Random Perturbations of 2-dimensional Hamiltonian Flows

Leonid Koralov

Department of Mathematics
Princeton University
Princeton, NJ 08544
koralov@math.princeton.edu

Abstract

We consider the motion of a particle in a periodic two dimensional flow perturbed by small (molecular) diffusion. The flow is generated by a divergence free zero mean vector field. The long time behavior corresponds to the behavior of the homogenized process - that is diffusion process with the constant diffusion matrix (effective diffusivity). We obtain the asymptotics of the effective diffusivity when the molecular diffusion tends to zero.

1 Introduction

Consider the following stochastic differential equation

$$dX_t^\varepsilon = v(X_t^\varepsilon)dt + \sqrt{\varepsilon} dW_t \ , \quad X_t^\varepsilon \in \mathbb{R}^2 \ . \tag{1}$$

Here $v(x)$ is an incompressible periodic vector field, W_t is a 2-dimensional Brownian motion, and ε (molecular diffusivity) is a small parameter. We further assume that the stream function $H(x_1, x_2)$, such that

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

is itself periodic in both variables, that is the integral of v over the periodicity cell is zero. For simplicity of notation assume that the period of H in each of the variables is equal to one.

It is well known (see for example [4]), that with ε fixed, the solution of (1) scales like a diffusion process with constant diffusion matrix when time goes to infinity. More precisely, there exists the limit, called the effective diffusivity,

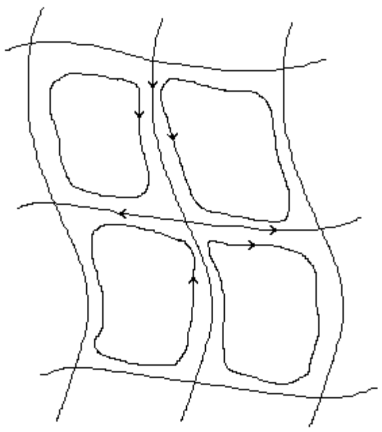
$$D^{ij}(\varepsilon) = \lim_{t\to\infty} \mathbf{E}_\lambda \frac{X_t^{\varepsilon i} X_t^{\varepsilon j}}{t} \ , \quad i,j = 1,2 \ ,$$

where i and j are the coordinates and λ is the initial distribution of the process X_t^ε, which we can take to be an arbitrary measure with compact support. The measure on $C([0,T],\mathbb{R}^2)$, induced by the process $\frac{1}{\sqrt{c}}X_{ct}^\varepsilon$, converges weakly, when $c \to \infty$, to the measure induced by the diffusion process with constant matrix $D(\varepsilon)$.

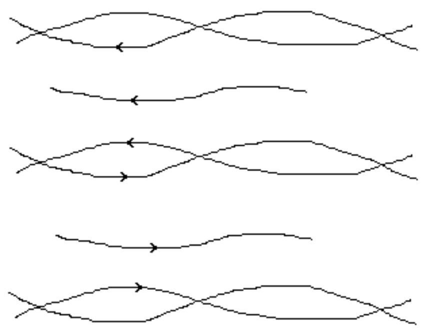
We are interested in the behavior of the effective diffusivity when the molecular diffusion ε tends to zero. Assume that all the critical points of H are non degenerate. We distinguish two qualitatively different cases, depending on the structure of the stream lines of the flow given by $v(x)$.

In the first case, there is a level set of H, which contains some of the saddle points, and which forms a lattice in $\mathbb{R}^2$, thus dividing the plane into bounded sets, invariant under the flow. A standard example of a cellular flow, which has been studied in several of the papers cited below, is the flow with the stream function $H(x_1,x_2) = \sin(2\pi x_1)\sin(2\pi x_2)$. In this particular example the separatrices (the level sets of H containing saddle points) form a rectangular lattice.

In the second case, there is more than one unbounded level set of H containing critical points, and thus there are 'open channels' in the flow, and some of the solutions of the equation $x'(t) = v(x(t))$ go off to infinity. An example of a flow with open channels is the flow with the stream function $H(x_1,x_2) = \sin(2\pi x_1)\sin(2\pi x_2) + 10\sin(2\pi x_2)$. Indeed, the horizontal axis $\{x_2 = 0\}$ is an unbounded stream line of the flow.



cellular flow



flow with open channels

Since $v(x)$ is periodic, we may consider $x'(t) = v(x(t))$ as the flow on the torus. The torus is then a union of the sepatatrices and a finite number of open domains, bounded by the separatrices, and invariant under the flow.

In [3] Fannjiang and Papanicolaou considered cellular flows for which the separatrices form a rectangular lattice on $\mathbb{R}^2$ and the stream function satisfies certain symmetry conditions. They showed that in this case

$$D^{ij}(\varepsilon) = (d^{ij} + o(1))\sqrt{\varepsilon}\ , \quad \text{as } \ \varepsilon \to 0\ , \tag{2}$$

that is the effective diffusivity is enhanced by a factor of order $\varepsilon^{-\frac{1}{2}}$ compared to case of the diffusion process $\sqrt{\varepsilon}W_t$ without the advection term. Moreover, they found the constant matrix d^{ij} explicitly. Their proof is based on a variational principle applied to a symmetric operator associated to the generator of the process X_t^ε. In [6] Heinze provided certain upper and lower estimates on the effective diffusivity in the case of cellular flows, for which the separatrices form a rectangular lattice on $\mathbb{R}^2$.

There are earlier physical papers ([1], [8], [9]), arguing that the asymptotics in (2) is true for particular flows. Our first result is the rigorous proof of this statement for general cellular flows.

Theorem 1.1. *Assume that an infinitely smooth periodic stream function $H(x_1, x_2)$ defines a cellular flow, and that its critical points are nondegenerate. Then the asymptotics of the effective diffusivity for the process (1) is given by (2).*

Let $\mathcal{L}_p$ be the noncompact connected level set of H. This level set contains some of the saddle points of H and forms a lattice in $\mathbb{R}^2$. Without loss of generality we may assume that $H(x) = 0$ when $x \in \mathcal{L}_p$. The corresponding level set on the torus will be denoted by $\mathcal{L}$.

The process X_t^ε consists of the 'fast' part, which is the periodic motion along the streamlines, and the 'slow' diffusion across them. The motion is almost periodic away from the separatrices. However, once the trajectory is in a sufficiently small neighborhood of the level set $\mathcal{L}_p$, it is likely to continue along it, and may go from cell to cell in a time much shorter than it would take the 'slow' diffusion to cover the same distance.

The rough outline of the proof of Theorem 1.1 is the following. We introduce a Markov chain, which can be viewed as a discrete time version of the process X_t^ε. The state space for the Markov chain is $\mathcal{L}$. Note, that due to the periodicity of H, the process X_t^ε can be viewed as a process on the torus. In order to define the transition probabilities, we introduce stopping times for the process X_t^ε. The stopping time τ_0^ε is the first time when X_t^ε hits $\mathcal{L}$, and τ_n^ε is defined as the first time after τ_{n-1}^ε when the process X_t^ε returns to $\mathcal{L}$, after having traveled 'past' a saddle point. The transition times of the Markov chain are random.

We show that the study of the asymptotics of the effective diffusivity can be reduced to the study of the asymptotics of transition probabilities and of the expectations of the transition times for the Markov chain. The limit of the transition probabilities as $\varepsilon \to 0$ is determined by the behavior of the process X_t^ε in an arbitrarily small neighborhood of $\mathcal{L}$. The asymptotics of the expectations of the transition times, on the contrary, is determined by the event that the trajectory of X_t^ε wanders away from the level set $\mathcal{L}$.

In order to study the transition times we use the results of Freidlin and Wentzell [5]. For a given stream function H they introduce a graph and a mapping g from the plane into the graph, such that each connected level curve of H gets mapped into a point on the graph, with the level sets containing critical points mapped into vertices. Then they demonstrate that the process $g(X^\varepsilon_{t/\varepsilon})$ on the graph converges to a limiting Markov process. The asymptotics of the expectations of the transition times for our Markov chain is related to the limiting process on the graph.

Now consider the flows with 'open channels'. Assuming that the channels are directed along the x_1 axis, we prove that

$$D^{11}(\varepsilon) = (d^{11} + o(1))\frac{1}{\varepsilon} \ , \quad \text{and} \quad D^{22}(\varepsilon) = (d^{22} + o(1))\varepsilon \ , \quad \text{as} \ \ \varepsilon \to 0 \ , \tag{3}$$

that is the diffusion across the channels is not qualitatively enhanced, compared with the process $\sqrt{\varepsilon}W_t$. The effective diffusivity in the direction of the flow is enhanced by a factor of order ε^{-2}.

Theorem 1.2. *Assume that an infinitely smooth periodic stream function $H(x_1, x_2)$ defines a flow with open channels, which are directed along the x_1 axis, and that its critical points are nondegenerate. Then the asymptotics of the effective diffusivity for the process (1) is given by (3).*

Remark Since the matrix of effective diffusivity is symmetric and positive definite, from Theorem 1.2 it follows that the off-diagonal terms $D^{12}(\varepsilon) = D^{21}(\varepsilon)$ are bounded uniformly in ε. However we do not make a statement here on their asymptotic behavior.

The proof of Theorem 1.2 is much simpler than that of Theorem 1.1, and is an easy application of the results of Freidlin and Wentzell [5].

The paper is organized as follows. In Section 2 we describe the construction of the discrete time Markov chain associated with a cellular flow, and relate the question of effective diffusivity for the process X^ε_t to the study of transition probabilities and transition times for the Markov chain. In Sections 3 and 4 we study of transition probabilities and transition times respectively. In Section 5 we prove Theorem 1.2. In Section 6 we prove several technical lemmas used in the previous sections.

2 Construction of the Discrete Time Markov Chain

Consider the set $\{x : |H(x)| < \varepsilon^{\alpha_1}\}$ on the torus, with some $\frac{1}{4} < \alpha_1 < \frac{1}{2}$, and let V^ε be the connected component of this set, which contains $\mathcal{L}$. Thus V^ε is a thin tube around $\mathcal{L}$, whose width, however, is much larger than a typical fluctuation of the process $H(X^\varepsilon_t)$ in fixed time (see the picture below).

Let U_i, $i = 1, ..., n$, be the connected components of $\mathbb{T}^2 \setminus \mathcal{L}$, and let A_i, $i = 1, ..., n$, be the saddle points, which belong to $\mathcal{L}$. While the numbers of the connected components and of the saddles are the same for topological reasons, their equality is not used in the proofs, and the numbering of U_i's is not related in any manner to that of A_i's. If there are points, which are carried to A_j by the flow $x' = v(x)$, and to A_i by the flow $x' = -v(x)$, then the set of such points (a subset of $\mathcal{L}$) is denoted by $\gamma(A_i, A_j)$. We assume (for the sake of simplicity of notation only) that $\gamma(A_i, A_i)$ is empty, that is the separatrices do not form "loops". In a neighborhood of each curve $\gamma(A_i, A_j)$ we may consider a smooth change of coordinates $(x_1, x_2) \to (H, \theta)$, where θ is defined by the conditions: $|\nabla\theta| = |\nabla H|$ on $\gamma(A_i, A_j)$, and $\nabla\theta \perp \nabla H$ (this way θ is defined up to multiplication by -1 and up to an additive constant). The same change of coordinates can be considered in $V^\varepsilon \cap \overline{U}_k$. In this case $\theta \in [0, \int_{\partial U_k} |\nabla H| dl]$, with the end points of the interval identified. Thus, if $A_i \in \partial U_k$, we define

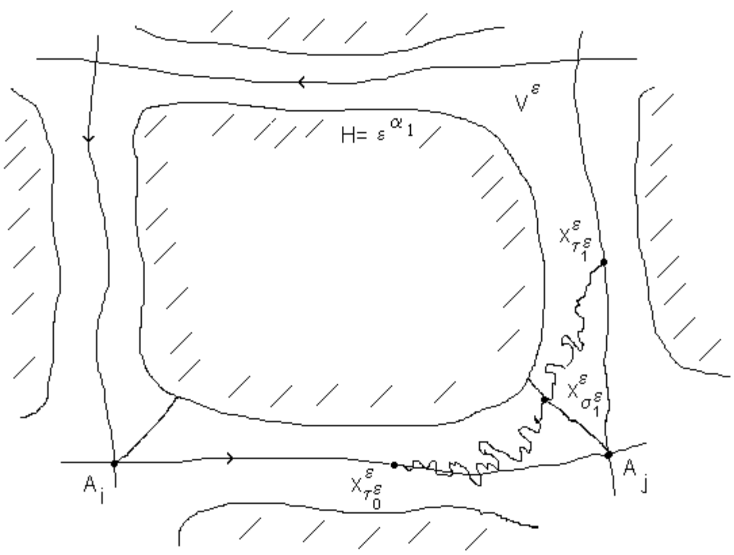
$$B(A_i, U_k) = \{x \in V^\varepsilon \cap \overline{U}_k : \theta(x) = \theta(A_i)\}.$$

Let $B(A_i) = \bigcup_{k: A_i \in \partial U_k} B(A_i, U_k)$.

Define the stopping times $\sigma_0^\varepsilon = 0$, $\tau_0^\varepsilon = \inf\{t : X_t^\varepsilon \in \mathcal{L}\}$. Then $\sigma_n^\varepsilon, \tau_n^\varepsilon$, $n \geq 1$ are defined inductively as follows. Assume that $X^\varepsilon_{\tau^\varepsilon_{n-1}} \in \gamma(A_i, A_j)$, and $i \neq j$. Then

$$\sigma_n^\varepsilon = \inf\{t \geq \tau_{n-1}^\varepsilon : X_t^\varepsilon \in \bigcup_{k \neq i} B(A_k) \bigcup \partial V^\varepsilon\}.$$

Thus, σ_n^ε is the first time after τ_{n-1}^ε that the process either exits V^ε, or goes past a saddle point different from A_i. Define $\tau_n^\varepsilon = \inf\{t \geq \sigma_n^\varepsilon : X_t^\varepsilon \in \mathcal{L}\}$. Let $\mathcal{L}^0 = \mathcal{L} \setminus \{A_i, i = 1, ..., n\}$. Since almost every trajectory of X_t^ε does not contain any of the points A_i, $X^\varepsilon_{\tau^\varepsilon_n}$ is a Markov chain with the state space $\mathcal{L}^0$. The stopping times τ_n^ε are the consecutive times when the process X_t^ε hits the separatrix $\mathcal{L}$ after exiting V^ε or after having passed past a saddle point. Note, that the case, when a point $x \in \gamma(A_i, A_j)$ travels, due to diffusion, against the flow $v(x)$, and returns to $B(A_i)$, does not count as having passed past a saddle point.



It is not difficult to see that $X^\varepsilon_{\tau^\varepsilon_n}$ satisfies the Doeblin condition, with the unique ergodic set. Therefore there exists a unique invariant measure $\mu^\varepsilon(dy)$ on $\mathcal{L}^0$ ([2]). Note that $(X^\varepsilon_{\tau^\varepsilon_n}, \tau^\varepsilon_n - \tau^\varepsilon_{n-1})$ also forms a Markov chain with the state space $\mathcal{L}^0 \times R_+$, which satisfies Doeblin condition for each ε, and has a unique ergodic set . Let $\widetilde{p}^\varepsilon_x(dy, dt)$ be the stochastic transition function for this chain (it only depends on the first component of the original point, as is reflected in the notation). Then $\widetilde{\mu}^\varepsilon(dy, dt) = \int_{\mathcal{L}^0} \widetilde{p}^\varepsilon_x(dy, dt)\mu^\varepsilon(dx)$ is the invariant measure. Since, due to the presence of diffusion, the distributions of the transition times have exponentially decreasing tails, we can apply the law of large numbers to the function $(x, t) \to t$ on $\mathcal{L}^0 \times R_+$ to obtain

$$\lim_{n\to\infty} \frac{\tau^\varepsilon_n}{n} = \int\int_{\mathcal{L}_0 \times R_+} t\widetilde{\mu}^\varepsilon(dy, dt) =$$

$$\int\int\int_{\mathcal{L}^0\times\mathcal{L}^0\times R_+} t\widetilde{p}^\varepsilon_x(dy, dt)\mu^\varepsilon(dx) = \int_{\mathcal{L}^0} \mathbf{E}_x \tau^\varepsilon_1 \mu^\varepsilon(dx) \quad \text{almost surely.} \tag{4}$$

In the arguments, which led to (4), we considered X^ε_t as a process on the torus. In order to keep track of the displacements of $X^\varepsilon_{\tau^\varepsilon_n}$, as the process on $\mathbb{R}^2$, we introduce another Markov chain, on the extended phase space $\mathcal{L}^0 \times \mathbb{Z}^2$. Let now $A_i, i = 1, ...$ be the saddle points of H on the plane. Then any $x_p \in \gamma(A_i, A_j) \subset \mathcal{L}_p$ can be uniquely identified with a pair (x, z), where $x \in \mathcal{L}^0$ and $z = ([A^1_i], [A^2_i]) \in \mathbb{Z}^2$. ($[A^1_i]$ and $[A^2_i]$ are the integer parts of the first and second coordinates of A_i). Thus, we have the mapping $\phi : \mathcal{L}_p \setminus \{A_i, i = 1, ...\} \to \mathcal{L}^0 \times \mathbb{Z}^2$. Let ϕ_1 and ϕ_2 be the components of this mapping. We define the Markov chain Y^ε_n as follows

$$Y^\varepsilon_n = (\phi_1(X^\varepsilon_{\tau^\varepsilon_n}); \phi_2(X^\varepsilon_{\tau^\varepsilon_n}) - \phi_2(X^\varepsilon_{\tau^\varepsilon_{n-1}})) \ .$$

Note that the second component $\phi_2(X^\varepsilon_{\tau^\varepsilon_n}) - \phi_2(X^\varepsilon_{\tau^\varepsilon_{n-1}})$ almost surely takes the values in some finite subset S of $\mathbb{Z}^2$. Thus Y^ε_n is a Markov chain on $\mathcal{L}^0 \times S$, where $S = \{s_1, ..., s_k\} \subset \mathbb{Z}^2$. It is not difficult to see that Y^ε_n satisfies Doeblin condition with the unique ergodic set. Applying the law of large numbers to the vector valued function $f(x, s) = s$, defined on $\mathcal{L}^0 \times S$, we obtain that there exists $m \in \mathbb{R}^2$, such that

$$\lim_{n\to\infty} \frac{X^\varepsilon_{\tau^\varepsilon_n}}{n} = \lim_{n\to\infty} \frac{\sum_{i=1}^n f(Y^\varepsilon_i)}{n} = m \quad \text{almost surely .}$$

Since $\frac{X^\varepsilon_t}{\sqrt{t}} \to N(0, D(\varepsilon))$ in distribution, and due to (4), we conclude that $m = 0$. Applying the central limit theorem to the same function $f(x, s)$, we obtain that there exists a matrix d^ε, such that

$$\lim_{n\to\infty} \frac{X^\varepsilon_{\tau^\varepsilon_n}}{\sqrt{n}} = \lim_{n\to\infty} \frac{\sum_{i=1}^n f(Y^\varepsilon_i)}{\sqrt{n}} = N(0, d^\varepsilon) \quad \text{in distribution .}$$

Due to (4), we have

$$D(\varepsilon) = d^\varepsilon / \int_{\mathcal{L}^0} \mathbf{E}_x \tau^\varepsilon_1 d\mu^\varepsilon(x) \ . \tag{5}$$

In sections 3 and 4 we shall obtain the asymptotics of the stochastic transition functions for the chain Y_n^ε, and of the functions $\mathbf{E}_x\tau_1^\varepsilon$. Assuming that this is accomplished, the next lemma will allow us to obtain the asymptotics of the effective diffusivity, using (5).

First we introduce the notations and formulate the assumptions needed for the lemma. Let M be a locally compact separable metric space. Let $C_b(M)$ be the set of bounded continuous functions on M. Let $p_\varepsilon(x,dy)$, $0 \le \varepsilon \le \varepsilon^0$ be a family of stochastic transition functions on M. Assume that

(A) The family of measures $p_0(x,dy)$, $x \in K$ is tight for any compact set K.

(B) $p_0(x,dy)$ is weakly Feller, that is $\int f(y)p_0(x,dy) \in C_b(M)$ if $f \in C_b(M)$.

(C) For any $f \in C_b(M)$ and any compact $K \subset M$,

$$\lim_{\varepsilon\to 0}\int f(y)p_\varepsilon(x,dy) = \int f(y)p_0(x,dy) \quad \text{uniformly in } x \in K \ .$$

(D) There exist unique invariant measures $\mu^\varepsilon(dy)$. There exist $\lambda > 0, c > 0$, such that

$$|p_\varepsilon^n(x,A) - \mu^\varepsilon(A)| \le ce^{-\lambda n} \quad \text{for all } x, A, \varepsilon \ .$$

(That is p_ε are uniformly exponentially mixing).

Let $g \in C_b(M,\mathbb{R}^2)$ be such that $\int g d\mu^\varepsilon = 0$ for all ε. Let Y_n^ε be the stationary Markov chain, with the stochastic transition function p_ε. Since Y_i^ε is exponentially mixing, the central limit theorem can be applied to $g(Y_i^\varepsilon)$, and thus $\frac{\sum_{i=1}^n g(Y_i^\varepsilon)}{\sqrt{n}}$ converges weakly as $n\to\infty$ to a mean-zero Gaussian distribution. We denote the covariance matrix of the limiting distribution by $d^\varepsilon(g)$.

Lemma 2.1. *Suppose that assumptions (A)-(D) hold. Then*
(a) $\mu^\varepsilon \to \mu^0$ weakly.
(b) If $f^\varepsilon(x)$ *are uniformly bounded,* $f^0(x) \in C_b(M)$, *and* $\lim_{\varepsilon\to 0} f^\varepsilon(x) = f^0(x)$ *uniformly on any compact, then*

$$\lim_{\varepsilon\to 0}\int f^\varepsilon d\mu^\varepsilon = \int f^0 d\mu^0 \ .$$

(c) If $g \in C_b(M,\mathbb{R}^2)$ *is such that* $\int g d\mu^\varepsilon = 0$ *for all* ε, *then*

$$d^\varepsilon(g) \to d^0(g) \ .$$

Proof: From (A) it follows that for each n the family of measures $p_0^n(x,dy), x \in K$ is tight for any compact set K. Let us assume that for a certain n, for any $f \in C_b(M)$, uniformly on any compact set K we have

$$\int f(y)p_\varepsilon^n(x,dy) - \int f(y)p_0^n(x,dy) \to 0 \ \text{ as } \varepsilon \to 0 \ . \tag{6}$$

Note that this is true for $n = 1$ by (C). Combining (6) and the fact that $p_0^n(x,dy), x \in K$ is tight we obtain that for any compact set K and for any $\delta > 0$ there is a compact set K_1 such that

$$p_\varepsilon^n(x,K_1) > 1-\delta, \ x \in K \tag{7}$$

for sufficiently small ε. Next we justify (6) for $n+1$ instead of n.

$$\int f(y)(p_\varepsilon^{n+1}(x,dy) - p_0^{n+1}(x,dy)) =$$

$$\int f_1(y)(p_\varepsilon^n(x,dy) - p_0^n(x,dy)) + \int\int f(z)(p_\varepsilon(y,dz) - p_0(y,dz))p_\varepsilon^n(x,dy) \ ,$$

where $f_1(x) = \int f(y)p_0(x,dy) \in C_b(M)$. The first term on the right hand side tends to zero by (6), while the second term tends to zero by (7) as $\int f(z)(p_\varepsilon(y,dz) - p_0(y,dz))$ is bounded and tends to zero uniformly on any compact. We therefore have established (6) for all n.

Let us prove part (a) of the lemma. Fix an arbitrary $x \in M$. Then for $f \in C_b(M)$ we have

$$\int f(y)d\mu^\varepsilon(y) = \lim_{n\to\infty}\int f(y)p_\varepsilon^n(x,dy) \ ,$$

and the limit is uniform in ε by (D). The weak convergence of μ^ε to μ^0 now follows from (6).

To prove part (b) we write

$$\int f^\varepsilon d\mu^\varepsilon - \int f^0 d\mu^0 = (\int f^0 d\mu^\varepsilon - \int f^0 d\mu^0) + (\int f^\varepsilon d\mu^\varepsilon - \int f^0 d\mu^\varepsilon) \ .$$

The difference of the first two terms on the right hand side tends to zero as $\mu^\varepsilon \to \mu^0$ weakly. The difference of the last two terms tends to zero since for any $\delta > 0$ there is a compact set K for which $\mu^\varepsilon(K) > 1-\delta$ for sufficiently small ε (since $\mu^\varepsilon \to \mu^0$ weakly), and $f^\varepsilon \to f^0$ on any compact set.

In order to prove part (c) of the lemma it is sufficient to consider $g \in C_b(M,\mathbb{R})$ (scalar valued). In this case

$$d^\varepsilon(g) = \mathbf{E}[(g(Y_0^\varepsilon))^2 + 2g(Y_0^\varepsilon)g(Y_1^\varepsilon) + 2g(Y_0^\varepsilon)g(Y_2^\varepsilon) + ...] \ .$$

Due to uniform mixing (D)

$$|\mathbf{E}(g(Y_0^\varepsilon)g(Y_n^\varepsilon))| \le e^{-\gamma n} \ ,$$

where γ does not depend on ε. In order to prove that $d^\varepsilon(g) \to d^0(g)$ we therefore only need to establish that

$$\mathbf{E}[g(Y_0^\varepsilon)g(Y_n^\varepsilon) - g(Y_0^0)g(Y_n^0)] \to 0 \tag{8}$$

for any fixed n. The left hand side of (8) can be written as

$$\int g(x)g(y)p_\varepsilon^n(x,dy)d\mu^\varepsilon(x) - \int g(x)g(y)p_0^n(x,dy)d\mu^0(x) \ .$$

Let $G_\varepsilon(x) = g(x)\int g(y)p_\varepsilon^n(x,dy)$. Then $G_\varepsilon(x) \to G_0(x)$ uniformly on any compact by (6), and the conclusion follows by part (b). □

3 The Limit of the Transition Probabilities

In this section we shall identify the limit of the transition probabilities for the chains Y_n^ε on $\mathcal{L}^0 \times S$ and verify the conditions (A)-(D) for these chains.

Recall the (H, θ) coordinates which we may consider inside each cell U_k near its boundary, that is in $U_k \cap V^\varepsilon$. Let

$$h = \varepsilon^{-\frac{1}{2}} H \ .$$

In order to find the limit of the transition probabilities we shall demonstrate that in a small neighborhood of $\mathcal{L}_p$ after a random change of time the process X_t^ε is well approximated by the process X_t with the generator $\frac{1}{2}\partial_{hh} + \partial_\theta$ in (h, θ) coordinates.

Let $x \in \gamma(A_i, A_j)$ be a point on $\mathcal{L}_p$. We introduce the stochastic transition function $p_0(x, dy), x, y \in \mathcal{L}_p$ as follows:

Let $x, y \in \partial U_k$ (if x and y do not belong to the boundary of the same cell, then $p_0(x, dy) = 0$). Consider the (h, θ) coordinates in $U_k \cap V^\varepsilon$, so that $\theta(x) = 0$, and θ increases in the direction of the flow. Since ∂U_k is a closed contour, points with coordinates (h, θ) and $(h, \theta + \int_{\partial U_k} |\nabla H| dl)$ are identified. Let $\theta(A_j)$ and $\theta(y)$ belong to $(0, \int_{\partial U_k} |\nabla H| dl]$, and consider the process X_t with the generator $\partial_\theta + \frac{1}{2}\partial_{hh}$ in (h, θ) coordinates, which starts at the origin (corresponding to the point x). Let τ be the time of the first exit from the following domain: $\mathcal{D}^0 = \{\theta < \theta(A_j)\} \cup \{\theta \geq \theta(A_j); h > 0\}$. Then define

$$p_0(x, dy) = \sum_{k: x, y \in \partial U_k} \sum_{m \geq 0} \mathrm{Prob}_x\{\theta(X_\tau) \in [\theta(y) + m \int_{\partial U_k} |\nabla H| dl, \theta(y + dy) + m \int_{\partial U_k} |\nabla H| dl]\}. \quad (9)$$

The summation over k is needed to account for the fact that x and y may both belong to the same edge $\gamma(A_i, A_j)$, in which case they both belong to the boundaries of two cells, and we need to consider two sets of (h, θ) coordinates.

The function $p_0(x, dy)$ is a stochastic transition function on $\mathcal{L}_p$, and it can be considered as a stochastic transition function on $\mathcal{L}^0 \times S$. It clearly satisfies conditions (A) and (B) preceding Lemma 2.1.

For $x, y \in \mathcal{L}_p$, and for the stopping times τ_n^ε defined in the previous section, let $p_\varepsilon(x, dy)$ be the transition function for the chain $X_{\tau_n^\varepsilon}^\varepsilon$ considered on $\mathcal{L}_p$:

$$p_\varepsilon(x, dy) = \mathrm{Prob}_x\{X_{\tau_1^\varepsilon}^\varepsilon \in [y, y + dy]\} \ .$$

Note that this definition is similar to that of $p_0(x, dy)$. Here, however, we do not use the (h, θ) coordinates since with small probability the process X_t^ε may travel outside of the domains U_k for which $x \in \gamma(A_i, A_j) \subseteq \partial U_k$ before time τ_1^ε (due to the presence of the small diffusion term X_t^ε may go 'past' the saddle point A_i, thus traveling to one of the neighboring domains before time τ_1^ε).

Lemma 3.1. *For any closed interval $I \subset \gamma(A_i, A_j)$, and any bounded continuous function f on $\mathcal{L}_p$,*

$$\lim_{\varepsilon \to 0} \int f(y) p_\varepsilon(x, dy) = \int f(y) p_0(x, dy) \quad \text{uniformly in } x \in I \ .$$

Notice that Lemma 3.1 implies the condition (C) for the chain Y_n^ε. Before we start the proof of Lemma 3.1 we state and prove the following preliminary lemma.

Lemma 3.2. *Let X_t^1 and X_t^2 be the following two diffusion processes on $\mathbb{R}^d$ with infinitely smooth coefficients:*

$$dX_t^1 = v(X_t^1)dt + a(X_t^1)dW_t + \varepsilon^2 v_1(X_t^1)dt + \varepsilon a_1(X_t^1)dW_t \ ,$$

$$dX_t^2 = v(X_t^2)dt + a(X_t^2)dW_t + \varepsilon^2 v_2(X_t^2)dt + \varepsilon a_2(X_t^2)dW_t \ ,$$

with $X_0^1 = X_0^2$. Suppose that for a certain constant L the following bound on the coefficients holds:

$$|\nabla v^i|, |\nabla a^{ij}|, |v_1^i|, |v_2^i|, |a_1^{ij}|, |a_2^{ij}| \leq L \ , \quad i,j = 1, ..., d \ ,$$

where i and j stand for the vector (matrix) entries of the coefficients. Let λ be the initial distribution for the processes X_0^1 and X_0^2. Then for some constant $K = K(L)$ and for any $t, \eta > 0$ we have

$$\text{Prob}_\lambda\{\sup_{0\leq s\leq t} |X_s^1 - X_s^2| \geq \eta\} \leq \frac{(e^{Kt}-1)\varepsilon^2}{\eta^2} \ .$$

Proof: Let us assume that $d = 1$ in order to avoid vector and matrix indices. By Ito's formula, for any stopping time $\tau \leq t$,

$$\mathbf{E}_\lambda |X_\tau^1 - X_\tau^2|^2 = \mathbf{E}_\lambda \int_0^\tau 2(X_s^1 - X_s^2)(v(X_s^1) - v(X_s^2) + \varepsilon^2[v_1(X_s^1) - v_2(X_s^2)])ds +$$

$$\mathbf{E}_\lambda \int_0^\tau (a(X_s^1) - a(X_s^2) + \varepsilon[a_1(X_s^1) - a_2(X_s^2)])^2 ds \ . \tag{10}$$

From the estimates on the coefficients and their derivatives it follows that the expression in (10) can be estimated as follows

$$\mathbf{E}_\lambda |X_\tau^1 - X_\tau^2|^2 \leq K(L)(\int_0^\tau \mathbf{E}_\lambda |X_s^1 - X_s^2|^2 ds + \varepsilon^2 t). \tag{11}$$

In particular, for $\tau = t$ we have

$$\mathbf{E}_\lambda |X_t^1 - X_t^2|^2 \leq K(L)(\int_0^t \mathbf{E}_\lambda |X_s^1 - X_s^2|^2 ds + \varepsilon^2 t).$$

Let $R(t) = \mathbf{E}_\lambda |X_t^1 - X_t^2|^2 + \varepsilon^2$. Then $R(t) \leq K(L) \int_0^t R(s)ds$, $R(0) = \varepsilon^2$. By Gronwall's Lemma applied to R(t) we have

$$\mathbf{E}_\lambda |X_t^1 - X_t^2|^2 \leq \varepsilon^2 (e^{K(L)t} - 1) \ .$$

Define the stopping time $\tau = \min\{s : |X_s^1 - X_s^2| \geq \eta\} \wedge t$. Then, by (11)

$$\eta^2 \text{Prob}_\lambda\{\max_{0\leq s\leq t} ||X_s^1 - X_s^2| \geq \eta\} \leq \mathbf{E}_\lambda |X_\tau^1 - X_\tau^2|^2 \leq$$

$$K(L)(\mathbf{E}_\lambda \int_0^\tau |X_s^1 - X_s^2|^2 ds + \varepsilon^2 t) \leq \varepsilon^2 (e^{K(L)t} - 1) ,$$

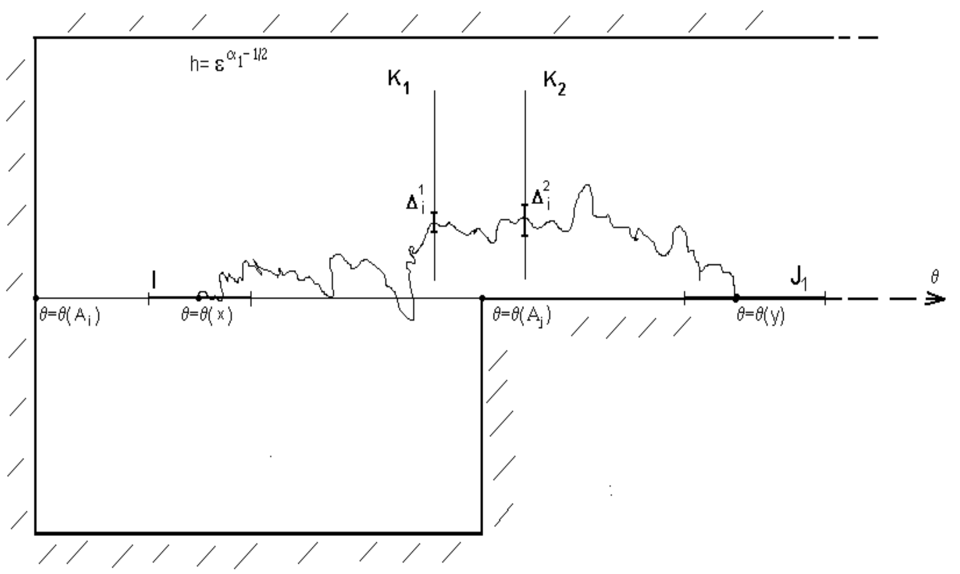
which yields the lemma. □

Proof of Lemma 3.1: Since the kernel $p_0(x, dy)$ is smooth in both variables and satisfies condition (A) preceding Lemma 2.1, to prove the uniform weak convergence stated in Lemma 3.1 it is sufficient to demonstrate that for an arbitrary closed interval $J \subset \gamma(A_j, A_l)$ and an arbitrary $\delta > 0$ there is $\varepsilon_0 > 0$ such that

$$p_\varepsilon(x, J) > p_0(x, J) - \delta \quad \text{for all} \quad x \in I, \quad \varepsilon < \varepsilon_0 . \tag{12}$$

Suppose that $I, J \subset \partial U_k$. (If I and J don't belong to the boundary of the same cell, then $p_0(x, J)$ is equal to zero.) For the sake of simplicity of notation let us assume that $I \subset \gamma(A_i, A_j)$ and $J \subset \gamma(A_j, A_l)$, that is I and J belong to the adjacent edges of $\mathcal{L}_p$. Without loss of generality we can assume that $h > 0$ in $U_k \cap V^\varepsilon$ for sufficiently small ε. We can consider the process X_t^ε in (h, θ) coordinates in the following domain (see the picture below)

$$\mathcal{D}^\varepsilon = \{\theta > \theta(A_i)\} \bigcap \{|h| < \varepsilon^{\alpha_1 - \frac{1}{2}}\} \bigcap \{\theta < \theta(A_j)\} \bigcup \{\theta \geq \theta(A_j), 0 < h < \varepsilon^{\alpha_1 - \frac{1}{2}}\} .$$



As above we consider the process X_t in (h, θ) coordinates in the domain

$$\mathcal{D}^0 = \{\theta < \theta(A_j)\} \bigcup \{\theta \geq \theta(A_j); h > 0\} .$$

Note that $p_\varepsilon(x, J)$ is estimated from below by the probability that X_t^ε leaves $\mathcal{D}^\varepsilon$ through any of the copies of J (which corresponds to X_t^ε making a finite number of rotations inside

$U_k \cap V^\varepsilon$, and then leaving U_k through the segment J). Let $J_0, J_1, ...$ be the copies of J in (h,θ) coordinates (J_{m+1} can be obtained from J_m by a shift by $\int_{\partial U_k} |\nabla H| dl$ along the θ axis). For an initial point $x \in \mathcal{D}^\varepsilon$ let $\widetilde{p}_\varepsilon(x, J_m)$ be the probability that X_t^ε leaves the domain $\mathcal{D}^\varepsilon$ through J_m. As stated above, for $x \in I$ we have

$$p_\varepsilon(x, J) \geq \sum_{m=0}^{\infty} \widetilde{p}_\varepsilon(x, J_m) \ . \tag{13}$$

Similarly, for $x \in \mathcal{D}^0$ let $\widetilde{p}_0(x, J_m)$ be the probability that X_t leaves the domain $\mathcal{D}^0$ through J_m. If $x \in I$ and J belong to different edges of ∂U_k (as we have assumed) then by (9) we have

$$p_0(x, J) = \sum_{m=0}^{\infty} \widetilde{p}_0(x, J_m) \ . \tag{14}$$

Since all the terms in the sums (13) and (14) are non-negative, in order to prove (12) it is sufficient to demonstrate that for each m and each $\delta > 0$ there is $\varepsilon_0 > 0$ such that

$$\widetilde{p}_\varepsilon(x, J_m) > \widetilde{p}_0(x, J_m) - \delta \ \text{ for all } \ x \in I, \quad \varepsilon < \varepsilon_0 \ . \tag{15}$$

For the sake of simplicity of notation we shall only prove (15) for $m = 0$.

The generator of the process X_t^ε is

$$L^\varepsilon f = \frac{\varepsilon}{2} \Delta f + v \nabla f \ ,$$

which in (h,θ) coordinates becomes

$$L^\varepsilon f = \frac{1}{2}(f''_{hh}|\nabla H|^2 + \varepsilon f''_{\theta\theta}|\nabla\theta|^2 + \sqrt{\varepsilon} f'_h \Delta H + \varepsilon f'_\theta \Delta\theta) + f'_\theta |\nabla H||\nabla\theta| \ .$$

Dividing all of the coefficients of the generator by the same function $|\nabla H||\nabla\theta|$ amounts to a random time change for the process X_t^ε, which does not affect any of the transition probabilities. We shall denote the time-changed process with the generator $\widetilde{L}^\varepsilon f = \frac{L^\varepsilon f}{|\nabla H||\nabla\theta|}$ also by X_t^ε. This process satisfies the equation

$$dX_t^\varepsilon = (1,0)\sqrt{\frac{|\nabla H|}{|\nabla\theta|}} dW_t^h + (0,1)\sqrt{\varepsilon}\sqrt{\frac{|\nabla\theta|}{|\nabla H|}} dW_t^\theta + (\frac{\sqrt{\varepsilon}}{2}\frac{\Delta H}{|\nabla\theta||\nabla H|}, \frac{\varepsilon}{2}\frac{\Delta\theta}{|\nabla\theta||\nabla H|} + 1)dt \ , \tag{16}$$

while

$$dX_t = (1,0)dW_t^h + (0,1)dt \ , \tag{17}$$

where W_t^h and W_t^θ are one dimensional Wiener processes in h and θ variables respectively. We can not apply Lemma 3.2 to (16) and (17) directly, as some of the coefficients on the right hand side of (16) may be unbounded near the saddle point ($h = 0, \theta = \theta(A_j)$). To circumvent this problem we shall take a sequence of steps (justified below), which will single out a small neighborhood of the saddle point, where yet another coordinate system will be considered.

Step 1. Let us take $\delta' > 0$ small enough, so that there exist $0 < h_1 < h_2$, such that for any $\theta_1 \in [\theta(A_j) - \delta', \theta(A_j)]$ the probability of the event that X_t passes through the interval $K = \{h_1 \le h \le h_2, \theta = \theta_1\}$ before leaving the domain $\mathcal{D}^0$ through J_0 differs from $\widetilde{p}_0(x, J_0)$ by less than $\frac{\delta}{10}$ for any $x \in I$. Due to the smoothness of the transition kernel of the process X_t, for some $\alpha < 1$ the interval K can be replaced by any set K' as long as K' is contained in K and has Lebesgue measure at least $(h_2 - h_1)\alpha$.

Step 2. If necessary make δ' from Step 1 smaller, so that $|\widetilde{p}_0(A, J_0) - \widetilde{p}_0(B, J_0)| < \frac{\delta}{10}$ whenever $|h(A) - h(B)| < \delta'$ and $A, B \in [h_1, h_2] \times [\theta(A_j) - \delta', \theta(A_j) + \delta']$.

Step 3. Take $\delta'' \le \delta'$ and let $K_1 = \{h_1 \le h \le h_2, \theta = \theta(A_j) - \delta''\}$, and $K_2 = \{h_1 \le h \le h_2, \theta = \theta(A_j) + \delta''\}$. Take δ'' sufficiently small so that whenever $A \in K_1$ the process X_t^ε starting at A passes through the set $\{h(A) - \frac{(1-\alpha)\delta'}{2} < h < h(A) + \frac{(1-\alpha)\delta'}{2}, \theta = \theta(A_j) + \delta''\}$ before leaving $\mathcal{D}^\varepsilon$ with probability at least $1 - \frac{\delta}{10}$ for small enough ε.

Step 4. Let us split the set K_2 into intervals Δ_i^2 , $i = 1, ..., r$ of length δ' (we can assume that each of the intervals included the endpoints). Construct on K_1 the intervals Δ_i^1 , $i = 1, ..., r$ of length $\alpha\delta'$, such that $h(\text{center}(\Delta_i^1)) = h(\text{center}(\Delta_i^2))$. Let τ_{K_1} be the first time when a process either exits $\mathcal{D}^\varepsilon(\mathcal{D}^0)$ or reaches K_1. Let us take ε_0 so small that

$$\text{Prob}_x\{X^\varepsilon_{\tau_{K_1}} \in \Delta_i^1\} \ge \text{Prob}_x\{X_{\tau_{K_1}} \in \Delta_i^1\} - \frac{\delta}{10r} \quad \text{for all} \quad x \in I, \quad \varepsilon < \varepsilon_0 \ .$$

Step 5. Let us take ε_0 so small that

$$\widetilde{p}_\varepsilon(A, J_0) > \widetilde{p}_0(A, J_0) - \frac{\delta}{10} \quad \text{for all} \quad A \in K_2, \quad \varepsilon < \varepsilon_0 \ .$$

Assuming that the Steps 1 - 5 are valid let us prove (15). By the Markov property

$$\widetilde{p}_\varepsilon(x, J_0) \ge \sum_{i=1}^r \text{Prob}_x\{X^\varepsilon_{\tau_{K_1}} \in \Delta_i^1\} \min_{A \in \Delta_i^1} \widetilde{p}_\varepsilon(A, J_0) \ .$$

By Steps 3, 5, and 2, the second factor on the right hand side can be estimated as follows:

$$\min_{A \in \Delta_i^1} \widetilde{p}_\varepsilon(A, J_0) \ge \min_{A \in \Delta_i^2} \widetilde{p}_\varepsilon(A, J_0) - \frac{\delta}{10} \ge \min_{A \in \Delta_i^2} \widetilde{p}_0(A, J_0) - \frac{2\delta}{10} \ge \max_{A \in \Delta_i^1} \widetilde{p}_0(A, J_0) - \frac{3\delta}{10} \ ,$$

while by Step 4

$$\text{Prob}_x\{X^\varepsilon_{\tau_{K_1}} \in \Delta_i^1\} \ge \text{Prob}_x\{X_{\tau_{K_1}} \in \Delta_i^1\} - \frac{\delta}{10r} \ .$$

Combining the above inequalities and using Step 1 we obtain

$$\widetilde{p}_\varepsilon(x, J_0) \ge \sum_{i=1}^r (\text{Prob}_x\{X_{\tau_{K_1}} \in \Delta_i^1\} - \frac{\delta}{10r})(\max_{A \in \Delta_i^1} \widetilde{p}_0(A, J_0) - \frac{3\delta}{10}) \ge$$

$$\sum_{i=1}^{r} \text{Prob}_x\{X_{\tau_{K_1}} \in \Delta_i^1\} \max_{A\in\Delta_i^1} \widetilde{p}_0(A, J_0) - \frac{4\delta}{10} \geq \widetilde{p}(x, J_0) - \frac{5\delta}{10} ,$$

which implies (15).

It remains to justify the construction in Steps 1-5. The validity of Steps 1 and 2 follows from the fact that the transition kernel of the process X_t is smooth. To justify Steps 4 and 5 it is sufficient to consider both processes X_t^ε and X_t in a compliment to a neighborhood of the saddle point, where Lemma 3.2 applies.

In order to justify Step 3 we note that by Morse Lemma in a neighborhood O_j of the saddle point A_j there is a smooth change of variables, such that in the new variables the stream function is $H(x_1, x_2) = x_1 x_2$, and the interior of U_k corresponds to the first quadrant $x_1, x_2 > 0$. In the new variables the generator of the process X_t^ε, after a random change of time, becomes $L^\varepsilon f = \varepsilon L_1 f + v_1 \nabla f$, where L_1 is a differential operator with first and second order terms, with bounded coefficients, and $v_1(x_1, x_2) = (-x_1, x_2)$. We shall consider the operator L^ε in the domain $\widetilde{\mathcal{D}}^\varepsilon = O_j \bigcap \{x_1 > 0; x_2 > 0; x_1 + x_2 > \varepsilon^{\frac{2}{3}}; x_1 x_2 < \varepsilon^{\frac{1}{3}}\}$. Make a further change of variables in $\widetilde{\mathcal{D}}^\varepsilon$:

$$(x_1, x_2) \to (u, v) = (\frac{x_1 x_2}{\sqrt{\varepsilon}}, x_2 - x_1) .$$

In the new variables, after dividing all the coefficients of the operator by $(x_1 + x_2)$, which amounts to a random change of time for the process, the operator can be written as

$$L^\varepsilon f = M^\varepsilon f + \frac{\partial f}{\partial v} , \tag{18}$$

where $M^\varepsilon f$ is a differential operator with first and second order terms. All the coefficients of M^ε can be made arbitrarily small in $\widetilde{\mathcal{D}}^\varepsilon$ by selecting a sufficiently small neighborhood O_j of the point A_j, and then taking ε to be sufficiently small.

The construction in Step 3 now follows from Lemma 3.2 by comparing the process whose generator is the operator (18) with the deterministic process with generator $\frac{\partial f}{\partial v}$.

Remark To verify condition (D) (uniform mixing) preceding Lemma 2.1 for the chain Y_n^ε it is sufficient to show (see [2], page 197) that there is an integer $n \geq 1$, an interval $I \subset \mathcal{L}^0 \times S$, and a constant $c > 0$, such that

$$p_\varepsilon^n(x, dy) \geq c\lambda(dy), \quad \text{and} \quad p_0^n(x, dy) \geq c\lambda(dy), \quad \text{for} \quad x \in \mathcal{L}^0 \times S, \ y \in I , \tag{19}$$

where $\lambda(dy)$ is the Lebesgue measure on I. The proof of estimate (19) is absolutely similar to that of Lemma 3.1.

4 The Asymptotics of the Transition Times

In this section we shall study the asymptotics of the integral $\int_{\mathcal{L}^0} \mathbf{E}_x \tau_1^\varepsilon d\mu^\varepsilon(x)$, which enters in the expression (5) for the effective diffusivity.

We shall demonstrate the following:

$$\mathbf{E}_x\tau_1^\varepsilon \le c\varepsilon^{-\frac{1}{2}} \ \text{ for all } \ x \in \mathcal{L}^0, \tag{20}$$

$$\lim_{\varepsilon\to 0} \varepsilon^{\frac{1}{2}}\mathbf{E}_x\tau_1^\varepsilon = f^0(x) \ \text{ uniformly in } \ x \in I, \tag{21}$$

where $f^0(x) \in C_b(\mathcal{L}^0)$ is a positive function and I is an arbitrary closed interval $I \subset \gamma(A_i, A_j)$. From parts (a) and (b) of Lemma 2.1 it then follows that

$$\int_{\mathcal{L}^0} \mathbf{E}_x\tau_1^\varepsilon d\mu^\varepsilon(x) = \varepsilon^{-\frac{1}{2}}\left(\int_{\mathcal{L}^0} f^0(x)d\mu^0(x) + o(1)\right) \ \text{ as } \ \varepsilon \to 0 \ , \tag{22}$$

where $\mu^0(x)$ is the invariant measure on $\mathcal{L}^0$ for the kernel $p_0(x, dy)$, defined in Section 3.

The proof of formulas (20) and (21) will rely on a sequence of lemmas stated below. We shall study separately the probability of the event that the process X_t^ε starting form $x \in \gamma(A_i, A_j)$ reaches ∂V^ε before time τ_1^ε, and the expectation of the time it takes for the process starting from ∂V^ε to reach $\mathcal{L}^0$.

Consider the process X_t^ε together with the process X_t, whose generator in (h, θ) coordinates is $\frac{1}{2}\partial_{hh} + \partial_\theta$ in the domain $\mathcal{D}_1^\varepsilon = \{\theta(A_i) < \theta < \theta(A_j); |h| < \varepsilon^{\alpha_1 - \frac{1}{2}}\}$. We follow the process X_t till it exits $\mathcal{D}_1^\varepsilon$. Let $P_0(x, dh)$ be the corresponding transition kernel. Thus $P_0(x, dh)$ coincides with a Gaussian distribution on $-\varepsilon^{\alpha_1 - \frac{1}{2}} < h < \varepsilon^{\alpha_1 - \frac{1}{2}}$, and has two point masses at $h = \pm\varepsilon^{\alpha_1 - \frac{1}{2}}$. Similarly let $P_\varepsilon(x, dh)$ be the transition kernel for the process X_t^ε, which starts at $x \in \gamma(A_i, A_j)$ and is stopped at the time σ_1^ε. We have the following:

Lemma 4.1. *For any continuous function $f : \mathbb{R} \to \mathbb{R}$, such that $|f(h)| \le 1 + |h|$, there exists $c > 0$, such that*

$$\int |f(h)|P_\varepsilon(x, dh) < c \ \text{ for } \ x \in \gamma(A_i, A_j) \ . \tag{23}$$

Furthermore, for any closed interval $I \subset \gamma(A_i, A_j)$,

$$\lim_{\varepsilon\to 0} \int f(h)(P_\varepsilon(x, dh) - P_0(x, dh)) = 0 \ , \ \text{ uniformly in } x \in I \ . \tag{24}$$

The proof of Lemma 4.1 is completely similar to that of Lemma 3.1.

We introduce the following notation: τ^{V^ε} is the first time the process X_t^ε leaves V^ε; similarly, τ^{U_k} and $\tau^{V^\varepsilon \cap U_k}$ are the first instances when X_t^ε leaves U_k and $V^\varepsilon \cap U_k$ respectively.

In order to estimate the probability that the process X_t^ε starting from $x \in \gamma(A_i, A_j)$ reaches ∂V^ε before time τ_1^ε we shall need the following

Lemma 4.2. *[5] There exists a constant $c > 0$, such that*

$$\mathbf{E}_x\tau^{V^\varepsilon} \le c\varepsilon^{2\alpha_1 - 1}|\ln\varepsilon| \ \text{ for any } \ x \in V^\varepsilon.$$

This Lemma is the same as Lemma 4.7 of [5] (it must be observed that the proof of Lemma 4.7 of [5] goes through for any $A_{24} < \frac{1}{2}$).

In the event that $X^\varepsilon_{\sigma^\varepsilon_1} \in V^\varepsilon \cap U_k$, after the stopping time σ^ε_1 the process X^ε_t may exit $V^\varepsilon \cap U_k$ either through ∂V^ε or through $\mathcal{L}^0$. The next lemma estimates the probability that the process exits the domain through ∂V^ε.

Lemma 4.3. *There exists $c > 0$, such that for any $x \in V^\varepsilon \cap U_k$*

$$|\mathrm{Prob}_x\{X^\varepsilon_{\tau^{V^\varepsilon \cap U_k}} \in \partial V^\varepsilon\} - h(x)\varepsilon^{\frac{1}{2}-\alpha_1}| \leq c\varepsilon^{\alpha_1}|\ln\varepsilon| \ . \tag{25}$$

Proof: Let L^ε be the generator of the process X^ε_t in the domain $V^\varepsilon \cap U_k$. Then the probability in (25) is equal to the solution $u(x)$ of the equation $L^\varepsilon u = 0$ in $V^\varepsilon \cap U_k$ with the boundary conditions $u|_{H=\varepsilon^{\alpha_1}} = 1$, $u|_{H=0} = 0$. Let $u_1(x) = u(x) - \frac{H(x)}{\varepsilon^{\alpha_1}}$. Then u_1 is the solution of the equation $L^\varepsilon u_1 = -L^\varepsilon \frac{H(x)}{\varepsilon^{\alpha_1}}$ with the boundary conditions $u_1|_{\partial(V^\varepsilon \cap U_k)} = 0$. By Lemma 4.2, since $\varepsilon^{-1}L^\varepsilon H(x)$ is bounded uniformly in ε, the solution u_1 is estimated as follows:

$$|u_1| \leq c_0\varepsilon^{1-\alpha_1}\mathbf{E}_x\tau^{V^\varepsilon \cap U_k} \leq c_1\varepsilon^{\alpha_1}|\ln\varepsilon| \ .$$

This implies the statement of the lemma. □

Using the Markov property of the process X^ε_t with respect to the stopping time σ^ε_1, we obtain that there is $c > 0$, such that for any $x \in \mathcal{L}^0$ we have the following:

$$\mathrm{Prob}_x\{\tau^{V^\varepsilon} < \tau^\varepsilon_1\} \leq \int_{-\infty}^{\infty} \sup_{\overline{x} \in V^\varepsilon : h(\overline{x}) = \overline{h}} \mathrm{Prob}_{\overline{x}}\{X^\varepsilon_{\tau^{V^\varepsilon \cap U(\overline{x})}} \in \partial V^\varepsilon\} P_\varepsilon(x, d\overline{h}) \leq c\varepsilon^{\frac{1}{2}-\alpha_1} \ , \tag{26}$$

where $U(\overline{x})$ is the domain which contains $\overline{x}$ (one of the domains U_k), and the second inequality is due to Lemma 4.3 and (23). Furthermore, due to (24) we can evaluate the asymptotics of the event $\{\tau^{V^\varepsilon} < \tau^\varepsilon_1;\ X^\varepsilon_{\tau^{V^\varepsilon}} \in \partial V^\varepsilon \cap U_k\}$ as follows:

$$\lim_{\varepsilon \to 0} \mathrm{Prob}_x\{\tau^{V^\varepsilon} < \tau^\varepsilon_1;\ X^\varepsilon_{\tau^{V^\varepsilon}} \in \partial V^\varepsilon \cap U_k\}\varepsilon^{\alpha_1 - \frac{1}{2}} = \lim_{\varepsilon \to 0}\int_0^\infty \overline{h}P_0(x, d\overline{h}) \ , \ \text{ uniformly in } x \in I. \tag{27}$$

The next lemma allows us to estimate the expectation of the time it takes for the process starting at ∂V^ε to return to $\mathcal{L}^0$.

Lemma 4.4. *For each of the domains U_k there exists a constant $c_k > 0$, such that*

$$\lim_{\varepsilon \to 0} \varepsilon^{1-\alpha_1}\mathbf{E}_x\tau^{U_k} = c_k \ \text{ uniformly in } \ x \in \partial V^\varepsilon \cap U_k. \tag{28}$$

Note that Lemma 4.4, together with (26) and (27) implies (20) and (21) since the expectation of the time it takes for the process to reach ∂V^ε can be estimated by Lemma 4.2. It remains to prove Lemma 4.4.

We introduce notations and state several technical lemmas needed for the proof of Lemma 4.4.

Since $H(x) = 0$ on ∂U_k, we may assume without loss of generality that $H(x) > 0$ inside U_k in a small neighborhood of U_k. Then there is a region $V \subset U_k$, whose boundary consists of ∂U_k and a level curve $\{H(x) = H_0\}$, and, by selecting a sufficiently small H_0, we can ensure that each level set of H in V is connected and there are no critical points of H in the closure of V other than on ∂U_k.

For $0 \leq H \leq H_0$ let us define the following functions:

$$a(H) = \int |\nabla H| dl, \quad b(H) = \int \frac{\Delta H}{|\nabla H|} dl, \quad q(H) = \int \frac{1}{|\nabla H|} dl \ , \tag{29}$$

in each case the integration is over the level set $\{H(x) = H, x \in V\}$. Let $r < \frac{H_0}{2}$ be a small constant, to be specified later. Consider the function $f(H)$, which solves the equation

$$a(H) f''(H) + b(H) f'(H) = -q(H) \ , \tag{30}$$

with boundary conditions $f(0) = f(2r) = 0$. While it not used here explicitly, we note the fact that the operator in the left hand side of (30) after dividing it by the function $2\,q(H)$ becomes the generator of the limiting diffusion process on the edge of the graph corresponding to the domain U_k (cf [5] and Section 5 of this article).

We need the following lemma, which will be proved in Section 6.

Lemma 4.5. *There is a function $g(r)$, which satisfies $\lim_{r \to 0} g(r) = 0$, such that $|f'(H)| < g(r)$ for all $0 < H < 2r$. Further, there is a constant $c > 0$ such that $|f''(H)| < c|\ln H|$ and $|f'''(H)| < \frac{c}{H}$.*

Let us select constants α_2 and α_3 such that $\alpha_1 < \alpha_2 < \alpha_3 < \frac{1}{2}$. Define the subsets V^A and V^B of V as follows:

$$V^A = \{x \in V; \varepsilon^{\alpha_2} < H(x) < r\}, \quad V^B = \{x \in V; \varepsilon^{\alpha_3} < H(x) < 2r\} \ .$$

Let τ^A be the first time the process X_t^ε leaves V^A, similarly τ^B is the first time the process leaves V^B. Let x_t be the deterministic process

$$dx_t = v(x_t) dt \ ,$$

and let $T(x)$ be the time it takes the process x_t starting at x to make one rotation along the level set, $T(x) = \inf_{t > 0}\{x_t = x\}$. The next lemma shows that for times of order $T(x)$ the process X_t^ε is in a certain sense close to the deterministic process x_t. The lemma is proved in Section 6.

Lemma 4.6. *For any $\delta > 0$ there is $\gamma > 0$ such that*

$$\mathrm{Prob}_x\{\sup_{s \leq T(x)} |H(X_s^\varepsilon) - H(x_s)| > \varepsilon^{\frac{1}{2} - \delta}\} < \varepsilon^\gamma \quad \text{for all} \quad x \in V^A. \tag{31}$$

There exist $\delta' > 0$ and $\gamma > 0$ such that

$$\mathrm{Prob}_x\{\sup_{s \leq T(x)} |X_s^\varepsilon - x_s| > \varepsilon^{\delta'}\} < \varepsilon^\gamma \quad \text{for all} \quad x \in V^A. \tag{32}$$

One of the main ingredients of the proof of Lemma 4.4 is the following lemma, which is a particular case of the main result (Theorem 2.3) of [5].

Lemma 4.7. *[5] There is a constant $c_k > 0$, such that on each level set $\{H(x) = r, x \in V\}$ we have*

$$\lim_{\varepsilon\to 0} \varepsilon \mathbf{E}_x \tau^{U_k} = c_k(1 + g_1(r)) \ , \tag{33}$$

the limit is uniform on each level set, and $g_1(r)$ satisfies $\lim_{r\to 0} g_1(r) = 0$.

Lemma 4.7 is different from Lemma 4.4 in that the initial point in (33) belongs to a fixed level set of H, while in (28) the initial point is asymptotically close to ∂U_k as $\varepsilon \to 0$.

Proof of Lemma 4.4: We shall demonstrate that there exists a function $g(r)$, such that $\lim_{r\to 0} g(r) = 0$, for which

$$\mathbf{E}_x \tau^A \le \varepsilon^{\alpha_1 - 1} g(r) \ , \tag{34}$$

uniformly in $x \in \{H(x) = \varepsilon^{\alpha_1}, x \in V\}$. Let us show that (34) is sufficient to prove the lemma. As in the proof of Lemma 4.3, from (34) it follows that

$$|\mathrm{Prob}_x\{H(X^\varepsilon_{\tau^A}) = r\} - \frac{\varepsilon^{\alpha_1}}{r}| \le c\frac{\varepsilon^{\alpha_1}}{r} g(r) \tag{35}$$

uniformly in $x \in \{H(x) = \varepsilon^{\alpha_1}, x \in V\}$. For $x \in \partial V^\varepsilon \cap U_k$ by the Markov property

$$\mathbf{E}_x \tau^{U_k} = \mathbf{E}_x \tau^A + \mathbf{E}_x(\mathbf{E}_{X^\varepsilon_{\tau^A}} \tau^{U_k}; H(X^\varepsilon_{\tau^A}) = r) + \mathbf{E}_x(\mathbf{E}_{X^\varepsilon_{\tau^A}} \tau^{U_k}; H(X^\varepsilon_{\tau^A}) = \varepsilon^{\alpha_2}) \ . \tag{36}$$

The first term on the right side of (36) is estimated from above by $\varepsilon^{\alpha_1 - 1} g(r)$ due to (34). The second term has the following asymptotics due to (35) and Lemma 4.7

$$|\mathbf{E}_x(\mathbf{E}_{X^\varepsilon_{\tau^A}} \tau^{U_k}; H(X^\varepsilon_{\tau^A}) = r) - c_k \varepsilon^{\alpha_1 - 1}| \le g_2(r) \varepsilon^{\alpha_1 - 1}$$

for sufficiently small ε, where $g_2(r)$ satisfies $\lim_{r\to 0} g_2(r) = 0$. The last term on the right side of (36) is estimated from above by $c\varepsilon^{\alpha_2 - 1}$ due to Lemmas 4.2 and 4.3, and the repeated use of (36). Therefore

$$|\mathbf{E}_x \tau^{U_k} - c_k \varepsilon^{\alpha_1 - 1}| \le \varepsilon^{\alpha_1 - 1} g_3(r)$$

for sufficiently small ε, and $g_3(r)$ which satisfies $\lim_{r\to 0} g_3(r) = 0$. Since r can be selected arbitrarily small we obtain $\lim_{\varepsilon\to 0} \frac{\mathbf{E}_x \tau^{U_k}}{\varepsilon^{\alpha_1 - 1}} = c_k$. It remains to prove (34).

Let $\sigma^B = \min\{\tau^B, T(x)\}$. We shall prove that for some $K > 0$ for all sufficiently small values of r

$$T(x) + \frac{K\mathbf{E}_x f(H(X^\varepsilon_{\sigma^B}))}{\varepsilon} \le \frac{K f(H(x))}{\varepsilon} \tag{37}$$

for all $x \in V^A$. From (37) it follows that

$$\mathbf{E}_x \tau^A \le \frac{K f(H(x))}{\varepsilon} \quad \text{for} \quad x \in V_A. \tag{38}$$

Due to the estimate on the derivative of f from Lemma 4.5 for $x \in \{H(x) = \varepsilon^{\alpha_1}, x \in V\}$ the right side of (38) is estimated from above by $Kg(r)\varepsilon^{\alpha_1 - 1}$, which implies (34). Now we need to prove (37).

Applying Ito's formula to $f(H(X_t^\varepsilon))$ we obtain

$$\frac{1}{\varepsilon}(\mathbf{E}_x f(H(X^\varepsilon_{\sigma^B})) - f(H(x))) = \frac{1}{2}\mathbf{E}_x \int_0^{\sigma^B} (f''(H(X_s^\varepsilon))|\nabla H(X_s^\varepsilon)|^2 + f'(H(X_s^\varepsilon))\Delta H(X_s^\varepsilon))ds \, ,$$

while from (30)

$$T(x) = -\int_0^{T(x)} (f''(H(x_s))|\nabla H(x_s)|^2 + f'(H(x_s))\Delta H(x_s))ds \, .$$

Thus, what we want to show is that there is a constant K_1 such that for all $x \in V^A$

$$\begin{aligned} \mathbf{E}_x \Big| \int_0^{\sigma^B} [f''(H(X_s^\varepsilon))|\nabla H(X_s^\varepsilon)|^2 - f''(H(x_s))|\nabla H(x_s)|^2] ds \Big| + \\ \mathbf{E}_x \Big| \int_0^{\sigma^B} [f'(H(X_s^\varepsilon))\Delta H(X_s^\varepsilon)) - f'(H(x_s))\Delta H(x_s))] ds \Big| + \\ \mathbf{E}_x \Big| \int_{\sigma^B}^{T(x)} [f''(H(X_s^\varepsilon))|\nabla H(X_s^\varepsilon)|^2 + f'(H(X_s^\varepsilon))\Delta H(X_s^\varepsilon))] ds \Big| \le K_1 T(x) \, . \end{aligned} \tag{39}$$

Since f' and ΔH are bounded, and K_1 can be taken arbitrarily large, it is sufficient to estimate only those of the terms in (39) which contain the second derivative of f. By Lemma 4.6 we have $\mathrm{Prob}_x\{\sigma^B < T(x)\} \le \varepsilon^\gamma$, and the second derivative of f can be estimated by Lemma 4.5. Therefore for the last term containing f'' we have

$$\mathbf{E}_x \Big| \int_{\sigma^B}^{T(x)} f''(H(X_s^\varepsilon))|\nabla H(X_s^\varepsilon)|^2 ds \Big| \le cT(x)|\ln(\varepsilon^{\alpha_3})|\mathrm{Prob}_x\{\sigma^B < T(x)\} \le cT(x) \, .$$

The estimate

$$\mathbf{E}_x \Big| \int_0^{\sigma^B} [f''(H(X_s^\varepsilon))|\nabla H(X_s^\varepsilon)|^2 - f''(H(x_s))|\nabla H(x_s)|^2] ds \Big| \le cT(x)$$

follows from Lemma 4.6 and the estimates on f'' and f''' of Lemma 4.5. This completes the proof of Lemma 4.4. □

Proof of Theorem 1.1: The effective diffusivity $D(\varepsilon)$ is related to the variance d^ε of the limit of the functional of the Markov chain, and to the integral of the expectation of the transition times via formula (5). As shown in Lemma 3.1 and in the Remark following it, Lemma 2.1 applies, and therefore there exists the limit $d^0 = \lim_{\varepsilon \to 0} d^\varepsilon$. The asymptotics of the integral $\int_{\mathcal{L}^0} \mathbf{E}_x \tau_1^\varepsilon d\mu^\varepsilon(x)$ is given by (22). This completes the proof of the theorem. □

5 The Case of the Open Channels

In this section it will be convenient to consider the process $\widetilde{X}_t^\varepsilon$, which is the same as X_t^ε, but only accelerated by the factor $\frac{1}{\varepsilon}$, that is $\widetilde{X}_t^\varepsilon = X_{t/\varepsilon}^\varepsilon$. This process satisfies the equation

$$d\widetilde{X}_t^\varepsilon = \frac{1}{\varepsilon} v(\widetilde{X}_t^\varepsilon)dt + dW_t, \qquad \widetilde{X}_t^\varepsilon \in \mathbb{R}^2.$$

Note, that as a process on the torus, $\widetilde{X}_t^\varepsilon$ is uniformly (in ε) exponentially mixing. Following [5] we consider the finite graph G which corresponds to the structure of the level sets of H on the torus.

The graph G is constructed as follows: we identify all the points which belong to each connected component of each level set of H. This way each of the domains U_k, bounded by the separatrices, gets mapped into an edge of the graph, while the separatrices themselves get mapped into the vertices. Let $e(\widetilde{X}_t^\varepsilon)$ label the edge of the graph and let $H(\widetilde{X}_t^\varepsilon)$ be the coordinate on the edge. Then the process $(e(\widetilde{X}_t^\varepsilon), H(\widetilde{X}_t^\varepsilon))$ can be considered as a process on the graph. It is proved in [5] (Theorem 2.2) that the process $(e(\widetilde{X}_t^\varepsilon), H(\widetilde{X}_t^\varepsilon))$ converges to a certain Markov process on the graph with continuous trajectories, which is exponentially mixing. We state the result here in less generality than in [5], but this is sufficient for our purposes.

Theorem 5.1. *[5] There is a Markov (diffusion) process Y_t on the graph G, which is exponentially mixing, and has continuous trajectories, such that for any $T > 0$, the process $(e(\widetilde{X}_t^\varepsilon), H(\widetilde{X}_t^\varepsilon))$ converges to Y_t weakly in $C([0,T], G)$.*

We are now in the position to prove Theorem 1.2.

Proof of Theorem 1.2: The displacement of the process $\widetilde{X}_t^\varepsilon$ in the direction x_1 (the direction of the channels) is given by

$$(\widetilde{X}_t^\varepsilon)^1 = \frac{1}{\varepsilon}\int_0^t v_1(\widetilde{X}_s^\varepsilon)ds + W_t^1,$$

where v_1 is the x_1 component of the velocity field. Therefore,

$$D^{11}(\varepsilon) = \lim_{t\to\infty} \frac{\mathbf{E}_\lambda(\frac{1}{\varepsilon}\int_0^t v_1(\widetilde{X}_s^\varepsilon)ds + W_t^1)^2}{(\frac{t}{\varepsilon})} = \frac{1}{\varepsilon}(\lim_{t\to\infty} \frac{\mathbf{E}_\lambda(\int_0^t v_1(\widetilde{X}_s^\varepsilon)ds)^2}{t} + o(1)) =$$

$$\frac{2}{\varepsilon}(\int_0^\infty \mathbf{E}_\lambda v_1(\widetilde{X}_0^\varepsilon)v_1(\widetilde{X}_s^\varepsilon)ds + o(1)) \ ,$$

where $\widetilde{X}_0^\varepsilon$ is distributed according to the invariant (Lebesgue) measure λ on $\mathbb{T}^2$. For a function $f \in C^\infty(\mathbb{T}^2)$, let $\overline{f}(e,H)$, $(e,H) \in G$ be the function defined on the graph, other than on the vertices, which is equal to the average of f over the corresponding connected component of the level set of H

$$\overline{f}(e,H) = \frac{\int_0^{T(x)} f(x_s)ds}{T(x)},$$

where x_t is the solution of the deterministic equation $dx_t = v(x_t)dt$, the initial point x belongs to the level set, and $T(x)$ is the time of one revolution around the level set. It is easily seen that for any initial point x which does not belong to any of the separatrices of H we have

$$\lim_{\varepsilon\to 0}\int_0^t \mathbf{E}_x f(\widetilde{X}_s^\varepsilon)ds = 0, \quad \text{ if } \ \overline{f}(e,H)\equiv 0.$$

Therefore,

$$\lim_{\varepsilon\to 0}\int_0^t \mathbf{E}_\lambda v_1(\widetilde{X}_0^\varepsilon)v_1(\widetilde{X}_s^\varepsilon)\, ds = \lim_{\varepsilon\to 0}\int_0^t \mathbf{E}_\lambda v_1(\widetilde{X}_0^\varepsilon)\overline{v}_1(e(\widetilde{X}_s^\varepsilon), H(\widetilde{X}_s^\varepsilon))\, ds =$$

$$\int_0^t \mathbf{E}_\mu \overline{v}_1(Y_0)\overline{v}_1(Y_s)\, ds\ , \tag{40}$$

where μ is the measure on G, which is invariant for the process Y_t.

The integrals

$$\int_t^\infty \mathbf{E}_\lambda v_1(\widetilde{X}_0^\varepsilon)v_1(\widetilde{X}_s^\varepsilon)\, ds$$

and

$$\int_t^\infty \mathbf{E}_\mu \overline{v}_1(Y_0)\overline{v}_1(Y_s)\, ds$$

can be made arbitrarily small by selecting sufficiently large t due to uniform mixing of the processes $\widetilde{X}_t^\varepsilon$ and Y_t. Therefore,

$$\lim_{\varepsilon\to 0}\int_0^\infty \mathbf{E}_\lambda v_1(\widetilde{X}_0^\varepsilon)v_1(\widetilde{X}_s^\varepsilon)\, ds = \int_0^\infty \mathbf{E}_\mu \overline{v}_1(Y_0)\overline{v}_1(Y_s)\, ds, \tag{41}$$

which shows that the asymptotics for $D^{11}(\varepsilon)$ is as stated in the theorem.

Now let us consider the asympotics for $D^{22}(\varepsilon)$. Note that $\overline{v}_2(e,H)\equiv 0$, thus the arguments leading to (41) do not provide the asymptotics of $D^{22}(\varepsilon)$. Let $P_1,\ldots,P_n$ be those of the separatrices of H on the torus which, when unfolded onto the plane, are non-compact. Let us select a point A_i on each of P_i. Let us introduce the sequence of stopping times τ_n, $n\geq 1$, which are the consecutive times when $\widetilde{X}_t^\varepsilon$ makes the transition to a different level set P_i. Thus $\widetilde{X}_{\tau_n}^\varepsilon$ is a Markov chain on the set $\{P_1,\ldots,P_n\}$. We can also consider the Markov chain

$$Z_n^\varepsilon = (\widetilde{X}_{\tau_n}^\varepsilon, \tau_n - \tau_{n-1}, \Delta_n)$$

on the extended phase space $\{P_1,\ldots P_n\}\times\mathbb{R}_+\times\mathbb{R}$. The third component Δ_n is defined as follows: If $\widetilde{X}_{\tau_n}^\varepsilon$ is considered on the plane, then

$$\Delta_n = A^2(n) - A^2(n-1),$$

where $A^2(n)$ is the x_2 coordinate of the point corresponding to the separatrix containing the point $\widetilde{X}_{\tau_n}^\varepsilon$.

Similarly we can introduce the stopping times η_n for the process Y_t on the graph, which are the consecutive times when Y_t visits different vertices $Q_i = H(P_i)$ of G, corresponding to the unbounded separatrices of H. Together with the Markov chain Y_{η_n} we can consider the chain

$$\widetilde{Z}_n = (Y_{\eta_n}, \eta_n - \eta_{n-1}, \widetilde{\Delta}_n)$$

on $\{Q_1, \dots Q_n\} \times \mathbb{R}_+ \times \mathbb{R}$, where $\widetilde{\Delta}_n$ is defined the same way as Δ_n.

Let μ^ε be the invariant measure for the chain Z_n^ε, and let $\widetilde{\mu}$ be the invariant measure for the chain $\widetilde{Z}_n$. Let f be the function defined on the state space of the chain Z_n^ε, which is equal to the third component: $f(x, \tau, \Delta) = \Delta$. The function $\widetilde{f}$ is defined the same way on $\{Q_1, \dots Q_n\} \times \mathbb{R}_+ \times \mathbb{R}$.

By the central limit theorem applied to the chain Z_n^ε, there is a number d^ε such that

$$\lim_{n \to \infty} \frac{\sum_{i=0}^n f(Z_i^\varepsilon)}{n} = N(0, d^\varepsilon).$$

Similarly,

$$\lim_{n \to \infty} \frac{\sum_{i=0}^n \widetilde{f}(\widetilde{Z}_i)}{n} = N(0, \widetilde{d}).$$

The effective diffusivity in the x_2 direction is then different form d^ε by the factor $\frac{1}{\varepsilon} \int \tau_1 d\mu^\varepsilon$,

$$D^{22}(\varepsilon) = \frac{\varepsilon d(\varepsilon)}{\int \tau_1 d\mu^\varepsilon}.$$

From Theorem 2.2 in [5] and the uniform mixing of the Markov chains Z_n^ε and $\widetilde{Z}_n$ it easily follows that

$$d(\varepsilon) \to d$$

and

$$\int \tau_1 d\mu^\varepsilon \to \int \eta_1 d\widetilde{\mu}.$$

This completes the proof of Theorem 1.2. □

6 Proof of the Technical Lemmas

Proof of Lemma 4.5: For any function $u \in C^\infty(U_k)$ we have

$$\frac{d}{dH} \int u dl = \int \frac{u \Delta H}{|\nabla H|^2} dl + \int \frac{\langle \nabla H, \nabla(\frac{u}{|\nabla H|}) \rangle}{|\nabla H|} dl \ , \tag{42}$$

where the integrals are over the level set $\{H(x) = H, x \in V\}$. In particular $b(H) = a'(H)$, and therefore equation (30) can be written as

$$(a(H) f'(H))' = -q(H) \ . \tag{43}$$

From the definition of the coefficients $a(H), b(H)$, and $q(H)$ it easily follows that

$$\lim_{H\to 0} a(H) = a_0 > 0\,;\;\; b(H) = O(|\ln H|) \text{ as } H \to 0\,;\;\; q(H) = O(|\ln H|) \text{ as } H \to 0\,. \quad (44)$$

Further, with the help of Morse Lemma and (42) it is easily seen that

$$b'(H) = O(\frac{1}{H}) \text{ as } H \to 0\,;\;\; q'(H) = O(\frac{1}{H}) \text{ as } H \to 0\,. \quad (45)$$

Let $H_m \in (0, 2r)$ be the point where $f(H)$ achieves its maximum, thus $f'(H_m) = 0$. From (43) it follows that

$$f'(H) = \frac{-\int_{H_m}^{H} q(s)ds}{a(H)}\,. \quad (46)$$

Thus, the estimate on the first derivative of f stated in the Lemma follows from (44). Rewrite (30) as

$$f''(H) = -\frac{q(H) + b(H)f'(H)}{a(H)}\,.$$

From (44) in now follows that $|f''(H)| \leq c|\ln H|$ for some $c > 0$. Differentiating both sides of (30) we obtain

$$f'''(H) = -\frac{q'(H) + b'(H)f'(H) + b(H)f''(H) + a'(H)f''(H)}{a(H)}\,.$$

The estimate on $f'''(H)$ now follows from the estimates on the first two derivatives and from (44) and (45). This completes the proof of the Lemma. □

Proof of Lemma 4.6: The proof is based on the use of Lemma 3.2. We can not however apply Lemma 3.2 to the pair of processes X_t^ε and x_t directly, since the rotation time $T(x)$ grows logarithmically in ε when $x \in V^A$.

Let us establish the following property of the deterministic flow x_t:

Let $0 = t_0 < t_1 < t_2 < ... < t_n$. Consider a process y_t, which solves the equation

$$dy_t = v(y_t)dt \quad (47)$$

on each of the segments $[t_0, t_1), [t_1, t_2), ..., [t_{n-1}, t_n]$, with a finite number of jump discontinuities $\lim_{t\to t_i+} y(t) - \lim_{t\to t_i-} y(t) = p_i$, $i = 1, ..., n-1$. Then for any positive c there are positive κ and δ' such that under the conditions

$$x_{t_0} = y_{t_0} \in V^A\,;\;\; \sum_{i=1}^{n-1} ||p_i|| < \varepsilon^{\frac{1}{2}-\kappa}\,;\;\; t_n - t_0 \leq c|\ln\varepsilon|$$

we have

$$\sup_{0\leq t\leq t_n} ||y_t - x_t|| < \varepsilon^{2\delta'}\,. \quad (48)$$

Note that it is sufficient to establish the following: for any pair of points a_0, b_0 such that $a_0 \in V^A$ and $||a_0 - b_0|| \le \varepsilon^{\frac{1}{2}-\kappa}$ we have

$$\sup_{0\le t\le c|\ln\varepsilon|} ||a_t - b_t|| < \varepsilon^{2\delta'+\kappa-\frac{1}{2}}||a_0 - b_0|| \ , \tag{49}$$

where a_t and b_t are the solutions for the deterministic flow (47). Let us take

$$\delta' = \kappa = \frac{1}{4}(\frac{1}{2} - \alpha_2) \ . \tag{50}$$

The time it takes for the trajectory of (47) to make one rotation along the level set $\{H(x) = H, x \in V\}$ is equal to $T(x) = q(H(x))$, where $q(H)$ was defined in (29) and is a smooth function for sufficiently small positive H, which satisfies

$$q(H) = O(|\ln H|) \ , \ q'(H) = O(\frac{1}{H}) \text{ as } H \to 0 \ . \tag{51}$$

The number of full rotations of the trajectory starting from a_0 is equal to $[\frac{t}{T(a_0)}]$ and the time it takes to make these rotations is equal to $[\frac{t}{T(a_0)}]T(a_0)$. It takes $[\frac{t}{T(a_0)}]T(b_0)$ to make the same number of rotations for the trajectory starting at b_0.

Due to (51) the difference is estimated as follows;

$$|[\frac{t}{T(a_0)}]T(a_0) - [\frac{t}{T(a_0)}]T(b_0)| \le \text{const}|\ln\varepsilon|\varepsilon^{-\alpha_2}||a_0 - b_0|| \ . \tag{52}$$

Here we used the facts that $t \le c|\ln\varepsilon|$ and that $H(a_0) \ge \varepsilon^{\alpha_2}$. Now consider the images of a_0 and b_0 under the flow (47) for time $t \le T(a_0)$. Using the reduction of the flow to a linear system in a neighborhood of each of the saddle points (Hartman-Grobman Theorem [7]), it is easy to show that

$$\sup_{0\le T(a_0)} ||a_t - b_t|| \le \text{const}||a_0 - b_0||\varepsilon^{-\alpha_2} \ .$$

Combining this with (52) and with the fact that the speed of motion in (47) is bounded, we obtain (49) with δ' and κ defined in (50). This in turn implies (48) as noted above.

Note that $0 < 2\delta' < \frac{1}{2} - \kappa$ and that by making κ smaller (if necessary) we can satisfy $0 < \kappa < \delta$, where δ is the same as in (31). Observe that for some $c > 0$ we have

$$T(x) < c|\ln\varepsilon| \ \text{ for all } \ x \in V^A. \tag{53}$$

Select the points $0 = t_0 < t_1 < t_2 < ... < t_n = T(x)$ in such a way that $\frac{\kappa}{2K}|\ln\varepsilon| \le |t_{i+1} - t_i| \le \frac{\kappa}{K}|\ln\varepsilon|$ for $i = 0, ..., n-1$, where K is the constant from Lemma 3.2 (applied to the pair of processes X^ε_t and x_t). By (53) there is the estimate $n \le \frac{2cK}{\kappa}$. Let y^ε_t be the piecewise continuous process, which is defined by the conditions: $y^\varepsilon_{t_i} = X^\varepsilon_{t_i}$ and $dy^\varepsilon_t = v(y^\varepsilon_t)dt$ on $[t_i, t_{i+1})$, $i = 0, ..., n-1$. By Lemma 3.2

$$\text{Prob}_x\{\sum_{i=0}^{n-1} \sup_{t\in[t_i,t_{i+1})} ||X^\varepsilon_t - y^\varepsilon_t|| > \varepsilon^{\frac{1}{2}-\kappa}\} \le (\frac{2cK}{\kappa})^3\varepsilon^\kappa \ . \tag{54}$$

Due to continuity of X_t^ε formula (54) provides an estimate on the sum of the jumps of the process y_t^ε. From (48) it now follows that

$$\text{Prob}_x\{||x_t - y_t^\varepsilon|| > \varepsilon^{2\delta'}\} \leq (\frac{2cK}{\kappa})^3 \varepsilon^\kappa \ . \tag{55}$$

This, together with (54) implies (32) for any $\gamma < \kappa$. Since $H(x_t)$ is constant and $H(y_t^\varepsilon)$ is piecewise constant, we have

$$\text{Prob}_x\{\sup_{t \leq T(x)} |H(X_t^\varepsilon) - H(x_t)| > \varepsilon^{\frac{1}{2}-\delta}\} \leq \text{Prob}_x\{\sum_{i=0}^{n-1} \sup_{t \in [t_i, t_{i+1})} ||X_t^\varepsilon - y_t^\varepsilon|| > \frac{\varepsilon^{\frac{1}{2}-\delta}}{\sup ||\nabla H||}\} \ .$$

This, together with (54) and the condition that $\kappa < \delta$ implies (31) for any $\gamma < \kappa$. This completes the proof of the Lemma. □

Acknowledgements: I am grateful to Prof. S. Molchanov for introducing me to this problem and for many useful discussions.

References

[1] S. Childress. *Alpha-effect in flux ropes and sheets*, Phys. Earth Planet Internat., 20 (1979), pp 172-180.

[2] J. L. Doob. *Stochastic Processes.* J. Wiley and Sons, 1953.

[3] A. Fannjiang and G. Papanicolaou. *Convection Enhanced Diffusion for Periodic Flows.* SIAM J. Appl. Math. Vol 54, No 2, pp 333-408, 1994.

[4] M. I. Freidlin. *Dirichlet Problem for Equations with Periodic Coefficients.* Probability Theory and Appl. 9, pp 133-139, 1964.

[5] M. I. Freidlin and A. D. Wentzell. *Random Perturbations of Hamiltonian Systems.* Memoirs of the American Math. Society. Vol 109, No 523, May 1994.

[6] S. Heinze. *Diffusion-advection in cellular flows with large Peclet numbers*, Max-Plank-Institute, preprint no 63, 2001.

[7] L. Perko, *Differential Equations and Dynamical Systems*, Springer-Verlag, 1991.

[8] B. Shraiman, *Diffusive transport in a Rayleigh-Benard convection cell*, Phys. Rev. A, 36 (1987), p 261.

[9] A. M. Soward *Fast dynamo action in steady flow.* J. Fluid Mech., 180 (1987), pp 267-295.