



Mathematical Approaches to the Problem of Noise-Induced Exit

Martin V. Day

ABSTRACT: We provide an overview of some of the methods that have been used to study sub-large deviations phenomena in the small noise exit problem. Both formal asymptotic methods and rigorous probabilistic methods are considered and compared.

KEYWORDS: Exit Problem; Small Noise.

1 Introduction

We consider here the basic problem of small noise induced exit of a finite dimensional system from an asymptotically stable equilibrium point. In \mathbb{R}^d let $\omega(t)$ be a standard Wiener process and consider the diffusion process

$$dx^\epsilon(t) = b(x^\epsilon(t)) dt + \epsilon^{1/2} \sigma(x^\epsilon(t)) d\omega(t), \quad (1.1)$$

starting at $x^\epsilon(0) \in D$, where D is a basin of attraction for an asymptotically stable critical point at 0 for the deterministic system

$$\dot{x}(t) = b(x(t)). \quad (1.2)$$

For an initial $x = x^\epsilon(0) \in D$ we follow $x^\epsilon(t)$ up to its first contact with ∂D :

$$\tau_D^\epsilon = \inf\{t > 0 : x^\epsilon(t) \in \partial D\}.$$

Our interest is in the asymptotic behavior ($\epsilon \downarrow 0$) of the *exit point distribution*

$$\mu_x^\epsilon(dy) = P_x[x^\epsilon(\tau_D^\epsilon) \in dy] \quad (1.3)$$

and the *mean exit time*

$$v_D^\epsilon(x) = E_x[\tau_D^\epsilon]. \quad (1.4)$$

The main offering of this paper is the overview in Section 3 of some of the approaches which have been used to obtain sub-large deviations results. By this we mean results that describe properties of (1.3) and (1.4) that cannot be determined from the usual large deviations results [17] describing

exponential dependence on ϵ . It is the mathematical concepts behind the methods that we discuss, rather than specific results which apply under diverse and various hypotheses on ∂D . Some of the approaches we describe have been given rigorous treatments in the literature, others are based only on formal asymptotic calculations. The latter have played an important role in the discovery and explanation of significant features, though they do not by themselves provide “proof” in the mathematical sense. Beyond theoretical and formal asymptotic approaches, analogue simulations have also been used to observe features of the exit problem firsthand; see [11] and references. More sophisticated digital simulations based on the importance sampling technique were conducted in [10].

A long-standing interest of Wendell Fleming’s has been small noise ($\epsilon \downarrow 0$) limits in stochastic control problems, particularly as regards the Hamilton-Jacobi-Bellman equation which describes the associated value function ϕ^ϵ , and its asymptotic relation to that of its deterministic ($\epsilon = 0$) counterpart. A logarithmic transformation $\phi^\epsilon = -\epsilon \log(u^\epsilon)$ typically converts solutions u^ϵ of linear equations (such as those associated with \mathcal{L}^ϵ below) into Hamilton-Jacobi-Bellman equations of this small noise type. This allows the small noise asymptotics of appropriate control problems to be used to obtain large deviations results for the exit problem. Fleming’s paper [14] broke new ground in this area. The advent of viscosity solution techniques has made these analyses possible without explicit recourse to control arguments; see [13] for instance. The 1990 paper of Perthame [29] is a development of these ideas specifically in the context of the exit problem as we are considering it here. We will comment more on that work in Section 3.2 below. This general approach is not limited to Wentzel-Freidlin type large deviations but can also be used for results on occupation measures, principal eigenvalues and other issues; see [15] and [16]. In fact the logarithmic connection to stochastic optimization problems has been developed by Dupuis and Ellis [9] into a general approach to large deviations analysis of broad scope.

2 Basic Results

We review in this section a number of results to which our discussion in Section 3 refers. Two special cases of the exit problem will be mentioned frequently. These depend on whether (1.2) crosses ∂D with a positive angle or is tangent to it. Let n denote the (unit) outward normal to ∂D . The *nontangential case* is succinctly described by $n \cdot b < 0$ at all points $y \in \partial D$. This was the case originally treated by the large deviations analysis in [17]. In many applications D is taken to be the maximal domain of attraction of the stable point at $x = 0$. This makes the *characteristic boundary*, $n \cdot b = 0$, a more natural hypothesis, though more difficult mathematically.

2.1 PDE Formulations

The generator of the diffusion process (1.1) is the differential operator

$$\mathcal{L}^\epsilon u(x) = \frac{\epsilon}{2} \sum_{i,j} a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} u(x) + \sum_i b_i(x) \frac{\partial}{\partial x_i} u(x). \quad (1.5)$$

The diffusion matrix is $a(x) = \sigma(x)\sigma(x)^T$. Typical regularity assumptions would require $a(\cdot), b(\cdot)$ to be C^3 , with a positive definite hypothesis on $a(\cdot)$. Similarly D is assumed to be bounded with C^3 boundary. However, we will not concern ourselves with precise hypotheses, since our goal is to provide only a heuristic overview.

The quantities of interest in the exit problem can be characterized in terms of appropriate boundary value problems associated with \mathcal{L}^ϵ . The exit point distribution (1.3) is intimately connected with the Dirichlet problem in D : given a continuous function $f(y)$ defined on ∂D , find $u(x) = u_f^\epsilon(x)$ solving

$$\mathcal{L}^\epsilon u(x) = 0, \quad x \in D; \quad u(y) = f(y), \quad y \in \partial D. \quad (1.6)$$

Indeed as is well-known,

$$u_f^\epsilon(x) = E_x[f(x^\epsilon(\tau_D^\epsilon))] = \int_{\partial D} f(y) \mu_x^\epsilon(dy).$$

With regard to the mean exit time (1.4), its distribution is linked to \mathcal{L}^ϵ through the Poisson equation

$$\mathcal{L}^\epsilon v(x) + \phi(x) = 0, \quad x \in D; \quad v(y) = 0, \quad y \in \partial D \quad (1.7)$$

by the solution formula

$$v(x) = E_x \left[\int_0^{\tau_D^\epsilon} \phi(x^\epsilon(t)) dt \right].$$

In particular, for $\phi \equiv 1$ we get the mean exit time, $v_D^\epsilon(x) = E_x[\tau_D^\epsilon]$.

2.2 Green's Identity

A principal tool in exploiting these PDE characterizations is Green's identity, which provides a fundamental relation between \mathcal{L}^ϵ and its formal adjoint:

$$\mathcal{L}^{\epsilon*} g(x) = \frac{\epsilon}{2} \sum_{i,j} \frac{\partial^2}{\partial x_i \partial x_j} [a_{ij}(x) f(x)] - \sum_i \frac{\partial}{\partial x_i} [b_i(x) f(x)]. \quad (1.8)$$

The unit outward normal and *conormal* at $y \in \partial D$ will be

$$n(y) = (n_i(y)), \quad \eta(y) = a(y)n(y).$$

Two operators defined on ∂D appear in (1.11) below. One is a scaled conormal derivative:

$$\mathcal{N}^\epsilon f(y) = \frac{\epsilon}{2} \frac{\partial}{\partial \eta} f(y) = \frac{\epsilon}{2} \sum_{i,j} n_i a_{ij} \frac{\partial}{\partial x_j} f(y). \quad (1.9)$$

The second is a sort of “boundary adjoint” to \mathcal{N}^ϵ :

$$\mathcal{N}^{\epsilon*} g(y) = \frac{\epsilon}{2} \sum_{i,j} n_i \frac{\partial}{\partial x_j} [a_{ij} g(y)] - \langle n, b \rangle \cdot g(y). \quad (1.10)$$

Green’s (second) identity, equation (6.5) of [27], is a fundamental relation among these operators:

$$\int_D g(x) \mathcal{L}^\epsilon f(x) - f(x) \mathcal{L}^{\epsilon*} g(x) dx = \int_{\partial D} g(y) \mathcal{N}^\epsilon f(y) - f(x) \mathcal{N}^{\epsilon*} g(y) dy. \quad (1.11)$$

2.3 Leveling Results

The exit point distributions μ_x^ϵ need not in general converge as $\epsilon \downarrow 0$. However the dependence on the initial point x vanishes in the limit. This is manifested in the leveling property of the solution u_f^ϵ to the Dirichlet problem (1.6): for any compact $K \subseteq D$,

$$\sup_{x,y \in K} |u_f^\epsilon(x) - u_f^\epsilon(y)| \rightarrow 0 \text{ as } \epsilon \downarrow 0.$$

This was proven in [8] for the nontangential case, but the concluding remarks show that it holds for characteristic boundaries as well. See [12] also. If in fact μ_x^ϵ does converge as $\epsilon \downarrow 0$ then the leveling property implies that the limit $\mu_x^\epsilon \Rightarrow \mu^0$ is independent of $x \in D$ and that for a given f there is a constant

$$C_f^0 = \int_{\partial D} f d\mu^0 \quad (1.12)$$

so that $u_f^\epsilon \rightarrow C_f^0$ uniformly on compacts. In this case finding μ^0 is equivalent to identifying a formula for C_f^0 .

A relative leveling result for the mean exit time is provided by Corollary 1 of [7]:

$$\lambda^\epsilon E_x[\tau_D^\epsilon] \rightarrow 1, \quad (1.13)$$

holding uniformly on any compact $K \subseteq D$. Here $\lambda^\epsilon > 0$ is the principal eigenvalue for \mathcal{L}^ϵ with Dirichlet boundary data on ∂D : $\mathcal{L}^\epsilon \phi^\epsilon + \lambda^\epsilon \phi^\epsilon = 0$ for $\phi^\epsilon > 0$ in D with $\phi^\epsilon = 0$ on ∂D . Because of (1.13) analysis of the principal eigenvalue is often considered equivalent to analysis of the mean exit time. The treatment in [7] only considered the nontangential case. To our knowledge (1.13) has not been worked out in the characteristic boundary case,

although it may be implied by some of the sophisticated work on exponential loss of memory and long-time asymptotics: [23], [24] and [30] for instance. As we will see in Section 3.2, (1.13) would be useful in the characteristic boundary case.

2.4 Hasminskii's Formula

Many of the approaches to sub-large deviations phenomena can be viewed as expressions involving an invariant density for some version of $x^\epsilon(t)$. We summarize here a general formula for the invariant distribution in terms of an embedded Markov chain. Although this formula goes back much further in the context of Markov chains, it was developed in the context of diffusion processes by Hasminskii in 1960. An exposition can be found in [18]; see Ch. IV, §4. We will use this formula below for several variants of the original $x^\epsilon(t)$ of (1.1), to produce rather explicit relations among the invariant distribution, exit point distribution and exit time. We describe it here in general for a process $x(t)$, which will refer to any of the variants actually used below.

The idea is to construct a sequence of stopping times

$$0 < \tau_1 < \tau_2 < \dots$$

associated with contact times of a (compact) subset Γ . This is typically done using a second disjoint subset Γ' : with $\tau_0 = 0$ define

$$\begin{aligned} \tau'_i &= \inf\{t > \tau_i : x(t) \in \Gamma'\} \\ \tau_{i+1} &= \inf\{t > \tau'_i : x(t) \in \Gamma\}. \end{aligned}$$

The essential technical features are that $E_x[\tau_1] < \infty$, uniformly over $x \in \Gamma$, and that there is some sort of separation between Γ and Γ' which insures $\tau_i < \tau_{i+1}$ with positive probability. The idea is to view the history of $x(t)$ as a sequence of cycles, $\tau_i < t \leq \tau_{i+1}$, punctuated by the Markov chain on Γ determined by

$$X_i = x(\tau_i).$$

Generally X_i will have a (unique) invariant measure ν on Γ , from which the (unique) invariant measure π for $x(t)$ is constructed by means of *Hasminskii's formula*:

$$\int f(x) \pi(dx) = \frac{1}{E_\nu[\tau_1]} \int_\Gamma E_z \left[\int_0^{\tau_1} f(x(s)) ds \right] \nu(dz). \quad (1.14)$$

3 Methods for Sub-Large Deviations Asymptotics

Large deviations results provide definitive descriptions of the exponential ϵ -dependence of the mean exit time and the exit point distribution. However,

in some situations there are interesting phenomena which can not be seen through this kind of exponential asymptotics. This is true particularly in cases of characteristic boundary (see [3]). For instance large deviations analysis will identify a special subset $\Gamma^* \subseteq \partial D$ on which the exit point distribution μ_x^ϵ must concentrate as $\epsilon \downarrow 0$. However the distribution of μ_x^ϵ within Γ^* is not discernible through large deviations results. In particular if ∂D is a limit cycle of (1.2) then Γ^* will necessarily be all of ∂D , so something other than large deviations is needed to study the exit point distribution. If Γ^* consists of isolated points there is the question of the relative weighings of these points in μ_x^ϵ as $\epsilon \downarrow 0$. Even if Γ is a single point, large deviations does not tell us anything about the asymptotic “shape” of μ_x^ϵ . The issue of skewing of the exit point distribution is a prime example. One may be even more interested in the effect of these features on the mean exit time, and less than exponential aspects of its ϵ -dependence. These are all examples of what we call “sub-large deviations” phenomena. It is the mathematical ideas that have been used to address these kind of questions that we overview in this section.

A variety of methods have been used to study sub-large deviations features. Some are probabilistic in nature, while others are based on PDE techniques. Our purpose in this section is to summarize a number of these methods, and indicate their relations to each other using Green’s identity and Hasminkskii’s formula. We have tried to use a consistent notation to do this. As a result the notation below is sometimes different that that of the original sources.

3.1 The Original Matkowsky-Schuss Method

The 1977 paper [25] of Matkowsky and Schuss introduced a fruitful formal approach to sub-large deviations results for the exit problem. Initially developed in the nontangential case $n \cdot b < 0$, it was subsequently applied to cases with characteristic boundary; see [26]. The basic idea is to use a solution z^ϵ of the adjoint equation $\mathcal{L}^{\epsilon*} z^\epsilon = 0$ together with u_f^ϵ of (1.6) in Green’s identity, resulting in

$$0 = \int_{\partial D} z^\epsilon \mathcal{N}^\epsilon u_f^\epsilon - f \mathcal{N}^{\epsilon*} z^\epsilon. \quad (1.15)$$

This is exploited by postulating suitable asymptotic forms for z^ϵ and u_f^ϵ :

$$z^\epsilon(x) \sim e^{\phi(x)/\epsilon} \sum \epsilon^{n/2} w_n(x). \quad (1.16)$$

and

$$u_f^\epsilon(x) \sim C_f^0 + (f(x) - C_f^0) e^{\zeta(x)/\epsilon} \quad (1.17)$$

The function $\zeta(x)$ should be 0 on ∂D but negative inside D , so that (1.17) produces a function which adjusts from the interior constant C_f^0 to the

boundary data f in a layer near ∂D . Assuming that the exit point distributions do converge, (1.17) agrees nicely with the leveling property (1.12). Careful asymptotic calculations proceed from (1.17) and (1.16). In the course of these calculations one finds that the term $f\mathcal{N}^{\epsilon*}z^\epsilon$ is negligible with respect to the first term in (1.15), and is therefore dropped from the calculations. The determination of C_f^0 is thus based on using (1.16) and (1.17) in

$$0 \sim \int_{\partial D} z^\epsilon \mathcal{N}^\epsilon u_f^\epsilon. \quad (1.18)$$

The resulting formulas involve terms that describe the distribution of μ^0 within the support set Γ^* , thus giving sub-large deviations information. For the mean exit time (1.4) we know that $\mathcal{L}^{\epsilon*}v_D^\epsilon = -1$ with $v_D^\epsilon(y) = 0$ on ∂D . Using these facts and z^ϵ in Green's identity implies

$$\int_D -z^\epsilon = \int_{\partial D} z^\epsilon \mathcal{N}^\epsilon v_D^\epsilon.$$

by postulating the asymptotic expression

$$v_D^\epsilon(x) \sim c^\epsilon e^{K/\epsilon} (1 - e^{\zeta/\epsilon}).$$

Expressions for c^ϵ are derived, which provide sub-large deviations information on the mean exit time.

3.2 Boundary Local Time Formulation

The success of [25] prompted several efforts to put those results on a rigorous basis: [19], [6], [29]. These developments all assumed the nontangential case, $n \cdot b < 0$. In [5] the probabilistic approach of [6] was refined for use in the context of characteristic boundaries. This is the method we describe in this section. It provides a probabilistic redevelopment of the original Matkowsky-Schuss method by interpreting $z^\epsilon(x)$ as the equilibrium density $p^\epsilon(x)$ of a reflected version of the diffusion $x^\epsilon(t)$. Shortly after [5], the exit conditioning approach described in Section 3.5 below emerged and provided a way to obtain qualitative results without needing precise, rigorous asymptotics of an equilibrium density. As a consequence these boundary local time ideas have not been exploited to the point of producing explicit sub-large deviations results. Even so, the method remains interesting as a probabilistic interpretation of the original Matkowsky-Schuss method. Consider the diffusion process $x^r(t)$ in \bar{D} which obeys (1.1) in D but with conormal reflection from ∂D . This can be characterized as the solution to the following stochastic differential equation with reflection term (see [1]):

$$dx^r(t) = b(x^r(t)) dt + \epsilon^{1/2} \sigma(x^r(t)) d\omega(t) - \frac{\epsilon}{2} \eta(x^r(t)) d\ell^\epsilon(t). \quad (1.19)$$

Here $\ell^\epsilon(t)$ is the *boundary local time* process, a continuous monotone process which increases only when $x^r(t) \in \partial D$: $d\ell^\epsilon(t) = 1_{\partial D}(x^r(t)) d\ell^\epsilon(t)$. We

will not go into the existence, uniqueness theory of (1.19). Suffice it to say that unique solutions exist with appropriate regularity properties, with $x^r(t) \in \bar{D}$ for all $t \geq 0$, satisfying a version of Itô's lemma: for $u \in C^2(\bar{D})$:

$$\begin{aligned} du(x^r(t)) &= \mathcal{L}^\epsilon u(x^r(t)) dt + \epsilon^{1/2} \nabla u(x^r(t)) \cdot \sigma(x^r(t)) d\omega(t) \\ &\quad - \mathcal{N}^\epsilon u(x^r(t)) d\ell^\epsilon(t). \end{aligned} \quad (1.20)$$

From this it is clear that, for smooth functions u , the generator of $x^r(t)$ is given by $\mathcal{L}^\epsilon u$ restricted to those functions with $\mathcal{N}^\epsilon u = 0$ on ∂D .

The reflected process $x^r(t)$ has a unique invariant density $p^\epsilon(x)$ (see [5]). We want to consider p^ϵ as the function z^ϵ in the Matkowsky-Schuss approach above. The invariance implies $0 = \int_D p^\epsilon \mathcal{L}^\epsilon u$ for all smooth u with $\mathcal{N}^\epsilon u = 0$ on ∂D . Applying this in Green's identity for u having compact support in D leads to $\mathcal{L}^{\epsilon*} p^\epsilon = 0$. Now Green's identity implies $0 = \int_{\partial D} u \mathcal{N}^{\epsilon*} p^\epsilon$ for all $\mathcal{N}^\epsilon u = 0$. From this follows the necessary boundary conditions for p^ϵ on ∂D :

$$\mathcal{N}^{\epsilon*} p^\epsilon = 0. \quad (1.21)$$

Moreover, applying Green's identity yet again to u_f^ϵ and p^ϵ tells us that

$$0 = \int_{\partial D} p^\epsilon \mathcal{N}^\epsilon u_f^\epsilon, \quad (1.22)$$

for all nice f defined on ∂D . Thus with p^ϵ as z^ϵ in the Matkowsky-Schuss development, (1.21) makes the negligibility of $f \mathcal{N}^{\epsilon*} z^\epsilon$ precise, and (1.22) is an exact version of (1.18).

Further probabilistic reasoning reveals a mechanism behind the asymptotic formulas resulting from this method. Let $C \subseteq D$ be a neighborhood of 0 inside D . Suppose we start $x^r(t)$ at $x^r(0) = y \in \partial D$ and follow it until its first contact with ∂C and then its subsequent return τ_1 to ∂D :

$$\begin{aligned} \tau_C^\epsilon &= \inf\{t > 0 : x^r(t) \in C\} \\ \tau_1 &= \inf\{t > \tau_C^\epsilon : x^r(t) \in \partial D\}. \end{aligned}$$

For $y \in \partial D$ and $g(x) \in C(\bar{D})$, define the *mean excursion operator*,

$$T^\epsilon g(y) = E_y[g(x^r(\tau_1))].$$

For smooth $f(y)$ on ∂D , define the *boundary generator*

$$\mathcal{G}^\epsilon f(y) = -\mathcal{N}^\epsilon u_f^\epsilon(y)$$

and the *boundary local time operator*

$$B^\epsilon f(y) = E_y\left[\int_0^{\tau_1} f(x^r(t)) d\ell^\epsilon\right],$$

both for $y \in \partial D$. Applying Itô's formula (1.20) to u_f^ϵ , integrating over $[0, \tau_1]$ and taking the expectation produces the relationship among these operators:

$$T^\epsilon f - f = B^\epsilon[\mathcal{G}^\epsilon f] \quad \text{on } \partial D. \quad (1.23)$$

The asymptotic behavior of the exit location distribution is given by the behavior of T^ϵ as $\epsilon \downarrow 0$. Suppose for instance that a limiting exit measure $\mu_x^\epsilon \Rightarrow \mu^0$ exists. Then we have

$$T^\epsilon f(y) \rightarrow \int_{\partial D} f d\mu^0,$$

with leveling results implying that the convergence is uniform in $y \in \partial D$. As we have described it, the Matkowsky-Schuss approach is essentially the application in (1.22) (or (1.18)) of appropriate asymptotic formulas for $\mathcal{G}^\epsilon f = -\mathcal{N}^\epsilon u_f^\epsilon$ and p^ϵ on ∂D . Now the primary goal of [5] was to establish an appropriate scaled limit for B^ϵ : with $c^\epsilon = \epsilon^{1/2}$:

$$c^\epsilon B \rightarrow B, \quad (1.24)$$

along with an inversion formula, B^{-1} . (The scaling factor $c^\epsilon = \epsilon^{1/2}$ is for non-degenerate characteristic boundary. For the nontangential case, $c^\epsilon = 1$ would be appropriate.) So instead of arguing from (1.22), we can use a scaled limit of (1.23) to find the relationship between μ^0 and limiting forms of the various operators. Proceeding purely speculatively, suppose that with some scaling factor k^ϵ the invariant density converges to a probability measure on ∂D :

$$k^\epsilon p^\epsilon(y) dy \Rightarrow \nu(dy). \quad (1.25)$$

Convergence of T^ϵ in (1.23) suggests $1/c^\epsilon$ as a scaling constant for \mathcal{G}^ϵ : $\frac{1}{c^\epsilon} \mathcal{G}^\epsilon \rightarrow \mathcal{G}$. The scaled limits of (1.23) and (1.22) would then be

$$\begin{aligned} \int_{\partial D} f d\mu^0 - f &= B[\mathcal{G}f], \\ \int_{\partial D} \mathcal{G}f d\nu &= 0. \end{aligned}$$

Taking B^{-1} of the first line and then integrating with respect to ν produces the formula

$$\int_{\partial D} f d\mu^0 = \frac{\int B^{-1}[f] d\nu}{\int B^{-1}[1] d\nu}. \quad (1.26)$$

This is of course purely formal. However it makes the important point that the asymptotic analysis of normal derivatives (\mathcal{G}^ϵ) ought to be replaceable by asymptotic analysis of the boundary local time operator, whose asymptotics have been worked out rigorously in [5].

The equation (1.22) at the heart of the Matkowsky-Schuss approach is an expression of the invariance of the density p^ϵ . Hasminskii's formula provides an alternate expression of this invariance, which links it directly to the boundary local time operator. This provides a rigorous approach to

the relationship (1.26) among the exit time, exit location and invariant density which our hypothetical calculation suggested. Moreover such an approach avoids the conormal derivative \mathcal{G}^ϵ and other differential operators altogether, using only mean path integral expressions. This reduces the exit problem completely to the study of (1.25), or its counterpart in nonconvergent situations. The idea is to use the excursions from ∂D to ∂C and back again as the basic cycles in Hasminskii's formula, and then form a sort of conormal derivative of the resulting version of (1.14). The basic formula that results (see [5] for details) is the following, holding for any $\epsilon > 0$:

$$\int_{\partial D} g(y) p^\epsilon(y) dy = \frac{\int_{\partial D} B^\epsilon g(y) \mu_D^\epsilon(dy)}{E_{\mu_D^\epsilon}[\tau_1]}, \quad \text{all } g \in C(\partial D). \quad (1.27)$$

Here μ_D^ϵ is the (unique) T^ϵ -invariant probability measure on ∂D . The leveling results tell us that $\mu_D^\epsilon \sim \mu_x^\epsilon$ for all $x \in D$. Note that (1.22) follows from this by taking $g = \mathcal{G}^\epsilon f$ and using (1.23): $\int_{\partial D} p^\epsilon \mathcal{G}^\epsilon f dy = c \int_{\partial D} (T^\epsilon f - f) d\mu_D^\epsilon = 0$. Thus (1.27) is an expression of the same relationship between p^ϵ and u_f^ϵ as before, but in which the three basic quantities p^ϵ , μ_D^ϵ and $E[\tau_1]$ appear explicitly.

To illustrate the use of (1.27) assume again the convergent case (1.25). We have a scaled limit (1.24) of B^ϵ . Then using $g \equiv 1$ in (1.27) we obtain the appropriate scaling of the mean excursion time:

$$\frac{c^\epsilon}{k^\epsilon} E_{\mu_D^\epsilon}[\tau_1] \rightarrow c = \int B[1] d\mu^0.$$

Using this in (1.27) gives the general limiting relationship:

$$c \int_{\partial D} g d\nu = \int_{\partial D} B[g] d\mu^0.$$

Replacing g with $B^{-1}[f]$ we find

$$\int f d\mu^0 = c \cdot \int B^{-1}[f] d\nu, \quad (1.28)$$

holding for all $f \in C(\partial D)$. Using $f \equiv 1$ gives the alternate formula $c = 1 / \int_{\partial D} B^{-1}[1] d\nu$, so that (1.28) confirms the speculative formula (1.26), when (1.25) holds. If $k^\epsilon p^\epsilon(y) dy$ does not converge, a similar analysis with the known asymptotics of B^ϵ will connect the possible limit points of μ^0 with those of $k^\epsilon p^\epsilon(y) dy$ on ∂D .

One could also obtain asymptotic expressions for the mean exit time. With a version of (1.13) for the characteristic boundary case one could justify writing

$$\begin{aligned} E_{\mu_D^\epsilon}[\tau_1] &= E_{\mu_D^\epsilon}[\tau_C^\epsilon] + E_{\mu_D^\epsilon}[E_{x^r(\tau_C^\epsilon)}[\tau_D^\epsilon]] \\ &= (o(1) + 1)E_0[\tau_D^\epsilon], \end{aligned}$$

for some $\delta > 0$. Therefore $E_0[\tau_D^\epsilon] \sim E_{\mu_D^\epsilon}[\tau_1] \sim \frac{k^\epsilon}{c^\epsilon} c$.

Perthame [29] gives a rigorous development of the Matkowsky-Schuss approach based on PDE-viscosity methods. Although it uses no probabilistic reasoning, there are a number of close parallels with the boundary local time ideas above that we wish to point out. First note that in place of Matkowsky and Schuss' $z^\epsilon(x)$ he uses $v_\epsilon > 0$ solving $\mathcal{L}^{\epsilon*} v_\epsilon = 0$ with boundary conditions $\mathcal{N}^{\epsilon*} v_\epsilon = 0$. Thus his v_ϵ corresponds exactly with the equilibrium density p^ϵ for the conormally reflected process $x^r(t)$. Perthame's hypothesis (H3) rules out fully characteristic ∂D . In the nontangential case $n \cdot b < 0$ it turns out that the scaled limit in (1.24) with $c^\epsilon = 1$ is just a multiplicative operator:

$$B[f] = \frac{-1}{n \cdot b} f; \quad B^{-1}[f] = -(n \cdot b) f.$$

(This was not derived in [5], but follows by applying the same methods in the simpler nontangential case.) In light of this we recognize Perthame's equation (7) as an approximate version of our (1.23): $\mathcal{G}^\epsilon f \sim B^{-1}[T^\epsilon f - f]$. (He attributes this to Kamin [19] originally.) His Theorem 3 is an exponential leveling result, with the limiting formula for the nontangential case.

3.3 Green's Function Approach

A new approach to formal asymptotic calculations was presented by Naeh, Klosek, Matkowsky and Schuss in [28]. In contrast to the method of Section 3.1 which involved asymptotic representations of solutions to both forward and backward equations, as well as some matching conditions, the newer method seemed more direct. It involves the asymptotic representation of a single quantity: the classical Green's function.

The Green's function $G(x, y) = G^\epsilon(x, y)$ provides an integral representation of the solution to the Poisson equation (1.7):

$$v(x) = \int_D G(x, y) \phi(y) dy. \quad (1.29)$$

The construction of G is a classical approach to the study of elliptic PDEs; see [27]. G is characterized as a solution of

$$\mathcal{L}^\epsilon_x G(x, y) = 0, \quad \mathcal{L}^{\epsilon*}_y G(x, y) = 0 \text{ for } x \neq y; \quad G(x, y) = 0 \text{ for } y \in \partial D,$$

with a singularity of appropriate type at $x = y$. (The subscripts on \mathcal{L}^ϵ_x and $\mathcal{L}^{\epsilon*}_y$ indicate which of the two variables of $G(x, y)$ the operators act on.) A classical formula extends (1.29) to solutions of the Poisson equation (1.7) with nonzero Dirichlet boundary data, $v(y) = f(y)$, $y \in \partial D$:

$$v(x) = \int_D G(x, y) \phi(y) dy - \int_{\partial D} [\mathcal{N}^{\epsilon*}_y G(x, y)] f(y) dy.$$

This is [27] equation (10.4) (after conversion to our notation). The function

$$K(x, y) = -\mathcal{N}^{\epsilon*}_y G(x, y) \quad (1.30)$$

is sometimes called the ‘‘Poisson kernel’’ (at least in the case of $\mathcal{L}^\epsilon = \frac{1}{2}\Delta$). It yields an integral representation of the solution to the Dirichlet problem (1.6)

$$u_f^\epsilon(x) = \int_{\partial D} K(x, y) f(y) dy,$$

and is therefore the density of the exit point distribution with respect to surface measure on ∂D :

$$\mu_x^\epsilon(dy) = K(x, y) dy.$$

In other words (1.30) is the density of $x^\epsilon(\tau_D^\epsilon)$ for $x^\epsilon(0) = x$.

The notation of [28] uses $p(x, y)$ in place of $G(x, y)$ and expresses (1.30) as

$$-\mathcal{N}^{\epsilon*}_y G(x, y) = J^G(x, y) \cdot n(y), \quad (1.31)$$

where $J^G = (J_i)$ is the *probability current density*:

$$J_i(x, y) = b_i(y)G(x, y) - \frac{\epsilon}{2} \sum_j \frac{\partial}{\partial x_j} [a_{ij}(y)G(x, y)], \quad (1.32)$$

So the point is that if we can produce the Green’s function $G(x, y)$, or a suitable asymptotic representation of it as $\epsilon \downarrow 0$, then the following formulas give the asymptotic behavior of the mean exit time and exit point distribution:

$$E_x[\tau_D^\epsilon] = \int_D G(x, y) dy \quad (1.33)$$

$$E_x[f(x^\epsilon(\tau_D^\epsilon))] = \int_{\partial D} J^G(x, y) \cdot n(y) f(y) dy \quad (1.34)$$

These are equations (2.32) and (2.41) of [28] respectively. They were developed in [28] with the case of singular diffusion in mind, but for nonsingular diffusion they follow from the classical theory of PDEs.

Formal asymptotic calculations proceed by supposing an asymptotic approximation to the Green’s function of the form

$$G(x, y) = p(y)q(x, y) \quad (1.35)$$

The factor $p(y)$ (called a *quasi-stationary density*) is taken to be a solution of $\mathcal{L}^{\epsilon*}p = 0$ of the form

$$p(y) = K^\epsilon(y)e^{-\psi(y)/\epsilon},$$

with appropriate normalization. The factor $q(x, y)$ is a boundary layer function with $q(x, y) = 0$ for $y \in \partial D$ but should be essentially constant for y in the interior of D (away from an asymptotically thin boundary layer).

In general (1.35) is only formal, unable to capture all the detailed mathematical structure of the true Green's function $G(x, y)$. For instance, the boundary layer function q will not capture the essential singularity of G at $x = y$. In fact as the asymptotic calculations in [28] proceed, no x -dependence of q is included. However expressions of the form (1.35) can be worked out explicitly in some special cases. Based on these, there is reasonable confidence that the asymptotic calculations stemming from (1.35) will lead to appropriate conclusions. However the conclusions of such an approach are limited by the presumptions and "ansatzes" which feed the calculations. They failed in particular to predict the cycling phenomena of exit to limit cycles. Once that type of possible behavior is recognized and built into the presumed asymptotic forms, formal methods become effective in studying its detailed structure; see [21].

We have explained (1.33) and (1.34) apart from any probabilistic interpretation. The development in [28] discusses them in terms of a process $x^+(t)$, which obeys the same stochastic dynamics (1.1) as $x^\epsilon(t)$ in D , but upon contact with ∂D instantaneously jumps to a specified $x_0 \in D$ and then resumes evolution in accord with (1.1). It is intuitively reasonable to expect that $x^+(t)$ has a unique stationary distribution π^+ in D . We merely wish to point out the relation of π^+ to G using Hasminskii's formula. The evolution of $x^+(t)$ can be decomposed into the sequence of independent cycles between successive jumps from ∂D to x_0 by using $\Gamma = \partial D$ and $\Gamma' = \{x_0\}$ in Section 2.4. The mean cycle length is $E_{x_0}[\tau_D^\epsilon]$. Hasminskii's formula identifies the stationary distribution according to

$$\int f(y) \pi^+(dy) = \frac{1}{E_{x_0}[\tau_D^\epsilon]} E_{x_0} \left[\int_0^{\tau_D^\epsilon} f(x^+(s)) ds \right].$$

Now we know that

$$E_{x_0} \left[\int_0^{\tau_D^\epsilon} f(x^+(s)) ds \right] = E_{x_0} \left[\int_0^{\tau_D^\epsilon} f(x^\epsilon(s)) ds \right] = \int_D G(x_0, y) f(y) dy.$$

Thus we see that

$$\pi^+(dy) = \frac{1}{E_{x_0}[\tau_D^\epsilon]} G(x_0, y) dy,$$

identifying the Green's function $G(x_0, y)$ as the (unnormalized) stationary density for $x^+(t)$ in D . This interpretation is described in [28].

3.4 Maier and Stein Principal Eigenfunction Method

R. Maier and D. Stein have written a number of papers exploring particular phenomena of the exit problem by means of formal asymptotic calculations.

Their real contribution is not so much the mathematical foundation of their method (described below) but how they have effectively used it to study some of the interesting sub-large deviations phenomena in the exit problem, particularly in cases for which the quasipotential function of large deviations analysis is nonsmooth. For instance in [22] and [20] they look at the situation in which the exit location concentrates at a saddle point on ∂D . They point out that singularities in the quasipotential are quite possible precisely at this saddle point, and that they effect the asymptotic description of the mean exit time. Their calculations offer a more detailed description of the ‘skewing’ phenomena of the exit location distribution than in other treatments, such as [3].

The calculations of Maier and Stein are based on an approach closely related to the Green’s function approach of Section 3.3. The central object is the principal adjoint eigenfunction ψ^ϵ for $\mathcal{L}^{\epsilon*}$ with Dirichlet conditions on ∂D :

$$\mathcal{L}^{\epsilon*}\psi^\epsilon + \lambda^\epsilon\psi^\epsilon = 0 \text{ in } D; \quad \psi^\epsilon(y) = 0 \text{ on } \partial D. \quad (1.36)$$

Here λ^ϵ is the principal eigenvalue mentioned in Section 2. General results imply that $\psi^\epsilon(x) > 0$ everywhere in D . We are free to specify a normalization of ψ^ϵ . The natural choice is

$$1 = \int_D \psi^\epsilon(x) dx, \quad (1.37)$$

allowing us to interpret $\psi^\epsilon(x)$ as a probability density in D , called the *quasi-stationary density* in [20]. They compute the probability current associated with ψ^ϵ :

$$J^\psi(y) \cdot n(y) = -\mathcal{N}^{\epsilon*}\psi^\epsilon(y)$$

and use this as the (unnormalized) density of exit points. The ‘reaction rate’ $\lambda^\epsilon \sim 1/E_x[\tau_D^\epsilon]$ describes the mean exit time and is given the formula

$$\lambda^\epsilon = \int_{\partial D} J^\psi(y) \cdot n(y) dy / \int_D \psi^\epsilon(x) dx. \quad (1.38)$$

We will see that (1.38) follows directly from Green’s formula. Consider any smooth u in \bar{D} . Since $\psi^\epsilon(y) = 0$ on ∂D , Green’s formula implies,

$$\lambda^\epsilon \int_D \psi^\epsilon u = - \int_D (\mathcal{L}^{\epsilon*}\psi^\epsilon)u = \int_D \psi^\epsilon \mathcal{L}^\epsilon u + \int_{\partial D} (-\mathcal{N}^{\epsilon*}\psi^\epsilon)u.$$

Taking $u \equiv 1$ and our normalization (1.37) yields (1.38) above. Using the solution $u = u_f^\epsilon$ of the Dirichlet problem we find

$$\int_D \psi^\epsilon u_f^\epsilon = \frac{1}{\lambda^\epsilon} \int_{\partial D} (-\mathcal{N}^{\epsilon*}\psi^\epsilon)f. \quad (1.39)$$

Equation (1.39) shows us that $-\frac{1}{\lambda^\epsilon} \mathcal{N}^{\epsilon*} \psi^\epsilon = \frac{1}{\lambda^\epsilon} J^\psi \cdot n$ is the density of the exit position of $x^\epsilon(\tau_D^\epsilon)$ if $x^\epsilon(0)$ is distributed within D according to density $\psi^\epsilon(x)$:

$$\int_D \psi^\epsilon u_f^\epsilon = \int_D E_x[f(x^\epsilon(\tau_D^\epsilon))] \psi^\epsilon(x) dx.$$

Recognizing $\psi^\epsilon(x)$ as the stationary density of the process $x^*(t)$ below, it seems clear that $\psi^\epsilon(x) dx \Rightarrow \delta_0(dx)$. More detailed large deviations results on ψ^ϵ should be possible. The leveling properties of u_f^ϵ show that $\int_D \psi^\epsilon u_f^\epsilon \rightarrow u_f^\epsilon$ as $\epsilon \downarrow 0$. Thus the accuracy of the approximation

$$E_0[f(x^\epsilon(\tau_D^\epsilon))] \approx \frac{1}{\lambda^\epsilon} \int_D J^\psi(y) \cdot n(y) f(y) dy$$

can be studied directly in terms of leveling and large deviations results. The probability current densities associated with $G(x, y)$ and $\psi^\epsilon(y)$ respectively should produce the same asymptotic results as $\epsilon \downarrow 0$. However, they are not identical for $\epsilon > 0$; they produce slightly different exit point distributions:

$$\begin{aligned} E_x[f(x^\epsilon(\tau_D^\epsilon))] &= \int_{\partial D} J^G(x, y) \cdot n(y) f(y) dy \\ \int_D E_x[f(x^\epsilon(\tau_D^\epsilon))] \psi^\epsilon(x) dy &= \frac{1}{\lambda^\epsilon} \int_{\partial D} J^\psi(y) \cdot n(y) f(y) dy. \end{aligned}$$

The second of these is an averaged version of the first. As the formal asymptotic calculations in [20] proceed, an asymptotic expression for ψ^ϵ is developed by setting λ^ϵ to 0 in (1.36) and using $\mathcal{L}^{\epsilon*} \psi^\epsilon = 0$ as an approximate equation. This is the same equation used for $p(y)$ of (1.35). Thus, as the asymptotic calculations are actually carried out, any distinction between the Green's function interpretation and the principal eigenfunction interpretation is lost.

We observe that the formulas arising from this eigenfunction approach can once again be interpreted in terms of Hasminskii's formula for an appropriately defined process. Let $x^*(t)$ be the stochastic process which follows (1.1) between times of contact with ∂D , but upon contact with ∂D is instantaneously restarted at a random point in D distributed with density ψ^ϵ . The successive times of contact with ∂D form the sequence of cycle times. The expected cycle time is given by

$$E[\tau_1] = \int_D E_x[\tau_D^\epsilon] \psi^\epsilon(x) dx = \int_D \psi^\epsilon v_D^\epsilon dx.$$

Notice that (using Green's formula)

$$\lambda^\epsilon \int_D \psi^\epsilon v_D^\epsilon = - \int_D (\mathcal{L}^{\epsilon*} \psi^\epsilon) v_D^\epsilon = - \int_D \psi^\epsilon \mathcal{L}^\epsilon v_D^\epsilon = \int_D \psi^\epsilon = 1.$$

Thus the mean cycle time for $x^*(t)$ is given exactly by

$$E[\tau_1] = 1/\lambda^\epsilon.$$

The solution $v(x)$ of the Poisson equation (1.7) provides the mean path integral over one cycle: $E[\int_0^{\tau_D^\epsilon} \phi(x^\epsilon(s)) ds] = \int_D \psi^\epsilon v$. Hasminskii's formula would identify the stationary distribution π^* for $x^*(t)$ as

$$\begin{aligned} \int \phi(x) \pi^*(dx) &= \frac{1}{E^*[\tau]} E^*[\int_0^\tau \phi(x^*(s)) ds] \\ &= \lambda^\epsilon \int_D \psi^\epsilon v \\ &= \int_D -(\mathcal{L}^{\epsilon*} \psi^\epsilon) v \\ &= \int_D -\psi^\epsilon \mathcal{L}^\epsilon v \\ &= \int_D \psi^\epsilon \phi. \end{aligned}$$

Thus ψ^ϵ is the stationary density of $x^*(t)$. We can view (1.38) and (1.39) as a method for determining the mean exit time and exit point distribution from the invariant density $\psi^\epsilon(x)$.

3.5 Exit Conditioning

We conclude with a brief mention of the method of exit conditioning: [4]. The approaches described so far relate strongly to Green's identity and quantities which are natural from a PDE approach to the problem. Exit conditioning is more exclusively probabilistic. It was developed expressly to treat cases of characteristic boundary and has been successful in producing qualitative results on sub-large deviations phenomena without using formal asymptotic expansions. The idea is to consider a subregion $G \subseteq D$ with $0 \in G$ for which the exit problem (τ_G^ϵ) is nontangential: $n \cdot b < 0$ on ∂G . Large deviations results are used to describe the exit point distribution from G , i.e. $\nu_x^\epsilon(dz) = P_x[x^\epsilon(\tau_G^\epsilon) \in dz]$. The exit point distribution $\mu_x^\epsilon(dy)$ from D is then represented in terms of ν_x^ϵ and a *conditional exit kernel* $Q^\epsilon(z, dy)$, $z \in \partial G$:

$$\mu^\epsilon(dy) = \int_{\partial G} c^\epsilon g(z) Q^\epsilon(z, dy) \nu^\epsilon(dz) + o(1). \quad (1.40)$$

Here g is a specific positive continuous function on ∂D , and the $o(1)$ accounts for dependence on the initial point, which we know vanishes as $\epsilon \downarrow 0$. To define Q^ϵ we use another subdomain C : $0 \in C \subseteq G \subseteq D$. One might think of $D \setminus C$ as a boundary strip. Define τ_C^ϵ to be the first time of contact with ∂C for $x^\epsilon(t)$, started at $x^\epsilon(0) = z \in \partial G$. Then

$$Q^\epsilon(z, A) = P_z[x^\epsilon(\tau_D^\epsilon) \in A | \tau_D^\epsilon < \tau_C^\epsilon]. \quad (1.41)$$

What makes (1.40) effective is the possibility of describing the asymptotic behavior of Q^ϵ in terms of qualitative features of (1.2). The details are more involved than we can summarize here. See [3] and [2] for a fuller description, and some results of this approach.

Acknowledgments: This article is an expansion and update of the author's presentation at the Workshop on Asymptotic Methods for Stochastic Sys-

tems - Theory and Practice, Univ. Maryland and U. S. Army, Oct. 25–27, 1987. Sections 3.1 and 3.2 in particular were originally presented there. We wish to thank Wendell Fleming for his suggestion in 1979 that we consider these problems, and his continued support and interest in our work on them.

References

- [1] R. F. Anderson and S. Orey (1976). *Small random perturbations of dynamical systems with reflecting boundary*, Nagoya Mathematical Journal , **60**, 189–216.
- [2] M. V. Day (1996). *Exit cycling for the van der Pol oscillator and quasi-potential calculations*, Journal of Dynamics and Differential Equations, **8**, 573–601.
- [3] M. V. Day (1994). *Cycling and skewing of exit measures for planar systems*, Stochastics and Stochastic Reports **48**, 227–247.
- [4] M. V. Day (1992). *Conditional exits for small noise diffusions*, The Annals of Probability **20**, 1385–1419.
- [5] M. V. Day (1989). *Boundary local time and small parameter exit problems with characteristic boundaries*, SIAM Journal on Mathematical Analysis **20**, 222–248.
- [6] M. V. Day (1987). *Recent progress on the small parameter exit problem*, Stochastics **20**, 121–150.
- [7] M. V. Day (1983). *On the exponential exit law in the small parameter exit problem*, Stochastics **8**, 297–323.
- [8] M. V. Day (1982). *Exponential leveling for stochastically perturbed dynamical systems*, SIAM Journal on Mathematical Analysis **13**, 532–540.
- [9] Paul Dupuis and Richard S. Ellis (1997). *A WEAK CONVERGENCE APPROACH TO THE THEORY OF LARGE DEVIATIONS*, J. Wiley, New York.
- [10] P. Dupuis and H. J. Kushner (1987). *Stochastic systems with small noise, analysis and simulation: a phase locked loop example*, SIAM Journal on Applied Mathematics **47**, 643–660.
- [11] M. I. Dykman, P. V. E. McClintock, V. N. Smelyanski, N. D. Stein, and N. G. Stocks (1992). *Optimal paths and the prehistory problem for large fluctuations in noise driven systems*, Physical Review Letters **68**, 2718–2721.

- [12] A. Eizenberg (1990). *The exponential leveling in elliptic singular perturbation problems with complicated attractors*, Journal d'Analyse Mathématique **55**, 229–249.
- [13] L. C. Evans and H. Ishii (1985). *A PDE approach to some asymptotic problems concerning random differential equations with small noise intensities*, Annales de l'Institut Henri Poincaré. Analyse Non Linéaire **2**, 1–20.
- [14] Wendell H. Fleming (1977/78). *Exit probabilities and optimal stochastic control*, Applied Mathematics and Optimization **4**, 329–346.
- [15] Wendell H. Fleming (1992). *Stochastic control and large deviations*, Future Tendencies in Computer Science, Control and Applied Mathematics, 291–300, Lecture Notes in Computer Science **653** Springer, Berlin.
- [16] Wendell H. Fleming and Shuenn-Jyi Sheu, *Asymptotics for the principal eigenvalue and eigenfunction of a nearly first order operator with large potential*, The Annals of Probability (to appear).
- [17] M. I. Freidlin and A. D. Wentzell (1984). RANDOM PERTURBATIONS OF DYNAMICAL SYSTEMS, Springer-Verlag, Berlin, New York.
- [18] R. Z. Hasminskii (1980). STOCHASTIC STABILITY OF DIFFERENTIAL EQUATIONS, Sijthoff and Noordhoff, Alphen aan den Rijn, the Netherlands.
- [19] S. Kamin (1978). *Elliptic perturbation of a first order operator with a singular point of attracting type*, Indiana University Mathematics Journal **27**, 935–952.
- [20] R. S. Maier and D. L. Stein (1997). *Limiting exit location distribution in the stochastic exit problem*, SIAM Journal on Applied Mathematics **57**, 752–790.
- [21] R. S. Maier and D. L. Stein (1996). *Oscillatory behavior of the rate of escape through an unstable limit cycle*, Physical Review Letters **77**, 4860–4863.
- [22] R. S. Maier and D. L. Stein (1996). *A scaling theory of bifurcations in the symmetric weak-noise escape problem*, Journal of Statistical Physics **83**, 291–357.
- [23] F. Martinelli, and E. Scoppola (1988). *Small random perturbations of dynamical systems: exponential loss of memory of the initial condition*, Communications in Mathematical Physics **120**, 25–69.

- [24] P. Mathieu (1995). *Spectra, exit times and long time asymptotics in the zero-white-noise limit*, Stochastics and Stochastics Reports **55**, 1–20.
- [25] B. J. Matkowsky and Z. Schuss (1977). *The exit problem for randomly perturbed dynamical systems*, SIAM Journal on Applied Mathematics **33**, 365–382.
- [26] B. J. Matkowsky, Z. Schuss and C. Tier (1983). *Diffusion across characteristic boundaries with critical points*, SIAM Journal on Applied Mathematics **43**, 673–695.
- [27] C. Miranda (1970). PARTIAL DIFFERENTIAL EQUATIONS OF ELLIPTIC TYPE, Springer-Verlag, Heidelberg.
- [28] T. Naeh, M. M. Kłosek, B. J. Matkowsky and Z. Schuss (1990). *A direct approach to the exit problem*, SIAM Journal on Applied Mathematics **50** 595–627.
- [29] B. Perthame (1990). *Perturbed dynamical systems with an attracting singularity and weak viscosity limits in Hamilton-Jacobi equations*, Transactions of the American Mathematical Society **317** 723–748.
- [30] M. Sigura (1995). *Metastable behaviors of diffusion processes with small parameter*, Journal of the Mathematical Society of Japan **47** 755–788.