ABSTRACT

Title of dissertation: ANOMALOUS DIFFUSION
IN STRONG CELLULAR FLOWS:
AVERAGING AND HOMOGENIZATION

Zsolt Pajor-Gyulai, Doctor of Philosophy, 2015

Dissertation directed by: Professor Leonid Koralov
Department of Mathematics

This thesis considers the possible limit behaviors of a strong Hamiltonian cellular flow that is subjected to a Brownian stochastic perturbation.

Three possible limits are identified. When long timescales are considered, the limit behavior is described by classical homogenization theory. In the intermediate (finite) time case, it is shown that the limit behavior is anomalously diffusive. This means that the limit is given by a Brownian motion that is time changed by the local time of a process on the graph which is associated with the structure of the unperturbed flow lines (Reeb graph) that one obtains by Freidlin-Wentzell type averaging. Finally, we consider the case when the motion starts close to, or on, the cell boundary and derive a limit for the displacement on timescales of order ε^{α} where $\alpha \in (0,1)$ (modulo a logarithmic correction to compensate for the slowdown of the flow near the saddle points of the Hamiltonian). The latter two cases are novel results obtained by the author and his collaborators ([1]).

We also consider two applications of the above results to associated partial differential equation (PDE) problems. Namely, we study a two-parameter averaging-homogenization type elliptic boundary value problem and obtain a precise description of the limit behavior of the solution as a function of the parameters using the well-known stochastic representation. Additionally, we study a similar parabolic Cauchy problem.

ANOMALOUS DIFFUSION IN STRONG CELLULAR FLOWS: AVERAGING AND HOMOGENIZATION

by

Zsolt Pajor-Gyulai

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
2015

Advisory Committee:
Professor Leonid Koralov, Chair/Advisor
Professor Sandra Cerrai
Professor Mark Freidlin
Professor Konstantina Trivisa
Professor Ilya Ryzhov

Dedication

I dedicate this thesis to my sister Dóra, the strongest fighter I have ever known, and Gergő, who always inspires me to be a better big brother.

Acknowledgments

I owe my gratitude to all the people who have made this thesis possible and because of whom my graduate experience has been one that I will cherish forever.

First and foremost, I would like to thank my advisor, Professor Leonid Koralov, for giving me an invaluable opportunity to work on challenging and extremely interesting projects over the past five years. He has always made himself available for help and advice, and there has never been an occasion when I have knocked on his door and he has not made time for me. It has been a pleasure to work with and learn from such an extraordinary individual.

I would also like to thank Professor Dimitry Dolgopyat, who to some extent served as my coadvisor, for the numerous discussions and invaluable advice guiding me on my journey to become a professional researcher. My MSc advisor at the Budapest University of Technology and Economics, Professor Domokos Szász, launched me on my carreer for which I will be eternally grateful. Withouth his endless patience and guidence, I would have never been able to get into graduate school. Gratitude is also due to Professor Sandra Cerrai, Professor Konstantina Trivisa, and Professor Mark Freidlin for teaching me excellent courses, agreeing to serve on my thesis committee, and sparing their invaluable time reviewing this manuscript. Special thanks go to Professor Alexander Barg who agreed to serve as the substitute Dean's represetative during my defense on a short notice. Finally, acknowledgement is due to Professor Martin Hairer for being a collaborator developing the results presented in this thesis.

I would also like to acknowledge help and support from some of the staff members of the Department of Mathematics at UMD. Celeste Regalado, Haydee Hidalgo, Linette Berry, Sharon Welton, Rynetha Gumbs and Jamie Carrigan were always ready to help me with the most tedious administrative problems.

My fellow graduate students at the Mathematics Department have enriched my graduate life in many ways and deserve a special mention. Mickey Salins and I had a joint project that produced an accepted paper (and has the promise for many more), the quality of which could not have been as high without his relentless attention to detail. Scott Smith, Sam Punshon-Smith and Andreas Roy were always ready to discuss probability or show me something new in PDEs. My interaction with them has been extremely fruitful.

I owe my deepest thanks to my girlfriend, Colleen Stock, who has provided me with the emotional background through the final years of my PhD and was always there for me at the darkest parts to point out the light. Without her love, these years would have been a great deal more difficult. I would also like to thank her for reading through the manuscript looking for typos and making sure that it was actually written in English.

Words cannot express the gratitude I owe to my family - my mother, my father and my stepfather who have always stood by me and guided me through life, and have pulled me through against impossible odds at times.

My housemates at my place of residence have been a crucial factor in my success. I would like to express my gratitude to Katelyn Dolan, Shirley Pon, Kaitilyn Shulman, Oliver Rourke, Daniel Butler and Benjamin Stewart for their friendship and support.

I would like to acknowledge financial support from the Graduate School in the form of the Graduate Deans Dissertation Fellowship which allowed me to have my last year free of teaching which enabled me to focus strongly on my research. I also owe gratitude to the Department of Mathematics for awarding me the Brin Fellowship which allowed me to teach only every other semester throughout most of the program.

It is impossible to remember all, and I apologize to those I have inadvertently left out.

Table of Contents

List of Figures

Chapter 1: Introduction

1.1 Cellular flow

Consider a smooth, periodic and incompressible vector field v in the plane. It is well-known that there is a smooth function H, called the Hamiltonian or the stream function, such that

$$v = \nabla^{\perp} H = (-\partial_{x_2} H, \partial_{x_1} H).$$

We assume that v is periodic and let us also assume for simplicity that the period is one in both directions. This implies that the most general form H can take is

$$H(x_1, x_2) = H_{per}(x_1, x_2) + a x_1 + b x_2$$

where H_{per} is periodic with period one in both directions. In this thesis, we are going to consider the case when H itself is periodic ($a = b = 0$). As a consequence, the integral of v over the domain of periodicity is zero which means that the vector field has no overall drift. We will denote this domain of periodicity by $\mathcal{T}$, which can be viewed as a unit square or, alternatively, as a torus.

Our main additional structural assumption is that the critical points of H are non-degenerate and that there is a level set of H (say $H = 0$ without loss of generality) that contains some of the saddle points and forms a lattice in $\mathbb{R}^2$, thus dividing the plane into bounded sets that are invariant under the flow (see Figure 1.1).

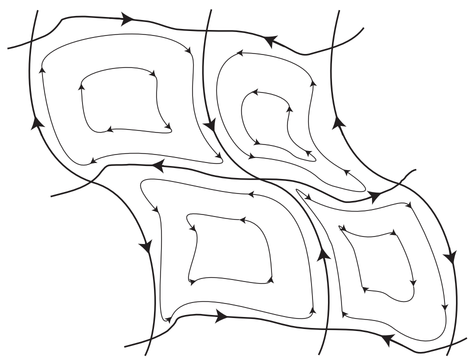

Figure 1.1: A period of the generic cellular flow

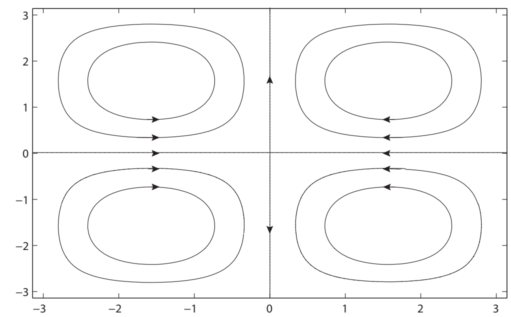

Figure 1.2: A period of the flow with stream function $H(x_1, x_2) = \sin(x_1)\sin(x_2)$

An example to keep in mind is given by $H(x_1, x_2) = \sin(x_1)\sin(x_2)$ (Figure 1.2). Cellular patterns occur sometimes in nature as well, such as the Rayleigh-Benard flow that occurs when a thin layer of fluid is heated from below and the warm liquid on the bottom exchanges places with the cold liquid on the top ([2]).

We call the family of mappings $x \to x_t^x$ defined by the ordinary differential equation

$$\dot{x}_t^x = v(x_t^x), \qquad x_0^x = x$$

the cellular flow associated to the vector field v.

1.2 Stochastically perturbed cellular flow

It is of primary interest to study how a flow described above behaves under a small stochastic perturbation which leads to the study of the family of stochastic differential equations

$$d\tilde{X}_t^{x,\varepsilon} = v(\tilde{X}_t^{x,\varepsilon})dt + \sqrt{\varepsilon}dW_t, \qquad X_0^{x,\varepsilon} = x \tag{1.1}$$

on some probability space $(\Omega, \mathcal{F}, \mathrm{P})$, where W_t is a two dimensional Brownian motion. We adopt the notation $\tilde{X}^{\mu,\varepsilon}$ to denote the process with a random initial condition distributed according to some measure μ. This convention will be used for all the processes appearing in this thesis.

Iit turns out, however, that on every finite time interval, the solution of 1.1 simply approaches the unperturbed flow.

Theorem 1.1 ([3] Chapter 1, Theorem 1.2)**.** *For any $T, \eta > 0$, we have*

$$\lim_{\varepsilon \downarrow 0} \mathrm{P}\left(\sup_{t \in [0,T]} |\tilde{X}_t^{x,\varepsilon} - x_t^x| > \eta \right) = 0$$

This is not surprising as the qualitative effect of the perurbation is a motion across the flow lines on a timescale of order ε^{-1}, which is much longer than the order one natural timescale of the deterministic motion. This means that if we want to study any non-trivial behavior (e.g. transitions between cells), we need to look at the process on longer timescales. Hence, we introduce $X_t^{x,\varepsilon} = \tilde{X}_{t/\varepsilon}^{x,\varepsilon}$ which leads to the family of stochastic differential equations

$$dX^{x,\varepsilon}(t) = \frac{1}{\varepsilon} v(X^{x,\varepsilon}(t))dt + dW_t, \qquad X^{x,\varepsilon}(0) = x. \tag{1.2}$$

This equation describes the behavior of tracer particles diffusing on the advective background of a strong flow described above.

The long time behavior of stochastic differential equations (SDE) like (1.2) has been studied by stochastic homogenization theory. For example, in [4], Freidlin proved that for any $\varepsilon > 0$, the diffusively rescaled process $\hat{X}_t^{x,\varepsilon,R} = X_{R^2t}^{x,\varepsilon}/R$ converges weakly to $\sqrt{D_{eff}(\varepsilon)}W'$, where W' is a two dimensional Brownian motion, and D_{eff} is the effective constant diffusivity matrix. Intuitively, the spatial rescaling can be thought of as an observer zooming out until the microscopic details of the cellular flow cannot be seen anymore and, for all intents and purposes, can be replaced by a homogeneous background. The outline of the proof of this result (presented as in [5]) is the following: Let χ be a periodic vector solution to the cell problem

$$-\Delta\chi + \frac{1}{\varepsilon}v\nabla\chi = -\frac{1}{\varepsilon}v.$$

Applying Itô's formula yields, after some elementary manipulations, that

$$X_t^{x,\varepsilon,R} - x = -\varepsilon\left[\chi\left(X_{t/\varepsilon^2}^{x,\varepsilon}\right) - \chi(x)\right] + \varepsilon\int_0^{t/\varepsilon^2}\sqrt{2}\,(I + \nabla\chi(X_s^{x,\varepsilon}))\,dW_s.$$

Since the corrector χ is bounded and independent of ε, the drift term converges to zero. By the ergodic theorem (note that since the flow is incompressible, the invariant measure of $X_t^{x,\varepsilon}$ is the Lebesgue measure), the quadratic variation of the diffusion term converges to $D_{eff}(\varepsilon)t$ where

$$D_{eff}(\varepsilon) = 2\delta_{i,j} + 2\int_{\mathcal{T}}\nabla\chi_i(x)\cdot\nabla\chi_j(x)dx.$$

Lévy's criterion now implies that the limit process is a Brownian motion with constant diffusion coefficient $\sqrt{D_{eff}(\varepsilon)}$.

The behavior of $D_{eff}(\varepsilon)$ when ε is small has been extensively studied in the literature under certain geometric restrictions (see e.g. the references in [5]). For the generic flow, it has been shown in [6] that there is some matrix D_0 such that

$$D_{eff}(\varepsilon) = \varepsilon^{-1/2}(D_0 + o(1)). \tag{1.3}$$

This in turn implies that the variance of $X_t^{x,\varepsilon}$ grows like $\mathcal{O}(t/\sqrt{\varepsilon})$ when t is large and ε is small. See [6], [7] and the references therein for further information on the long time behavior, while here we simply mention that these results have a wide variety of applications ranging from flame propagation to swimming (e.g [8], [9], [10], [11], [12], [13]).

In this thesis, our goal is to understand what happens in the small ε limit first when t is of order one and then on some even shorter timescales of order ε^α with $\alpha \in (0,1)$. In the former case, we obtain a limit theorem as $\varepsilon \downarrow 0$ provided that $X_t^{x,\varepsilon}$ is considered on spatial scales of order $\varepsilon^{-1/4}$, and identify the limiting process as a time changed Brownian motion. The time change arising in the construction of the limiting process is non-trivial and can be described as the local time of a diffusion process on a certain graph which we describe in Chapter 2.

On the other hand, the case when time is of order ε^α might seem trivial at first glance. Indeed, if the process starts from a generic point inside the cells, it will simply make a few rotations along the flowlines. However, if it starts close enough to (or on) the separatrix (also called heteroclinic orbits), some non-trivial movement is immediately possible. It was found that if we consider $X_t^{x,\varepsilon}$ on spatial scales of order $\varepsilon^{-(1-\alpha)/4}$ and time scales of order $\varepsilon^\alpha \log \varepsilon$ (which contains a logarithmic term to compensate for the slowdown of the deterministic flow around the saddles of H), the limiting process is once again a time changed Brownian motion. However, the time change now is arising from the local time of a Brownian motion on a similar graph as before but with infinite edges.

In both cases, the trajectories are diffusive, but the variance grows slower than proportional to t. This behavior is called anomalous diffusion and it was conjectured by W. Young ([14], [15]). The intuitive reasoning is that the time the process spends locked inside a cell is essentially wasted in terms of spatial movements. On the other hand, once the process is close to the separatrix, it can make an excursion involving many cell changes following a random walk pattern. This suggests Brownian limiting trajectories, but we need to time-change them with some quantity that keeps track of how much time the process spends in an active state around the separatrix. The variance of $X_t^{x,\varepsilon}$ over these intermediate and short time scales were shown to be proportional to $\sqrt{t}$ rigorously by G. Iyer and A. Novikov([5]).

Finally, we remark that it is also known what happens when $\alpha = 1$. In [16], Bakhtin showed that starting from a heteroclinic orbit, the process $X^{x,\varepsilon}_{\varepsilon|\log\varepsilon|t}$ converges in distribution in a certain special topology to a process that spends all the time on the set of saddle points and jumps instantaneously between them along the heteroclinic trajectories.

1.3 Connection to partial differential equations

Let $D_R \subseteq \mathbb{R}^2$ be obtained from a bounded smooth domain D by stretching it by a factor R. Consider the elliptic Dirichlet problem

$$\frac{1}{2}\Delta u^{\varepsilon,R} + \frac{1}{\varepsilon} v \nabla u^{\varepsilon,R} = -f\left(\frac{x}{R}\right) \text{ in } D_R, \qquad u^{\varepsilon,R}|_{\partial D_R} = 0, \tag{1.4}$$

where f is a bounded continuous function on D and v is a cellular flow as described above. For simplicity, assume that D contains the origin. This equation for example describes the concentration of some particles that are injected at rate f after which they diffuse on the strong convective background of the cellular flow when this concentration is kept at zero on the boundary of D_R.

There are two parameters in this problem: ε measures the inverse of the strength of the vector field, while R measures the size of the domain. For fixed R (for example when D_R coincides with exactly one cell) and $\varepsilon \downarrow 0$, solution to (1.4) becomes constant on stream lines. Indeed, multiplying by ε and letting $\varepsilon \downarrow 0$ formally gives us $v\nabla u = 0$. The precise values of the asymptotics of the solution on each streamline are determined by an ODE corresponding to the structure of the level sets according to classical averaging results [3].

If, on the other hand, ε is fixed and $R \uparrow \infty$, then the asymptotic behavior of u can be obtained by homogenization (e.g. [17–19]), i.e., by solving an elliptic problem on D with appropriately chosen constant coefficients.

It was shown in [20] that averaging and homogenization can also be used to study the two-parameter asymptotics in certain regimes. Namely, if $R^4 \log^2 R \leq c/(\varepsilon \log^2 \varepsilon)$ for some constant c as $1/\varepsilon, R \uparrow \infty$, then averaging theory applies. On the other hand, if $R^{4-\alpha} \geq 1/\varepsilon$ for some positive α, then homogenization type behavior is observed. The methods in [20] are analytic, based on investigating the asymptotic behavior of the principal Dirichlet eigenvalue of the elliptic operator, and it seems unlikely that they can be directly applied near the transition regime. To our knowledge, only numerical results were available in the intermediate cases [21, 22] up until now.

In this thesis, we study the two-parameter asymptotics using a probabilistic approach and we prove that the crossover from homogenization to averaging occurs when R is precisely of order $\varepsilon^{-1/4}$. In order to achieve this, we study the family of two dimensional diffusion processes 1.2 and then use the well known representation formula

$$u^{\varepsilon,R}(x) = \mathrm{E}\int_0^{\tau_{\partial D_R}(X^{x,\varepsilon}_\cdot)} f(X^{x,\varepsilon}_s/R)\,ds, \tag{1.5}$$

where $\tau_{\partial D_R}(\omega)$ is the first hitting of the boundary of D_R by the trajectory $\omega \in \mathcal{C}([0,T];\mathbb{R}^2)$. This is in accordance with the fact that the essence of the averaging and transition regimes can be captured by the mechanism of the exit of the process X^ε_t from D_R (see [20]). Let us remark that the case of non-zero boundary conditions can also be studied this way if we complement our results with the ones on the exit locations in [16].

We will also observe a similar phenomenon in connection with the corresponding parabolic problem

$$u^{\varepsilon,R}_t(x,t) = \frac{1}{2}\Delta u^{\varepsilon,R}(x,t) + \frac{1}{\varepsilon}v(x)\nabla u^{\varepsilon,R}(x,t), \qquad \tilde{u}^\varepsilon(x,0) = f(x/R). \tag{1.6}$$

where f is a continuous function that vanishes at infinity. Namely, we are going to show that $R \approx \varepsilon^{-1/4}$ is once again when the transition occurs between the averaging and homogenization regimes using the representation formula

$$u^{\varepsilon,R}_t(x,t) = \mathrm{E}f\left(\frac{X^{x,\varepsilon}_t}{R}\right).$$

Chapter 2: The main results

2.1 Diffusions on graphs

In this section, we describe the processes that give rise to the time change discussed in Section 1.2.

It is well-known that there is a graph G naturally associated to the structure of the level sets of H (see Figure 2.1). Namely, let $\mathcal{L} = \{x \in \mathbb{R}^2; H(x) = 0\}$ be the connected level set of H that contains a periodic array of saddle points, and denote the corresponding level set on the torus by $\mathcal{L}_{\mathcal{T}}$. Let A_i, for $i = 1, \ldots, n$, be the saddle points of H in $\mathcal{L}_{\mathcal{T}}$. Then $\mathcal{L}$ (or $\mathcal{L}_{\mathcal{T}}$) is the union of heteroclinic orbits connecting the A_i's and will be referred to as the separatrix. For notational simplicity, we assume that there are no homoclinic orbits, i.e. ones that connect a saddle to itself. Also, let U_i, for $i = 1, \ldots, n$, be the connected components of $\mathcal{T} \backslash \mathcal{L}_{\mathcal{T}}$. (There is no particular connection between the numbering of the U_i's and that of the A_i's. However, by Euler's theorem, there is actually the same number of them). For convenience, we also assume that there are no saddle points of H inside any U_i. The graph G will then have an interior vertex O and n edges connecting O with the exterior vertices corresponding to the extrema of H. Every other point on an edge corresponds to the appropriate connected component of a level curve of $|H|$. Accordingly, $|H|$ will serve as a local coordinate on each edge I_i which gives G a natural metric structure. In topology terminology, this is known as the Reeb graph of H.

Define

$$\Gamma : \mathcal{T} \to G, \qquad \Gamma(x) = (i, |H(x)|) \ \text{ if } x \in \overline{U}_i,$$

to be the mapping that takes U_i into an edge I_i of the graph in such a way that the entire set $\mathcal{L}_{\mathcal{T}}$ is mapped into O, the extrema inside each U_i are mapped into the corresponding exterior vertices, and each connected component of a level set of H is mapped into one point on the corresponding edge of the graph. Note that Γ is well defined as $\partial U_k \subseteq \mathcal{L}_{\mathcal{T}}$. Naturally, Γ can be extended periodically to the entire plane. We will refer to a generic point on the graph as $y = (i, z)$ with the identification $(1, O) \equiv ... \equiv (n, O)$.

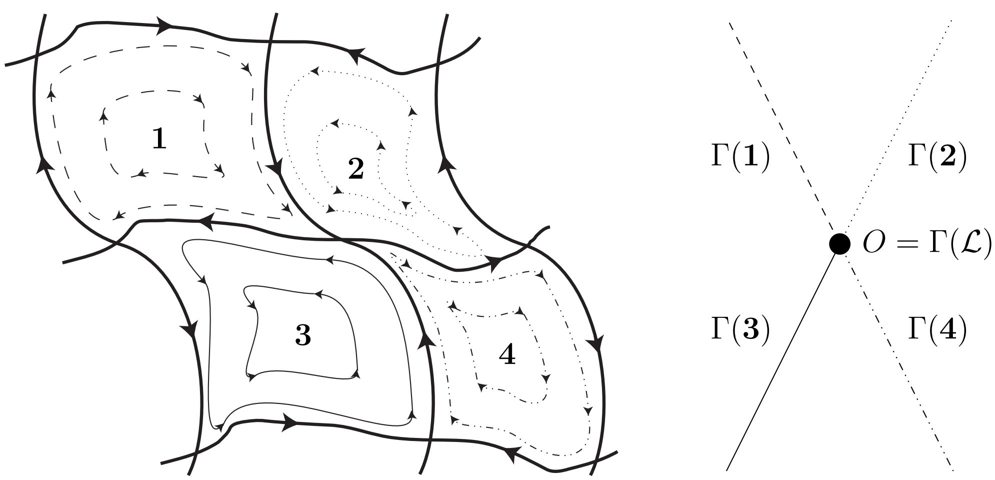

Figure 2.1: The graph corresponding to the structure of the level sets of H on $\mathcal{T}$.

It was shown in [3, Chapter 8] that the non-Markovian processes $\Gamma(X_t^{x,\varepsilon})$ converge in distribution, as $\varepsilon \downarrow 0$, to a diffusion on G. Let us describe this limiting process briefly. On the ith edge of the graph, the process is a diffusion with generator

$$\mathcal{A}_i = \frac{a(i,z)}{2}\frac{d^2}{dz^2} + b(i,z)\frac{d}{dz},$$

where the coefficients are determined by the Hamiltonian. The behavior of the process at the interior vertex can also be described in terms of H. More precisely, for a set of constants $\alpha_i > 0$ with $\sum_{i=1}^n \alpha_i = 1$, we can define an operator $\mathcal{A}$ on the domain $D(\mathcal{A})$ that consists of the functions F that satisfy:

a) $F \in \mathcal{C}(G)$ and furthermore $F \in \mathcal{C}^2(I_i)$ for each edge i,

b) $\mathcal{A}_i F(z)$, $z \in I_i$, which is defined on the union of the interiors of all the edges, can be extended to a continuous function on G,

c) $\sum_{i=1}^n \alpha_i D_i F(O) = 0$, where $D_i F(O)$ is the one-sided interior derivative of F along the edge I_i.

We then define the operator $\mathcal{A}$ by $\mathcal{A}F|_{I_i} = \mathcal{A}_i F|_{I_i}$. Below, we are going to write $y = (i,z)$ to refer to a point on G. As shown in [23], $\mathcal{A}$ generates a Fellerian Markov family Y_t^y on G. With these notations at hand, we can recall the following theorem also known as the averaging principle.

Theorem 2.1. *Freidlin-Wentzell (1994) The measures on $\mathcal{C}([0,\infty);G)$ induced by the processes $\Gamma(X_t^{x,\varepsilon})$ converge weakly to the one induced by the process $Y_t^{\Gamma(x)}$, provided*

$$a(i,z) = \frac{1}{T_i(z)}\int_{\gamma_i(z)} |\nabla H(x)| dl, \qquad b(i,z) = \frac{1}{2T_i(z)}\int_{\gamma_i(z)} \frac{\Delta H(x)}{|\nabla H(x)|} dl,$$

where $\gamma_i(z) = \Gamma^{-1}(i, y)$ *and*

$$T_i(z) = \int_{\gamma_i(y)} \frac{1}{|\nabla H(x)|} dl$$

is the period of the unperturbed motion on $\gamma_i(z)$. *The constants* $\{\alpha_i\}_{i=1}^n$ *are given by*

$$\alpha_i = \int_{\partial U_i} |\nabla H(x)| dl.$$

Note that the classical Freidlin-Wentzell theory requires $H(x) \to \infty$ as $|x| \to \infty$. Nevertheless, adapting the results for the compact setting on $\mathcal{T}$ is trivial. For a recent treatment, see [24] or the more general [25].

For the case when short times are considered, we need a different variant of the averaging principle. Namely, on timescales of order $\varepsilon^\alpha |\log \varepsilon|$, the process starting from $\mathcal{L}_\mathcal{T}$ only has time to wander to a distance of order $\varepsilon^{\alpha/2}$ away. This suggests that the behavior of $\varepsilon^{-\alpha/2} H(X_t^{x,\varepsilon})$ on these timescales might be non-trivial. Indeed, let $\bar{G}$ be a similar graph as G except that $\bar{G}$ has semi-infinite edges and define the mapping

$$\bar{\Gamma}^\varepsilon : \mathcal{T} \to \bar{G}, \qquad \bar{\Gamma}^\varepsilon(x) = (i, \varepsilon^{-\alpha/2}|H(x)|) \ \text{ if } x \in \overline{U}_i.$$

Also let μ^ε be a family of probability measures on $\mathbb{R}^2$ such that there is another probability measure ν on $\bar{G}$ so that the pushforward of μ^ε under Γ^ε converges weakly to ν, i.e.

$$\mu^\varepsilon \circ (\bar{\Gamma}^\varepsilon)^{-1} \Rightarrow \nu$$

as $\varepsilon \downarrow 0$. The next theorem, proved in Section 4.2, asserts that the image of the process $X_t^{x,\varepsilon}$, on the short timescales, converges to a generalized skew-Brownian motion on $\bar{G}$.

Theorem 2.2. *The measures on* $\mathcal{C}([0,\infty);\bar{G})$ *induced by the processes* $\bar{\Gamma}^{\varepsilon}(Z_t^{\mu^{\varepsilon},\varepsilon})$ *where* $Z_t^{x,\varepsilon} = X^{\mu^{\varepsilon},\varepsilon}\left(\frac{\alpha\varepsilon^{\alpha}|\log\varepsilon|t}{2}\right)$, *converge weakly to the one induced by the process* $\bar{Y}_t^{\nu}$ *which is a graph diffusion as described above with*

$$a(i,z) = C_i, \qquad b(i,z) = 0,$$

where

$$C_i = \frac{\alpha}{2}\oint_{\partial U_i} |\nabla H(x)| dl \cdot \lim_{\varepsilon\downarrow 0} \frac{|\log\varepsilon|}{T_i(\varepsilon^{\alpha/2})}, \tag{2.1}$$

and α_i *is as in Theorem 2.1.*

It is well-known from the theory of Hamiltonian systems that $T(z) \sim C|\log z|$ and therefore the limit in (2.1) exists and does not depend on α. Since $a(i,z)$ only depends on i, we simply have a skew Brownian motion in a generalized sense with constant diffusivity C_i on each edge. We emphasize that Theorem 2.2 is a new result that does not trivially follow from Theorem 2.1.

Recall that our goal is to obtain a quantity that captures the amount of time $X_t^{x,\varepsilon}$ spends around $\mathcal{L}$. In both cases, this is exactly the amount of time $\Gamma(X_t^{x,\varepsilon})$ (or $\bar{\Gamma}^{\varepsilon}(X_t^{x,\varepsilon})$) spends in a neighborhood of the interior vertex of G (or $\bar{G}$). This motivates the relevance of the following notion.

Definition 2.1. *The local time of a diffusion* Y^{y_0} *on a graph* G *is the unique non-negative random field*

$$L^{y_0} = \{L_t^{y_0}(y) \,:\, (t,y) \in [0,\infty) \times G\}$$

such that the following hold:

1. *The mapping* $(t,y) \to L_t^{y_0}(y)$ *is measurable, and* $L_t^{y_0}(y)$ *is adapted.*

2. *For each* $y \in G$*, the mapping* $t \to L_t^{y_0}(y)$ *is non-decreasing and constant on each open interval where* $Y_t^{y_0} \neq y$*.*

3. *For every Borel measurable* $f : G \to [0,\infty)$*, we have*

$$\int_0^t f(Y_s^{y_0})a(Y_s^{y_0})ds = 2\int_G f(y)L_t^{y_0}(y)dy \qquad a.s.$$

4. $L_t^{y_0}(y)$ *is almost surely jointly continuous in* t *and* y *for* $y \neq O$*, while*

$$L_t^{y_0}(O) = \sum_{i=1}^{n} \lim_{y \to O,\ y \in I_i} L_t^{y_0}(y)\ .$$

The existence and uniqueness of local time for diffusions on the real line is relatively well studied. These standard results, together with a straightforward modification of the discussion in Section 2 of [26], give the existence and uniqueness for the local time on the graph. Note that $a^{-1}(\cdot)$ is locally integrable near the interior vertex in both cases, which is sufficient for the method of [26] to work.

2.2 Main results

We are now ready to state our limit theorems for $X_t^{x,\varepsilon}$. For a positive definite symmetric matrix Q, let $\tilde{W}_t^Q$ be a two dimensional Brownian motion with covariance matrix Q. Assume that the families of processes Y_t^y and $\tilde{W}_t^Q$ are independent, and consider the process $\tilde{W}_{L_t^y}^Q$, where $L_t^y = L_t^y(O)$ is the local time of Y_t^y at the interior vertex.

Theorem 2.3. *There exists a strictly positive definite matrix Q such that the law of the process $\varepsilon^{1/4} X_t^{x,\varepsilon}$ converges, as $\varepsilon \downarrow 0$, to that of $\tilde{W}^Q_{L_t^{\Gamma(x)}}$.*

We remark that since G is compact and the interior vertex is accessible, $Y_t^{\Gamma(x)}$ is a positively recurrent process. Consequently, the law of large numbers applied to the additive functional $L_t^{\Gamma(x)}$ implies

$$\frac{L_t^{\Gamma(x)}}{t} \to \rho(O)$$

where ρ is the invariant density of Y_t^y which can be obtained as the unique normalized solution of the adjoint equation $\mathcal{A}^*\rho = 0$. (Strictly speaking, the law of large numbers has to be applied to the occupation measure $\int_A L_t^{\Gamma(x)}(y)dy = \int_0^t \chi_{\{Y_t^{\Gamma(x)} \in A\}} dt$.) This means that for large values of t, the variance will grow approximately as $\rho(O)t$. On the other hand, after hitting the interior vertex for the first time (which one can control by solving the appropriate ordinary differential equations corresponding to $\mathcal{A}$), the graph process will locally have the same path properties as the Brownian motion. This implies that the expected local time (and hence the variance) will grow proportionally to $\sqrt{t}$ establishing the conjectured anomalous diffusion behavior immediately after the hitting of $\mathcal{L}$. This is in accordance with the variance estimates in [5].

The anomalous diffusion is even more apparent if we zoom in on what happens after hitting the separatrix for the first time. To study this, it would be enough to let the process start from the separatrix, but for the sake of generality, we will only require that this starting point is at distance no more than of order $\varepsilon^{\alpha/2}$ from $\mathcal{L}$. This means that on timescales of order $\varepsilon^\alpha |\log \varepsilon|$, the separatrix can be reached due to the fluctuations of the noise.

Theorem 2.4. *Let the initial point be distributed according to a measure μ^ε, where $\mu^\varepsilon \circ (\Gamma^\varepsilon)^{-1}$ converges weakly to some probability measure ν on $\bar{G}$. If $\bar{L}_t^\nu$ is the local time of $\bar{Y}_t^\nu$ at the interior vertex, then there exists a strictly positive definite matrix Q such that the laws of the processes*

$$\varepsilon^{\frac{1-\alpha}{4}} X^{\mu^\varepsilon,\varepsilon}\left(\frac{\alpha\varepsilon^\alpha|\log\varepsilon|}{2}t\right)$$

converge, as $\varepsilon \downarrow 0$, to that of $\tilde{W}^Q_{\bar{L}_t^\nu}$ where $\tilde{W}_t^Q$ and $\bar{Y}_t^\nu$ are independent processes.

As we mentioned above, the logarithmic correction in the time scale is necessary to compensate for the slow down of the deterministic component around the saddle points of H.

We remark that if there are only one type of cell, $\bar{L}^y$ simply becomes a constant multiple of the Brownian local time, and the limit process is the so called fractional kinetic process of index 1/2 which arises as scaling limits of randomly trapped random walks with heavy tail trapping times in [27]. The connection is intuitively explained by noting that the time of an excursion of away from the interior vertex (when X is trapped inside a cell) is the excursion of a Brownian motion, and its length is accordingly heavy tailed with index 1/2.

Also note that by well-known Brownian formulas, $\mathrm{E}\bar{L}_t^O = c\sqrt{t}$ with some constant $c > 0$ which yields a variance for the limit process that is proportional to $\sqrt{t}$ for all times. This is once again an anomalous diffusion type behavior, and it is in accordance with the variance estimates in [5].

2.3 Implications for partial differential equations

Let us state the results on the partial differential equation problems introduced in Section 1.3, starting with the elliptic case (1.4).

The rough intuition is as follows: in the averaging regime ($\varepsilon \ll R^{-4}$), the process $X_t^{x,\varepsilon}$ revolves many times close to the flow lines within one cell, but once the separatrix is reached, the exit from D_R happens quickly. This follows from the typical $\varepsilon^{-1/4}$ fluctuation of the limiting process as the local time immediately becomes non-zero after the process reaches the boundary.

On the other hand, in the homogenization regime ($\varepsilon \gg R^{-4}$), the boundary is far away, and the process visits the interiors of many cells before the exit from D_R. This gives enough time for the process $L_t^{\Gamma(x)}$ to start growing nearly linearly in t, and therefore an overall Brownian behavior of $X_t^{x,\varepsilon}$ to set in. The mean exit time becomes infinite in the limit.

In the intermediate transition regime ($\varepsilon \approx R^{-4}$), the time required to leave D_R remains finite and is of the same order as the local time, although $L_t^{\Gamma(x)}$ is not directly proportional to t in this regime.

We will apply Theorem 2.3 in order to obtain the following asymptotic results for the solution of equation (1.4). The precise statement is as follows:

Theorem 2.5. *Let $\varepsilon \downarrow 0$ and $R = R(\varepsilon) \uparrow \infty$ in (1.4).*

1. *(Averaging regime) If $R\varepsilon^{1/4} \downarrow 0$, then*

$$u^{\varepsilon,R}(x) \to f(0) \cdot \mathrm{E}\tau_O(Y_\cdot^{\Gamma(x)}) \, ,$$

where τ_O is the first time when a process on G hits the interior vertex.

2. *(Transition regime) If $R\varepsilon^{1/4} \to C \in (0,\infty)$, then*

$$u^{\varepsilon,R}(x) \to \mathrm{E} \int_0^{\tau_{\partial D}} f\left(\tilde{W}^{Q/C^2}_{L_t^{\Gamma(x)}}\right) dt \ ,$$

with Q as in Theorem 2.3, where $\tau_{\partial D}$ is the first time the process $\tilde{W}^{Q/C^2}_{L_t^{\Gamma(x)}}$ hits the boundary of D.

3. *(Homogenization regime) There is a constant $c > 0$ such that if $R\varepsilon^{1/4} \uparrow \infty$, then*

$$(\varepsilon^{1/2}R^2)^{-1}u^{\varepsilon,R}(x) \to \mathrm{E} \int_0^{\tau_{\partial D}} f(\tilde{W}^{cQ}_t)\, dt \ , \tag{2.2}$$

where $\tilde{W}^{cQ}_t$ is a Brownian motion with covariance cQ and $\tau_{\partial D}$ is the first time the process $\tilde{W}^{cQ}_t$ hits the boundary of D.

Remark 2.1. *Note that there is no x dependence on the right hand side of (2.2). If we scale the problem back to the original domain D and then normalize appropriately, the above result gives us that the limit is the solution of a constant coefficient Dirichlet problem on D evaluated at the origin. To get the values of this solution at another point x, we must apply the result to the shifted domain $D - x$. This way we can prove that*

$$(\varepsilon^{1/2}R^2)^{-1}u^{\varepsilon,R}(Rx) \to \mathrm{E} \int_0^{\tau_{\partial D}} f(x + \tilde{W}^{cQ}_t)\, dt \qquad \text{as} \ \ \varepsilon \downarrow 0, R \uparrow \infty,$$

which contains the classical homogenization result. Here $\tau_{\partial D}$ is the first time when the process $x + \tilde{W}^{cQ}_t$ hits the boundary of D.

Remark 2.2. *Although it is not an aim of the present paper, Theorem 2.3 can also be used to derive asymptotics for PDEs with a periodic right hand side. These techniques are suitable for investigating equations with non-zero boundary data as well when combined with the results of [16].*

Let us consider next the parabolic problem (1.6). We will see once again that in the averaging regime, the limit of the solution is the solution of an equation that has one less spatial dimension while in the homogenization regime we obtain an effective equation that has the same dimensions as the one before taking the limit.

The intuition behind the results is very similar to the elliptic case. In the averaging regime, reaching the separatrix immediately implies that the process is of distance $\mathcal{O}(\varepsilon^{-1/4})$ from the origin which is much larger than R, and therefore $X_t^{x,\varepsilon}/R$ is outside the region where f is significant. On the other hand, in the homogenization regime, we pick up contributions from the entire life of the process. After rescaling time and space appropriately, the major contribution comes from the long time behavior of $\tilde{W}^{Q}_{L_t^{\Gamma(x)}}$ which is simply Brownian. The precise results are summarised in the next theorem.

Theorem 2.6. *Let $\varepsilon \downarrow 0$ and $R = R(\varepsilon) \uparrow \infty$ in (1.6).*

1. *(Averaging regime) If $R\varepsilon^{1/4} \downarrow 0$, then*

$$u^{\varepsilon,R}(x,t) \to f(0) \cdot \mathrm{P}\left(\tau_O(Y_\cdot^{\Gamma(x)} \geq t\right) ,$$

where τ_O is the first time when a process on G hits the interior vertex.

2. *(Transition regime) If* $R\varepsilon^{1/4} \to C \in (0,\infty)$*, then*

$$u^{\varepsilon,R}(x,t) \to \mathrm{E} f(\tilde{W}^{Q/C^2}_{L_t^{\Gamma(x)}}) \ ,$$

with Q as in Theorem 2.3.

3. *(Homogenization regime) There is a constant* $c > 0$ *such that if* $R\varepsilon^{1/4} \uparrow \infty$*, then*

$$\varepsilon^{-1/2} R^{-2} u^{\varepsilon,R}(Rx, \varepsilon^{1/2} R^2 t) \to \mathrm{E} f(x + \tilde{W}^{cQ}_t) \ . \tag{2.3}$$

Remark 2.3. *It is not proved strictly speaking, however, it is clear that both in Theorem 2.5 and Theorem 2.6, we have* $c = \rho(O)$*. Comparing this with (1.3), we get*

$$D_{eff}(\varepsilon) = \varepsilon^{-1/2}(\rho(0)Q + o(1)).$$

In both problems, the transition case is interesting. Formally, it was derived in [14] and [15] that the corresponding parabolic equation involves a fractional time derivative of order 1/2. The precise mathematical treatment is a future goal of the author.

Chapter 3: Intermediate timescales

3.1 Displacement when the process is near the separatrix

In this section, we study the behavior of the process when it is close to the separatrix. The process spends most of the time in the interiors of the cells where no cell changes are possible. However, when the process leaves the cell interior, rapid displacement occurs along the separatrix. We will show what happens during one excursion, i.e., between the time when the process hits the separatrix and the time when it goes back to the interior of the domain (the exact meaning of the latter will be explained below).

First, we need some notations. For any two saddle points, introduce $\gamma(A_i, A_j)$ as the set of points in $\mathcal{L}_{\mathcal{T}}$ that get taken to A_j by the flow $\dot{x} = v(x)$ and to A_i by the flow $\dot{x} = -v(x)$. Since we assumed that the separatrices do not form loops, we always have $\gamma(A_i, A_i) = \emptyset$.

In a neighborhood of each curve $\gamma(A_i, A_j)$, we can consider a smooth coordinate change $(x_1, x_2) \to (H, \theta)$ defined by the conditions $|\nabla\theta| = |\nabla H|$ and $\nabla\theta \perp \nabla H$ on $\gamma(A_i, A_j)$. This way θ is defined up to multiplication by -1 and up to an additive constant.

Let $V^\delta = \{x \in \mathbb{R}^2 : |H(x)| \leq \delta\}$. If δ is sufficiently small, we can make a continuous coordinate change $(x_1, x_2) \to (H, \theta)$ in $V^\delta \cap \overline{U}_k$. Here θ takes values in $[0, \int_{\partial U_k} |\nabla H| dl]$, with the endpoints of the interval identified, and satisfies the conditions (a) θ is smooth in a neighborhood of $\gamma(A_i, A_j)$ for each A_i, A_j such that $\gamma(A_i, A_j) \subset \overline{U}_k$, (b) $|\nabla\theta| = |\nabla H|$ on $\gamma(A_i, A_j)$, and (c) θ is constant on curves perpendicular to the level sets of H. Note that this way θ is defined uniquely up to the curve corresponding to $\theta = 0$ and the direction in which θ increases. Using these new coordinates, we can define what it means for the process to pass a saddle point. Namely, let

$$B(A_i, U_k) = \{x \in V^\delta \cap \overline{U}_k \,:\, \theta(x) = \theta(A_i)\}\,, \qquad B(A_i) = \bigcup_{k: A_i \in \partial U_k} B(A_i, U_k).$$

Observe that $B(A_i, U_k)$ is a curve in U_k transversal to the flow with an endpoint being the saddle point A_i.

Let $\pi : \mathbb{R}^2 \to \mathcal{T}$ be the quotient map from the plane to the torus and, for simplicity, let us denote $\pi(V^\delta)$ by V^δ again. Introduce the stopping times $\alpha_0^{x,\delta,\varepsilon} = 0$, $\beta_0^{x,\delta,\varepsilon} = \inf\{t \ge 0 \,:\, X_t^{x,\varepsilon} \in \mathcal{L}\}$ and recursively define $\alpha_n^{x,\delta,\varepsilon}$ and $\beta_n^{x,\delta,\varepsilon}$ as follows. Given $\beta_{n-1}^{x,\delta,\varepsilon}$, find i and j such that $\pi\left(X^{x,\varepsilon}_{\beta_{n-1}^{x,\delta,\varepsilon}}\right) \in \gamma(A_i, A_j)$. Then we define

$$\alpha_n^{x,\delta,\varepsilon} = \inf\left\{t \ge \beta_{n-1}^{x,\delta,\varepsilon} \,:\, \pi(X_t^{x,\varepsilon}) \in \bigcup_{k \ne i} B(A_k) \cup \partial V^\delta\right\},$$

$$\beta_n^{x,\delta,\varepsilon} = \inf\{t \ge \alpha_n^{x,\delta,\varepsilon} : X_t^{x,\varepsilon} \in \mathcal{L}\}.$$

In other words, $\alpha_n^{x,\delta,\varepsilon}$ is the first time after $\beta_{n-1}^{x,\delta,\varepsilon}$ that the process either hits ∂V^δ, or goes past a saddle point different from the one behind $X^{x,\varepsilon}_{\beta_{n-1}^{x,\delta,\varepsilon}}$.

We introduce another pair of sequences of stopping times corresponding to successive visits to $\mathcal{L}$ and ∂V^δ. Namely, let $\mu_0^{x,\delta,\varepsilon} = 0$, $\sigma_0^{x,\delta,\varepsilon} = \beta_0^{x,\delta,\varepsilon}$, and recursively define

$$\mu_n^{x,\delta,\varepsilon} = \inf\{t \geq \sigma_{n-1}^{x,\delta,\varepsilon} \,:\, X_t^{x,\varepsilon} \in \partial V^\delta\}, \quad \sigma_n^{x,\delta,\varepsilon} = \inf\{t \geq \mu_n^{x,\delta,\varepsilon} \,:\, X_t^{x,\varepsilon} \in \mathcal{L}\}.$$

Let

$$S_n^{x,\delta,\varepsilon} = X_{\sigma_n^{x,\delta,\varepsilon}}^{x,\varepsilon} - X_{\sigma_{n-1}^{x,\delta,\varepsilon}}^{x,\varepsilon},\ n \geq 1, \quad T_n^{x,\delta,\varepsilon} = \sigma_n^{x,\delta,\varepsilon} - \mu_n^{x,\delta,\varepsilon},\ n \geq 0,$$

be the displacement between successive visits to $\mathcal{L}$ and the time spent on the n-th downcrossing of V^δ, respectively. We will use the following notion of uniform weak convergence of probability measures.

Definition 3.1. *Given two families of random variables, $f^{x,\varepsilon}$ and g^x, with values in a metric space M and indexed by a parameter x, we will say that $f^{x,\varepsilon}$ converge to g^x in distribution uniformly in x if*

$$\mathrm{E}\varphi(f^{x,\varepsilon}) \to \mathrm{E}\varphi(g^x)\ ,$$

as $\varepsilon \to 0$, uniformly in x for each $\varphi \in \mathcal{C}_b(M)$.

Let $\eta^{x,\delta,\varepsilon}$ be the random vector with values in $\{1, \ldots, n\}$ defined by

$$\eta^{x,\delta,\varepsilon} = i \ \text{ if } \ X_{\mu_1^{x,\delta,\varepsilon}}^{x,\varepsilon} \in U_i, \quad i = 1, \ldots, n,$$

i.e., $\eta^{x,\delta,\varepsilon} = i$ if the process ends up in U_i after the first upcrossing of V^δ. The main result of this section is the following theorem:

Theorem 3.1. *There are a 2×2 non-degenerate matrix Q, a vector $(p_1, \ldots, p_n)$, and functions $a(\delta)$, $b_1(\delta), \ldots, b_n(\delta)$ that go to zero as $\delta \to 0$, such that*

$$(\varepsilon^{1/4} S_1^{x,\delta,\varepsilon}, \eta^{x,\delta,\varepsilon}) \to \big(\sqrt{\delta}(1 + a(\delta))\sqrt{\xi} N(0, Q), \eta^\delta\big) \tag{3.1}$$

in distribution as $\varepsilon \downarrow 0$*, uniformly in* $x \in \mathcal{L}$ *for all sufficiently small* $\delta > 0$*.* ξ *is an exponential random variable with parameter one,* N *is a two dimensional normal with covariance matrix* Q*, independent of* ξ*, and* η^δ *is a random vector with values in* $\{1, \dots, n\}$ *independent of* ξ *and* N *such that* $\mathrm{P}(\eta^\delta = i) = p_i + b_i(\delta)$*.*

Before proving Theorem 3.1, let us briefly discuss one implication. Let $T^{x,\varepsilon} := T_0^{x,\delta,\varepsilon}$ be the time it takes the process starting at x to reach the separatrix. Let $\bar{T}^y$ be the time it takes the limiting process Y_t^y on the graph to reach the vertex O. By the averaging principle [23], $T^{x,\varepsilon} \to \bar{T}^{\Gamma(x)}$ in distribution uniformly in $x \in \mathcal{T}$. This, together with Theorem 3.1 and the strong Markov property of the process imply the following lemma.

Lemma 3.1. *For fixed* m *and* δ*, the random vectors*

$$(T_0^{x,\delta,\varepsilon}, \varepsilon^{1/4} S_1^{x,\delta,\varepsilon}, T_1^{x,\delta,\varepsilon}, \varepsilon^{1/4} S_2^{x,\delta,\varepsilon}, \dots, T_{m-1}^{x,\delta,\varepsilon}, \varepsilon^{1/4} S_m^{x,\delta,\varepsilon})$$

converge, as $\varepsilon \downarrow 0$*, to a random vector with independent components. The limiting distribution for each of the components* $\varepsilon^{1/4} S_1^{x,\delta,\varepsilon}, \dots, \varepsilon^{1/4} S_m^{x,\delta,\varepsilon}$ *is given by Theorem 3.1, i.e., it is equal to the distribution of* $\sqrt{\delta}(1 + a(\delta))\sqrt{\xi} N(0, Q)$*. The limiting distribution of* $T_0^{x,\delta,\varepsilon}$ *is the distribution of* $\bar{T}^{\Gamma(x)}$*. The limiting distribution for each of the components* $T_1^{x,\delta,\varepsilon}, \dots, T_{m-1}^{x,\delta,\varepsilon}$ *is equal to the distribution of* $\bar{T}^\zeta$*, where* ζ *is a random initial point for the process on the graph, chosen to be at distance* δ *from the vertex* O*, in such a way that* ζ *belongs to the* i*-th edge with probability* $p_i + b_i(\delta)$*.*

Proof. By the averaging principle ([23]), $T_0^{x,\delta,\varepsilon} \to \bar{T}^{\Gamma(x)}$ in distribution uniformly in $x \in \mathcal{T}$. The convergence of other components of the random vector to their respective limits follows from Theorem 3.1. The independence of the components of the limiting vector immediately follows from the strong Markov property of the process $X_t^{x,\varepsilon}$ and the fact that the convergence in Theorem 3.1 is uniform with respect to x. $\square$

We will prove Theorem 3.1 by proving a more abstract lemma on Markov chains with a small probability of termination at each step, and demonstrating that the conditions of the lemma are satisfied in the situation of Theorem 3.1.

Let M be a metric space which can be written as a disjoint union

$$M = X \sqcup C_1 \sqcup \ldots \sqcup C_n ,$$

where the sets C_i are closed. Assume also that X is a σ-locally compact separable subspace, i.e., locally compact that is the union of countably many compact subspaces. Let $p_\varepsilon(x, dy)$, $0 \leq \varepsilon \leq \varepsilon^0$, be a family of transition probabilities on M and let $g \in \mathcal{C}_b(M, \mathbb{R}^2)$. Later, $p_\varepsilon(x, dy)$ will come up as transition probabilities of a certain discrete time process associated to $X_t^{x,\varepsilon}$. We assume that the following properties hold:

1. $p_0(x, X) = 1$ for all $x \in M$ and $p_\varepsilon(x, X) = 1$ for all $x \in M \backslash X$.

2. $p_0(x, dy)$ is weakly Feller, meaning the map $x \mapsto \int_M f(y) p_0(x, dy)$ belongs to $\mathcal{C}_b(M)$ if $f \in \mathcal{C}_b(M)$.

3. There exist bounded continuous functions $h_1, \ldots, h_n : X \to [0, \infty)$ such that

$$\varepsilon^{-\frac{1}{2}} p_\varepsilon(x, C_i) \to h_i(x), \quad \text{uniformly in } x \in K \text{ if } K \subseteq X \text{ is compact},$$

while $\sup_{x \in X} |\varepsilon^{-\frac{1}{2}} p_\varepsilon(x, C_i)| \leq c$ for some positive constant c. We also have

$$J(x) := h_1(x) + \ldots + h_n(x) > 0 \quad \text{for } x \in X. \tag{3.2}$$

4. $p_\varepsilon(x, dy)$ converges weakly to $p_0(x, dy)$ as $\varepsilon \to 0$, uniformly in $x \in K$ for $K \subseteq X$ compact.

5. The transition functions satisfy a strong Doeblin condition uniformly in ε. Namely, there exist a probability measure η on X, a constant $a > 0$, and an integer $m > 0$ such that

$$p_\varepsilon^m(x, A) \geq a\eta(A) \quad \text{for} \quad x \in M,\ A \in \mathcal{B}(X),\ \varepsilon \in [0, \varepsilon_0].$$

It then follows that for every ε, there is a unique invariant measure $\lambda^\varepsilon(dy)$ on M for $p_\varepsilon(x, dy)$, and the associated Markov chain is uniformly exponentially mixing, i.e., there are $\Lambda > 0, c > 0$, such that

$$|p_\varepsilon^k(x, A) - \lambda^\varepsilon(A)| \leq ce^{-\Lambda k} \quad \text{for all } x \in M,\ A \in \mathcal{B}(M),\ \varepsilon \in [0, \varepsilon_0].$$

6. The function g is such that $\int_M g \, d\lambda^\varepsilon = 0$ for each $\varepsilon \in [0, \varepsilon_0]$.

Lemma 3.2. *Suppose that assumptions 1–6 above are satisfied and let $Z_k^{x,\varepsilon}$ be the Markov chain on M starting at x, with transition function p_ε. Let $\tau = \tau(x, \varepsilon)$ be the first time when the chain reaches the set $C = C_1 \sqcup \ldots \sqcup C_n$. Let $e(Z_k^{x,\varepsilon}) = i$ if $Z_k^{x,\varepsilon} \in C_i$. Then*

$$\left(\varepsilon^{\frac{1}{4}} (g(Z_1^{x,\varepsilon}) + \ldots + g(Z_\tau^{x,\varepsilon})), e(Z_\tau^{x,\varepsilon}) \right) \to (F_1, F_2) \tag{3.3}$$

in distribution, uniformly in $x \in X$, where F_1 takes values in $\mathbb{R}^2$, F_2 takes values in $\{1, \ldots, n\}$, and F_1 and F_2 are independent. The random variable F_1 is distributed *as* $(\xi/\int_X J d\lambda^0)^{\frac{1}{2}} N(0,\bar{Q})$*, where* ξ *is exponential with parameter one independent of* $N(0,\bar{Q})$ *and* $\bar{Q}$ *is the matrix such that*

$$(g(Z_1^{x,0}) + \ldots + g(Z_k^{x,0}))/\sqrt{k} \to N(0,\bar{Q}) \quad \text{in distribution as } k \to \infty.$$

The random variable F_2 *satisfies* $\mathrm{P}(F_2 = i) = \int_X h_i \, d\lambda^0 / \int_X J \, d\lambda^0$*,* $i = 1, \ldots, n$*.*

Before we proceed with the proof of Lemma 3.2, let us show that it does indeed imply Theorem 3.1.

Proof of Theorem 3.1. Let $\mathcal{L}_0 = \mathcal{L} \backslash \{A \in \mathbb{R}^2 : \pi(A) \in \{A_i, i = 1, \ldots, n\}\}$. Define $\bar{M} = \mathcal{L}_0 \sqcup \partial V^\delta$. Let us define a family of transition functions $\bar{p}_\varepsilon(x, dy)$ on $\bar{M}$. For $x \in \mathcal{L}_0$, we define $\bar{p}_\varepsilon(x, dy)$ as the distribution of $X_\tau^{x,\varepsilon}$ with $\tau = \mu_1^{x,\delta,\varepsilon} \wedge \beta_1^{x,\delta,\varepsilon}$. In other words, it is the measure induced process that stops when it reaches either the boundary of V^δ or the separatrix after passing by a saddle point. For $x \in \partial V^\delta$, let $\bar{p}_\varepsilon(x, dy)$ coincide with the distribution of $X_{\bar{\tau}}^{x,\varepsilon}$ with $\bar{\tau} = \beta_0^{x,\delta,\varepsilon}$, i.e., the measure induced by the process that stops when it reaches the separatrix. Since almost every trajectory of $X_t^{x,\varepsilon}$ that starts outside of the set of saddle points does not contain saddle points, $\bar{p}_\varepsilon$ is indeed a stochastic transition function. Let $\bar{Z}_k^{x,\varepsilon}$ be the corresponding Markov chain starting at $x \in \bar{M}$.

While we introduced $\bar{M}$ as a subset of $\mathbb{R}^2$, it is more convenient to keep track of $\pi(\bar{Z}_k^{x,\varepsilon})$ and the latest displacement separately. Let $\varphi : \bar{M} \to M := \pi(\bar{M}) \times \mathbb{Z}^2$ map $x \in \bar{M}$ into $(\pi(x), ([x_1], [x_2]))$, where $[x_1]$ and $[x_2]$ are the integer parts of the first and second coordinates of x. Define the Markov chain $Z_k^{x,\varepsilon}$ on M via

$$Z_0^{\pi(x),\varepsilon} = (\pi(x), 0), \quad Z_k^{\pi(x),\varepsilon} = (\varphi_1(\bar{Z}_k^{x,\varepsilon}), \varphi_2(\bar{Z}_k^{x,\varepsilon}) - \varphi_2(\bar{Z}_{k-1}^{x,\varepsilon})), \quad k \geq 1.$$

Let $X = \pi(\mathcal{L}_0) \times \mathbb{Z}^2 = (\mathcal{L}_\mathcal{T} \backslash \{A_1, \ldots, A_n\}) \times \mathbb{Z}^2$ and $C_i = (\pi(\partial V^\delta) \cap U_i) \times \mathbb{Z}^2$. Thus $M = X \sqcup C_1 \sqcup \ldots \sqcup C_n$ as required. The transition functions $p_\varepsilon(x, dy)$ are defined as the transition functions for the Markov chain $Z_k^{x,\varepsilon}$.

For $x = (q, \xi) \in M$, define $g((q, \xi)) = \xi \in \mathbb{Z}^2$, which corresponds to the integer part of the displacement during the last step if the chain is viewed as a process on $\mathbb{R}^2$, where only the integer parts of the initial and end points are counted. From the definition of the stopping times $\beta_k^{x,\delta,\varepsilon}$, it follows that $\varphi_2(\bar{Z}_k^{x,\varepsilon}) - \varphi_2(\bar{Z}_{k-1}^{x,\varepsilon})$ can only take a finite number of values (roughly speaking, the process $X_t^{x,\varepsilon}$ makes transitions from one domain of periodicity to a neighboring one or to itself between the times $\beta_k^{x,\delta,\varepsilon}$ and $\beta_{k+1}^{x,\delta,\varepsilon}$). Therefore, $g(Z_k^{\pi(x),\varepsilon})$ is bounded almost surely, uniformly in x and k. Also, it is continuous in the product topology of $\pi(\bar{M}) \times \mathbb{Z}^2$.

The paper [6] contains some detailed results on the behavior of the process $X_t^{x,\varepsilon}$ near the separatrix. The main idea behind those results is that the process can be considered in (H, θ) coordinates in the vicinity of $\mathcal{L}$. In those coordinates, after an appropriate re-scaling, the limiting process (as $\varepsilon \to 0$) is easily identified.

Note that in [6], the width of the separatrix region is of order ε^{α_1} with some $\alpha_1 \in (1/4, 1/2)$, while here, it is of width δ. The results we are about to refer to can all be easily seen to hold with ε^{α_1} replaced by δ, our current case being simpler.

The existence of the limit of the transition functions p_ε in the sense of Assumption (4) was justified in [6, Lemma 3.1]. This limit is denoted by p_0. An explicit formula for the density of p_0 was also provided ([6, formula (9)]), which implies that Assumption (2) is satisfied. Observe that the probability of $\beta_1^{x,\delta,\varepsilon}$ being less than $\mu_1^{x,\delta,\varepsilon}$ tends to one as $\varepsilon \downarrow 0$ uniformly in $x \in \mathcal{L}$ by [6, formula (26)]. This implies

Property (1).

Let us sketch the proof of the Doeblin condition (5). Fix $a_1, a_2, a_3 \in \gamma(A_i, A_j) \subset \overline{U}_k$ with some A_i, A_j, and U_k. The points are ordered in the direction of the flow v. Let γ' be the part of $\gamma(A_i, A_j)$ that lies between a_2 and a_3. Let $J = \{(H, \theta) \in V^\delta \bigcap U_k : \sqrt{\varepsilon} \leq H \leq 2\sqrt{\varepsilon}, \theta = \theta(a_1)\}$. We can assume that a_1, a_2, and a_3 are chosen in such a way that φ_2 is constant on $J \bigcup \gamma'$. It is not difficult to show that there is $m > 0$ such that

$$\mathrm{P}\left(\varphi_2(X_t^{x,\varepsilon}) = \varphi_2(x), \quad X_t^{x,\varepsilon} \in J \ \text{ for some } \alpha_m^{x,\delta,\varepsilon} < t < \beta_m^{x,\delta,\varepsilon}\right) > c > 0$$

for all $x \in \mathcal{L}$. Roughly speaking, this statement means that the process has a positive chance of going to a particular curve at a distance $\sqrt{\varepsilon}$ from the separatrix, transversal to the flow lines, prior to passing by m saddle points. This is not surprising since the motion consists of advection with speed of order $1/\varepsilon$ and diffusion of order one. The proof follows along the same lines as the proof of Lemma 3.1. in [6]. Now the distribution of $X^{x,\varepsilon}_{\beta^{x,\delta,0}}$ has a component with density strictly bounded from below on γ', uniformly in $x \in J$, as follows from (63) in [24]. This implies the Doeblin condition for $Z_k^{x,\varepsilon}$.

With our definition of g,

$$\int_M g(x)\, d\lambda^\varepsilon(x) \left(\int_M \mathrm{E}\overline{\tau}^x \, d\lambda^\varepsilon(x)\right)^{-1} = \lim_{t\to\infty} (\mathrm{E}X_t^{x,\varepsilon}/t),$$

where $\overline{\tau}^x$ is the random transition time for our Markov chain, and the right hand side is the effective drift for the original process starting from an arbitrary point x. Note that $\lim_{t\to\infty}(\mathrm{E}X_t^{x,\varepsilon}/t) = \int_{\mathcal{T}} v(x)dx = 0$, which implies Property (6).

Property (3) follows from [6, Lemma 4.1 and Lemma 4.3]. Indeed, the former lemma describes the asymptotics of the distribution of $H(X^{x,\varepsilon}_{\alpha_1^{x,\delta,\varepsilon}})$, while the latter one describes the probability of the process starting at x to exit the boundary layer before reaching the separatrix, assuming that $H(x)$ is fixed. The two lemmas, together with the Markov property of the process, imply Property (3). The functions $h_i(x) = h_i^\delta(x)$ depend on δ and can be identified as

$$h_i^\delta(x) = \lim_{\varepsilon \to 0} \varepsilon^{-1/2} \mathrm{P}\Big(\text{the process starting at } X^{x,\varepsilon}_{\alpha_1^{x,\delta,\varepsilon}} \text{ reaches } \partial V^\delta \cap U_i \text{ before reaching } \mathcal{L}\Big).$$

From [6, Lemma 4.1 and Lemma 4.3] (with δ now playing the role of ε^{α_1}) it follows that

$$\int_X h_i^\delta(x)\, d\lambda^0(x) = \delta^{-1}(\bar{p}_i + \bar{b}_i(\delta)), \quad i = 1, \ldots, n,$$

where $\bar{p}_i > 0$ and $\bar{b}_i(\delta) \to 0$ as $\delta \to 0$. Now Lemma 3.2 implies that Theorem 3.1 holds with

$$Q = \bar{Q}/(\bar{p}_1 + \ldots + \bar{p}_n)\ , \qquad p_i = \bar{p}_i/(\bar{p}_1 + \ldots + \bar{p}_n)\ .$$

Finally, let us show that $\bar{Q}$ is non-degenerate. Assuming by contradiction that this is not the case, there is a unit vector $e \in \mathbb{R}^2$ such that the function $\bar{g} = \langle e, g \rangle : X \to \mathbb{R}$ has the property that

$$\big(\bar{g}(Z_1^{x,0}) + \ldots + \bar{g}(Z_k^{x,0})\big)/\sqrt{k} \to 0\ , \tag{3.4}$$

in distribution as $k \to \infty$. It follows from $\int_X \bar{g}\, d\lambda_0 = 0$ and from exponential mixing that the sum

$$G(x) = \sum_{k=0}^{\infty} \mathrm{E}\bar{g}(Z_k^{x,0})$$

converges in $L^2(X, \lambda^0)$.

Let z_k denote the process which is $Z_k^{x,0}$ started from the invariant distribution λ_0. It follows from [28, Thm 11] that under our assumption (3.4),

$$0 = \mathrm{E}G^2(z_k) - \mathrm{E}([\mathrm{E}G(Z_1^{x,\varepsilon})]|_{x=z_k})^2. \tag{3.5}$$

By the definition of G, we have the following identity:

$$\bar{g}(z_k) = U_{k+1} + G(z_k) - G(z_{k+1}), \tag{3.6}$$

where $U_{k+1} = U(z_k, z_{k+1}) = G(z_{k+1}) - [\mathrm{E}G(Z_1^{x,\varepsilon})]|_{x=z_k}$. It is straightforward to see that

$$\mathrm{E}U_{k+1}^2 = \mathrm{E}G^2(z_{k+1}) - \mathrm{E}([\mathrm{E}G(Z_1^{x,\varepsilon})]|_{x=z_k})^2 = 0$$

by (3.5). This implies that $U_{k+1} = 0$ almost everywhere with respect to λ_0. Combining this fact with $k = 0$ and (3.6), we get that

$$\bar{g}(x) = G(x) - G(Z_1^{x,0}) \ ,$$

almost surely for λ^0-almost all x. Recall that $x \in X$ can be written as $x = (q, \xi)$, where $q \in \pi(\mathcal{L}_0)$ and $\xi \in \mathbb{Z}^2$. Since $Z_1^{x,0}$ does not depend on ξ, while $\bar{g}(x) = \langle e, \xi\rangle$, we can write $G(x) = \tilde{G}(q) + \langle e, \xi\rangle$ for some function $\tilde{G}$. Thus

$$\tilde{G}(q) = \tilde{G}\big((Z_1^{x,0})^1\big) + \big\langle e, (Z_1^{x,0})^2\big\rangle \ , \tag{3.7}$$

where $(Z_1^{x,0})^1 \in \pi(\mathcal{L}_0)$ and $(Z_1^{x,0})^2 \in \mathbb{Z}^2$. Thus for λ^0-almost all x, we have $\tilde{G}(q) = \tilde{G}\big((Z_1^{x,0})^1\big)$ almost surely on the event $\langle e, (Z_1^{x,0})^2\rangle = 0$. Let $\overline{\lambda}_0$ denote the projection of λ^0 onto $\pi(\mathcal{L}_0)$. An explicit expression for the density of p_0 (found in formula (9) of [6]) implies that $(Z_k^{x,0})^1$, $k \geq 1$, has density with respect to the Lebesgue measure on $\pi(\mathcal{L}_0)$, and the density is bounded from below for sufficiently large k. Therefore

$\bar{\lambda}_0$ is equivalent with the Lebesgue measure and the distribution of $(Z_1^{x,0})^1$ is absolutely continuous with respect to $\overline{\lambda}^0$ for each x. Therefore, by the Markov property, $\tilde{G}(q) = \tilde{G}\big((Z_k^{x,0})^1\big)$ almost surely on the event $\langle e, (Z_1^{x,0})^2\rangle = \ldots = \langle e, (Z_k^{x,0})^2\rangle = 0$, for λ^0-almost all x. For sufficiently large k, the (sub-probability) distribution of $(Z_k^{x,0})^1$ restricted to this event has a positive density with respect to $\overline{\lambda}^0$. (The latter statement is a consequence of the geometry of the flow. Roughly speaking, given two points on the separatrix that belong to the same cell of periodicity, the process $\bar{Z}_k^{x,0}$ can go with positive probability from the first point to an arbitrary neighborhood of the second point without leaving the cell of periodicity.) Therefore, $\tilde{G}$ is λ_0-almost everywhere constant. By (3.7), this implies that $\langle e, (Z_1^{x,0})^2\rangle = 0$ for λ_0-almost all x. Again by the Markov property, $\langle e, (Z_k^{x,0})^2\rangle = 0$ for λ_0-almost all x for each k. Observe, however, that the process $\bar{Z}_k^{x,0}$ starting at an arbitrary point x on the separatrix, has a positive probability of going to any other cell of periodicity if k is sufficiently large. This yields a contradiction, and thus $\bar{Q}$ is non-degenerate. $\square$

Now let us turn to the proof of Lemma 3.2. Let

$$\Omega = \{\omega = (x, x_1, \ldots, x_k; i) : k \geq 0,\ x, x_1, \ldots, x_k \in X, i \in \{1, \ldots, n\}\}$$

be the space of sequences that start at $x \in X$ and end when the sequence enters $C = C_1 \sqcup \ldots \sqcup C_n$, at which point only the index of the set that the sequence enters is taken into account. The Markov chain $Z_k^{x,\varepsilon}$ together with the stopping time τ determine a probability measure μ_ε on Ω, namely,

$$\mu_\varepsilon(x, A_1, \ldots, A_k; i) = \int_{A_1} \ldots \int_{A_k} p_\varepsilon(x, dx_1) p_\varepsilon(x_1, dx_2) \cdots p_\varepsilon(x_{k-1}, dx_k) p_\varepsilon(x_k, C_i),$$

where $A_1, \ldots, A_k \in \mathcal{B}(X)$. We introduce another probability measure on Ω via

$$\nu_\varepsilon(x, A_1, \ldots, A_k, i) =$$
$$= \int_{A_1} \cdots \int_{A_k} e^{-\sqrt{\varepsilon}(J(x)+\ldots+J(x_{k-1}))} \frac{p_\varepsilon(x, dx_1)}{p_\varepsilon(x, X)} \cdots \frac{p_\varepsilon(x_{k-1}, dx_k)}{p_\varepsilon(x_{k-1}, X)} \frac{(1 - e^{-\sqrt{\varepsilon}J(x_k)})h_i(x_k)}{J(x_k)}$$

where $J(x)$ was defined in (3.2). More precisely, we consider a Markov chain $\tilde{Z}_k^{x,\varepsilon}$ on the state space X with transition function $\tilde{p}_\varepsilon(x, dy) = p_\varepsilon(x, dy)/p_\varepsilon(x, X)$. We can adjoin the states $\{1, \ldots, n\}$ to the space X and assume that at each step the process may get killed by entering a terminal state i with probability $(1 - e^{-\sqrt{\varepsilon}J(x_k)})\frac{h_i(x_k)}{J(x_k)}$, $i = 1, \ldots, n$. Let σ be the number of steps after which the process is killed. To clarify our notations, let us stress that $\tilde{Z}_k^{x,\varepsilon}$ is a conservative Markov chain, and the killing is expressed through the presence of the random variable σ defined on the same probability space. Then $\nu_\varepsilon(x, A_1, \ldots, A_k, i)$ is the probability that the chain starting at x visits the sets $A_1, \ldots, A_k$ and then enters the terminal state i. With a slight abuse of notation we can view σ as a random variable on Ω as well.

We will prove in Lemma 3.4, that we can replace the measure μ^ε with ν^ε in a certain sense. First, however, we need to derive a few properties of $\tilde{Z}^{x,\varepsilon}$. Note that it inherits the strong Doeblin property, which holds uniformly in ε, i.e.,

$$\tilde{p}_\varepsilon^m(x, A) \geq a\eta(A) \quad \text{for} \;\; x \in X, \;\; A \in \mathcal{B}(X), \;\; \varepsilon \in [0, \varepsilon_0].$$

This implies the uniform exponential mixing, i.e., there are $\Lambda > 0, c > 0$, such that

$$|\tilde{p}_\varepsilon^k(x, A) - \tilde{\lambda}^\varepsilon(A)| \leq ce^{-\Lambda k} \quad \text{for all } x \in X, \;\; A \in \mathcal{B}(X), \;\; \varepsilon \in [0, \varepsilon_0],$$

where $\tilde{p}_\varepsilon$ is the transition function for the chain and $\tilde{\lambda}^\varepsilon$ is the invariant measure associated with the transition function $\tilde{p}_\varepsilon(x, A)$.

Lemma 3.3. *Let $g \in \mathcal{C}_b(M, \mathcal{R})$ satisfy Assumption (6). For each $\alpha > 0$, we have*

$$\left| \int_X g d\tilde{\lambda}^\varepsilon \right| \leq C\varepsilon^{1/2-\alpha} \tag{3.8}$$

for some constant C and each $\varepsilon \in [0, \varepsilon_0]$.

Proof. By the exponential mixing,

$$\left| \int_X g(y)\tilde{p}^k_\varepsilon(x, dy) - \int_X g(y)\tilde{\lambda}^\varepsilon(dy) \right| + \left| \int_M g(y)p^k_\varepsilon(x, dy) - \int_M g(y)\lambda^\varepsilon(dy) \right| \leq c_1 e^{-\Lambda k}$$

for $x \in X$, $\varepsilon \in (0, \varepsilon_0]$. It is also easy to see by induction that

$$\left| \int_X g(y)\tilde{p}^k_\varepsilon(x, dy) - \int_M g(y)p^k_\varepsilon(x, dy) \right| \leq c_2\sqrt{\varepsilon}k. \tag{3.9}$$

Now we can take $k = [\varepsilon^{-\alpha}]$ in these two inequalities, proving (3.8) since $\int_M g(y)\lambda^\varepsilon(dy) = 0$.

□

The last two inequalities of the above proof with g replaced by an arbitrary bounded continuous function f imply that

$$\int_X f(y)\tilde{\lambda}^\varepsilon(dy) - \int_M f(y)\lambda^\varepsilon(dy) \to 0 \text{ as } \varepsilon \downarrow 0.$$

We also know that $\lambda^\varepsilon(M \setminus X) \to 0$ and $\lambda^\varepsilon \Rightarrow \lambda^0$ as $\varepsilon \downarrow 0$, as immediately follows from the properties of p_ε (the latter statement can be also found in Lemma 2.1 in [6]). Therefore,

$$\int_X f(y)\tilde{\lambda}^\varepsilon(dy) - \int_X f(y)\lambda^0(dy) \to 0 \text{ as } \varepsilon \downarrow 0,$$

that is $\tilde{\lambda}_\varepsilon \Rightarrow \lambda_0$ as $\varepsilon \downarrow 0$.

Lemma 3.4. *For every $\delta > 0$ there is $\varepsilon' > 0$ such that for $\varepsilon \leq \varepsilon'$ there is a set Ω_ε with $\nu_\varepsilon(\Omega_\varepsilon) \geq 1 - \delta$ such that $d\mu_\varepsilon/d\nu_\varepsilon \in (1 - \delta, 1 + \delta)$ on Ω_ε.*

Proof. To choose the set Ω_ε, note that

$$\nu^\varepsilon(\sigma = k) = \tilde{\mathrm{E}}\left[e^{-\sqrt{\varepsilon}(J(\tilde{Z}_1^{x,\varepsilon})+\ldots+J(\tilde{Z}_{k-1}^{x,\varepsilon}))}(1 - e^{-\sqrt{\varepsilon}J(\tilde{Z}_k^{x,\varepsilon})})\right].$$

Using the law of large numbers for the Markov chain $\tilde{Z}^{x,\varepsilon}$, which can be applied uniformly in ε due to the uniform mixing (a consequence of assumption 5.), and the boundedness of J (a consequence of assumption 3.), we conclude that for every $\eta > 0$, there is a k_0 independent of ε such that

$$\tilde{\mathrm{P}}\left(\left|\frac{1}{k}\sum_{j=0}^{k-1} J(\tilde{Z}_j^{x,\varepsilon}) - J_\varepsilon\right| \geq \eta\right) \leq \eta$$

for $k \geq k_0$, where $J_\varepsilon = \int_X J(u)d\lambda^\varepsilon(u)$. Therefore

$$\nu^\varepsilon(\sigma < a/\sqrt{\varepsilon}) \leq \nu^\varepsilon(\sigma < k_0) + \eta + (1 - e^{-\sqrt{\varepsilon}\sup_{u\in X} J(u)})\sum_{k=k_0}^{[a/\sqrt{\varepsilon}]} e^{-\sqrt{\varepsilon}(kJ_\varepsilon - k\eta)}.$$

Since $J_\varepsilon \to J_0 > 0$, and since η was arbitrary, we have $\nu^\varepsilon(\sigma < a/\sqrt{\varepsilon}) < \delta/8$ (for all sufficiently small ε) if a is small enough. Similarly one can show that $\nu^\varepsilon(\sigma > b/\sqrt{\varepsilon}) < \delta/8$ if we choose b to be sufficiently large. We set $\Omega_\varepsilon^1 = \{\sqrt{\varepsilon}\sigma \in [a,b]\}$. Note that $\nu_\varepsilon(\Omega_\varepsilon^1) \geq 1 - \delta/4$. Also note that

$$\nu^\varepsilon(\sigma = k, h_i(x_k) < \eta; i) = \tilde{\mathrm{E}}\left[e^{-\sqrt{\varepsilon}\sum_{j=0}^{k-1} J(\tilde{Z}_j^{x,\varepsilon})}(1 - e^{-\sqrt{\varepsilon}J(\tilde{Z}_k^{x,\varepsilon})})\chi_{\{h_i(\tilde{Z}_k^{x,\varepsilon})<\eta\}}\frac{h_i(\tilde{Z}_k^{x,\varepsilon})}{J(\tilde{Z}_k^{x,\varepsilon})}\right].$$

Using the inequality $x^{-1}(1 - e^{-cx}) < c$ for $x, c > 0$, this is less than or equal to $\eta\sqrt{\varepsilon}$. This means that if $\eta > 0$ is choosen small enough, then

$$\nu^\varepsilon(\sqrt{\varepsilon}\sigma \in [a,b], h_i(x_\sigma) < \eta; i) < \delta/4n \quad \text{for each } i = 1, \ldots, n.$$

We set $\Omega_\varepsilon^2 = \bigcup_{i=1}^n \{\sqrt{\varepsilon}\sigma \in [a,b], h_i(x_\sigma) < \eta; i\}$.

Fix $\gamma > 0$ to be specified later. Let $K_0 \subset X$ be a compact set such that $\lambda^0(X \setminus K_0) < \gamma/3$. This is possible by the σ-compactness of X. Take an open set $U \subseteq X$ such that $K_0 \subseteq U$ and $K = \bar{U}$ is compact, which is possible by local compactness of X. Note that $\lambda^0(X \setminus U) < \gamma/3$. By the weak law of large numbers (which holds uniformly in ε due to the uniform mixing),

$$\sup_{\varepsilon \in (0,\varepsilon_0]} \tilde{\mathrm{P}} \left(\left| \frac{1}{N} \sum_{j=0}^{N-1} \chi_{\{\tilde{Z}_j^{x,\varepsilon} \notin K\}} - \tilde{\lambda}^\varepsilon(X \setminus K) \right| > \gamma/3 \right) < \delta/4$$

for large enough N. Elementary properties of weak convergence imply

$$\tilde{\lambda}^\varepsilon(X \setminus K) \leq \tilde{\lambda}^\varepsilon(X \setminus U) \leq \lambda^0(X \setminus U) + \gamma/3 < 2\gamma/3$$

for small enough ε. This means that the set

$$\Omega_\varepsilon^3 = \left\{ \sqrt{\varepsilon}\sigma \in [a,b], \sum_{j=0}^{\sigma-1} \chi_{\{x_j \notin K\}} \geq \gamma \frac{2b}{a} \sigma \right\} \subseteq$$
$$\subseteq \left\{ \sqrt{\varepsilon}\sigma \in [a,b], \sum_{j=0}^{[b/\sqrt{\varepsilon}]} \chi_{\{x_j \notin K\}} \geq \gamma([b/\sqrt{\varepsilon}] + 1) \right\}$$

has $\nu^\varepsilon(\Omega_\varepsilon^3) < \delta/4$ if ε is sufficiently small.

Similarly, by the ergodic theorem, one can show, by possibly making K larger, that $\Omega_\varepsilon^4 = \{\sqrt{\varepsilon}\sigma \in [a,b], x_\sigma \notin K\}$ has $\nu^\varepsilon(\Omega_\varepsilon^4) < \delta/4$ for sufficiently small ε. Therefore $\Omega_\varepsilon = \Omega_\varepsilon^1 \setminus (\Omega_\varepsilon^2 \cup \Omega_\varepsilon^3 \cup \Omega_\varepsilon^4)$ has $\nu^\varepsilon(\Omega_\varepsilon) > 1 - \delta$.

Observe that

$$\frac{d\mu_\varepsilon}{d\nu_\varepsilon}(x, x_1, \ldots, x_k, i) = \frac{p_\varepsilon(x, X) \cdots p_\varepsilon(x_{k-1}, X)}{e^{-\sqrt{\varepsilon}(J(x)+\ldots+J(x_{k-1}))}} \frac{p_\varepsilon(x_k, C_i)}{1 - e^{-\sqrt{\varepsilon}J(x_k)}} \frac{J(x_k)}{h_i(x_k)} \quad \text{on} \quad \Omega_\varepsilon.$$

By the definition of Ω_ε^1, it suffices to consider $k(\varepsilon) \in [a/\sqrt{\varepsilon}, b/\sqrt{\varepsilon}]$. By the definition of h_i and J, the product of the last two fractions converges to 1 uniformly as $\varepsilon \downarrow 0$ (here we use the definition of Ω_ε^2, Ω_ε^4, and Assumption (3)). Also note that

$$\left| \prod_{j=0}^{k(\varepsilon)-1} p_\varepsilon(x_j, X) - e^{-\sqrt{\varepsilon}\sum_{j=0}^{k(\varepsilon)-1} J(x_j)} \right| =$$

$$= \left| \prod_{j=0}^{k(\varepsilon)-1} \left(1 - \sqrt{\varepsilon} \sum_{i=1}^{n} \varepsilon^{-1/2} p_\varepsilon(x_j, C_i)\right) - \prod_{j=0}^{k(\varepsilon)-1} e^{-\sqrt{\varepsilon} J(x_j)} \right|.$$

Using the fact that $|\prod a_i - \prod b_i| \le \sum |a_i - b_i|$ when $|a_i|, |b_i| \le 1$ and the boundedness part of Assumption (3), this is less than or equal to

$$\sum_{j=0}^{k(\varepsilon)-1} \left| 1 - \sqrt{\varepsilon} \sum_{i=1}^{n} \varepsilon^{-1/2} p_\varepsilon(x_j, C_i) - e^{-\sqrt{\varepsilon} J(x_j)} \right| \le$$

$$\le \sqrt{\varepsilon} \sum_{j=0}^{k(\varepsilon)-1} \sum_{i=1}^{n} \left| h_i(x_j) - \varepsilon^{-1/2} p_\varepsilon(x_j, C_i) \right| + o(1),$$

where we used the Taylor expansion of the exponential. Note that by Assumption (3), we have for small enough ε that

$$\sqrt{\varepsilon} \sum_{j=0}^{k(\varepsilon)-1} \left| h_i(x_j) - \varepsilon^{-1/2} p_\varepsilon(x_j, C_i) \right| \le \frac{2bc}{k(\varepsilon)} \sum_{j=0}^{k(\varepsilon)-1} \chi_{\{x_j \notin K\}} + b\gamma < \gamma b(4bc/a + 1),$$

where the definition of Ω_ε^3 was used in the last inequality. Since γ was arbitrary and

$$\sqrt{\varepsilon} \sum_{j=0}^{k(\varepsilon)-1} J(x_j) \le nc\sqrt{\varepsilon} k(\varepsilon) \le ncb,$$

we have shown that

$$\left| \prod_{j=0}^{k(\varepsilon)-1} p_\varepsilon(x_j, X) \Big/ e^{-\sqrt{\varepsilon} \sum_{j=0}^{k(\varepsilon)-1} J(x_j)} - 1 \right| < \delta$$

for small enough ε provided that $k(\varepsilon) \in [a/\sqrt{\varepsilon}, b/\sqrt{\varepsilon}]$, which implies the desired result. □

Proof of Lemma 3.2. Using Lemma 3.4, we restate Lemma 3.2 in terms of the Markov chain $\tilde{Z}_k^{x,\varepsilon}$. Note first that we can restrict the function g (originally defined on M) to the space X at the expense that the average of g is not zero anymore but satisfies (3.8) instead.

Recall that $\bar{Q}$ is the matrix such that

$$(g(Z_1^{x,0}) + \ldots + g(Z_k^{x,0}))/\sqrt{k} \to N(0, \bar{Q})$$

in distribution as $k \to \infty$. Let $\bar{Q}(\varepsilon)$ be such that

$$\Big(g(\tilde{Z}_1^{x,0}) + \ldots + g(\tilde{Z}_k^{x,0}) - k \int_X g d\tilde{\lambda}_\varepsilon\Big)/\sqrt{k} \to N(0, \bar{Q}(\varepsilon))$$

in distribution as $k \to \infty$. From (3.9) with $k = 1$ and g replaced by an arbitrary bounded continuous function f on X, it follows that $\tilde{p}_\varepsilon(x, dy) \stackrel{\varepsilon \to 0}{\Rightarrow} p_0(x, dy)$ uniformly in $x \in K$ for $K \subseteq X$ compact, since we assumed that the same convergence holds for $p_\varepsilon(x, dy)$. This and the strong Doeblin property for $\tilde{p}_\varepsilon(x, dy)$ easily imply that $\bar{Q}(\varepsilon) \to \bar{Q}$ as $\varepsilon \downarrow 0$ (this was proved in Lemma 2.1 (c) of [6] under an additional assumption that $\int_X g d\tilde{\lambda}^\varepsilon = 0$, which is now replaced by (3.8)).

We still have the functions h_i defined on X, and we assume that the chain terminates by entering the state $i \in \{1, \ldots, n\}$ with probability $(1 - e^{-\sqrt{\varepsilon} J(x)}) h_i(x)/J(x)$. Let σ be the time when the chain terminates. Let the random variable $\tilde{e}$ be equal to i if the process terminates by entering the state i. Since the function g is bounded, omitting one last term in the sum on the left hand side of (3.3) does not affect the limiting distribution. Now we can recast (3.3) as follows:

$$\Big(\varepsilon^{\frac{1}{4}} \big(g(\tilde{Z}_1^{x,\varepsilon}) + \ldots + g(\tilde{Z}_\sigma^{x,\varepsilon})\big), \tilde{e}\Big) \to (F_1, F_2)$$

in distribution. Fix $t \in \mathbb{R}^2$. For $i \in \{1, \ldots, n\}$, we have that

$$\tilde{\mathrm{E}}\left(e^{i\left\langle \varepsilon^{1/4} \sum_{j=1}^{\sigma} g(\tilde{Z}_j^{x,\varepsilon}), t\right\rangle}; \tilde{e} = i\right) =$$
$$= \tilde{\mathrm{E}}\left(e^{i\left\langle \varepsilon^{1/4} \sum_{j=1}^{\sigma} g(\tilde{Z}_j^{x,\varepsilon}), t\right\rangle}; \tilde{e} = i; \left[a/\sqrt{\varepsilon}\right] \le \sigma \le \left[b/\sqrt{\varepsilon}\right]\right) + \delta(a, b, \varepsilon),$$

where

$$|\delta(a,b,\varepsilon)| = \left|\tilde{\mathrm{E}}\left(e^{i\left\langle \varepsilon^{1/4}\sum_{j=1}^{\sigma} g(\tilde{Z}_j^{x,\varepsilon}),t\right\rangle}; \tilde{e}=i; \sigma < \left[a/\sqrt{\varepsilon}\right] \text{ or } \sigma > \left[b/\sqrt{\varepsilon}\right]\right)\right| \leq$$

$$\leq \nu^{\varepsilon}\left(\sigma < \left[a/\sqrt{\varepsilon}\right] \text{ or } \sigma > \left[b/\sqrt{\varepsilon}\right]\right),$$

which was shown in the proof of Lemma 3.4 to converge to zero as $a \to 0, b \to \infty$ uniformly in ε. Let ξ be an exponential random variable with parameter one on some probability space $(\Omega', \mathcal{F}', P')$ independent of the process. By summing over different possible values of σ,

$$\tilde{\mathrm{E}}\left(e^{i\left\langle \varepsilon^{1/4}\sum_{j=1}^{\sigma} g(\tilde{Z}_j^{x,\varepsilon}),t\right\rangle}; \tilde{e}=i\right) = \delta(a,b,\varepsilon) + \tag{3.10}$$
$$+\sum_{k=[a/\sqrt{\varepsilon}]}^{[b/\sqrt{\varepsilon}]} \tilde{\mathrm{E}}\left(\frac{h_i(\tilde{Z}_k^{x,\varepsilon})}{J(\tilde{Z}_k^{x,\varepsilon})} e^{i\left\langle \varepsilon^{1/4}\sum_{j=1}^{k} g(\tilde{Z}_j^{x,\varepsilon}),t\right\rangle} \mathrm{P}'\left(\sqrt{\varepsilon}\sum_{j=0}^{k-1} J(\tilde{Z}_j^{x,\varepsilon}) < \xi \leq \sqrt{\varepsilon}\sum_{j=0}^{k} J(\tilde{Z}_j^{x,\varepsilon})\right)\right).$$

where we used the definition of ν_ε and the fact that

$$\mathrm{P}'(c < \xi \leq d) = e^{-c}(1 - e^{-(d-c)}). \tag{3.11}$$

Note that by the law of large numbers, (3.11), and the uniform exponential mixing property of $\tilde{Z}^{x,\varepsilon}$,

$$\tilde{\mathrm{E}}\sum_{k=[a/\sqrt{\varepsilon}]}^{[b/\sqrt{\varepsilon}]}\left|\mathrm{P}'\left(\sum_{j=0}^{k-1} J(\tilde{Z}_j^{x,\varepsilon}) < \frac{\xi}{\sqrt{\varepsilon}} < \sum_{j=0}^{k} J(\tilde{Z}_j^{x,\varepsilon})\right) - \sqrt{\varepsilon}e^{-k\tilde{J}_\varepsilon\sqrt{\varepsilon}} J(\tilde{Z}_k^{x,\varepsilon})\right| \to 0 \tag{3.12}$$

as $\varepsilon \to 0$ uniformly in $0 < a < b$, where $\tilde{J}_\varepsilon = \int_X J(u)d\tilde{\lambda}^\varepsilon(u) > 0$. Note that the fact that there are $\mathcal{O}(1/\sqrt{\varepsilon})$ terms in the sum is not a problem since the contribution from each term is $\mathcal{O}(\varepsilon)$. Observe that $h_i(x)/J(x) \leq 1$ and therefore the factor proceeding P' on the right hand side of (3.10) is bounded. Therefore, due to (3.12), the main term in (3.10) can be replaced by

$$\sqrt{\varepsilon}\sum_{k=[a/\sqrt{\varepsilon}]}^{[b/\sqrt{\varepsilon}]} \tilde{\mathrm{E}}\left(h_i(\tilde{Z}_k^{x,\varepsilon}) e^{i\left\langle \varepsilon^{1/4}\sum_{j=1}^{k} g(\tilde{Z}_j^{x,\varepsilon}),t\right\rangle}\right) e^{-k\tilde{J}_\varepsilon\sqrt{\varepsilon}}. \tag{3.13}$$

Uniform exponential mixing also tells us that there is a constant C such that for every $0 < k_0 < k$ we have

$$\left| \tilde{\mathrm{E}} \left(h_i(\tilde{Z}_k^{x,\varepsilon}) e^{i\left\langle \varepsilon^{1/4} \sum_{j=1}^{k-k_0} g(\tilde{Z}_j^{x,\varepsilon}),t \right\rangle} \right) - \tilde{\mathrm{E}} \left(h_i(\tilde{Z}_k^{x,\varepsilon}) \right) \tilde{\mathrm{E}} \left(e^{i\left\langle \varepsilon^{1/4} \sum_{j=1}^{k-k_0} g(\tilde{Z}_j^{x,\varepsilon}),t \right\rangle} \right) \right| < Ce^{-\Lambda k_0} \ . \tag{3.14}$$

It is easy to see that fixing $k_0 > 0$, i.e., dropping finitely many terms from the sum in the exponent in (3.13) does not change the limit (it only introduces an overall error term of order $\varepsilon^{1/4}$).

Since we have uniform exponential mixing in ε for the the transition function $\tilde{p}_\varepsilon(x, dy)$ (i.e. for the process $\tilde{Z}_k^{x,\varepsilon}$), and from the fact that $\tilde{\lambda}^\varepsilon \Rightarrow \lambda^0$, it follows that

$$\sup_{k \in [[a/\sqrt{\varepsilon}],[b/\sqrt{\varepsilon}]]} \left| \tilde{\mathrm{E}} \left(h_i(\tilde{Z}_k^{x,\varepsilon}) \right) - \int_X h_i(u) d\lambda^0(u) \right| \to 0 \ , \tag{3.15}$$

as $\varepsilon \downarrow 0$. Choosing $\alpha < 1/4$, it follows from (3.8) that

$$\sup_{k \in [[a/\sqrt{\varepsilon}],[b/\sqrt{\varepsilon}]]} \left| \tilde{\mathrm{E}} \left(e^{i\left\langle \varepsilon^{1/4} \sum_{j=1}^{k-k_0} g(\tilde{Z}_j^{x,\varepsilon}),t \right\rangle} \right) - \tilde{\mathrm{E}} \left(e^{i\left\langle \varepsilon^{1/4} \sum_{j=1}^{k-k_0} \left(g(\tilde{Z}_j^{x,\varepsilon}) - \int_X g d\tilde{\lambda}^\varepsilon \right),t \right\rangle} \right) \right| \to 0 \ ,$$

as $\varepsilon \downarrow 0$. On the other hand, we have the following version of the central limit theorem:

$$\sup_{k \in [[a/\sqrt{\varepsilon}],[b/\sqrt{\varepsilon}]]} \left| \tilde{\mathrm{E}} \left(e^{i\left\langle \varepsilon^{1/4} \sum_{j=1}^{k-k_0} \left(g(\tilde{Z}_j^{x,\varepsilon}) - \int_X g d\tilde{\lambda}^\varepsilon \right),t \right\rangle} \right) - \tilde{\mathrm{E}} e^{i\left\langle \sqrt{k}\varepsilon^{1/4} \cdot N(0,\bar{Q}),t \right\rangle} \right| \to 0 \ ,$$

as $\varepsilon \downarrow 0$, which holds thanks to the uniform strong Doeblin property and the fact that $\bar{Q}(\varepsilon) \to \bar{Q}$ as $\varepsilon \downarrow 0$.

Combining this with (3.10), (3.13), (3.14), and (3.15), and using the fact that $\tilde{J}_\varepsilon \to J_0$, we obtain that

$$\limsup_{\varepsilon \downarrow 0} \left| \tilde{\mathrm{E}} \left(e^{i\left\langle \varepsilon^{1/4} \sum_{j=1}^{k} g(\tilde{Z}_j^{x,\varepsilon}),t \right\rangle} ; \tilde{e} = i \right) - \frac{\int_X h_i d\lambda^0}{\int_X J d\lambda_0} \int_0^\infty \tilde{\mathrm{E}} e^{i\sqrt{s}\langle N(0,\bar{Q}),t \rangle} J_0 e^{-sJ_0} ds \right| \leq ce^{-\Lambda k_0}.$$

Since t and k_0 were arbitrary, this implies the desired result. $\square$

We close this section by stating a technical lemma that gives us control over how far away the process wanders during an upcrossing. Its proof relies on the same arguments as the proof of Lemma 3.2 considering the maximum of $\sum_{j=1}^{k} g(\tilde{Z}_j^{x,\varepsilon})$ until σ and using the invariance principle for Markov chains.

Lemma 3.5. *For each $\eta > 0$, there is $\delta_0 > 0$ such that*

$$\limsup_{\varepsilon \downarrow 0} \sup_{x \in \mathbb{R}^2} \mathrm{P}\Big(\varepsilon^{1/4} \sup_{0 \le t \le \sigma_1^{x,\delta,\varepsilon}} |X_t^{x,\varepsilon} - x| > \eta\Big) < \eta$$

whenever $0 < \delta \le \delta_0$.

3.2 Proof of Theorem 2.3

The first step in the proof of Theorem 2.3 is to show tightness of the family of measures induced by $\varepsilon^{1/4}(X_t^{x,\varepsilon} - x)$, $0 < \varepsilon \le 1$, $x \in \mathbb{R}^2$. We will then show the convergence of one-dimensional distributions. The convergence of finite-dimensional distributions (and therefore the statement of the theorem) will then follow from the Markov property.

Define $D_t^{y,\delta}$ to be the number of downcrossings from level δ to 0 by the trajectory of the process Y_t^y up until time t, where we start counting after the first visit to the vertex. Namely, set $\theta_0^\delta = 0$, $\tau_0^\delta = \inf\{t \ge 0 : Y_t^y = 0\}$, and recursively define

$$\theta_n^\delta = \inf\{t \ge \tau_{n-1}^\delta : Y_t^y = \delta\}, \quad \tau_n^\delta = \inf\{t \ge \theta_n^\delta : Y_t^y = 0\}, \quad n \ge 1,$$

where $|Y_t^y|$ is the Euclidean distance of $Y^{t,y}$ from the interior vertex O. Finally, let $D_t^{y,\delta} = \sup\{n \ge 0 : \tau_n^\delta \le t\}$.

Lemma 3.6. *We have*

$$\lim_{\delta\downarrow 0} \mathrm{E}|\delta D_t^{y,\delta} - L_t^y| = 0$$

for each $t > 0$ and $y \in G$.

The proof of this result is almost identical to [26, Section 2], the only difference being the replacement of the condition $a(i,y) \geq c > 0$ by the local integrability of $(a(i,y))^{-2}$ (and hence of $(a(i,y))^{-1}$) at the interior vertex. As already noted earlier, this is indeed the case here since our graph process arises from the averaging of a Hamiltonian, see [3, Chapter 8], so that $a^{-2}(i,y)$ only diverges logarithmically as $y \to 0$.

For the proof of tightness, we are going to need the following two simple results.

Lemma 3.7. *Let Z_i be a sequence of independent zero mean variables with a common distribution Z, such that all the moments are finite. Then there exists a universal constant C such that*

$$\mathrm{P}\Big(l^{-1/2} \max_{1\leq m\leq l} |Z_1 + \ldots + Z_m| > K\Big) \leq C\frac{\mathrm{E}|Z_i|^{10}}{K^{10}} ,$$

for all $K > 0$.

Proof. By taking the 10th power and using Chebyshev's inequality,

$$\mathrm{P}\left(\max_{1\leq m\leq l} |Z_1 + \ldots + Z_m| \geq K\sqrt{l}\right) \leq \frac{1}{K^{10} l^5}\mathrm{E}\max_{1\leq m\leq l} |Z_1 + \ldots + Z_m|^{10} . \quad (3.16)$$

Since the Z_i are independent centered random variables,

$$\mathrm{E}(Z_1 + ... + Z_l)^{10} = \sum_{i_1,...,i_{10}=1}^{l} \mathrm{E}Z_{i_1}...Z_{i_{10}} =$$

$$+ \sum_{m_1+...+m_5=10,\ m_i\neq 1} C(l, m_1, ..., m_5)\mathrm{E}Z^{m_1} \cdot ... \cdot \mathrm{E}Z^{m_5},$$

where $|C(l, m_1, ..., m_5)| \leq Cl^5$ for some constant $C > 0$. By Holder's inequality, the sum is bounded by $Cl^5\mathrm{E}|Z|^{10}$ with a possibly different constant $C > 0$. The partial sums of the Z_i-s form a martingale so that, by Doob's maximal inequality,

$$\sup_{l\geq 1}\left(l^{-5}\mathrm{E}\max_{1\leq m\leq l}|Z_1+\ldots+Z_m|^{10}\right) \leq \left(\frac{10}{9}\right)^{10}\sup_{l\geq 1}\mathrm{E}\left|\frac{Z_1+\ldots+Z_l}{\sqrt{l}}\right|^{10} \leq C\mathrm{E}|Z|^{10} .$$

The claim now follows at once. □

Lemma 3.8. *We have*

$$\limsup_{t\to 0}\mathrm{E}(L_t^0/t^{1/2})^n < \infty$$

for every $n \in \mathbb{N}$.

Proof. By Lemma 2.3 in [26] with $F(y) = |y - O|$ being the distance of $y \in G$ from the interior vertex, we get that

$$|Y_t^0| = \int_0^t a(i(s), Y_s^0)dW_s + \int_0^t b(i(s), Y_s^0)\, ds + L_t^0 .$$

By the uniqueness of the Skorokhod-reflection, see e.g. [29, Section 3.6.C], we have the representation

$$L_t^0 = \max_{0\leq s\leq t}\left(-\int_0^s a(i(s), Y_s^0)\, dW_s - \int_0^s b(i(s), Y_s^0)\, ds\right) . \tag{3.17}$$

This implies that there is a standard Brownian motion B such that

$$\left(\frac{L_t^0}{t^{1/2}}\right)^n \leq C\left(\max_{0\leq s\leq t}|B_{\frac{1}{t}\int_0^s (a(i(s),Y_s^0))^2 ds}| + t^{-1/2}\int_0^t |b(i(s), Y_s^0)|\, ds\right)^n ,$$

and thus the proof is finished by noting that a and b are bounded on the graph. □

Lemma 3.9. *The family of measures induced by the processes* $\{\varepsilon^{1/4}(X_t^{x,\varepsilon} - x)\}_{0<\varepsilon\leq 1, x\in\mathbb{R}^2}$ *is tight.*

Proof. By the Markov property, it is sufficient to prove that for each $\eta > 0$ there are $r \in (0,1)$ and $\varepsilon_0 > 0$ such that

$$\mathrm{P}\Big(\sup_{0 \le t \le r} |\varepsilon^{1/4}(X_t^{x,\varepsilon} - x)| > \eta \Big) \le r\eta \ , \tag{3.18}$$

for all $\varepsilon \le \varepsilon_0$ and $x \in \mathbb{R}^2$.

Take $Z = \sqrt{\xi} N(0,Q)$ and let Z_1^δ, Z_2^δ, etc. be independent identically distributed. Assume that their distribution coincides with the distribution of $\sqrt{\delta}(1 + a(\delta))Z$, where $a(\delta)$ is the same as in the right hand side of (4.1).

Applying Lemma 3.7 with $K = \eta k^{-1/2}/4$, we see that for a given $\eta > 0$, there are $k_0 \in (0,1)$ and $\delta_1 > 0$ such that

$$\mathrm{P}\Big(\max_{1 \le m \le k/\delta} |Z_1^\delta + \ldots + Z_m^\delta| > \eta/4 \Big) \le k^4 \eta/4 \ , \tag{3.19}$$

whenever $k \in (0, k_0)$ and $\delta \in (0, \delta_1)$. From (3.19) and Lemma 3.1, it follows that there is $\varepsilon_1(k,\delta) > 0$ such that

$$\mathrm{P}\Big(\max_{1 \le m \le k/\delta} \varepsilon^{1/4} |S_1^{x,\delta,\varepsilon} + \ldots + S_m^{x,\delta,\varepsilon}| > \eta/3 \Big) \le k^4 \eta/3 \ , \tag{3.20}$$

provided that $\varepsilon \le \varepsilon_1(k,\delta)$. It is not difficult to see that this estimate and those below are uniform in x. Combining (3.20) and Lemma 3.5, it now follows that there is $\varepsilon_2(k,\delta) > 0$ such that

$$\mathrm{P}\Big(\sup_{0 \le t \le \sigma_{[k/\delta]}^{x,\delta,\varepsilon}} \varepsilon^{1/4} |X_t^{x,\varepsilon} - x| > \eta/2 \Big) \le k^4 \eta/2 \ . \tag{3.21}$$

provided that $\varepsilon \le \varepsilon_2(k,\delta)$.

Note that by Lemma 3.6 for a given $\eta > 0$, we can find $r > 0$ and $\delta_2 = \delta_2(r) > 0$

such that

$$\sup_{y\in G} \mathrm{P}(D_r^{y,\delta} \geq r^{1/4}/\delta) < \sup_{y\in G} \mathrm{P}(L_r^y \geq r^{1/4}) + \eta r/4 \leq r^2 \mathrm{E}(L_r^0/r^{1/2})^8 + \eta r/4 \leq \eta r/3 \tag{3.22}$$

if $\delta \leq \delta_2$, where the second inequality follows from the Chebyshev inequality and the strong Markov property, while the last inequality follows from Lemma 3.8. As a consequence of Lemma 3.1, we see that there is $\varepsilon_3(r,\delta)$ such that

$$\mathrm{P}\Big(\sigma^{x,\delta,\varepsilon}_{[r^{1/4}/\delta]} < r\Big) \leq \mathrm{P}(D_r^{y,\delta} \geq r^{1/4}/\delta) + \eta r/6 \tag{3.23}$$

if $\varepsilon \leq \varepsilon_3(r,\delta)$.

Clearly,

$$\mathrm{P}\Big(\sup_{0\leq t\leq r} |\varepsilon^{1/4}(X_t^{x,\varepsilon} - x)| > \eta\Big) \leq \mathrm{P}\Big(\sigma^{x,\delta,\varepsilon}_{[r^{1/4}/\delta]} < r\Big) + \mathrm{P}\Big(\sup_{0\leq t\leq \sigma^{x,\delta,\varepsilon}_{[r^{1/4}/\delta]}} \varepsilon^{1/4}|X_t^{x,\varepsilon} - x| > \eta\Big)$$

so that, choosing $r > 0$ sufficiently small, combining (3.21) with $k = r^{1/4}$, (3.22), and (3.23) with $\delta < \min(\delta_1, \delta_2)$ and $\varepsilon < \min(\varepsilon_1(k,\delta), \varepsilon_2(k,\delta), \varepsilon_3(r,\delta))$, we obtain (3.18), which implies tightness. □

For the proof of convergence of one-dimensional distributions, we are going to need a lemma that is a straightforward consequence of tightness.

Lemma 3.10. *For $\eta > 0$ and $f \in \mathcal{C}_b(\mathbb{R}^2)$ that is uniformly continuous, we can find an $r > 0$ such that*

$$\sup_{\varepsilon\in(0,1]} |\mathrm{E}f(\varepsilon^{1/4}(X^{x,\varepsilon}_{\tau''} - x)) - \mathrm{E}f(\varepsilon^{1/4}(X^{x,\varepsilon}_{\tau'} - x))| < \eta, \tag{3.24}$$

$$|\mathrm{E}f(\tilde{W}^Q_{\tau''}) - \mathrm{E}f(\tilde{W}^Q_{\tau'})| < \eta \tag{3.25}$$

for each pair of stopping times $\tau' \leq \tau''$ that satisfy $\mathrm{P}(\tau'' > \tau' + r) \leq r$.

Proof. By the tightness result above, for each $\alpha > 0$ we can find $r > 0$ such that

$$\sup_{x\in\mathbb{R}^2} \mathrm{P}\Big(\varepsilon^{1/4} \sup_{0\le t\le r} |X_t^{x,\varepsilon} - x| > \alpha\Big) < \alpha.$$

Using that f is uniformly continuous, we can choose $\alpha(\eta)$ small enough so that we can write

$$\mathrm{E}|f(\varepsilon^{1/4}(X_{\tau''}^{x,\varepsilon} - x)) - f(\varepsilon^{1/4}(X_{\tau'}^{x,\varepsilon} - x)| < \frac{\eta}{3} + \mathrm{P}(\varepsilon^{1/4}|X_{\tau''}^{x,\varepsilon} - X_{\tau'}^{x,\varepsilon}| > \alpha)$$

After conditioning on $X_{\tau'}^{x,\varepsilon}$ and using the strong Markov property, the second term is seen to be bounded from above by

$$\sup_{x\in\mathbb{R}^2} \mathrm{P}\Big(\varepsilon^{1/4} \sup_{0\le t\le r} |X_t^{x,\varepsilon} - x)| > \alpha\Big) + \mathrm{P}(\tau'' - \tau' > r) \le \alpha + r,$$

which finishes the proof of (3.24) once α and r are chosen to be small enough. The proof of (3.25) is similar. □

Let us fix $t > 0$, $f \in \mathcal{C}_b(\mathbb{R}^2)$ uniformly continuous, and $\eta > 0$. To show the convergence of one-dimensional distributions, it suffices to prove that

$$|\mathrm{E}f(\varepsilon^{1/4}(X_t^{x,\varepsilon} - x)) - \mathrm{E}f(\tilde{W}^Q_{L_t^{\Gamma(x)}})| < \eta \tag{3.26}$$

for all sufficiently small ε. As we discussed in the introduction, the main contribution to $X_t^{x,\varepsilon}$ (found in the first term on the left hand side of (3.26)) comes from the excursions between $\mathcal{L}$ and ∂V^δ, i.e., the upcrossings of V^δ. Also, the local time in the second term on the left hand side of (3.26) can be related to the number of excursions (i.e., upcrossings) between the interior vertex and the set $\Gamma(\{x : |H(x)| = \delta\})$ on the graph G that happen before time t. These two observations will lead us to the proof of (3.26).

In order to choose an appropriate value for δ, we need the following lemma (a simple generalization of the central limit theorem).

Lemma 3.11. *Suppose that N_δ are $\mathbb{N}$-valued random variables independent of the family $\{Z_i^\delta\}$ that satisfy $\mathrm{E}N_\delta \le C/\delta$ for some $C > 0$. Let $f \in \mathcal{C}_b(\mathbb{R}^2)$ and let $\tilde{W}_t^Q$ be a Brownian motion with covariance Q, independent of $\{N_\delta\}$. Then*

$$\mathrm{E}f(Z_1^\delta + \ldots + Z_{N_\delta}^\delta) - \mathrm{E}f(\tilde{W}^Q_{\delta N_\delta}) \to 0 \quad as\ \delta \downarrow 0.$$

Let $e^\delta(t)$ be the (random) time that elapses before the time spent by the process $Y_\cdot^y$, aside from the upcrossings, equals t, i.e.,

$$e^\delta(t) = t + \sum_{n=1}^{\infty}(\theta_n^\delta \wedge e^\delta(t) - \tau_{n-1}^\delta \wedge e^\delta(t)).$$

In other words, we stop a 'special' clock every time the process hits the vertex O, and re-start it once the process reaches the level set $\{|y| = \delta\}$. Then $e^\delta(t)$ is the actual time that elapses when the special clock reaches time t. Let $N_\delta = N_t^{y,\delta}$ be the number of upcrossings of the interval $[0, \delta]$ by the process $Y_\cdot^y$ prior to time $e^\delta(t)$.

Similarly, let $e^{\delta,\varepsilon}(t)$ be the time that elapses before the time spent by the process $X_t^{x,\varepsilon}$, aside from the upcrossings, equals t. Let $N_t^{x,\delta,\varepsilon}$ be the number of upcrossings by the process $X_t^{x,\varepsilon}$ prior to time $e^{\delta,\varepsilon}(t)$.

Lemma 3.12. *We have $e^\delta(t) \to t$ and $\delta(N_t^{y,\delta} - D_t^{y,\delta}) \to 0$ in L^1 as $\delta \downarrow 0$ for each $y \in G$.*

Proof. The first statement implies that most of the time is spent on downcrossings rather than upcrossings. Its proof is contained in the proof of Lemma 2.2 in [26].

The second statement follows from the first one together with the Markov property of the process and Lemmas 3.6 and 3.8. □

From Lemmas 3.12 and 3.6 it follows that the conditions of Lemma 3.11 are satisfied with our choice of N_δ. We can therefore choose $\delta_0 > 0$ such that

$$\sup_{y \in G} \left| \mathrm{E} f(Z_1^\delta + \ldots + Z^\delta_{N_t^{\Gamma(x),\delta}}) - \mathrm{E} f\left(\tilde{W}^Q_{\delta N_t^{\Gamma(x),\delta}}\right) \right| \leq \eta/10 \tag{3.27}$$

whenever $\delta \leq \delta_0$.

Choose r is such that (3.24) and (3.25) in Lemma 3.10 hold with $\eta/10$ instead of η. Also, use Lemma 3.6 and Lemma 3.12 to choose $\delta < \delta_0$ sufficiently small so that

$$\left| \mathrm{E} f\left(\tilde{W}^Q_{\delta D_t^{\Gamma(x),\delta}}\right) - \mathrm{E} f\left(\tilde{W}^Q_{L_t^{\Gamma(x)}}\right) \right| < \eta/10 \tag{3.28}$$

and

$$\mathrm{P}(\delta N_t^{\Gamma(x),\delta} > \delta D_t^{\Gamma(x),\delta} + r) \leq r, \quad \mathrm{P}(e^\delta(t) > t + r) \leq r/2 \ .$$

From the weak convergence of the processes, the latter implies that there is $\varepsilon_0 > 0$ such that

$$\mathrm{P}(e^{\delta,\varepsilon}(t) > t + r) \leq r$$

for $\varepsilon < \varepsilon_0$. By Lemma 3.10, these inequalities imply that

$$|\mathrm{E} f(\varepsilon^{1/4}(X^{x,\varepsilon}_{e^{\delta,\varepsilon}(t)} - x)) - \mathrm{E} f(\varepsilon^{1/4}(X^{x,\varepsilon}_t - x))| < \eta/10 \ , \tag{3.29}$$

and

$$\left| \mathrm{E} f\left(\tilde{W}^Q_{\delta N_t^{\Gamma(x),\delta}}\right) - \mathrm{E} f\left(\tilde{W}^Q_{\delta D_t^{\Gamma(x),\delta}}\right) \right| < \eta/10 \ . \tag{3.30}$$

In what follows δ is fixed at this value.

Choose N large enough so that

$$|\mathrm{E}f(Z_1^\delta + \ldots + Z^\delta_{N_t^{\Gamma(x),\delta}}) - \mathrm{E}f(Z_1^\delta + \ldots + Z^\delta_{N_t^{\Gamma(x),\delta}\wedge N})| < \eta/10 \tag{3.31}$$

and by possibly increasing N, let $\varepsilon_1 > 0$ be such that

$$|\mathrm{E}f(\varepsilon^{1/4}(X^{x,\varepsilon}_{e^{\delta,\varepsilon}(t)} - x)) - \mathrm{E}f(\varepsilon^{1/4}(X^{x,\varepsilon}_{e^{\delta,\varepsilon}(t)\wedge\sigma_N^{x,\delta,\varepsilon}} - x))| < \eta/10 \tag{3.32}$$

for all $\varepsilon \le \varepsilon_1$. The latter can be done by noting that by Lemma 3.1, for every α, one can select an N such that

$$\mathrm{P}(\sigma_N^{x,\delta,\varepsilon} \le e^{\delta,\varepsilon}(t)) < \alpha \tag{3.33}$$

for every small enough ε. Indeed,

$$\mathrm{P}(\sigma_N^{x,\delta,\varepsilon} \le e^{\delta,\varepsilon}(t)) = \mathrm{P}(T_1^{x,\delta,\varepsilon} + \ldots + T_N^{x,\delta,\varepsilon} \le t).$$

For fixed N and δ, the random variable $T_1^{x,\delta,\varepsilon} + \ldots + T_N^{x,\delta,\varepsilon}$ converges in distribution to some random variable $\tilde{\tau}_N^\delta$ as $\varepsilon \downarrow 0$. Choose N large enough so that

$$\mathrm{P}(\tilde{\tau}_N^\delta \le t) < \alpha/2,$$

which implies (3.33). Both N and δ are fixed now.

By Lemma 3.1, there is $\varepsilon_2(\delta) > 0$ such that

$$|\mathrm{E}f(\varepsilon^{1/4}(S_1^{x,\delta,\varepsilon} + \ldots + S^{x,\delta,\varepsilon}_{N_t^{\Gamma(x),\delta,\varepsilon}\wedge N})) - \mathrm{E}f(Z_1^\delta + \ldots + Z^\delta_{N_t^{\Gamma(x),\delta}\wedge N})| < \eta/10 \tag{3.34}$$

if $\varepsilon \le \varepsilon_2$. It it here where we used the fact that the displacements during upcrossings become independent, in the limit of $\varepsilon \downarrow 0$, from the times spent on downcrossings. We also have that there is an $\varepsilon_3 > 0$ such that

$$|\mathrm{E}f(\varepsilon^{1/4}(S_1^{x,\delta,\varepsilon} + \ldots + S^{x,\delta,\varepsilon}_{N_t^{\Gamma(x),\delta,\varepsilon}\wedge N})) - \mathrm{E}f(\varepsilon^{1/4}(X^{x,\varepsilon}_{e^{\delta,\varepsilon}(t)\wedge\sigma_N^{x,\delta,\varepsilon}} - x))| < \eta/10 \tag{3.35}$$

for all $\varepsilon < \varepsilon_3$.

Collecting (3.29), (3.32), (3.35), (3.34), (3.31), (3.27), (3.30), and (3.28), we obtain (3.26) for $\varepsilon \le \min\{\varepsilon_0, \varepsilon_1, \varepsilon_2, \varepsilon_3\}$, which completes the proof of Theorem 2.3.

□

Remark 3.1. *It is not difficult to show (and it indeed follows from the proof) that convergence in Theorem 2.3 is uniform in* $x \in \mathbb{R}^2$.

Chapter 4: Short timescales

4.1 Proof of Theorem 2.4

The idea of the proof is to establish a setup similar to the one we had for the intermediate time case so that the arguments in Section 3.2 can be repeated. To do this, let us define $V^{\delta,\varepsilon} = \{x \in \mathbb{R}^2 : |H(x)| \leq \delta\varepsilon^{\alpha/2}\}$ and consider the (H, θ) coordinates in $V^{\delta,\varepsilon} \cap \overline{U}_k$. Once again, $\theta \in [0, \int_{\partial U_k} |\nabla H| dl]$ and the endpoints of the interval are identified. Using these new coordinates, we can define the relevant quantities just as in Chapter 3. For the sake of completeness, we repeat this definitions in the present context.

$$B(A_i, U_k) = \{x \in V^{\delta,\varepsilon} \cap \overline{U}_k \,:\, \theta(x) = \theta(A_i)\}\,, \qquad B(A_i) = \bigcup_{k: A_i \in \partial U_k} B(A_i, U_k).$$

Let $\pi : \mathbb{R}^2 \to \mathcal{T}$ be the quotient map from the plane to the torus and, for simplicity, let us denote $\pi(V^{\delta,\varepsilon})$ by $V^{\delta,\varepsilon}$ again. Let $Z_t^{x,\varepsilon} = X^{x,\varepsilon}\left(\frac{\alpha\varepsilon^{\alpha}|\log\varepsilon| t}{2}\right)$ and introduce the stopping times $\alpha_0^{x,\delta,\varepsilon} = 0$, $\beta_0^{x,\delta,\varepsilon} = \inf\{t \geq 0 \,:\, Z_t^{x,\varepsilon} \in \mathcal{L}\}$ and recursively define

$$\alpha_n^{x,\delta,\varepsilon} = \inf\left\{t \geq \beta_{n-1}^{x,\delta,\varepsilon} \,:\, \pi(Z^{x,\varepsilon}(t)) \in \bigcup_{k \neq i} B(A_k) \cup \partial V^{\delta,\varepsilon}\right\} \qquad \text{if } \pi(Z^{x,\varepsilon}(\beta_{n-1}^{x,\delta,\varepsilon})) \in \gamma(A_i, A_j)$$

and $\beta_n^{x,\delta,\varepsilon} = \inf\{t \geq \alpha_n^{x,\delta,\varepsilon} : X^{x,\varepsilon}(t) \in \mathcal{L}\}$. In other words, $\alpha_n^{x,\delta,\varepsilon}$ is the first time after $\beta_{n-1}^{x,\delta,\varepsilon}$ that the process either hits $\partial V^{\delta,\varepsilon}$, or goes past a saddle point different from the one behind $Z^{x,\varepsilon}(\beta_{n-1}^{x,\delta,\varepsilon})$.

We introduce another pair of sequences of stopping times corresponding to successive visits to $\mathcal{L}$ and ∂V^ε. Namely, let $\mu_0^{x,\delta,\varepsilon} = 0$, $\sigma_0^{x,\delta,\varepsilon} = \beta_0^{x,\delta,\varepsilon}$, and recursively define

$$\mu_n^{x,\delta,\varepsilon} = \inf\{t \geq \sigma_{n-1}^{x,\delta,\varepsilon} \,:\, Z^{x,\varepsilon}(t) \in \partial V^{\delta,\varepsilon}\}, \quad \sigma_n^{x,\delta,\varepsilon} = \inf\{t \geq \mu_n^{x,\delta,\varepsilon} \,:\, Z^{x,\varepsilon}(t) \in \mathcal{L}\}.$$

Let

$$S_n^{x,\delta,\varepsilon} = Z^{x,\varepsilon}(\sigma_n^{x,\delta,\varepsilon}) - Z^{x,\varepsilon}(\sigma_{n-1}^{x,\delta,\varepsilon}),\ n \geq 1, \quad T_n^{x,\delta,\varepsilon} = \sigma_n^{x,\delta,\varepsilon} - \mu_n^{x,\delta,\varepsilon},\ n \geq 0,$$

be the displacement between successive visits to $\mathcal{L}$ and the time spent on the n-th downcrossing of $V^{\delta,\varepsilon}$, respectively.

Let $\eta^{x,\delta,\varepsilon}$ be the random vector with values in $\{1, \ldots, n\}$ defined by

$$\eta^{x,\delta,\varepsilon} = i \ \text{ if } \ Z^{x,\varepsilon}(\mu_1^{x,\delta,\varepsilon}) \in U_i, \quad i = 1, \ldots, n,$$

i.e., $\eta^{x,\delta,\varepsilon} = i$ if the process ends up in U_i after the first upcrossing of $V^{\delta,\varepsilon}$.

Our first task is to describe how far $Z^{x,\varepsilon}(t)$ can travel from $\mathcal{L}$ before hitting $V^{\delta,\varepsilon}$, and we do that by adapting Theorem 3.1 to the current situation.

Theorem 4.1. *There is a 2×2 non-degenerate matrix Q and a vector $(p_1, \ldots, p_n)$ such that*

$$(\varepsilon^{\frac{1-\alpha}{4}} S_1^{x,\delta,\varepsilon}, \eta^{x,\delta,\varepsilon}) \to (\sqrt{\delta\xi} N(0,Q), \eta) \quad \text{in distribution as } \varepsilon \downarrow 0 \qquad (4.1)$$

uniformly in $x \in \mathcal{L}$. Here ξ is an exponential random variable with parameter one, N is a two dimensional normal with covariance matrix Q, independent of ξ, and η is a random vector with values in $\{1, \ldots, n\}$ independent of ξ and N such that $\mathrm{P}(\eta = i) = p_i$.

This is a simple modification of Theorem 3.1, the main difference being that after reaching $B(A_i)$, the probability of getting absorbed in the cell interior before going back to $\mathcal{L}$ is $\sim \varepsilon^{(1-\alpha)/2}$.

It is not hard to deduce the following consequence of Theorem 2.2.

Lemma 4.1. *Denote by $T^{x,\delta,\varepsilon}$ the time it takes for $Z^{x,\varepsilon}(t)$ to reach the separatrix. Also let $\bar{T}^y$ be the analogous quantity for $\bar{Y}^y(t)$, i.e. the time it takes for $Y^y(t)$ to reach the interior vertex O. Then*

$$T^{\mu^\varepsilon,\delta,\varepsilon} \overset{\varepsilon\to 0}{\Rightarrow} \bar{T}^\nu$$

where μ^ε and ν are measures satisfying the requirements of Theorem 2.2.

Theorems 4.1 and Lemma 4.1 combined with the strong Markov property imply

Corollary 4.1. *For fixed m, the random vectors*

$$(T_0^{\mu^\varepsilon,\delta,\varepsilon}, \varepsilon^{\frac{1-\alpha}{4}} S_1^{\mu^\varepsilon,\delta,\varepsilon}, T_1^{\mu^\varepsilon,\delta,\varepsilon}, \varepsilon^{\frac{1-\alpha}{4}} S_2^{\mu^\varepsilon,\delta,\varepsilon}, \ldots, T_{m-1}^{\mu^\varepsilon,\delta,\varepsilon}, \varepsilon^{\frac{1-\alpha}{4}} S_m^{\mu^\varepsilon,\delta,\varepsilon})$$

converge, as $\varepsilon \downarrow 0$, to a random vector with independent components. The limiting distribution for each of the components $\varepsilon^{(1-\alpha)/4} S_1^{\mu^\varepsilon,\delta,\varepsilon}, \ldots, \varepsilon^{(1-\alpha)/4} S_m^{\mu^\varepsilon,\delta,\varepsilon}$ is given by Theorem 4.1, i.e., it is equal to the distribution of $\sqrt{\xi} N(0,Q)$. The limiting distribution of $T_0^{\mu^\varepsilon,\delta,\varepsilon}$ is the distribution of $\bar{T}^\nu$, while the limiting distribution for each of the components $T_1^{x,\delta,\varepsilon}, \ldots, T_{m-1}^{x,\varepsilon}$ is equal to the distribution of $\bar{T}^\zeta$, where ζ is a random initial point for the process on the graph, chosen to be at distance δ from the vertex O, in such a way that ζ belongs to the i-th edge with probability p_i.

Using Corollary 4.1, the proof of Theorem 2.4 is the same as that of the intermediate time case in Section 3.2.

4.2 The averaging principle on the short time scales

4.2.1 Convergence to the limit

Let $\bar{Y}_t^{\mu^\varepsilon,\varepsilon} = \Gamma^\varepsilon(Z_t^{\mu^\varepsilon,\varepsilon})$. In order to prove Theorem 2.2, we will show that the unique weak limit point of $\bar{Y}^{\mu^\varepsilon,\varepsilon}$ is the solution of the martingale problem of the operator $\mathcal{A}$. The proof is very similar to the verification of the original averaging principle in [3] and we follow the approach presented in [24].

Let $\bar{Y}^{x,\varepsilon} = \bar{Y}^{\delta_x,\varepsilon}$.

Theorem 4.2. *Let $\mathcal{A}$ be the generator of a diffusion on $\bar{G}$ (as defined in Chapter 2) and $\Psi \subseteq \mathcal{D}(\mathcal{A})$ be a set that separates measures on $\bar{G}$. Also, let D be a subset of $\mathcal{D}(A)$ large enough such that $\Psi \subseteq (\lambda - A)(D)$ for every $\lambda > 0$. Assume that for any $f \in D$, $T > 0$, $K \subseteq \bar{G}$ compact and any $\eta > 0$, we have*

$$\sup_{x \in (\bar{\Gamma}^\varepsilon)^{-1}(K)} \left| \mathrm{E} f(\bar{Y}_T^{x,\varepsilon}) - f(\bar{\Gamma}^\varepsilon(x)) - \int_0^T \mathcal{A} f(\bar{Y}_t^{x,\varepsilon}) dt \right| \to 0 \tag{4.2}$$

as $\varepsilon \to 0$. Then, if the family $\{\bar{Y}^{\mu^\varepsilon,\varepsilon}\}_{\varepsilon \in (0,\varepsilon_0]}$ is tight, $\bar{Y}^{\mu^\varepsilon,\varepsilon}(t)$ converges weakly in $\mathcal{C}([0,\infty), \bar{G})$ to the unique solution of the martingale problem associated to the operator $\mathcal{A}$ and the initial measure ν.

As in the case of the corresponding result of Freidlin and Wentzell [3], this result can be proved using tightness, the strong Markov property and by taking

$\Psi = \cap_{i=1}^n \mathcal{C}_b^2(I_i) \cap \mathcal{C}_b(G)$ and $D = \cap_{i=1}^n \mathcal{C}_b^4(I_i) \cap \mathcal{D}(A)$. As the verification is completely analogous, we omit the proof.

Tightness will be proved in Section 4.2.4, the rest of this section is devoted to proving (4.2). It is clear that it suffices to show this with $K = \{d(O,y) \leq H_0\}$ for some H_0.

We first state several lemmas needed in the proof of (4.2) the first of which tells us that the process does not wander too far into the cell interior over any finite time interval. It will be proved in Section 4.2.4.

Lemma 4.2. *For every $\eta > 0, T > 0$, and $H_0 > 0$, there is a constant $H_1 > 0$ and $\varepsilon_0 > 0$ such that we have*

$$\sup_{|H(x)| \leq \varepsilon^{\alpha/2} H_0} \mathrm{P}\left(\sup_{t \in [0,T]} |H(Z_t^{x,\varepsilon})| \geq \varepsilon^{\alpha/2} H_1 \right) < \eta \tag{4.3}$$

whenever $\varepsilon < \varepsilon_0$.

Let $H_1 > 0$ and $\lambda_{H_1}^{x,\varepsilon}$ be the first time $Z_t^{x,\varepsilon}$ reaches the set

$$\gamma(\varepsilon^{\alpha/2} H_1) = \{H(z) = \varepsilon^{\alpha/2} H_1\}.$$

Then (4.3) can be reformulated as

$$\sup_{|H(x)| \leq \varepsilon^{\alpha/2} H_0} \mathrm{P}\left(\lambda_{H_1}^{x,\varepsilon} \leq T \right) < \eta \tag{4.4}$$

Let $\beta \in (\alpha/2, \alpha \wedge 1/2)$, and let $\bar{\gamma}_k = \gamma_k(\varepsilon^\beta) = \gamma(\varepsilon^\beta) \cap U_k$. Define $\bar{\gamma} = \cup_k \bar{\gamma}_k = \gamma(\varepsilon^\beta)$. This is a level set that is farther from $\mathcal{L}$ than the typical fluctuation of $H(Z_t^{x,\varepsilon})$ in finite time, but close enough so that the process will make infinitely many travels between $\bar{\gamma}$ and $\mathcal{L}$. Let $\tau^{x,\varepsilon}$ be the first time when $Z_t^{x,\varepsilon}$ reaches $\bar{\gamma}$ and $\kappa^{x,\varepsilon}$ when it first reaches $\mathcal{L}$. The following lemma gives us estimates on the expectation of these stopping times.

Lemma 4.3. *For sufficiently small* ε, *we have*

$$\sup_{x\in\mathcal{L}} \mathrm{E}\tau^{x,\varepsilon} \le \varepsilon^{2\beta-\alpha}. \tag{4.5}$$

Moreover, for every $H_0 > 0$, *there is a* $K \in (0,\infty)$ *such that*

$$\sup_{|H(x)|\le\varepsilon^{\alpha/2}H_0} \mathrm{E}\lambda_{H_0}^{x,\varepsilon} < K \tag{4.6}$$

for sufficiently small ε.

Here, (4.6) follows from Lemma 4.4 [6] by the appropriate time change. Similarly, (4.5) follows Lemma 4.2 in [6]. It also follows from Lemma 4.4 in [6] that

$$\sup_{x\in\bar{\gamma}} \mathrm{E}\kappa^{x,\varepsilon} = k_1\frac{\varepsilon^{\beta-\alpha}}{\log\varepsilon}(1+o(1)). \tag{4.7}$$

However, this formula blows up and therefore is of limited use in this case.

The following estimate is singled out as a separate lemma as it does not immediately follow from the previous literature and will be proved.

Lemma 4.4. *For any* $H_0 > 0$ *and small enough* ε,

$$\sup_{x\in\bar{\gamma}} \mathrm{E}\kappa^{x,\varepsilon} \wedge \lambda_{H_0}^{x,\varepsilon} = \mathcal{O}(\varepsilon^{\beta-\alpha/2})$$

Using κ and τ, we can define the following sequence of stopping times $\kappa_0^{x,\varepsilon} = 0$, $\tau_1^{x,\varepsilon} = \tau^{x,\varepsilon}$, and inductively define

$$\tau_n^{x,\varepsilon} = \inf\{t > \kappa_{n-1}^{x,\varepsilon} : Z_t^{x,\varepsilon} \in \bar{\gamma}\}, \qquad \kappa_n^{x,\varepsilon} = \inf\{t > \tau_n : Z_t^{x,\varepsilon} \in \mathcal{L}\}.$$

From here therein, we adopt the convention of only writing the indices referring to x, ε only once whenever these stopping times are involved in a more complicated formula. Using these times, we can define the discrete time Markov chains $\xi_n^1 = Z^{x,\varepsilon}(\tau_n)$ and $\xi_n^2 = Z^{x,\varepsilon}(\kappa_n)$ with state space γ and $\mathcal{L}$ respectively. We denote the transition operators by $P_1^\varepsilon(x, dy)$ and $P_2^\varepsilon(x, dy)$ respectively.

Lemma 4.5. *There exists a $c \in (0, 1)$, $\varepsilon_0 > 0$, $n_0 > 0$ and probability measures $\nu_{\mathcal{L}}^\varepsilon$ and ν_γ^ε on $\mathcal{L}$ and $\bar{\gamma}$ respectively such that*

$$\sup_{x\in\bar{\gamma}} d_{TV}((P_1^\varepsilon)^n(x, dy), \nu_{\bar{\gamma}}^\varepsilon(dy)) \leq c^n, \qquad \sup_{x\in\mathcal{L}} d_{TV}((P_2^\varepsilon)^n(x, dy), \nu_{\mathcal{L}}^\varepsilon(dy)) \leq c^n \quad (4.8)$$

where d_{TV} is the total variation distance.

The proof of this result is completely analogous to the one presented in Section 7 of [24]. It is also true that there is a constant c such that

$$\lim_{\varepsilon\downarrow 0} \nu_{\bar{\gamma}}^\varepsilon(\gamma_k) = c\alpha_i \quad (4.9)$$

We will estimate contributions to (4.2) of three different types: until the first hitting of $\bar{\gamma}$, on intervals $[\tau_i^{x,\varepsilon}, \kappa_i^{x,\varepsilon}]$ (downcrossings), and on intervals $[\kappa_i^{x,\varepsilon}, \tau_{i+1}^{x,\varepsilon}]$ (upcrossings). This is acheived by the following lemma, which will be proved in Section 4.2.3.

Lemma 4.6. *For any $f \in D$, we have that the following estimates hold as $\varepsilon \downarrow 0$ for $H_0 > 0$.*

1.

$$\sup_{\varepsilon^{\alpha/2}|H(x)|\leq H_0} \left| E\left[f(\bar{Y}^{x,\varepsilon}_{\tau\wedge\lambda^{x,\varepsilon}_{H_0}}) - f(\Gamma^\varepsilon(x)) - \int_0^{\tau^{x,\varepsilon}\wedge\lambda^{x,\varepsilon}_{H_0}} \mathcal{A}f(\bar{Y}^{x,\varepsilon}_s)ds \right] \right| \to 0 \quad (4.10)$$

2.

$$\sup_{x\in\bar\gamma}\left|E\left[f(\bar Y^{x,\varepsilon}_{\kappa\wedge\lambda_{H_0}})-f(\Gamma^\varepsilon(x))-\int_0^{\kappa^{x,\varepsilon}\wedge\lambda^{x,\varepsilon}_{H_0}}\mathcal{A}f(\bar Y^{x,\varepsilon}_s)ds\right]\right|=o\left(\varepsilon^{\beta-\alpha/2}\right)\quad(4.11)$$

3.

$$\left|E\left[f(\bar Y^{\nu_{\mathcal{L}},\varepsilon}_{\tau})-f(\bar Y^{\nu_{\mathcal{L}},\varepsilon}_0)-\int_0^{\tau^{\nu_{\mathcal{L}},\varepsilon}}\mathcal{A}f(\bar Y^{\nu_{\mathcal{L}},\varepsilon}_s)ds\right]\right|=o\left(\varepsilon^{\beta-\alpha/2}\right)\quad(4.12)$$

Finally, we need to have control over the number of these transitions between $\mathcal{L}$ and $\bar\gamma$ which we can acheive by the following Lemma also proved in Section 4.2.3.

Lemma 4.7. *There is a constant $r>0$, such that for all sufficiently small ε, we have*

$$\sup_{x\in\bar\gamma}\mathrm{E}e^{-\kappa^{x,\varepsilon}}\le 1-r\varepsilon^{\beta-\alpha/2}$$

Moreover, by the Markov property, we have

$$\sup_{\varepsilon^{\alpha/2}|H(x)|\le H_0}\mathrm{E}e^{-\kappa^{x,\varepsilon}_n}\le\left(1-r\varepsilon^{\beta-\alpha/2}\right)^n\quad(4.13)$$

Proof of (4.2). Let $f\in D$, $T>0$, and $\eta>0$ fixed. We will argue that whenever ε is sufficiently small, the supremum on the left hand side of (4.2) is less than η.

First we want to exclude the possibility that the process can wander too far into the cell interior. More precisely, the difference of (4.2) and

$$\sup_{|H(x)|\le\varepsilon^{\alpha/2}H_0}\left|\mathrm{E}f(\bar Y^{x,\varepsilon}_{T\wedge\lambda_{H_1}})-f(\Gamma^\varepsilon(x))-\int_0^{T\wedge\lambda^{x,\varepsilon}_{H_1}}\mathcal{A}f(\bar Y^{x,\varepsilon}_t)dt\right|\quad(4.14)$$

is less than

$$(||f||+T||\mathcal{A}f||)\sup_{|H(x)|\le\varepsilon^{\alpha/2}H_0}\mathrm{P}(\lambda^{x,\varepsilon}_{H_1}\le T)$$

which can be made $\eta/10$ by (4.4). Therefore, it remains to prove (4.14).

Let $\tilde{\tau}^{x,\varepsilon}$ be the first stopping time $\tau_n^{x,\varepsilon}$ that is larger than $T \wedge \lambda_{H_1}^{x,\varepsilon}$, meaning

$$\tilde{\tau}_{H_1}^{x,\varepsilon} = (\min_{n:\tau_n \geq T} \tau_n^{x,\varepsilon}) \wedge \lambda_{H_1}^{x,\varepsilon}$$

If we replace $T \wedge \lambda_{H_1}^{x,\varepsilon}$ in (4.14) by $\tilde{\tau}_{H_1}^{x,\varepsilon}$, the error we make can be estimated using the strong Markov property of $Z_t^{x,\varepsilon}$ at time $T \wedge \lambda_{H_1}^{x,\varepsilon}$ as

$$\left| \mathrm{E}\left[f(\bar{Y}_{\tilde{\tau}_{H_1}}^{x,\varepsilon}) - f(\bar{Y}_{T\wedge\lambda_{H_1}}^{x,\varepsilon})) - \int_{T\wedge\lambda_{H_1}^{x,\varepsilon}}^{\tilde{\tau}^{x,\varepsilon}} \mathcal{A}f(\bar{Y}_t^{x,\varepsilon})dt \right] \right| \leq$$
$$\leq \sup_{|H(x)| \leq \varepsilon^{\alpha/2} H_1} \left| \mathrm{E}\left[f(\bar{Y}_{\tau\wedge\lambda_{H_1}^{x,\varepsilon}}^{x,\varepsilon}) - f(\Gamma^\varepsilon(x)) - \int_0^{\tau^{x,\varepsilon}\wedge\lambda_{H_1}^{x,\varepsilon}} \mathcal{A}f(\bar{Y}_t^{x,\varepsilon})dt \right] \right|.$$

This is less than $\eta/10$ for every $|H(x)| \leq \varepsilon^{\alpha/2} H_0$ by (4.10). Therefore, it remains to prove

$$\sup_{|H(x)| \leq \varepsilon^{\alpha/2} H_0} \left| \mathrm{E} f(\bar{Y}_{\tilde{\tau}_{H_1}}^{x,\varepsilon}) - f(\Gamma^\varepsilon(x)) - \int_0^{\tilde{\tau}_{H_1}^{x,\varepsilon}} \mathcal{A}f(\bar{Y}_t^{x,\varepsilon})dt \right| \leq \frac{4\eta}{5}.$$

We will acheive this by breaking up the interval $[0, \tilde{\tau}_{H_1}]$ into what happens before the process first reaches $\bar{\gamma}$, and the successive series of downcrossings and upcrossings afterwards. More precisely,

$$\mathrm{E}\left[f(\bar{Y}_{\tilde{\tau}_{H_1}}^{x,\varepsilon}) - f(\Gamma^\varepsilon(x)) - \int_0^{\tilde{\tau}_{H_1}^{x,\varepsilon}} \mathcal{A}f(\bar{Y}_t^{x,\varepsilon})dt \right] =$$
$$= \mathrm{E}\left[f(\bar{Y}_{\tau\wedge\lambda_{H_1}}^{x,\varepsilon}) - f(\Gamma^\varepsilon(x)) - \int_0^{\tau^{x,\varepsilon}\wedge\lambda_{H_1}^{x,\varepsilon}} \mathcal{A}f(\bar{Y}_t^{x,\varepsilon})dt \right] +$$
$$+ \sum_{n=1}^{\infty} \mathrm{E}\chi_{\{\tau_n^{x,\varepsilon} < \tilde{\tau}_{H_1}^{x,\varepsilon}\}} \mathrm{E}\left[f(\bar{Y}_{\kappa\wedge\lambda_{H_1}}^{y,\varepsilon}) - f(\Gamma^\varepsilon(y)) - \int_0^{\kappa^{y,\varepsilon}\wedge\lambda_{H_1}^{y,\varepsilon}} \mathcal{A}f(\bar{Y}_t^{y,\varepsilon})dt \right]_{y=Z_{\tau_n}^{x,\varepsilon}} +$$
$$+ \sum_{n=1}^{\infty} \mathrm{E}\chi_{\{\kappa_n^{x,\varepsilon} < \tilde{\tau}_{H_1}^{x,\varepsilon}\}} \mathrm{E}\left[f(\bar{Y}_{\tau}^{y,\varepsilon}) - f(\Gamma^\varepsilon(y)) - \int_0^{\tau^{y,\varepsilon}} \mathcal{A}f(\bar{Y}_t^{y,\varepsilon})dt \right]_{y=Z_{\kappa_n}^{x,\varepsilon}},$$

provided that the sums converge absolutely (which follows from the arguments below). The supremum of the first term over the region where $|H(x)| \leq \varepsilon^{\alpha/2} H_0$ is less than $\eta/5$ by (4.10) if ε is sufficiently small. To finish the proof, we have to estimate the two infinite sums.

Note that by (4.13),

$$\mathrm{E}\chi_{\{\tau_n^{x,\varepsilon}<\tilde{\tau}_{H_1}^{x,\varepsilon}\}} \le \mathrm{E}\chi_{\{\kappa_{n-1}^{x,\varepsilon}<T\}} \le \mathrm{E}\chi_{\{e^{-\kappa_{n-1}^{x,\varepsilon}}>e^{-T}\}} \le e^T\left(1-r\varepsilon^{\beta-\alpha/2}\right)^{n-1},$$

whenever ε is sufficiently small and $|H(x)|\le \varepsilon^{\alpha/2}H_0$. Taking the sum in n leads to

$$\sup_{|H(x)|\le\varepsilon^{\alpha/2}H_0}\sum_{n=1}^{\infty}\mathrm{E}\chi_{\{\tau_n^{x,\varepsilon}<\tilde{\tau}^{x,\varepsilon}\}} \le K(r,T)\varepsilon^{-(\beta-\alpha/2)}.$$

On the other hand, 4.11 implies that for sufficiently small ε, we have

$$\sup_{x\in\bar{\gamma}}\left|\mathrm{E}\left[f(\bar{Y}^{x,\varepsilon}_{\kappa\wedge\lambda_{H_1}})-f(\Gamma^\varepsilon(x))-\int_0^{\kappa^{x,\varepsilon}\wedge\lambda^{x,\varepsilon}_{H_1}}\mathcal{A}f(\bar{Y}^{x,\varepsilon}_t)dt\right]\right| < \frac{\eta}{5K}\varepsilon^{\beta-\alpha/2}.$$

This implies that

$$\sup_{|H(x)|\le\varepsilon^{\alpha/2}H_0}\left|\sum_{n=1}^{\infty}\mathrm{E}\chi_{\{\tau_n^{x,\varepsilon}<\tilde{\tau}^{x,\varepsilon}\}}\mathrm{E}\left[f(\bar{Y}^{y,\varepsilon}_{\kappa\wedge\lambda_{H_1}})-f(\Gamma^\varepsilon(y))-\int_0^{\kappa^{y,\varepsilon}\wedge\lambda_{H_1}}\mathcal{A}f(\bar{Y}^{y,\varepsilon}_t)dt\right]_{y=Z^{x,\varepsilon}_{\tau_n}}\right| \le$$
$$\sup_{x\in\bar{\gamma}}\left|\mathrm{E}\left[f(\bar{Y}^{x,\varepsilon}_{\kappa\wedge\lambda_{H_1}})-f(\Gamma^\varepsilon(x))-\int_0^{\kappa^{x,\varepsilon}\wedge\lambda^{x,\varepsilon}_{H_1}}\mathcal{A}f(\bar{Y}^{x,\varepsilon}_t)dt\right]\right|\sup_{|H(x)|\le\varepsilon^{\alpha/2}H_0}\sum_{n=1}^{\infty}\mathrm{E}\chi_{\{\tau_n^{x,\varepsilon}<\tilde{\tau}^{x,\varepsilon}\}} \le \frac{\eta}{5}.$$

The same argument allows us to write

$$\sup_{|H(x)|\le\varepsilon^{\alpha/2}H_0}\left|\sum_{n=1}^{\infty}\mathrm{E}\chi_{\{\kappa_n^{x,\varepsilon}<\tilde{\tau}^{x,\varepsilon}\}}\mathrm{E}\left[f(\bar{Y}^{\nu_{\mathcal{L}},\varepsilon}_\tau)-f(\bar{Y}^{\nu_{\mathcal{L}},\varepsilon}_0)-\int_0^{\tau^{\nu_{\mathcal{L}},\varepsilon}}\mathcal{A}f(\bar{Y}^{\nu_{\mathcal{L}},\varepsilon}_t)dt\right]\right| < \frac{\eta}{5},$$

so we are done if we can justify starting from the invariant measure $\nu_{\mathcal{L}}$ in the second expectation. The absolute value of the difference

$$\sum_{n=1}^{\infty}\mathrm{E}\chi_{\{\kappa_n^{x,\varepsilon}<\tilde{\tau}^{x,\varepsilon}\}}\mathrm{E}\left[f(\bar{Y}^{y,\varepsilon}_\tau)-f(\Gamma^\varepsilon(y))-\int_0^{\tau^{y,\varepsilon}}\mathcal{A}f(\bar{Y}^{y,\varepsilon}_t)dt\right]_{y=Z^{x,\varepsilon}_{\kappa_n}} -$$
$$-\sum_{n=1}^{\infty}\mathrm{E}\chi_{\{\kappa_n^{x,\varepsilon}<\tilde{\tau}^{x,\varepsilon}\}}\mathrm{E}\left[f(\bar{Y}^{\nu_{\mathcal{L}},\varepsilon}_\tau)-f(\bar{Y}^{\nu_{\mathcal{L}},\varepsilon}_0)-\int_0^{\tau^{\nu_{\mathcal{L}},\varepsilon}}\mathcal{A}f(\bar{Y}^{\nu_{\mathcal{L}},\varepsilon}_t)dt\right]$$

can be bounded from above by

$$\sup_{x\in\mathcal{L}}\left|\mathrm{E}\left[f(\bar{Y}^{x,\varepsilon}_\tau)-f(\Gamma^\varepsilon(x))-\int_0^{\tau^{x,\varepsilon}}\mathcal{A}f(\bar{Y}^{x,\varepsilon}_t)dt\right]\right|\sum_{n=1}^{\infty}\sup_{x\in\mathcal{L}}d_{TV}((P_2^\varepsilon)^n(x,dy),\nu^\varepsilon_{\mathcal{L}}(dy))$$

which is less than $\eta/5$ for small enough ε by (4.10) and (4.8). This finishes the proof. □

4.2.2 Logarithmic decay of the averaged diffusion coefficient

In this section, we are going to prove the following lemma which captures the inverse logarithmic decay of the averaged diffusion coefficient given by Theorem 2.1 near the separatrix.

Lemma 4.8. *Let $C_i = a(i,z)$ from (2.1) Then for every $H_0 > 0$, every smooth $g : \mathbb{R}^2 \to \mathbb{R}$ with $\sup_{t\in[0,\infty)} g(t,x) < \infty$ for each $x \in \mathbb{R}$, there is an $\varepsilon_0 > 0$ such that*

1.

$$\mathrm{E}\int_0^{\kappa^{x,\varepsilon}\wedge\lambda^{x,\varepsilon}_{H_0}} g(s,\varepsilon^{-\alpha/2}H(Z^{x,\varepsilon}_t))\left[\frac{\alpha}{2}|\log\varepsilon||\nabla H(Z^{x,\varepsilon}_s)|^2 - C_{i(x)}\right] ds = o(\varepsilon^{\beta-\alpha/2}). \tag{4.15}$$

uniformly for all $x \in \bar{\gamma}$,

2.

$$\mathrm{E}\int_0^{\tau^{x,\varepsilon}\wedge\lambda^{x,\varepsilon}_{H_0}} g(s,\varepsilon^{-\alpha/2}H(Z^{x,\varepsilon}_t))\left[\frac{\alpha}{2}|\log\varepsilon||\nabla H(Z^{x,\varepsilon}_s)|^2 - C_{i(x)}\right] ds = o(1). \tag{4.16}$$

uniformly in $|H(x)| \in [\varepsilon^\beta, H_0\varepsilon^{\alpha/2}]$,

whenever $\varepsilon < \varepsilon_0$.

Proof. We sketch the proof of the first part, the verification of the second statement is similar. Let $x^\varepsilon_t(x)$ be the solution of the deterministic equation

$$\dot{x}^\varepsilon = \varepsilon^{\alpha-1}|\log\varepsilon|\nabla^\perp H(x^\varepsilon) \qquad x^\varepsilon_t(x) = x$$

and let $T^\varepsilon(x) = \inf\{t > 0 : x_t^\varepsilon(x) = x\}$. It is not hard to verify that if $H(x) = \varepsilon^\gamma$,

$$\left| \int_0^{T^\varepsilon(x)} g(t_0 + t, \varepsilon^{-\alpha/2} H(x_t)) \left[\frac{\alpha}{2} |\log \varepsilon| |\nabla H(x_t)|^2 - C_i \right] dt \right| \leq C T^\varepsilon(x) \left| \frac{\alpha}{2\gamma} - 1 \right|,$$

for any $t_0 > 0$. Using that $T^\varepsilon(x) = c\gamma\varepsilon^{1-\alpha}(1 + o(1))$ for $H(x) = \varepsilon^\gamma$, which follows from the Hamiltonian nature of the system, and Lemma 3.3 of [24], a similar bound can be shown to hold for the process $Z^{x,\varepsilon}$, namely

$$\left| \mathrm{E} \int_0^{T^\varepsilon(x)} g(t_0 + t, \varepsilon^{-\alpha/2} H(Z_t^{x,\varepsilon})) \left[\frac{\alpha}{2} |\log \varepsilon| |\nabla H(Z_t^{x,\varepsilon})|^2 - C_i \right] dt \right| \leq C T^\varepsilon(x) \left| \frac{\alpha}{2\gamma} - 1 \right|. \tag{4.17}$$

We are going to use (4.17) with γ close enough to $\alpha/2$. This allows us to conclude that the contribution to the integral in (4.15) when the process is close to levels of order $\varepsilon^{\alpha/2}$ is much less than the time it spends there.

By Itô's formula, it is not hard to show that

$$H(Z_{T^\varepsilon}^{x,\varepsilon}) - H(x) = \sqrt{\varepsilon} \int_0^{T^\varepsilon(x)} \nabla H(x_s^\varepsilon) dW_s + \mathcal{O}(\varepsilon).$$

The integral has a centered Gaussian distribution with variance

$$\int_0^{T(x)} |\nabla H(x_s^\varepsilon)|^2 ds = \oint_{\gamma(H(x))} |\nabla H(y)| dl = \oint_{\partial U_k} |\nabla H(y)| dl + \mathcal{O}(\varepsilon^{\alpha/2}),$$

for $x \in V^{\delta,\varepsilon}$. Consequently, there is a standard normal random variable $\mathcal{N}$ such that

$$H(Z_{T^\varepsilon}^{x,\varepsilon}) - H(x) = \sqrt{\varepsilon} \sqrt{\oint_{\partial U_k} |\nabla H(y)| dl} \cdot \mathcal{N} + \mathcal{O}\left(\varepsilon^{\frac{2+\alpha}{4}} \right) \mathcal{N}.$$

Using that

$$T^\varepsilon(x) = c\varepsilon^{1-\alpha} \frac{\log |H(x)|}{\log \varepsilon},$$

an invariance principle suggests that $H(Z_t^{x,\varepsilon})$ can be approximated by the process

$$dH_t^\varepsilon = c \sqrt{\oint_{\partial U_k} |\nabla H(y)| dl} \sqrt{\frac{|\log \varepsilon|}{|\log |H_t^\varepsilon||}} dW_t.$$

with $H_0^\varepsilon = H(Z_0^{x,\varepsilon})$ up until it exits from $U_k \cap V^{\delta,\varepsilon}$. Let us introduce $\tilde{H}_t^\varepsilon = \varepsilon^{-\alpha/2} H_t^\varepsilon$ and note that it solves

$$d\tilde{H}_t^\varepsilon = c \frac{|\log \varepsilon|}{\alpha |\log \varepsilon|/2 + |\log \tilde{H}_t^\varepsilon|} dW_t \qquad \tilde{H}_0^\varepsilon = \varepsilon^{-\alpha/2} H(Z_0^{x,\varepsilon}).$$

It follows from this that for small values of ε, the dynamics is approximately Brownian. Brownian formulas imply that, when $H(Z_0^{x,\varepsilon})$, the expected exit time of $Z_t^{x,\varepsilon}$ from $U_k \cap V^{\delta,\varepsilon}$ is $\mathcal{O}(\varepsilon^{\beta-\alpha/2})$. Similarly, we can conclude that the time spent by $|H(Z_t^{x,\varepsilon})|$ in some interval $[0, \varepsilon^\gamma]$ for any $\gamma > \alpha/2$ converges to zero. These two facts combined with (4.17) imply (4.15).

□

This lemma also yields the following corollary, which we will need to prove tightness.

Lemma 4.9. *For every $H_0, T > 0$, we have that there is a constand $C > 0$ and an $\varepsilon_0 > 0$ such that*

$$\sup_{|H(x)| \le \varepsilon^{\alpha/2} H_0} \mathrm{E} \int_0^{T \wedge \lambda_{H_1}^{x,\varepsilon}} |\nabla H(Z_t^{x,\varepsilon})|^2 dt \le \frac{CT}{|\log \varepsilon|}$$

whenever $\varepsilon < \varepsilon_0$.

Proof. Let $g \equiv 1$. For $|H(x)| \le \varepsilon^{\alpha/2} H_0$ and Let $\tilde{\tau}_{H_0}$ as in, we have

$$\begin{aligned}
\mathrm{E} \int_0^{T \wedge \lambda_{H_1}^{x,\varepsilon}} |\nabla H(Z_t^{x,\varepsilon})|^2 dt \le \mathrm{E} \int_0^{\tau^{x,\varepsilon} \wedge \lambda_{H_0}^{x,\varepsilon}} |\nabla H(Z_t^{x,\varepsilon})|^2 dt + \\
+ \sum_{n=1}^{\infty} \mathrm{E} \left[\chi_{\{\tau_n^{x,\varepsilon} < \tilde{\tau}_{H_0}^{x,\varepsilon}\}} \mathrm{E} \int_0^{\kappa^{y,\varepsilon} \wedge \lambda_{H_0}^{y,\varepsilon}} |\nabla H(Z_t^{x,\varepsilon})|^2 dt \right]_{y = Z_{\tau_n}^{x,\varepsilon}} + \\
+ \sum_{n=1}^{\infty} \mathrm{E} \left[\chi_{\{\kappa_n^{x,\varepsilon} < \tilde{\tau}_{H_0}^{x,\varepsilon}\}} \mathrm{E} \int_0^{\tau^{y,\varepsilon}} |\nabla H(Z_t^{x,\varepsilon})|^2 dt \right]_{y = Z_{\kappa_n}^{x,\varepsilon}}.
\end{aligned}$$

It is straightforward to see using Lemma 4.8 that all terms are $\mathcal{O}(1/|\log\varepsilon|)$ times the expected value of the length of the interval of integration. The claim follows now in a straightforward way. $\square$

Remark 4.1. *Note that the assertions of Lemma 4.9 and therefore Lemma 4.9 remain valid if we let H_0 depend on ε provided $H_0(\varepsilon) \leq C|\log\varepsilon|)$ for some $C > 0$ for small enough ε.*

4.2.3 Proof of the necessary estimates

In this section, we prove Lemma 4.4, Lemma 4.7, and Lemma 4.6 relying on Lemma 4.8.

Proof of Lemma 4.4. We apply Itô's formula to $(\varepsilon^{-\alpha/2}H(Z_t^{x,\varepsilon}))^2 - C_i t$ at time $t = \kappa^{x,\varepsilon} \wedge \lambda_{H_0}^{x,\varepsilon}$ and then take expectations with $x \in \bar{\gamma}$ to get

$$H_0^2 \mathrm{P}(\lambda_{H_0}^{x,\varepsilon} \leq \kappa^{x,\varepsilon}) - C_i \mathrm{E}\kappa^{x,\varepsilon} \wedge \lambda_{H_0}^{x,\varepsilon} = \tag{4.18}$$

$$= \varepsilon^{2\beta-\alpha} + \int_0^{\kappa^{x,\varepsilon}\wedge\lambda_{H_0}^{x,\varepsilon}} \left[\frac{\alpha}{2}|\log\varepsilon||\nabla H(Z_t^{x,\varepsilon})|^2 - C_i\right] dt + \mathcal{O}(\varepsilon^{\alpha/2}|\log\varepsilon|)\mathrm{E}\kappa^{x,\varepsilon} \wedge \lambda_{H_0}^{x,\varepsilon}. \tag{4.19}$$

It follows from Lemma 4.3 in [6] that for some $c > 0$, we have

$$\mathrm{P}(\lambda_{H_0}^{x,\varepsilon} \leq \kappa^{x,\varepsilon}) \leq \frac{1}{H_0}\varepsilon^{\beta-\alpha/2} + c\varepsilon^{\alpha/2}|\log\varepsilon| \tag{4.20}$$

Since $\alpha/2 < \beta < \alpha$, this is easily seen to be $\mathcal{O}(\varepsilon^{\beta-\alpha/2})$.

Rearranging (4.18), using (4.20), and (4.15) with $g \equiv 0$, gives

$$\mathrm{E}\kappa^{x,\varepsilon} \wedge \lambda_{H_0}^{x,\varepsilon} = \frac{\mathcal{O}(\varepsilon^{\beta-\alpha/2})}{C_i + \mathcal{O}(\varepsilon^{\alpha/2}|\log\varepsilon|)}$$

which proves the claim. $\square$

Proof of Lemma 4.7. Applying Itô-s formula for $\exp\left(-\sqrt{\frac{2}{C_i\varepsilon^\alpha}}H(Z_t^{x,\varepsilon})-t\right)$ at time $t=\kappa^{x,\varepsilon}\wedge\lambda_{H_0}^{x,\varepsilon}$ and taking expectations gives us, after simple manipulations, that for $x\in\bar{\gamma}$

$$\mathrm{E}e^{-\kappa^{x,\varepsilon}\wedge\lambda_{H_0}^{x,\varepsilon}}=\mathrm{E}e^{-\kappa^{x,\varepsilon}\wedge\lambda_{H_0}^{x,\varepsilon}}\chi_{\{\lambda_{H_0}^{x,\varepsilon}\le K\}}\left(1-\varepsilon^{-\sqrt{\frac{2}{C_i}}H_0}\right)+e^{-\sqrt{\frac{2}{C_i}}\varepsilon^{\beta-\alpha/2}}+$$
$$+\mathrm{E}\int_0^{\kappa^{x,\varepsilon}\wedge\lambda_{H_0}^{x,\varepsilon}}e^{-\sqrt{\frac{2}{C_i\varepsilon^\alpha}}H(Z_s^{x,\varepsilon})-t}\left[\frac{\alpha}{2C_i}|\log\varepsilon||\nabla H(Z_s^{x,\varepsilon})|^2-1\right]dt+\mathcal{O}(\varepsilon^{\alpha/2}|\log\varepsilon|).$$

Here, we used that $\mathrm{E}\kappa^{x,\varepsilon}\wedge\lambda_{H_0}^{x,\varepsilon}\le\mathrm{E}\lambda_{H_0}^{x,\varepsilon}<K$ for some $K>0$ by (4.6). After elementary manipulations, we see that

$$Ee^{-\kappa^{x,\varepsilon}\wedge\lambda_{H_0}^{x,\varepsilon}}=e^{-\sqrt{\frac{2}{C_i}}\varepsilon^{\beta-\alpha/2}}+\left(1-\varepsilon^{-\sqrt{\frac{2}{C_i}}H_0}\right)\mathrm{P}(\lambda_0^{x,\varepsilon}\le\kappa^{x,\varepsilon})+o(\varepsilon^{\beta-\alpha/2})$$

by $\beta<\alpha$ and the application of (4.15) with $g(s,z)=e^{-\sqrt{2/C_i}z-s}$. (4.20) implies that this can be written as

$$Ee^{-\kappa^{x,\varepsilon}\wedge\lambda_{H_0}^{x,\varepsilon}}=e^{-\sqrt{\frac{2}{C_i}}\varepsilon^{\beta-\alpha/2}}+\frac{1-\varepsilon^{-\sqrt{\frac{2}{C_i}}H_0}}{H_0}\varepsilon^{\beta-\alpha/2}+o(\varepsilon^{\beta-\alpha/2}).$$

The result follows from this formula by elementary considerations using that the function $x^{-1}(1-e^{-x})$ is strictly between 0 and 1 for every $x>0$ and that $\kappa^{x,\varepsilon}\ge\kappa^{x,\varepsilon}\wedge\lambda_{H_0}^{x,\varepsilon}$. □

Proof of Lemma 4.6. Let us prove (4.12) first. By splitting up the expectation of $f(Y_\tau^{\nu_\mathcal{L},\varepsilon})$ with respect to which edge it belongs to, one can write

$$E\left[f(Y_\tau^{\nu_\mathcal{L},\varepsilon})-f(\bar{Y}_0^{\nu_\mathcal{L},\varepsilon})-\int_0^{\tau^{\nu_\mathcal{L},\varepsilon}}\mathcal{A}f(\bar{Y}_s^{\nu_\mathcal{L},\varepsilon})ds\right]=$$
$$=\varepsilon^{\beta-\alpha/2}\sum_{i=1}^n\nu_\mathcal{L}^\varepsilon(\bar{\gamma}_i)D_iF(0)+||\mathcal{A}f||\mathrm{E}\tau^{x,\varepsilon}.$$

The sum in the first term converges to some constant times $\sum_{i=1}^n\alpha_iD_iF(0)=0$ by (4.9). On the other hand, the second term is of order $\varepsilon^{2\beta-\alpha}$ by (4.5) which proves the claim.

To prove (4.10), note that for $|H(x)| \leq \varepsilon^\beta$, we have $\tau^{x,\varepsilon} \leq \lambda_{H_0}^{x,\varepsilon}$. Using this and (4.5), we can write

$$\sup_{|H(x)|\leq\varepsilon^\beta} \left| \mathrm{E}\left[f(\bar{Y}^{x,\varepsilon}_{\tau\wedge\lambda_{H_0}}) - f(\Gamma^\varepsilon(x)) - \int_0^{\tau^{x,\varepsilon}\wedge\lambda_{H_0}^{x,\varepsilon}} \mathcal{A}f(\bar{Y}^{x,\varepsilon}_s)ds \right] \right| \leq 2||f'||\varepsilon^{\beta-\alpha/2} + ||\mathcal{A}f||\varepsilon^{2\beta-\alpha}.$$

We still have to show that (4.10) holds for starting points $|H(x)| \in [\varepsilon^\beta, \varepsilon^{\alpha/2}H_0]$. Let $f_x(y) = f(y, i(x))$. Using Itô's formula for $f_x(\varepsilon^{-\alpha/2}H(Z_t^{x,\varepsilon}))$ at time $\tau^{x,\varepsilon} \wedge \lambda_{H_0}^{x,\varepsilon}$ and taking expectations gives for $|H(x)| \in [\varepsilon^\beta, \varepsilon^{\alpha/2}H_0]$

$$\begin{aligned}
&\left| \mathrm{E}\left[f(\bar{Y}^{x,\varepsilon}_{\tau\wedge\lambda_{H_0}}) - f(\Gamma^\varepsilon(x)) - \int_0^{\tau^{x,\varepsilon}\wedge\lambda_{H_0}^{x,\varepsilon}} \mathcal{A}f(\bar{Y}^{x,\varepsilon}_s)ds \right] \right| = \\
&= \left| \mathrm{E}\left[f_x(\varepsilon^{-\alpha/2}H(Z^{x,\varepsilon}_{\tau\wedge\lambda_{H_0}})) - f_x(\Gamma^\varepsilon(x)) - \int_0^{\tau^{x,\varepsilon}\wedge\lambda_{H_0}^{x,\varepsilon}} \mathcal{A}f_x(\varepsilon^{-\alpha/2}H(Z^{x,\varepsilon}_s))ds \right] \right| \leq \\
&\leq \left| \mathrm{E}\int_0^{\tau^{x,\varepsilon}\wedge\lambda_{H_0}^{x,\varepsilon}} \frac{1}{2}f_x''(\varepsilon^{-\alpha/2}H(x)) \left[\frac{\alpha}{2}|\log\varepsilon||\nabla H(Z^{x,\varepsilon}_s)|^2 - C_i\right] ds \right| + \mathcal{O}(\varepsilon^{\alpha/2}\log\varepsilon)\mathrm{E}\lambda_{H_0}^{x,\varepsilon}.
\end{aligned}$$

Now (4.10) follows from $\beta < \alpha$, (4.16) with $g(s,z) = f_x''(z)/2$, and (4.6).

The proof of (4.11) is similar to this last case. Again by Itô's formula and taking expectations,

$$\begin{aligned}
&\left| \mathrm{E}\left[f(\bar{Y}^{x,\varepsilon}_{\kappa\wedge\lambda_{H_0}}) - f(\Gamma^\varepsilon(x)) - \int_0^{\kappa^{x,\varepsilon}\wedge\lambda_{H_0}^{x,\varepsilon}} \mathcal{A}f(\bar{Y}^{x,\varepsilon}_s)ds \right] \right| \leq \\
&\leq \left| \mathrm{E}\int_0^{\kappa^{x,\varepsilon}\wedge\lambda_{H_0}^{x,\varepsilon}} \frac{1}{2}f_x''(\varepsilon^{-\alpha/2}H(x)) \left[\frac{\alpha}{2}|\log\varepsilon||\nabla H(Z^{x,\varepsilon}_s)|^2 - C_i\right] ds \right| + \mathcal{O}(\varepsilon^{\alpha/2}\log\varepsilon)\mathrm{E}\lambda_{H_0}^{x,\varepsilon}.
\end{aligned}$$

Using $\beta < \alpha$ and (4.6), the last term is $o(\varepsilon^{\beta-\alpha/2})$. To show that the first term is also of the same order, we use (4.15) with $g(s,z) = f_x''(z)/2$. □

4.2.4 Tightness

Lemma 4.10. *The family of processes $Y^{\mu^\varepsilon,\varepsilon}$ is tight.*

We will now prove Lemma 4.2 which we recall for convenience.

Lemma 4.11. *For every $\eta > 0, T > 0$ and $H_0 > 0$, there are $H_0', \varepsilon_0 > 0$ such that we have*

$$\sup_{|H(x)|\le\varepsilon^{\alpha/2}H_0} \mathrm{P}\left(\max_{t\in[0,T]} |H(Z_t^{x,\varepsilon})| \ge \varepsilon^{\alpha/2}H_0'\right) < \eta \tag{4.21}$$

whenever $\varepsilon < \varepsilon_0$.

Proof. Applying Itô's formula for $H(Z_t^{x,\varepsilon})$, we get

$$\varepsilon^{-\alpha/2}H(Z_t^{x,\varepsilon}) =$$
$$= \varepsilon^{-\alpha/2}H(x) + \sqrt{\frac{\alpha\log\varepsilon}{2}}\int_0^t \nabla H(Z_s^{x,\varepsilon})dW_s + \frac{\alpha\varepsilon^{\alpha/2}\log\varepsilon}{4}\int_0^t \Delta H(Z_s^{x,\varepsilon})ds.$$

This implies that if $\varepsilon_1 > 0$ is such that $\varepsilon^{\alpha/2}\log\varepsilon < h$ for every $\varepsilon < \varepsilon_1$ and $H_0' > H_0 + \alpha||\Delta H||h/4$ then

$$\sup_{|H(x)|\le\varepsilon^{\alpha/2}H_0} \mathrm{P}\left(\max_{t\in[0,T]} |H(Z_t^{x,\varepsilon})| \ge \varepsilon^{\alpha/2}H_0'\right) \le$$
$$\le \sup_{|H(x)|\le\varepsilon^{\alpha/2}H_0} \mathrm{P}\left(\max_{t\in[0,T]} \left|\int_0^t \nabla H(Z_s^{x,\varepsilon})dW_s\right| \ge \frac{\sqrt{2}\left(H_0' - H_0 - \frac{\alpha||\Delta H||}{4}h\right)}{\sqrt{\alpha|\log\varepsilon|}}\right).$$

holds for every $\varepsilon < \varepsilon_1$. By the martingale moment inequality and Lemma 4.9, there is an ε_0 such that

$$\sup_{|H(x)|\le\varepsilon^{\alpha/2}H_0} \mathrm{P}\left(\max_{t\in[0,T]} |H(Z_t^{x,\varepsilon})| \ge \varepsilon^{\alpha/2}H_0'\right) \le \frac{\alpha|\log\varepsilon|\sup_{|H(x)|\le\varepsilon^{\alpha/2}H_0}\mathrm{E}\int_0^T|\nabla H(Z_t^{x,\varepsilon})|^2dt}{2(H_0'-H_0-\frac{\alpha||\Delta H||h}{4})^2} \tag{4.22}$$

We will show that the numerator is bounded as $\varepsilon \downarrow 0$ and therefore the lemma is proved if we choose H_0' large enough. To do this, let $H_1^\varepsilon = H_2|\log\varepsilon|$ for some $H_2 > 0$. We can write

$$\mathrm{E}\int_0^T |\nabla H(Z_t^{x,\varepsilon})|^2dt \le \mathrm{E}\int_0^{T\wedge\lambda_{H_1(\varepsilon)}^{x,\varepsilon}} |\nabla H(Z_t^{x,\varepsilon})|^2dt + T||\nabla H||^2\mathrm{P}(\lambda_{H_1(\varepsilon)}^{x,\varepsilon} \le T). \tag{4.23}$$

By the same analysis as the one led to (4.22), we have

$$\sup_{|H(x)|\le\varepsilon^{\alpha/2}H_0} \mathrm{P}(\lambda^{x,\varepsilon}_{H_1(\varepsilon)} \le T) \le \frac{\alpha|\log\varepsilon|\mathrm{E}\int_0^T|\nabla H(Z^{x,\varepsilon}_t)|^2dt}{2(H_2|\log\varepsilon| - H_0 - \frac{\alpha||\Delta H||h}{4})^2} \le$$

$$\le \frac{\alpha T|\log\varepsilon|||\nabla H||^2}{2(H_2|\log\varepsilon| - H_0 - \frac{\alpha||\Delta H||h}{4})^2} \le \frac{C}{|\log\varepsilon|}.$$

Combining (4.22) with this, Lemma 4.9, Remark 4.1, and (4.23) proves the claim.

□

We are going to use Theorem 2.1 from [23], which is in turn a variant of Theorem 1.4.6 in [30].

Theorem 4.3. *Assume that for every compact $K \subseteq \bar{G}$ and sufficiently small ρ, there is a constant A_ρ such that for every $a \in K$, there exists a function f^a_ρ on G such that $f^a_\rho(a) = 1$, $f^a_\rho(y) = 0$ for $d(y,a) \ge \rho$, $0 \le f^a_\rho(y) \le 1$ everywhere, and $f^a_\rho(Y^{\mu^\varepsilon,\varepsilon}(t)) + A_\rho t$ is a submartingale for all ε. Then, if (4.21) is also satisfied the family $\{Y^{\mu^\varepsilon,\varepsilon}\}_{\varepsilon\in(0,\varepsilon_0]}$ is tight.*

Proof of Lemma 4.10. Inspired by the proof of Lemma 3.2 in Chapter 8 of [23], let h be a smooth function on $[0,1]$ such that $h(0) = 1$, $h(1) = 0$ and $0 \le h \le 1$ everywhere, and define

$$f^a_\rho(y) = \begin{cases} h(5d(y,a)/\rho) & \text{if} \quad d(a,O) > 2\rho/5 \\ h(5d(y,O)/\rho - 2) & \text{if} \quad d(a,O) \le 2\rho/5 \end{cases}$$

Note that f^a_ρ satisfies the requirements in Theorem 4.3.

Also observe that $g^{\rho,\varepsilon}_a(x) = f^a_\rho(\Gamma^\varepsilon(x)) = f^a_\rho(\varepsilon^{-\alpha/2}H(x))$ is twice continuously differentiable (as $h'(0) = h''(0) = h'(-2) = h''(-2) = 0$) and its gradient is orthogonal to the flow-lines. Therefore by Itô's formula, we get

$$f_\rho^a(Y^{\mu^\varepsilon,\varepsilon}) = M^{\varepsilon,a}(t) + \frac{\alpha}{4}\int_0^t \partial_{yy} f_\rho^a(Y^{\mu^\varepsilon,\varepsilon})|\nabla H(Z^{\mu^\varepsilon,\varepsilon})|^2 \log\varepsilon ds +$$
$$+ \int_0^t \partial_y f_\rho^a(\varepsilon^{-\alpha/2} H(Z_s^\varepsilon))\Delta H(Z^{\mu^\varepsilon,\varepsilon})\varepsilon^{\alpha/2}\log\varepsilon ds,$$

where $M^{\varepsilon,a}$ is a martingale. Note that the integrand in the second term is bounded for small enough ε by some constant A_ρ^1. Using Lemma 4.9, the expectation of the first integral is also bounded by some $A_\rho^2 t$. Therefore, by the strong Markov property,

$$\mathrm{E}[f_\rho^a(Y^{\mu^\varepsilon,\varepsilon}(t))|\mathcal{F}_s] \geq f_\rho^a(Y^{\mu^\varepsilon,\varepsilon}(s)) - (A_\rho^1 + A_\rho^2)(t-s)$$

As $\mu^\varepsilon \circ (\Gamma^\varepsilon)^{-1}$, $Y^{\mu^\varepsilon,\varepsilon}(0)$ is tight and thus the proof is completed by Theorem 4.3 with $A_\rho = A_\rho^1 + A_\rho^2$.

$\square$

Chapter 5: Proofs of the PDE results

5.1 The elliptic problem

Proof of Theorem 2.5. Part 1. By the representation formula,

$$u^{\varepsilon,R}(x) = \mathrm{E}\int_0^{\tau_{\partial D_R}(X_\cdot^{x,\varepsilon})} f(X_s^{x,\varepsilon}/R)\, ds \ ,$$

which can be decomposed as

$$\mathrm{E}\int_0^{\tau_{\mathcal{L}}(X_\cdot^{x,\varepsilon})} f(X_s^{x,\varepsilon}/R) ds + \mathrm{E}\int_{\tau_{\mathcal{L}}(X_\cdot^{x,\varepsilon})}^{\tau_{\partial D_R}(X_\cdot^{x,\varepsilon})} f(X_s^{x,\varepsilon}/R)\, ds \ ,$$

where $\tau_{\mathcal{L}}$ is the first time the process hits the separatrix. The first term can easily be seen to converge by the averaging principle (Theorem 2.1) to $f(0)\mathrm{E}\tau_O(Y_\cdot^{\Gamma(x)})$, and thus it remains to show that the second term converges to zero. It suffices to show that $\mathrm{E}(\tau_{\partial D_R}(X_\cdot^{x,\varepsilon}) - \tau_{\mathcal{L}}(X_\cdot^{x,\varepsilon})) \to 0$ as $\varepsilon \to 0$.

With a slight abuse of notation, let $\mathcal{T}$ be the copy of the domain of periodicity that contains the origin. Recall that $\mathcal{L}_{\mathcal{T}}$ is the projection of $\mathcal{L}$ on the torus. Equivalently, we can view it as a set on the plane that is the intersection of $\mathcal{L}$ and $\mathcal{T}$. Thus it is sufficient to show that

$$\sup_{x\in\mathcal{L}_{\mathcal{T}}} \mathrm{E}\tau_{\partial D_R}(X_\cdot^{x,\varepsilon}) \to 0 \quad \text{as } \varepsilon \downarrow 0,\ R = R(\varepsilon)\ . \tag{5.1}$$

We claim that

$$\sup_{x\in\mathcal{L}_\mathcal{T}} \mathrm{P}(\tau_{\partial D_R}(X_\cdot^{x,\varepsilon}) > K) \to 0 \quad \text{as } \varepsilon \downarrow 0,\ R = R(\varepsilon) \tag{5.2}$$

for each $K > 0$, and that there is $\varepsilon_0 > 0$ such that

$$\sup_{\varepsilon\in(0,\varepsilon_0]} \sup_{x\in\mathbb{R}^2} \mathrm{P}(\tau_\mathcal{L}(X_\cdot^{x,\varepsilon}) > 1) < 1\ . \tag{5.3}$$

The latter easily follows from the averaging principle (see [3], Chapter 8), while the former will be justified below.

Note that

$$\sup_{x\in\mathcal{L}_\mathcal{T}} \mathrm{E}\tau_{\partial D_R}(X_\cdot^{x,\varepsilon}) \le \int_0^\infty \sup_{x\in\mathcal{L}_\mathcal{T}} \mathrm{P}(\tau_{\partial D_R}(X_\cdot^{x,\varepsilon}) > K)\, dK\ .$$

By (5.2), the integrand tends to zero for each K. Also note that the integrand decays exponentially in K, uniformly in ε, as follows from (5.2), (5.3), and the Markov property of the process. This justifies (5.1).

We still need to prove (5.2). For a given value of $\delta > 0$ and all sufficiently small ε, we have

$$\tau_{\partial D_R}(X_\cdot^{x,\varepsilon}) \le \tau_{B(0,\delta)}(\varepsilon^{1/4} X_\cdot^{x,\varepsilon})\ ,$$

where $\tau_{B(0,\delta)}$ is the time to reach the boundary of the ball of radius δ centered at the origin. By Theorem 2.3,

$$\mathrm{P}(\tau_{B(0,\delta)}(\varepsilon^{1/4} X_\cdot^{x,\varepsilon}) > K) \to \mathrm{P}(\tau_{B(0,\delta)}(\tilde{W}^Q_{L^0_\cdot}) > K) \quad \text{as } \varepsilon \downarrow 0\ ,$$

since the boundary of the event on the right hand side has probability zero. It remains to note that we can make the right hand side arbitrarily small by choosing a sufficiently small δ. This is possible since $\mathrm{P}(L^0_t > 0) = 1$ for each $t > 0$ (as follows from (3.17) and the elementary properties of the Brownian motion).

Part 2. Let us first assume that $f \geq 0$. Observe that for each $t > 0$ we have

$$\mathrm{E}\int_0^{\tau_{\partial D_R}(X^{x,\varepsilon}_\cdot)\wedge t} f(R^{-1}X^{x,\varepsilon}_s)ds = \mathrm{E}\int_0^{\tau_{\partial D}(R^{-1}X^{x,\varepsilon}_\cdot)\wedge t} f\left(R^{-1}X^{x,\varepsilon}_s\right) ds =: \mathrm{E}I^t_f(R^{-1}X^{x,\varepsilon}_\cdot).$$

By Theorem 2.3, the processes $R^{-1}X^{x,\varepsilon}_\cdot$ converge weakly to $C^{-1}W^Q_{L^{\Gamma(x)}_\cdot}$. Since I^t_f is bounded and is continuous almost surely with respect to the measure induced by $C^{-1}W^Q_{L^{\Gamma(x)}_\cdot}$, we have

$$\mathrm{E}\int_0^{\tau_{\partial D_R}(X^{x,\varepsilon}_\cdot)\wedge t} f(X^{x,\varepsilon}_s/R)\,ds \to \mathrm{E}\int_0^{\tau_{\partial D}(C^{-1}W^Q_{L^{\Gamma(x)}_\cdot})\wedge t} f(C^{-1}W^Q_{L^{\Gamma(x)}_s})\,ds \quad \text{as } \varepsilon\downarrow 0. \tag{5.4}$$

As in the proof of Part 1, we have that $\mathrm{P}(\tau_{\partial D_R}(X^{x,\varepsilon}_\cdot) > K)$ decays exponentially in K, uniformly in ε, which justifies the fact that we can take $t = \infty$ in (5.4). The general case follows by taking $f = f_+ - f_-$.

Part 3. The PDE result easily follows from the weak convergence of the corresponding processes. More precisely, let $\bar{X}^{x,\varepsilon}_t = R^{-1}(\varepsilon)X^{x,\varepsilon}_{\varepsilon^{1/2}R(\varepsilon)^2 t}$. We need to show that

$$\bar{X}^{x,\varepsilon}_\cdot \Rightarrow \tilde{W}^{cQ}_\cdot \qquad \text{as } \varepsilon\downarrow 0. \tag{5.5}$$

It follows from [6] that

$$\frac{\varepsilon^{1/4}X^{x,\varepsilon}_{k\cdot}}{\sqrt{k}} \Rightarrow \tilde{W}^{D(\varepsilon)}_\cdot \qquad \text{as } k\to\infty, \tag{5.6}$$

where $D(\varepsilon) = D_0 + o(1)$ and D_0 is a constant multiple of Q. (Strictly speaking, the result in [6] concerns the finite dimensional distributions, but the generalization to the functional CLT is standard in this situation.) Moreover, it is not difficult to show (by following the proof in [6] and using arguments similar to those in the the proof of Lemma 3.2) that the convergence is uniform in ε. Therefore, (5.6) implies (5.5) with $cQ = D_0$. □

5.2 The parabolic problem

Proof of Theorem 2.6. Consider first Part 1 and note that

$$u^{\varepsilon,R}(x,t) = \mathrm{E}\left(\frac{X_t^{x,\varepsilon}}{R}\right) = \mathrm{E}f\left(\frac{X_t^{x,\varepsilon}}{R}\right)\chi_{\{t>\tau_{\mathcal{L}}\}} + f(0)\mathrm{P}(\tau_{\mathcal{L}} \geq t) + o(1)$$

where the second term on the right hand side converges to $f(0)\mathrm{P}(\tau_O \geq t)$ by Theorem 2.1. To show that the first term converges to zero, we can clearly assume without loss of generality that $x \in \mathcal{L}$. Pick an $\eta > 0$. Then, as f vanishes at infinity, there exists a $K > 0$ such that $|f(x)| < \eta/2$ whenever $|x| > K$. With this, we can write

$$\mathrm{E}f\left(\frac{X_t^{x,\varepsilon}}{R}\right)\chi_{\{t>\tau_{\mathcal{L}}\}} \leq \frac{\eta}{2} + ||f||_\infty \mathrm{P}(X_t^{x,\varepsilon} \leq KR).$$

For any $\delta > 0$, we have $\varepsilon^{1/4}R < \delta$ if ε is small enough. For such an ε, we have

$$\mathrm{P}(X_t^{x,\varepsilon} \leq KR) \leq \mathrm{P}(|\varepsilon^{1/4}X_t^{x,\varepsilon}| < K\delta) \to \mathrm{P}(\tilde{W}^Q_{L_t^O} < K\delta)$$

as $\varepsilon \downarrow 0$ by Theorem 2.3. Since $P(L_t^O > 0) = 1$ for each $t > 0$, the right hand side can be made less than $\eta/2$ by choosing δ small enough. Since η was arbitrary, the result follows.

Part 2 is proved easily by noticing that Theorem 2.3 and the fact that f is continuous and bounded implies

$$u^{\varepsilon,R}(x,t) = \mathrm{E}f\left(\frac{X_t^{x,\varepsilon}}{R}\right) = \mathrm{E}f\left(\frac{\varepsilon^{1/4}X_t^{x,\varepsilon}}{\varepsilon^{1/4}R}\right) \to \mathrm{E}f\left(W^{Q/C^2}_{L_t^{\Gamma(x)}}\right).$$

Part 3 follows similarly from (5.5) and then referring to Remark 2.1. □

Bibliography

[1] M. Hairer, L. Koralov, and Zs. Pajor-Gyulai. From averaging to homogenization in cellular flows - an exact description of the transition. *Accepted for publication in Annales de l'Institut Henri Poincaré*, 2014.

[2] B. Shraiman. Diffusive transport in rayleigh-benard convection cells. *Phys. Rev. A*, 36:261–267, 1987.

[3] M. Freidlin and A. Wentzell. *Random Perturbations of Dynamical Systems*. Grundlehren der mathematischen Wissenschaften. Springer, 2012.

[4] M. Freidlin. The dirichlet problem for an equation with periodic coefficients depending on a small parameter. (9):133–139, 1964.

[5] G. Iyer and A. Novikov. Anomalous diffusion in fast cellular flows at intermediate time scales.

[6] L. Koralov. Random perturbations of 2-dimensional hamiltonian flows. *Probability Theory and Related Fields*, 129(1):37–62, 2004.

[7] V. Kaloshin, Dolgopyat D., and L. Koralov. Long time behavior of periodic stochastic flows. *In XIVth International congress on Mathematical Physics*, pages 290–295, 2005.

[8] Y.-Y. Liu, J. Xin, and Y. Yu. Asymptotics for turbulent flame speeds of the viscous g-equation enhanced by cellular and shear flows. *Arch. Ration. Mech. Anal.*, 202(2):461–492, 2011.

[9] J. Nolen, J. Xin, and Y. Yu. Bounds on front speeds for inviscid and viscous g-equations. *Methods Appl. Anal*, 16(4):507–520, 2009.

[10] T. Shaw, J.-L. Thiffeault, and C. Doering. Stirring up trouble: multi-scale mixing measures for steady scalar sources. *Phys. D*, 231(2):143–164, 2007.

[11] R. Sowers. Random perturbation of two-dimensional pseudoperiodic flows. *Illinois J. Math.*, 50(1-4):853–959, 2006.

[12] G. Taylor. Dispersion of soluble matter in solvent flowing slowly through a tube. *Proc. R. Soc. Lond. A*, 219(1137):186–203, 1953.

[13] J.-L. Thiffeault and S. Childress. Stirring by swimming bodies. *Phys. Lett. A*, 374(34):3487–3490, 2010.

[14] W. Young. Arrested shear dispersion and other models of anomalous diffusion. *J. Fluid Mech.*, 193:129–149, Aug 1988.

[15] W. Young, A. Pumir, and Y. Pomeau. Anomalous diffusion of tracer in convection rolls. *Phys. Fluids A*, 1(3):462–469, 1989.

[16] Y. Bakhtin. Noisy heteroclinic networks. *Probab. Theory Relat. Fields*, 150:1–42, 2011.

[17] A. Bensoussan, J.L. Lions, and G. Papanicolaou. *Asymptotic Analysis for Periodic Structures.* Studies in Mathematics and its Applications. Elsevier Science, 1978.

[18] V. Zhikov, S. Kozlov, and O. Oleĭnik. *Homogenization of differential operators and integral functionals.* Springer-Verlag, 1994.

[19] G. Pavliotis and A. Stuart. *Multiscale Methods: Averaging and Homogenization.* Texts in Applied Mathematics. Springer, 2008.

[20] G. Iyer, T. Komorowski, A. Novikov, and L. Ryzhik. From homogenization to averaging in cellular flows. *Annales de l'Institut Henri Poincare (C) Non Linear Analysis*, (0):–, 2013.

[21] G. Iyer and K. Zygalakis. Numerical studies of homogenization under a fast cellular flow. *Multiscale Modeling and Simulation*, 10(3):1046–1058, 2012.

[22] G. Pavliotis, A. Stuart, and K. Zygalakis. Calculating effective diffusivities in the limit of vanishing molecular diffusion. *Journal of Computational Physics*, 228(4):1030 – 1055, 2009.

[23] M. Freidlin and A. Wentzell. Diffusion processes on graphs and the averaging principle. *The Annals of Probability*, 21(4):2215–2245, 10 1993.

[24] D. Dolgopyat and L. Koralov. Averaging of hamiltonian flows with an ergodic component. *Ann. Probab.*, 36(6):1999–2049, 2008.

[25] D. Dolgopyat and L. Koralov. Averaging of incompressible flows on two-dimensional surfaces. *J. Amer. Math. Soc.*, 26(2):427–449, 2008.

[26] M. Freidlin and S-J. Sheu. Diffusion processes on graphs: stochastic differential equations, large deviation principle. *Probability Theory and Related Fields*, 116(2):181–220, 2000.

[27] G. Ben Arous and J. Černý. Scaling limit for trap models on $\mathbb{Z}^d$. *The Annals of Probability*, 35(6):2356–2384, 11 2007.

[28] E. Bolthausen and I. Goldsheid. Lingering random walks in random environment on a strip. *Communications in Mathematical Physics*, 278(1):253–288, 2008.

[29] I. Karatzas and S. Shreve. *Brownian Motion and Stochastic Calculus*. Graduate Texts in Mathematics. Springer New York, 1991.

[30] D. Strook and S. Varadhan. Multidimensional diffusion processes. 1971.