ABSTRACT

Title of dissertation: ASYMPTOTIC PROBLEMS IN STOCHASTIC PROCESSES AND DIFFERENTIAL EQUATIONS

Wenqing Hu, Doctor of Philosophy, 2013

Dissertation directed by: Professor Mark Freidlin
Department of Mathematics

We study some asymptotic problems in stochastic processes and in differential equations. We consider Smoluchowski-Kramers approximations with variable and vanishing friction. We also consider the problem around second order elliptic equations with a small parameter.

ASYMPTOTIC PROBLEMS IN STOCHASTIC PROCESSES AND DIFFERENTIAL EQUATIONS

by

Wenqing Hu

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2013

Advisory Committee:
Professor Mark Freidlin, Chair/Advisor
Professor Sandra Cerrai
Professor Dmitry Dolgopyat
Professor Leonid Koralov
Professor Issak Mayergoyz

Foreword

In Chapter 1 we consider the small mass asymptotics (Smoluchowski-Kramers approximation) for the Langevin equation with a variable friction coefficient. The limit of the solution in the classical sense does not exist in this case. We study a modification of the Smoluchowski-Kramers approximation. Some applications of the Smoluchowski-Kramers approximation to problems with fast oscillating or discontinuous coefficients are considered. This is joint work with Mark Friedlin.

In Chapter 2 we consider the small mass asymptotic (Smoluchowski-Kramers approximation) for the Langevin equation with a variable friction coefficient. The friction coefficient is assumed to be vanishing within certain region. We introduce a regularization for this problem and study the limiting motion for the 1-dimensional case and a multidimensional model problem. The limiting motion is a Markov process on a projected space. We specify the generator and boundary condition of this limiting Markov process and prove the convergence. This is joint work with Mark Freidlin and Alexander Wentzell.

In Chapter 3 we consider the Neumann problem with a small parameter

$$\left(\frac{1}{\varepsilon}L_0 + L_1\right) u^{\varepsilon}(x) = f(x) \text{ for } x \in G, \ \left.\frac{\partial u^{\varepsilon}(x)}{\partial \gamma^{\varepsilon}(x)}\right|_{\partial G} = 0 \ .$$

The operators L_0 and L_1 are self-adjoint second order operators. We assume that L_0 has a non-negative characteristic form and L_1 is strictly elliptic. The reflection is with respect to inward co-normal unit vector $\gamma^{\varepsilon}(x)$. The behavior of $\lim\limits_{\varepsilon \downarrow 0} u^{\varepsilon}(x)$ is effectively described via the solution of an ordinary differential equation on a tree. We calculate the differential operators inside the edges of this tree and the gluing condition at the root. Our approach is based on an analysis of the corresponding diffusion processes. This is joint work with Mark Freidlin.

In conclusion we will explain how formulas and theorems are numbered. For example, Theorem 3.2.1 is the first theorem in the second section in Chapter 3. Inside Chapter 3, it is written as Theorem 2.1 only. Formulas and figures are numbered in a similar fashion.

Dedicated to my parents.

Acknowledgments

I thank my advisor Professor Mark Freidlin for helping me so much and for teaching me so much in my graduate studies over the years.

I thank members of my committee and many other professors at University of Maryland, College Park and at other universities for their help in my graduate studies over the years.

I thank my parents and all my friends for helping me so much over the many difficult or even desperate years of life.

TABLE OF CONTENTS

LIST OF FIGURES

1. SMOLUCHOWSKI-KRAMERS APPROXIMATION IN THE CASE OF VARIABLE FRICTION.

1.1 Introduction

The Langevin equation

$$\mu\ddot{\boldsymbol{q}}_t^\mu = \boldsymbol{b}(\boldsymbol{q}_t^\mu) - \lambda\dot{\boldsymbol{q}}_t^\mu + \sigma(\boldsymbol{q}_t^\mu)\dot{\boldsymbol{W}}_t \ , \ \boldsymbol{q}_0^\mu = \boldsymbol{q} \in \mathbb{R}^n \ , \dot{\boldsymbol{q}}_0^\mu = \boldsymbol{p} \in \mathbb{R}^n \ , \tag{1.1}$$

describes the motion of a particle of mass μ in a force field $\boldsymbol{b}(\boldsymbol{q})$, $\boldsymbol{q} \in \mathbb{R}^n$, subjected to random fluctuations and to a friction proportional to the velocity. Here $\boldsymbol{W}_t$ is the standard Wiener process in $\mathbb{R}^n$, $\lambda > 0$ is the friction coefficient. The vector field $\boldsymbol{b}(\boldsymbol{q})$ and the matrix function $\sigma(\boldsymbol{q})$ are assumed to be continuously differentiable and bounded together with their first derivatives. The matrix $a(\boldsymbol{q}) = (a_{ij}(\boldsymbol{q})) = \sigma(\boldsymbol{q})\sigma^*(\boldsymbol{q})$ is assumed to be non-degenerate.

Put $\boldsymbol{p}_t^\mu = \dot{\boldsymbol{q}}_t^\mu$. Then (1.1) can be written as a first order system:

$$\begin{cases} \dot{\boldsymbol{q}}_t^\mu = \boldsymbol{p}_t^\mu \\ \dot{\boldsymbol{p}}_t^\mu = \dfrac{1}{\mu}\boldsymbol{b}(\boldsymbol{q}_t^\mu) - \dfrac{\lambda}{\mu}\boldsymbol{p}_t^\mu + \dfrac{1}{\mu}\sigma(\boldsymbol{q}_t^\mu)\dot{\boldsymbol{W}}_t \end{cases} \tag{1.2}$$

The diffusion process $(\boldsymbol{p}_t^\mu, \boldsymbol{q}_t^\mu) = \boldsymbol{X}_t^\mu$ in $\mathbb{R}^{2n}$ is governed by the generator L:

$$Lu(\boldsymbol{p},\boldsymbol{q}) = \frac{1}{2\mu^2}\sum_{i,j=1}^n a_{ij}(\boldsymbol{q})\frac{\partial^2 u}{\partial p_i \partial p_j} + \frac{1}{\mu}(\boldsymbol{b}(\boldsymbol{q}) - \lambda\boldsymbol{p})\cdot\nabla_{\boldsymbol{p}}u + \boldsymbol{p}\cdot\nabla_{\boldsymbol{q}}u \ .$$

Note that, since functions $\boldsymbol{q}_t^\mu$ are continuously differentiable with probability one,

$$\int_0^t \sigma_{ij}(\boldsymbol{q}_s^\mu)dW_s^j = \sigma_{ij}(\boldsymbol{q}_t^\mu)W_t^j - \int_0^t W_s^j\,(\nabla_{\boldsymbol{q}}\sigma_{ij}(\boldsymbol{q}_s^\mu)\cdot\boldsymbol{p}_s^\mu)\,ds \ .$$

This allows to consider equations (1.2) for each trajectory $\boldsymbol{W}_t$ individually, and there is no necessity in the introduction of a stochastic integral. In particular, if (1.2) is considered as a stochastic differential equation, stochastic integrals in the Itô and in the Stratonovich sense coincide: $\int_0^t \sigma(\boldsymbol{q}_s^\mu)d\boldsymbol{W}_s = \int_0^t \sigma(\boldsymbol{q}_s^\mu)\circ d\boldsymbol{W}_s$.

It is assumed usually that the friction coefficient λ is constant. Under this assumption, one can prove that $\boldsymbol{q}_t^\mu$ converges in probability as $\mu \downarrow 0$ uniformly on each finite time interval $[0,T]$ to an n-dimensional diffusion process $\boldsymbol{q}_t$: for any $\kappa, T > 0$ and any $\boldsymbol{p}_0^\mu = \boldsymbol{p} \in \mathbb{R}^n$ fixed,

$$\lim_{\mu\downarrow 0}\mathbf{P}\left(\max_{0\le t\le T}|\boldsymbol{q}_t^\mu - \boldsymbol{q}_t|_{\mathbb{R}^d} > \kappa\right) = 0 \ .$$

Here $\boldsymbol{q}_t$ is the solution of equation

$$\dot{\boldsymbol{q}}_t = \frac{1}{\lambda}\boldsymbol{b}(\boldsymbol{q}_t) + \frac{1}{\lambda}\sigma(\boldsymbol{q}_t)\dot{\boldsymbol{W}}_t \ , \ \boldsymbol{q}_0 = \boldsymbol{q}_0^\mu = \boldsymbol{q} \in \mathbb{R}^n \ . \tag{1.3}$$

The stochastic term in (1.3) should be understood in the Itô sense.

The approximation of $\boldsymbol{q}_t^\mu$ by $\boldsymbol{q}_t$ for $0 < \mu << 1$ is called the Smoluchowski-Kramers approximation. This is the main justification for replacement of the second order equation (1.1) by the first order equation (1.3). The price for such a simplification, in particular, consists of certain non-universality of equation (1.3): The white noise in (1.1) is an idealization of a more regular stochastic process $\dot{\boldsymbol{W}}_t^\delta$ with correlation radius $\delta << 1$ converging to $\dot{\boldsymbol{W}}_t$ as $\delta \downarrow 0$. Let $\boldsymbol{q}_t^{\mu,\delta}$ be the solution of equation (1.1) with $\dot{\boldsymbol{W}}_t$ replaced by $\dot{\boldsymbol{W}}_t^\delta$. Then limit of $\boldsymbol{q}_t^{\mu,\delta}$ as $\mu, \delta \downarrow 0$ depends on the relation between μ and δ. Say, if first $\delta \downarrow 0$ and then $\mu \downarrow 0$, the stochastic integral in (1.3) should be understood in the Itô sense; if first $\mu \downarrow 0$ and then $\delta \downarrow 0$, $\boldsymbol{q}_t^{\mu,\delta}$ converges to the solution of (1.3) with stochastic integral in the Stratonovich sense. (See, for instance, [8].)

Consider now the case of a variable friction coefficient $\lambda = \lambda(\boldsymbol{q})$. We assume that $\lambda(\boldsymbol{q})$ has continuous bounded derivatives and $0 < \lambda_0 \leq \lambda(\boldsymbol{q}) \leq \Lambda < \infty$. It turns out, as we will see in the next section, that in this case the solution $\boldsymbol{q}_t^\mu$ of (1.1) does not converge, in general, to the solution of (1.3) with $\lambda = \lambda(\boldsymbol{q})$, so that the Smoluchowski-Kramers approximation should be modified. In order to do this, we consider equation (1.1) with $\dot{\boldsymbol{W}}_t$ replaced by $\dot{\boldsymbol{W}}_t^\delta$ described above:

$$\mu \ddot{\boldsymbol{q}}_t^{\mu,\delta} = \boldsymbol{b}(\boldsymbol{q}_t^{\mu,\delta}) - \lambda(\boldsymbol{q}_t^{\mu,\delta})\dot{\boldsymbol{q}}_t^{\mu,\delta} + \sigma(\boldsymbol{q}_t^{\mu,\delta})\dot{\boldsymbol{W}}_t^\delta \ , \ \boldsymbol{q}_0^{\mu,\delta} = \boldsymbol{q} \ , \ \dot{\boldsymbol{q}}_0^{\mu,\delta} = \boldsymbol{p} \ . \tag{1.4}$$

We prove that after such a regularization, the solution of (1.4) has a limit $\widetilde{\boldsymbol{q}}_t^\delta$ as $\mu \downarrow 0$, and $\widetilde{\boldsymbol{q}}_t^\delta$ is the unique solution of the equation obtained from (1.4) as $\mu = 0$:

$$\dot{\widetilde{\boldsymbol{q}}}_t^\delta = \frac{1}{\lambda(\widetilde{\boldsymbol{q}}_t^\delta)}\boldsymbol{b}(\widetilde{\boldsymbol{q}}_t^\delta) + \frac{1}{\lambda(\widetilde{\boldsymbol{q}}_t^\delta)}\sigma(\widetilde{\boldsymbol{q}}_t^\delta)\dot{\boldsymbol{W}}_t^\delta \ , \ \widetilde{\boldsymbol{q}}_0^\delta = \boldsymbol{q} \ . \tag{1.5}$$

Now we can take $\delta \downarrow 0$ in (1.5). As the result we get the equation

$$\dot{\widehat{\boldsymbol{q}}}_t = \frac{1}{\lambda(\widehat{\boldsymbol{q}}_t)}\boldsymbol{b}(\widehat{\boldsymbol{q}}_t) + \frac{1}{\lambda(\widehat{\boldsymbol{q}}_t)}\sigma(\widehat{\boldsymbol{q}}_t) \circ \dot{\boldsymbol{W}}_t \ , \ \widehat{\boldsymbol{q}}_0 = \boldsymbol{q} \ , \tag{1.6}$$

where the stochastic term should be understood in the Stratonovich sense. So the regularization leads to a modified Smoluchowski-Kramers equation (1.6). We prove this in Section 3.

Some applications of the Smoluchowski-Kramers approximation are considered in Sections 4 and 5: the case of fast oscillating in the space variable, periodic or stochastic, friction coefficient is studied; gluing condition at the discontinuity points of the friction coefficient are considered.

Notations. We use $|\bullet|_{\mathbb{R}^d}$ to denote the standard Euclidean norm in $\mathbb{R}^d$. When $d = 1$ we set $|\bullet|_{\mathbb{R}^1} = |\bullet|$. For a vector-valued function $\boldsymbol{f}(\boldsymbol{x}) = (f_1(\boldsymbol{x}), ..., f_d(\boldsymbol{x}))$, $\boldsymbol{x} \in \mathbb{R}^d$, we set $\|\boldsymbol{f}\|_\infty = \max\limits_{1 \leq i \leq d} \|f_i\|_\infty = \max\limits_{1 \leq i \leq d} \sup\limits_{\boldsymbol{x} \in \mathbb{R}^d} |f_i(\boldsymbol{x})|$. All the vectors are marked with either bold letters or with an arrow on it.

1.2 Some estimates. The classical Smoluchowski-Kramers approximation does not work for variable friction coefficients

We consider the following system

$$\mu\ddot{\boldsymbol{q}}_t^\mu = \boldsymbol{b}(\boldsymbol{q}_t^\mu) - \lambda(\boldsymbol{q}_t^\mu)\dot{\boldsymbol{q}}_t^\mu + \dot{\boldsymbol{W}}_t \ , \ \boldsymbol{q}_0^\mu = \boldsymbol{q} \in \mathbb{R}^d \ , \ \dot{\boldsymbol{q}}_0^\mu = \boldsymbol{p} \in \mathbb{R}^d. \tag{2.1}$$

Here $\infty > \Lambda \geq \lambda(\bullet) \geq \lambda_0 > 0$ is a function of $\boldsymbol{q}_t^\mu$. We assume that function $\lambda(\bullet)$ and the vector field $\boldsymbol{b}(\bullet)$ are continuously differentiable and bounded together with their first derivatives. The process $\boldsymbol{W}_t$ is the standard Wiener process in $\mathbb{R}^d$. For simplicity of calculations we consider here the case when the diffusion matrix $a(\bullet)$ is the identity (compare with (1.1)).

Let $\boldsymbol{p}_t^\mu = \dot{\boldsymbol{q}}_t^\mu$, we have, that (2.1) is equivalent to the system

$$\left\{\begin{array}{l} \dot{\boldsymbol{q}}_t^\mu = \boldsymbol{p}_t^\mu \ , \\ \dot{\boldsymbol{p}}_t^\mu = \dfrac{1}{\mu}\boldsymbol{b}(\boldsymbol{q}_t^\mu) - \dfrac{\lambda(\boldsymbol{q}_t^\mu)}{\mu}\boldsymbol{p}_t^\mu + \dfrac{1}{\mu}\dot{\boldsymbol{W}}_t \ . \end{array}\right. \tag{2.2}$$

Then

$$\frac{d}{dt}\left(e^{\frac{1}{\mu}\int_0^t \lambda(\boldsymbol{q}_s^\mu)ds}\boldsymbol{p}_t^\mu\right) = e^{\frac{1}{\mu}\int_0^t \lambda(\boldsymbol{q}_s^\mu)ds}\left(\dot{\boldsymbol{p}}_t^\mu + \frac{1}{\mu}\lambda(\boldsymbol{q}_t^\mu)\boldsymbol{p}_t^\mu\right) = e^{\frac{1}{\mu}\int_0^t \lambda(\boldsymbol{q}_s^\mu)ds}\left(\frac{1}{\mu}\boldsymbol{b}(\boldsymbol{q}_t^\mu) + \frac{1}{\mu}\dot{\boldsymbol{W}}_t\right) ,$$

and

$$e^{\frac{1}{\mu}\int_0^t \lambda(\boldsymbol{q}_s^\mu)ds}\boldsymbol{p}_t^\mu - \boldsymbol{p} = \frac{1}{\mu}\int_0^t e^{\frac{1}{\mu}\int_0^s \lambda(\boldsymbol{q}_r^\mu)dr}\boldsymbol{b}(\boldsymbol{q}_s^\mu)ds + \frac{1}{\mu}\int_0^t e^{\frac{1}{\mu}\int_0^s \lambda(\boldsymbol{q}_r^\mu)dr}d\boldsymbol{W}_s \ . \tag{2.3}$$

For notational convenience we introduce the function $A(\mu,t) = \int_0^t \lambda(\boldsymbol{q}_s^\mu)ds$. It is clear that $t\Lambda \geq A(\mu,t) \geq t\lambda_0$. Using (2.3) we have, that

$$\boldsymbol{p}_t^\mu = e^{-\frac{1}{\mu}A(\mu,t)}\left(\boldsymbol{p} + \frac{1}{\mu}\int_0^t e^{\frac{1}{\mu}A(\mu,s)}\boldsymbol{b}(\boldsymbol{q}_s^\mu)ds + \frac{1}{\mu}\int_0^t e^{\frac{1}{\mu}A(\mu,s)}d\boldsymbol{W}_s\right) \ .$$

Therefore we have

$$\begin{aligned} \boldsymbol{q}_t^\mu &= \boldsymbol{q} + \int_0^t \boldsymbol{p}_s^\mu ds \\ &= \boldsymbol{q} + \boldsymbol{p}\int_0^t e^{-\frac{1}{\mu}A(\mu,s)}ds + \frac{1}{\mu}\int_0^t e^{-\frac{1}{\mu}A(\mu,s)}\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\right)ds + \\ &\quad + \frac{1}{\mu}\int_0^t e^{-\frac{1}{\mu}A(\mu,s)}\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}d\boldsymbol{W}_r\right)ds \\ &= \boldsymbol{q} + \boldsymbol{\alpha}(\mu) + \boldsymbol{\beta}(\mu) + \boldsymbol{\gamma}(\mu) \ . \end{aligned} \tag{2.4}$$

Here $\boldsymbol{\alpha}(\mu), \boldsymbol{\beta}(\mu), \boldsymbol{\gamma}(\mu)$ are three (vector) functions in the right hand side of (2.4):

$$\boldsymbol{\alpha}(\mu) = \boldsymbol{p} \int_0^t e^{-\frac{1}{\mu}A(\mu,s)} ds \ ,$$

$$\boldsymbol{\beta}(\mu) = \frac{1}{\mu} \int_0^t e^{-\frac{1}{\mu}A(\mu,s)} \left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)} \boldsymbol{b}(\boldsymbol{q}_r^\mu) dr \right) ds \ ,$$

$$\boldsymbol{\gamma}(\mu) = \frac{1}{\mu} \int_0^t e^{-\frac{1}{\mu}A(\mu,s)} \left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)} d\boldsymbol{W}_r \right) ds \ .$$

In the following we will use the relation

$$\frac{d}{dt}\left(e^{-\frac{1}{\mu}A(\mu,t)}\right) = -\frac{1}{\mu} e^{-\frac{1}{\mu}A(\mu,t)} \frac{dA(\mu,t)}{dt} = -\frac{1}{\mu} e^{-\frac{1}{\mu}A(\mu,t)} \lambda(\boldsymbol{q}_t^\mu) \ . \tag{2.5}$$

We will also use the estimates

$$\frac{\mu}{c\Lambda}(1-e^{-\frac{c\Lambda t}{\mu}}) = \int_0^t e^{-\frac{c\Lambda s}{\mu}} ds \leq \int_0^t e^{-\frac{c}{\mu}A(\mu,s)} ds \leq \int_0^t e^{-\frac{c\lambda_0 s}{\mu}} ds = \frac{\mu}{c\lambda_0}(1-e^{-\frac{c\lambda_0 t}{\mu}}) \leq \frac{\mu}{c\lambda_0} \ , \tag{2.6}$$

$$\begin{aligned} \frac{\mu}{c\Lambda}(1-e^{-\frac{c\Lambda t}{\mu}}) &= \int_0^t e^{-\frac{c\Lambda(t-s)}{\mu}} ds \leq \int_0^t e^{-\frac{c}{\mu}(A(\mu,t)-A(\mu,s))} ds \leq \\ &\leq \int_0^t e^{-\frac{c\lambda_0(t-s)}{\mu}} ds = \frac{\mu}{c\lambda_0}(1-e^{-\frac{c\lambda_0 t}{\mu}}) \leq \frac{\mu}{c\lambda_0} \ . \end{aligned} \tag{2.7}$$

Here c is a positive constant.

We get in this section some bounds for $\boldsymbol{\alpha}(\mu)$, $\boldsymbol{\beta}(\mu)$, $\boldsymbol{\gamma}(\mu)$ which show, in particular, that the classical Smoluchowski-Kramers approximation does not hold in the case of variable friction. These bounds also will be used to obtain a modified Smoluchowski-Kramers approximation.

2.1. Estimates of $\boldsymbol{\alpha}(\mu)$.

The estimates of $\boldsymbol{\alpha}(\mu)$ is relatively simple since we have $\boldsymbol{\alpha}(\mu) = \boldsymbol{p} \int_0^t e^{-\frac{1}{\mu}A(\mu,s)} ds$ and $A(\mu,s) \geq \lambda s$. Therefore $|\boldsymbol{\alpha}(\mu)|_{\mathbb{R}^d} \to 0$ as $\mu \downarrow 0$.

2.2. Estimates of $\boldsymbol{\beta}(\mu)$.

We have, by (2.5), that

$$
\begin{aligned}
&\boldsymbol{\beta}(\mu)=\frac{1}{\mu}\int_0^t e^{-\frac{1}{\mu}A(\mu,s)}\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\right)ds\\
&=\frac{1}{\mu}\int_0^t\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\right)\left(-\frac{\mu}{\lambda(\boldsymbol{q}_s^\mu)}\right)d(e^{-\frac{1}{\mu}A(\mu,s)})\\
&=\int_0^t\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\right)\left(-\frac{1}{\lambda(\boldsymbol{q}_s^\mu)}\right)d(e^{-\frac{1}{\mu}A(\mu,s)})\\
&=-\frac{e^{-\frac{1}{\mu}A(\mu,s)}}{\lambda(\boldsymbol{q}_s^\mu)}\int_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\Bigg|_{s=0}^{s=t}+\int_0^t e^{-\frac{1}{\mu}A(\mu,s)}d\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\frac{1}{\lambda(\boldsymbol{q}_s^\mu)}\right)\\
&=-\frac{e^{-\frac{1}{\mu}A(\mu,t)}}{\lambda(\boldsymbol{q}_t^\mu)}\int_0^t e^{\frac{1}{\mu}A(\mu,s)}\boldsymbol{b}(\boldsymbol{q}_s^\mu)ds+\int_0^t\frac{\boldsymbol{b}(\boldsymbol{q}_s^\mu)}{\lambda(\boldsymbol{q}_s^\mu)}ds+\int_0^t e^{-\frac{1}{\mu}A(\mu,s)}\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\right)d\left(\frac{1}{\lambda(\boldsymbol{q}_s^\mu)}\right)\\
&=\boldsymbol{R}_{\boldsymbol{\beta}}(\mu)+\int_0^t\frac{\boldsymbol{b}(\boldsymbol{q}_s^\mu)}{\lambda(\boldsymbol{q}_s^\mu)}ds+(\vec{II})\ .
\end{aligned}
$$

It is easy to see that

$$|\boldsymbol{R}_{\boldsymbol{\beta}}(\mu)|_{\mathbb{R}^d}\leq\frac{\|\boldsymbol{b}\|_\infty}{\lambda_0}\int_0^t e^{-\frac{\lambda_0}{\mu}(t-s)}ds=\frac{\|\boldsymbol{b}\|_\infty}{\lambda_0}\frac{\mu}{\lambda_0}\ .$$

We also have

$$
\begin{aligned}
&(\vec{II})=-\int_0^t e^{-\frac{1}{\mu}A(\mu,s)}\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\right)\frac{1}{\lambda^2(\boldsymbol{q}_s^\mu)}\nabla\lambda(\boldsymbol{q}_s^\mu)\cdot\boldsymbol{p}_s^\mu ds\\
&=-\int_0^t e^{-\frac{1}{\mu}A(\mu,s)}\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\right)\frac{1}{\lambda^2(\boldsymbol{q}_s^\mu)}\nabla\lambda(\boldsymbol{q}_s^\mu)\cdot\\
&\qquad\cdot e^{-\frac{1}{\mu}A(\mu,s)}\left(\boldsymbol{p}+\frac{1}{\mu}\int_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr+\frac{1}{\mu}\int_0^s e^{\frac{1}{\mu}A(\mu,r)}d\boldsymbol{W}_r\right)ds\\
&=(\vec{II}_1)+(\vec{II}_2)+(\vec{II}_3)\ .
\end{aligned}
$$

Here

$$(\vec{II}_1)=-\int_0^t\frac{e^{-\frac{2}{\mu}A(\mu,s)}}{\lambda^2(\boldsymbol{q}_s^\mu)}\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\right)\nabla\lambda(\boldsymbol{q}_s^\mu)\cdot\boldsymbol{p}ds\ ,$$

$$(\vec{II}_2)=-\frac{1}{\mu}\int_0^t\frac{e^{-\frac{2}{\mu}A(\mu,s)}}{\lambda^2(\boldsymbol{q}_s^\mu)}\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\right)\nabla\lambda(\boldsymbol{q}_s^\mu)\cdot\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\right)ds\ ,$$

$$(\vec{II}_3)=-\frac{1}{\mu}\int_0^t\frac{e^{-\frac{2}{\mu}A(\mu,s)}}{\lambda^2(\boldsymbol{q}_s^\mu)}\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\right)\nabla\lambda(\boldsymbol{q}_s^\mu)\cdot\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}d\boldsymbol{W}_r\right)ds\ .$$

We conclude that

$$
\begin{aligned}
&|(\vec{II}_1)|_{\mathbb{R}^d}\leq\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}|\boldsymbol{p}|_{\mathbb{R}^d}\|\boldsymbol{b}\|_\infty\int_0^t e^{-\frac{\lambda_0 s}{\mu}}\left(\int_0^s e^{-\frac{(s-r)\lambda_0}{\mu}}dr\right)ds\\
&\leq\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}|\boldsymbol{p}|_{\mathbb{R}^d}\|\boldsymbol{b}\|_\infty\frac{\mu^2}{\lambda_0^2}\ ;
\end{aligned}
$$

$$
\begin{aligned}
|(\vec{II_2})|_{\mathbb{R}^d} &\leq \frac{1}{\mu}\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\|\boldsymbol{b}\|_\infty^2\int_0^t\left(\int_0^s e^{-\frac{(s-r)\lambda_0}{\mu}}dr\right)^2 ds \\
&\leq \frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\|\boldsymbol{b}\|_\infty^2\frac{\mu t}{\lambda_0^2}\ ;
\end{aligned}
$$

$$
\begin{aligned}
&\mathbf{E}|(\vec{II_3})|_{\mathbb{R}^d}^2 \leq \left(\frac{1}{\mu}\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\|\boldsymbol{b}\|_\infty\right)^2\mathbf{E}\left|\int_0^t e^{-\frac{2}{\mu}A(\mu,s)}\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}dr\right)\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}d\boldsymbol{W}_r\right)ds\right|_{\mathbb{R}^d}^2 \\
&= \left(\frac{1}{\mu}\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\|\boldsymbol{b}\|_\infty\right)^2\mathbf{E}\left|\int_0^t\left(\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))}dr\right)\left(\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))}d\boldsymbol{W}_r\right)ds\right|_{\mathbb{R}^d}^2 \\
&\leq \left(\frac{1}{\mu}\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\|\boldsymbol{b}\|_\infty\right)^2\mathbf{E}\left(\int_0^t\left(\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))}dr\right)^2 ds\right)\left(\int_0^t\left|\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))}d\boldsymbol{W}_r\right|_{\mathbb{R}^d}^2 ds\right) \\
&\leq \left(\frac{1}{\mu}\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\|\boldsymbol{b}\|_\infty\right)^2\left(\int_0^t\left(\int_0^s e^{-\frac{(s-r)\lambda_0}{\mu}}dr\right)^2 ds\right)\left(\int_0^t\mathbf{E}\left|\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))}d\boldsymbol{W}_r\right|_{\mathbb{R}^d}^2 ds\right) \\
&\leq \left(\frac{1}{\mu}\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\|\boldsymbol{b}\|_\infty\right)^2\left(\int_0^t\left(\int_0^s e^{-\frac{(s-r)\lambda_0}{\mu}}dr\right)^2 ds\right)\left(\int_0^t\left(\int_0^s e^{-\frac{2(s-r)\lambda_0}{\mu}}dr\right)ds\right) \\
&\leq \left(\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\|\boldsymbol{b}\|_\infty\right)^2\left(\frac{t}{\lambda_0}\right)^2\left(\frac{\mu t}{2\lambda_0}\right)\ .
\end{aligned}
$$

Combining these estimates we see that $\mathbf{E}|(\vec{II})|_{\mathbb{R}^d}^2 \to 0$ as $\mu \downarrow 0$. This implies that $\mathbf{E}\left|\boldsymbol{\beta}(\mu) - \int_0^t \frac{\boldsymbol{b}(\boldsymbol{q}_s^\mu)}{\lambda(\boldsymbol{q}_s^\mu)}ds\right|_{\mathbb{R}^d}^2 \to 0$ as $\mu \downarrow 0$.

2.3. Estimates of $\gamma(\mu)$ - the reason why the classical Smoluchowski-Kramers approximation does not work.

We will show that $\mathbf{E}\left|\gamma(\mu) - \int_0^t \frac{1}{\lambda(\boldsymbol{q}_s^\mu)}d\boldsymbol{W}_s\right|_{\mathbb{R}^d}^2$, in general, does not tend to 0 as $\mu \downarrow 0$. Therefore the Smoluchowski-Kramers approximation does not work in the case of purely white noise perturbation.

We have, by (2.5), that

$$
\begin{aligned}
\boldsymbol{\gamma}(\mu) &= \frac{1}{\mu}\int_0^t e^{-\frac{1}{\mu}A(\mu,s)}\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}d\boldsymbol{W}_r\right)ds \\
&= \frac{1}{\mu}\int_0^t \left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}d\boldsymbol{W}_r\right)\left(-\frac{\mu}{\lambda(\boldsymbol{q}_s^\mu)}\right)d(e^{-\frac{1}{\mu}A(\mu,s)}) \\
&= -\left[\frac{\int_0^t e^{\frac{1}{\mu}A(\mu,s)}d\boldsymbol{W}_s}{\lambda(\boldsymbol{q}_t^\mu)}e^{-\frac{1}{\mu}A(\mu,t)} - \int_0^t e^{-\frac{1}{\mu}A(\mu,s)}d\left(\frac{1}{\lambda(\boldsymbol{q}_s^\mu)}\int_0^s e^{\frac{1}{\mu}A(\mu,r)}d\boldsymbol{W}_r\right)\right] \\
&= -\frac{\int_0^t e^{\frac{1}{\mu}A(\mu,s)}d\boldsymbol{W}_s}{\lambda(\boldsymbol{q}_t^\mu)}e^{-\frac{1}{\mu}A(\mu,t)} + \int_0^t \frac{1}{\lambda(\boldsymbol{q}_s^\mu)}d\boldsymbol{W}_s + \\
&\qquad + \int_0^t e^{-\frac{1}{\mu}A(\mu,s)}\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}d\boldsymbol{W}_r\right)d\left(\frac{1}{\lambda(\boldsymbol{q}_s^\mu)}\right) \\
&= \boldsymbol{R}_{\boldsymbol{\gamma}}(\mu) + \int_0^t \frac{1}{\lambda(\boldsymbol{q}_s^\mu)}d\boldsymbol{W}_s + (\vec{III}) \ .
\end{aligned}
$$

It is easy to check that

$$
\mathbf{E}|\boldsymbol{R}_{\boldsymbol{\gamma}}(\mu)|_{\mathbb{R}^d}^2 \le \frac{1}{\lambda_0^2}\int_0^t e^{-\frac{2\lambda_0(t-s)}{\mu}}ds \le \frac{\mu}{2\lambda_0^3} \ .
$$

We have

$$
\begin{aligned}
(\vec{III}) &= \int_0^t e^{-\frac{1}{\mu}A(\mu,s)}\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}d\boldsymbol{W}_r\right)\left(-\frac{1}{\lambda^2(\boldsymbol{q}_s^\mu)}\right)\nabla\lambda(\boldsymbol{q}_s^\mu)\cdot\boldsymbol{p}_s^\mu ds \\
&= (\vec{III}_1) + (\vec{III}_2) + (\vec{III}_3)
\end{aligned}
$$

where

$$
(\vec{III}_1) = -\int_0^t \frac{e^{-\frac{2}{\mu}A(\mu,s)}}{\lambda^2(\boldsymbol{q}_s^\mu)}\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}d\boldsymbol{W}_r\right)\nabla\lambda(\boldsymbol{q}_s^\mu)\cdot\boldsymbol{p}ds \ ,
$$

$$
(\vec{III}_2) = -\frac{1}{\mu}\int_0^t \frac{e^{-\frac{2}{\mu}A(\mu,s)}}{\lambda^2(\boldsymbol{q}_s^\mu)}\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}d\boldsymbol{W}_r\right)\nabla\lambda(\boldsymbol{q}_s^\mu)\cdot\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}\boldsymbol{b}(\boldsymbol{q}_r^\mu)dr\right)ds \ ,
$$

$$
(\vec{III}_3) = -\frac{1}{\mu}\int_0^t \frac{e^{-\frac{2}{\mu}A(\mu,s)}}{\lambda^2(\boldsymbol{q}_s^\mu)}\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}d\boldsymbol{W}_r\right)\nabla\lambda(\boldsymbol{q}_s^\mu)\cdot\left(\int_0^s e^{\frac{1}{\mu}A(\mu,r)}d\boldsymbol{W}_r\right)ds \ ,
$$

We can estimate

$$\mathbf{E}|(I\vec{II}_1)|^2_{\mathbb{R}^d} \leq \left(\frac{|\boldsymbol{p}|_{\mathbb{R}^d}\|\nabla\lambda\|_\infty}{\lambda_0^2}\right)^2 \mathbf{E}\left|\int_0^t e^{-\frac{1}{\mu}A(\mu,s)}\left(\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))}d\boldsymbol{W}_r\right)ds\right|^2_{\mathbb{R}^d}$$
$$\leq \left(\frac{|\boldsymbol{p}|_{\mathbb{R}^d}\|\nabla\lambda\|_\infty}{\lambda_0^2}\right)^2 \mathbf{E}\left(\int_0^t e^{-\frac{2}{\mu}A(\mu,s)}ds\right)\left(\int_0^t \left|\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))}d\boldsymbol{W}_r\right|^2_{\mathbb{R}^d} ds\right)$$
$$\leq \left(\frac{|\boldsymbol{p}|_{\mathbb{R}^d}\|\nabla\lambda\|_\infty}{\lambda_0^2}\right)^2 \left(\int_0^t e^{-\frac{2\lambda_0 s}{\mu}}ds\right)\left(\int_0^t\left(\int_0^s e^{-\frac{2\lambda_0(s-r)}{\mu}}dr\right)ds\right)$$
$$\leq \left(\frac{|\boldsymbol{p}|_{\mathbb{R}^d}\|\nabla\lambda\|_\infty}{\lambda_0^2}\right)^2 \left(\frac{\mu}{2\lambda_0}\right)\left(\frac{\mu t}{2\lambda_0}\right) .$$

The term $(I\vec{II}_2)$ could be estimated in the same way as $(I\vec{I}_3)$:

$$\mathbf{E}|(I\vec{II}_2)|^2_{\mathbb{R}^d} \leq \left(\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\|\boldsymbol{b}\|_\infty\right)^2\left(\frac{t}{\lambda_0}\right)^2\left(\frac{\mu t}{2\lambda_0}\right) .$$

But in general one cannot estimate $\mathbf{E}|(I\vec{II}_3)|^2$ up to a term which goes to 0 as $\mu\downarrow 0$. As an example, let $\Lambda = \|\lambda\|_\infty$ and let us suppose that for $0\leq t\leq T<\infty$ we have $\nabla\lambda(\boldsymbol{q}_t^\mu)=\mathbf{e}_1$. Here $\mathbf{e}_1$ is the unit basis vector $\mathbf{e}_1=(1,0,...,0)$ in $\mathbb{R}^d$. Let W_r^k be the k-th $(1\leq k\leq d)$ component of the Wiener process $\boldsymbol{W}_r$. We have, for $0<t\leq T$:

$$\mathbf{E}|(I\vec{II}_3)|_{\mathbb{R}^d} \geq \frac{1}{\mu\Lambda^2}\mathbf{E}\left|\int_0^t\left(\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))}d\boldsymbol{W}_r\right)\left(\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))}dW_r^1\right)ds\right|_{\mathbb{R}^d}$$
$$= \frac{1}{\mu\Lambda^2}\mathbf{E}\left[\left(\int_0^t\left(\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))}dW_r^1\right)^2 ds\right)^2 + \right.$$
$$\left. + \sum_{k=2}^d\left(\int_0^t\left(\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))}dW_r^k\right)\left(\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))}dW_r^1\right)ds\right)^2\right]^{\frac{1}{2}}$$
$$\geq \frac{1}{\mu\Lambda^2}\mathbf{E}\left(\int_0^t\left(\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))}dW_r^1\right)^2 ds\right)$$
$$= \frac{1}{\mu\Lambda^2}\left(\int_0^t\left(\int_0^s \mathbf{E}e^{-\frac{2}{\mu}(A(\mu,s)-A(\mu,r))}dr\right)ds\right)$$
$$\geq \frac{1}{\mu\Lambda^2}\left(\int_0^t\left(\int_0^s e^{-\frac{2}{\mu}\Lambda(s-r)}dr\right)ds\right)$$
$$= \frac{1}{\mu\Lambda^2}\frac{\mu}{2\Lambda}\int_0^t(1-e^{-\frac{2\Lambda s}{\mu}})ds = \frac{t}{2\Lambda^3}-\frac{\mu}{4\Lambda^4}(1-e^{-\frac{2\Lambda t}{\mu}}) ,$$

which does not tend to 0 as $\mu\downarrow 0$. Since $\mathbf{E}|(I\vec{II}_3)|^2_{\mathbb{R}^d} \geq (\mathbf{E}|(I\vec{II}_3)|_{\mathbb{R}^d})^2$, we see that $\mathbf{E}|(I\vec{II}_3)|^2_{\mathbb{R}^d}$ does not go to 0 as $\mu\downarrow 0$. Now we have

$$\mathbf{E}\left|\boldsymbol{\gamma}(\mu)-\int_0^t\frac{1}{\lambda(\boldsymbol{q}_s^\mu)}d\boldsymbol{W}_s\right|^2_{\mathbb{R}^d} \geq \frac{1}{4}\mathbf{E}|(I\vec{II}_3)|^2_{\mathbb{R}^d}-\mathbf{E}|\boldsymbol{R}_{\boldsymbol{\gamma}}(\mu)|^2_{\mathbb{R}^d}-\mathbf{E}|(I\vec{II}_1)|^2_{\mathbb{R}^d}-\mathbf{E}|(I\vec{II}_2)|^2_{\mathbb{R}^d} .$$

Therefore $\mathbf{E}\left|\boldsymbol{\gamma}(\mu)-\int_0^t \frac{1}{\lambda(\boldsymbol{q}_s^\mu)}d\boldsymbol{W}_s\right|_{\mathbb{R}^d}^2$ is uniformly bounded from below by a positive constant as $\mu\downarrow 0$.

We can check now that the process $\boldsymbol{q}_t^\mu$, $0\le t\le T$, does not converge as $\mu\downarrow 0$ to the process $\boldsymbol{q}_t$, $\boldsymbol{q}_0=\boldsymbol{q}$. We have

$$\begin{aligned}\boldsymbol{q}_t^\mu &= \boldsymbol{q}+\int_0^t \frac{\boldsymbol{b}(\boldsymbol{q}_s^\mu)}{\lambda(\boldsymbol{q}_s^\mu)}ds+\int_0^t \frac{1}{\lambda(\boldsymbol{q}_s^\mu)}d\boldsymbol{W}_s+\\ &\qquad +\boldsymbol{\alpha}(\mu)+\left(\boldsymbol{\beta}(\mu)-\int_0^t \frac{\boldsymbol{b}(\boldsymbol{q}_s^\mu)}{\lambda(\boldsymbol{q}_s^\mu)}ds\right)+\left(\boldsymbol{\gamma}(\mu)-\int_0^t \frac{1}{\lambda(\boldsymbol{q}_s^\mu)}d\boldsymbol{W}_s\right)\ ,\\ \boldsymbol{q}_t &= \boldsymbol{q}+\int_0^t \frac{\boldsymbol{b}(\boldsymbol{q}_s)}{\lambda(\boldsymbol{q}_s)}ds+\int_0^t \frac{1}{\lambda(\boldsymbol{q}_s)}d\boldsymbol{W}_s\ .\end{aligned}$$

Suppose that we have, for any $\kappa, T>0$ and any $\boldsymbol{p}_0^\mu=\boldsymbol{p}\in\mathbb{R}^d$ fixed, that

$$\lim_{\mu\downarrow 0}\mathbf{P}\left(\max_{0\le t\le T}|\boldsymbol{q}_t^\mu-\boldsymbol{q}_t|_{\mathbb{R}^d}^2\ge\kappa\right)=0\ .$$

We have, for some $A>0$ independent of μ and κ, that

$$\begin{aligned}&\mathbf{E}\left|(\boldsymbol{q}_t^\mu-\boldsymbol{q}_t)-\int_0^t\left(\frac{\boldsymbol{b}(\boldsymbol{q}_s^\mu)}{\lambda(\boldsymbol{q}_s^\mu)}-\frac{\boldsymbol{b}(\boldsymbol{q}_s)}{\lambda(\boldsymbol{q}_s)}\right)ds-\int_0^t\left(\frac{1}{\lambda(\boldsymbol{q}_s^\mu)}-\frac{1}{\lambda(\boldsymbol{q}_s)}\right)d\boldsymbol{W}_s\right|_{\mathbb{R}^d}^2\\ &\le A\mathbf{E}\max_{0\le s\le t}|\boldsymbol{q}_s^\mu-\boldsymbol{q}_s|_{\mathbb{R}^d}^2\\ &\le A\left[\mathbf{P}\left(\max_{0\le s\le t}|\boldsymbol{q}_s^\mu-\boldsymbol{q}_s|_{\mathbb{R}^d}^2\ge\kappa\right)\cdot\mathbf{E}\max_{0\le s\le t}|\boldsymbol{q}_s^\mu-\boldsymbol{q}_s|_{\mathbb{R}^d}^2+\mathbf{P}\left(\max_{0\le s\le t}|\boldsymbol{q}_s^\mu-\boldsymbol{q}_s|_{\mathbb{R}^d}^2<\kappa\right)\cdot\kappa\right]\\ &\le A[\kappa+o(\mu,\kappa)]\ ,\end{aligned}$$

since $\mathbf{E}\max\limits_{0\le s\le t}|\boldsymbol{q}_s^\mu-\boldsymbol{q}_s|_{\mathbb{R}^d}^2<\infty$. Here the term $o(\mu,\kappa)$ converges to 0 as $\mu\downarrow 0$ for every fixed $\kappa>0$. Fix $\kappa>0$, let $\mu\downarrow 0$, we see that

$$\lim_{\mu\downarrow 0}\mathbf{E}\left|(\boldsymbol{q}_t^\mu-\boldsymbol{q}_t)-\int_0^t\left(\frac{\boldsymbol{b}(\boldsymbol{q}_s^\mu)}{\lambda(\boldsymbol{q}_s^\mu)}-\frac{\boldsymbol{b}(\boldsymbol{q}_s)}{\lambda(\boldsymbol{q}_s)}\right)ds-\int_0^t\left(\frac{1}{\lambda(\boldsymbol{q}_s^\mu)}-\frac{1}{\lambda(\boldsymbol{q}_s)}\right)d\boldsymbol{W}_s\right|_{\mathbb{R}^d}^2\le A\kappa\ .$$

Since $\kappa>0$ is arbitrary, we see that

$$\lim_{\mu\downarrow 0}\mathbf{E}\left|(\boldsymbol{q}_t^\mu-\boldsymbol{q}_t)-\int_0^t\left(\frac{\boldsymbol{b}(\boldsymbol{q}_s^\mu)}{\lambda(\boldsymbol{q}_s^\mu)}-\frac{\boldsymbol{b}(\boldsymbol{q}_s)}{\lambda(\boldsymbol{q}_s)}\right)ds-\int_0^t\left(\frac{1}{\lambda(\boldsymbol{q}_s^\mu)}-\frac{1}{\lambda(\boldsymbol{q}_s)}\right)d\boldsymbol{W}_s\right|_{\mathbb{R}^d}^2=0\ .$$

On the other hand, let us suppose that $\nabla\lambda(\boldsymbol{q}_t^\mu)=\mathbf{e}_1$ for $0\le t\le T<\infty$. Here $\mathbf{e}_1$ is the unit basis vector $\mathbf{e}_1=(1,0,...,0)$ in $\mathbb{R}^d$. We have

$$\mathbf{E}\left|\boldsymbol{\alpha}(\mu)+\left(\boldsymbol{\beta}(\mu)-\int_0^t \frac{\boldsymbol{b}(\boldsymbol{q}_s^\mu)}{\lambda(\boldsymbol{q}_s^\mu)}ds\right)+\left(\boldsymbol{\gamma}(\mu)-\int_0^t \frac{1}{\lambda(\boldsymbol{q}_s^\mu)}d\boldsymbol{W}_s\right)\right|_{\mathbb{R}^d}^2$$
$$\geq \frac{1}{3}\mathbf{E}\left|\boldsymbol{\gamma}(\mu)-\int_0^t \frac{1}{\lambda(\boldsymbol{q}_s^\mu)}d\boldsymbol{W}_s\right|_{\mathbb{R}^d}^2-\mathbf{E}\left|\boldsymbol{\alpha}(\mu)\right|_{\mathbb{R}^d}^2-\mathbf{E}\left|\boldsymbol{\beta}(\mu)-\int_0^t \frac{\boldsymbol{b}(\boldsymbol{q}_s^\mu)}{\lambda(\boldsymbol{q}_s^\mu)}ds\right|_{\mathbb{R}^d}^2 .$$

It follows from our estimates that this leads to a contradiction.

1.3 Regularization via approximation of the Wiener process

We could regularize the problem via *approximation of the Wiener process.* To this end we introduce the process

$$\boldsymbol{W}_t^\delta=\frac{1}{\delta}\int_0^\infty \boldsymbol{W}_s\rho\left(\frac{s-t}{\delta}\right)ds=\frac{1}{\delta}\int_0^\delta \boldsymbol{W}_{s+t}\rho\left(\frac{s}{\delta}\right)ds ,$$

where $\rho(\bullet)$ is a smooth C^∞ function whose support is contained in the interval $[0,1]$ such that

$$\int_0^1 \rho(s)ds=1 .$$

One can prove that (see [2] and the references there)

$$\lim_{\delta\downarrow 0}\mathbf{E}\max_{t\in[0,T]}|\boldsymbol{W}_t^\delta-\boldsymbol{W}_t|_{\mathbb{R}^d}^2=0 .$$

We have

$$\dot{\boldsymbol{W}}_t^\delta=-\frac{1}{\delta}\int_0^1 \boldsymbol{W}_{t+\delta r}\dot{\rho}(r)dr .$$

We can then introduce the following regularization of our problem: first we consider the system

$$\mu\ddot{\boldsymbol{q}}_t^{\mu,\delta}=\boldsymbol{b}(\boldsymbol{q}_t^{\mu,\delta})-\lambda(\boldsymbol{q}_t^{\mu,\delta})\dot{\boldsymbol{q}}_t^{\mu,\delta}+\dot{\boldsymbol{W}}_t^\delta\ ,\ \boldsymbol{q}_0^{\mu,\delta}=\boldsymbol{q}\in\mathbb{R}^d\ ,\ \dot{\boldsymbol{q}}_0^{\mu,\delta}=\boldsymbol{p}\in\mathbb{R}^d. \tag{3.1}$$

Equivalently it is the first order system

$$\begin{cases} \dot{\boldsymbol{q}}_t^{\mu,\delta}=\boldsymbol{p}_t^{\mu,\delta}\ , \\ \dot{\boldsymbol{p}}_t^{\mu,\delta}=\dfrac{1}{\mu}\boldsymbol{b}(\boldsymbol{q}_t^{\mu,\delta})-\dfrac{\lambda(\boldsymbol{q}_t^{\mu,\delta})}{\mu}\boldsymbol{p}_t^{\mu,\delta}+\dfrac{1}{\mu}\dot{\boldsymbol{W}}_t^\delta\ . \end{cases} \tag{3.2}$$

We can proceed with the estimates similar to the previous sections. Making use of the formula

$$\dot{\boldsymbol{W}}_t^\delta=-\frac{1}{\delta}\int_0^1 \boldsymbol{W}_{t+\delta r}\dot{\rho}(r)dr \tag{3.3}$$

we could prove that all the terms

$$\mathbf{E}|\boldsymbol{\alpha}(\mu)|_{\mathbb{R}^d}\ ,\ \mathbf{E}\left|\boldsymbol{\beta}(\mu)-\int_0^t \frac{\boldsymbol{b}(\boldsymbol{q}_s^\mu)}{\lambda(\boldsymbol{q}_s^\mu)}ds\right|_{\mathbb{R}^d}\ ,\ \mathbf{E}\left|\boldsymbol{\gamma}(\mu)-\int_0^t \frac{1}{\lambda(\boldsymbol{q}_s^\mu)}d\boldsymbol{W}_s^\delta\right|_{\mathbb{R}^d}$$

goes to zero as $\mu\downarrow 0$. (To be precise, we should write $\boldsymbol{\alpha}(\mu,\delta)$, $\boldsymbol{\beta}(\mu,\delta)$ and $\boldsymbol{\gamma}(\mu,\delta)$ to indicate the dependence on δ, but for brevity we neglect that.) In particular, with $\delta>0$ fixed, we can estimate the term $(\vec{III_3})$ up to a term which tends to 0 as $\mu\downarrow 0$. We have

$$\begin{aligned}
&\mathbf{E}|(\vec{III_3})|_{\mathbb{R}^d}\le \frac{1}{\mu}\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\int_0^t \mathbf{E}\left|\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))}\dot{\boldsymbol{W}}_r^\delta dr\right|_{\mathbb{R}^d}^2 ds\\
&=\frac{1}{\mu}\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\int_0^t \frac{1}{\delta^2}\mathbf{E}\left|\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))}\left(\int_0^1 \boldsymbol{W}_{r+\delta m}\dot{\rho}(m)dm\right)dr\right|_{\mathbb{R}^d}^2 ds\\
&=\frac{1}{\mu}\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\int_0^t \frac{1}{\delta^2}\mathbf{E}\left|\int_0^1 \dot{\rho}(m)\boldsymbol{W}_{r+\delta m}dm\int_0^s e^{-\frac{1}{\mu}(A(\mu,s)-A(\mu,r))}dr\right|_{\mathbb{R}^d}^2 ds\\
&\le \frac{1}{\mu}\frac{\|\nabla\lambda\|_\infty}{\lambda_0^2}\int_0^t \frac{1}{\delta^2}\left(\max_{0\le m\le 1}|\dot{\rho}(m)|\right)^2\mathbf{E}\left(\max_{0\le l\le s+\delta}|\boldsymbol{W}_l|_{\mathbb{R}^d}\right)^2\left(\int_0^s e^{-\frac{\lambda_0(s-r)}{\mu}}dr\right)^2 ds\\
&\le \mu\frac{\|\nabla\lambda\|_\infty}{\lambda_0^4}\frac{t}{\delta^2}\left(\max_{0\le m\le 1}|\dot{\rho}(m)|\right)^2\mathbf{E}\left(\max_{0\le l\le t+\delta}|\boldsymbol{W}_l|_{\mathbb{R}^d}\right)^2 .
\end{aligned}$$

Therefore, for fixed $\delta>0$, we have $\mathbf{E}|(\vec{III_3})|_{\mathbb{R}^d}\to 0$ as $\mu\downarrow 0$. By (2.4), we get:

$$\begin{aligned}
\boldsymbol{q}_t^{\mu,\delta}=\boldsymbol{q}+\int_0^t \frac{\boldsymbol{b}(\boldsymbol{q}_s^{\mu,\delta})}{\lambda(\boldsymbol{q}_s^{\mu,\delta})}ds+\int_0^t \frac{1}{\lambda(\boldsymbol{q}_s^{\mu,\delta})}d\boldsymbol{W}_s^\delta+&\\
+\boldsymbol{\alpha}(\mu)+\left(\boldsymbol{\beta}(\mu)-\int_0^t \frac{\boldsymbol{b}(\boldsymbol{q}_s^{\mu,\delta})}{\lambda(\boldsymbol{q}_s^{\mu,\delta})}ds\right)+\left(\boldsymbol{\gamma}(\mu)-\int_0^t \frac{1}{\lambda(\boldsymbol{q}_s^{\mu,\delta})}d\boldsymbol{W}_s^\delta\right)&\ .
\end{aligned}\tag{3.4}$$

Let the process $\tilde{\boldsymbol{q}}_t^\delta$ be governed by the equation

$$\dot{\tilde{\boldsymbol{q}}}_t^\delta=\frac{\boldsymbol{b}(\tilde{\boldsymbol{q}}_t^\delta)}{\lambda(\tilde{\boldsymbol{q}}_t^\delta)}+\frac{1}{\lambda(\tilde{\boldsymbol{q}}_t^\delta)}\dot{\boldsymbol{W}}_t^\delta\ ,\ \tilde{\boldsymbol{q}}_0^\delta=\boldsymbol{q}\in\mathbb{R}^d\ .\tag{3.5}$$

Then

$$\tilde{\boldsymbol{q}}_t^\delta=\boldsymbol{q}+\int_0^t \frac{\boldsymbol{b}(\tilde{\boldsymbol{q}}_s^\delta)}{\lambda(\tilde{\boldsymbol{q}}_s^\delta)}ds+\int_0^t \frac{1}{\lambda(\tilde{\boldsymbol{q}}_s^\delta)}d\boldsymbol{W}_s^\delta\ .\tag{3.6}$$

Let $M(t,\delta,\mu)=\mathbf{E}\max\limits_{0\le s\le t}|\boldsymbol{q}_s^{\mu,\delta}-\tilde{\boldsymbol{q}}_s^\delta|_{\mathbb{R}^d}$. By (3.4) and (3.6), using estimate (3.3), we have

$$M(t,\delta,\mu)\le K_1\int_0^t M(s,\delta,\mu)ds+K_2(t,\delta)\int_0^t M(s,\delta,\mu)ds+o_\mu(1)\ .$$

Here $o_\mu(1)$ is a term which goes to 0 as $\mu\downarrow 0$. The positive constant K_1 is independent of μ, δ and t. The positive constant $K_2=K_2(t,\delta)$ may depend on t and δ, but is independent of μ. Now we use the Bellman-Gronwall inequality:

$$M(t, \delta, \mu) \leq o_\mu(1) \exp((K_1 + K_2(t, \delta))t) \ .$$

We conclude that for any $\delta, \kappa, T > 0$ fixed and any $\boldsymbol{p}_0^{\mu,\delta} = \boldsymbol{p}$ fixed,

$$\lim_{\mu \downarrow 0} \mathbf{P} \left(\max_{0 \leq t \leq T} |\boldsymbol{q}_t^{\mu,\delta} - \widetilde{\boldsymbol{q}}_t^\delta|_{\mathbb{R}^d} > \kappa \right) = 0 \ .$$

Now we can take $\delta \downarrow 0$. Using Theorem 6.7.2 from [22] we get the following result.

Theorem 3.1. *We have, as $\delta \downarrow 0$, that*

$$\lim_{\delta \to 0} \mathbf{E} \max_{t \in [0,T]} |\widetilde{\boldsymbol{q}}_t^\delta - \widehat{\boldsymbol{q}}_t|_{\mathbb{R}^d} = 0 \ ,$$

where $\widehat{\boldsymbol{q}}_t$ is the solution of the problem

$$\dot{\widehat{\boldsymbol{q}}}_t = \frac{\boldsymbol{b}(\widehat{\boldsymbol{q}}_t)}{\lambda(\widehat{\boldsymbol{q}}_t)} + \frac{1}{\lambda(\widehat{\boldsymbol{q}}_t)} \circ \dot{\boldsymbol{W}}_t \ , \ \widehat{\boldsymbol{q}}_0 = \boldsymbol{q} \in \mathbb{R}^d \ . \tag{3.6}$$

Here the stochastic term is understood in the Stratonovich sense.

In the general case

$$\mu \ddot{\boldsymbol{q}}_t^{\mu,\delta} = \boldsymbol{b}(\boldsymbol{q}_t^{\mu,\delta}) - \lambda(\boldsymbol{q}_t^{\mu,\delta}) \dot{\boldsymbol{q}}_t^{\mu,\delta} + \sigma(\boldsymbol{q}_t^{\mu,\delta}) \dot{\boldsymbol{W}}_t^\delta \ , \ \boldsymbol{q}_0^{\mu,\delta} = \boldsymbol{q} \ , \ \dot{\boldsymbol{q}}_0^{\mu,\delta} = \boldsymbol{p} \ , \tag{3.7}$$

where the matrix $\sigma(\bullet)$ satisfy assumptions made in Section 1, we have, similarly, that for any $\delta, \kappa, T > 0$ fixed and any $\boldsymbol{p}_0^{\mu,\delta} = \boldsymbol{p}$ fixed,

$$\lim_{\mu \downarrow 0} \mathbf{P} \left(\max_{0 \leq t \leq T} |\boldsymbol{q}_t^{\mu,\delta} - \widetilde{\boldsymbol{q}}_t^\delta|_{\mathbb{R}^d} > \kappa \right) = 0 \ .$$

The process $\widetilde{\boldsymbol{q}}_t^\delta$ is governed by the equation

$$\dot{\widetilde{\boldsymbol{q}}}_t^\delta = \frac{\boldsymbol{b}(\widetilde{\boldsymbol{q}}_t^\delta)}{\lambda(\widetilde{\boldsymbol{q}}_t^\delta)} + \frac{\sigma(\widetilde{\boldsymbol{q}}_t^\delta)}{\lambda(\widetilde{\boldsymbol{q}}_t^\delta)} \dot{\boldsymbol{W}}_t^\delta \ , \ \widetilde{\boldsymbol{q}}_0^\delta = \boldsymbol{q} \in \mathbb{R}^d \ . \tag{3.8}$$

And we conclude with

Theorem 3.2. *Under the assumptions mentioned above,*

$$\lim_{\delta \to 0} \mathbf{E} \max_{t \in [0,T]} |\widetilde{\boldsymbol{q}}_t^\delta - \widehat{\boldsymbol{q}}_t|_{\mathbb{R}^d} = 0 \ ,$$

where $\widehat{\boldsymbol{q}}_t$ is the solution of the problem

$$\dot{\widehat{\boldsymbol{q}}}_t = \frac{\boldsymbol{b}(\widehat{\boldsymbol{q}}_t)}{\lambda(\widehat{\boldsymbol{q}}_t)} + \frac{\sigma(\widehat{\boldsymbol{q}}_t)}{\lambda(\widehat{\boldsymbol{q}}_t)} \circ \dot{\boldsymbol{W}}_t \ , \ \widehat{\boldsymbol{q}}_0 = \boldsymbol{q} \in \mathbb{R}^d \ . \tag{3.9}$$

1.4 One dimensional case

In the case of one space variable, Smoluchowski-Kramers approximation leads to an one-dimensional diffusion process q_t which is defined by the following stochastic differential equation written in the Itô form:

$$\dot{q}_t = \frac{b(q_t)}{\lambda(q_t)} - \frac{\lambda'(q_t)}{2\lambda^3(q_t)} + \frac{1}{\lambda(q_t)}\dot{W}_t \ , \ q_0 = q \in \mathbb{R}^1 \ . \tag{4.1}$$

Put

$$\begin{aligned} u(q) &= \int_0^q \lambda(x) \exp\left(-2\int_0^x b(y)\lambda(y)dy\right) dx \ , \\ v(q) &= 2\int_0^q \lambda(x) \exp\left(2\int_0^x b(y)\lambda(y)dy\right) dx \ . \end{aligned} \tag{4.2}$$

Since $\lambda(x) > 0$, $u(q)$ and $v(q)$ are strictly increasing functions. Following [7] we introduce an operator $D_v D_u$, where D_u means the differentiation with respect to the monotone function $u(q)$: $D_u f(q) = \lim\limits_{h\to 0} \dfrac{f(x+h)-f(x)}{u(x+h)-u(x)}$; the operator D_v is defined in a similar way. One can check that $D_v D_u$ is the generator of the diffusion process q_t defined by (4.1).

Suppose now that the friction coefficient $\lambda(q) = \lambda_\varepsilon(q)$ depends on a parameter $\varepsilon > 0$. We assume that, for each $\varepsilon \in (0,1]$, $\lambda_\varepsilon(q)$ has a bounded continuous derivative $\lambda'_\varepsilon(q)$, and $0 < \underline{\lambda} \le \lambda_\varepsilon(q) \le \overline{\lambda} < \infty$. Let $u_\varepsilon(q)$ and $v_\varepsilon(q)$ be the functions defined by (4.2) when $\lambda(q)$ is replaced by $\lambda_\varepsilon(q)$.

Consider the stochastic process $q_t^{\mu,\delta,\varepsilon}$ in $\mathbb{R}^1$ defined by the equation

$$\mu\ddot{q}_t^{\mu,\delta,\varepsilon} = b(q_t^{\mu,\delta,\varepsilon}) - \lambda^\varepsilon(q_t^{\mu,\delta,\varepsilon})q_t^{\mu,\delta,\varepsilon} + \dot{W}_t^\delta \ , \ q_0^{\mu,\delta,\varepsilon} = q \ , \ \dot{q}_0^{\mu,\delta,\varepsilon} = p \ . \tag{4.3}$$

where $\dot{W}_t^\delta$ is, as before, a "smoothed" white noise converging to $\dot{W}_t$ as $\delta \downarrow 0$.

Theorem 4.1. *Assume that the function $\lambda_\varepsilon(q)$ converge weakly as $\varepsilon \downarrow 0$ on each finite interval $[\alpha,\beta] \subset \mathbb{R}^1$ to a function $\overline{\lambda}(q)$ (maybe, discontinuous). Then processes $q_t^{\mu,\delta,\varepsilon}$ converge weakly on each finite time interval to the diffusion process $\overline{q}_t$ governed by the generator $D_{\overline{v}}D_{\overline{u}}$ (where $\overline{u}(q)$ and $\overline{v}(q)$ defined by (4.2) with $\lambda = \overline{\lambda}(q)$) as, first $\mu \downarrow 0$, then $\delta \downarrow 0$, and then $\varepsilon \downarrow 0$.*

Proof. According to Section 3, processes $q_t^{\mu,\delta,\varepsilon}$ converge weakly as first $\mu \downarrow 0$ and then $\delta \downarrow 0$ to the process $\widehat{q}_t^\delta$ which solves equation (4.1) with $\lambda(q) = \lambda^\varepsilon(q)$. It follows from our assumptions that functions $u_\varepsilon(q)$ and $v_\varepsilon(q)$ converge as $\varepsilon \downarrow 0$ to functions $\overline{u}(q)$ and $\overline{v}(q)$ respectively for each $q \in \mathbb{R}^1$. The functions $\overline{u}(q)$ and $\overline{v}(q)$ are continuous and strictly increasing. Therefore ([7]) a diffusion process $\overline{q}_t$ exists governed by $D_{\overline{v}}D_{\overline{u}}$. As shown in [19], convergence of $u_\varepsilon(q)$ and $v_\varepsilon(q)$ as $\varepsilon \downarrow 0$ to $\overline{u}(q)$ and $\overline{v}(q)$ respectively implies weak convergence of processes q_t^ε to the process corresponding to $D_{\overline{v}} D_{\overline{u}}$ as $\varepsilon \downarrow 0$. $\square$

Theorem 4.2. *Let $\lambda_\varepsilon(q) = \widetilde{\lambda}\left(\frac{q}{\varepsilon}\right)$. Assume that one of the following conditions is satisfied:*

1. $\widetilde{\lambda}(q)$ is a continuously differentiable positive 1-periodic function;

2. $\widetilde{\lambda}(q)$ is an ergodic stationary process (independent of the process W_t in (4.3)) with continuously differentiable trajectories and $0 < \lambda_- \leq \widetilde{\lambda}(q) \leq \lambda_+ < \infty$ for some constants λ_-, λ_+.

Put $\overline{\lambda} = \int_0^1 \widetilde{\lambda}(q)dq$ if condition 1 is satisfied, and $\overline{\lambda} = \mathbf{E}\widetilde{\lambda}(q)$ is condition 2 is satisfied.

Then the process $q_t^{\mu,\delta,\varepsilon}$ defined by (5.3) converge weakly when first $\mu \downarrow 0$ and then $\varepsilon \downarrow 0$ to the process $\overline{q}_t$ defined by the equation

$$\overline{q}_t = \frac{1}{\overline{\lambda}} b(\overline{q}_t) + \frac{1}{\overline{\lambda}} \dot{W}_t \ , \ \overline{q}_0 = q \ .$$

Proof of this theorem follows from Theorem 4.1 since each of conditions 1 and 2 implies conditions of Theorem 4.1 and $\overline{\lambda}(q) = \overline{\lambda}$. $\square$

Assume now that $\lambda_\varepsilon(q)$ is a bounded and separated from zero uniformly in $\varepsilon \in (0,1]$ positive function such that $\lim_{\varepsilon \downarrow 0} \lambda_\varepsilon(q) = \lambda_1$ for $q < 0$, and $\lim_{\varepsilon \downarrow 0} \lambda_\varepsilon(q) = \lambda_2$ for $q > 0$. Assume that $\lambda_\varepsilon(q)$ is continuously differentiable for each $\varepsilon > 0$. Let $\widehat{\lambda}(q)$ be the step function equal to λ_1 for $q \leq 0$ and to λ_2 for $q > 0$. Let functions $\widehat{u}(q)$ and $\widehat{v}(q)$ be defined by formula (4.2) with $\lambda(q) = \widehat{\lambda}(q)$; $\widehat{u}(q)$ and $\widehat{v}(q)$ are continuous strictly increasing functions. Denote by $\widehat{q}_t$ the diffusion process in $\mathbb{R}^1$ governed by the generator $A = D_{\widehat{v}} D_{\widehat{u}}$. The process $\widehat{q}_t$ behaves as $\frac{1}{\lambda_1} W_t$ on the negative part of axis q and as $\frac{1}{\lambda_2} W_t$ on the positive part. Its behavior at $q = 0$ is defined by the domain of definition $\mathfrak{D}_A$ of the generator A: a continuous bounded function $f(q)$, $q \in \mathbb{R}^1$, twice continuously differentiable at $q \in \{\mathbb{R}^1 \setminus \{q = 0\}\}$ belongs to $\mathfrak{D}_A$ if and only if left and right derivatives at $q = 0$, $f'_-(0)$ and $f'_+(0)$ respectively, satisfy the equality $\frac{1}{\lambda_1} f'_-(0) = \frac{1}{\lambda_2} f'_+(0)$ and $Af(q)$ is continuous.

It is easy to see that functions $u_\varepsilon(q)$ and $v_\varepsilon(q)$ defined by (4.2) with $\lambda(q) = \lambda_\varepsilon(q)$ converge as $\varepsilon \downarrow 0$ to $\widehat{u}(q)$ and $\widehat{v}(q)$ respectively for each $q \in \mathbb{R}^1$. This implies the following result.

Theorem 4.3. *Let the friction coefficient $\lambda_\varepsilon(q)$ satisfies the conditions mentioned* *above. Then the stochastic process $q_t^{\mu,\delta,\varepsilon}$ defined by (4.3) converges weakly to the diffusion process $\widehat{q}_t$ in $\mathbb{R}^1$ governed by $A = D_{\widehat{v}} D_{\widehat{u}}$ as first $\mu \downarrow 0$, then $\delta \downarrow 0$, and then $\varepsilon \downarrow 0$.*

This means, roughly speaking, that, if the friction coefficient is close to the step-function $\widehat{\lambda}(q)$, then process q_t^{μ}, for $0 < \mu << 1$, can be approximated by the diffusion process $\widehat{q}_t$.

1.5 Multidimensional case

In this section we consider the problem of fast oscillating periodic environment in multidimensional case. We consider the system

$$\mu \ddot{\boldsymbol{q}}_t^{\mu,\delta,\varepsilon} = \boldsymbol{b}\left(\frac{\boldsymbol{q}_t^{\mu,\delta,\varepsilon}}{\varepsilon}\right) - \lambda\left(\frac{\boldsymbol{q}_t^{\mu,\delta,\varepsilon}}{\varepsilon}\right)\dot{\boldsymbol{q}}_t^{\mu,\delta,\varepsilon} + \dot{\boldsymbol{W}}_t^{\delta} \ , \ \boldsymbol{q}_0^{\mu,\delta,\varepsilon} = \boldsymbol{q} \in \mathbb{R}^d \ , \ \dot{\boldsymbol{q}}_0^{\mu,\delta,\varepsilon} = \boldsymbol{p} \in \mathbb{R}^d \ . \quad (5.1)$$

Here as in Section 3 the process $\boldsymbol{W}_t^{\delta}$ is the approximation of the Wiener process in $\mathbb{R}^d$. We make the same assumptions about the functions $\lambda(\bullet)$ and $\boldsymbol{b}(\bullet)$ as in Section 2. In addition we assume that the functions $\lambda(\bullet)$ and $\boldsymbol{b}(\bullet)$ are 1-periodic, i.e. $\lambda(\boldsymbol{x}+\mathbf{e}_k) = \lambda(\boldsymbol{x})$ and $\boldsymbol{b}(\boldsymbol{x}+\mathbf{e}_k) = \boldsymbol{b}(\boldsymbol{x})$ for $\boldsymbol{x} \in \mathbb{R}^d$ and $\mathbf{e}_k = (0, 0, ..., 1(k\text{-th coordinate}), ..., 0)$, $1 \leq k \leq d$. Under this assumption our system (5.1) could be regarded as a system on the d-torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$. Fix $\varepsilon > 0$, we can proceed as in Section 3 to see that first as $\mu \downarrow 0$ then as $\delta \downarrow 0$ the process $\boldsymbol{q}_t^{\mu,\delta,\varepsilon}$ converges in probability to the process $\boldsymbol{q}_t^{\varepsilon}$ subjected to

$$\dot{\boldsymbol{q}}_t^{\varepsilon} = \frac{\boldsymbol{b}\left(\dfrac{\boldsymbol{q}_t^{\varepsilon}}{\varepsilon}\right)}{\lambda\left(\dfrac{\boldsymbol{q}_t^{\varepsilon}}{\varepsilon}\right)} + \frac{1}{\lambda\left(\dfrac{\boldsymbol{q}_t^{\varepsilon}}{\varepsilon}\right)} \circ \dot{\boldsymbol{W}}_t \ , \boldsymbol{q}_0^{\varepsilon} = \boldsymbol{q} \in \mathbb{R}^d \ .$$

The above equation, written in the form of Itô integral, will be

$$\dot{\boldsymbol{q}}_t^{\varepsilon} = \frac{\boldsymbol{b}\left(\dfrac{\boldsymbol{q}_t^{\varepsilon}}{\varepsilon}\right)}{\lambda\left(\dfrac{\boldsymbol{q}_t^{\varepsilon}}{\varepsilon}\right)} - \frac{1}{2\varepsilon}\frac{\nabla\lambda\left(\dfrac{\boldsymbol{q}_t^{\varepsilon}}{\varepsilon}\right)}{\lambda^3\left(\dfrac{\boldsymbol{q}_t^{\varepsilon}}{\varepsilon}\right)} + \frac{1}{\lambda\left(\dfrac{\boldsymbol{q}_t^{\varepsilon}}{\varepsilon}\right)}\dot{\boldsymbol{W}}_t \ , \boldsymbol{q}_0^{\varepsilon} = \boldsymbol{q} \in \mathbb{R}^d \ . \quad (5.2)$$

The generator corresponding to (5.2) is the second order differential operator

$$L^{\varepsilon}u(\boldsymbol{x}) = \left(\frac{\boldsymbol{b}\left(\dfrac{\boldsymbol{x}}{\varepsilon}\right)}{\lambda\left(\dfrac{\boldsymbol{x}}{\varepsilon}\right)} - \frac{1}{2\varepsilon}\frac{\nabla\lambda\left(\dfrac{\boldsymbol{x}}{\varepsilon}\right)}{\lambda^3\left(\dfrac{\boldsymbol{x}}{\varepsilon}\right)}\right) \cdot \nabla u(\boldsymbol{x}) + \frac{1}{2}\frac{1}{\lambda^2\left(\dfrac{\boldsymbol{x}}{\varepsilon}\right)}\Delta u(\boldsymbol{x}) \ . \quad (5.3)$$

Our goal is to study the homogenization properties of (5.3) for general multidimensional case. Homogenization problems are considered by many authors. However, we provide here an elementary probabilistic way of doing this. Our method follows [10] and [11] (pp. 104-106).

Let us first make a change of variable $\dfrac{\boldsymbol{q}}{\varepsilon} = \boldsymbol{y}$ and $\dfrac{t}{\varepsilon^2} = s$. The process $\boldsymbol{y}_s^\varepsilon = \dfrac{1}{\varepsilon}\boldsymbol{q}_t^\varepsilon$ corresponds to the generator

$$A^\varepsilon = \frac{1}{2\lambda^2(\boldsymbol{y})}\Delta_{\boldsymbol{y}} - \frac{\nabla\lambda(\boldsymbol{y})}{2\lambda^3(\boldsymbol{y})}\cdot\nabla_{\boldsymbol{y}} + \varepsilon\frac{\boldsymbol{b}(\boldsymbol{y})}{\lambda(\boldsymbol{y})}\cdot\nabla_{\boldsymbol{y}} \ .$$

We regard $\boldsymbol{y}_s^\varepsilon$ as a process on $\mathbb{T}^d$. Then we have the bound

$$\left|\mathbf{E}_{\boldsymbol{q}/\varepsilon} f(\boldsymbol{y}_s^\varepsilon) - \int_{\mathbb{T}^d} f(\boldsymbol{x})\mu^\varepsilon(\boldsymbol{x})d\boldsymbol{x}\right| < Ke^{-as} \ .$$

Here $K>0$ and $a>0$ are independent of ε for small ε. The function f is bounded and measurable. The function $\mu^\varepsilon(\boldsymbol{x})$ is the density of the unique invariant measure of $\boldsymbol{y}_s^\varepsilon$ on $\mathbb{T}^d$ and $\displaystyle\int_{\mathbb{T}^d}\mu^\varepsilon(\boldsymbol{x})d\boldsymbol{x} = 1$. We have

$$\lim_{\varepsilon\downarrow 0}\mu^\varepsilon(\boldsymbol{x}) = \mu(\boldsymbol{x}) \ , \ \lim_{\varepsilon\downarrow 0}\int_{\mathbb{T}^d} f(\boldsymbol{x})\mu^\varepsilon(\boldsymbol{x})d\boldsymbol{x} = \int_{\mathbb{T}^d} f(\boldsymbol{x})\mu(\boldsymbol{x})d\boldsymbol{x}$$

for $f\in C(\mathbb{T}^d)$ and $\mu(\boldsymbol{x})$ the unique invariant measure for the process with generator A^0 on $\mathbb{T}^d$ and $\displaystyle\int_{\mathbb{T}^d}\mu(\boldsymbol{x})d\boldsymbol{x} = 1$. Combining these estimates we have, that for any n, for any $t\geq\delta>0$, there exist $\varepsilon_0(n,\delta)>0$ such that for any $0<\varepsilon<\varepsilon_0(n,\delta)$, we have

$$\left|\mathbf{E}_{\boldsymbol{q}} f\left(\frac{\boldsymbol{q}_t^\varepsilon}{\varepsilon}\right) - \int_{\mathbb{T}^d} f(\boldsymbol{x})\mu(\boldsymbol{x})d\boldsymbol{x}\right| < \frac{1}{n} \ .$$

This implies that for any $f\in C(\mathbb{T}^d)$,

$$\lim_{\varepsilon\downarrow 0}\sup_{t\geq\delta}\left|\mathbf{E}_{\boldsymbol{q}} f\left(\frac{\boldsymbol{q}_t^\varepsilon}{\varepsilon}\right) - \int_{\mathbb{T}^d} f(\boldsymbol{x})\mu(\boldsymbol{x})d\boldsymbol{x}\right| = 0 \ .$$

Finally we calculate the density $\mu(\boldsymbol{x})$. Since

$$A^0 = \frac{1}{2\lambda^2(\boldsymbol{y})}\Delta_{\boldsymbol{y}} - \frac{\nabla\lambda(\boldsymbol{y})}{2\lambda^3(\boldsymbol{y})}\cdot\nabla_{\boldsymbol{y}} = \frac{1}{2\lambda^2(\boldsymbol{y})}(\Delta_{\boldsymbol{y}} - \nabla(\ln\lambda(\boldsymbol{y}))\cdot\nabla_{\boldsymbol{y}}) \ ,$$

we see that $\mu(\boldsymbol{x}) = C\lambda(\boldsymbol{x})$ with $C = \left(\displaystyle\int_{\mathbb{T}^d}\lambda(\boldsymbol{x})d\boldsymbol{x}\right)^{-1}$ and we have the following result:

Lemma 5.1. *For any $f\in C(\mathbb{T}^d)$, we have*

$$\lim_{\varepsilon\downarrow 0}\sup_{t\geq\delta}\left|\mathbf{E}_{\boldsymbol{q}} f\left(\frac{\boldsymbol{q}_t^\varepsilon}{\varepsilon}\right) - \frac{\displaystyle\int_{\mathbb{T}^d} f(\boldsymbol{x})\lambda(\boldsymbol{x})d\boldsymbol{x}}{\displaystyle\int_{\mathbb{T}^d}\lambda(\boldsymbol{x})d\boldsymbol{x}}\right| = 0 \ . \tag{5.4}$$

Corollary. *For any bounded continuous function $f(\boldsymbol{x})$ on $\mathbb{T}^d$, $\boldsymbol{q} \in \mathbb{T}^d$ we have*

$$\mathbf{E}_{\boldsymbol{q}}\left[\int_0^t f\left(\frac{\boldsymbol{q}_s^\varepsilon}{\varepsilon}\right)ds - \frac{t\displaystyle\int_{\mathbb{T}^d} f(\boldsymbol{x})\lambda(\boldsymbol{x})d\boldsymbol{x}}{\displaystyle\int_{\mathbb{T}^d}\lambda(\boldsymbol{x})d\boldsymbol{x}}\right]^2 \to 0$$

as $\varepsilon \downarrow 0$, for $0 < t < \infty$.

The *proof* of this corollary follows the same proof of the corollary after Lemma 1 in [10].

Now let us consider auxiliary functions $N_k(\boldsymbol{y})$, $k = 1, ..., d$, which are the periodic bounded solutions (i.e., on $\mathbb{T}^d$) of the equations

$$\frac{1}{2\lambda^2(\boldsymbol{y})}\Delta_{\boldsymbol{y}} N_k(\boldsymbol{y}) - \frac{\nabla_{\boldsymbol{y}}\lambda(\boldsymbol{y})}{2\lambda^3(\boldsymbol{y})}\cdot\nabla_{\boldsymbol{y}} N_k(\boldsymbol{y}) = A^0(N_k(\boldsymbol{y})) = \frac{1}{2\lambda^3(\boldsymbol{y})}\frac{\partial\lambda}{\partial y_k}(\boldsymbol{y}) \ , \ \boldsymbol{y} \in \mathbb{T}^d \ . \quad (5.5)$$

The solvability of this equation comes from the fact that $(A^0)^*\lambda(\boldsymbol{y}) = 0$ and $\displaystyle\int_{\mathbb{T}^d}\frac{1}{2\lambda^3(\boldsymbol{y})}\frac{\partial\lambda}{\partial y_k}(\boldsymbol{y})\lambda(\boldsymbol{y})d\boldsymbol{y} = 0$. The boundedness of solution comes from our assumptions on the function $\lambda(\bullet)$. Now we apply Itô's formula:

$$\begin{aligned}
&\varepsilon N_k\left(\frac{\boldsymbol{q}_t^\varepsilon}{\varepsilon}\right) - \varepsilon N_k\left(\frac{\boldsymbol{q}}{\varepsilon}\right) \\
&= \varepsilon\left[\int_0^t \nabla N_k\left(\frac{\boldsymbol{q}_s^\varepsilon}{\varepsilon}\right)\cdot\frac{1}{\varepsilon}\left(\frac{\boldsymbol{b}}{\lambda}\left(\frac{\boldsymbol{q}_s^\varepsilon}{\varepsilon}\right) - \frac{1}{2\varepsilon}\frac{\nabla\lambda}{\lambda^3}\left(\frac{\boldsymbol{q}_s^\varepsilon}{\varepsilon}\right) + \frac{\dot{\boldsymbol{W}}_s}{\lambda\left(\frac{\boldsymbol{q}_s^\varepsilon}{\varepsilon}\right)}\right)ds + \right. \\
&\qquad\qquad \left. +\frac{1}{2}\int_0^t \Delta N_k\left(\frac{\boldsymbol{q}_s^\varepsilon}{\varepsilon}\right)\frac{1}{\varepsilon^2}\frac{1}{\lambda^2\left(\frac{\boldsymbol{q}_s^\varepsilon}{\varepsilon}\right)}ds\right] \qquad (5.6)\\
&= \int_0^t \nabla N_k\left(\frac{\boldsymbol{q}_s^\varepsilon}{\varepsilon}\right)\cdot\left(\frac{\boldsymbol{b}}{\lambda}\left(\frac{\boldsymbol{q}_s^\varepsilon}{\varepsilon}\right) + \frac{\dot{\boldsymbol{W}}_s}{\lambda\left(\frac{\boldsymbol{q}_s^\varepsilon}{\varepsilon}\right)}\right)ds + \frac{1}{2\varepsilon}\int_0^t\frac{\frac{\partial\lambda}{\partial y_k}\left(\frac{\boldsymbol{q}_s^\varepsilon}{\varepsilon}\right)}{\lambda^3\left(\frac{\boldsymbol{q}_s^\varepsilon}{\varepsilon}\right)}ds \ .
\end{aligned}$$

Let $\boldsymbol{N}(\boldsymbol{y}) = (N_1(\boldsymbol{y}), ..., N_d(\boldsymbol{y}))$. Using (5.5) we have

$$\boldsymbol{q}_t^\varepsilon - \boldsymbol{q} = \int_0^t \left(\frac{\boldsymbol{b}}{\lambda}\left(\frac{\boldsymbol{q}_s^\varepsilon}{\varepsilon}\right) + \frac{\dot{\boldsymbol{W}}_s}{\lambda\left(\frac{\boldsymbol{q}_s^\varepsilon}{\varepsilon}\right)} \right) ds + \int_0^t (D\boldsymbol{N})\left(\frac{\boldsymbol{q}_s^\varepsilon}{\varepsilon}\right) \left(\frac{\boldsymbol{b}}{\lambda}\left(\frac{\boldsymbol{q}_s^\varepsilon}{\varepsilon}\right) + \frac{\dot{\boldsymbol{W}}_s}{\lambda\left(\frac{\boldsymbol{q}_s^\varepsilon}{\varepsilon}\right)} \right) ds + \\ -\varepsilon \left(\boldsymbol{N}\left(\frac{\boldsymbol{q}_t^\varepsilon}{\varepsilon}\right) - \boldsymbol{N}\left(\frac{\boldsymbol{q}}{\varepsilon}\right) \right) ;$$

Here $(D\boldsymbol{N})(\boldsymbol{y}) = \left(\dfrac{\partial N_i}{\partial y_j}\right)_{1\le i,j\le d}$, $\boldsymbol{y} = (y_1, ..., y_d) \in \mathbb{T}^d$.

Therefore using the corollary after Lemma 1, we see that $\boldsymbol{q}_t^\varepsilon$ converges weakly to a process $\boldsymbol{q}_t$, $\boldsymbol{q}_0 = \boldsymbol{q} \in \mathbb{R}^d$ governed by the operator

$$\overline{L} = \frac{1}{2}\sum_{i,j=1}^d \bar{a}_{ij}\frac{\partial^2}{\partial y_i \partial y_j} + \sum_{i=1}^d \bar{b}_i \frac{\partial}{\partial y_i} \ , \tag{5.7}$$

with coefficients

$$\bar{a}_{ij} = \int_{\mathbb{T}^d} \left(\frac{\nabla N_i(\boldsymbol{y}) \cdot \nabla N_j(\boldsymbol{y})}{\lambda(\boldsymbol{y})} + \frac{1}{\lambda(\boldsymbol{y})}\left(\frac{\partial N_j}{\partial y_i}(\boldsymbol{y}) + \frac{\partial N_i}{\partial y_j}(\boldsymbol{y}) \right) + \delta_{ij}\frac{1}{\lambda(\boldsymbol{y})} \right) d\boldsymbol{y} \Big/ \left(\int_{\mathbb{T}^d} \lambda(\boldsymbol{y}) d\boldsymbol{y} \right) \ ,$$

$$\bar{b}_i = \frac{\displaystyle\int_{\mathbb{T}^d} b_i(\boldsymbol{y}) d\boldsymbol{y}}{\displaystyle\int_{\mathbb{T}^d} \lambda(\boldsymbol{y}) d\boldsymbol{y}} + \sum_{k=1}^d \frac{\displaystyle\int_{\mathbb{T}^d} b_k(\boldsymbol{y}) \frac{\partial N_i}{\partial y_k}(\boldsymbol{y}) d\boldsymbol{y}}{\displaystyle\int_{\mathbb{T}^d} \lambda(\boldsymbol{y}) d\boldsymbol{y}} \ . \tag{5.8}$$

Here $\delta_{ij} = 1$ if $i = j$, and $\delta_{ij} = 0$ otherwise.

We could simplify the expression for $\bar{a}_{ij}$: using (5.5) we get

$$\begin{aligned} \bar{a}_{ij} &= \int_{\mathbb{T}^d} \left(\frac{\nabla N_i(\boldsymbol{y}) \cdot \nabla N_j(\boldsymbol{y})}{\lambda(\boldsymbol{y})} + \frac{1}{\lambda(\boldsymbol{y})}\left(\frac{\partial N_j}{\partial y_i}(\boldsymbol{y}) + \frac{\partial N_i}{\partial y_j}(\boldsymbol{y}) \right) + \delta_{ij}\frac{1}{\lambda(\boldsymbol{y})} \right) d\boldsymbol{y} \Big/ \left(\int_{\mathbb{T}^d} \lambda(\boldsymbol{y}) d\boldsymbol{y} \right) \\ &= \int_{\mathbb{T}^d} \left(\operatorname{div}\left(\frac{N_i(\boldsymbol{y})}{\lambda(\boldsymbol{y})} \nabla N_j(\boldsymbol{y}) \right) - \frac{N_i(\boldsymbol{y})}{\lambda(\boldsymbol{y})} \Delta N_j(\boldsymbol{y}) - N_i(\boldsymbol{y}) \nabla N_j(\boldsymbol{y}) \cdot \nabla \left(\frac{1}{\lambda(\boldsymbol{y})} \right) + \right. \\ &\qquad \left. + \frac{1}{\lambda(\boldsymbol{y})}\left(\frac{\partial N_j}{\partial y_i}(\boldsymbol{y}) + \frac{\partial N_i}{\partial y_j}(\boldsymbol{y}) \right) + \delta_{ij}\frac{1}{\lambda(\boldsymbol{y})} \right) d\boldsymbol{y} \Big/ \left(\int_{\mathbb{T}^d} \lambda(\boldsymbol{y}) d\boldsymbol{y} \right) \\ &= \int_{\mathbb{T}^d} \left(\frac{\partial}{\partial y_j}\left(N_i(\boldsymbol{y}) \frac{1}{\lambda(\boldsymbol{y})} \right) - \frac{1}{\lambda(\boldsymbol{y})} \frac{\partial N_i}{\partial y_j}(\boldsymbol{y}) + \right. \\ &\qquad \left. + \frac{1}{\lambda(\boldsymbol{y})}\left(\frac{\partial N_j}{\partial y_i}(\boldsymbol{y}) + \frac{\partial N_i}{\partial y_j}(\boldsymbol{y}) \right) + \delta_{ij}\frac{1}{\lambda(\boldsymbol{y})} \right) d\boldsymbol{y} \Big/ \left(\int_{\mathbb{T}^d} \lambda(\boldsymbol{y}) d\boldsymbol{y} \right) \\ &= \frac{\displaystyle\int_{\mathbb{T}^d} \frac{\partial N_j}{\partial y_i}(\boldsymbol{y}) \frac{1}{\lambda(\boldsymbol{y})} d\boldsymbol{y}}{\displaystyle\int_{\mathbb{T}^d} \lambda(\boldsymbol{y}) d\boldsymbol{y}} + \delta_{ij} \frac{\displaystyle\int_{\mathbb{T}^d} \frac{1}{\lambda(\boldsymbol{y})} d\boldsymbol{y}}{\displaystyle\int_{\mathbb{T}^d} \lambda(\boldsymbol{y}) d\boldsymbol{y}} \ . \end{aligned} \tag{5.9}$$

So we have

Theorem 5.1. *As $\varepsilon \downarrow 0$, the process $\boldsymbol{q}_t^\varepsilon$ converges weakly to a process $\boldsymbol{q}_t$, $\boldsymbol{q}_0 = \boldsymbol{q} \in \mathbb{R}^d$ governed by the operator (5.7) with coefficients given by (5.8) and (5.9).*

This Theorem implies a homogenization result for the process $\boldsymbol{q}_t^{\mu,\delta,\varepsilon}$ defined by equation (5.1).

2. SMALL MASS ASYMPTOTIC FOR THE MOTION WITH VANISHING FRICTION.

2.1 Introduction

The Langevin equation

$$\mu \ddot{\boldsymbol{q}}_t^\mu = \boldsymbol{b}(\boldsymbol{q}_t^\mu) - \lambda \dot{\boldsymbol{q}}_t^\mu + \sigma(\boldsymbol{q}_t^\mu)\dot{\boldsymbol{W}}_t \ , \ \boldsymbol{q}_0^\mu = \boldsymbol{q} \in \mathbb{R}^n \ , \dot{\boldsymbol{q}}_0^\mu = \boldsymbol{p} \in \mathbb{R}^n \ , \tag{1.1}$$

describes the motion of a particle of mass μ in a force field $\boldsymbol{b}(\boldsymbol{q})$, $\boldsymbol{q} \in \mathbb{R}^n$, subjected to random fluctuations and to a friction proportional to the velocity. Here $\boldsymbol{W}_t$ is the standard Wiener process in $\mathbb{R}^n$, $\lambda > 0$ is the friction coefficient. The vector field $\boldsymbol{b}(\boldsymbol{q})$ and the matrix function $\sigma(\boldsymbol{q})$ are assumed to be continuously differentiable and bounded together with their first derivatives. The matrix $a(\boldsymbol{q}) = (a_{ij}(\boldsymbol{q})) = \sigma(\boldsymbol{q})\sigma^*(\boldsymbol{q})$ is assumed to be non-degenerate.

It is assumed usually that the friction coefficient λ is a positive constant. Under this assumption, one can prove that $\boldsymbol{q}_t^\mu$ converges in probability as $\mu \downarrow 0$ uniformly on each finite time interval $[0, T]$ to an n-dimensional diffusion process $\boldsymbol{q}_t$: for any $\kappa, T > 0$ and any $\boldsymbol{p}_0^\mu = \boldsymbol{p} \in \mathbb{R}^n$, $\boldsymbol{q}_0^\mu = \boldsymbol{q} \in \mathbb{R}^n$ fixed,

$$\lim_{\mu\downarrow 0} \mathbf{P}\left(\max_{0\le t\le T} |\boldsymbol{q}_t^\mu - \boldsymbol{q}_t|_{\mathbb{R}^d} > \kappa \right) = 0 \ .$$

Here $\boldsymbol{q}_t$ is the solution of equation

$$\dot{\boldsymbol{q}}_t = \frac{1}{\lambda}\boldsymbol{b}(\boldsymbol{q}_t) + \frac{1}{\lambda}\sigma(\boldsymbol{q}_t)\dot{\boldsymbol{W}}_t \ , \ \boldsymbol{q}_0 = \boldsymbol{q}_0^\mu = \boldsymbol{q} \in \mathbb{R}^n \ . \tag{1.2}$$

The stochastic term in (1.2) should be understood in the Itô sense.

The approximation of $\boldsymbol{q}_t^\mu$ by $\boldsymbol{q}_t$ for $0 < \mu << 1$ is called the Smoluchowski-Kramers approximation. This is the main justification for replacement of the second order equation (1.1) by the first order equation (1.2). The price for such a simplification, in particular, consists of certain non-universality of equation (1.2): The white noise in (1.1) is an idealization of a more regular stochastic process $\dot{\boldsymbol{W}}_t^\delta$ with correlation radius $\delta << 1$ converging to $\dot{\boldsymbol{W}}_t$ as $\delta \downarrow 0$. Let $\boldsymbol{q}_t^{\mu,\delta}$ be the solution of equation (1.1) with $\dot{\boldsymbol{W}}_t$ replaced by $\dot{\boldsymbol{W}}_t^\delta$. Then limit of $\boldsymbol{q}_t^{\mu,\delta}$ as $\mu, \delta \downarrow 0$ depends on the relation between μ and δ. Say, if first $\delta \downarrow 0$ and then $\mu \downarrow 0$, the stochastic integral in (1.2) should be understood in the Itô sense; if first $\mu \downarrow 0$ and then $\delta \downarrow 0$, $\boldsymbol{q}_t^{\mu,\delta}$ converges to the solution of (1.2) with stochastic integral in the Stratonovich sense. (See, for instance, [8].)

We considered in [12] the case of a variable friction coefficient $\lambda = \lambda(\boldsymbol{q})$. We assumed in that work that $\lambda(\boldsymbol{q})$ is smooth and $0 < \lambda_0 \le \lambda(\boldsymbol{q}) \le \Lambda < \infty$. It turns out that in this case the solution $\boldsymbol{q}_t^\mu$ of (1.1) does not converge, in general, to the solution of (1.2) with $\lambda = \lambda(\boldsymbol{q})$, so that the Smoluchowski-Kramers approximation should be modified. In order to do this, we considered in [12] equation (1.1) with $\dot{\boldsymbol{W}}_t$ replaced by $\dot{\boldsymbol{W}}_t^\delta$ described above:

$$\mu\ddot{\boldsymbol{q}}_t^{\mu,\delta}=\boldsymbol{b}(\boldsymbol{q}_t^{\mu,\delta})-\lambda(\boldsymbol{q}_t^{\mu,\delta})\dot{\boldsymbol{q}}_t^{\mu,\delta}+\sigma(\boldsymbol{q}_t^{\mu,\delta})\dot{\boldsymbol{W}}_t^{\delta}\ ,\ \boldsymbol{q}_0^{\mu,\delta}=\boldsymbol{q}\ ,\ \dot{\boldsymbol{q}}_0^{\mu,\delta}=\boldsymbol{p}\ . \tag{1.3}$$

It was proved in [12] that after such a regularization, the solution of (1.3) has a limit $\boldsymbol{q}_t^{\delta}$ as $\mu\downarrow 0$, and $\boldsymbol{q}_t^{\delta}$ is the unique solution of the equation obtained from (1.3) as $\mu=0$:

$$\dot{\boldsymbol{q}}_t^{\delta}=\frac{1}{\lambda(\boldsymbol{q}_t^{\delta})}\boldsymbol{b}(\boldsymbol{q}_t^{\delta})+\frac{1}{\lambda(\boldsymbol{q}_t^{\delta})}\sigma(\boldsymbol{q}_t^{\delta})\dot{\boldsymbol{W}}_t^{\delta}\ ,\ \boldsymbol{q}_0^{\delta}=\boldsymbol{q}\ . \tag{1.4}$$

Now we can take $\delta\downarrow 0$ in (1.4). As the result we get the equation

$$\dot{\boldsymbol{q}}_t=\frac{1}{\lambda(\boldsymbol{q}_t)}\boldsymbol{b}(\boldsymbol{q}_t)+\frac{1}{\lambda(\boldsymbol{q}_t)}\sigma(\boldsymbol{q}_t)\circ\dot{\boldsymbol{W}}_t\ ,\ \boldsymbol{q}_0=\boldsymbol{q}\ , \tag{1.5}$$

where the stochastic term should be understood in the Stratonovich sense. We have, for any $\delta,\kappa,T>0$ fixed and any $\boldsymbol{p}_0^{\mu,\delta}=\boldsymbol{p}$ fixed, that

$$\lim_{\mu\downarrow 0}\mathbf{P}\left(\max_{0\le t\le T}|\boldsymbol{q}_t^{\mu,\delta}-\boldsymbol{q}_t^{\delta}|_{\mathbb{R}^d}>\kappa\right)=0\ ,$$

and we have

$$\lim_{\delta\to 0}\mathbf{E}\max_{t\in[0,T]}|\boldsymbol{q}_t^{\delta}-\boldsymbol{q}_t|_{\mathbb{R}^d}=0\ .$$

So the regularization leads to a modified Smoluchowski-Kramers equation (1.5).

In this chapter we study a further generalization of the problem considered in [12]. Keeping the assumptions on uniform boundedness and smoothness of $\lambda(\bullet)$, we drop the assumption that $0<\lambda_0\le\lambda(\boldsymbol{q})$ and instead assume that $\lambda(\boldsymbol{q})=0$ for $\boldsymbol{q}\in[G]\subset\mathbb{R}^n$ and $\lambda(\boldsymbol{q})>0$ for $\boldsymbol{q}\in\mathbb{R}^n\backslash[G]$. Here G is a domain in $\mathbb{R}^n$ and $[G]$ its closure in the standard Euclidean metric. For simplicity of presentation we assume in the rest of this paper that $\sigma(\bullet)$ is the identity matrix. (In Section 3 we further assume that $\boldsymbol{b}(\bullet)=\mathbf{0}$.) In order to use the results of [12] we introduce a further regularization of problem (1.5). We consider the problem

$$\dot{\boldsymbol{q}}_t^{\varepsilon}=\frac{1}{\lambda(\boldsymbol{q}_t^{\varepsilon})+\varepsilon}\boldsymbol{b}(\boldsymbol{q}_t^{\varepsilon})+\frac{1}{\lambda(\boldsymbol{q}_t^{\varepsilon})+\varepsilon}\circ\dot{\boldsymbol{W}}_t\ ,\ \boldsymbol{q}_0^{\varepsilon}=\boldsymbol{q}\ ,\ \varepsilon>0 \tag{1.6}$$

and we study the limit of $\boldsymbol{q}_t^{\varepsilon}$ as $\varepsilon\downarrow 0$. This limiting process can be regarded as a limiting process of the system

$$\mu\ddot{\boldsymbol{q}}_t^{\mu,\delta,\varepsilon}=\boldsymbol{b}(\boldsymbol{q}_t^{\mu,\delta,\varepsilon})-(\lambda(\boldsymbol{q}_t^{\mu,\delta,\varepsilon})+\varepsilon)\dot{\boldsymbol{q}}_t^{\mu,\delta,\varepsilon}+\dot{\boldsymbol{W}}_t^{\delta}\ ,\ \boldsymbol{q}_0^{\mu,\delta,\varepsilon}=\boldsymbol{q}\ ,\ \dot{\boldsymbol{q}}_0^{\mu,\delta,\varepsilon}=\boldsymbol{p} \tag{1.7}$$

as first $\mu\downarrow 0$ then $\delta\downarrow 0$ and then $\varepsilon\downarrow 0$.

System (1.6), in Itô's form, can be written as follows:

$$\dot{\boldsymbol{q}}_t^{\varepsilon}=\frac{1}{\lambda(\boldsymbol{q}_t^{\varepsilon})+\varepsilon}\boldsymbol{b}(\boldsymbol{q}_t^{\varepsilon})-\frac{\nabla\lambda(\boldsymbol{q}_t^{\varepsilon})}{2(\lambda(\boldsymbol{q}_t^{\varepsilon})+\varepsilon)^3}+\frac{1}{\lambda(\boldsymbol{q}_t^{\varepsilon})+\varepsilon}\dot{\boldsymbol{W}}_t\ ,\ \boldsymbol{q}_0^{\varepsilon}=\boldsymbol{q}\ . \tag{1.8}$$

However, as will be shown later, for non-compact region $[G]$, it is sometimes more convenient to consider the projection of the above system onto another space $\mathfrak{X}$. (In particular, in Section 3 the space $\mathfrak{X}$ is a cylinder $\mathfrak{X} = S^1 \times [a-1, b+1]$ for $a < 0, b > 0$.) Let us work with system (1.8) on $\mathfrak{X}$ and compact region $[G]$. It turns out that, in the limit, to get a Markov process with continuous trajectories, one has to glue all the points of $[G]$ and form a projected space $\mathfrak{C}$. Let the projection map be $\boldsymbol{\pi} : \mathfrak{X} \to \mathfrak{C}$. We will prove, for the 1-dimensional case (Section 2) and a multidimensional model problem (Section 3), that the processes $\widetilde{\boldsymbol{q}}_t^\varepsilon = \boldsymbol{\pi}(\boldsymbol{q}_t^\varepsilon)$ converge weakly as $\varepsilon \downarrow 0$ to a continuous strong Markov process $\widetilde{\boldsymbol{q}}_t$ on $\mathfrak{C}$. We will characterize the generator of this Markov process and specify its boundary condition. In particular, we will show that as $\varepsilon > 0$ is very small, certain mixing within $[G]$ is likely to happen for the process $\boldsymbol{q}_t^\varepsilon$. This mixing is the key mechanism that leads to our special boundary condition. We expect that (see Section 4), within the region that the friction is vanishing, similar mixing phenomenon will happen for the general multidimensional case.

It is worth mentioning here that some related problems are considered in [27], [28], [30] and [31]. It is also interesting to note that the limiting process for our two dimensional model problem (see Section 3) shares some common feature with the so called Walsh's Brownian motion (see, for example [1]).

However, at this stage we are not able to prove, in the most general multidimensional case (except for the 2-d model problem in Section 3), the convergence of $\widetilde{\boldsymbol{q}}_t^\varepsilon = \boldsymbol{\pi}(\boldsymbol{q}_t^\varepsilon)$ in (1.8) to some Markov process $\widetilde{\boldsymbol{q}}_t$. We will formulate a conjecture about this in Section 4.

2.2 One dimensional case

Let us consider in this section the 1-dimensional case. Besides the usual assumptions made in Section 1 we suppose that our friction $\lambda(\bullet)$ satisfies $\lambda(q) > 0$ for $q \in (-\infty, -1) \cup (1, \infty)$. Let $\lambda(q) = 0$ for $q \in [-1, 1]$. Equation (1.8) now takes the following form:

$$\dot{q}_t^\varepsilon = \frac{b(q_t^\varepsilon)}{\lambda(q_t^\varepsilon) + \varepsilon} - \frac{\lambda'(q_t^\varepsilon)}{2(\lambda(q_t^\varepsilon) + \varepsilon)^3} + \frac{1}{\lambda(q_t^\varepsilon) + \varepsilon} \dot{W}_t \ , \ q_0^\varepsilon = q_0 \in \mathbb{R} \ . \tag{2.1}$$

We suppose that $q_0 \in [a-1, b+1]$ for some $a < 0 < b$. The process q_t^ε is supposed to be stopped once it hits $q = a-1$ or $q = b+1$.

Our goal is to study the asymptotic behavior of (2.1) as $\varepsilon \downarrow 0$. To this end we shall write the process (2.1) as a strong Markov process subject to a generalized second order differential operator in the form $D_{v^\varepsilon} D_{u^\varepsilon}$ (see [7], [5], [26]). We have

$$u^\varepsilon(q) = \int_0^q (\lambda(x) + \varepsilon) \exp\left(-2\int_0^x b(y)(\lambda(y) + \varepsilon) dy\right) dx \ , \tag{2.2}$$

$$v^{\varepsilon}(q) = 2\int_0^q (\lambda(x)+\varepsilon)\exp\left(2\int_0^x b(y)(\lambda(y)+\varepsilon)dy\right)dx\ . \tag{2.3}$$

For fixed $\varepsilon > 0$, the functions u^{ε} and v^{ε} are strictly increasing functions in their arguments. As $\varepsilon \downarrow 0$, they will converge uniformly on finite intervals to the functions u and v defined by

$$u(q) = \int_0^q \lambda(x)\exp\left(-2\int_0^x b(y)\lambda(y)dy\right)dx\ , \tag{2.4}$$

$$v(q) = 2\int_0^q \lambda(x)\exp\left(2\int_0^x b(y)\lambda(y)dy\right)dx\ . \tag{2.5}$$

The functions u and v are strictly increasing outside the interval $[-1,1]$ and have constant stretches on $[-1,1]$.

Consider a projection map π: we let $\pi([-1,1]) = 0$ and $\pi(q) = q+1$ for $q < -1$ and $\pi(q) = q-1$ for $q > 1$. Consider the process $\widetilde{q}_t^{\varepsilon} = \pi(q_t^{\varepsilon})$. Process $\widetilde{q}_t^{\varepsilon}$ for fixed $\varepsilon > 0$, in general, is *not* a Markov process.

Let us define two functions $\widetilde{u}$ and $\widetilde{v}$ as follows: $\widetilde{u}(\widetilde{q}) = u(\widetilde{q}-1)$ for $\widetilde{q} < 0$ and $\widetilde{u}(\widetilde{q}) = u(\widetilde{q}+1)$ for $\widetilde{q} > 0$ and $\widetilde{u}(0) = u(1) = u(-1) = 0$; $\widetilde{v}(\widetilde{q}) = v(\widetilde{q}-1)$ for $\widetilde{q} < 0$ and $\widetilde{v}(\widetilde{q}) = v(\widetilde{q}+1)$ for $\widetilde{q} > 0$ and $\widetilde{v}(0) = v(1) = v(-1) = 0$. Here the functions u and v are defined in (2.4), (2.5). The functions $\widetilde{u}$ and $\widetilde{v}$ are continuous strictly increasing functions on $[a,b]$.

Define a Markov process $\widetilde{q}_t$ on $[a,b]$ as follows. The generator A of $\widetilde{q}_t$ is $A = D_{\widetilde{v}}D_{\widetilde{u}}$. The domain of definition $D(A)$ of operator A consists of all functions f that are continuous on $[a,b]$, are twice continuously differentiable in $\widetilde{q} \in [a,b]\backslash\{0\}$, with finite limit $\lim\limits_{\widetilde{q}\to 0} Af(\widetilde{q})$ (taken as the value of $Af(0)$) and finite one-sided limits $\lim\limits_{\delta\downarrow 0}\dfrac{f(\delta)-f(0)}{\widetilde{u}(\delta)-\widetilde{u}(0)} \equiv D_{\widetilde{u}}^{+}f(0) = D_{\widetilde{u}}^{-}f(0) \equiv \lim\limits_{\delta\downarrow 0}\dfrac{f(0)-f(-\delta)}{\widetilde{u}(0)-\widetilde{u}(-\delta)}$. Also we have $\lim\limits_{\widetilde{q}\to a} Af(\widetilde{q}) = \lim\limits_{\widetilde{q}\to b} Af(\widetilde{q}) = 0$ (taken as the value of $Af(a)$ and $Af(b)$).

Lemma 2.1. *There exists the Markov process $\widetilde{q}_t$ on $[a,b]$.*

Proof. The existence of such a process could be checked similarly as in [18, Section 2]. For the sake of completeness and comparison with results in the next section we shall check it here. To this end we use an equivalent formulation of the Hille-Yosida theorem (see [18, Section 2] also [32, Theorem 2]). We check three conditions.

• The domain $D(A)$ is dense in the space $\mathbf{C}([a,b])$. This is because we can approximate every continuous function f with one that is constant in a neighborhood of 0.

After that in the interior part of the intervals $[a, 0)$ and $(0, b]$, at a positive distance from 0, with a smooth function. The approximating smooth function satisfy our boundary conditions since $Af(0) = D^+_{\widetilde{u}} f(0) = D^-_{\widetilde{u}} f(0) = 0$.

• The maximum principle: if $f \in D(A)$ and the function f reaches its maximum at a point $x_0 \in [a, b]$, then $Af(x_0) \le 0$. If $x_0 \ne 0$ we have $f'(x_0) = 0$ and $f''(x_0) \le 0$ and

$$D_{\widetilde{v}} D_{\widetilde{u}} f(x_0) = \frac{f''(x_0)}{\widetilde{v}'(x_0)\widetilde{u}'(x_0)} - \frac{\widetilde{u}''(x_0)}{\widetilde{v}'(x_0)(\widetilde{u}'(x_0))^2} f'(x_0) \le 0 \ .$$

If the maximum is achieved at 0, we consider the expansion

$$f(x) = f(0) + D_{\widetilde{u}} f(0)(\widetilde{u}(x) - \widetilde{u}(0)) + (Af(0) + o(1)) \int_0^x (\widetilde{v}(y) - \widetilde{v}(0)) d\widetilde{u}(y) \ .$$

The last integral is $O(\widetilde{u}(x)\widetilde{v}(x))$ as $x \to 0$. Since $D^-_{\widetilde{u}} f(0) \ge 0$ and $D^+_{\widetilde{u}} f(0) \le 0$, by our boundary conditions at 0 we get $D_{\widetilde{u}} f(0) = 0$. This implies that $Af(0) \le 0$.

• Existence of solution $f \in D(A)$ of $\lambda f - Af = F$ for all $F \in \mathbf{C}([a, b])$. On each of the intervals $[a, 0)$ and $(0, b]$ the general solution of equation $\lambda f - D_{\widetilde{v}} D_{\widetilde{u}} f = F$, $F \in \mathbf{C}([a, b])$ can be written as

$$f^{\pm}(q) = \widehat{f}^{\pm}(q) + G^{\pm}(q) \ .$$

Here $\widehat{f}^{\pm}(q)$ satisfy the equation $\lambda \widehat{f}^{\pm} - D_{\widetilde{v}} D_{\widetilde{u}} \widehat{f}^{\pm} = F$, $\widehat{f}^{+}(0+) = 0$ (or $\widehat{f}^{-}(0-) = 0$), $D^+_{\widetilde{u}} \widehat{f}^{+}(0) = 0$ (or $D^-_{\widetilde{u}} \widehat{f}^{-}(0) = 0$) and $G^{\pm}(q)$ satisfy the equation $\lambda G^{\pm} - D_{\widetilde{v}} D_{\widetilde{u}} G^{\pm} = 0$, $G^{+}(0+) = k_1^{+}$ (or $G^{-}(0-) = k_1^{-}$), $D^+_{\widetilde{u}} G^{+}(0) = k_2^{+}$ (or $D^-_{\widetilde{u}} G^{-}(0) = k_2^{-}$). Here $k_1^{\pm}$ and $k_2^{\pm}$ are constants. Our boundary condition gives $k_1^{+} = k_1^{-}$ and $k_2^{+} = k_2^{-}$. The boundary condition $D_{\widetilde{u}} D_{\widetilde{v}} f^{+}(a) = D_{\widetilde{u}} D_{\widetilde{v}} f^{-}(b) = 0$ singles out a unique $f \in D(A)$. □

We have

Theorem 2.1. *As $\varepsilon \downarrow 0$, for fixed $T > 0$, the process $\widetilde{q}_t^{\varepsilon}$ converges weakly in the space $\mathbf{C}_{[0,T]}([a, b])$ to the process $\widetilde{q}_t$.*

The proof of this Theorem is based on an application of the machinery developed in [14, Ch.8], [17] and [18]. We shall use the following lemma, which is the Lemma 3.1 of [14, Ch.8, page 301]. We formulate it here in the terminology that meets our purpose.

Lemma 2.2. *Let M be a metric space; Y, a continuous mapping $M \mapsto Y(M)$, $Y(M)$ being a complete separable metric space. Let $(X_t^{\varepsilon}, \mathbf{P}_x^{\varepsilon})$ be a family of Markov processes in M; suppose that the process $Y(X_t^{\varepsilon})$ has continuous trajectories. Let $(y_t, \mathbf{P}_y)$ be a Markov process with continuous paths in $Y(M)$ whose infinitesimal operator is A*

with domain of definition $D(A)$. Let $T > 0$. Let us suppose that the space $\mathbf{C}_{[0,T]}(Y(M))$ of continuous functions on $[0,T]$ with values in $Y(M)$ is taken as the sample space, so that the distribution of the process in the space of continuous functions is simply $\mathbf{P}_y$. Let Ψ be a subset of the space $\mathbf{C}_{[0,\infty)}(Y(M))$ such that for measures μ_1, μ_2 on $Y(M)$ the equality $\int F d\mu_1 = \int F d\mu_2$ for all $F \in \Psi$ implies $\mu_1 = \mu_2$. Let D be the subset of $D(A)$ such that for every $F \in \Psi$ and $\lambda > 0$ the equation $\lambda f - Af = F$ has a solution $f \in D$.

Suppose that for every $x \in M$ the family of distributions $\mathbf{Q}_x^\varepsilon$ of $Y(X_\bullet^\varepsilon)$ in the space $\mathbf{C}_{[0,T]}(Y(M))$ corresponding to the probabilities of $\mathbf{P}_x^\varepsilon$ is weakly pre-compact; and that for every compact $K \subset Y(M)$, for every $f \in D$ and every $\lambda > 0$,

$$\mathbf{E}_x^\varepsilon \int_0^\infty e^{-\lambda t}[\lambda f(Y(X_t^\varepsilon)) - Af(Y(X_t^\varepsilon))]dt \to f(Y(x))$$

as $\varepsilon \downarrow 0$ uniformly in $x \in Y^{-1}(K)$.

Then $\mathbf{Q}_x^\varepsilon$ converges weakly as $\varepsilon \downarrow 0$ to the probability measure $\mathbf{P}_{Y(x)}$.

Proof of Theorem 2.1. Making use of Lemma 2.2, we take the metric space $M = [a-1, b+1]$ and the mapping $Y = \pi$. The space $Y(M) = \pi([a-1, b+1]) = [a, b]$. We take the process q_t^ε as $(X_t^\varepsilon, \mathbf{P}_x^\varepsilon)$. We take the process $\widetilde{q}_t$ as $(y_t, \mathbf{P}_y)$.

Let Ψ be the space of all continuous bounded functions in $[a, b]$ which are once continuously differentiable inside $[a, 0)$ and $(0, b]$, with bounded derivatives. The space $D \subset D(A)$ consists of those functions $f \in D(A)$ such that they are continuous and bounded in $[a, b]$ and are three times continuously differentiable inside $[a, 0)$ and $(0, b]$, with bounded derivatives up to the third order.

Pre-compactness of the family of distributions of the process $\{\widetilde{q}_\bullet^\varepsilon\}_{\varepsilon > 0}$ is checked in Lemma 2.4. What remains to do is to check that for every compact $K \subset [a, b]$, for every $f \in D$ and every $\lambda > 0$,

$$\mathbf{E}_{q_0}\left[\int_0^\infty e^{-\lambda t}[\lambda f(\pi(q_t^\varepsilon)) - Af(\pi(q_t^\varepsilon))]dt - f(\pi(q_0))\right] \to 0$$

as $\varepsilon \downarrow 0$ uniformly in $q_0 \in \pi^{-1}(K)$. This is done in Lemma 2.5. This finishes the proof of Theorem 2.1. □

For positive δ small enough, let $G(\delta) = [a-1, -1-\delta] \cup [1+\delta, b+1]$. Let $0 < \delta' < \delta$. Let $C(\delta') = \{-1-\delta', 1+\delta'\}$. We introduce a sequence of stopping times $\tau_0 \le \sigma_0 < \tau_1 < \sigma_1 < \tau_2 < \sigma_2 < ...$ by

$$\tau_0 = 0 \ , \ \sigma_n = \min\{t \ge \tau_n, q_t^\varepsilon \in G(\delta)\} \ , \ \tau_n = \min\{t > \sigma_{n-1} : q_t^\varepsilon \in C(\delta')\} \ .$$

This is well-defined up to some σ_k $(k \geq 0)$ such that

$$\mathbf{P}_{q^\varepsilon_{\sigma_k}}(q^\varepsilon_{t+\sigma_k} \text{ hits } a-1 \text{ or } b+1 \text{ before it hits } -1-\delta' \text{ or } 1+\delta') = 1 \ .$$

We will then define $\tau_{k+1} = \min\{t > \sigma_k : q^\varepsilon_t = a-1 \text{ or } b+1\}$. And we define $\tau_{k+1} < \sigma_{k+1} = \tau_{k+1}+1 < \tau_{k+2} = \tau_{k+1}+2 < \sigma_{k+2} = \tau_{k+1}+3 < ...$ and so on.

We have $\lim\limits_{n\to\infty} \tau_n = \lim\limits_{n\to\infty} \sigma_n = \infty$. And we have obvious relations $q^\varepsilon_{\tau_n} \in C(\delta')$, $q^\varepsilon_{\sigma_n} \in C(\delta)$ for $1 \leq n \leq k$ (as long as $k \geq 1$, if $k=0$ the process may start from $G(\delta)$ and goes directly to $a-1$ or $b+1$ without touching $C(\delta')$ and is stopped there, or it may start from $(-1-\delta, 1+\delta)$, reaches $\{-1-\delta, 1+\delta\}$ first and then goes directly to $a-1$ or $b+1$ without touching $C(\delta')$ and is stopped there). Also, for $n \geq k+1$ we have $q^\varepsilon_{\tau_n} = q^\varepsilon_{\sigma_n} = a-1$ or $b+1$. If $q^\varepsilon_0 = q_0 \in G(\delta)$, then we have $\sigma_0 = 0$ and τ_1 is the first time at which the process q^ε_t reaches $C(\delta')$ or $\{a-1, b+1\}$.

Now we check weak pre-compactness of the family of distributions of the processes $\{\widetilde{q}^\varepsilon_t\}_{\varepsilon>0}$. To this end we need the following lemma, which is Lemma 5.1 in [18]. We formulate it using our terminology.

Lemma 2.3. *Let $\widetilde{q}^{\varepsilon,\delta}_\bullet$ for every $\varepsilon > 0$, $\delta > 0$, be a random element in $\mathbf{C}_{[0,T]}([a,b])$ such that $\max\limits_{0\leq t\leq T} |\widetilde{q}^\varepsilon_t - \widetilde{q}^{\varepsilon,\delta}_t| \leq \delta$ on the whole probability space. If for every positive δ the family of distributions of $\widetilde{q}^{\varepsilon,\delta}_\bullet$, $\varepsilon > 0$, is tight, then the family of distributions of $\widetilde{q}^\varepsilon_\bullet$ is pre-compact.*

Now we have

Lemma 2.4. *The family of distributions of $\{\widetilde{q}^\varepsilon_\bullet\}_{\varepsilon>0}$ is pre-compact.*

Proof. Let $\delta' = \delta/2$ so that we need only one parameter δ. Between the times σ_{i-1} and τ_i the process q^ε_t is either in $[a, -1-\delta/2)$ or in $(1+\delta/2, b]$, and for $\sigma_{i-1} \leq t < t' < \tau_i$ we have $|\widetilde{q}^\varepsilon_t - \widetilde{q}^\varepsilon_{t'}| = |q^\varepsilon_t - q^\varepsilon_{t'}|$. Since we have

$$q^\varepsilon_t - q^\varepsilon_{t'} = \int_t^{t'} \left[\frac{b(q^\varepsilon_s)}{\lambda(q^\varepsilon_s)+\varepsilon} - \frac{\lambda'(q^\varepsilon_s)}{2(\lambda(q^\varepsilon_t)+\varepsilon)^3}\right] ds + \int_t^{t'} \frac{1}{\lambda(q^\varepsilon_s)+\varepsilon} dW_s \ ,$$

we can estimate

$$\mathbf{E}|q^\varepsilon_t - q^\varepsilon_{t'}|^4 \leq K(\delta)|t-t'|^2 \ .$$

The constant $K(\delta)$ is independent of ε provided that ε is small. Now we let

$$Z^{\varepsilon,\delta}_t = \int_0^t \mathbf{1}_{G(\delta/2)}(q^\varepsilon_s) \left[\frac{b(q^\varepsilon_s)}{\lambda(q^\varepsilon_s)+\varepsilon} - \frac{\lambda'(q^\varepsilon_s)}{2(\lambda(q^\varepsilon_t)+\varepsilon)^3}\right] ds + \int_0^t \mathbf{1}_{G(\delta/2)}(q^\varepsilon_s) \frac{1}{\lambda(q^\varepsilon_s)+\varepsilon} dW_s \ .$$

From the above estimate we see that $Z_t^{\varepsilon,\delta}$ for fixed $\delta > 0$ is tight. The trajectories of these stochastic processes satisfy the Hölder condition $|Z_t^{\varepsilon,\delta} - Z_{t'}^{\varepsilon,\delta}| \le H^{\varepsilon,\delta}|t-t'|^{1/5}$ where $H^{\varepsilon,\delta}$ are random variables with $\mathbf{E}(H^{\varepsilon,\delta})^4$ bounded by the same $K(\delta)$.

For $i \ge 1$ if $q_{\tau_i}^\varepsilon \in C(\delta/2)$ and $q_{\sigma_i}^\varepsilon \in C(\delta)$ then between the times τ_i and σ_i $(\le T)$ the process q_t^ε travels a distance at least $\delta/2$ and at least this distance in $G(\delta/2)$ on the same interval either $[a, -1-\delta/2)$ or $(1+\delta/2, b]$. By our estimate on Hölder continuity of $Z_t^{\varepsilon,\delta}$ this implies that $\sigma_i - \tau_i \ge \left(\dfrac{\delta}{4H^{\varepsilon,\delta}}\right)^5$, $i \ge 1$. If $q_{\tau_i}^\varepsilon \in \{a-1, b+1\}$ then by our definition of the stopping time $\sigma_i = \tau_i + 1$ we can choose δ small enough such that the above inequality also holds.

Now we shall define the process $\widetilde{q}_t^{\varepsilon,\delta}$ as follows.

• For $\sigma_{i-1} \le t \le \tau_i$ we take $\widetilde{q}_t^{\varepsilon,\delta} = \widetilde{q}_t^\varepsilon$.

• For $\tau_0 \le t \le \sigma_0$ we take $\widetilde{q}_t^{\varepsilon,\delta} = \widetilde{q}_{\sigma_0}^\varepsilon$. This gives $\max\limits_{\tau_0 \le t \le \sigma_0} |\widetilde{q}_t^{\varepsilon,\delta} - \widetilde{q}_t^\varepsilon| = \max\limits_{\tau_0 \le t \le \sigma_0} |\widetilde{q}_{\sigma_0}^\varepsilon - \widetilde{q}_t^\varepsilon| \le \delta$.

• If $\tau_i < T < \sigma_i$ we take $\widetilde{q}_t^{\varepsilon,\delta} = \widetilde{q}_{\tau_i}^\varepsilon$ for $\tau_i \le t \le T$. This gives $\max\limits_{\tau_i \le t \le T} |\widetilde{q}_t^{\varepsilon,\delta} - \widetilde{q}_t^\varepsilon| = \max\limits_{\tau_i \le t \le T} |\widetilde{q}_{\tau_i}^\varepsilon - \widetilde{q}_t^\varepsilon| \le \delta/2$.

• If $\sigma_i \le T$. In this case if $\widetilde{q}_{\tau_i}^\varepsilon$ and $\widetilde{q}_{\sigma_i}^\varepsilon$ are within a distance $\le \delta$ from 0, we define $\widetilde{q}^{\varepsilon,\delta}_{\frac{\tau_i+\sigma_i}{2}} = 0$,

$$\widetilde{q}_t^{\varepsilon,\delta} = \left(1 - \frac{2(t-\tau_i)}{\sigma_i - \tau_i}\right) \widetilde{q}_{\tau_i}^\varepsilon \text{ for } \tau_i \le t \le \frac{\tau_i + \sigma_i}{2} ,$$

$$\widetilde{q}_t^{\varepsilon,\delta} = -\left(1 - \frac{2(t-\tau_i)}{\sigma_i - \tau_i}\right) \widetilde{q}_{\sigma_i}^\varepsilon \text{ for } \frac{\tau_i + \sigma_i}{2} \le t \le \sigma_i .$$

Since this is just a linear interpolation it is clear that in this case we have $\max\limits_{\tau_i \le t \le \sigma_i} |\widetilde{q}_t^{\varepsilon,\delta} - \widetilde{q}_t^\varepsilon| \le 2\delta$. Within this time interval $\tau_i \le t < t' \le \sigma_i$, $i \ge 1$ we have

$$|\widetilde{q}_t^{\varepsilon,\delta} - \widetilde{q}_{t'}^{\varepsilon,\delta}| \le \frac{\delta}{|\sigma_i - \tau_i|}|t-t'| \le \frac{\delta}{(\min\limits_{i\ge 1} \frac{1}{2}|\sigma_i - \tau_i|)^{1/5}}|t-t'|^{1/5} \le 2^{11/5} H^{\varepsilon,\delta}|t-t'|^{1/5} .$$

Another possibility is that $q_{\sigma_i}^\varepsilon = q_{\tau_i}^\varepsilon = a-1$ or $b+1$. In this case we define $\widetilde{q}_t^{\varepsilon,\delta} = \widetilde{q}_t^\varepsilon$ for $\tau_i \le t < \sigma_i$.

On the whole interval $0 \le t < t' \le T$ we have $|\widetilde{q}_t^{\varepsilon,\delta} - \widetilde{q}_{t'}^{\varepsilon,\delta}| \le (2^{11/5}+2)H^{\varepsilon,\delta}|t'-t|^{1/5}$ for $|t'-t| \le \left(\dfrac{\delta}{4H^{\varepsilon,\delta}}\right)^5$. This means that for fixed $\delta > 0$ we have the tightness of the family of distributions of $\widetilde{q}_t^{\varepsilon,\delta}$ in the space $\mathbf{C}_{[0,T]}([a,b])$. Since we checked $\max\limits_{0\le t\le T} |\widetilde{q}_t^{\varepsilon,\delta} - \widetilde{q}_t^\varepsilon| \le 2\delta$, by using Lemma 2.3 with 2δ instead of δ we get the pre-compactness of the family of distributions of $\widetilde{q}_t^\varepsilon$ in $\mathbf{C}_{[0,T]}([a,b])$. $\square$

The proof of the next Lemma 2.5 is based on Lemmas 2.6-2.10. Within the proof of this lemma and the auxiliary Lemmas 2.6-2.10, we will take $\varepsilon \downarrow 0$, $\delta = \delta(\varepsilon) \downarrow 0$,

$\delta' = \delta'(\varepsilon) \downarrow 0$ in an asymptotic order such that $0 < \varepsilon << \delta' << \delta$. Although not very precise, but for simplicity of presentation we will just refer this choice of order as first $\varepsilon \downarrow 0$, then $\delta' \downarrow 0$ and then $\delta \downarrow 0$. It could be checked that such an order of taking limit does not alter the validity of the result.

Throughout the rest of this section and next section when we use symbols U, V, M_i, C_i, A_i, etc., they are referring to some positive constants. We will not point out this explicitly unless some special properties of the implied constants are stressed. Also we sometimes use the same letter for constants in different estimates.

Lemma 2.5. *For every compact $K \subset [a, b]$, for every $f \in D$ and every $\lambda > 0$,*

$$\mathbf{E}_{q_0}\left[\int_0^{\infty} e^{-\lambda t}[\lambda f(\pi(q_t^{\varepsilon})) - Af(\pi(q_t^{\varepsilon}))]dt - f(\pi(q_0))\right] \to 0$$

as $\varepsilon \downarrow 0$ uniformly in $q_0 \in \pi^{-1}(K)$.

Proof. The above expectation can be written as

$$\begin{aligned}
&\mathbf{E}_{q_0}\left[\sum_{n=0}^{\infty}\left[\int_{\tau_n}^{\sigma_n} e^{-\lambda t}[\lambda f(\pi(q_t^{\varepsilon})) - Af(\pi(q_t^{\varepsilon}))]dt + e^{-\lambda\sigma_n} f(\pi(q_{\sigma_n}^{\varepsilon})) - e^{-\lambda\tau_n} f(\pi(q_{\tau_n}^{\varepsilon}))\right] + \right.\\
&\left.\sum_{n=0}^{\infty}\left[\int_{\sigma_n}^{\tau_{n+1}} e^{-\lambda t}[\lambda f(\pi(q_t^{\varepsilon})) - Af(\pi(q_t^{\varepsilon}))]dt + e^{-\lambda\tau_{n+1}} f(\pi(q_{\tau_{n+1}}^{\varepsilon})) - e^{-\lambda\sigma_n} f(\pi(q_{\sigma_n}^{\varepsilon}))\right]\right]\\
&= \mathbf{E}_{q_0}\left[\sum_{n=0}^{\infty} e^{-\lambda\tau_n}\psi_1^{\varepsilon}(q_{\tau_n}^{\varepsilon}) + \sum_{n=0}^{\infty} e^{-\lambda\sigma_n}\psi_2^{\varepsilon}(q_{\sigma_n}^{\varepsilon})\right] ,
\end{aligned} \tag{2.6}$$

where

$$\psi_1^{\varepsilon}(q) = \mathbf{E}_q\left[\int_0^{\sigma_0} e^{-\lambda t}[\lambda f(\pi(q_t^{\varepsilon})) - Af(\pi(q_t^{\varepsilon}))]dt + e^{-\lambda\sigma_0} f(\pi(q_{\sigma_0}^{\varepsilon}))\right] - f(\pi(q)) \ , \tag{2.7}$$

$$\psi_2^{\varepsilon}(q) = \mathbf{E}_q\left[\int_{\sigma_0}^{\tau_1} e^{-\lambda t}[\lambda f(\pi(q_t^{\varepsilon})) - Af(\pi(q_t^{\varepsilon}))]dt + e^{-\lambda\tau_1} f(\pi(q_{\tau_1}^{\varepsilon}))\right] - f(\pi(q)) \ . \tag{2.8}$$

We used the strong Markov property of q_t^{ε}. Since for $n \geq k+1$ we have $\psi_1^{\varepsilon}(q_{\tau_n}^{\varepsilon}) = \psi_2^{\varepsilon}(q_{\sigma_n}^{\varepsilon}) = 0$ we can assume that the function ψ_2^{ε} is taken at a point on $G(\delta)\backslash\{a-1, b+1\}$ and the expectation is determined by the values of the process q_t^{ε} in one of the intervals either $(1+\delta', b+1]$ or $[a-1, -1-\delta')$. We will prove, in Lemma 2.6, that under our specified asymptotic order we can have $|\psi_2^{\varepsilon}(q)| \leq (\widetilde{u}(\delta) - \widetilde{u}(-\delta))^2$ as $\varepsilon \downarrow 0$.

We can assume that the function ψ_1^{ε} is taken at a point in $[-1-\delta', 1+\delta']$ (in the case when $n = 0$ and $q_0^{\varepsilon} \in G(\delta)$, we also have $\psi_1^{\varepsilon}(q_0) = 0$). We can write

$$\begin{aligned}
&\psi_1^\varepsilon(q)\\
&= \left(\mathbf{E}_q f(\pi(q^\varepsilon_{\sigma_0})) - f(\pi(q))\right) - \mathbf{E}_q(1-e^{-\lambda\sigma_0})f(\pi(q^\varepsilon_{\sigma_0})) + \mathbf{E}_q\int_0^{\sigma_0} e^{-\lambda t}[\lambda f(\pi(q_t^\varepsilon)) - Af(\pi(q_t^\varepsilon)))]dt\\
&= (I)^\varepsilon(q) + (II)^\varepsilon(q) + (III)^\varepsilon(q)\ .
\end{aligned} \tag{2.9}$$

We are going to prove, in Lemma 2.8, that for $q \in [-1-\delta', 1+\delta']$, for a function $f \in D$ we can have the estimate $|(I)^\varepsilon(q)| \le M_1(\widetilde{u}(\delta) - \widetilde{u}(-\delta))^2$.

In Lemma 2.9 we will show that $\mathbf{E}_q\sigma_0 \le M_1(\widetilde{u}(\delta) - \widetilde{u}(-\delta))(\widetilde{v}(\delta) - \widetilde{v}(-\delta))$ and $\mathbf{E}_q(1-e^{-\lambda\sigma_0}) \le M_1(\widetilde{u}(\delta) - \widetilde{u}(-\delta))(\widetilde{v}(\delta) - \widetilde{v}(-\delta))$ so that $|(II)^\varepsilon(q)| + |(III)^\varepsilon(q)| < M_1(\widetilde{u}(\delta) - \widetilde{u}(-\delta))(\widetilde{v}(\delta) - \widetilde{v}(-\delta))$ for $q \in [-1-\delta', 1+\delta']$.

These estimates show that

$$|\psi_1^\varepsilon(q)| < (\widetilde{u}(\delta) - \widetilde{u}(-\delta))^2 + M_1(\widetilde{u}(\delta) - \widetilde{u}(-\delta))(\widetilde{v}(\delta) - \widetilde{v}(-\delta))$$

for all $q \in [-1-\delta', 1+\delta']$.

As we only consider the arguments $q^\varepsilon_{\tau_n}$ of ψ_1^ε in (2.6) being in $[-1-\delta', 1+\delta']$ starting with $n=1$ (otherwise $\psi_1^\varepsilon = 0$), we have, by strong Markov property of q_t^ε, that

$$\begin{aligned}
&\left|\mathbf{E}_{q_0}\sum_{n=1}^\infty e^{-\lambda\tau_n}\psi_1^\varepsilon(q^\varepsilon_{\tau_n})\right|\\
&\le (\widetilde{u}(\delta) - \widetilde{u}(-\delta))^2 + M_1(\widetilde{u}(\delta) - \widetilde{u}(-\delta))(\widetilde{v}(\delta) - \widetilde{v}(-\delta))\sum_{n=1}^\infty \mathbf{E}_{q_0}e^{-\lambda\tau_n}\\
&\le (\widetilde{u}(\delta) - \widetilde{u}(-\delta))^2 + M_1(\widetilde{u}(\delta) - \widetilde{u}(-\delta))(\widetilde{v}(\delta) - \widetilde{v}(-\delta))\sum_{n=1}^\infty \left(\sup_{q\in G(\delta)} \mathbf{E}_q e^{-\lambda\tau_1}\right)^{n-1} .
\end{aligned}$$

We will show, in Lemma 2.10, that $\mathbf{E}_q e^{-\lambda\tau_1} < 1 - M_2\widetilde{u}(\delta)\wedge(-\widetilde{u}(-\delta))$ for all $q \in G(\delta)$. Since as $\delta \downarrow 0$ we have $0 < M_2 \le \left|\dfrac{\widetilde{u}(\delta)}{-\widetilde{u}(-\delta)}\right| \le M_3 < \infty$, we have

$$\begin{aligned}
&\left|\mathbf{E}_{q_0}\sum_{n=1}^\infty e^{-\lambda\tau_n}\psi_1^\varepsilon(q^\varepsilon_{\tau_n})\right|\\
&\le ((\widetilde{u}(\delta) - \widetilde{u}(-\delta))^2 + M_1(\widetilde{u}(\delta) - \widetilde{u}(-\delta))(\widetilde{v}(\delta) - \widetilde{v}(-\delta)))\frac{1}{M_2(\widetilde{u}(\delta))\wedge(-\widetilde{u}(-\delta))} \to 0
\end{aligned}$$

as $\delta \downarrow 0$. For $n=0$ the expectation $\mathbf{E}_{q_0}\psi_1^\varepsilon(q_0^\varepsilon)$ is small as ε is small.

For the second term in (2.6) we can estimate

$$\begin{aligned}
&\left|\sum_{n=0}^\infty \mathbf{E}_q e^{-\lambda\sigma_n}\psi_2^\varepsilon(q^\varepsilon_{\sigma_n})\right| \le \sum_{n=0}^\infty \mathbf{E}_q e^{-\lambda\sigma_n}|\psi_2^\varepsilon(q)| \le \sum_{n=0}^\infty \mathbf{E}_q e^{-\lambda\tau_n}|\psi_2^\varepsilon(q)|\\
&\le (1 + \frac{M_4}{(\widetilde{u}(\delta))\wedge(-\widetilde{u}(-\delta))})(\widetilde{u}(\delta) - \widetilde{u}(-\delta))^2
\end{aligned}$$

which converges to 0 as $\varepsilon \downarrow 0$. This proves this lemma. $\square$

Lemma 2.6. *We have, for $q \in G(\delta)$, as ε is small, that* $|\psi_2^\varepsilon(q)| \leq (\widetilde{u}(\delta) - \widetilde{u}(-\delta))^2$.

Proof. For the initial point $q \in G(\delta)$ and the time interval $0 \leq t \leq \tau_1$ the trajectory of q_t^ε is traveling in one of the intervals either $[1+\delta', 1+b]$ or $[a-1, -1-\delta']$. Without loss of generality let us assume that $q \in [1+\delta, 1+b]$ and we are traveling in the interval $[1+\delta', 1+b]$. Let $\widetilde{q} = \pi(q)$. Let $B(\widetilde{q}) = b(\widetilde{q}+1)$ and $\Lambda(\widetilde{q}) = \lambda(\widetilde{q}+1)$. Let us extend the function $\Lambda(\bullet)$ to the whole line $\mathbb{R}$. The extended function $\widehat{\Lambda}(\bullet)$ is smooth, bounded, with uniformly bounded derivatives and such that $\widehat{\Lambda}(x) \geq \min\limits_{q \in [1+\delta', 1+b]} \lambda(q)$, $\widehat{\Lambda}(x) = \lambda(1+x)$ for $x \in [\delta', b]$.

Let the process $\widehat{\widetilde{q}}_t^\varepsilon$ be subject to the stochastic differential equation

$$\dot{\widehat{\widetilde{q}}}_t^\varepsilon = \frac{B(\widehat{\widetilde{q}}_t^\varepsilon)}{\widehat{\Lambda}(\widehat{\widetilde{q}}_t^\varepsilon) + \varepsilon} - \frac{\widehat{\Lambda}'(\widehat{\widetilde{q}}_t^\varepsilon)}{2(\widehat{\Lambda}(\widehat{\widetilde{q}}_t^\varepsilon) + \varepsilon)^3} + \frac{1}{\widehat{\Lambda}(\widehat{\widetilde{q}}_t^\varepsilon) + \varepsilon} \dot{W}_t \ , \ \widehat{\widetilde{q}}_0^\varepsilon = \widetilde{q} \ , 0 \leq t < \infty \ .$$

We introduce a stochastic process $\widehat{\widetilde{q}}_t$, $\widehat{\widetilde{q}}_0 = \widetilde{q}$ with generator $\widehat{A}$, subject to the stochastic differential equation

$$\dot{\widehat{\widetilde{q}}}_t = \frac{B(\widehat{\widetilde{q}}_t)}{\widehat{\Lambda}(\widehat{\widetilde{q}}_t)} - \frac{\widehat{\Lambda}'(\widehat{\widetilde{q}}_t)}{2\widehat{\Lambda}^3(\widehat{\widetilde{q}}_t)} + \frac{1}{\widehat{\Lambda}(\widehat{\widetilde{q}}_t)} \dot{W}_t \ , \ \widehat{\widetilde{q}}_0 = \widetilde{q} \ , 0 \leq t < \infty \ .$$

Notice that the modified generator $\widehat{A}$ agrees with A before the process $\widetilde{q}_t^\varepsilon$ reaches $\widetilde{q}_{\tau_1}^\varepsilon$. And before the time τ_1 the process $\widehat{\widetilde{q}}_t^\varepsilon$ agrees with the process $\widetilde{q}_t^\varepsilon$. Therefore we have,

$$\psi_2^\varepsilon(q) = \mathbf{E}_{\widetilde{q}} \left[\int_0^{\tau_1} e^{-\lambda t} [\lambda f(\widehat{\widetilde{q}}_t^\varepsilon) - \widehat{A} f(\widehat{\widetilde{q}}_t^\varepsilon)] dt - e^{-\lambda \tau_1} f(\widehat{\widetilde{q}}_{\tau_1}^\varepsilon) \right] - f(\widetilde{q}) \ .$$

It is clear by Itô's formula that we have (also see, [16, Section 2]), for the stopping time τ_1,

$$\mathbf{E}_{\widetilde{q}} \left[\int_0^{\tau_1} e^{-\lambda t} [\lambda f(\widehat{\widetilde{q}}_t) - \widehat{A} f(\widehat{\widetilde{q}}_t)] dt - e^{-\lambda \tau_1} f(\widehat{\widetilde{q}}_{\tau_1}) \right] - f(\widetilde{q}) = 0 \ .$$

Notice that the function $f \in D \subset D(A)$ is three times continuously differentiable in $[\delta', b]$. This gives the estimate that for some positive $U, V > 0$ and $T = T(\varepsilon)$ we have

$$
\begin{aligned}
&|\psi_2^\varepsilon(q)| \\
&= \left|\mathbf{E}_{\widetilde{q}}\int_0^{\tau_1} e^{-\lambda t}[\lambda(f(\widehat{\widetilde{q}}_t^\varepsilon)-f(\widehat{\widetilde{q}}_t))-(\widehat{A}f(\widehat{\widetilde{q}}_t^\varepsilon)-\widehat{A}f(\widehat{\widetilde{q}}_t))]dt - e^{-\lambda\tau_1}(f(\widehat{\widetilde{q}}_{\tau_1}^\varepsilon)-f(\widehat{\widetilde{q}}_{\tau_1}))\right| \\
&\le \mathbf{E}_{\widetilde{q}}\left(\int_0^{T(\varepsilon)} \lambda e^{-\lambda t}dt\,(\mathrm{Lip}(f))\cdot|\widehat{\widetilde{q}}_t^\varepsilon-\widehat{\widetilde{q}}_t|+\right. \\
&\qquad\left.\int_0^{T(\varepsilon)} e^{-\lambda t}dt\,(\mathrm{Lip}(Af))\cdot|\widehat{\widetilde{q}}_t^\varepsilon-\widehat{\widetilde{q}}_t|+(\mathrm{Lip}(f))\cdot|\widehat{\widetilde{q}}_{\tau_1}^\varepsilon-\widehat{\widetilde{q}}_{\tau_1}|\mathbf{1}\,(\tau_1\le T(\varepsilon))\right)+ \\
&\qquad\quad V\mathbf{P}(\tau_1\ge T(\varepsilon)) \\
&\le U\left(\max_{0\le t\le T(\varepsilon)}\mathbf{E}_{\widetilde{q}}|\widehat{\widetilde{q}}_t^\varepsilon-\widehat{\widetilde{q}}_t|\right)+V\mathbf{P}(\tau_1\ge T(\varepsilon)) \\
&\le U\max_{0\le t\le T(\varepsilon)}\left(\mathbf{E}_{\widetilde{q}}|\widehat{\widetilde{q}}_t^\varepsilon-\widehat{\widetilde{q}}_t|^2\right)^{1/2}+V\mathbf{P}(\tau_1\ge T(\varepsilon))\ .
\end{aligned}
$$

By the integral form of the stochastic differential equations of the processes $\widehat{\widetilde{q}}_t^\varepsilon$ and $\widehat{\widetilde{q}}_t$ we have

$$
\begin{aligned}
&|\widehat{\widetilde{q}}_t^\varepsilon-\widehat{\widetilde{q}}_t|^2 \\
&\le C\left(\left|\int_0^t\left[\left(\frac{B(\widehat{\widetilde{q}}_s^\varepsilon)}{\widehat{\Lambda}(\widehat{\widetilde{q}}_s^\varepsilon)+\varepsilon}-\frac{\widehat{\Lambda}'(\widehat{\widetilde{q}}_s^\varepsilon)}{2(\widehat{\Lambda}(\widehat{\widetilde{q}}_s^\varepsilon)+\varepsilon)^3}\right)-\left(\frac{B(\widehat{\widetilde{q}}_s^\varepsilon)}{\widehat{\Lambda}(\widehat{\widetilde{q}}_s^\varepsilon)}-\frac{\widehat{\Lambda}'(\widehat{\widetilde{q}}_s^\varepsilon)}{2(\widehat{\Lambda}(\widehat{\widetilde{q}}_s^\varepsilon))^3}\right)\right]ds\right|^2+\right. \\
&\left|\int_0^t\left[\left(\frac{B(\widehat{\widetilde{q}}_s^\varepsilon)}{\widehat{\Lambda}(\widehat{\widetilde{q}}_s^\varepsilon)}-\frac{\widehat{\Lambda}'(\widehat{\widetilde{q}}_s^\varepsilon)}{2(\widehat{\Lambda}(\widehat{\widetilde{q}}_s^\varepsilon))^3}\right)-\left(\frac{B(\widehat{\widetilde{q}}_s)}{\widehat{\Lambda}(\widehat{\widetilde{q}}_s)}-\frac{\widehat{\Lambda}'(\widehat{\widetilde{q}}_s)}{2(\widehat{\Lambda}(\widehat{\widetilde{q}}_s))^3}\right)\right]ds\right|^2+ \\
&\left.\left|\int_0^t\left[\frac{1}{\widehat{\Lambda}(\widehat{\widetilde{q}}_s^\varepsilon)+\varepsilon}-\frac{1}{\widehat{\Lambda}(\widehat{\widetilde{q}}_s^\varepsilon)}\right]dW_s\right|^2+\left|\int_0^t\left[\frac{1}{\widehat{\Lambda}(\widehat{\widetilde{q}}_s^\varepsilon)}-\frac{1}{\widehat{\Lambda}(\widehat{\widetilde{q}}_s)}\right]dW_s\right|^2\right)\ .
\end{aligned}
$$

Let $\alpha(\lambda)$ be the Lipschitz constant of $\dfrac{1}{x}$ $(x>\lambda)$, $\beta(\lambda)$ that of $\dfrac{1}{2x^3}$ $(x>\lambda)$, $\gamma(\delta')$ that of $\dfrac{B(\widehat{q})}{\widehat{\Lambda}(q)}-\dfrac{\widehat{\Lambda}'(q)}{2\widehat{\Lambda}(q)^3}$ $(q\ge\delta')$, $\mu(\delta')$ that of $\dfrac{1}{\widehat{\Lambda}(q)}$ $(q\ge\delta')$. Let $m(\delta')\equiv\min\limits_{x\in[\delta',b]}\Lambda(x)$.

We can estimate

$$
\begin{aligned}
&\mathbf{E}_{\widetilde{q}}\left|\int_0^t\left[\left(\frac{B(\widehat{\widetilde{q}}_s^\varepsilon)}{\widehat{\Lambda}(\widehat{\widetilde{q}}_s^\varepsilon)+\varepsilon}-\frac{\widehat{\Lambda}'(\widehat{\widetilde{q}}_s^\varepsilon)}{2(\widehat{\Lambda}(\widehat{\widetilde{q}}_s^\varepsilon)+\varepsilon)^3}\right)-\left(\frac{B(\widehat{\widetilde{q}}_s^\varepsilon)}{\widehat{\Lambda}(\widehat{\widetilde{q}}_s^\varepsilon)}-\frac{\widehat{\Lambda}'(\widehat{\widetilde{q}}_s^\varepsilon)}{2(\widehat{\Lambda}(\widehat{\widetilde{q}}_s^\varepsilon))^3}\right)\right]ds\right|^2 \\
&\le A_1(t^2\varepsilon^2[\alpha^2(m(\delta'))+\beta^2(m(\delta'))])\ ,
\end{aligned}
$$

$$
\begin{aligned}
&\mathbf{E}_{\widetilde{q}}\left|\int_0^t\left[\left(\frac{B(\widehat{\widetilde{q}}_s^\varepsilon)}{\widehat{\Lambda}(\widehat{\widetilde{q}}_s^\varepsilon)}-\frac{\widehat{\Lambda}'(\widehat{\widetilde{q}}_s^\varepsilon)}{2(\widehat{\Lambda}(\widehat{\widetilde{q}}_s^\varepsilon))^3}\right)-\left(\frac{B(\widehat{\widetilde{q}}_s)}{\widehat{\Lambda}(\widehat{\widetilde{q}}_s)}-\frac{\widehat{\Lambda}'(\widehat{\widetilde{q}}_s)}{2(\widehat{\Lambda}(\widehat{\widetilde{q}}_s))^3}\right)\right]ds\right|^2 \\
&\le A_2t\gamma^2(\delta')\int_0^t\mathbf{E}_{\widetilde{q}}|\widehat{\widetilde{q}}_s^\varepsilon-\widehat{\widetilde{q}}_s|^2ds\ ,
\end{aligned}
$$

$$\mathbf{E}_{\widetilde{q}}\left|\int_0^t\left[\frac{1}{\widehat{\Lambda}(\widehat{\widetilde{q}}_s^{\varepsilon})+\varepsilon}-\frac{1}{\widehat{\Lambda}(\widehat{\widetilde{q}}_s^{\varepsilon})}\right]dW_s\right|^2\leq\int_0^t\varepsilon^2\alpha^2(m(\delta'))ds=\varepsilon^2t\alpha^2(m(\delta'))\ ,$$

$$\mathbf{E}_{\widetilde{q}}\left|\int_0^t\left[\frac{1}{\widehat{\Lambda}(\widehat{\widetilde{q}}_s^{\varepsilon})}-\frac{1}{\widehat{\Lambda}(\widehat{\widetilde{q}}_s)}\right]dW_s\right|^2\leq\mu^2(\delta')\int_0^t\mathbf{E}_{\widetilde{q}}|\widehat{\widetilde{q}}_s^{\varepsilon}-\widehat{\widetilde{q}}_s|^2ds\ .$$

We have, by using the above estimates, with a possible change of the constant C, that

$$\mathbf{E}_{\widetilde{q}}|\widehat{\widetilde{q}}_t^{\varepsilon}-\widehat{\widetilde{q}}_t|^2\leq C\left(t\varepsilon^2(t(\alpha^2(m(\delta'))+\beta^2(m(\delta')))+\alpha^2(m(\delta')))+(t\gamma^2(\delta')+\mu^2(\delta'))\int_0^t\mathbf{E}_{\widetilde{q}}|\widehat{\widetilde{q}}_s^{\varepsilon}-\widehat{\widetilde{q}}_s|^2ds\right)\ .$$

By Bellman-Gronwall inequality we have

$$\mathbf{E}_{\widetilde{q}}|\widehat{\widetilde{q}}_t^{\varepsilon}-\widehat{\widetilde{q}}_t|^2\leq Ct\varepsilon^2(t(\alpha^2(m(\delta'))+\beta^2(m(\delta')))+\alpha^2(m(\delta')))\exp\left(C(t\gamma^2(\delta')+\mu^2(\delta'))t\right)\ .$$

As we can check that $|\alpha(m(\delta'))|\leq\dfrac{1}{m^2(\delta')}$, $\beta(m(\delta'))\leq\dfrac{A_3}{m^4(\delta')}$, $\gamma(\delta')\leq\dfrac{A_3}{m^4(\delta')}$ and $|\mu(\delta')|\leq\dfrac{A_3}{m^2(\delta')}$, this gives, as δ' is small, that

$$\begin{aligned}&\max_{0\leq t\leq T(\varepsilon)}\left(\mathbf{E}_{\widetilde{q}}|\widehat{\widetilde{q}}_t^{\varepsilon}-\widehat{\widetilde{q}}_t|^2\right)^{1/2}\leq\\&\qquad\leq CT(\varepsilon)\varepsilon(\alpha^2(m(\delta'))+\beta^2(m(\delta'))+\frac{\alpha^2(m(\delta'))}{T(\varepsilon)})^{1/2}\exp\left(C(T(\varepsilon)\gamma^2(\delta')+\mu^2(\delta'))T(\varepsilon)\right)\\&\qquad\leq CT(\varepsilon)\frac{\varepsilon}{\min\limits_{q\in[1+\delta',1+b]}\lambda^4(q)}\exp\left(CT^2(\varepsilon)\frac{1}{\min\limits_{q\in[1+\delta',1+b]}\lambda^8(q)}\right)\ .\end{aligned}$$

Noticing that by strong Markov property $\mathbf{P}(\tau_1\geq T(\varepsilon))\leq K\exp(-pT(\varepsilon))$ for some $p>0,K>0$, we see that

$$|\psi_2^{\varepsilon}(q)|\leq CT(\varepsilon)\frac{\varepsilon}{\min\limits_{q\in[1+\delta',1+b]}\lambda^4(q)}\exp\left(CT^2(\varepsilon)\frac{1}{\min\limits_{q\in[1+\delta',1+b]}\lambda^8(q)}\right)+V\exp(-pT(\varepsilon))\ .$$

Let us choose $T(\varepsilon)=\sqrt{\ln\ln\dfrac{1}{\varepsilon}}$. We will then have

$$|\psi_2^{\varepsilon}(q)|\leq C\left(\ln\ln\frac{1}{\varepsilon}\right)^{1/2}\frac{\varepsilon}{\min\limits_{q\in[1+\delta',1+b]}\lambda^4(q)}\left(\ln\frac{1}{\varepsilon}\right)^{\frac{C}{\min\limits_{q\in[1+\delta',1+b]}\lambda^8(q)}}+V\exp(-p\sqrt{\ln\ln\frac{1}{\varepsilon}}))\ .$$

For fixed $\delta' > 0$, one can choose ε small enough such that

$$|\psi_2^\varepsilon(q)| \leq \frac{U_0\varepsilon^\kappa}{\min\limits_{q\in[1+\delta',1+b]\cup[-1+a,-1-\delta']} \lambda^4(q)} + U_0 \exp(-p\sqrt{\ln\ln\frac{1}{\varepsilon}})$$

for some $U_0 > 0$, $p > 0$ and $0 < \kappa < 1$. As we choose first $\varepsilon \downarrow 0$ and then $\delta' \downarrow 0$, this gives that as ε is small we have $|\psi_2^\varepsilon(q)| \leq (\widetilde{u}(\delta) - \widetilde{u}(-\delta))^2$. $\square$

Lemma 2.7. *We have, as $\varepsilon, \delta, \delta'$ are small, for $q \in [-1-\delta', 1+\delta']$ and $C > 0$, that*

$$\left|\mathbf{P}_q(\pi(q^\varepsilon_{\sigma_0}) = \delta) - \frac{\widetilde{u}(0) - \widetilde{u}(-\delta)}{\widetilde{u}(\delta) - \widetilde{u}(-\delta)}\right| \leq \frac{\widetilde{u}(\delta') - \widetilde{u}(0) + C\varepsilon}{\widetilde{u}(\delta) - \widetilde{u}(-\delta)} \ ,$$

$$\left|\mathbf{P}_q(\pi(q^\varepsilon_{\sigma_0}) = -\delta) - \frac{\widetilde{u}(\delta) - \widetilde{u}(0)}{\widetilde{u}(\delta) - \widetilde{u}(-\delta)}\right| \leq \frac{\widetilde{u}(\delta') - \widetilde{u}(0) + C\varepsilon}{\widetilde{u}(\delta) - \widetilde{u}(-\delta)} \ .$$

Proof. Let $\widetilde{q} = \pi(q) \in [-\delta', \delta']$. We have, for bounded positive functions $C_1(\delta, \varepsilon)$, $C_2(\delta, \varepsilon)$ and positive constants C_1, C_2, C, that

$$\begin{aligned}
&\left|\mathbf{P}_q(\pi(q^\varepsilon_{\sigma_0}) = \delta) - \frac{\widetilde{u}(0) - \widetilde{u}(-\delta)}{\widetilde{u}(\delta) - \widetilde{u}(-\delta)}\right| \\
&= \left|\frac{u^\varepsilon(q) - u^\varepsilon(-1-\delta)}{u^\varepsilon(1+\delta) - u^\varepsilon(-1-\delta)} - \frac{\widetilde{u}(0) - \widetilde{u}(-\delta)}{\widetilde{u}(\delta) - \widetilde{u}(-\delta)}\right| \\
&= \left|\frac{\widetilde{u}(0) - \widetilde{u}(-\delta) + \widetilde{u}(\widetilde{q}) - \widetilde{u}(0) + C_1(\delta,\varepsilon)\varepsilon}{\widetilde{u}(\delta) - \widetilde{u}(-\delta) + C_2(\delta,\varepsilon)\varepsilon} - \frac{\widetilde{u}(0) - \widetilde{u}(-\delta)}{\widetilde{u}(\delta) - \widetilde{u}(-\delta)}\right| \\
&\leq \frac{(\widetilde{u}(\widetilde{q}) - \widetilde{u}(0) + C_1\varepsilon)(\widetilde{u}(\delta) - \widetilde{u}(-\delta)) + C_2\varepsilon(\widetilde{u}(0) - \widetilde{u}(-\delta))}{(\widetilde{u}(\delta) - \widetilde{u}(-\delta))^2} \\
&\leq \frac{\widetilde{u}(\delta') - \widetilde{u}(0) + C\varepsilon}{\widetilde{u}(\delta) - \widetilde{u}(-\delta)} \ .
\end{aligned}$$

The estimate of $\mathbf{P}_q(\pi(q^\varepsilon_{\sigma_0}) = -\delta)$ is similar. $\square$

Lemma 2.8. *We have, as ε are small, for $q \in [-1-\delta', 1+\delta']$, that $|(I)^\varepsilon(q)| \leq C(\widetilde{u}(\delta) - \widetilde{u}(-\delta))^2$.*

Proof. We have, using Lemma 2.7, that

$$
\begin{aligned}
&|(I)^\varepsilon(q)| \\
&= |\mathbf{E}_q f(\pi(q^\varepsilon_{\sigma_0})) - f(\pi(q))| \\
&= |(f(\delta)-f(0))\mathbf{P}_q(\pi(q^\varepsilon_{\sigma_0})=\delta) - (f(0)-f(-\delta))\mathbf{P}_q(\pi(q^\varepsilon_{\sigma_0})=-\delta) + (f(0)-f(\pi(q)))| \\
&\leq \left|(f(\delta)-f(0))\frac{\widetilde{u}(0)-\widetilde{u}(-\delta)}{\widetilde{u}(\delta)-\widetilde{u}(-\delta)} - (f(0)-f(-\delta))\frac{\widetilde{u}(\delta)-\widetilde{u}(0)}{\widetilde{u}(\delta)-\widetilde{u}(-\delta)}\right| + \\
&\quad C_4\frac{\widetilde{u}(\delta')-\widetilde{u}(0)+M\varepsilon}{\widetilde{u}(\delta)-\widetilde{u}(-\delta)} + C_5(\widetilde{u}(\delta')-\widetilde{u}(0)) \\
&= \left|\frac{(\widetilde{u}(0)-\widetilde{u}(-\delta))(\widetilde{u}(\delta)-\widetilde{u}(0))}{\widetilde{u}(\delta)-\widetilde{u}(-\delta)}\left(\frac{f(\delta)-f(0)}{\widetilde{u}(\delta)-\widetilde{u}(0)} - \frac{f(0)-f(-\delta)}{\widetilde{u}(0)-\widetilde{u}(-\delta)}\right)\right| + \\
&\quad C_4\frac{\widetilde{u}(\delta')-\widetilde{u}(0)+M\varepsilon}{\widetilde{u}(\delta)-\widetilde{u}(-\delta)} + C_5(\widetilde{u}(\delta')-\widetilde{u}(0)) \\
&\leq C_3(\widetilde{u}(\delta)-\widetilde{u}(-\delta))^2 + C_4\frac{\widetilde{u}(\delta')-\widetilde{u}(0)+M\varepsilon}{\widetilde{u}(\delta)-\widetilde{u}(-\delta)} + C_5(\widetilde{u}(\delta')-\widetilde{u}(0)) \ .
\end{aligned}
$$

We have used our gluing condition $D^+_{\widetilde{u}}f(0) = D^-_{\widetilde{u}}f(0)$. Now we choose first $\varepsilon \downarrow 0$ then $\delta' \downarrow 0$, we get, as ε is small, that $|(I)^\varepsilon(q)| \leq C(\widetilde{u}(\delta)-\widetilde{u}(-\delta))^2$. $\square$

Lemma 2.9. *As $\varepsilon, \delta, \delta'$ are small, for $q \in [-1-\delta', 1+\delta']$ we have,*

$$
\mathbf{E}_q\sigma_0 \leq C(\widetilde{u}(\delta)-\widetilde{u}(-\delta))(\widetilde{v}(\delta)-\widetilde{v}(-\delta)) \ , \ \mathbf{E}_q(1-e^{-\lambda\sigma_0}) \leq C(\widetilde{u}(\delta)-\widetilde{u}(-\delta))(\widetilde{v}(\delta)-\widetilde{v}(-\delta)) \ .
$$

Proof. We apply the well known formula for the expected exit time (see, for example [29, Chapter VII, Theorem 3.6]) and we have

$$
\mathbf{E}_q\sigma_0 = \int_{-1-\delta}^{1+\delta} G^\varepsilon(q,r)dv^\varepsilon(r) \ ,
$$

where the Green function

$$
G^\varepsilon(q,r) = \begin{cases} \dfrac{(u^\varepsilon(q)-u^\varepsilon(-1-\delta))(u^\varepsilon(1+\delta)-u^\varepsilon(r))}{u^\varepsilon(1+\delta)-u^\varepsilon(-1-\delta)} \text{ for } -1-\delta \leq q \leq r \leq 1+\delta \ , \\ \dfrac{(u^\varepsilon(r)-u^\varepsilon(-1-\delta))(u^\varepsilon(1+\delta)-u^\varepsilon(q))}{u^\varepsilon(1+\delta)-u^\varepsilon(-1-\delta)} \text{ for } -1-\delta \leq r \leq q \leq 1+\delta \ , \\ 0 \text{ otherwise } . \end{cases}
$$

Therefore it is easy to estimate

$$
\begin{aligned}
&\mathbf{E}_q\sigma_0 \\
&\leq (u^\varepsilon(1+\delta)-u^\varepsilon(-1-\delta))(v^\varepsilon(1+\delta)-v^\varepsilon(-1-\delta)) \\
&\leq (\widetilde{u}(\delta)-\widetilde{u}(-\delta)+C_6\varepsilon)(\widetilde{v}(\delta)-\widetilde{v}(-\delta)+C_7\varepsilon) \\
&\leq C(\widetilde{u}(\delta)-\widetilde{u}(-\delta))(\widetilde{v}(\delta)-\widetilde{v}(-\delta))
\end{aligned}
$$

as desired.

This helps us to find

$$\mathbf{E}_q(1-e^{-\lambda\sigma_0}) = \lambda\mathbf{E}_q\left[\int_0^{\sigma_0} e^{-\lambda s}ds\right] \leq \lambda\mathbf{E}_q\sigma_0 \leq C(\widetilde{u}(\delta)-\widetilde{u}(-\delta))(\widetilde{v}(\delta)-\widetilde{v}(-\delta)) \ .$$

□

Lemma 2.10. *For $q \in G(\delta)$ and δ sufficiently small, we have*

$$\lim_{\delta'\downarrow 0}\lim_{\varepsilon\downarrow 0}\mathbf{E}_q e^{-\lambda\tau_1} \leq 1 - C(\widetilde{u}(\delta)) \wedge (-\widetilde{u}(-\delta)) \ .$$

Proof. Without loss of generality let $q \in [1+\delta, 1+b]$. The expected value $M^\varepsilon(q) = \mathbf{E}_q e^{-\lambda\tau_1}$ is the solution of the differential equation $D_{v^\varepsilon}D_{u^\varepsilon}M^\varepsilon(q) = \lambda M^\varepsilon(q)$, $M^\varepsilon(1+\delta') = M^\varepsilon(1+b) = 1$.

There exist two solutions $f_1^\lambda(q)$, $f_2^\lambda(q)$ of the equation $D_vD_uf = \lambda f$ with $f_1^\lambda(1) = f_2^\lambda(1+b) = 1$ and $f_1^\lambda(1+b) = f_2^\lambda(1) = 0$. The derivatives $D_uf_1^\lambda(x)$, $D_uf_2^\lambda(x)$ are increasing functions, $-\infty < \lim\limits_{q\downarrow 1} D_u(f_1^\lambda+f_2^\lambda)(q) < 0$, $0 < \lim\limits_{q\uparrow 1+b} D_u(f_1^\lambda+f_2^\lambda)(q) < \infty$ (see [7], [26]).

We shall make use of Lemma 2.6. Since $q \in [1+\delta, 1+b]$ we see that $\sigma_0 = 0$. Lemma 2.6 tells us that, for $k = 1, 2$, we have

$$\lim_{\varepsilon\downarrow 0}\left|\mathbf{E}_q\left[\int_0^{\tau_1} e^{-\lambda t}[\lambda f_k^\lambda(q_t^\varepsilon) - D_vD_uf_k^\lambda(q_t^\varepsilon)]dt + e^{-\lambda\tau_1}f_k^\lambda(q_{\tau_1}^\varepsilon)\right] - f_k^\lambda(q)\right| \leq (\widetilde{u}(\delta)-\widetilde{u}(-\delta))^2 \ .$$

Taking into account the definitions of f_1^λ, f_2^λ we see that the above inequality gives

$$\left|\lim_{\varepsilon\downarrow 0}\mathbf{E}_q e^{-\lambda\tau_1}f_k^\lambda(q_{\tau_1}^\varepsilon) - f_k^\lambda(q)\right| \leq (\widetilde{u}(\delta)-\widetilde{u}(-\delta))^2 \ .$$

Since $f_k^\lambda(q_{\tau_1}^\varepsilon) = f_k^\lambda(1+\delta')$ when $q_{\tau_1}^\varepsilon = 1+\delta'$ and $f_k^\lambda(q_{\tau_1}^\varepsilon) = f_k^\lambda(1+b)$ when $q_{\tau_1}^\varepsilon = 1+b$, we see that for some $K > 0$ we have

$$\left|\lim_{\varepsilon\downarrow 0}\mathbf{E}_q e^{-\lambda\tau_1} - \frac{(f_2^\lambda(1+b)-f_2^\lambda(1+\delta'))f_1^\lambda(q) + (f_1^\lambda(1+\delta')-f_1^\lambda(1+b))f_2^\lambda(q)}{f_1^\lambda(1+\delta')f_2^\lambda(1+b) - f_1^\lambda(1+b)f_2^\lambda(1+\delta')}\right| \leq K(\widetilde{u}(\delta)-\widetilde{u}(-\delta))^2 \ .$$

(The expression

$$\frac{(f_2^\lambda(1+b)-f_2^\lambda(1+\delta'))f_1^\lambda(q) + (f_1^\lambda(1+\delta')-f_1^\lambda(1+b))f_2^\lambda(q)}{f_1^\lambda(1+\delta')f_2^\lambda(1+b) - f_1^\lambda(1+b)f_2^\lambda(1+\delta')}$$

is the solution of the equation $\lambda f(q) = D_vD_uf$ with $f(1+\delta') = f(1+b) = 1$.)

This gives

$$\left|\lim_{\delta'\downarrow 0}\lim_{\varepsilon\downarrow 0}\mathbf{E}_q(1-e^{-\lambda\tau_1})-[1-(f_1^\lambda(q)+f_2^\lambda(q))]\right|\leq K(\widetilde{u}(\delta)-\widetilde{u}(-\delta))^2 \ .$$

Taking into account that $-\infty<\lim\limits_{q\downarrow 1}D_u(f_1^\lambda+f_2^\lambda)(q)<0$, $0<\lim\limits_{q\uparrow 1+b}D_u(f_1^\lambda+f_2^\lambda)(q)<\infty$ we see from the above estimate that

$$\lim_{\delta'\downarrow 0}\lim_{\varepsilon\downarrow 0}\mathbf{E}_q(1-e^{-\lambda\tau_1})\geq C(\widetilde{u}(\delta))$$

for $q\in[1+\delta,1+b]$ and δ sufficiently small. The case of $\widetilde{u}(-\delta)$ is handled in a similar way. $\square$

2.3 A two dimensional model problem

In this section we discuss a two dimensional model problem. We work with a Smoluchowski-Kramers approximation in the plane $\mathbb{R}^2$. Let us suppose that the friction coefficient $\lambda(\bullet)$ depends on the y variable only: $\lambda(x,y)=\lambda(y)$. Suppose for $y\in[-1,1]$ we have $\lambda(y)=0$. For $y\notin[-1,1]$ we have $\lambda(y)>0$. For simplicity of presentation we also assume that the drift is zero: $\boldsymbol{b}(\bullet)=\mathbf{0}$. All the other assumptions about $\lambda(\bullet)$ are the same as was made in Section 1.

In addition, we assume that for $\varepsilon>0$,

$$\int_{-\varepsilon-1}^{-1}\frac{1}{\lambda(y)}dy=\int_1^{1+\varepsilon}\frac{1}{\lambda(y)}dy=\infty \ .$$

(In the case that both integrals converge the proof of Lemma 3.1 repeat that in the case of both integrals divergent but we do not know anything about the case of one integral convergent and the other divergent.)

As we already introduced in equation (1.8) of Section 1, we are actually considering the stochastic differential equation for the position of the particle $\boldsymbol{q}_t^\varepsilon\in\mathbb{R}^2$ as follows:

$$\dot{\boldsymbol{q}}_t^\varepsilon=-\frac{\nabla\lambda(\boldsymbol{q}_t^\varepsilon)}{2(\lambda(\boldsymbol{q}_t^\varepsilon)+\varepsilon)^3}+\frac{1}{\lambda(\boldsymbol{q}_t^\varepsilon)+\varepsilon}\dot{\boldsymbol{W}}_t \ , \ \boldsymbol{q}_0^\varepsilon=\boldsymbol{q}_0\in\mathbb{R}^2 \ , \ \varepsilon>0 \ . \tag{3.1}$$

By taking into account our assumption on the friction coefficient λ we can write the above equation in coordinate form. Let $\boldsymbol{q}_t^\varepsilon=(x_t^\varepsilon,y_t^\varepsilon)$. Let $\boldsymbol{W}_t=(W_t^1,W_t^2)$. We have

$$\begin{cases}\dot{x}_t^\varepsilon=\dfrac{1}{\lambda(y_t^\varepsilon)+\varepsilon}\dot{W}_t^1 \ , \ x_0^\varepsilon=x_0\in\mathbb{R} \ , \\ \dot{y}_t^\varepsilon=-\dfrac{\lambda'(y_t^\varepsilon)}{2(\lambda(y_t^\varepsilon)+\varepsilon)^3}+\dfrac{1}{\lambda(y_t^\varepsilon)+\varepsilon}\dot{W}_t^2 \ , \ y_0^\varepsilon=y_0\in\mathbb{R} \ .\end{cases} \tag{3.2}$$

Let $a < 0 < b$ be given. Throughout this section we will assume that our process $\boldsymbol{q}_t^\varepsilon$ is stopped once it exits from the domain $\{(x,y) \in \mathbb{R}^2 : a-1 \leq y \leq b+1\}$. We therefore suppose that $y_0 \in [a-1, b+1]$.

Note that, similarly as in Section 2, the process y_t^ε is a strong Markov process subject to a generalized second order differential operator in the form $D_{v^\varepsilon(y)}D_{u^\varepsilon(y)}$ where

$$u^\varepsilon(y) = \int_0^y (\lambda(s)+\varepsilon)ds \ , \ v^\varepsilon(y) = 2\int_0^y (\lambda(s)+\varepsilon)ds \ . \tag{3.3}$$

Let

$$u(y) = \int_0^y \lambda(s)ds \ , \ v(y) = 2\int_0^y \lambda(s)ds \ . \tag{3.4}$$

We have the obvious relation $u^\varepsilon(y) = u(y) + \varepsilon y$ and $v^\varepsilon(y) = v(y) + 2\varepsilon y$.

Let us identify points in the x direction $x \sim x + 2\pi$. Therefore we get a process on the cylinder $S^1 \times [a-1, b+1]$, stopped once it hits the boundary $\{y = a-1 \text{ or } b+1\}$. Let

$$\begin{cases} \theta_t^\varepsilon = x_t^\varepsilon \mod 2\pi \ , \\ y_t^\varepsilon = y_t^\varepsilon \ . \end{cases}$$

In the rest of this section we refer to the process $\boldsymbol{q}_t^\varepsilon$ as the one on a cylinder: $\boldsymbol{q}_t^\varepsilon = (\theta_t^\varepsilon, y_t^\varepsilon)$ is on the cylinder $S^1 \times [a-1, b+1]$. When we speak about the process $\boldsymbol{q}_t^\varepsilon$ on the domain $\{(x,y) \in \mathbb{R}^2 : a-1 \leq y \leq b+1\} \subset \mathbb{R}^2$ we will instead refer to the coordinate representation $(x_t^\varepsilon, y_t^\varepsilon)$.

Let $\mathfrak{C}$ be the product $S^1 \times [a,b]$ with all points $S^1 \times \{0\}$ identified, forming the point $\mathfrak{o}$. A generic point on $\mathfrak{C}$ will be denoted $\widetilde{\boldsymbol{q}} = (\theta, \widetilde{y})$ where $\theta \in S^1$ and $\widetilde{y} \in [a,b]$. All points $(\theta, 0)$ correspond to $\mathfrak{o}$.

Let us consider the following projection map $\boldsymbol{\pi} : S^1 \times [a-1, b+1] \to \mathfrak{C}$. We let

$$\boldsymbol{\pi}(\theta, y) = \begin{cases} (\theta, y-1) \ , & \text{for } 1 < y \leq b+1 \ ; \\ (\theta, y+1) \ , & \text{for } a-1 \leq y < -1 \ ; \\ \mathfrak{o} \ , & \text{for } -1 \leq y \leq 1 \ . \end{cases} \tag{3.5}$$

Let $\boldsymbol{\pi}(\boldsymbol{q}_t^\varepsilon) = \widetilde{\boldsymbol{q}}_t^\varepsilon = (\theta_t^\varepsilon, \widetilde{y}_t^\varepsilon)$. We see that $\widetilde{y}_t^\varepsilon = \pi(y_t^\varepsilon)$ where π is the projection map introduced in Section 2.

Let, as in Section 2, $\widetilde{u}(\widetilde{y}) = u(\widetilde{y}-1)$ for $\widetilde{y} < 0$ and $\widetilde{u}(\widetilde{y}) = u(\widetilde{y}+1)$ for $\widetilde{y} > 0$ and $\widetilde{u}(0) = u(1) = u(-1)$; $\widetilde{v}(\widetilde{y}) = v(\widetilde{y}-1)$ for $\widetilde{y} < 0$ and $\widetilde{v}(\widetilde{y}) = v(\widetilde{y}+1)$ for $\widetilde{y} > 0$ and $\widetilde{v}(0) = \widetilde{v}(1) = \widetilde{v}(-1)$. The functions $\widetilde{u}(\widetilde{y})$ and $\widetilde{v}(\widetilde{y})$ are continuous strictly increasing functions on $[a,b]$. Let $\widetilde{\lambda}(\widetilde{y}) = \lambda(\widetilde{y}-1)$ for $\widetilde{y} < 0$ and $\widetilde{\lambda}(\widetilde{y}) = \lambda(\widetilde{y}+1)$ for $\widetilde{y} > 0$ and $\widetilde{\lambda}(0) = 0$.

Let A be the operator given, for $\widetilde{y} \neq 0$, by the formula

$$Af(\theta, \widetilde{y}) = D_{\widetilde{u}(\widetilde{y})} D_{\widetilde{v}(\widetilde{y})} f + \frac{1}{\widetilde{\lambda}^2(\widetilde{y})} \frac{\partial^2}{\partial \theta^2} f \ . \tag{3.6}$$

Let $D(A)$ be the subset of the space $\mathbf{C}(\mathfrak{C})$ consisting of functions $f(\widetilde{\boldsymbol{q}})$ for which $Af(\theta, \widetilde{y})$ is defined and continuous for $\widetilde{y} \neq 0$, the derivatives in it being continuous; such that finite limits

$$\lim_{\theta' \to \theta, \widetilde{y} \to 0-} D_{\widetilde{u}(\widetilde{y})} f(\theta', \widetilde{y}) \ , \quad \lim_{\theta' \to \theta, \widetilde{y} \to 0+} D_{\widetilde{u}(\widetilde{y})} f(\theta', \widetilde{y}) \ , \tag{3.7}$$

exist;

$$\lim_{\theta' \to \theta, \widetilde{y} \to 0} Af(\theta', \widetilde{y}) \tag{3.8}$$

exists and does not depend on θ;

$$\lim_{\theta' \to \theta, \widetilde{y} \to a} Af(\theta', \widetilde{y}) = \lim_{\theta' \to \theta, \widetilde{y} \to b} Af(\theta', \widetilde{y}) = 0 \ ; \tag{3.9}$$

and

$$\int_0^{2\pi} \lim_{\theta' \to \theta, \widetilde{y} \to 0-} D_{\widetilde{u}(\widetilde{y})} f(\theta', \widetilde{y}) d\theta = \int_0^{2\pi} \lim_{\theta' \to \theta, \widetilde{y} \to 0+} D_{\widetilde{u}(\widetilde{y})} f(\theta', \widetilde{y}) d\theta \ . \tag{3.10}$$

It is worth mentioning here that the above condition (3.10) in the definition of $D(A)$ can be replaced by the condition that $\lim_{\theta' \to \theta, \widetilde{y} \to 0-} D_{\widetilde{u}(\widetilde{y})} f(\theta', \widetilde{y})$ and $\lim_{\theta' \to \theta, \widetilde{y} \to 0+} D_{\widetilde{u}(\widetilde{y})} f(\theta', \widetilde{y})$ not depending on θ and coinciding. In this case the proof of Lemma 3.1 remains the same.

Let us define, for $f \in D(A)$, $Af(\theta, a)$ and $Af(\theta, b)$ as the limits (3.9) and $Af(\mathfrak{o})$ as the limit (3.8). The operator A defined on $D(A)$ is a linear operator $D(A) \mapsto \mathbf{C}(\mathfrak{C})$.

Lemma 3.1. *The closure $\overline{A|_{D(A)}}$ of the operator $A|_{D(A)}$ exists and is the infinitesimal operator of a Markov semigroup on $\mathbf{C}(\mathfrak{C})$.*

(The corresponding Markov process $\widetilde{\boldsymbol{q}}_t$ stops after reaching the boundary of $\mathfrak{C}$ ($\widetilde{y} = a$ or b).)

Proof. We use the Hille-Yosida theorem and we check the following:

• The domain $D(A)$ is dense in $\mathbf{C}(\mathfrak{C})$.

This is because we can approximate every function g in $\mathbf{C}(\mathfrak{C})$ by a function f which is smooth, close to g outside a neighborhood of $\mathfrak{o}$ and is equal to $g(\mathfrak{o})$ in the neighborhood of $\mathfrak{o}$. This function f satisfies our restrictions on $D(A)$ and can approximate the function g with respect to the norm of $\mathbf{C}(\mathfrak{C})$ as we choose the neighborhood of $\mathfrak{o}$ small enough.

• The operator $A|_{D(A)}$ satisfies the maximum principle: for $f \in D(A)$, if this function reaches its maximum value at a point $\widetilde{\boldsymbol{q}} \in \mathfrak{C}$ we have $Af(\widetilde{\boldsymbol{q}}) \leq 0$.

Indeed, for $\widetilde{\boldsymbol{q}} = (\theta, a)$ or (θ, b), we have $Af(\widetilde{\boldsymbol{q}}) = 0$. If $\widetilde{\boldsymbol{q}} = (\theta, \widetilde{y})$, $\widetilde{y} \neq 0$ the first partial derivatives at $\widetilde{\boldsymbol{q}}$ are equal to 0 and $\frac{\partial^2}{\partial \theta^2} f(\theta, \widetilde{y}) \leq 0$, $D_{\widetilde{v}(\widetilde{y})} D_{\widetilde{u}(\widetilde{y})} \leq 0$. Finally, if $\widetilde{\boldsymbol{q}} = \mathfrak{o}$ we have the left-hand derivative $D^-_{\widetilde{u}(\widetilde{y})} f(\theta, 0) \geq 0$, the right-hand derivative $D^+_{\widetilde{u}(\widetilde{y})} f(\theta, 0) \leq 0$ and by (3.10) both these derivatives are equal to 0. It follows then that the limit as $\widetilde{y} \to 0$ of the second $\widetilde{y}$-derivative is non-positive for all $\theta \in S^1$. Since the integral over S^1 of the second θ derivative is equal to 0 for all $\widetilde{y} \neq 0$, taking into account that $Af(\mathfrak{o})$ is equal to the limit (3.8), we have that $Af(\mathfrak{o}) \leq 0$.

It follows from the maximum principle that for $\lambda > 0$ the operator $\lambda I - A|_{D(A)}$ does not send to zero any function that is not equal 0, and this linear operator has an inverse (that is not defined on the whole $\mathbf{C}(\mathfrak{C})$), with $\|(\lambda I - A|_{D(A)})^{-1}\| \leq \lambda^{-1}$. Every bounded linear operator does have a closure (which is just its extension by continuity), and with it the operators $\lambda I - A|_{D(A)}$ and $A|_{D(A)}$ also have closures.

• Finally, to check that we can apply Hille-Yosida theorem to the closure $\overline{A|_{D(A)}}$ we have only to check that the bounded operator $(\lambda I - A|_{D(A)})^{-1}$ is defined on a dense set. That is, for a dense subset of $F \in \mathbf{C}(\mathfrak{C})$ there exists a solution $f \in D(A)$ of the equation

$$\lambda f - Af = F\ . \tag{3.11}$$

Let us take $F(\theta, \widetilde{y}) = e^{in\theta} G(\widetilde{y})$, defining $F(\mathfrak{o})$ as its limit as $\widetilde{y} \to 0$. Of course for $n \neq 0$ we have to have $\lim\limits_{\widetilde{y} \to 0} G(\widetilde{y})$ (which limit we'll take as the value $G(0)$) equal to 0.

We shall look for the solution $f \in D(A)$ of the equation (3.11) in the form $f(\theta, \widetilde{y}) = e^{in\theta} g(\widetilde{y})$ (again, for $n \neq 0$ it should be $g(0) = \lim\limits_{\widetilde{y} \to 0} g(\widetilde{y}) = 0$).

The differential equation for $g(\widetilde{y})$ following from (3.11) is the ordinary differential equation

$$(\lambda + \frac{n^2}{\widetilde{\lambda}^2(\widetilde{y})}) g(\widetilde{y}) - D_{\widetilde{v}(\widetilde{y})} D_{\widetilde{u}(\widetilde{y})} g(\widetilde{y}) = G(\widetilde{y})\ , \tag{3.12}$$

and it should be solved with the boundary conditions $\dfrac{n^2}{\widetilde{\lambda}^2(a)} g(a) - D_{\widetilde{v}(\widetilde{y})} D_{\widetilde{u}(\widetilde{y})} g(a) = \dfrac{n^2}{\widetilde{\lambda}^2(b)} g(b) - D_{\widetilde{v}(\widetilde{y})} D_{\widetilde{u}(\widetilde{y})} g(b) = 0$, $D^-_{\widetilde{u}(\widetilde{y})} g(0) = D^+_{\widetilde{u}(\widetilde{y})} g(0)$ and for $n \neq 0$, $g(0) = 0$. From the boundary conditions we get at once $g(a) = \lambda^{-1} G(a)$ and $g(b) = \lambda^{-1} G(b)$.

For $n = 0$ the equation (3.12) with the boundary conditions $D_{\widetilde{u}(\widetilde{y})} D_{\widetilde{v}(\widetilde{y})} g(a) = D_{\widetilde{u}(\widetilde{y})} D_{\widetilde{v}(\widetilde{y})} g(b) = 0$ and the gluing condition $D^-_{\widetilde{u}(\widetilde{y})} g(0) = D^+_{\widetilde{u}(\widetilde{y})} g(0)$ is just the ordinary differential equation for a one-dimensional diffusion process that has been considered infinitely many times, and it has a solution for every $G \in \mathbf{C}[a, b]$. Let us go to the case $n \neq 0$. We are going to consider the intervals $[a, 0)$ and $(0, b]$ separately; what follows is about the interval $(0, b]$.

Similarly to how it is done in, e.g.[7], we can prove that there exist two non-negative solutions $\xi_1(\widetilde{y})$ and $\xi_2(\widetilde{y})$ of the equation

$$(\lambda + \frac{n^2}{\widetilde{\lambda}^2(\widetilde{y})}) \xi_i(\widetilde{y}) - D_{\widetilde{v}(\widetilde{y})} D_{\widetilde{u}(\widetilde{y})} \xi_i(\widetilde{y}) = 0\ ,\ 0 < \widetilde{y} \leq b\ , \tag{3.13}$$

the first one increasing and the second one decreasing, $\xi_1(0) = \xi_2(b) = 0$, $\xi_1(b) < \infty$, $\xi_2(0+) = \infty$. The derivatives $D_{\widetilde{u}(\widetilde{y})}\xi_i(\widetilde{y})$ are increasing, $D_{\widetilde{u}(\widetilde{y})}\xi_1(0) = 0$, $D_{\widetilde{u}(\widetilde{y})}\xi_2(b) < 0$.

It is easily checked that the Wronskian

$$W(\widetilde{y}) = \det \begin{pmatrix} D_{\widetilde{u}(\widetilde{y})}\xi_1(\widetilde{y}) & D_{\widetilde{u}(\widetilde{y})}\xi_2(\widetilde{y}) \\ \xi_1(\widetilde{y}) & \xi_2(\widetilde{y}) \end{pmatrix}$$

(both summands $D_{\widetilde{u}(\widetilde{y})}\xi_1(\widetilde{y}) \cdot \xi_2(\widetilde{y})$ and $-D_{\widetilde{u}(\widetilde{y})}\xi_2(\widetilde{y}) \cdot \xi_1(\widetilde{y})$ are positive) does not depend on $\widetilde{y}$: $W(\widetilde{y}) \equiv W > 0$.

Now we define, for $\widetilde{y} \in [0, b]$,

$$\widetilde{g}(\widetilde{y}) = \frac{1}{W}\left[\xi_2(\widetilde{y})\int_0^{\widetilde{y}} \xi_1(z) \cdot G(z) d\widetilde{v}(z) + \xi_1(\widetilde{y})\int_{\widetilde{y}}^b \xi_2(z) \cdot G(z) d\widetilde{v}(z)\right] . \tag{3.14}$$

It is easily checked that $\lambda\widetilde{g}(\widetilde{y}) - A\widetilde{g}(\widetilde{y}) = G(\widetilde{y})$ for $0 < \widetilde{y} \le b$.

Of course

$$|\widetilde{g}(\widetilde{y})| \le \frac{\|G\|}{W}\left[\xi_2(\widetilde{y})\int_0^{\widetilde{y}} \xi_1(z) d\widetilde{v}(z) + \xi_1(\widetilde{y})\int_{\widetilde{y}}^b \xi_2(z) d\widetilde{v}(z)\right] . \tag{3.15}$$

Let us check that this goes to 0 as $y \to 0+$.

We have:

$$\xi_i(z) = \frac{D_{\widetilde{v}(\widetilde{y})}D_{\widetilde{u}(\widetilde{y})}\xi_i(z)}{\lambda + n^2/\widetilde{\lambda}^2(z)}$$

so the first summand in the brackets in (3.15) is less or equal

$$\xi_2(\widetilde{y}) \cdot \frac{D_{\widetilde{u}(\widetilde{y})}\xi_1(\widetilde{y}) - D_{\widetilde{u}(\widetilde{y})}\xi_1(0)}{\min\limits_{0\le z\le\widetilde{y}}[\lambda + n^2/\widetilde{\lambda}^2(z)]} = \frac{\xi_2(\widetilde{y}) \cdot D_{\widetilde{u}(\widetilde{y})}\xi_1(\widetilde{y})}{\min\limits_{0\le z\le\widetilde{y}}[\lambda + n^2/\widetilde{\lambda}^2(z)]} < \frac{W}{\min\limits_{0\le z\le\widetilde{y}}[\lambda + n^2/\widetilde{\lambda}^2(z)]} ,$$

and it goes to zero as $\widetilde{y} \to 0+$.

The second summand in (3.15) is less or equal

$$\xi_1(\widetilde{y}) \cdot \frac{D_{\widetilde{u}(\widetilde{y})}\xi_2(c) - D_{\widetilde{u}(\widetilde{y})}\xi_2(\widetilde{y})}{\min\limits_{\widetilde{y}\le z\le c}[\lambda + n^2/\widetilde{\lambda}^2(z)]} + \xi_1(\widetilde{y}) \cdot \frac{D_{\widetilde{u}(\widetilde{y})}\xi_2(b) - D_{\widetilde{u}(\widetilde{y})}\xi_2(c)}{\min\limits_{c\le z\le b}[\lambda + n^2/\widetilde{\lambda}^2(z)]} , \tag{3.16}$$

where $\widetilde{y} < c < b$. The first term in (3.16) is less or equal

$$\frac{-\xi_1(y) \cdot D_{\widetilde{u}(\widetilde{y})}\xi_2(\widetilde{y})}{\min\limits_{\widetilde{y}\le z\le c}[\lambda + n^2/\widetilde{\lambda}^2(z)]} \le \frac{W}{\min\limits_{\widetilde{y}\le z\le c}[\lambda + n^2/\widetilde{\lambda}^2(z)]} ,$$

and it can be made arbitrarily small by choosing a positive c close enough to 0. The second term in (3.16), for a fixed $c > 0$, converges to 0 as $\widetilde{y} \to 0+$. So we get that $\lim\limits_{\widetilde{y}\to 0+} \widetilde{g}(\widetilde{y}) = 0$.

Now we are going to find $D_{\widetilde{u}(\widetilde{y})}\widetilde{g}(0+)$. We have:

$$D_{\widetilde{u}(\widetilde{y})}\widetilde{g}(\widetilde{y}) = \frac{1}{W}\left[D_{\widetilde{u}(\widetilde{y})}\xi_1(\widetilde{y})\int_{\widetilde{y}}^{b}\xi_2(z)\cdot G(z)d\widetilde{v}(z) + D_{\widetilde{u}(\widetilde{y})}\xi_2(\widetilde{y})\int_0^{\widetilde{y}}\xi_1(z)\cdot G(z)d\widetilde{v}(z)\right] . \tag{3.17}$$

The first integral here is equal to $\int_{\widetilde{y}}^{c} + \int_{c}^{b}$, and it is not greater than

$$\|G\| \cdot [\xi_2(\widetilde{y}) \cdot \widetilde{v}(c) + \xi_2(c) \cdot \widetilde{v}(b)] ,$$

and the first summand is not greater than

$$\|G\|/W \cdot [W \cdot \widetilde{v}(c) + \xi_2(c) \cdot \widetilde{v}(b) \cdot D_{\widetilde{u}(\widetilde{y})}\xi_1(\widetilde{y})] .$$

By choosing $c \in (0, b)$ close enough to 0 we make $\widetilde{v}(c)$ arbitrarily small; and we know $D_{\widetilde{u}(\widetilde{y})}\xi_1(\widetilde{y}) \to 0$ as $\widetilde{y} \to 0+$. So the first summand in (3.17) goes to 0 as $\widetilde{y} \to 0+$.

The second summand in (3.17) does not exceed in absolute value

$$\|G\| \cdot \xi_1(\widetilde{y}) \cdot |D_{\widetilde{u}(\widetilde{y})}\xi_2(\widetilde{y})| \cdot \widetilde{v}(\widetilde{y}) \leq \|G\| \cdot W \cdot \widetilde{v}(\widetilde{y}) \to 0 \quad (\widetilde{y} \to 0+) .$$

Now we are looking for the solution $g(\widetilde{y})$ of the equation (3.12) with the boundary conditions under this formula in the form $g(\widetilde{y}) = \widetilde{g}(\widetilde{y}) + C \cdot \xi_1(\widetilde{y})$. For the undetermined coefficient C we get one linear equation, and it does have a solution since $\xi_1(b) \neq 0$.

The same way we get, for $n \neq 0$, a solution $g(\widetilde{y})$ for $\widetilde{y} < 0$ with $g(0-) = D_{\widetilde{u}(\widetilde{y})}g(0-) = 0$, $g(a) = \mu^{-1}G(a)$.

So we get a solution $f \in D(A)$ of the equation (3.11) for every function $F(\theta, \widetilde{y}) = \sum_{n=-N}^{N} e^{in\theta} \cdot G_n(\widetilde{y})$, $G_n(\widetilde{y}) \in \mathbf{C}[a, b]$, such that $G_n(0) = 0$ for $n \neq 0$ (we take $f(\mathfrak{o}) = G_0(0)$). The set of such functions is dense in $\mathbf{C}(\mathfrak{C})$ so that the closure operator $\overline{(\lambda I - A|_{D(A)})^{-1}}$ is defined on the whole $\mathbf{C}(\mathfrak{C})$ which finishes the proof. $\square$

Let $\widetilde{\boldsymbol{q}}_t$ be the Markov process corresponding to $\overline{A|_{D(A)}}$, whose existence was proved in Lemma 3.1. We prove the following

Theorem 3.1. *As $\varepsilon \downarrow 0$, for fixed $T > 0$, the process $\widetilde{\boldsymbol{q}}_t^{\varepsilon} = \boldsymbol{\pi}(\boldsymbol{q}_t^{\varepsilon})$ converges weakly in the space $\mathbf{C}_{[0,T]}(\mathfrak{C})$ to the process $\widetilde{\boldsymbol{q}}_t$.*

The proof is again based on an application of Lemma 2.2.

Proof of Theorem 3.1. Making use of Lemma 2.2, we take the metric space $M = S^1 \times [a-1, b+1]$ with standard metric. The mapping $Y = \boldsymbol{\pi}$. The space $Y(M) = \mathfrak{C}$

is endowed with the metric d, defined as follows. For any two points $(\theta_1, \widetilde{y}_1)$ and $(\theta_2, \widetilde{y}_2)$ on $\mathfrak{C}$ with $\widetilde{y}_1, \widetilde{y}_2$ having the same sign we let $d((\theta_1, \widetilde{y}_1), (\theta_2, \widetilde{y}_2))$ be the Euclidean distance between points $(|\widetilde{y}_1| \cos \theta_1, |\widetilde{y}_1| \sin \theta_1)$ and $(|\widetilde{y}_2| \cos \theta_2, |\widetilde{y}_2| \sin \theta_2)$ in $\mathbb{R}^2$; if $\widetilde{y}_1$ and $\widetilde{y}_2$ have different sign we take $d((\theta_1, \widetilde{y}_1), (\theta_2, \widetilde{y}_2)) = d((\theta_1, \widetilde{y}_1), \mathfrak{o}) + d(\mathfrak{o}, (\theta_2, \widetilde{y}_2))$. With respect to this metric the space $\mathfrak{C}$ is a complete separable metric space. We take the process $(X_t^\varepsilon, \mathbf{P}_x^\varepsilon)$ as $\boldsymbol{q}_t^\varepsilon$ and the process $(y_t, \mathbf{P}_y)$ is taken as $\widetilde{\boldsymbol{q}}_t$.

For the uniqueness of solution of martingale problem we set the space Ψ be the space of all continuous functions on $\mathfrak{C}$ which has the form $F(\theta, \widetilde{y}) = \sum\limits_{n=-N}^{N} e^{in\theta} \cdot G_n(\widetilde{y})$, $G_n \in \mathbf{C}[a, b]$ is continuously differentiable inside $[a, 0)$ and $(0, b]$, also $G_n(0) = 0$ for $n \neq 0$. We take $f(\mathfrak{o}) = G_0(0)$. It is proved in the proof of Lemma 3.1 that the equation $\lambda f - Af = F$ always has a solution $f \in D \subset D(A)$ for all $F \in \Psi$ and $\lambda > 0$. The space D contains those functions $f \in \mathbf{C}(\mathfrak{C})$ that are bounded and are three times continuously differentiable inside $\mathfrak{C}^+ \equiv \{(\theta, \widetilde{y}) \in \mathfrak{C} : a < \widetilde{y} < 0\}$ and $\mathfrak{C}^- \equiv \{(\theta, \widetilde{y}) \in \mathfrak{C} : 0 < \widetilde{y} < b\}$.

We will state pre-compactness of family of distributions of processes $\widetilde{\boldsymbol{q}}_t^\varepsilon$ in Lemma 3.2. What remains to do is to check that for every compact $K \subset \mathfrak{C}$ and for every $f \in D$ and every $\lambda > 0$ we have

$$\mathbf{E}_{\boldsymbol{q}_0} \left[\int_0^\infty e^{-\lambda t} [\lambda f(\boldsymbol{\pi}(\boldsymbol{q}_t^\varepsilon)) - Af(\boldsymbol{\pi}(\boldsymbol{q}_t^\varepsilon))] dt - f(\boldsymbol{\pi}(\boldsymbol{q}_0)) \right] \to 0$$

as $\varepsilon \downarrow 0$ uniformly in $\boldsymbol{q}_0 \in \boldsymbol{\pi}^{-1}(K)$. The proof of this is essentially the same as the proof we did in Lemma 2.5, based on the following auxiliary Lemmas 3.9 (for the proof of convergence for processes near $\mathfrak{o}$) and 3.10 (for the proof of convergence for processes away from $\mathfrak{o}$) and the auxiliary Lemmas 2.9 and 2.10 (for the estimates on the exit times, notice that the stopping times σ_n and τ_n we will work with in this section are essentially the same stopping times that we worked with in Section 2 since we are discussing about a model problem). We omit the details in the proof. $\square$

Let κ be a real number with small absolute value. Let $G(\kappa) = \{(\theta, y) \in S^1 \times [a-1, b+1] : a - 1 \leq y \leq -1 - \kappa \text{ or } 1 + \kappa \leq y \leq b + 1\}$. Let $C^+(\kappa) = \{(\theta, y) \in S^1 \times [a-1, b+1] : y = 1 + \kappa\}$ and $C^-(\kappa) = \{(\theta, y) \in S^1 \times [a-1, b+1] : y = -1 - \kappa\}$. Let $C(\kappa) = C^+(\kappa) \cup C^-(\kappa)$. Let $\delta > \delta' > 0$ be small. We shall introduce a sequence of stopping times $\tau_0 \leq \sigma_0 < \tau_1 < \sigma_1 < \tau_2 < \sigma_2 < ...$ by

$$\tau_0 = 0 \ , \ \sigma_n = \min\{t \geq \tau_n, \boldsymbol{q}_t^\varepsilon \in G(\delta)\} \ , \ \tau_n = \min\{t \geq \sigma_{n-1}, \boldsymbol{q}_t^\varepsilon \in C(\delta')\} \ .$$

This is well-defined up to some σ_k $(k \geq 0)$ such that

$$\mathbf{P}_{y_{\sigma_k}^\varepsilon}(y_{t+\sigma_k}^\varepsilon \text{ hits } a-1 \text{ or } b+1 \text{ before it hits } -1-\delta' \text{ or } 1+\delta') = 1 \ .$$

We will then define $\tau_{k+1} = \min\{t > \sigma_k : y_t^\varepsilon = a-1 \text{ or } b+1\}$. And we define $\tau_{k+1} < \sigma_{k+1} = \tau_{k+1}+1 < \tau_{k+2} = \tau_{k+1}+2 < \sigma_{k+2} = \tau_{k+1}+3 < ...$ and so on.

We have $\lim\limits_{n\to\infty} \tau_n = \lim\limits_{n\to\infty} \sigma_n = \infty$. And we have obvious relations $\boldsymbol{q}_{\tau_n}^\varepsilon \in C(\delta')$, $\boldsymbol{q}_{\sigma_n}^\varepsilon \in C(\delta)$ for $1 \le n \le k$ (as long as $k \ge 1$, if $k=0$ the process may start from $G(\delta)$ and goes directly to $S^1 \times \{a-1\}$ or $S^1 \times \{b+1\}$ without touching $C(\delta')$ and is stopped there, or it may start from $S^1 \times (-1-\delta, 1+\delta)$, reaches $C(\delta)$ first and then goes directly to $S^1 \times \{a-1\}$ or $S^1 \times \{b+1\}$ without touching $C(\delta')$ and is stopped there). Also, for $n \ge k+1$ we have $\boldsymbol{q}_{\tau_n}^\varepsilon = \boldsymbol{q}_{\sigma_n}^\varepsilon \in S^1 \times \{a-1\}$ or $S^1 \times \{b+1\}$. If $\boldsymbol{q}_0^\varepsilon = \boldsymbol{q}_0 \in G(\delta)$, then we have $\sigma_0 = 0$ and τ_1 is the first time at which the process $\boldsymbol{q}_t^\varepsilon$ reaches $C(\delta')$ or $S^1 \times \{a-1\}$ or $S^1 \times \{b+1\}$.

Note that these stopping times are the same as those defined in Section 2 since our process y_t^ε is essentially the process q_t^ε in Section 2.

The pre-compactness of the family $\{\widetilde{\boldsymbol{q}}_t^\varepsilon\}_{\varepsilon>0}$ in $\mathbf{C}_{[0,T]}(\mathfrak{C})$ for $0 < T < \infty$ is proved in the same way as in the one-dimensional case. We shall make use of the technical Lemma 2.3 with $\widetilde{q}_\bullet^{\varepsilon,\delta}$ and $\widetilde{q}_\bullet^\varepsilon$ replaced by $\widetilde{\boldsymbol{q}}_\bullet^{\varepsilon,\delta}$ and $\widetilde{\boldsymbol{q}}_\bullet^\varepsilon$ and the space $\mathbf{C}_{[0,T]}(\mathfrak{C})$ instead of $\mathbf{C}_{[0,T]}([a,b])$. We omit the proof of the next lemma.

Lemma 3.2. *The family of distributions of $\{\widetilde{\boldsymbol{q}}_t^\varepsilon\}_{\varepsilon>0}$ is pre-compact in $\mathbf{C}_{[0,T]}(\mathfrak{C})$.*

The next few lemmas establish the estimates on the asymptotic joint law of the processes $(y_t^\varepsilon, \theta_t^\varepsilon)$ at first exit from a small neighborhood of the domain within which the friction vanishes. This is the key part to the proof of Theorem 3.1.

Let $\delta'' > 0$ be small. We consider the process $\boldsymbol{q}_t^\varepsilon$ starting from $\boldsymbol{q}_0^\varepsilon = \boldsymbol{q}_0 \in S^1 \times [-1-\delta', 1+\delta']$. Let us introduce another sequence of stopping times $\alpha_1 < \beta_1 < \alpha_2 < \beta_2 < ... < \alpha_{n(\varepsilon)}$ by

$$\alpha_1 = \min\{0 \le t < \sigma_0 : \boldsymbol{q}_t^\varepsilon \in C(0)\} \ , \ \beta_1 = \min\{\alpha_1 < t < \sigma_0 : \boldsymbol{q}_t^\varepsilon \in C(-\delta'')\} \ ,$$

and for $k \ge 2$ we define

$$\alpha_k = \min\{\beta_{k-1} < t < \sigma_0 : \boldsymbol{q}_t^\varepsilon \in C(0)\} \ , \ \beta_k = \min\{\alpha_k < t < \sigma_0 : \boldsymbol{q}_t^\varepsilon \in C(-\delta'')\} \ .$$

Here we take the convention that the minimum over an empty set is ∞. The number $n(\varepsilon)$ is a non-negative integer-valued random variable such that $\alpha_{n(\varepsilon)} < \infty$ and $\beta_{n(\varepsilon)} = \infty$. If $\alpha_1 = \infty$ we set $n(\varepsilon) = 0$.

Lemma 3.3. *For $\boldsymbol{q}_0 \in G(\delta')$ we have*

$$\mathbf{P}_{\boldsymbol{q}_0}(\alpha_1 < \infty) \ge 1 - \max\left(\frac{\widetilde{u}(\delta') + \varepsilon\delta'}{\widetilde{u}(\delta) + \varepsilon\delta}, \frac{-\widetilde{u}(-\delta') + \varepsilon\delta'}{-\widetilde{u}(-\delta) + \varepsilon\delta}\right) \ . \tag{3.18}$$

Proof. If $1 \leq y_0^\varepsilon = y_0 \leq 1+\delta'$ we have

$$\mathbf{P}_{\boldsymbol{q}_0}(\alpha_1<\infty) = \frac{u^\varepsilon(1+\delta)-u^\varepsilon(y)}{u^\varepsilon(1+\delta)-u^\varepsilon(1)} \geq \frac{u^\varepsilon(1+\delta)-u^\varepsilon(1+\delta')}{u^\varepsilon(1+\delta)-u^\varepsilon(1)} = 1-\frac{\widetilde{u}(\delta')+\varepsilon\delta'}{\widetilde{u}(\delta)+\varepsilon\delta} \ .$$

If $-1-\delta' \leq y_0^\varepsilon = y_0 \leq -1$ we have

$$\mathbf{P}_{\boldsymbol{q}_0}(\alpha_1<\infty) = \frac{u^\varepsilon(y)-u^\varepsilon(-1-\delta')}{u^\varepsilon(-1)-u^\varepsilon(-1-\delta)} \geq \frac{u^\varepsilon(-1-\delta')-u^\varepsilon(-1-\delta)}{u^\varepsilon(-1)-u^\varepsilon(-1-\delta)} = 1-\frac{-\widetilde{u}(-\delta')+\varepsilon\delta'}{-\widetilde{u}(-\delta)+\varepsilon\delta} \ .$$

If $-1 < y_0^\varepsilon = y_0 < 1$ we have $\mathbf{P}_{\boldsymbol{q}_0}(\alpha_1<\infty)=1$. $\square$

Lemma 3.4. *For $\boldsymbol{q}_0 \in G(\delta')$ we have*

$$\mathbf{P}_{\boldsymbol{q}_0}(\beta_1<\infty|\alpha_1<\infty) \geq 1-\max\left(\frac{\varepsilon\delta''}{\widetilde{u}(\delta)+\varepsilon(\delta+\delta'')}, \frac{\varepsilon\delta''}{-\widetilde{u}(-\delta)+\varepsilon(\delta+\delta'')}\right) \ . \qquad (3.19)$$

Proof. If $y_{\alpha_1}^\varepsilon = 1$ we have

$$\mathbf{P}_{\boldsymbol{q}_0}(\beta_1<\infty|\alpha_1<\infty) = \frac{u^\varepsilon(1+\delta)-u^\varepsilon(1)}{u^\varepsilon(1+\delta)-u^\varepsilon(1-\delta'')} = 1-\frac{\varepsilon\delta''}{\widetilde{u}(\delta)+\varepsilon(\delta+\delta'')} \ .$$

If $y_{\alpha_1}^\varepsilon = -1$ we have

$$\mathbf{P}_{\boldsymbol{q}_0}(\beta_1<\infty|\alpha_1<\infty) = \frac{u^\varepsilon(-1)-u^\varepsilon(-1-\delta)}{u^\varepsilon(-1+\delta'')-u^\varepsilon(-1-\delta)} = 1-\frac{\varepsilon\delta''}{-\widetilde{u}(-\delta)+\varepsilon(\delta+\delta'')} \ .$$

$\square$

Let $M(\varepsilon)\to\infty$ as $\varepsilon\downarrow 0$ be an integer. The exact asymptotics of $M(\varepsilon)$ will be specified later. We prove

Lemma 3.5. *For $\boldsymbol{q}_0 \in G(\delta')$ we have*

$$\mathbf{P}_{\boldsymbol{q}_0}(n(\varepsilon)\geq M(\varepsilon)|\alpha_1<\infty) \geq \left[1-\max\left(\frac{\varepsilon\delta''}{\widetilde{u}(\delta)+\varepsilon(\delta+\delta'')}, \frac{\varepsilon\delta''}{-\widetilde{u}(-\delta)+\varepsilon(\delta+\delta'')}\right)\right]^{M(\varepsilon)-1} \ . \qquad (3.20)$$

Proof. This is because trajectories of $\boldsymbol{q}_t^\varepsilon$ between times $\alpha_i \leq t < \alpha_{i+1}$ are independent and by iteratively using Lemma 3.4 we get the desired result. $\square$

Lemma 3.6. *We have*

$$\alpha_{i+1} - \beta_i \geq \varepsilon^2 \left(\frac{\delta''}{H_i}\right)^5 \tag{3.21}$$

with H_i being i.i.d. positive random variables with $\mathbf{E}(H_i)^4 < \infty$ for $i = 1, 2, ..., n(\varepsilon) - 1$.

Proof. This is a result of the Hölder continuity of the standard Wiener trajectory $|W_t - W_s| \leq H_i |t-s|^{1/5}$ and the fact that between times $\beta_i \leq t < \alpha_{i+1}$ the process y_t^ε is a time-changed Wiener process $\dfrac{1}{\varepsilon} W_t$ traveling at least a distance of δ''. $\square$

Let us define an auxiliary function

$$\begin{aligned}
&\Omega(\varepsilon, \delta, \delta', \delta'', M(\varepsilon)) \\
&\equiv 2\left[1 - \left[1 - \max\left(\frac{\varepsilon\delta''}{\widetilde{u}(\delta) + \varepsilon(\delta+\delta'')}, \frac{\varepsilon\delta''}{-\widetilde{u}(-\delta) + \varepsilon(\delta+\delta'')}\right)\right]^{M(\varepsilon)-1} + \right. \\
&\qquad\qquad \left. 2\max\left(\frac{\widetilde{u}(\delta') + \varepsilon\delta'}{\widetilde{u}(\delta) + \varepsilon\delta}, \frac{-\widetilde{u}(-\delta') + \varepsilon\delta'}{-\widetilde{u}(-\delta) + \varepsilon\delta}\right)\right] .
\end{aligned}$$

Lemma 3.7. *For $\boldsymbol{q}_0 \in G(\delta')$ and for some $A > 0$, $\kappa > 0$ and $C > 0$, there exists $\varepsilon_0 > 0$ such that for all $0 < \varepsilon < \varepsilon_0$, for any $0 \leq \theta_1 \leq \theta_2 \leq 2\pi$ we have*

$$\begin{aligned}
&\left|\mathbf{P}_{\boldsymbol{q}_0}(\theta_{\sigma_0}^\varepsilon \in [\theta_1, \theta_2], y_{\sigma_0}^\varepsilon = 1 + \delta) - \frac{\theta_2 - \theta_1}{2\pi}\mathbf{P}_{\boldsymbol{q}_0}(y_{\sigma_0}^\varepsilon = 1 + \delta)\right| \\
&\leq C\exp(-A(\delta'')^5 \kappa M(\varepsilon)) + 2\Omega(\varepsilon, \delta, \delta', \delta'', M(\varepsilon))
\end{aligned}$$

and

$$\begin{aligned}
&\left|\mathbf{P}_{\boldsymbol{q}_0}(\theta_{\sigma_0}^\varepsilon \in [\theta_1, \theta_2], y_{\sigma_0}^\varepsilon = -1 - \delta) - \frac{\theta_2 - \theta_1}{2\pi}\mathbf{P}_{\boldsymbol{q}_0}(y_{\sigma_0}^\varepsilon = -1 - \delta)\right| \\
&\leq C\exp(-A(\delta'')^5 \kappa M(\varepsilon)) + 2\Omega(\varepsilon, \delta, \delta', \delta'', M(\varepsilon)) .
\end{aligned}$$

Proof. As we have

$$x_t^\varepsilon = \int_0^t \frac{1}{\lambda(y_t^\varepsilon) + \varepsilon} dW_s^1 = W^1\left(\int_0^t \frac{ds}{(\lambda(y_s^\varepsilon) + \varepsilon)^2}\right) ,$$

we set $T^\varepsilon(t) = \displaystyle\int_0^t \frac{ds}{(\lambda(y_s^\varepsilon) + \varepsilon)^2}$. Using Lemma 3.6 for $\boldsymbol{q}_0 \in G(\delta')$ the random time $T^\varepsilon(\sigma_0)$ can be estimated from below by

$$T^\varepsilon(\sigma_0) \geq \int_0^{\sigma_0} \frac{ds}{(\lambda(y_s^\varepsilon) + \varepsilon)^2} \geq \frac{1}{\varepsilon^2}\int_0^{\sigma_0} \mathbf{1}_{\{-1 \leq y_s^\varepsilon \leq 1\}} ds \geq \frac{1}{\varepsilon^2}\sum_{i=1}^{n(\varepsilon)-1}(\alpha_{i+1} - \beta_i) \geq (\delta'')^5 \sum_{i=1}^{n(\varepsilon)-1} \frac{1}{(H_i)^5} .$$

(If $n(\varepsilon) = 0, 1$ the sum is supposed to be 0.)

And we also notice that the random time $T^\varepsilon(\sigma_0)$ only depends on the behavior of the process y_t^ε and is therefore independent of the Wiener process W_t^1 in the stochastic differential equation $\dot{x}_t^\varepsilon = \dfrac{1}{\lambda(y_t^\varepsilon)+\varepsilon}\dot{W}_t^1$ (see (3.2)). For the same reason the random variables $y_{\sigma_0}^\varepsilon$, $n(\varepsilon)$ and α_1 are of course also independent of W_t^1.

As we have the elementary inequality $\left(\mathbf{E}\dfrac{1}{(H_i)^5}\right)^{1/5}(\mathbf{E}(H_i)^4)^{1/4} \geq \left(\mathbf{E}\dfrac{1}{H_i}\right)(\mathbf{E}H_i) \geq 1$, we have, by Strong Law of Large Numbers

$$\lim_{\varepsilon\downarrow 0}\frac{1}{M(\varepsilon)-1}\sum_{i=1}^{M(\varepsilon)-1}\frac{1}{(H_i)^5} = \mathbf{E}\left(\frac{1}{(H_i)^5}\right) \geq \frac{1}{(\mathbf{E}(H_i)^4)^{5/4}} \geq c > 0 \ \text{ a. s.}$$

for some constant $c > 0$. (We can always assume that H_i is uniformly bounded from below by a positive constant so that $\left(\mathbf{E}\dfrac{1}{(H_i)^5}\right) < \infty$ and we can apply SLLN.)

Now we see that we can find some $\varepsilon_0 > 0$ such that for all $0 < \varepsilon < \varepsilon_0$ we will have

$$\mathbf{P}_{\boldsymbol{q}_0}(T^\varepsilon(\sigma_0) \geq (\delta'')^5\kappa M(\varepsilon) | n(\varepsilon) \geq M(\varepsilon), \alpha_1 < \infty) = 1$$

for some constant $\kappa > 0$.

This gives

$$\begin{aligned}&\mathbf{P}_{\boldsymbol{q}_0}(T^\varepsilon(\sigma_0) \geq (\delta'')^5\kappa M(\varepsilon), y_{\sigma_0}^\varepsilon = 1+\delta | n(\varepsilon) \geq M(\varepsilon), \alpha_1 < \infty)\\ &= \mathbf{P}_{\boldsymbol{q}_0}(y_{\sigma_0}^\varepsilon = 1+\delta | n(\varepsilon) \geq M(\varepsilon), \alpha_1 < \infty)\ .\end{aligned}$$

Recall that we have $\theta_{\sigma_0}^\varepsilon = x_{\sigma_0}^\varepsilon \mod 2\pi = W^1_{T^\varepsilon(\sigma_0)} \mod 2\pi$. Using this, the independence of $T^\varepsilon(\sigma_0)$, $y_{\sigma_0}^\varepsilon$, α_1 and $n(\varepsilon)$ with W_t^1, and the above estimates we have, as $0 < \varepsilon < \varepsilon_0$, that

$$\begin{aligned}&\mathbf{P}_{\boldsymbol{q}_0}(\theta_{\sigma_0}^\varepsilon \in [\theta_1,\theta_2], y_{\sigma_0}^\varepsilon = 1+\delta | n(\varepsilon) \geq M(\varepsilon), \alpha_1 < \infty)\\ &= \int_0^\infty \mathbf{P}_{\boldsymbol{q}_0}(T^\varepsilon(\sigma_0) \in dt, y_{\sigma_0}^\varepsilon = 1+\delta | n(\varepsilon) \geq M(\varepsilon), \alpha_1 < \infty)\mathbf{P}_{\boldsymbol{q}_0}(W_t^1 \mod 2\pi \in [\theta_1,\theta_2])\\ &= \int_{(\delta'')^5\lambda M(\varepsilon)}^\infty \mathbf{P}_{\boldsymbol{q}_0}(T^\varepsilon(\sigma_0) \in dt, y_{\sigma_0}^\varepsilon = 1+\delta | n(\varepsilon) \geq M(\varepsilon), \alpha_1 < \infty)\mathbf{P}_{\boldsymbol{q}_0}(W_t^1 \mod 2\pi \in [\theta_1,\theta_2])\ .\end{aligned}$$

Since we have the exponential decay

$$\left|\mathbf{P}(W_t^1 \mod 2\pi \in [\theta_1,\theta_2]) - \frac{\theta_2-\theta_1}{2\pi}\right| < C\exp(-At)$$

for some $C > 0$ and $A > 0$, we could estimate

$$\begin{aligned}&\big|\mathbf{P}_{\boldsymbol{q}_0}(\theta_{\sigma_0}^\varepsilon \in [\theta_1,\theta_2]\ ,\ y_{\sigma_0}^\varepsilon = 1+\delta | n(\varepsilon) \geq M(\varepsilon), \alpha_1 < \infty) -\\ &\qquad \frac{\theta_2-\theta_1}{2\pi}\,\mathbf{P}_{\boldsymbol{q}_0}(y_{\sigma_0}^\varepsilon = 1+\delta | n(\varepsilon) \geq M(\varepsilon), \alpha_1 < \infty)\big|\\ &< C\exp(-A(\delta'')^5\kappa M(\varepsilon))\end{aligned}$$

for $0 < \varepsilon < \varepsilon_0$.

Notice that we have, by using Lemmas 3.5 and 3.3,

$$
\begin{aligned}
&\left|\mathbf{P}_{\boldsymbol{q}_0}(\theta^\varepsilon_{\sigma_0} \in [\theta_1, \theta_2], y^\varepsilon_{\sigma_0} = 1+\delta) - \mathbf{P}_{\boldsymbol{q}_0}(\theta^\varepsilon_{\sigma_0} \in [\theta_1, \theta_2]\ ,\ y^\varepsilon_{\sigma_0} = 1+\delta | n(\varepsilon) \geq M(\varepsilon), \alpha_1 < \infty)\right| \\
&= \left|\mathbf{P}_{\boldsymbol{q}_0}(\theta^\varepsilon_{\sigma_0} \in [\theta_1, \theta_2], y^\varepsilon_{\sigma_0} = 1+\delta | n(\varepsilon) \geq M(\varepsilon), \alpha_1 < \infty)\mathbf{P}(n(\varepsilon) \geq M(\varepsilon), \alpha_1 < \infty) - \right. \\
&\qquad \left. \mathbf{P}_{\boldsymbol{q}_0}(\theta^\varepsilon_{\sigma_0} \in [\theta_1, \theta_2]\ ,\ y^\varepsilon_{\sigma_0} = 1+\delta | n(\varepsilon) \geq M(\varepsilon), \alpha_1 < \infty)\right| + \mathbf{P}_{\boldsymbol{q}_0}(n(\varepsilon) < M(\varepsilon)) + \mathbf{P}_{\boldsymbol{q}_0}(\alpha_1 = \infty) \\
&\leq 2(\mathbf{P}_{\boldsymbol{q}_0}(n(\varepsilon) < M(\varepsilon)) + \mathbf{P}_{\boldsymbol{q}_0}(\alpha_1 = \infty)) \\
&\leq 2(\mathbf{P}_{\boldsymbol{q}_0}(n(\varepsilon) < M(\varepsilon) | \alpha_1 < \infty) + 2\mathbf{P}_{\boldsymbol{q}_0}(\alpha_1 = \infty)) \\
&\leq 2\left[1 - \left[1 - \max\left(\frac{\varepsilon\delta''}{\widetilde{u}(\delta) + \varepsilon(\delta+\delta'')}, \frac{\varepsilon\delta''}{-\widetilde{u}(-\delta) + \varepsilon(\delta+\delta'')}\right)\right]^{M(\varepsilon)-1} + \right. \\
&\qquad\qquad \left. 2\max\left(\frac{\widetilde{u}(\delta') + \varepsilon\delta'}{\widetilde{u}(\delta) + \varepsilon\delta}, \frac{-\widetilde{u}(-\delta') + \varepsilon\delta'}{-\widetilde{u}(-\delta) + \varepsilon\delta}\right)\right] \\
&= \Omega(\varepsilon, \delta, \delta', \delta'', M)\ .
\end{aligned}
$$

By the same argument we can estimate

$$
\left|\frac{\theta_2 - \theta_1}{2\pi}\mathbf{P}_{\boldsymbol{q}_0}(y^\varepsilon_{\sigma_0} = 1+\delta) - \frac{\theta_2 - \theta_1}{2\pi}\mathbf{P}_{\boldsymbol{q}_0}(y^\varepsilon_{\sigma_0} = 1+\delta | n(\varepsilon) \geq M(\varepsilon), \alpha_1 < \infty)\right| \leq \Omega(\varepsilon, \delta, \delta', \delta'', M)\ .
$$

Summing up these estimates we have

$$
\begin{aligned}
&\left|\mathbf{P}_{\boldsymbol{q}_0}(\theta^\varepsilon_{\sigma_0} \in [\theta_1, \theta_2], y^\varepsilon_{\sigma_0} = 1+\delta) - \frac{\theta_2 - \theta_1}{2\pi}\mathbf{P}_{\boldsymbol{q}_0}(y^\varepsilon_{\sigma_0} = 1+\delta)\right| \\
&\leq \left|\mathbf{P}_{\boldsymbol{q}_0}(\theta^\varepsilon_{\sigma_0} \in [\theta_1, \theta_2], y^\varepsilon_{\sigma_0} = 1+\delta) - \mathbf{P}_{\boldsymbol{q}_0}(\theta^\varepsilon_{\sigma_0} \in [\theta_1, \theta_2]\ ,\ y^\varepsilon_{\sigma_0} = 1+\delta | n(\varepsilon) \geq M(\varepsilon), \alpha_1 < \infty)\right| + \\
&\quad \left|\mathbf{P}_{\boldsymbol{q}_0}(\theta^\varepsilon_{\sigma_0} \in [\theta_1, \theta_2]\ ,\ y^\varepsilon_{\sigma_0} = 1+\delta | n(\varepsilon) \geq M(\varepsilon), \alpha_1 < \infty) - \right. \\
&\qquad\qquad\qquad \left. \frac{\theta_2 - \theta_1}{2\pi}\ \mathbf{P}_{\boldsymbol{q}_0}(y^\varepsilon_{\sigma_0} = 1+\delta | n(\varepsilon) \geq M(\varepsilon), \alpha_1 < \infty)\right| + \\
&\quad \left|\frac{\theta_2 - \theta_1}{2\pi}\mathbf{P}_{\boldsymbol{q}_0}(y^\varepsilon_{\sigma_0} = 1+\delta) - \frac{\theta_2 - \theta_1}{2\pi}\mathbf{P}_{\boldsymbol{q}_0}(y^\varepsilon_{\sigma_0} = 1+\delta | n(\varepsilon) \geq M(\varepsilon), \alpha_1 < \infty)\right| \\
&\leq 2\Omega(\varepsilon, \delta, \delta', \delta'', M) + C\exp(-A(\delta'')^5\kappa M(\varepsilon))\ ,
\end{aligned}
$$

as desired. The other inequality is established in a similar way. $\square$

Combining Lemma 3.7 and Lemma 2.7 we can have

Lemma 3.8. *For $\boldsymbol{q}_0 \in G(\delta')$ and for some $A > 0$, $\kappa > 0$ and $C_1, C_2 > 0$, there exists $\varepsilon_0 > 0$ such that for all $0 < \varepsilon < \varepsilon_0$, for any $0 \leq \theta_1 \leq \theta_2 \leq 2\pi$ we have*

$$
\begin{aligned}
&\left|\mathbf{P}_{\boldsymbol{q}_0}(\theta^\varepsilon_{\sigma_0} \in [\theta_1, \theta_2], y^\varepsilon_{\sigma_0} = 1+\delta) - \frac{\theta_2 - \theta_1}{2\pi}\frac{\widetilde{u}(0) - \widetilde{u}(-\delta)}{\widetilde{u}(\delta) - \widetilde{u}(-\delta)}\right| \\
&\leq C_1\exp(-A(\delta'')^5\kappa M(\varepsilon)) + 2\Omega(\varepsilon, \delta, \delta', \delta'', M(\varepsilon)) + \frac{\widetilde{u}(\delta') - \widetilde{u}(0) + C_2\varepsilon}{\widetilde{u}(\delta) - \widetilde{u}(-\delta)} \equiv \rho(\varepsilon)\ ,
\end{aligned}
\tag{3.22}
$$

and

$$\left|\mathbf{P}_{\boldsymbol{q}_0}(\theta^\varepsilon_{\sigma_0}\in[\theta_1,\theta_2], y^\varepsilon_{\sigma_0}=-1-\delta)-\frac{\theta_2-\theta_1}{2\pi}\frac{\widetilde{u}(\delta)-\widetilde{u}(0)}{\widetilde{u}(\delta)-\widetilde{u}(-\delta)}\right|$$
$$\leq C_1\exp(-A(\delta'')^5\kappa M(\varepsilon))+2\Omega(\varepsilon,\delta,\delta',\delta'',M(\varepsilon))+\frac{\widetilde{u}(\delta')-\widetilde{u}(0)+C_2\varepsilon}{\widetilde{u}(\delta)-\widetilde{u}(-\delta)}\equiv\rho(\varepsilon)\ . \tag{3.23}$$

Now let us specify the asymptotic order of $M(\varepsilon)\to\infty$, $\delta=\delta(\varepsilon)\to 0$, $\delta'=\delta'(\varepsilon)\to 0$ and $\delta''=\delta''(\varepsilon)\to 0$ as $\varepsilon\downarrow 0$. Since for $0<\kappa<1$ we have the elementary estimate $1-(1-\kappa)^n=\kappa(1+(1-\kappa)+...+(1-\kappa)^{n-1})\leq\kappa n$ we can estimate

$$\begin{aligned}&\Omega(\varepsilon,\delta,\delta',\delta'',M(\varepsilon))\\&\leq 2\left[M(\varepsilon)\cdot\max\left(\frac{\varepsilon\delta''}{\widetilde{u}(\delta)+\varepsilon(\delta+\delta'')},\frac{\varepsilon\delta''}{-\widetilde{u}(-\delta)+\varepsilon(\delta+\delta'')}\right)+\right.\\&\qquad\left.2\max\left(\frac{\widetilde{u}(\delta')+\varepsilon\delta'}{\widetilde{u}(\delta)+\varepsilon\delta},\frac{-\widetilde{u}(-\delta')+\varepsilon\delta'}{-\widetilde{u}(-\delta)+\varepsilon\delta}\right)\right]\ .\end{aligned}$$

We shall choose $\delta''=\delta''(\varepsilon)<<\delta$ and $M(\varepsilon)$ such that the requirements of Lemmas 2.6, 2.7 and 2.8 hold. At the same time, we need

$$(\delta'')^5M(\varepsilon)\gtrsim\ln\frac{1}{(\widetilde{u}(\delta)-\widetilde{u}(-\delta))^2} \tag{3.24}$$

and

$$M(\varepsilon)\frac{\varepsilon\delta''}{\widetilde{u}(\delta)\wedge(-\widetilde{u}(-\delta))}\lesssim(\widetilde{u}(\delta)-\widetilde{u}(-\delta))^2\ . \tag{3.25}$$

To this end we let $M(\varepsilon)=\ln\left(\dfrac{1}{\varepsilon}\right)$ and $\delta''=\left(\dfrac{(\frac{1}{\widetilde{u}(\delta)-\widetilde{u}(-\delta)})\ln(\frac{1}{\widetilde{u}(\delta)-\widetilde{u}(-\delta)})^2}{\ln(\frac{1}{\varepsilon})}\right)^{1/5}$. At the same time we keep our asymptotic order of choice of ε, δ and δ' as in Section 2. This means that we need

$$\varepsilon\left(\ln\left(\frac{1}{\varepsilon}\right)\right)^{4/5}\frac{1}{\widetilde{u}(\delta)-\widetilde{u}(-\delta)}\ln\left(\frac{1}{\widetilde{u}(\delta)-\widetilde{u}(-\delta)}\right)^2\lesssim(\widetilde{u}(\delta)-\widetilde{u}(-\delta))^2\ .$$

It could be checked that this is possible to make (3.24) and (3.25) to hold. We formulate this as a corollary.

Corollary 3.1. *Let $\boldsymbol{q}_0\in G(\delta')$. Under the above specified asymptotic order we have, there exist $\varepsilon_0>0$ such that for any $0<\varepsilon<\varepsilon_0$ we have*

$$\left|\mathbf{P}_{\boldsymbol{q}_0}(\theta^\varepsilon_{\sigma_0}\in[\theta_1,\theta_2], y^\varepsilon_{\sigma_0}=1+\delta)-\frac{\theta_2-\theta_1}{2\pi}\frac{\widetilde{u}(0)-\widetilde{u}(-\delta)}{\widetilde{u}(\delta)-\widetilde{u}(-\delta)}\right|\leq C\cdot(\widetilde{u}(\delta)-\widetilde{u}(-\delta))^2\ , \tag{3.26}$$

$$\left|\mathbf{P}_{\boldsymbol{q}_0}(\theta^\varepsilon_{\sigma_0}\in[\theta_1,\theta_2], y^\varepsilon_{\sigma_0}=-1-\delta)-\frac{\theta_2-\theta_1}{2\pi}\frac{\widetilde{u}(\delta)-\widetilde{u}(0)}{\widetilde{u}(\delta)-\widetilde{u}(-\delta)}\right|\leq C\cdot(\widetilde{u}(\delta)-\widetilde{u}(-\delta))^2\ . \tag{3.27}$$

Lemma 3.9. *For any $\boldsymbol{q} \in G(\delta')$ and for any $\rho > 0$ there exist $\varepsilon_0 = \varepsilon_0(\rho)$ such that for any $0 < \varepsilon < \varepsilon_0$, for any $f \in D(A)$ we have, for some $K > 0$*

$$|\mathbf{E}_{\boldsymbol{q}} f(\boldsymbol{\pi}(\boldsymbol{q}^\varepsilon_{\sigma_0})) - f(\boldsymbol{\pi}(\boldsymbol{q}))| < K(\widetilde{u}(\delta) - \widetilde{u}(-\delta))^2 \ . \tag{3.28}$$

Proof. We have, using Corollary 3.1, that

$$\begin{aligned}
&|\mathbf{E}_{\boldsymbol{q}} f(\boldsymbol{\pi}(\boldsymbol{q}^\varepsilon_{\sigma_0})) - f(\boldsymbol{\pi}(\boldsymbol{q}))| \\
&= \left|\mathbf{E}_{\boldsymbol{q}} f(\theta^\varepsilon_{\sigma_0}, \pi(y^\varepsilon_{\sigma_0})) - f(\boldsymbol{\pi}(\boldsymbol{q}))\right| \\
&= \left|\int_0^{2\pi} f(\theta,\delta)\mathbf{P}_{\boldsymbol{q}}(\theta^\varepsilon_{\sigma_0} \in d\theta, y^\varepsilon_{\sigma_0} = 1+\delta) + \int_0^{2\pi} f(\theta,-\delta)\mathbf{P}_{\boldsymbol{q}}(\theta^\varepsilon_{\sigma_0} \in d\theta, y^\varepsilon_{\sigma_0} = -1-\delta) - f(\boldsymbol{\pi}(\boldsymbol{q}))\right| \\
&\leq \left|\frac{1}{2\pi}\int_0^{2\pi} \frac{\widetilde{u}(0)-\widetilde{u}(-\delta)}{\widetilde{u}(\delta)-\widetilde{u}(-\delta)} f(\theta,\delta)d\theta + \frac{1}{2\pi}\int_0^{2\pi} \frac{\widetilde{u}(\delta)-\widetilde{u}(0)}{\widetilde{u}(\delta)-\widetilde{u}(-\delta)} f(\theta,-\delta)d\theta - f(\boldsymbol{\pi}(\boldsymbol{q}))\right| + K_1(\widetilde{u}(\delta)-\widetilde{u}(-\delta))^2 \\
&= \left|\frac{1}{2\pi}\int_0^{2\pi} \frac{\widetilde{u}(0)-\widetilde{u}(-\delta)}{\widetilde{u}(\delta)-\widetilde{u}(-\delta)} (f(\theta,\delta)-f(\mathfrak{o}))d\theta - \right. \\
&\qquad \left. \frac{1}{2\pi}\int_0^{2\pi} \frac{\widetilde{u}(\delta)-\widetilde{u}(0)}{\widetilde{u}(\delta)-\widetilde{u}(-\delta)} (f(\mathfrak{o})-f(\theta,-\delta))d\theta + (f(\mathfrak{o})-f(\boldsymbol{\pi}(\boldsymbol{q})))\right| + K_1(\widetilde{u}(\delta)-\widetilde{u}(-\delta))^2 \\
&\leq \left|\frac{(\widetilde{u}(0)-\widetilde{u}(-\delta))(\widetilde{u}(\delta)-\widetilde{u}(0))}{\widetilde{u}(\delta)-\widetilde{u}(-\delta)}\left(\frac{1}{2\pi}\int_0^{2\pi}\frac{f(\theta,\delta)-f(\mathfrak{o})}{\widetilde{u}(\delta)-\widetilde{u}(0)}d\theta - \frac{1}{2\pi}\int_0^{2\pi}\frac{f(\mathfrak{o})-f(\theta,-\delta)}{\widetilde{u}(0)-\widetilde{u}(-\delta)}d\theta\right)\right| + \\
&\qquad |f(\mathfrak{o})-f(\boldsymbol{\pi}(\boldsymbol{q}))| + K_1(\widetilde{u}(\delta)-\widetilde{u}(-\delta))^2 \\
&\leq K(\widetilde{u}(\delta)-\widetilde{u}(-\delta))^2
\end{aligned}$$

for some $K_1 > 0$ and $K > 0$. We have used the gluing condition (3.10) and our specified choice of asymptotic order of δ, δ' and ε. $\square$

Lemma 3.10. *We have, as $\varepsilon, \delta, \delta'$ are small, for $\boldsymbol{q}_0 \in G(\delta)$, that*

$$\left|\mathbf{E}_{\boldsymbol{q}_0}\left[\int_{\sigma_0}^{\tau_1} e^{-\lambda t}[\lambda f(\boldsymbol{\pi}(\boldsymbol{q}^\varepsilon_t)) - Af(\boldsymbol{\pi}(\boldsymbol{q}^\varepsilon_t))]dt + e^{-\lambda\tau_1} f(\boldsymbol{\pi}(\boldsymbol{q}^\varepsilon_{\tau_1}))\right] - f(\boldsymbol{\pi}(\boldsymbol{q}_0))\right| \leq (\widetilde{u}(\delta)-\widetilde{u}(-\delta))^2 \ . \tag{3.29}$$

The proof of this Lemma is essentially the same proof in Lemma 2.6 modified into a two-dimensional version and we omit it.

Finally we would like to mention that our boundary condition given in this section also appears naturally in other model problems. As an example let consider the following system:

$$\begin{cases} x^\varepsilon_t = \displaystyle\int_0^t \frac{1}{\lambda(y^\varepsilon_t)+\varepsilon} dW^1_t \ , \\ y^\varepsilon_t = |W^2_t| \ . \end{cases} \tag{3.30}$$

Here $\lambda(\bullet)$ is a smooth function on $\mathbb{R}_+$ that vanishes at 0 and is strictly positive in $(0,\infty)$; W_t^1 and W_t^2 are two independent standard Wiener processes on $\mathbb{R}$. Let the process $z_t^\varepsilon = (x_t^\varepsilon, y_t^\varepsilon)$ on $\mathbb{R}\times\mathbb{R}_+$ be stopped once it hits the boundary $\{(x,y)\in\mathbb{R}^2 : y = R\}$ for some $R>0$. Let $\theta_t^\varepsilon = x_t^\varepsilon \mod 2\pi$. Let $\boldsymbol{\pi} : S^1\times\mathbb{R}_+\to\mathbb{R}^2$ be the mapping defined by $\boldsymbol{\pi}(\theta,y) = (y\cos\theta, y\sin\theta)$. For each fixed $\varepsilon>0$, the process $w_t^\varepsilon = (\theta_t^\varepsilon, y_t^\varepsilon)$ is a diffusion process on $S^1\times[0,R]$ with normal reflection at the boundary $\{(\theta,y) : y=0\}$ and is stopped once it hits the other boundary $\{(\theta,y) : y = R\}$. Let $m_t^\varepsilon = \boldsymbol{\pi}(w_t^\varepsilon)$ (i.e., we glue all points $\{(\theta,y) : y=0\}$). The process m_t^ε moves within the disk $B(R) = \{m\in\mathbb{R}^2 : |m|_{\mathbb{R}^2}\le R\}$ and is stopped once it hits the boundary. In general, this process is *not* a Markov process. But we expect that, as $\varepsilon\downarrow 0$, this process w_t^ε will converge weakly to a Markov process w_t on $B(R)$ with generator A and the domain of definition $D(A)$, defined as follows: The operator A at points (θ,r) (we use polar coordinates, that is, a point $(x,y)\in\mathbb{R}^2$ is represented by $(r\cos\theta, r\sin\theta)$) with $r\neq 0$ is defined by

$$Af(\theta,r) = \frac{1}{2\lambda^2(r)}\frac{\partial^2}{\partial\theta^2}f(\theta,r) + \frac{1}{2}\frac{\partial^2}{\partial r^2}f(\theta,r)\ . \tag{3.31}$$

The domain of definition $D(A)$ of the operator A consists of those continuous functions f on $B(R)$ for which $Af(\theta,r)$ is defined and continuous for $r\neq 0$, the derivative in r being continuous; such that finite limit

$$\lim_{\theta'\to\theta, r\to 0+}\frac{\partial f}{\partial r}(\theta',r) \tag{3.32}$$

exists;

$$\lim_{\theta'\to\theta, r\to 0+} Af(\theta',r) \tag{3.33}$$

exists and does not depend on θ;

$$\lim_{\theta'\to\theta, r\to R-} Af(\theta',r) = 0\ ; \tag{3.34}$$

and

$$\int_0^{2\pi}\lim_{\theta'\to\theta, r\to 0+}\frac{\partial f}{\partial r}(\theta',r)d\theta = 0\ . \tag{3.35}$$

We define, for $f\in D(A)$, $Af(\theta,R)$ as the limit (3.34) and $Af(O)$ as the limit (3.33).

The weak convergence of w_t^ε to w_t in $\mathbf{C}_{[0,T]}(B(R))$ described above shall be a result of fast motion x_t^ε running at the local time of the slow motion y_t^ε on the boundary $\{(x,y)\in\mathbb{R}\times\mathbb{R}_+ : y=0\}$. The proof of this result shall follow the same method of this section.

3. ON SECOND ORDER ELLIPTIC EQUATIONS WITH A SMALL PARAMETER.

3.1 Introduction

Let G be a bounded domain in $\mathbb{R}^d$ with the smooth boundary ∂G,

$$L_k u(x) = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}^{(k)}(x)\frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b_i^{(k)}(x)\frac{\partial u}{\partial x_i}\ ,\ k=0,1\ ,\ x\in\mathbb{R}^d\ .$$

The coefficients are assumed to be smooth enough, say, in $\mathbf{C}^{(2)}(\mathbb{R}^d)$, i.e., having continuous second derivatives.

Boundary problems for the operator $L_\varepsilon = L_0 + \varepsilon L_1$ in the domain G and corresponding initial-boundary problems for the equation $\dfrac{\partial u^\varepsilon(t,x)}{\partial t} = L_\varepsilon u^\varepsilon$, $t>0$, $x\in G$, are considered. The operator L_ε is assumed to be elliptic for $\varepsilon>0$. One can study the limiting behavior of solutions of stationary problems as $\varepsilon\downarrow 0$ and the limiting behavior of solutions of initial-boundary problems as $\varepsilon\downarrow 0$ and $t\to\infty$.

If the operator L_0 is elliptic in $G\cup\partial G$, the problem is simple: u^ε converges to the solution of corresponding problem for the operator L_0. In the case of degenerate operator L_0, situation is more complicated, and the question was considered in numerous papers. First, the case of first order operator L_0 was considered: $L_0 = b^{(0)}(x)\cdot\nabla$, $b^{(0)}(x) = (b_1^{(0)}(x),...,b_d^{(0)}(x))$. N.Levinson [25] showed in 1950-th that, if the characteristics of L_0 (e.g., trajectories of the dynamical system $\dot{X}_t = b^{(0)}(X_t)$ in $\mathbb{R}^d$) leave the domain G in finite time and cross the boundary in a regular way, then the solution of the Dirichlet problem $L_\varepsilon u^\varepsilon = 0$, $x\in G$, $u^\varepsilon(x)|_{\partial G} = \psi(x)$, converges as $\varepsilon\downarrow 0$ to the solution of degenerate equation $L_0 u^0(x) = 0$, $x\in G$, with the boundary condition $\psi(x)$ ($\psi(x)$ is assumed to be continuous) on the part of ∂G through which the characteristics leave the domain. Such a solution $u^0(x)$ is unique.

Most of subsequent results concerning this problem were obtained by probabilistic methods. With each operator L_ε, $\varepsilon\geq 0$, one can (see [9], notice that the coefficients of $a_{ij}^{(k)}(x)$ are in $\mathbf{C}^{(2)}(\mathbb{R}^d)$) connect a diffusion process $\widetilde{X}_t^\varepsilon$ in $\mathbb{R}^d$ defined by the stochastic differential equation

$$\begin{aligned}&\dot{\widetilde{X}}_t^\varepsilon = b^{(0)}(\widetilde{X}_t^\varepsilon) + \varepsilon b^{(1)}(\widetilde{X}_t^\varepsilon) + \sigma^{(0)}(\widetilde{X}_t^\varepsilon)\dot{W}_t^0 + \sqrt{\varepsilon}\sigma^{(1)}(\widetilde{X}_t^\varepsilon)\dot{W}_t^1\ ,\\ &\widetilde{X}_0^\varepsilon = x\in\mathbb{R}^d\ ,\ t>0\ ,\ \sigma^{(k)}(x)(\sigma^{(k)}(x))^* = (a_{ij}^{(k)}(x)) = a^{(k)}(x)\ ,\ k=0,1\ .\end{aligned}$$

Here W_t^0 and W_t^1 are independent Wiener processes in $\mathbb{R}^d$. Then the solution of the Dirichlet problem for the equation $L_\varepsilon u^\varepsilon(x) = 0$, $x\in G$, and of the initial boundary problem for $\dfrac{\partial u^\varepsilon(t,x)}{\partial t} = L_\varepsilon u^\varepsilon(t,x)$ can be represented as expectations of corresponding functionals of $\widetilde{X}_t^\varepsilon$. The trajectories $\widetilde{X}_t^\varepsilon$, in a sense, play the same role as characteristics in the case of first order operator L_0. Using these representations and studying limiting behavior of process $\widetilde{X}_t^\varepsilon$ one can describe the limiting behavior of the boundary problems (see [11], [14]).

If problems with the Neumann boundary conditions are considered, one can use the corresponding diffusion process with reflection on the boundary (see, for instance, [11, §2.5]). Various cases of first order operators L_0 not satisfying Levinson's conditions were examined using the probabilistic approach (see [11], [14] and the references therein).

If the operator L_0 has terms with second derivatives, one can introduce a generalized Levinson condition ([11, §4.2]). Under this condition the equation $L_0 u^0(x) = 0$, $x \in G$, with appropriate Dirichlet type boundary conditions has a unique solution, and the solution $u^\varepsilon(x)$ of the Dirichlet problem for equation $L_\varepsilon u^\varepsilon(x) = 0$, $x \in G$ converges to $u^0(x)$ as $\varepsilon \downarrow 0$. The difference with the classical Levinson case is just in the rate of convergence: under mild additional assumptions $|u^\varepsilon(x) - u^0(x)| < \varepsilon^\gamma$ for some $\gamma > 0$ and $0 < \varepsilon << 1$, but for any $\gamma' > 0$ one can find L_0 with infinitely differentiable coefficients non-degenerating on ∂G such that $|u^\varepsilon(x) - u^0(x)|$ is greater than $\varepsilon^{\gamma'}$ at a point $x \in G$ and $0 < \varepsilon << 1$.

A convenient way to specify the degeneration of L_0 is given by the conservation laws. A function $H(x)$ is called a first integral for the process X_t^0 corresponding to L_0 if $\mathbf{P}_x(X_t^0 \in S(H(x))) = 1$ for all $t \geq 0$ and $x \in \mathbb{R}^d$, where $S(z) = \{y \in \mathbb{R}^d : H(y) = z\}$; here and below the subscript $x \in \mathbb{R}^d$ in the probability $\mathbf{P}_x$ or expected value $\mathbf{E}_x$ means that the trajectory of the process starts at the point x.

We consider in this chapter self-adjoint operators L_0 and L_1:

$$L_k u(x) = \frac{1}{2} \nabla \cdot (a^{(k)}(x) \nabla u(x)) \ .$$

Then a smooth function $H(x)$ is a first integral for the process $\widetilde{X}_t^0$ (for the corresponding operator L_0) if and only if $a^{(0)}(x) \nabla H(x) \equiv 0$. In general, the process $\widetilde{X}_t^0$ can have several independent smooth first integrals. To restrict ourselves to the case of one smooth first integral we assume that $\mathbf{e} \cdot (a^{(0)}(x)\mathbf{e}) \geq \underline{a}(x) |\mathbf{e}|^2_{\mathbb{R}^d}$ for each $\mathbf{e} \in \mathbb{R}^d$ such that $\mathbf{e} \cdot \nabla H(x) = 0$: It is assumed that $\underline{a}(x)$ is smooth and strictly positive if x is not a critical point of $H(x)$; if x_0 is a critical point, $a^{(0)}(x_0) = 0$ and $\underline{a}(x_0) = 0$.

To be specific we consider the Neumann problem

$$\frac{1}{\varepsilon} L_\varepsilon u^\varepsilon = \left(\frac{1}{\varepsilon} L_0 + L_1 \right) u^\varepsilon(x) = f(x) \ , \ \ \frac{\partial u^\varepsilon(x)}{\partial \gamma^\varepsilon(x)} \bigg|_{\partial G} = 0 \ ; \tag{1.1}$$

$\gamma^\varepsilon(x)$ here is the inward co-normal unit vector to ∂G corresponding to L_ε. Let X_t^ε be the process in $G \cup \partial G$ governed by the operator inside G with reflection along the co-normal to ∂G. Since L_ε is self-adjoint, the Lebesgue measure is invariant for the process X_t^ε, and the problem (1.1) is solvable for continuous $f(x)$ such that $\displaystyle\int_G f(x)dx = 0$. Together with the last condition, we assume that L_1 is not degenerate in $G \cup \partial G$, so that to single out a unique solution of (1.1), we shall fix the value of $u^\varepsilon(x)$ at a point $x_O \in G \cup \partial G$ which is fixed the same for all $\varepsilon > 0$. We let $u^\varepsilon(x_O) = 0$.

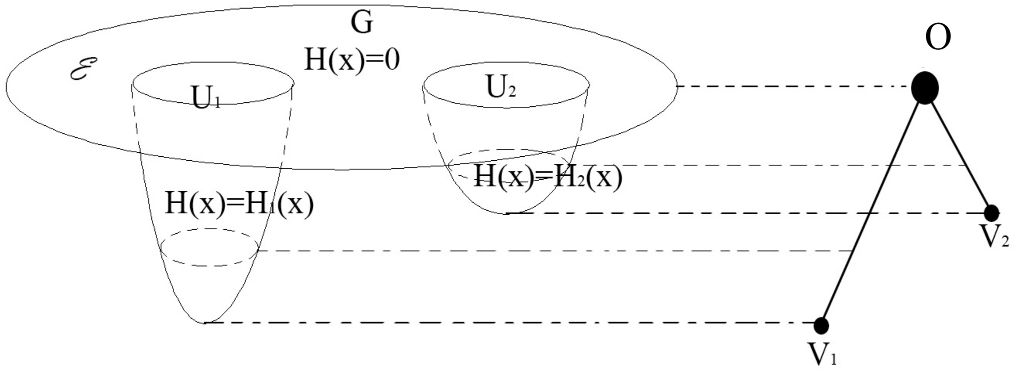

Fig. 3.1.

Then the solution of problem (1.1) can be written in the form (see, for instance [11])

$$u^\varepsilon(x) = -\int_0^\infty \mathbf{E}_x f(X_t^\varepsilon)dt + \int_0^\infty \mathbf{E}_{x_O} f(X_t^\varepsilon)dt\ . \tag{1.2}$$

If the first integral $H(x)$ has in $G \cup \partial G$ no critical points, one can describe the $\lim_{\varepsilon\downarrow 0} u^\varepsilon(x)$ in the way similar to [18]: One shall introduce a graph $\mathbb{G}$ corresponding to the set of connected components of the intersections of the level sets of $H(x)$ within G. A boundary problem on $\mathbb{G}$ with appropriate gluing conditions at the vertices can be formulated, and the solution of this problem defines $\lim_{\varepsilon\downarrow 0} u^\varepsilon(x)$. If the function $H(x)$ has saddle points inside G, additional branchings in the graph appear. The gluing conditions at these new vertices can be calculated using the results of [13].

All mentioned above results concern the case when the rank of $a^{(0)}(x)$ is constant and equal to $d-1$ for all $x \in G \cup \partial G$ except the critical points of $H(x)$. In this paper, we consider the case when L_0 is non-degenerate in a connected subdomain $\mathcal{E} \subset G$, and we let $H(x)$ be equal to a constant on $\mathcal{E}$. Outside $\mathcal{E}$ the first integral $H(x)$ has a finite number of critical points (see Fig.1). For convenience of presentation, we shall then introduce several first integrals $H_k(x)$ ($k = 1, ..., r$) for each of the connected components $U_1, ..., U_r$ on which L_0 is degenerate. We shall let $H(x) = H_k(x)$ for $x \in U_k$. A more concrete setup of the problem is in Section 2. Existence of the domain $\mathcal{E}$ where the operator L_0 is not degenerate leads to more general gluing conditions. The limiting process on the graph spends a positive time at the vertex corresponding to $\mathcal{E}$.

Let $S(z) = \{x \in G \cup \partial G : H(x) = z\}$. The graph $\mathbb{G}$ is the result of identification of points of each connected component of every level set $S(z)$. Let $\mathfrak{Y} : G \cup \partial G \to \mathbb{G}$ be the identification mapping. We call $\mathfrak{Y}(x)$ the projection of x onto $\mathbb{G}$. We consider the projection $Y_t^\varepsilon = \mathfrak{Y}(X_t^\varepsilon)$ of the process X_t^ε on $\mathbb{G}$ and prove that processes Y_t^ε on $\mathbb{G}$ converge weakly in the space of continuous functions $[0, T] \to \mathbb{G}$ to a diffusion process

Y_t on $\mathbb{G}$. The process Y_t is defined by a family of differential operators, one on each edge, and by gluing conditions at the vertices. We calculate the operators and the gluing condition. The function $u^0(x) = \lim\limits_{\varepsilon\downarrow 0} u^\varepsilon(x)$ is constant on each connected component of every level set of $H(x)$: $u^0(x) = v(\mathfrak{Y}(x))$. We formulate a boundary problem for the function $v(y)$, $y \in \mathbb{G}$, which has a unique solution, and actually can be solved explicitly.

The organization of this paper is as follows: Section 2 sets up the problem and gives the main results. Section 3 is devoted to the proof of the main results in Section 2. Section 4 proves auxiliary results needed in Section 3.

3.2 Main results

Let us first speak about our assumptions.

Suppose we have a bounded domain $G \subset \mathbb{R}^d$, with smooth boundary ∂G. Let L_0 be a self-adjoint operator

$$L_0 u(x) = \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial}{\partial x_i}\left(a_{ij}^{(0)}(x)\frac{\partial u(x)}{\partial x_j}\right) = \frac{1}{2}\nabla\cdot(a^{(0)}(x)\nabla u(x)) \ .$$

Let $U_1, ..., U_r$ be several regions inside G. They are simply connected open sets and they do not intersect each other. Let us assume, that the matrix $a^{(0)}(x) \equiv (a_{ij}^{(0)}(x))_{1\le i,j\le d}$ is strictly elliptic on $\mathcal{E} = [G]\backslash(\cup_{k=1}^r [U_k])$ (here $[D]$ is the closure of a domain D). For $x \in [\mathcal{E}]$, the coefficients $a_{ij}^{(0)}(x), 1 \le i,j \le d$ are assumed to be in $\mathbf{C}^{(3)}([\mathcal{E}])$.

Let us discuss the case when $x \in \cup_{k=1}^r [U_k]$. For each $k = 1, ..., r$ and $x \in [U_k]$, the coefficients $a_{ij}^{(0)}(x), 1 \le i,j \le d$ are assumed to be in $\mathbf{C}^{(3)}([U_k])$. We assume that the matrix $(a_{ij}^{(0)}(x))_{1\le i,j\le d}$ is degenerate on $\cup_{k=1}^r [U_k]$. To specify this degeneration, we assume that within each $[U_k]$ there is a function H_k which is a first integral of the (degenerate) operator L_0, i.e., $a^{(0)}(x)\nabla H_k(x) = 0$ for $x \in [U_k]$. Let H_k have only one minimum m_k inside U_k. (We can always make this assumption since if m_k is a maximum we work with $-H_k$ instead of H_k.) Let $x_k(m_k)$ be the point in U_k corresponding to the minimum m_k. This minimum is assumed to be non-degenerate, i.e., the matrix $\left(\dfrac{\partial^2 H_k}{\partial x_i \partial x_j}(x_k(m_k))\right)_{1\le i,j\le d}$ is positive definite. Since the choice of H_k is up to a constant we can assume that $H_k = 0$ on ∂U_k. For $h \in (m_k, 0]$ the level surfaces $C_k(h) = \{x \in U_k : H_k(x) = h\}$ of the functions H_k inside U_k are closed surfaces of dimension $(d-1)$ and the operator L_0 is non-degenerate on $C_k(h)$. Let $\gamma_k = \partial U_k = C_k(0)$. A non-degeneracy condition of $a^{(0)}(x)$ on $C_k(h)$ is assumed: for any vector $\mathbf{e} \in \mathbb{R}^d$ such that $\mathbf{e}\cdot\nabla H_k = 0$ we have $\mathbf{e}\cdot(a^{(0)}(x)\mathbf{e}) \ge \underline{a}(x)|\mathbf{e}|_{\mathbb{R}^d}^2$ for some $\underline{a}(x) > 0$ and $x \ne x_k(m_k)$. We set $a^{(0)}(x_k(m_k)) = 0$ and $C_k(m_k) = \{x_k(m_k)\}$. We assume that the level surfaces $C_k(h)$ for $h \in (m_k, 0]$ divide $U_k\backslash\{x_k(m_k)\}$ into pieces of $(d-1)$-dimensional surfaces.

For simplicity of presentation we will assume that $\nabla H_k(x) \neq 0$ for $x \in \gamma_k$. One can introduce a global first integral $H(x)$ on $[G]$ as in Section 1: $H(x) = H_k(x)$ for $x \in U_k$ and $H(x) = 0$ for $x \in [\mathcal{E}]$. However, this function $H(x)$ is not smooth at $\cup_{k=1}^r \gamma_k$ but this is only a result of non-essential technical assumptions.

Let $\gamma = \cup_{k=1}^r \gamma_k$. We assume that the order of degeneracy is given by the condition that for a certain unit vector field $\mathbf{e}_d(x)$ in a small neighborhood of $\cup_{k=1}^r [U_k]$ we have

$$\text{const}_1 \cdot \text{dist}^2(x,\gamma) \leq \mathbf{e}_d(x) \cdot (a^{(0)}(x)\mathbf{e}_d(x)) \leq \text{const}_2 \cdot \text{dist}^2(x,\gamma)$$

for some const_1, $\text{const}_2 > 0$. The distance $\text{dist}(x,\gamma)$ is the Euclidean distance between x and γ. The vector field $\mathbf{e}_d(x) = \dfrac{\nabla H_k}{|\nabla H_k|_{\mathbb{R}^d}}$ for $x \in \gamma_k$.

In particular, our assumptions imply that the matrix $a^{(0)}(x)$ has rank d in $\mathcal{E}$ and rank $(d-1)$ in $\cup_{k=1}^r [U_k]$. However, the coefficients $a_{ij}^{(0)}(x)$, $1 \leq i,j \leq d$ are only in $\mathbf{C}^{(1)}$ for $x \in \gamma$. We notice that in this case results of [9] do not apply. We shall then assume that there is a decomposition $a^{(0)}(x) = \sigma^{(0)}(x)(\sigma^{(0)}(x))^*$ for all $x \in [G]$. The square matrix $\sigma^{(0)}(x)$ has bounded Lipschitz continuous terms.

We shall assume, that the operator L_1 governing the perturbation is self-adjoint and strictly elliptic within $[G]$:

$$L_1 u(x) = \frac{1}{2}\sum_{i,j=1}^d \frac{\partial}{\partial x_i}\left(a_{ij}^{(1)}(x)\frac{\partial u(x)}{\partial x_j}\right) = \frac{1}{2}\nabla \cdot (a^{(1)}(x)\nabla u(x)) \ .$$

Again we denote the matrix $a^{(1)}(x) \equiv (a_{ij}^{(1)}(x))_{1\leq i,j\leq d}$ and we assume that the terms $a_{ij}^{(1)}(x)$ are in class $\mathbf{C}^{(2)}(\mathbb{R}^d)$. In this case results of [9] apply and we have $a^{(1)}(x) = \sigma^{(1)}(x)(\sigma^{(1)}(x))^*$ for all $x \in [G]$. The square matrix $\sigma^{(1)}(x)$ have bounded Lipschitz continuous terms.

Let us put a Neumann boundary condition with respect to co-normal unit vector $\gamma^\varepsilon(x)$ pointing inward on ∂G. Let X_t^ε be the diffusion process in $[G]$, corresponding to the operator $\dfrac{1}{\varepsilon}L_0 + L_1$ inside G with co-normal reflection at ∂G. We see that Lebesgue measure is invariant for the process X_t^ε.

Let us then speak about the results.

We construct a graph $\mathbb{G}$ as follows. The graph $\mathbb{G}$ has r edges $I_1, ..., I_r$ joined together at one vertex O. Let the other endpoint of I_k be V_k. Let us write $I_k = [m_k, 0]$. The coordinate (k, H_k) is a global coordinate on $\mathbb{G}$. The root O corresponds to all $(k, 0)$ for $k = 1, ..., r$. Let us introduce an identification map $\mathfrak{Y} : [G] \to \mathbb{G}$: for $x \in [\mathcal{E}]$ we have $\mathfrak{Y}(x) = O$ and for $x \in U_k$ we have $\mathfrak{Y}(x) = (k, H_k(x))$. Let the process $Y_t^\varepsilon = \mathfrak{Y}(X_t^\varepsilon)$. We are going to prove, that as $\varepsilon \downarrow 0$ the processes Y_t^ε converge weakly in the space $\mathbf{C}_{[0,T]}(\mathbb{G})$ to a Markov process Y_t on $\mathbb{G}$.

The process Y_t is defined as follows. It is a diffusion process on the graph $\mathbb{G}$ with generator A and the domain of definition $D(A)$. Inside each I_k it is governed by an operator $\mathcal{L}_k$ defined as

$$\mathcal{L}_k f(k, H_k) = \frac{1}{2} M_k^{-1}(H_k) \frac{d}{dH_k} \left(M_k(H_k) \overline{a^{(1)}}(H_k) \frac{df}{dH_k} \right) .$$

Here

$$\overline{a^{(1)}}(h) = M_k^{-1}(h) \int\limits_{C_k(h)} \frac{(a^{(1)}(x)\nabla H_k(x), \nabla H_k(x))}{|\nabla H_k(x)|_{\mathbb{R}^d}} d\sigma ,$$

and normalizing factor

$$M_k(h) = \int\limits_{C_k(h)} \frac{d\sigma}{|\nabla H_k(x)|_{\mathbb{R}^d}} .$$

The notation $d\sigma$ denotes the integral with respect to the area element on $C_k(h)$.

We set $Af = \mathcal{L}_k f$ for $(k, H_k) \in (I_k)$ ((I_k) is the interior of the interval I_k). Let the limit $\lim\limits_{(k,H_k)\to O} Af(k, H_k)$ be finite and independent of k. This limit is set to be $Af(O)$. The domain of definition $D(A)$ of the operator A consists of those functions f that are twice continuously differentiable inside each I_k having the limit $\lim\limits_{H_k\to 0} \dfrac{\partial f}{\partial H_k}(k, H_k)$. These functions satisfy the gluing condition at the vertex O:

$$0 = \text{Volume}(\mathcal{E}) \cdot Af(O) + \frac{1}{2} \sum_{k=1}^{r} p_k \cdot \lim_{H_k \to 0} \frac{\partial f}{\partial H_k}(k, H_k) . \tag{2.1}$$

Here Volume$(\mathcal{E})$ is the d-dimensional volume of the domain $\mathcal{E}$ and

$$p_k = \int_{\gamma_k} \frac{(a^{(1)}(x)\nabla H_k(x), \nabla H_k(x))}{|\nabla H_k(x)|_{\mathbb{R}^d}} d\sigma .$$

For the exterior vertices $V_1, ..., V_r$ no additional assumptions are to be imposed on the behavior of the function f in the domain $D(A)$.

It was proved in [16] the the process Y_t exists and is a strong Markov process on the graph $\mathbb{G}$.

We have

Theorem 2.1. *As $\varepsilon \downarrow 0$ the processes Y_t^ε converge weakly to Y_t in $\mathbf{C}_{[0,T]}(\mathbb{G})$.*

Let μ_x^ε be the distribution of the trajectory $Y_t^\varepsilon = \mathfrak{Y}(X_t^\varepsilon)$ starting from a point $x \in [G]$ in the space $\mathbf{C}_{[0,T]}(\mathbb{G})$: for each Borel subset $B \subseteq \mathbf{C}_{[0,T]}(\mathbb{G})$ we set $\mu_x^\varepsilon(B) = \mathbf{P}_{X_0^\varepsilon = x}(Y_\bullet^\varepsilon \in B)$. Similarly, for each $y \in \mathbb{G}$ we let μ_y^0 be the distribution of Y_t in the space $\mathbf{C}_{[0,T]}(\mathbb{G})$ with $\mu_y^0(B) = \mathbf{P}_{Y_0 = y}(Y_\bullet \in B)$. Theorem 2.1 can be reformulated as

Theorem 2.2. *For every $x \in [G]$ and every $T > 0$ the distribution μ_x^ε converges weakly to $\mu^0_{\mathfrak{Y}(x)}$ as $\varepsilon \downarrow 0$. For every bounded continuous functional F on $\mathbf{C}_{[0,T]}(\mathbb{G})$ we have*

$$\mathbf{E}_{X_0^\varepsilon = x} F(Y_\bullet^\varepsilon) \to \mathbf{E}_{Y_0 = \mathfrak{Y}(x)} F(Y_\bullet)$$

as $\varepsilon \downarrow 0$.

The process Y_t^ε can be viewed as the slow component of the process X_t^ε. The fast component Z_t^ε of X_t^ε is a process governed by the operator $\dfrac{1}{\varepsilon} L_0$. The process Z_t^ε moves on $\mathfrak{Y}^{-1}(y)$ for each $y \in \mathbb{G}$: it is moving on $[\mathcal{E}]$ when $y = O$ and it is moving on $C_k(H_k)$ when $y = (k, H_k)$. Since Lebesgue measure is invariant for the process X_t^ε, the fast component Z_t^ε, as $\varepsilon > 0$ is small, has, approximately, a distribution with density $\dfrac{1}{\text{Volume}(\mathcal{E})}$ on $[\mathcal{E}]$ (with respect to Lebesgue measure on $\mathbb{R}^d$) and $\dfrac{1}{M_k(H_k)} \dfrac{1}{|\nabla H_k|_{\mathbb{R}^d}}$ on $C_k(H_k)$ (with respect to the area element $d\sigma$ on $C_k(H_k)$). Using this we can formulate the above two theorems in terms of differential equations:

Theorem 2.3. *Consider the Neumann problem*

$$\frac{1}{\varepsilon} L_\varepsilon u^\varepsilon(x) = \left(\frac{1}{\varepsilon} L_0 + L_1\right) u^\varepsilon(x) = f(x) \text{ for } x \in G \ , \ \left.\frac{\partial u^\varepsilon(x)}{\partial \gamma^\varepsilon(x)}\right|_{x \in \partial G} = 0$$

with a Hölder continuous function $f(x)$ satisfying $\int_G f(x)dx = 0$. Let $u^\varepsilon(x_O) = 0$ for some $x_O \in G$. Then we have

$$\lim_{\varepsilon \downarrow 0} u^\varepsilon(x) = v(\mathfrak{Y}(x))$$

where $v(y)$ is a continuous function on $\mathbb{G}$ such that

$$\mathcal{L}_k v(y) = -\bar{f}(y) \text{ for } y \in (I_k) \ , \ k = 1, ..., r \ .$$

Here

$$\bar{f}(y) = \frac{1}{\text{Volume}(\mathcal{E})} \int_{\mathcal{E}} f(x)dx$$

when $y = O$ and

$$\bar{f}(y) = \frac{1}{M_k(H_k)} \int_{C_k(H_k)} f(x) \frac{d\sigma}{|\nabla H_k(x)|_{\mathbb{R}^d}}$$

when $y = (k, H_k)$. The function $v(y)$ satisfies the gluing condition (2.1) and $v(\mathfrak{Y}(x_O)) = 0$. Such a function $v(y)$ is unique.

3.3 Proof of Theorem 2.1

The **Proof** of Theorem 2.1 follows the arguments of [14], [17], [3] and [4].

Heuristically, the idea of [4] can be explained as follows. The process X_t^ε moves within $[G]$ and has Lebesgue measure as its invariant measure. Since the process X_t^ε has a "fast" component governed by the operator $\dfrac{1}{\varepsilon}L_0$, it will spend a positive amount of time proportional to Volume($\mathcal{E}$) within $\mathcal{E}$ as $\varepsilon \downarrow 0$. As we project X_t^ε onto the graph $\mathbb{G}$ and the whole ergodic component $\mathcal{E}$ corresponds to O, the limiting process Y_t has a boundary condition with a "delay" at O. (We recommend a nice article [23] and a brief summary [24, §5.7] about this boundary condition.) Our gluing condition (2.1) ensures that the process Y_t has an invariant measure on $\mathbb{G}$ that agrees with the Lebesgure measure on $[G]$. We also refer to [14, Ch.8, pp. 347–350] for an explanation of this.

Let us first introduce some notations. Below we will often suppress the small parameter ε and it could be understood directly from the context. Let $\overline{\gamma}_k = C_k(-\varepsilon^{1/2})$ and $\overline{\gamma} = \cup_{k=1}^r \overline{\gamma}_k$. Let σ be the first time when the process X_t^ε hits γ. Let τ be the first time when the process X_t^ε hits $\overline{\gamma}$. Let $\sigma_0 = \sigma$. Let τ_n be the first time following σ_n when the process reaches $\overline{\gamma}$. For $n \geq 1$ let σ_n be the first time after τ_{n-1} when the process X_t^ε reaches γ.

Let $\sigma^* \in \{\sigma_0, \sigma_1, ...\}$ and we denote by $m_{\sigma^*}^x$ the measure on γ induced by $X_{\sigma^*}^\varepsilon$ starting at x. That is,

$$m_{\sigma^*}^x(A) = \mathbf{P}_x(X_{\sigma^*}^\varepsilon \in A) \ , \ A \in \mathcal{B}(\gamma) \ .$$

Let $\nu(\bullet)$ be the invariant measure of the induced chain $X_{\sigma_n}^\varepsilon$ on γ. The key lemma of [4] is the following

Lemma 3.1. *Let $x \in [\mathcal{E}]$. For each $\delta > 0$ and all sufficiently small ε there is a stopping time σ^* which may depend on δ, ε and x and such that*

$$\mathbf{E}_x \sigma^* \leq \delta \ , \tag{3.1}$$

$$\sup_{x \in \gamma} Var(m_{\sigma^*}^x(dy) - \nu(dy)) \leq \delta \ , \tag{3.2}$$

where Var is the total variation of the signed measure.

Our proof of this lemma is a bit simpler than that of [4].

Proof. We will prove, in Lemma 4.11 that $X_{\sigma_n}^\varepsilon$ satisfies the Doeblin condition on γ uniformly in ε. This implies that one can choose an N depending on δ but independent of ε such that the distribution of $X_{\sigma_N}^\varepsilon$ is δ-close to the invariant measure $\nu(\bullet)$ on γ. That is, as we set $\sigma^* = \sigma_N$ the condition (3.2) is satisfied.

We are going to prove in Lemmas 4.8, 4.9, 4.10, respectively, that

$$\lim_{\varepsilon\downarrow 0}\sup_{x\in[\mathcal{E}]}\mathbf{E}_x\sigma = 0\ , \tag{3.3}$$

$$\lim_{\varepsilon\downarrow 0}\sup_{x\in\gamma}\mathbf{E}_x\tau = 0\ , \tag{3.4}$$

$$\lim_{\varepsilon\downarrow 0}\sup_{x\in\overline{\gamma}}\mathbf{E}_x\sigma = 0\ , \tag{3.5}$$

uniformly in ε. We can write $\sigma_N = \sum\limits_{i=1}^{N}[(\sigma_i - \tau_{i-1}) + (\tau_{i-1} - \sigma_{i-1})] + \sigma_0$. For each $i = 1, ..., N$ the random variable $\sigma_i - \tau_{i-1}$ has the same distribution as σ for the process X_t^ε starting at some point on $\overline{\gamma}$; similarly, the random variable $\tau_{i-1} - \sigma_{i-1}$ has the same distribution as τ for the process X_t^ε starting at some point on γ. The results (3.3), (3.4), (3.5) imply that as ε is small (notice that N is fixed at this stage) we can choose $\sigma^* = \sigma_N$ and the condition (3.1) is also satisfied. $\square$

Proof of Theorem 2.1. The proof of Theorem 2.1 is the same as the proof of Lemma 2.1 (including the proof of Lemma 3.4) stated in [4] using the above Lemma 3.1. For the sake of completeness let us briefly repeat it here. Reasoning as in [3], [4], [14], [17], it suffices to prove that for a function $f \in D(A)$, for every $T > 0$ and uniformly in $x \in [G]$ we have

$$\mathbf{E}_x\left[f(H(X_T^\varepsilon)) - f(H(X_0^\varepsilon)) - \int_0^T Af(H(X_s^\varepsilon))ds\right] \to 0$$

as $\varepsilon \downarrow 0$. Here $H(x) = H_k(x)$ if $x \in U_k$ and $H(x) = 0$ if $x \in [\mathcal{E}]$. Let us replace the time interval $[0, T]$ by a larger one $[0, \widetilde{\sigma}]$, where $\widetilde{\sigma}$ is the first of the stopping times σ_n which is greater than or equal to T: $\widetilde{\sigma} = \min\limits_{n:\sigma_n > T}\sigma_n$. Let $\widetilde{\sigma} = \sigma_{\widetilde{n}+1}$. We have

$$\begin{aligned}
&\mathbf{E}_x\left[f(H(X_T^\varepsilon)) - f(H(X_0^\varepsilon)) - \int_0^T Af(H(X_s^\varepsilon))ds\right] \\
&= \mathbf{E}_x\left[f(H(X_\sigma^\varepsilon)) - f(H(X_0^\varepsilon)) - \int_0^\sigma Af(H(X_s^\varepsilon))ds\right] + \mathbf{E}_x\sum_{k=0}^{\widetilde{n}}\int_{\sigma_k}^{\sigma_{k+1}} Af(H(X_s^\varepsilon))ds - \\
&\quad -\mathbf{E}_x\mathbf{E}_{X_T^\varepsilon}\left[f(H(X_\sigma^\varepsilon)) - f(H(X_0^\varepsilon)) - \int_0^\sigma Af(H(X_s^\varepsilon))ds\right] \\
&= (I) + (II) - (III)\ .
\end{aligned}$$

If $x \in \cup_{k=1}^r U_k$ we have $|(I)| \to 0$ uniformly in x as $\varepsilon \downarrow 0$ due to averaging principle. If $x \in [\mathcal{E}]$ then $|(I)| \to 0$ uniformly in x as $\varepsilon \downarrow 0$ due to Lemma 4.8. In a similar way we see that $|(III)| \to 0$ uniformly in $x \in [G]$ as $\varepsilon \downarrow 0$.

Let $\alpha_k = \displaystyle\int_{\sigma_k}^{\sigma_{k+1}} Af(H(X_s^\varepsilon))ds$ and let $\beta_k = \displaystyle\sum_{n=0}^{\infty} \mathbf{E}_x(\alpha_{k+n}|\mathcal{F}_k)$ ($\mathcal{F}_k$ is the filtration generated by the process X_t^ε for $t \leq \sigma_k$). We have $\mathbf{E}_x(\alpha_k - \beta_k + \beta_{k+1}|\mathcal{F}_k) = 0$ and therefore $\left(\displaystyle\sum_{k=1}^{n}(\alpha_k - \beta_k + \beta_{k+1}), \mathcal{F}_{n+1}\right)$, $n \geq 0$ is a martingale. Using the optimal sampling theorem we have

$$\mathbf{E}_x \sum_{k=0}^{\widetilde{n}} \alpha_k = \mathbf{E}_x \sum_{k=0}^{\widetilde{n}} (\alpha_k - \beta_k + \beta_{k+1}) + \mathbf{E}_x(\beta_0 - \beta_{\widetilde{n}+1}) = \mathbf{E}_x(\beta_0 - \beta_{\widetilde{n}+1}) \ .$$

The above argument shows that for the proof of $|(II)| \to 0$ uniformly in $x \in [G]$ as $\varepsilon \downarrow 0$ it suffices to prove $\sup\limits_{x\in\gamma} \left|\displaystyle\sum_{n=0}^{\infty} \mathbf{E}_x \alpha_n\right| \to 0$ uniformly in $x \in \gamma$ as $\varepsilon \downarrow 0$.

Let us first show that $\mathbf{E}_\nu \alpha_0 = 0$. By Lemma 4.11 the Markov chain $X_{\sigma_n}^\varepsilon$, $n \geq 0$ on γ is ergodic and has invariant measure ν. Therefore we have $\lim\limits_{n\to\infty} \dfrac{\sigma_n}{n} = \mathbf{E}_\nu \sigma_1$. By ergodicity of the process X_t^ε and self-adjointness of L_0 and L_1 we also have

$$\lim_{t\to\infty} \mathbf{E}_\nu \frac{1}{t} \int_0^t Af(H(X_s^\varepsilon))ds = \int_{[G]} Af(H(x))dx \ .$$

These two equalities imply that $\mathbf{E}_\nu \alpha_0 = \mathbf{E}_\nu \displaystyle\int_0^{\sigma_1} Af(H(X_s^\varepsilon))ds = (\mathbf{E}_\nu \sigma_1) \cdot \int_{[G]} Af(H(x))dx$.

We have

$$\int_{[G]} Af(H(x))dx = \text{Volume}(\mathcal{E}) \cdot Af(O) + \sum_{k=1}^{r} \int_{I_k} M_k(H_k)\mathcal{L}_k f(H_k)dH_k \ .$$

(The notations agree with those in the definition of the process Y_t.)

Since

$$\int_{I_k} M_k(H_k)\mathcal{L}_k f(H_k)dH_k = \int_{I_k} \frac{d}{dH_k}\left(M_k(H_k)\overline{a^{(1)}}(H_k)\frac{df}{dH_k}\right)dH_k = \frac{1}{2}p_k \cdot \lim_{H_k\to 0} \frac{df}{dH_k}(k, H_k) \ ,$$

we can use our boundary condition (2.1) to have $\displaystyle\int_{[G]} Af(H(x))dx = 0$ and therefore $\mathbf{E}_\nu \alpha_0 = 0$.

From the fact that $\mathbf{E}_\nu \alpha_0 = 0$ one first derives that $\sup\limits_{x\in\gamma} \mathbf{E}_x \alpha_n$ decays to 0 exponentially fast and therefore $\sup\limits_{x\in\gamma} \left|\sum\limits_{n=0}^{\infty} \mathbf{E}_x \alpha_n\right| < \infty$. It also gives, for $x \in \gamma$, that, for $\sigma^* \in \{\sigma_1, \sigma_2, ...\}$ we have

$$\left|\sum_{n=0}^{\infty} \mathbf{E}_x \alpha_n\right| \leq \|Af\|_\infty \cdot \mathbf{E}_x \sigma^* + \text{Var}(m_{\sigma^*}^x - \nu) \cdot \sup_{x\in\gamma} \left|\sum_{n=0}^{\infty} \mathbf{E}_x \alpha_n\right| \ .$$

Using Lemma 3.1 we see that for any $\delta > 0$ we can choose σ^* such that

$$\sup_{x\in\gamma}\left|\sum_{n=0}^{\infty}\mathbf{E}_x\alpha_n\right| \leq \|Af\|_{\infty}\cdot\delta+\delta\cdot\sup_{x\in\gamma}\left|\sum_{n=0}^{\infty}\mathbf{E}_x\alpha_n\right| \ ,$$

which proves that $\sup\limits_{x\in\gamma}\left|\sum\limits_{n=0}^{\infty}\mathbf{E}_x\alpha_n\right| \to 0$ uniformly in $x\in\gamma$ as $\varepsilon\downarrow 0$. This implies that $|(II)|\to 0$ uniformly in $x\in[G]$ as $\varepsilon\downarrow 0$ and Theorem 2.1 follows. $\square$

3.4 *Auxiliary results needed in the proof of Theorem 2.1*

We establish in this section all the auxiliary results needed in Section 3 for the proof of Theorem 2.1.

Let us make some further geometric constructions. Since we assumed that all these U_k's do not intersect each other we see that for sufficiently small neighborhoods of these U_k's they also do not intersect each other. Without loss of generality let us speak about one of these U_k's. We remind that the matrix $a^{(0)}(x)=(a_{ij}^{(0)}(x))_{1\leq i,j\leq d}$ is non-negative definite inside $[G]$ and has rank d on $[G]\backslash\cup_{k=1}^r[U_k]$ and rank $(d-1)$ on $\cup_{k=1}^n[U_k]$. Since the operator L_0 is non-degenerate on $C_k(h)$ for $h\in(m_k,0]$ we see that $a^{(0)}(x)\nabla H_k=0$ on $C_k(h)$ and $\mathbf{e}\cdot(a^{(0)}(x)\mathbf{e})\geq\underline{a}(x)|\mathbf{e}|^2_{\mathbb{R}^d}$ for any unit vector $\mathbf{e}$ tangent to $C_k(h)$. Here $\underline{a}(x)>0$ for $x\in C_k(h)$ and $h\in(m_k,0]$. The eigenvalue $\lambda(x)=0$ for $a^{(0)}(x)$, $x\in C_k(0)=\gamma_k$ is simple and is the smallest one in the spectrum of $a^{(0)}(x)$. For $x\in\gamma_k$ the family of eigen-polynomials $p(\lambda;x)=\det(\lambda I-a^{(0)}(x))$ pass through the origin. They are transversal (i.e. not tangent) to the axis $p=0$. The transversality is preserved under a small perturbation. From here one can see that the eigenvalue $\lambda(x)$ will remain simple and is still the smallest one in the spectrum of all the matrices $a^{(0)}(x)$ as x belongs to a small neighborhood of U_k. We then see from implicit function theorem that this eigenvalue $\lambda(x)$ is a $\mathbf{C}^{(3)}$ function in a small enough neighborhood of U_k. As a consequence, the unit eigenvector $\mathbf{e}_d(x)$ (for different k it is different vector fields but for simplicity of notation we ignore that k in our notation) corresponding to this smallest eigenvalue is a $\mathbf{C}^{(3)}$ vector field in a neighborhood of U_k, with $\mathbf{e}_d(x)=\dfrac{\nabla H_k(x)}{|\nabla H_k(x)|_{\mathbb{R}^d}}$ for $x\in\gamma_k$. Let $X^x(t)$ be the integral curve of this vector field. We let $\dfrac{dX^x(t)}{dt}=\mathbf{e}_d(X^x(t))$, $X^x(0)=x\in\gamma_k$. As we are working within a small neighborhood of γ_k and $\mathbf{e}_d(x)$ in this neighborhood is a $\mathbf{C}^{(3)}$ vector field, being transversal to γ_k when $x\in\gamma_k$, we see that for $t\in[0,\overline{h}]$ with $\overline{h}$ sufficiently small the points $X^x(t)$ for fixed t and all $x\in\gamma_k$ form a surface $\mathbf{C}^{(3)}$ diffeomorphic to γ_k. In this way we obtain an extension of H_k to a neighborhood of U_k by letting $H_k(X^x(t))=t$ for $t\in[0,\overline{h}]$. The Euclidean distance from a point $X^x(t)$ to γ_k is $\geq\underline{d}\cdot t$ for some $\underline{d}>0$. Let us denote by $C_k(+t)$ the level surface $\{H_k=+t\}$ for $t\in[0,\overline{h}]$. Let $\underline{\gamma}_k=C_k(+\varepsilon^{1/4})$ and $\underline{\underline{\gamma}}_k=C_k(+2\varepsilon^{1/4})$. Let $\underline{\gamma}=\cup_{k=1}^r\underline{\gamma}_k$

and $\underline{\underline{\gamma}} = \cup_{k=1}^r \underline{\underline{\gamma}}_k$. We can take ε small such that all $\underline{\underline{\gamma}}_k$'s do not intersect each other and do not touch ∂G. We denote by $\mathcal{E}(\varepsilon^{1/4})$ those points of $x \in \mathcal{E}$ which lie outside the union of the neighborhoods of the U_k's bounded by $\underline{\gamma}_k$, and we denote $\mathcal{E}(2\varepsilon^{1/4})$ in a similar way.

We shall denote, for $x \in \mathcal{E}(\varepsilon^{1/4})$, the stopping time $\sigma(\varepsilon^{1/4})$ to be the time when X_t^ε first hits $\underline{\gamma}$. Notice that by our assumption, for a point $x \in \mathcal{E}(\varepsilon^{1/4})$ we have

$$\sum_{i,j=1}^d a_{ij}^{(0)}(x)\xi_i\xi_j \geq \text{const} \cdot \varepsilon^{1/2} \sum_{i,j=1}^d \xi_i^2$$

for all $(\xi_1, ..., \xi_d) \in \mathbb{R}^d$ and some const > 0.

Lemma 4.1. *For any $0 < \varkappa < 1/2$, for any ε small enough we have*

$$\sup_{x \in [\mathcal{E}(\varepsilon^{1/4})]} \mathbf{E}_x \sigma(\varepsilon^{1/4}) \leq C\varepsilon^{1/2-\varkappa}$$

for some $C > 0$.

Proof. Our argument follows from [20, Ch.6]. Let $u^\varepsilon(x,t) = \mathbf{P}_x(\sigma(\varepsilon^{1/4}) > t)$. Then $u^\varepsilon(x,t)$ solves the problem

$$\begin{cases} \dfrac{\partial u^\varepsilon}{\partial t} = \left(\dfrac{1}{\varepsilon}L_0 + L_1\right) u^\varepsilon \ , \\ u^\varepsilon(y,0) = 1 \text{ for } y \in \mathcal{E}(\varepsilon^{1/4}) \ , \\ u^\varepsilon(y,t) = 0 \text{ for } y \in \underline{\gamma} \text{ and } t > 0 \ , \\ \dfrac{\partial u^\varepsilon}{\partial \gamma^\varepsilon}(y,t) = 0 \text{ for } y \in \partial G \ . \end{cases}$$

Let $\varphi(x) = e^{\alpha R} - e^{\alpha x_1}$ for some $\alpha > 0$. Here $R > 0$ is so chosen that $R \geq 2x_1$ for all $x = (x_1, ..., x_d) \in [G]$. We have $\varphi(x) \geq 0$ for $x \in [G]$. We have

$$\begin{aligned} &\left(\frac{1}{\varepsilon}L_0 + L_1 - \frac{\partial}{\partial t}\right)\varphi(x) \\ &= -\frac{1}{2\varepsilon}a_{11}^{(0)}(x)\alpha^2 e^{\alpha x_1} - \frac{1}{2\varepsilon}\frac{\partial a_{11}^{(0)}}{\partial x_1}\alpha e^{\alpha x_1} - \frac{1}{2}a_{11}^{(1)}(x)\alpha^2 e^{\alpha x_1} - \frac{1}{2}\frac{\partial a_{11}^{(1)}}{\partial x_1}\alpha e^{\alpha x_1} \ . \end{aligned}$$

One can choose α large enough but independent of ε such that

$$\left(\frac{1}{\varepsilon}L_0 + L_1 - \frac{\partial}{\partial t}\right)\varphi \leq -\frac{P}{\varepsilon^{1/2}}$$

with $P = \inf\limits_{x \in [G]} e^{\alpha x_1}$.

Let $P_0 = \inf\limits_{x \in [G]} \varphi(x)$ and $P_1 = \sup\limits_{x \in [G]} \varphi(x)$. Consider an auxiliary function

$$\psi(x,t)=\varepsilon\frac{\varphi(x)}{P}+\varepsilon\frac{\varphi(x)}{P_0}+A\frac{\varphi(x)}{P_0}e^{-\beta(t/\varepsilon-\rho)} .$$

Here $\rho>0$ is a small constant (which can be chosen arbitrarily small) such that $u^\varepsilon(y,t)=0$ for $y\in\underline{\gamma}$ and $t>\rho\varepsilon$. The constants $A>0$ and $\beta>0$ are to be chosen later.

We can calculate

$$\begin{aligned}&\left(\frac{1}{\varepsilon}L_0+L_1-\frac{\partial}{\partial t}\right)\psi\\ &\le -\varepsilon^{1/2}-\frac{P}{P_0}\varepsilon^{1/2}-\frac{P}{\varepsilon^{1/2}}\frac{A}{P_0}e^{-\beta(t/\varepsilon-\rho)}+\frac{\beta}{\varepsilon}A\frac{P_1}{P_0}e^{-\beta(t/\varepsilon-\rho)} .\end{aligned}$$

Letting $\beta=\beta(\varepsilon)=\dfrac{\varepsilon^{1/2}P}{P_1}$ we have

$$\psi(x,t)=\varepsilon\frac{\varphi(x)}{P}+\varepsilon\frac{\varphi(x)}{P_0}+A\frac{\varphi(x)}{P_0}e^{-\frac{P\varepsilon^{1/2}}{P_1}(t/\varepsilon-\rho)} .$$

We have

$$\left(\frac{1}{\varepsilon}L_0+L_1-\frac{\partial}{\partial t}\right)\psi\le-\varepsilon^{1/2} .$$

Also $\psi(x,t)>\varepsilon$ for $x\in\underline{\gamma}$ and $t\ge 0$.

Let $A>\sup\limits_{x\in[G]}|u^\varepsilon(x,\rho\varepsilon)|$ so that $\psi(x,\rho\varepsilon)>A>\sup\limits_{x\in[G]}|u^\varepsilon(x,\rho\varepsilon)|$ for $x\in[G]$.

Since

$$\left(\frac{1}{\varepsilon}L_0+L_1-\frac{\partial}{\partial t}\right)(\pm u^\varepsilon)=0$$

and $\pm u^\varepsilon(x,t)=0<\varepsilon<\psi(x,t)$ for $x\in\underline{\gamma}$ and $t>0$, by comparison, we have

$$|u^\varepsilon(x,t)|\le\psi(x,t)\le A_1\varepsilon+A_2\exp(-A_3\frac{t}{\varepsilon^{1/2}}+\varepsilon^{1/2}A_4)$$

for some constants $A_1,A_2,A_3,A_4>0$ and $x\in[G]$, $t>\rho\varepsilon$.

This implies

$$\sup_{x\in[\mathcal{E}(\varepsilon^{1/4})]}\mathbf{P}_x(\sigma(\varepsilon^{1/4})>\varepsilon^{1/2-\varkappa})\le A_1\varepsilon+A_2\exp(-A_3\varepsilon^{-\varkappa}+\varepsilon^{1/2}A_4) .$$

By strong Markov property of the process X_t^ε we see that

$$\begin{aligned}
&\mathbf{E}_x\sigma(\varepsilon^{1/4})\\
&=\int_0^\infty \mathbf{P}_x(\sigma(\varepsilon^{1/4})>t)dt\\
&\leq \varepsilon^{1/2-\varkappa}\sum_{n=0}^\infty \mathbf{P}_x(\sigma(\varepsilon^{1/4})>n\varepsilon^{1/2-\varkappa})\\
&\leq \varepsilon^{1/2-\varkappa}\sum_{n=0}^\infty (\sup_{x\in[\mathcal{E}(\varepsilon^{1/4})]}\mathbf{P}_x(\sigma(\varepsilon^{1/4})>\varepsilon^{1/2-\varkappa}))^n\\
&=\frac{\varepsilon^{1/2-\varkappa}}{1-\sup\limits_{x\in[\mathcal{E}(\varepsilon^{1/4})]}\mathbf{P}_x(\sigma(\varepsilon^{1/4})>\varepsilon^{1/2-\varkappa})}\\
&\leq \frac{\varepsilon^{1/2-\varkappa}}{1-A_1\varepsilon-A_2\exp(-A_3\varepsilon^{-\varkappa}+\varepsilon^{1/2}A_4)}\leq C\varepsilon^{1/2-\varkappa}
\end{aligned}$$

for ε small enough. This implies the statement of the Lemma. $\square$

We shall denote by $\mathcal{S}_k([0,\varepsilon^{1/4}])$ the closed set bounded by the surfaces $\underline{\gamma}_k$ and γ_k and by $\mathcal{S}([0,\varepsilon^{1/4}])=\cup_{k=1}^r\mathcal{S}_k([0,\varepsilon^{1/4}])$. We denote $\mathcal{S}_k([0,2\varepsilon^{1/4}])$ and $\mathcal{S}([0,2\varepsilon^{1/4}])$ in a similar way by replacing $\underline{\gamma}_k$ by $\underline{\underline{\gamma}}_k$.

Following the geometric construction stated before Lemma 4.1, for $\varepsilon>0$ small enough, and each $k=1,...,r$, at any point $x\in\mathcal{S}_k([0,2\varepsilon^{1/4}])$ one can introduce an orthonormal frame $\{\mathbf{e}_j(x)\}_{j=1}^d$ smoothly depending on $x\in\mathcal{S}([0,2\varepsilon^{1/4}])$ such that $\mathbf{e}_d(x)=\dfrac{\nabla H_k(x)}{|\nabla H_k(x)|_{\mathbb{R}^d}}$ and $\mathbf{e}_j(x)\cdot(a^{(0)}(x)\mathbf{e}_j(x))\geq \underline{a}|\mathbf{e}_j(x)|^2_{\mathbb{R}^d}=\underline{a}$ for some $\underline{a}>0$ and $j=1,...,d-1$. Also $a^{(0)}(x)\mathbf{e}_d(x)=\lambda(x)\mathbf{e}_d(x)$. The eigenvalue $\lambda(x)$ is in $\mathbf{C}^{(3)}(\mathcal{S}_k([0,2\varepsilon^{1/4}]))$ with $\lambda|_{\gamma_k}=0$ and $\lambda(x)>0$ for $x\in\mathcal{S}_k([0,2\varepsilon^{1/4}])\backslash\gamma_k$. Furthermore, for ε small enough we have $C_1\cdot\text{dist}^2(x,\gamma_k)\leq\lambda(x)\leq C_2\cdot\text{dist}^2(x,\gamma_k)$ for some $C_1,C_2>0$ and $x\in\mathcal{S}([0,2\varepsilon^{1/4}])$.

Within the rest of this section implied positive constants denoted by C_i's will not be explicitly pointed out unless necessary. Also, sometimes we use the same symbol C to denote different implied positive constants which are not important.

Let us introduce a new coordinate $(\varphi_1^k,...,\varphi_{d-1}^k,H_k)$ in $\mathcal{S}_k([0,2\varepsilon^{1/4}])$. We take $H_k=H_k(x)$, which is the extended first integral of H_k to $\mathcal{S}_k([0,2\varepsilon^{1/4}])$; and we take $(\varphi_1^k,...,\varphi_{d-1}^k)=(\varphi_1^k(x),...,\varphi_{d-1}^k(x))$ to be the coordinate for a point $\varphi^k(x)=(\varphi_1^k(x),...,\varphi_{d-1}^k(x))$ on γ_k. The point $\varphi^k(x)\in\gamma_k$ is such that $X^{\varphi^k(x)}(H_k(x))=x$ for the flow $X^x(t)$ introduced in the geometric construction before Lemma 4.1. In the more or less simpler case we can arrange the coordinate $(\varphi_1^k(x),...,\varphi_{d-1}^k(x),H_k(x))$ in such a way that $(\mathbf{e}_1(x),...,\mathbf{e}_d(x))$ is the orthonormal frame corresponding to axis curves of this new coordinate system. (We will discuss the general case a bit later.) The metric tensor of this new coordinate system is given by $ds^2=E_1(x)(d\varphi_1^k(x))^2+...+E_{d-1}(x)(d\varphi_{d-1}^k(x))^2+E_d(x)(dH_k(x))^2$. Here the functions $0<C_3<E_1(x),...,E_d(x)<C_4<\infty$ are in class $\mathbf{C}^{(3)}(\mathcal{S}_k([0,2\varepsilon^{1/4}]))$ with bounded derivatives. We notice that by our geometric construction we have $C_5\cdot H_k^2(x)\leq\lambda(x)\leq C_6\cdot H_k^2(x)$.

The theory of orthogonal curvilinear coordinate system (see, for example, [33, Ch.14]) tells us that for a differentiable function f on $\mathcal{S}_k([0,2\varepsilon^{1/4}])$ we have

$$\nabla f(x)=\sum_{i=1}^{d-1}\frac{1}{\sqrt{E_i(x)}}\frac{\partial f}{\partial \varphi_i^k}(x)\mathbf{e}_i(x)+\frac{1}{\sqrt{E_d(x)}}\frac{\partial f}{\partial H_k}(x)\mathbf{e}_d(x)\ ,$$

and for a differentiable vector field $\mathbf{B}(x)=\sum\limits_{i=1}^{d}B^i(x)\mathbf{e}_i(x)$ on $\mathcal{S}_k([0,2\varepsilon^{1/4}])$ we have

$$\nabla\cdot\mathbf{B}(x)=\frac{1}{\sqrt{\prod_{i=1}^d E_i(x)}}\left[\sum_{i=1}^{d-1}\frac{\partial}{\partial\varphi_i^k}\left(\sqrt{\frac{\prod_{j=1}^d E_j(x)}{E_i(x)}}B^i(x)\right)+\frac{\partial}{\partial H_k}\left(\sqrt{\frac{\prod_{j=1}^d E_j(x)}{E_d(x)}}B^d(x)\right)\right]\ .$$

Consider a function (so called "barrier function", see [21] and [11, Ch.3]) $u_k(x)\in\mathbf{C}^{(2)}(\mathcal{S}_k([0,2\varepsilon^{1/4}]))$ which depends only on H_k and is a constant on each level surface $\{H_k=\text{const}\}$. We can write $u_k(x)=u_k(H_k)$ and we apply the above two formulas to get

$$\begin{aligned}&\left(\frac{1}{\varepsilon}L_0+L_1\right)u_k(x)\\&=\left(\frac{1}{2\varepsilon}\nabla\cdot(a^{(0)}(x)\nabla u_k(x))+\frac{1}{2}\nabla\cdot(a^{(1)}(x)\nabla u_k(x))\right)\\&=\frac{1}{\sqrt{\prod_{i=1}^d E_i(x)}}\left[\frac{1}{2}\frac{\partial}{\partial H_k}\left(\sqrt{\frac{\prod_{i=1}^d E_i(x)}{E_d^2(x)}}\left(\frac{\lambda(x)}{\varepsilon}+\mu_d(x)\right)\frac{du_k}{dH_k}(H_k)\right)\right.\\&\qquad\left.+\sum_{i=1}^{d-1}\frac{1}{2}\frac{\partial}{\partial\varphi_i^k}\left(\sqrt{\frac{\prod_{i=1}^d E_i(x)}{E_d(x)E_i(x)}}\mu_i(x)\right)\cdot\frac{du_k}{dH_k}(H_k)\right]\ .\end{aligned}$$

Here the functions $\mu_1(x),...,\mu_d(x)$ are defined via the relation $a^{(1)}(x)\mathbf{e}_d(x)=\mu_1(x)\mathbf{e}_1(x)+...+\mu_d(x)\mathbf{e}_d(x)$. These functions are in $\mathbf{C}^{(3)}(\mathcal{S}([0,2\varepsilon^{1/4}]))$ with bounded derivatives. Notice that since L_1 is strictly elliptic, the matrix $a^{(1)}(x)$ is positive definite, and therefore the function $\mu_d(x)$ is uniformly bounded from below by a certain positive constant.

For simplicity of notation let us define $A(x)=\sqrt{\prod_{i=1}^d E_i(x)}$ and $A_i(x)=\dfrac{A(x)}{\sqrt{E_i(x)E_d(x)}}$ for $i=1,...,d$. These functions are strictly positive (with uniform lower bound) in $\mathbf{C}^{(2)}(\mathcal{S}([0,2\varepsilon^{1/4}]))$ with bounded derivatives. Under this notation we can write

$$\begin{aligned}&\left(\frac{1}{\varepsilon}L_0+L_1\right)u_k(x)\\&=\frac{1}{A(x)}\left[\frac{1}{2}\frac{\partial}{\partial H_k}\left(A_d(x)\left(\frac{\lambda(x)}{\varepsilon}+\mu_d(x)\right)\frac{du_k}{dH_k}(H_k)\right)+\sum_{i=1}^{d-1}\frac{1}{2}\frac{\partial}{\partial\varphi_i^k}\left(A_i(x)\mu_i(x)\right)\cdot\frac{du_k}{dH_k}(H_k)\right]\ .\end{aligned}$$

As a further simplification we shall define

$$\frac{1}{2}A_d(x)\lambda(x)=K_1(x)\ ,$$

$$\frac{1}{2}A_d(x)\mu_d(x) = K_2(x) \ ,$$

$$\sum_{i=1}^{d-1}\frac{1}{2}\frac{\partial}{\partial \varphi_i^k}(A_i(x)\mu_i(x)) = K_3(x) \ .$$

We have

$$\left(\frac{1}{\varepsilon}L_0 + L_1\right)u_k(x) = \frac{1}{A(x)}\left[\frac{\partial}{\partial H_k}\left(\left(\frac{K_1(x)}{\varepsilon} + K_2(x)\right)\frac{du_k}{dH_k}(H_k)\right) + K_3(x)\frac{du_k}{dH_k}(H_k)\right] . \tag{4.1}$$

For a point $x \in \mathcal{S}_k([0, 2\varepsilon^{1/4}])$ and ε small enough we have

$$C_7 H_k^2(x) \leq K_1(x) \leq C_8 H_k^2(x) \ ; \tag{4.2}$$

$$C_9 H_k(x) \leq \frac{\partial}{\partial H_k}(K_1(x)) \leq C_{10} H_k(x) \ ; \tag{4.3}$$

$$0 < C_{11} < K_2(x) < C_{12} < \infty \ ; \tag{4.4}$$

$$\left|\frac{\partial}{\partial H_k}(K_2(x))\right| \leq C_{13} < \infty \ ; \tag{4.5}$$

$$|K_3(x)| \leq C_{14} < \infty \ . \tag{4.6}$$

We also notice, that since we are working in a small neighborhood $\mathcal{S}_k([0, 2\varepsilon^{1/4}])$, the functions $A_d(x) = A_d(\varphi_1^k, ..., \varphi_{d-1}^k, H_k)$ and $\lambda(x) = \lambda(\varphi_1^k, ..., \varphi_{d-1}^k, H_k)$ have Taylor expansions

$$A_d(\varphi_1^k, ..., \varphi_{d-1}^k, H_k) = A_d(\varphi_1^k, ..., \varphi_{d-1}^k, 0) + O(H_k) \ ,$$

$$\lambda(\varphi_1^k, ..., \varphi_{d-1}^k, H_k) = \frac{1}{2}\frac{\partial^2 \lambda}{\partial H_k^2}(\varphi_1^k, ..., \varphi_{d-1}^k, 0)H_k^2 + O(H_k^3) \ .$$

Therefore we see that for $x \in \mathcal{S}_k([0, 2\varepsilon^{1/4}])$ we have

$$K_1(x) = C_k(\varphi_1^k, ..., \varphi_{d-1}^k)H_k^2 + O(H_k^3) \tag{4.7}$$

with a certain positive function $C_k(\varphi_1^k, ..., \varphi_{d-1}^k)$.

In the general case the axis curve corresponding to H_k will be orthogonal to those corresponding to the φ_i^k's, but the axis curves corresponding to the φ_i^k's are not necessarily orthogonal. The calculation will be more bulky since the metric tensor have cross terms with respect to the coordinate φ_i^k's, but the essence is the same as it is only important to have the axis curves corresponding to H_k being orthogonal to those corresponding to the φ_i^k's. To be more precise, let $(g_{ij})_{1\leq i,j\leq d}$ be the metric tensor corresponding to the coordinate system $(\varphi_1^k, ..., \varphi_{d-1}^k, H_k)$. We introduce a frame $\mathbf{e}_1(x), ..., \mathbf{e}_d(x)$. Here $\mathbf{e}_i(x)$ is the unit tangent vector on the axis curve corresponding to φ_i^k for $1 \leq i \leq d-1$; $\mathbf{e}_d(x)$ is the unit tangent vector on the axis curve corresponding to H_k. We have $g_{id} = g_{di} = 0$ for $i = 1, ..., d-1$ and $g_{dd} > 0$. Let $(g^{ij})_{1\leq i,j\leq d}$ be the dual tensor, i.e., $(g^{ij})_{1\leq i,j\leq d}$ is the inverse matrix of $(g_{ij})_{1\leq i,j\leq d}$. We have $g^{id} = g^{di} = 0$ for $1 \leq i \leq d-1$ and $g^{dd} = \dfrac{1}{g_{dd}}$. For $u_k = u_k(H_k)$ we have

$$\nabla u_k(x) = \frac{1}{\sqrt{g_{dd}(x)}} \frac{du_k}{dH_k} \mathbf{e}_d(x) \ ,$$

and

$$a^{(0)}(x)\nabla u_k(x) = \frac{\lambda(x)}{\sqrt{g_{dd}(x)}} \frac{du_k}{dH_k} \mathbf{e}_d(x) \ ;$$

$$a^{(1)}(x)\nabla u_k(x) = \frac{\mu_d(x)}{\sqrt{g_{dd}(x)}} \frac{du_k}{dH_k} \mathbf{e}_d(x) + \frac{1}{\sqrt{g_{dd}(x)}} \frac{du_k}{dH_k} (\mu_1(x)\mathbf{e}_1(x) + ... + \mu_{d-1}(x)\mathbf{e}_{d-1}(x)) \ .$$

Here, as before, we have $a^{(1)}(x)\mathbf{e}_d(x) = \mu_1(x)\mathbf{e}_1(x) + ... + \mu_d(x)\mathbf{e}_d(x)$. We shall then apply a general formula that for a vector field $\mathbf{B}(x) = \sum\limits_{i=1}^{d} B^i(x)\mathbf{e}^i(x)$ we have

$$\nabla \cdot \mathbf{B}(x) = \frac{1}{\sqrt{g(x)}} \sum_{i=1}^{d-1} \frac{\partial}{\partial \varphi_i^k} (B^i(x)\sqrt{g^{ii}(x)}\sqrt{g(x)}) + \frac{1}{\sqrt{g(x)}} \frac{\partial}{\partial H_k} (B^d(x)\sqrt{g^{dd}(x)}\sqrt{g(x)}) \ .$$

Here $g(x) = \det(g_{ij}(x))$. The basis $\mathbf{e}^1(x), ..., \mathbf{e}^d(x)$ is the reciprocal basis (normalized) dual to $\mathbf{e}_1(x), ..., \mathbf{e}_d(x)$, i.e., $(\mathbf{e}_i, \mathbf{e}^j)_{(g_{ij})} = \delta_{ij}$ with respect to the inner product $(\bullet, \bullet)_{(g_{ij})}$ defined by the metric tensor (g_{ij}). By the fact that the metric tensor has no cross terms between H_k and φ_i^k's, we actually have $\mathbf{e}^d(x) = \mathbf{e}_d(x)$ and $\text{span}\{\mathbf{e}_1(x), ..., \mathbf{e}_{d-1}(x)\} = \text{span}\{\mathbf{e}^1(x), ..., \mathbf{e}^{d-1}(x)\}$.

We then see that the operator $\dfrac{1}{\varepsilon}L_0 + L_1$ applied to $u_k(x) = u_k(H_k)$ will result in a formula which is the same as (4.1). The functions $K_1(x)$, $K_2(x)$ and $K_3(x)$ will somehow be different but they still satisfy the conditions (4.2) – (4.7).

Let $\zeta([0, 2\varepsilon^{1/4}])$ be the first time when the process X_t^ε, starting from a point $x \in \mathcal{S}([0, 2\varepsilon^{1/4}])$, hits γ or $\underline{\underline{\gamma}}$.

Lemma 4.2. *We have*

$$\sup_{x \in \mathcal{S}([0,2\varepsilon^{1/4}])} \mathbf{E}_x \zeta([0, 2\varepsilon^{1/4}]) \leq \ C\varepsilon^{3/4}$$

for some $C > 0$.

Proof. Let

$$K_4(H_k) = \min_{x \in \mathcal{S}_k([0,2\varepsilon^{1/4}]), H_k(x) = H_k} \left(\frac{K_1(x)}{\varepsilon} + K_2(x) \right) \ .$$

By (4.2) and (4.4) we can estimate

$$C_{15} \left(\frac{H_k^2}{\varepsilon} + 1 \right) \leq K_4(H_k) \leq C_{16} \left(\frac{H_k^2}{\varepsilon} + 1 \right) \tag{4.8}$$

for $C_{15}, C_{16} > 0$.

Let

$$K_5(x) = \frac{1}{K_4(H_k(x))}\left(\frac{K_1(x)}{\varepsilon} + K_2(x)\right) \geq 1 \ .$$

The function $K_5(x)$ is a bounded function with bounded derivatives for $x \in \mathcal{S}([0, 2\varepsilon^{1/4}])$.

Let the barrier function $u_k^{(1)}(x) = u_k^{(1)}(H_k)$ be defined by

$$u_k^{(1)}(H_k) = \int_0^{H_k} \frac{K_6(\varepsilon) - y}{K_4(y)} dy$$

with

$$K_6(\varepsilon) = \left(\int_0^{2\varepsilon^{1/4}} \frac{dy}{K_4(y)}\right)^{-1} \left(\int_0^{2\varepsilon^{1/4}} \frac{ydy}{K_4(y)}\right) \ .$$

It is easy to check that

$$u_k^{(1)}(0) = u_k^{(1)}(2\varepsilon^{1/4}) = 0 \ .$$

We can estimate $K_6(\varepsilon) \leq 2\varepsilon^{1/4}$ and we have, by (4.8), that

$$\int_0^{H_k} \frac{dy}{K_4(y)} \leq C_{17} \int_0^{H_k} \frac{dy}{\frac{y^2}{\varepsilon} + 1} \leq C_{17}\varepsilon^{1/2} \arctan(H_k\varepsilon^{-1/2}) \leq C_{18}\varepsilon^{1/2} \ . \tag{4.9}$$

This gives the estimates

$$0 \leq u_k^{(1)}(H_k) \leq C_{19}\varepsilon^{3/4} \tag{4.10}$$

and

$$\left|\frac{du_k^{(1)}}{dH_k}(H_k)\right| \leq C_{20}\varepsilon^{1/4} \tag{4.11}$$

for $0 \leq H_k \leq 2\varepsilon^{1/4}$. Apply (4.1) to the function $u_k^{(1)}$ we can see, using (4.11), that,

$$\begin{aligned}
&\left(\frac{1}{\varepsilon}L_0 + L_1\right) u_k^{(1)}(x) \\
&= \frac{1}{A(x)}\left[\frac{\partial}{\partial H_k}\left(K_5(x)K_4(H_k(x))\frac{du_k^{(1)}}{dH_k}(H_k(x))\right) + K_3(x)\frac{du_k^{(1)}}{dH_k}(H_k(x))\right] \\
&\leq \frac{1}{A(x)}\left[\frac{\partial}{\partial H_k}\left((K_6(\varepsilon) - H_k(x))K_5(x)\right) + C_{21}\varepsilon^{1/4}\right] \\
&= \frac{1}{A(x)}\left[-K_5(x) + \frac{\partial}{\partial H_k}(K_5(x))(K_6(\varepsilon) - H_k(x)) + C_{21}\varepsilon^{1/4}\right] \leq -C_{22}
\end{aligned} \tag{4.12}$$

for $x \in \mathcal{S}([0, 2\varepsilon^{1/4}])$ and ε small enough.

We notice that this process X_t^ε before hitting γ or $\underline{\underline{\gamma}}$ is restricted to one of the $\mathcal{S}_k([0, 2\varepsilon^{1/4}])$'s and the bound (4.12) can be made uniform in k.

Now we apply Itô's formula to the function $u_k^{(1)}$ constructed above up to the stopping time $\zeta([0,2\varepsilon^{1/4}])$. Taking expectation we get

$$u_k^{(1)}(x) = -\int_0^{\zeta([0,2\varepsilon^{1/4}])} \mathbf{E}_x \left(\frac{1}{\varepsilon}L_0 + L_1\right) u_k^{(1)}(X_s^\varepsilon)ds \geq C_{22}\mathbf{E}_x\zeta([0,2\varepsilon^{1/4}]) \ . \tag{4.13}$$

From (4.10) and (4.13) we see that the statement of this Lemma follows. □

Lemma 4.3. *For $x \in \underline{\gamma}$ we have*

$$\mathbf{P}_x(X^\varepsilon_{\zeta([0,2\varepsilon^{1/4}])} \in \gamma) \geq C\varepsilon^{1/4}$$

for some $C > 0$.

Proof. Let

$$K_7(H_k) = \max_{x\in\mathcal{S}_k([0,2\varepsilon^{1/4}]),H_k(x)=H_k} \frac{\dfrac{\partial}{\partial H_k}\left(\dfrac{K_1(x)}{\varepsilon} + K_2(x)\right) + K_3(x)}{\dfrac{K_1(x)}{\varepsilon} + K_2(x)} \ .$$

Let, for a fixed $H_k \in [0, 2\varepsilon^{1/4}]$, the above maximum be achieved at a point

$$\varphi^k = (\varphi_1^k(H_k), ..., \varphi_{d-1}^k(H_k), H_k) \ .$$

By Lemma 4.4 we have

$$\left|K_7(H_k) - \frac{2C_k(\varphi^k)H_k}{C_k(\varphi^k)H_k^2 + \varepsilon K_2(\varphi^k, H_k)}\right| \leq C_{23} \ . \tag{4.14}$$

Let the barrier function $u_k^{(2)}(x) = u_k^{(2)}(H_k)$ be defined by

$$u_k^{(2)}(H_k) = 1 - \frac{\displaystyle\int_0^{H_k} \exp\left(-\int_0^y K_7(z)dz\right)dy}{\displaystyle\int_0^{2\varepsilon^{1/4}} \exp\left(-\int_0^y K_7(z)dz\right)dy} \ .$$

It is easy to see that we have

$$u_k^{(2)}(0) = 1 \ , \ u_k^{(2)}(2\varepsilon^{1/4}) = 0 \ .$$

Apply formula (4.1) we can see that

$$\begin{aligned}
&\left(\frac{1}{\varepsilon}L_0 + L_1\right)u_k^{(2)}(x) \\
&= \frac{1}{A(x)}\left[\left(\frac{K_1(x)}{\varepsilon} + K_2(x)\right)\frac{d^2u_k^{(2)}}{dH_k^2}(H_k) + \left(\frac{\partial}{\partial H_k}\left(\frac{K_1(x)}{\varepsilon} + K_2(x)\right) + K_3(x)\right)\frac{du_k^{(2)}}{dH_k}(H_k)\right] \\
&\geq 0 \ .
\end{aligned} \tag{4.15}$$

However, by (4.14), (4.4) and the property of $C_k(\varphi^k)$ in (4.7) we can estimate

$$\int_{H_k}^{2\varepsilon^{1/4}} \exp\left(-\int_0^y K_7(z)dz\right) dy \geq C_{24} \int_{H_k}^{2\varepsilon^{1/4}} \exp\left(-\int_0^y \frac{2z}{z^2+C_{25}\varepsilon}dz\right) dy \ , \tag{4.16}$$

$$\int_0^{2\varepsilon^{1/4}} \exp\left(-\int_0^y K_7(z)dz\right) dy \leq C_{26} \int_0^{2\varepsilon^{1/4}} \exp\left(-\int_0^y \frac{2z}{z^2+C_{27}\varepsilon}dz\right) dy \ . \tag{4.17}$$

By Lemma 4.5, (4.16) and (4.17) we see that

$$u_k^{(2)}(\varepsilon^{1/4}) \geq C_{28}\varepsilon^{1/4} \ . \tag{4.18}$$

This bound (4.18) can actually be made uniform in k. We can apply Itô's formula to the function $u_k^{(2)}$ constructed above up to the stopping time $\zeta([0,2\varepsilon^{1/4}])$. Taking expectation we get

$$\mathbf{P}_x(X^\varepsilon_{\zeta([0,2\varepsilon^{1/4}])} \in \gamma) - u_k^{(2)}(\varepsilon^{1/4}) = \int_0^{\zeta([0,2\varepsilon^{1/4}])} \mathbf{E}_x\left(\frac{1}{\varepsilon}L_0 + L_1\right) u_k^{(2)}(X_s^\varepsilon)ds \geq 0 \tag{4.19}$$

for $x \in \underline{\gamma}$. Now (4.18) and (4.19) imply the statement of this Lemma. □

Lemma 4.4. *For a fixed $H_k \in [0, 2\varepsilon^{1/4}]$ and the corresponding φ^k defined as in the proof of Lemma 4.3, we have*

$$\left| K_7(H_k) - \frac{2C_k(\varphi^k)H_k}{C_k(\varphi^k)H_k^2 + \varepsilon K_2(\varphi^k, H_k)} \right| \leq C$$

for some $C > 0$.

Proof. Using (4.7), we can write

$$K_7(H_k) = \frac{2C_k(\varphi^k)H_k + O(H_k^2) + \varepsilon K_{2,3}(\varphi^k, H_k)}{C_k(\varphi^k)H_k^2 + O(H_k^3) + \varepsilon K_2(\varphi^k, H_k)} \ .$$

Here $K_{2,3}(\varphi^k, H_k)$ is a bounded function. We then have

$$
\begin{aligned}
&\left|K_7(H_k) - \frac{2C_k(\varphi^k)H_k}{C_k(\varphi^k)H_k^2 + \varepsilon K_2(\varphi^k, H_k)}\right| \\
&= \left|\frac{2C_k(\varphi^k)H_k + O(H_k^2) + \varepsilon K_{2,3}(\varphi^k, H_k)}{C_k(\varphi^k)H_k^2 + O(H_k^3) + \varepsilon K_2(\varphi^k, H_k)} - \frac{2C_k(\varphi^k)H_k}{C_k(\varphi^k)H_k^2 + \varepsilon K_2(\varphi^k, H_k)}\right| \\
&\le \left|\frac{O(H_k^2) + \varepsilon K_{2,3}(\varphi^k, H_k)}{C_k(\varphi^k)H_k^2 + O(H_k^3) + \varepsilon K_2(\varphi^k, H_k)}\right| + \\
&\quad + \left|\frac{2C_k(\varphi^k)H_k}{C_k(\varphi^k)H_k^2 + O(H_k^3) + \varepsilon K_2(\varphi^k, H_k)} - \frac{2C_k(\varphi^k)H_k}{C_k(\varphi^k)H_k^2 + \varepsilon K_2(\varphi^k, H_k)}\right| \\
&= \left|\frac{O(H_k^2) + \varepsilon K_{2,3}(\varphi^k, H_k)}{C_k(\varphi^k)H_k^2 + O(H_k^3) + \varepsilon K_2(\varphi^k, H_k)}\right| + \\
&\quad + \left|\frac{2C_k(\varphi^k)H_k}{C_k(\varphi^k)H_k^2 + \varepsilon K_2(\varphi^k, H_k)} \cdot \frac{O(H_k^3)}{C_k(\varphi^k)H_k^2 + O(H_k^3) + \varepsilon K_2(\varphi^k, H_k)}\right| \le C\ .
\end{aligned}
$$

□

Lemma 4.5. *We have*

$$\int_{\varepsilon^{1/4}}^{2\varepsilon^{1/4}} \exp\left(-\int_0^y \frac{2z}{z^2 + C\varepsilon}dz\right) dy \ge C_{29}\varepsilon^{3/4}\ ,$$

$$C_{31}\varepsilon^{1/2} \ge \int_0^{C_{32}\varepsilon^{1/4}} \exp\left(-\int_0^y \frac{2z}{z^2 + C\varepsilon}dz\right) dy \ge C_{30}\varepsilon^{1/2}\ .$$

Proof. Evaluating the integrals, we have

$$\int_0^y \frac{2z}{z^2 + C\varepsilon}dz = \ln\left(\frac{y^2 + C\varepsilon}{C\varepsilon}\right)\ ,$$

$$\int_a^b \exp\left(-\int_0^y \frac{2z}{z^2 + C\varepsilon}dz\right) dy = \sqrt{C}\varepsilon^{1/2}\left(\arctan(\frac{b}{\sqrt{C}\varepsilon^{1/2}}) - \arctan(\frac{a}{\sqrt{C}\varepsilon^{1/2}})\right)\ .$$

If $a = 0$ and $b = C_{32}\varepsilon^{1/4}$ we already get the second inequality of this Lemma. Now suppose $a = \varepsilon^{1/4}$ and $b = 2\varepsilon^{1/4}$. We shall make use of an asymptotic expansion of $\arctan(y)$ as $y \to \infty$:

$$\arctan(y) = \frac{\pi}{2} - \frac{1}{y} + O(\frac{1}{y^2}) \ \text{ as } \ y \to \infty\ .$$

This gives

$$
\begin{aligned}
&\int_{\varepsilon^{1/4}}^{2\varepsilon^{1/4}} \exp\left(-\int_0^y \frac{2z}{z^2 + C\varepsilon}dz\right) dy \\
&\ge C_{33}\varepsilon^{1/2}\left(-\frac{1}{2\varepsilon^{-1/4}} + \frac{1}{\varepsilon^{-1/4}} + O(\varepsilon^{1/2})\right) \ge C_{29}\varepsilon^{3/4}\ .
\end{aligned}
$$

□

Lemma 4.6. *We have*

$$\lim_{\varepsilon\downarrow 0}\sup_{x\in\underline{\gamma}}\mathbf{E}_x\sigma = 0$$

uniformly in ε.

Proof. Lemmas 4.1, 4.2 and 4.3 imply the statement of this Lemma. For $x\in\underline{\gamma}$ we have

$$\begin{aligned}
&\mathbf{E}_x\sigma\\
&= \mathbf{E}_x\zeta([0,2\varepsilon^{1/4}])\mathbf{1}(X^\varepsilon_{\zeta([0,2\varepsilon^{1/4}])}\in\gamma) + \mathbf{E}_x(\zeta([0,2\varepsilon^{1/4}]) + \mathbf{E}_{X^\varepsilon_{\zeta([0,2\varepsilon^{1/4}])}}\sigma)\mathbf{1}(X^\varepsilon_{\zeta([0,2\varepsilon^{1/4}])}\in\underline{\underline{\gamma}})\\
&\le \sup_{x\in\mathcal{S}([0,2\varepsilon^{1/4}])}\mathbf{E}_x\zeta([0,2\varepsilon^{1/4}]) + \mathbf{E}_x(\sup_{y\in\underline{\underline{\gamma}}}\mathbf{E}_y\sigma(\varepsilon^{1/4}) + \sup_{x\in\underline{\gamma}}\mathbf{E}_x\sigma)\mathbf{1}(X^\varepsilon_{\zeta([0,2\varepsilon^{1/4}])}\in\underline{\underline{\gamma}})\\
&\le \sup_{x\in\mathcal{S}([0,2\varepsilon^{1/4}])}\mathbf{E}_x\zeta([0,2\varepsilon^{1/4}]) + (\sup_{y\in\underline{\underline{\gamma}}}\mathbf{E}_y\sigma(\varepsilon^{1/4}) + \sup_{x\in\underline{\gamma}}\mathbf{E}_x\sigma)\mathbf{P}_x(X^\varepsilon_{\zeta([0,2\varepsilon^{1/4}])}\in\underline{\underline{\gamma}})\ .
\end{aligned}\tag{4.20}$$

Taking a sup over all $x\in\underline{\gamma}$ we get

$$\sup_{x\in\underline{\gamma}}\mathbf{E}_x\sigma \le \frac{\sup_{x\in\mathcal{S}([0,2\varepsilon^{1/4}])}\mathbf{E}_x\zeta([0,2\varepsilon^{1/4}]) + \sup_{y\in\underline{\underline{\gamma}}}\mathbf{E}_y\sigma(\varepsilon^{1/4})\cdot\mathbf{P}_x(X^\varepsilon_{\zeta([0,2\varepsilon^{1/4}])}\in\underline{\underline{\gamma}})}{\mathbf{P}_x(X^\varepsilon_{\zeta([0,2\varepsilon^{1/4}])}\in\gamma)}\ .$$

Using Lemmas 4.1, 4.2 and 4.3 we see that the statement of this Lemma follows. (We choose $\varkappa = 1/8$ in Lemma 4.1.) □

Lemma 4.7. *We have*

$$\lim_{\varepsilon\downarrow 0}\sup_{x\in\mathcal{S}([0,\varepsilon^{1/4}])}\mathbf{E}_x\sigma = 0$$

uniformly in ε.

Proof. This is a consequence of Lemmas 4.1, 4.2 and 4.6. □

Lemma 4.8. *We have*

$$\lim_{\varepsilon\downarrow 0}\sup_{x\in[\mathcal{E}]}\mathbf{E}_x\sigma = 0$$

uniformly in ε.

Proof. This is a consequence of Lemmas 4.1 and 4.7. $\square$

We shall denote by $\mathcal{S}_k([-\varepsilon^{1/2},\varepsilon^{1/4}])$ the closed set bounded by the surfaces $\overline{\gamma}_k$ and $\underline{\gamma}_k$ and by $\mathcal{S}([-\varepsilon^{1/2},\varepsilon^{1/4}])=\cup_{k=1}^r\mathcal{S}_k([-\varepsilon^{1/2},\varepsilon^{1/4}])$. We notice that by the same reason as before, a coordinate $(\varphi_1^k,...,\varphi_{d-1}^k,H_k)$ exists in $\mathcal{S}_k([-\varepsilon^{1/2},\varepsilon^{1/4}])$. We denote $\mathcal{S}_k([0,\varepsilon^{1/4}])$, $\mathcal{S}([0,\varepsilon^{1/4}])$ (replacing $\overline{\gamma}_k$ by γ_k) and $\mathcal{S}_k([-\varepsilon^{1/2},0])$, $\mathcal{S}([-\varepsilon^{1/2},0])$ (replacing $\underline{\gamma}_k$ by γ_k) in a similar way.

Lemma 4.9. *We have*

$$\lim_{\varepsilon\downarrow 0}\sup_{x\in\gamma}\mathbf{E}_x\tau=0$$

uniformly in ε.

Proof. The proof of this lemma is very similar to and is a bit simpler than that of Lemma 4.6. We shall construct two barrier functions $u_k^{(3)}$ (for estimating the exit time from $\mathcal{S}_k([-\varepsilon^{-1/2},\varepsilon^{1/4}]))$ and $u_k^{(4)}$ (for the probability of hitting $\overline{\gamma}$).

For the construction of $u_k^{(3)}$ all the arguments of Lemma 4.2 can be carried here with γ replaced by $\overline{\gamma}$, $\underline{\gamma}$ replaced by γ and $\underline{\underline{\gamma}}$ replaced by $\underline{\gamma}$. We are working now with $\mathcal{S}_k([-\varepsilon^{1/2},\varepsilon^{1/4}])$ and $H_k\in[-\varepsilon^{1/2},\varepsilon^{1/4}]$. We apply formula (4.1) with the change of the estimates (4.2)$-$(4.6) as follows: when $x\in\mathcal{S}_k([0,\varepsilon^{1/4}])$ there is no change in the estimates; when $x\in\mathcal{S}_k([-\varepsilon^{1/2},0])$ we replace (4.2) and (4.3) by $K_1(x)=\dfrac{\partial}{\partial H_k}(K_1(x))=0$ and (4.4)$-$(4.6) remain the same. The function $K_4(x)$ is then defined in a same way as in Lemma 4.2 with an estimate $0<C_{34}\leq K_4(H_k)\leq C_{35}<\infty$ for $x\in\mathcal{S}_k([-\varepsilon^{1/2},0])$. Again we let

$$u_k^{(3)}(H_k)=\int_{-\varepsilon^{1/2}}^{H_k}\frac{K_6(\varepsilon)-y}{K_4(y)}dy$$

with

$$K_6(\varepsilon)=\left(\int_{-\varepsilon^{1/2}}^{\varepsilon^{1/4}}\frac{dy}{K_4(y)}\right)^{-1}\left(\int_{-\varepsilon^{1/2}}^{\varepsilon^{1/4}}\frac{ydy}{K_4(y)}\right).$$

It is then checked that $u_k(-\varepsilon^{1/2})=u_k(\varepsilon^{1/4})=0$ and $K_6(\varepsilon)\leq 2\varepsilon^{1/4}$. The estimate (4.9) is still working for $H_k\in[-\varepsilon^{1/2},\varepsilon^{1/4}]$. The estimates (4.10), (4.11) and (4.12) are still working. Let $\zeta([-\varepsilon^{1/2},\varepsilon^{1/4}])$ be the first time when the process X_t^ε starting from a point $x\in\gamma$, first hits $\underline{\gamma}$ or $\overline{\gamma}$. A similar statement of (4.13) is then obtained. We have

$$\lim_{\varepsilon\downarrow 0}\sup_{x\in\mathcal{S}([-\varepsilon^{1/2},\varepsilon^{1/4}])}\mathbf{E}_x\zeta([-\varepsilon^{1/2},\varepsilon^{1/4}])=0 \tag{4.21}$$

uniformly in ε.

The estimate of the hitting probability is a bit simpler. We construct a barrier function $u_k^{(4)}$ similarly as in Lemma 4.3. The function $K_7(H_k)$ is defined as in Lemma 4.3. But now we have the property that $|K_7(H_k)| \leq C$ for $H_k \in [-\varepsilon^{1/2}, 0]$. We let

$$u_k^{(4)}(H_k) = 1 - \frac{\displaystyle\int_{-\varepsilon^{1/2}}^{H_k} \exp\left(-\int_{-\varepsilon^{1/2}}^{y} K_7(z)dz\right) dy}{\displaystyle\int_{-\varepsilon^{1/2}}^{\varepsilon^{1/4}} \exp\left(-\int_{-\varepsilon^{1/2}}^{y} K_7(z)dz\right) dy} .$$

We have $u_k^{(4)}(-\varepsilon^{1/2}) = 1$, $u_k^{(4)}(\varepsilon^{1/4}) = 0$. We have

$$C_{36}\varepsilon^{1/2} \leq \int_{-\varepsilon^{1/2}}^{0} \exp\left(-\int_{-\varepsilon^{1/2}}^{y} K_7(z)dz\right) dy \leq C_{37}\varepsilon^{1/2} ,$$

and by the second inequality in Lemma 4.5 we see that

$$C_{38}\varepsilon^{1/2} \leq \int_{0}^{\varepsilon^{1/4}} \exp\left(-\int_{-\varepsilon^{1/2}}^{y} K_7(z)dz\right) dy \leq C_{39}\varepsilon^{1/2} .$$

These estimates ensure that an analogue of (4.19) works, but the lower bound is a positive constant, and hence situation is a bit simpler. We have

$$1 \geq \mathbf{P}_x(X^{\varepsilon}_{\zeta([-\varepsilon^{1/2},\varepsilon^{1/4}])} \in \overline{\gamma}) \geq u_k^{(4)}(0) \geq C_{40} > 0 . \tag{4.22}$$

uniformly in $x \in \gamma$ and $\varepsilon > 0$. The results (4.21), (4.22) and Lemma 4.8, combined with a similar analysis of (4.20) in Lemma 4.6, give the statement of this Lemma. □

Lemma 4.10. *We have*

$$\lim_{\varepsilon \downarrow 0} \sup_{x \in \overline{\gamma}} \mathbf{E}_x \sigma = 0$$

uniformly in ε.

Proof. This is a result in the same essence of Lemma 3.2 (formula (10)) of [4]. □

Lemma 4.11. *The process* $X^{\varepsilon}_{\sigma_n}$ *satisfies the Doeblin condition on* γ *uniformly in* ε.

Proof. For each fixed $\varepsilon > 0$ we have the ergodicity of the process X_t^{ε}. Uniformly in ε the Doeblin condition is satisfied for the process X_t^{ε} in $[\mathcal{E}]$ and each of these $[U_k]$'s for $k = 1, ..., r$. As we have Lemmas 4.8, 4.9 and 4.10, we see that the statement of this Lemma follows. □

BIBLIOGRAPHY

[1] Barlow, M., Pitman, J., Yor, M., On Walsh's Brownian motion, *Séminaire de Probabilitiés XXIII*, Springer Lecture Notes in Mathematics, **1372** (1989), pp. 275 – 293.

[2] Cerrai, S., Freidlin, M., Small mass asymptotics of a charged particle in a magnetic field and long-time influence of small perturbations, *Journal of Statistical Physics*, submitted.

[3] Dolgopyat, D., Koralov, L., Averaging of Hamiltonian flows with an ergodic component, *Annals of Probability*, **36** (2008), pp. 1999 – 2049.

[4] Dolgopyat, D., Koralov, L., Averaging of incompressible flows on two dimensional surfaces, preprint, submitted to *Journal of AMS*.

[5] Dynkin, E., One dimensional continuous strong Markov processes, *Theory of Probability and Its Applications* (English translation), **4**, 1, 1959, pp. 1 – 52.

[6] Ethier, S., Kurtz, T., *Markov processes, characterization and convergence*, John Wiley and Sons, London, 1986.

[7] Feller, W., Generalized second-order differential operators and their lateral conditions, *Ill. Journal of Math.*, **1** (1957), pp. 459 – 504.

[8] Freidlin, M., Some Remarks on the Smoluchowski-Kramers Approximation, *Journal of Statistical Physics*, **117**, No.314, pp. 617 – 634, 2004.

[9] Freidlin, M., On the factoriztion of a non-negative definite matrix, *Theory of Probability and its Applications*, **13** (1968), pp. 354 – 356 (English translation).

[10] Freidlin, M., Dirichlet's problem for an equation with periodic coefficients depending on a small parameter, *Theory of Probability and Its Applications*, **9**, pp.133-139, 1964.

[11] Freidlin, M., *Functional integration and partial differential equations*, Princeton University Press, 1985.

[12] Freidlin, M., Hu, W., Smoluchowski-Kramers approximation in the case of variable friction, *Journal of Mathematical Sciences*, **79**, No.1, November 2011, pp. 184 – 207, translated from *Problems in Mathematical Analysis*, **61**, October 2011 (in Russian).

[13] Freidlin, M., Weber, M., Random perturbations of dynamical systems and diffusion processes with conservation laws, *Probability Theory and Related Fields*, **128**, pp. 441 – 466, 2004.

[14] Freidlin, M., Wentzell, A., *Random Perturbations of Dynamical Systems*, 2-nd edition, Springer, 1998.

[15] Freidlin, M., Wentzell, A., Diffusion processes on graphs and the averaging principle, *Annals of Probability*, **21**, 4, 1993, pp. 2215 – 2245.

[16] Freidlin, M., Wentzell, A., Diffusion processes on graphs and the averaging principle, *Annals of Probability*, **21**, 4, 1993, pp. 2215 – 2245.

[17] Freidlin, M., Wentzell, A., Random Perturbations of Hamiltonian Systems, *Mem. of AMS*, **109** (1994).

[18] Freidlin, M., Wentzell, A., On the Neumann problem for PDE's with a small parameter and corresponding diffusion processes, *Probability Theory and Related Fields*, **152**, pp. 101 – 140, 2012.

[19] Freidlin, M., Wentzell, A., Necessary and Sufficient Conditions for Weak Convergence of One-Dimensional Markov Processes. *The Dynkin Festschrift: Markov Processes and their Applications* , Birkhauser, (1994) pp.95-109.

[20] Freidman, A., *Partial differential equations of parabolic type*, Englewood Cliffs, N.J., Prentice-Hall, 1964.

[21] Has'minskii, R.Z., Diffusion processes and elliptic differential equations degenerating on the boundary of the region, *Theory of Probability and its Applications*, **3** (1958), pp. 430 – 461 (English translation).

[22] Ikeda, N., Watanabe, S., *Stochastic Differential Equations and Diffusion Processes*, second edition, North-Holland Publishing Company, 1989.

[23] Itô, K., McKean, H.P. Jr., Brownian motion on a half line, *Illinois Journal of Mathematics*, **7**, pp. 181 – 231, 1963.

[24] Itô, K., McKean, H.P. Jr., *Diffusion processes and their sample paths*, Second Edition, Springer, 1974.

[25] Levinson, N., The first boundary-value problem for $\varepsilon\Delta u + Au_x + Bu_y + Cu = 0$ for small ε, *Annals of Mathematics*, **51**, pp. 428 − 445, 1950.

[26] Mandl, P., *Analytical Treatment of One-dimensional Markov Processes*, Springer, 1968.

[27] Molchanov, S., On a problem in the theory of diffusion processes, *Theory of Probability and Its Applications* (English translation), **9**, 1964, pp. 472 − 477.

[28] Molchanov, S., Ostrovskii, E., Symmetric stable processes as traces of degenerate diffusion processes, *Theory of Probability and Its Applications* (English translation), **14**, 1969, pp. 128 − 131.

[29] Revuz, D., Yor, M., *Continuous Martingales and Brownian Motion*, 3.ed., Springer, 1999.

[30] Ueno, T., The diffusion satisfying Wentzell's boundary condition and the Markov process on the boundary I, *Proc. Japan Acad.*, **36**, 10, 1960, pp. 533 − 538.

[31] Ueno, T., The diffusion satisfying Wentzell's boundary condition and the Markov process on the boundary II, *Proc. Japan Acad.*, **36**, 10, 1960, pp. 625 − 629.

[32] Wentzell, A., On boundary conditions for multidimensional diffusion processes, *Theory of Probability and Its Applications* (English translation), **4**, 2, 1959, pp. 164 − 177.

[33] Zorich, V.A., *Mathematical Analysis*, Part II, 4th corrected edition, Moscow, MCCME, 2002 (in Russian), Chinese translation, Higher Education Press, Beijing, 2006.