$ is equal to the right hand side of (5.7.3). □

PROOF OF (5.7.4). Since $\{y\colon\ d \le R(y) \le r(N)\}$ is compact and does not contain any ω-limit points of the dynamical system $x'(t) = b^*(x(t))$, this follows from [25, Lemma 3.2, page 24]. □

6. Proof of the exit result. In this section we shall complete the proofs of Theorem 3.16 and Corollary 3.19.

REMARK. [See (3.13), Definition 3.14, Assumption 3.15, and Theorem 3.16 for notation.] Since it follows from Remark 3.7 that for $x \in \{y\colon\ R(y) = r(D)\}$, $P_x^0[T'_{D^*}(x(\cdot)) = 0] = 1$, Assumption 3.15 is the same as, for $x \in D$,

$$(6.1)\quad P_x^0\big[T'_D(x(\cdot)) = T_D(x(\cdot));\ T_D(x(\cdot)) < \infty\big] = P_x^0\big[T_D(x(\cdot)) < \infty\big].$$

A sufficient condition under which this holds will be given at the end of this section.

We show first

PROPOSITION 6.2. *Let D^* be as in Definition 3.14. If $x \in D^*$ and $\mu_x^{*\varepsilon}$ and μ_x^* denote, respectively, the measures induced on ∂D^* by $x^\varepsilon(\cdot)$ and $x^0(\cdot)$, then for every $x \in D^*$,*

$$\mu_x^{*\varepsilon} \Rightarrow \mu_x^* \quad \textit{as } \varepsilon \to 0.$$

The following lemmas are needed in the proof of Proposition 6.2.

LEMMA 6.3. *For fixed $x \in \mathfrak{R}^d$, let P_x^0 be the measure induced on $C([0,\infty), \mathfrak{R}^d)$ by the solution of (1.1). There exists a constant $T(D)$ so that for every x in the closure of D^*,*

$$P_x^0(T_{D^*}(x) \le T(D)) = 1.$$

PROOF. See Remark 3.7. □

The proofs of the following two lemmas are elementary.

LEMMA 6.4. *Suppose $G \subset \mathfrak{R}^d$ is open. Suppose that $g_n \to g$ in $C([0,T], \mathfrak{R}^d)$. Then*

$$\liminf_{n\to\infty} T_G(g_n) \ge T_G(g).$$

If also $T'_G(g) \le T$,

$$\limsup_{n\to\infty} T'_G(g_n) \le T'_G(g).$$

LEMMA 6.5. *Suppose $G \subset \mathfrak{R}^d$ is open. Suppose $g \in C([0,T], \mathfrak{R}^d)$ and $T'_G(g) = T_G(g) < T$. Then if $g_n \to g$ in $C([0,T], \mathfrak{R}^d)$, $g_n(T_G(g_n)) \to g(T_G(g))$ as $n \to \infty$.*

The next result follows by the standard argument using Gronwall's inequality.

LEMMA 6.6. *Suppose $T > 0$. Let $x^0(\cdot)$ and $x^\varepsilon(\cdot)$ denote the solutions of* (1.1) *and* (1.3), *respectively. Then for any $x \in \Re^d$,*

$$E_x\Big[\big\{\sup\{|x^\varepsilon(t) - x^0(t)|: 0 \le t \le 2T\}\big\}^2\Big] \to 0 \quad \text{as } \varepsilon \to 0.$$

PROOF OF PROPOSITION 6.2. From Lemma 6.6 and Skorohod's theorem, it follows that one can construct random variables $x^{\sim\varepsilon}(\cdot)$ and $x^{\sim 0}(\cdot)$ taking values in $C([0,2T(D)], \Re^d)$ so that a.s. $x^{\sim\varepsilon}(\cdot) \to x^{\sim 0}(\cdot)$ in $C([0,2T(D)], \Re^d)$ as $\varepsilon \to 0$. From this and Lemmas 6.3 and 6.5, it follows that a.s. $x^{\sim\varepsilon}(T_{D^*}(x^{\sim\varepsilon})) \to x^{\sim 0}(T_{D^*}(x^{\sim 0}))$ as $\varepsilon \to 0$. Since $x^\varepsilon(\cdot)$ and $x^0(\cdot)$ have the same distribution as $x^{\sim\varepsilon}(\cdot)$ and $x^{\sim 0}(\cdot)$, this implies that if $x \in D^*$, $P_x[x^\varepsilon(T_{D^*}(x^\varepsilon)) \to x^0(T_{D^*}(x^0))$ as $\varepsilon \to 0] = 1$. From this, the result of Proposition 6.2 follows. □

PROOF OF THEOREM 3.16. From (3.11) and Remark 3.7, it follows that if $y \in K$, $P_y^0(T_D'(x(\cdot)) > 0) = 1$. Hence it is implicit in Assumption 3.15 [(6.1)] that if $x \in D$, then $P_x^0(x(T_D) \in K) = 0$.

If $N \subset \partial D$, and $x \in D^*$, then by conditioning on $\mathbf{F}_{T_D}$ and using the strong Markov property, one can show that

$$\mu_x^\varepsilon(N) = \mu_x^{*\varepsilon}(N) + E_x\Big[\chi_{\{y:\, R(y)=r(D)\}}(x^\varepsilon(T_{D^*}))P_{x^\varepsilon(T_{D^*})}(x^\varepsilon(T_D) \in N)\Big].$$

Lemma 3.12 enables us to handle the second term if either $N \cap K$ is empty or if $K \subset N^\circ$: $P_y^\varepsilon(x^\varepsilon(T_D) \in N)$ approaches 0 or 1 as $\varepsilon \to 0$ uniformly over $\{y\colon R(y) = r(D)\}$ according as $N \cap K = \varnothing$ or $K \subset N^\circ$. Therefore, using Proposition 6.2, we deduce that for $x \in D^*$,

(6.7) $\quad$ if $N \cap K = \varnothing$, then $\mu_x^\varepsilon(N) \to \mu_x^*(N)$ as $\varepsilon \to 0$,

(6.8) $\quad$ if $K \subset N^\circ$, then $\mu_x^\varepsilon(N) \to \mu_x^*(N) + \mu_x^*(\{y\colon R(y) = r(D)\})$ as $\varepsilon \to 0$.

In particular, if K consists of a single point y_0, then (6.7) and (6.8) imply that for $x \in D^*$, as $\varepsilon \to 0$, $\mu_x^\varepsilon(N)$ approaches $\mu_x^*(N) + \mu_x^*(\{y\colon R(y) = r(D)\})\chi_N(y_0)$. Therefore, as $\varepsilon \to 0$,

$$\mu_x^\varepsilon \Rightarrow \mu_x^* + \mu_x^*(\{y\colon R(y) = r(D)\})\delta_{y_0}.$$

This proves Theorem 3.16 for $x \in D^*$. If $x \in D$ and $R(x) = r(D)$, then $\mu_x^* = \delta_x$, so $\mu_x^*|_{\partial D} = 0$. If $R(x) < r(D)$, then $P_x(T_{D^*}(x(\cdot)) < \infty) = 0$, $\mu_x^* = 0$. In both cases the result follows from Lemma 3.12. □

PROOF OF COROLLARY 3.19. It follows from Lemma 6.3 that the Dirichlet problem $\mathbf{L}u^* = 0$ in D^*; $u^*|_{\partial D^*} = g^*$ has a solution. For $\varepsilon > 0$, define $u^\varepsilon(x)$ as in (3.20). It follows from standard results that u^ε solves the boundary value problem: $\mathbf{L}^\varepsilon u^\varepsilon = 0$ in D; $u^\varepsilon|_{\partial D} = g$. The solution to this Dirichlet problem need not be unique (see [15, page 28]). However, the process $x^\varepsilon(\cdot)$ hits zero with probability zero so the singularity of τ at zero makes no difference. (3.21) follows immediately from Theorem 3.16. □

A sufficient condition for (6.1). As in (3.13), T_D and T_D' denote the exit times from D and $\bar{D}$, respectively.

DEFINITION 6.9. If $y \in \partial D$, then y is a *regular point* of ∂D if $P_y^0(T_D' > 0) = 0$. Otherwise y is an *irregular point.*

If $y \in \partial D$, let $n(y)$ be the exterior normal to ∂D at y, and let $n^*(y)$ denote the exterior normal to $\{x: R(x) < R(y)\}$. Under the assumptions of Section 3, for $y \in \partial D$, $n^*(y) = \nabla R(y)/|\nabla R(y)|$.

In [20], Stroock and Varadhan gave necessary conditions for a boundary point to be irregular. It follows from their results that in the present setting all boundary points y are regular except perhaps if $n(y) = n^*(y)$. Depending on the relative curvatures of the two surfaces ∂D and $\partial\{x: R(x) = R(y)\}$, such points may be irregular. In particular, all the points of the set K are irregular.

If we can show that starting from any point $x \in \Re^d$, the probability of hitting any other point $y \in \Re^d$ is zero, then under the following hypothesis, it will follow that (6.1) holds.

HYPOTHESIS 6.10. *The set* $\{y \in \partial D: n(y) = n^*(y)\}$ *is countable.*

We need to make one further assumption:

ASSUMPTION 6.11. Assumption on $c(\cdot)$. From now on we shall assume that the following condition is satisfied. For $x \in \Re^d$, let $P(x)$ denote the orthogonal projection onto the range of $c(x)$. Assume that there is a constant $\lambda > 0$ so that for every $x \in \Re^d$ and every $\xi \in \Re^d$,

$$\lambda |P(x)\xi|^2 \le c(x)P(x)\xi \cdot P(x)\xi \le \lambda^{-1}|P(x)\xi|^2.$$

LEMMA 6.12. *For* $y \in \Re^d$, *define* $T_y = \inf\{t \ge 0: x(t) = y\}$. *Under Assumption* 6.11, *if* $x \ne y$, *then*

$$P_x^0(T_y < \infty) = 0.$$

PROOF. The proof, which we omit, follows standard lines: It involves the construction of functions $g_n(x, t)$ such that if $\mathbf{L}_t = \mathbf{L} + \partial/\partial t$, then $\mathbf{L}_t g_n(x, t) \le 0$ outside a neighbourhood of y, and $g_n(x, t) \to \infty$ as $x \to y$ and $n \to \infty$. □

REFERENCES

[1] AZENCOTT, R. (1980). Grandes deviations et applications. *Ecole d'Été de Probabilités de Saint-Flour, VIII—1978. Lecture Notes in Math.* **774** 1–176. Springer, Berlin.

[2] BEZUIDENHOUT, C. (1985). Small random perturbations of stochastic systems. Ph.D. Thesis, Univ. of Minnesota.

[3] BEZUIDENHOUT, C. (1987). One-dimensional Ventsel–Freidlin theory and its analogue for singular perturbations of degenerate diffusions. *Stochastics* **20** 151–167.

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

[5] FLEMING, W. H. (1978). Exit probabilities and optimal stochastic control. *Appl. Math. Optim.* 4 329–346.

[6] FLEMING, W. H. and RISHEL, R. W. (1975). *Deterministic and Stochastic Optimal Control.* Springer, New York.

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

[8] FREIDLIN, M. I. and GARTNER, J. (1978). A new contribution to the question of large deviations for random processes. *Vestnik Moskov. Univ. Ser. I Mat. Meh.* **5** 52–59.

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

[10] FRIEDMAN, A. (1976). *Stochastic Differential Equations and Applications*. Academic, New York.

[11] GARTNER, J. (1978). On action functional for diffusion processes with small parameter in a part of higher derivatives. Preprint, Zentralinstitut für Mathematik und Mechanik, Berlin, GDR (in Russian).

[12] GIRSANOV, I. V. (1960). On transforming a certain class of stochastic processes by absolutely continuous substitution of measures. *Theory Probab. Appl.* **5** 285–301.

[13] IKEDA, N. and WATANABE, S. (1981). *Stochastic Differential Equations and Diffusion Processes*. North-Holland, Amsterdam.

[14] ITÔ, K. (1975). Stochastic differentials. *Appl. Math. Optim.* **1** 374–381.

[15] PUCCI, C. (1966). Limitazione per soluzioni de equazione ellittichi. *Ann. Mat. Pura Appl.* **74** 15–30.

[16] SCHILDER, M. (1966). Some asymptotic formulas for Wiener integrals. *Trans. Amer. Math. Soc.* **125** 63–85.

[17] SPIVAK, M. (1979). *A Comprehensive Introduction to Differential Geometry* **1**. Publish or Perish, Berkeley, Calif.

[18] STROOCK, D. W. (1971). On the growth of stochastic integrals. *Z. Wahrsch. verw. Gebiete* **18** 340–344.

[19] STROOCK, D. (1984). *An Introduction to the Theory of Large Deviations*. Springer, New York.

[20] STROOCK, D. W. and VARADHAN, S. R. S. (1972). On degenerate elliptic-parabolic operators of second order and their associated diffusions. *Comm. Pure Appl. Math.* **25** 651–713.

[21] STROOCK, D. W. and VARADHAN, S. R. S. (1972). On the support of diffusion processes, with applications to the strong maximum principle. *Proc. Sixth Berkeley Symp. Math. Statist. Probab.* **3** 333–360. Univ. California Press.

[22] STROOCK, D. W. and VARADHAN, S. R. S. (1979). *Multidimensional Diffusion Processes*. Springer, New York.

[23] VARADHAN, S. R. S. (1980). *Lectures on Diffusion Problems and Differential Equations*. Tata Institute of Fundamental Research, Bombay, India.

[24] VARADHAN, S. R. S. (1984). *Large Deviations and Applications*. SIAM, Philadelphia.

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

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WISCONSIN
MADISON, WISCONSIN 53706