JOURNAL OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 26, Number 2, April 2013, Pages 427–449
S 0894-0347(2012)00755-3
Article electronically published on October 22, 2012

AVERAGING OF INCOMPRESSIBLE FLOWS ON TWO-DIMENSIONAL SURFACES

DMITRY DOLGOPYAT AND LEONID KORALOV

1. Introduction

As a motivating example, consider the stochastic differential equation on the torus

$$dX_t^\varepsilon = \frac{1}{\varepsilon} v(X_t^\varepsilon)dt + dW_t, \quad X_t^\varepsilon \in \mathbb{T}^2, \tag{1.1}$$

with the initial condition $X_0^\varepsilon = X_0$ that does not depend on ε and is independent of the Brownian motion. Here $v(x)$ is a smooth incompressible periodic vector field, W_t is a two-dimensional Brownian motion, and ε is a small parameter. For simplicity of notation assume that the period of v in each of the variables is equal to one and that v is infinitely smooth. Let $H(x_1, x_2)$ be the stream function of the flow, that is,

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

Since v is periodic, we can write H as

$$H(x_1, x_2) = H_0(x_1, x_2) + ax_1 + bx_2,$$

where H_0 is periodic. We shall assume that

(A1) all the critical points of H are nondegenerate and

(A2) a and b are rationally independent.

It is known ([1]) that in this case the structure of the streamlines of v on the torus is as follows. There are finitely may domains U_k, $k = 1, \ldots, r$, bounded by the separatrices of H, such that the trajectories of the dynamical system $\dot{X}_t = v(X_t)$ in U_k are either periodic or tend to a point where the vector field is equal to zero. The trajectories form one ergodic class outside of the domains U_k. More precisely, let $\mathcal{E} = \mathbb{T}^2 \backslash \mathrm{Cl}(\bigcup_{k=1}^r U_k)$. Here $\mathrm{Cl}(\cdot)$ stands for the closure of a set. Then the dynamical system is ergodic on $\mathcal{E}$ (and is, in fact, mixing for a set of rotation numbers of full measure (see [11])).

Although H itself is not periodic, we can consider its critical points as points on the torus, since ∇H is periodic. All the maxima and the minima of H are located inside the domains U_k.

At first, let us assume that each of the domains U_k, $k = 1, \ldots, r$, contains a single critical point M_k (a maximum or a minimum of H). Let A_k, $k = 1, \ldots, r$ be

Received by the editors February 15, 2011 and, in revised form, July 28, 2012.
2010 *Mathematics Subject Classification.* Primary 70K65, 60J60; Secondary 34E10, 37E35.
While working on this article, both authors were supported by the NSF grant DMS-0854982.

the saddle points of H, such that A_k is on the boundary of U_k. Let

$$U = \bigcup_{k=1}^{r} U_k, \quad \gamma_k = \partial U_k, \quad \gamma = \bigcup_{k=1}^{r} \gamma_k.$$

We denote the set $\{x \in \mathrm{Cl}(U_k) : H(x) - H(A_k) = h\}$ by $\gamma_k(h)$. Thus $\gamma_k(0) = \gamma_k$ is the boundary of U_k. Let $p_k = \pm\frac{1}{2}\int_{\gamma_k} |\nabla H| dl$, where the + sign is taken if M_k is a local minimum for H restricted to U_k and $-$ is taken otherwise.

Consider the graph $\mathbb{G}$ that consists of r edges I_k, $k = 1, \ldots, r$ (segments labeled by k), where each segment is either $[H(M_k) - H(A_k), 0]$ (if M_k is a minimum) or $[0, H(M_k) - H(A_k)]$ (if M_k is a maximum). All the edges share a common vertex (the origin) which will be denoted by V. Thus a point on the graph can be determined by specifying k (the number of the edge) and the coordinate on the edge. We define the mapping $h : \mathbb{T}^2 \to \mathbb{G}$ as

$$h(x) = \begin{cases} 0 \quad \text{if } x \in \mathrm{Cl}(\mathcal{E}), \\ (k, H(x) - H(A_k)) \quad \text{if } x \in U_k. \end{cases}$$

We shall use the notation h_k for the coordinate on I_k.

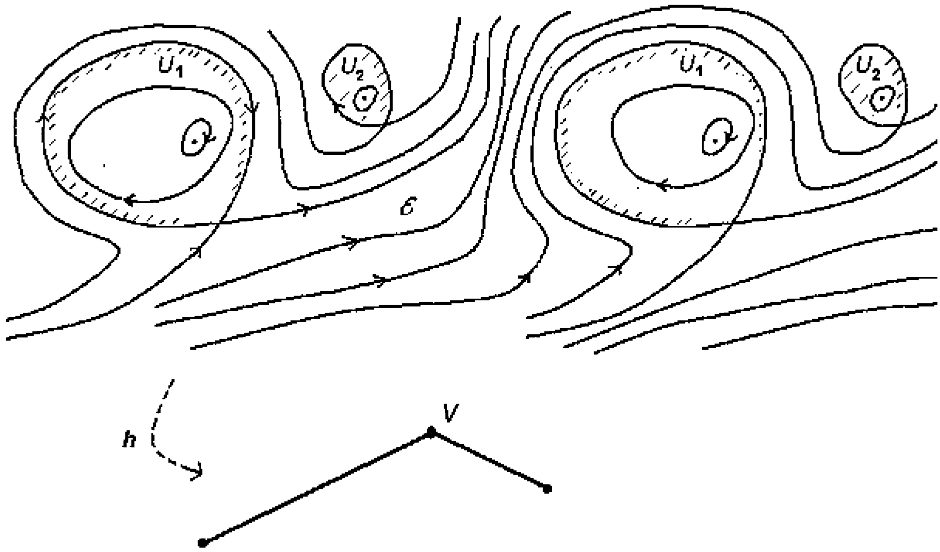

FIGURE 1. A periodic vector field and the corresponding graph $\mathbb{G}$.

Figure 1 shows an example of the flow lines of a vector field on the torus (viewed as a periodic vector field on $\mathbb{R}^2$) and the corresponding graph.

Consider the process Y_t on $\mathbb{G}$ which is defined via its generator $\mathcal{L}$ as follows. First, for each k we define the differential operator $L_k f = a_k(h_k) f'' + b_k(h_k) f'$ on the interior of I_k, where the coefficients a_k and b_k are given by

$$a_k(h_k) = \frac{1}{2}\left(\int_{\gamma_k(h_k)} \frac{1}{|\nabla H|} dl\right)^{-1} \int_{\gamma_k(h_k)} |\nabla H| dl$$

and

$$b_k(h_k) = \frac{1}{2}\left(\int_{\gamma_k(h_k)} \frac{1}{|\nabla H|} dl\right)^{-1} \int_{\gamma_k(h_k)} \frac{\Delta H}{|\nabla H|} dl.$$

The domain of $\mathcal{L}$ consists of those functions $f \in C(\mathbb{G})$ with the following properties:

(a) They are twice continuously differentiable in the interior of each of the edges.

(b) They have the limits $\lim_{h_k \to 0} L_k f(h_k)$ and $\lim_{h_k \to (H(M_k)-H(A_k))} L_k f(h_k)$ at the endpoints of each of the edges. Moreover, the value of the limit

$$q(f) = \lim_{h_k \to 0} L_k f(h_k)$$

is the same for all edges.

(c) They have the limits $\lim_{h_k \to 0} f'(h_k)$, and

$$\sum_{k=1}^{r} p_k \lim_{h_k \to 0} f'(h_k) = \mathrm{Area}(\mathcal{E}) q(f).$$

For functions f which satisfy the above three properties, we define $\mathcal{L}f = L_k f$ in the interior of each edge and as the limit of $L_k f$ at the endpoints of I_k.

It is well known (see [9]) that there exists a strong Markov process on $\mathbb{G}$ with continuous trajectories, with the generator $\mathcal{L}$. The measure on $C([0,\infty), \mathbb{G})$ (here and below $C([0,\infty), \mathbb{G})$ denotes the space of continuous functions from $[0,\infty)$ to $\mathbb{G}$) is uniquely defined by the operator and the initial distribution of the process.

We will prove the following theorem.

Theorem 1.1. *Let X_t^ε be given by* (1.1). *Then the measure on $C([0,\infty), \mathbb{G})$ induced by the process $Y_t^\varepsilon = h(X_t^\varepsilon)$ converges weakly to the measure induced by the process with the generator $\mathcal{L}$ with the initial distribution $h(X_0)$.*

We observe that irrationality of $\rho = a/b$ is necessary for Theorem 1.1, since if ρ is rational, then the restriction of X to $\mathcal{E}$ is periodic rather than ergodic and the graph corresponding to this system will be different. (We note that the limiting process for systems without ergodic components was described in [10].) It has been conjectured by M. Freidlin [8] that Theorem 1.1 holds for all irrational values of ρ. Our earlier paper [6] proves this result for ρ's which cannot be approximated too well by rationals. On the other hand, Sowers [18] shows that in the opposite case of ρ's which are very well approximable by rationals, the result is also true. There still remained a gap between the sets of rotation numbers considered in [6] and in [18]. The current paper establishes the result for general rotation numbers using the arguments that are much less technical than those of [6] or [18]. One of the important ingredients in the proof is an estimate on the time it takes for the process to exit the ergodic component. Here we use the results of [7] and [20] (which are also closely related to [2] and [3]).

The same result remains true if some of the periodic components contain more than one critical point. For example, Figure 2 shows a modification of the system depicted in Figure 1, with U_1 now containing one saddle point and the graph corresponding to the flow. In fact, we will consider a more general situation. Namely, let M be a compact two-dimensional C^∞-surface endowed with a C^∞-area form $\boldsymbol{\omega}$. Let λ be the associated area measure. Let v be a C^∞-incompressible vector field on M. (We refer the reader to [13, Chapter 14] and [17] for the discussion

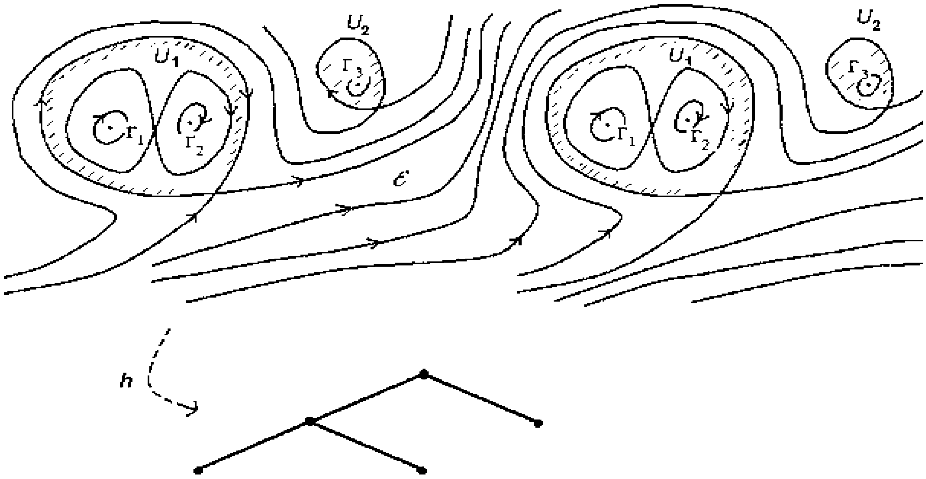

FIGURE 2. The graph $\mathbb{G}$ corresponding to the case when U_1 contains a saddle point.

of applications of incompressible flows on surfaces to geometry and physics.) We assume that v is typical in the sense that

(B1) all the equilibrium points of v are nondegenerate and

(B2) there are no saddle connections.

Instead of (1.1), we now consider the process X_t^ε with generator

$$L^\varepsilon = \frac{1}{\varepsilon} L_v + L_D \tag{1.2}$$

where L_v is the directional derivative along v and L_D is the generator of a non-degenerate diffusion process on M with C^∞-coefficients. We assume that L_D is selfadjoint in $L^2(\lambda)$, that is, λ is invariant measure for X_t^ε for all ε (this assumption will be relaxed later; see Theorem 6.1 in Section 6). We assume that $X_0^\varepsilon = X_0$, where the point X_0 does not depend on ε and is independent of the Brownian motion.

Let G_t^v be the flow generated by v. For $x \in M$, let $\omega(x)$ denote the ω-limits set of x. Under our assumption on v, there are three types of possible limit sets (see [13, Proposition 14.6.2 and Corollary 14.5.12]): equilibrium points, periodic orbits, and a finite number of sets $\mathcal{E}_i$, $i = 1, \ldots, n$, of positive measure. The sets $\mathcal{E}_i$ have the following properties:

(a1) $\mathcal{E}_i = \mathrm{Cl}(\mathrm{Int}(\mathcal{E}_i))$.

(a2) The flow on $\mathcal{E}_i$ is isomorphic to a special flow over an interval exchange transformation (an interval exchange transformation is an invertible piecewise isometry of a segment).

The property (a2) implies ([13, Lemma 14.5.7]) that $(\mathcal{E}_i, G_t^v, \lambda)$ has at most finitely many ergodic components, that is, for each i there are finitely many non-intersecting invariant sets $E_i^1, \ldots, E_i^{m_i} \subseteq \mathcal{E}_i$ such that $\lambda(\mathcal{E}_i \setminus (E_i^1 \cup \cdots \cup E_i^{m_i})) = 0$ and the flow (E_i^j, G_t^v, λ) is ergodic for each j.

We also need a fact ([16, Theorem 3.1.7]) that the set of periodic orbits can be divided into finitely many periodic components U_k where each component belongs to one of the four following classes:

(b1) a cylinder bounded by two separatrix loops;

(b2) a disc bounded by one separatrix loop and having one elliptic equilibrium point inside;

(b3) a sphere $\mathbb{S}^2$ containing two equilibrium points inside (in which case $M = U = \mathbb{S}^2$);

(b4) a torus $\mathbb{T}^2$ (in which case $M = U = \mathbb{T}^2$).

In this paper we exclude the latter two cases since the results we are going to prove are well known in those cases (see [12]).

Let $\mathbb{G}$ be the set of all the ω-limit sets of the flow generated by v. We can regard $\mathbb{G}$ as a topological space where two points (ω-limit sets) are close if the minimal distance between them is small. Then $\mathbb{G}$ is a graph with each edge corresponding to a periodic component of the flow and the vertices corresponding to either equilibrium points or the sets $\mathcal{E}_i$. The restriction of the flow to each of the periodic components is a Hamiltonian flow with Hamiltonian that will be denoted by H. We denote the mapping of M into $\mathbb{G}$ by h.

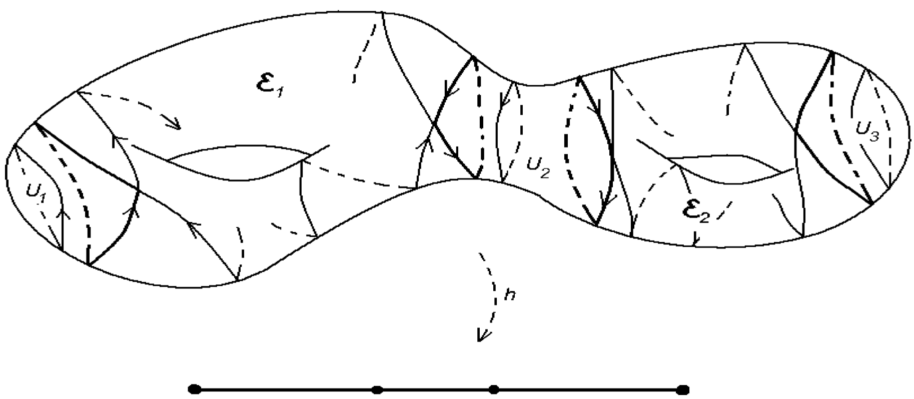

FIGURE 3. Flow lines on a surface and the corresponding graph $\mathbb{G}$.

Figure 3 shows an example of the flow lines of an incompressible vector field on a compact surface and the corresponding graph. More complicated graphs can be obtained by taking one or more surfaces, cutting out the disks along the periodic trajectories, and gluing the resulting parts together. For example, if we take two copies of the surface shown in Figure 3, cut off the right and left tips, and glue the surfaces together, then the graph corresponding to the resulting surface will be a cycle with four vertices (Figure 4(a)). Similarly, cutting the surface corresponding to Figure 2 along the curves Γ_1 and Γ_2 and gluing these curves together will lead to the graph depicted in Figure 4(b). Taking two copies of the surface from Figure 2, cutting each along the curves Γ_1, Γ_2, and Γ_3, and gluing the two copies together will lead to the graph depicted in Figure 4(c).

In order to describe the generator of the limiting process, we introduce a system of local coordinates on $\mathbb{G}$. Namely, given a vertex V of $\mathbb{G}$, let $I_1, \ldots, I_{r_V}$ be the

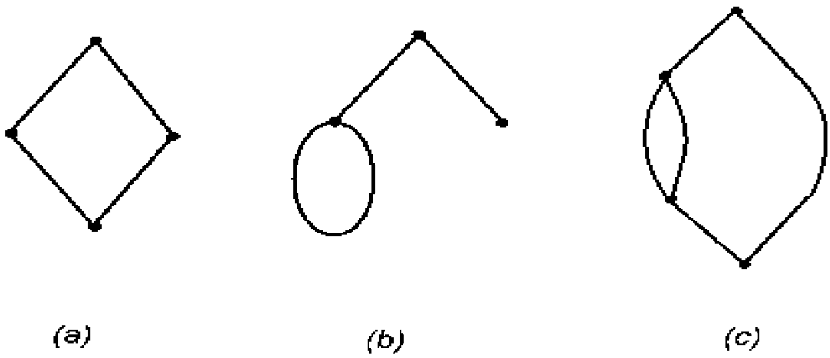

FIGURE 4. Different examples of the graph $\mathbb{G}$.

segments obtained by following the edges of $\mathbb{G}$ starting at V and removing the second endpoint on each edge. If V serves as both endpoints for an edge of $\mathbb{G}$ (as in Figure 4(b)), then this edge is represented twice in the list $I_1, \ldots, I_{r_V}$. We define the star-shaped graph $\mathbb{G}_V$ by taking V as the central vertex, taking the segments $I_1, \ldots, I_{r_V}$ which share a common endpoint V, and adjoining the second endpoint to each of them, with the second endpoint being distinct for each segment. (This is done so that we get two separate edges in $\mathbb{G}_V$ for each edge of $\mathbb{G}$ with both endpoints in V.) The resulting closed segments will still be denoted by $I_1, \ldots, I_{r_V}$. The graphs $\mathbb{G}_V$ play the role of local charts for the graph $\mathbb{G}$.

Let U_k be the periodic component of the flow corresponding to the edge I_k (it is possible that $U_{k_1} = U_{k_2}$ with $k_1 \neq k_2$ if there is an edge of $\mathbb{G}$ with both endpoints in V) and let γ_k be the part of the boundary of U_k corresponding to the endpoint of I_k that is the vertex V. Let A_k be the saddle point on γ_k. As in the case of the torus, we can introduce the coordinates h_k on I_k, $1 \leq k \leq r_V$, thus identifying I_k with the set of values of $H(x) - H(A_k)$ for $x \in \mathrm{Cl}(U_k)$, $k = 1, \ldots, r_V$. We will write $I_k = [m_k, 0]$ if A_k is a local maximum for H restricted to U_k and $I_k = [0, m_k]$ if A_k is a local minimum for H restricted to U_k. (As mentioned above, if the graph $\mathbb{G}$ contains a loop edge (such as in Figure 4(b)), this edge corresponds to two edges of $\mathbb{G}_V$. On one of those edges the coordinate will take values from zero to some number $\overline{h} > 0$, while on the other edge the coordinate will take values from $-\overline{h}$ to zero.)

As in the case of the torus, we denote the set $\{x \in \mathrm{Cl}(U_k) : H(x) - H(A_k) = h_k\}$ by $\gamma_k(h_k)$, thus obtaining $\gamma_k = \gamma_k(0)$. Introduce coordinates (x, y) on U_k so that $\boldsymbol{\omega} = dxdy$. Since λ is invariant, we can write the generator in the form

$$L_D(f) = \frac{1}{2}\mathrm{div}(\alpha \nabla f).$$

Let $L_k f(h_k) = a_k(h_k) f'' + b_k(h_k) f'$ be the differential operator on the interior of I_k with the coefficients

$$a_k(h_k) = \frac{1}{2}\left(\int_{\gamma_k(h_k)} \frac{1}{|\nabla H|} dl\right)^{-1} \int_{\gamma_k(h_k)} \frac{\langle \alpha \nabla H, \nabla H \rangle}{|\nabla H|} dl \tag{1.3}$$

and

$$b_k(h_k) = \frac{1}{2}\left(\int_{\gamma_k(h_k)} \frac{1}{|\nabla H|}dl\right)^{-1} \int_{\gamma_k(h_k)} \frac{2\langle u, \nabla H\rangle + \alpha \cdot H''}{|\nabla H|}dl, \tag{1.4}$$

where $\alpha \cdot H''(x) = \sum_{1\le i,j\le 2} \alpha_{ij}(x) H''_{x_i x_j}(x)$ and

$$u_i = ((\alpha_{1i})'_{x_1} + (\alpha_{2i})'_{x_2})/2, \quad i = 1, 2.$$

Alternatively, we can introduce the action angle coordinates (h_k, θ_k) on U_k, so that $h_k \in I_k$, $\theta_k \in [0, 1]$ is a periodic coordinate on the circle, and the equation of the unperturbed motion takes the form

$$\dot{h}_k = 0, \qquad \dot{\theta}_k = \rho(h_k).$$

In this case L_D can be represented as

$$L_D f = \hat{a}(h_k, \theta_k)\partial^2_{h_k} f + \hat{b}(h_k, \theta_k)\partial_{h_k} f + \tilde{L}f,$$

where $\tilde{L}f$ contains terms involving the derivatives of f with respect to θ and the mixed derivatives. Then

$$a_k = \int_0^1 \hat{a}(h_k, \theta_k) d\theta_k, \qquad b_k = \int_0^1 \hat{b}(h_k, \theta_k) d\theta_k.$$

Let

$$p_k^V = \pm\frac{1}{2}\int_{\gamma_k} \frac{\langle \alpha \nabla H, \nabla H\rangle}{|\nabla H|}dl = \pm\frac{1}{2}\left|\int_{U_k} \operatorname{div}(\alpha \nabla H)(x)dx\right|, \tag{1.5}$$

where the $+$ sign is taken if A_k is a local minimum for H restricted to U_k and $-$ is taken otherwise.

Consider the process Y_t on $\mathbb{G}$ which is defined via its generator $\mathcal{L}$. The domain of $\mathcal{L}$ consists of those functions $f \in C(\mathbb{G})$ which have the following properties when considered on $\mathbb{G}_V$ for each V:

(a) f is twice continuously differentiable in the interior of each of the edges of $\mathbb{G}_V$.

(b) There are the limits $\lim_{h_k \to 0} L_k f(h_k)$ and $\lim_{h_k \to m_k} L_k f(h_k)$ at the endpoints of each of the edges $I_1, \dots, I_{r_V}$. Moreover, the value of the limit $q^V(f) = \lim_{h_k \to 0} L_k f(h_k)$ is the same for all edges.

(c) If V corresponds to a set $\mathcal{E} \in \{\mathcal{E}_1, \dots, \mathcal{E}_n\}$ of positive measure, then there are the limits $\lim_{h_k \to 0} f'(h_k)$, and

$$\sum_{k=1}^{r_V} p_k^V \lim_{h_k \to 0} f'(h_k) = \lambda(\mathcal{E}) q^V(f). \tag{1.6}$$

If V corresponds to a saddle point, then there are the limits $\lim_{h_k \to 0} f'(h_k)$, and

$$\sum_{k=1}^{r_V} p_k^V \lim_{h_k \to 0} f'(h_k) = 0.$$

If V corresponds to an equilibrium point that is not a saddle, then no additional conditions at the endpoint are imposed (we do not need to impose the boundary conditions at vertices corresponding to elliptic equilibria since by [10] such vertices are inaccessible).

For functions f which satisfy the above three properties, we define $\mathcal{L}f = L_k f$ in the interior of each edge and as the limit of $L_k f$ at the endpoints of I_k. It is clear that the value of $L_k f$ does not depend on the choice of the $\mathbb{G}_V$ containing a certain edge, and therefore $\mathcal{L}f$ is defined globally on $\mathbb{G}$.

Let Y_t be the Markov process on $\mathbb{G}$ with continuous trajectories, with the generator $\mathcal{L}$. Our main result is the following theorem.

Theorem 1.2. *Let X_t^ε be given by* (1.2) *and assume that the possible ω-limit sets for the vector field v are either nondegenerate critical points, periodic orbits forming finitely many components of type* (b1) *or* (b2)*, or sets $\mathcal{E}_i$ satisfying* (a1) *and* (a2)*. Then the measure on $C([0,\infty),\mathbb{G})$ induced by the process $Y_t^\varepsilon = h(X_t^\varepsilon)$ converges weakly to the measure induced by the process with the generator $\mathcal{L}$ with the initial distribution $h(X_0)$.*

Note that the conditions of the theorem are satisfied if v is either a vector field on the torus satisfying (A1) and (A2) or a vector field on any surface satisfying (B1) and (B2). In particular, Theorem 1.1 follows from Theorem 1.2

The layout of the paper is the following. In Section 2 we reduce Theorem 1.2 to the case when $\mathbb{G}$ is star-shaped and has only one accessible vertex. This vertex can correspond to either a transitive component $\mathcal{E}_i$ or a saddle point. The more difficult case where the vertex corresponds to a transitive component is considered in Sections 3 and 4. In order to prove Theorem 1.2 in this case we'll introduce stopping times that will mark successive visits by the process to the set $\gamma = \bigcup_k \gamma_k$. Each new stopping time is counted after the process reaches a certain curve inside $U = \bigcup_k U_k$ that is close to γ. One of the main ingredients of the proof is a lemma that shows that the discretized process is sufficiently mixing. This lemma will be proved in Section 4. The lemma, in turn, relies on the estimate of the time it takes for the solution of (1.1) to leave the ergodic component and enter U_k for a given k that will also be derived in Section 4. In Section 3 we prove the main theorem while assuming that the lemma holds. The case where the vertex corresponds to a saddle point is easier and the needed modifications in the proof are discussed in Section 5. (This case has been studied before in [10] but our approach allows us to give a shorter proof.) Finally in Section 6 we describe several applications of Theorem 1.2 following [5].

2. Localization

Here we reduce the proof of our main result to the case where $\mathbb{G}$ is star-shaped, that is, there is a vertex V such that each edge I_k joins V with another vertex V_k. All vertices except V will correspond to elliptic equilibrium points and so they will be inaccessible ([10]).

Let $D(\mathcal{L})$ denote the domain of the operator $\mathcal{L}$. We need the following lemma.

Lemma 2.1. *For any function $f \in D(\mathcal{L})$ and any $T > 0$ we have*

$$\mathbb{E}_x[f(h(X_T^\varepsilon)) - f(h(X_0^\varepsilon)) - \int_0^T \mathcal{L}f(h(X_s^\varepsilon))ds] \to 0 \quad \text{as } \varepsilon \to 0, \tag{2.1}$$

uniformly in $x \in M$.

An analogous lemma was used in the monograph of Freidlin and Wentzell [10] to justify the convergence of the process Y_t^ε to the limiting process on the graph. The main idea, roughly speaking, is to use the tightness of the family Y_t^ε and then to show that the limiting process (along any subsequence) is a solution of the martingale problem, corresponding to the operator $\mathcal{L}$.

The main difference between our case and that of [10] is the presence of ergodic components. However, all the arguments used to prove the main theorem based on (2.1) remain the same. Thus, upon referring to Lemma 3.1 of [10], it is enough to prove our Lemma 2.1 above.

Recall that for each vertex V we have the star-shaped graph $\mathbb{G}_V$ (see Section 1), and the coordinates on its edges $I_1, \ldots, I_{r_V}$ are denoted by $h_1, \ldots, h_{r_V}$. The lengths of the edges are $|m_k|$, $1 \leq k \leq r_V$. Let γ_k', $1 \leq k \leq r_V$, be the curves defined by

$$\gamma_k' = \{x \in U_k : |h_k(x) - h_k(A_k)| = |m_k|/3\}.$$

Note that a separate sequence of curves is defined for each vertex.

We inductively define the following sequence of stopping times. Let $\eta_0 = 0$. If $X^\varepsilon_{\eta_i} \in \gamma_k$, then η_{i+1} is the first time following η_i when the process visits one of the curves γ_k' (possibly, with different k). If $X^\varepsilon_{\eta_i} \in \gamma_k'$, then η_{i+1} is the first time following η_i when the process visits one of the curves γ_k (possibly, with different k and V). If neither $X^\varepsilon_{\eta_i} \in \gamma_k$ nor $X^\varepsilon_{\eta_i} \in \gamma_k'$ is the case for any (V, k) (which is only possible if $i = 0$), then η_1 is defined to be the first time X^ε_t visits either one of the curves γ_k or γ_k' (for some V and k).

By applying the results of [12] it is easy to see that for each δ and T there exists an N such that

$$\mathbb{P}_x(\eta_N < T) < \delta.$$

Therefore, by the strong Markov property of X^ε_t, it suffices to prove that

$$\mathbb{E}_x[f(h(X^\varepsilon_{\min(\eta_1,T)})) - f(h(X^\varepsilon_0)) - \int_0^{\min(\eta_1,T)} \mathcal{L}f(h(X^\varepsilon_s))ds] \to 0 \text{ as } \varepsilon \to 0 \tag{2.2}$$

uniformly in x. We can assume, without loss of generality, that x does not belong to a periodic component, since otherwise (2.2) follows from the classical averaging principle ([12]). Therefore, x projects to a vertex V that corresponds to one of the sets $\mathcal{E}_i$ or to a saddle point. In this case we can construct an auxiliary surface M', as follows. Let M_0 be the set of points in M that can be connected to x by a curve that does not cross any of the curves γ_k', $1 \leq k \leq r_V$. The graph corresponding to the unperturbed dynamics on M_0 is star-shaped, but M_0 is not compact. Its boundary is the union of the curves γ_k', $1 \leq k \leq r_V$. Take domains $D_1, \ldots, D_{r_V}$ diffeomorphic to discs. On each disc we take a vector field preserving the circles centered at the origin, so that the dynamics on each of the circles is a rotation. The domains and the vector fields can be chosen in such a way that they can be glued to γ_1',..., γ_{r_V}' so that the vector field on the resulting surface is smooth and incompressible. The resulting surface will be denoted by M'. For example, if M is the surface depicted in Figure 3 and V is the vertex corresponding to $\mathcal{E}_1$, then M' will be a torus depicted in Figure 5. Similarly, if V corresponds to a saddle point, then M' will be a sphere.

The coefficients of the operator L_D can also be continued from M_0 to M' as smooth functions. The graph corresponding to the dynamics on M' is star-shaped, as required, and all the vertices of the graph other than V are inaccessible. Clearly, if (2.2) holds for M', then it also holds for the original surface with x that gets mapped to V.

It remains to prove (2.2) for the process on M'. Note that (2.2) is different from (2.1) in that now we have $\min(\eta_1, T)$ instead of T. Let us show that (2.1) in fact implies (2.2) for processes (on any surface and in particular) on M'.

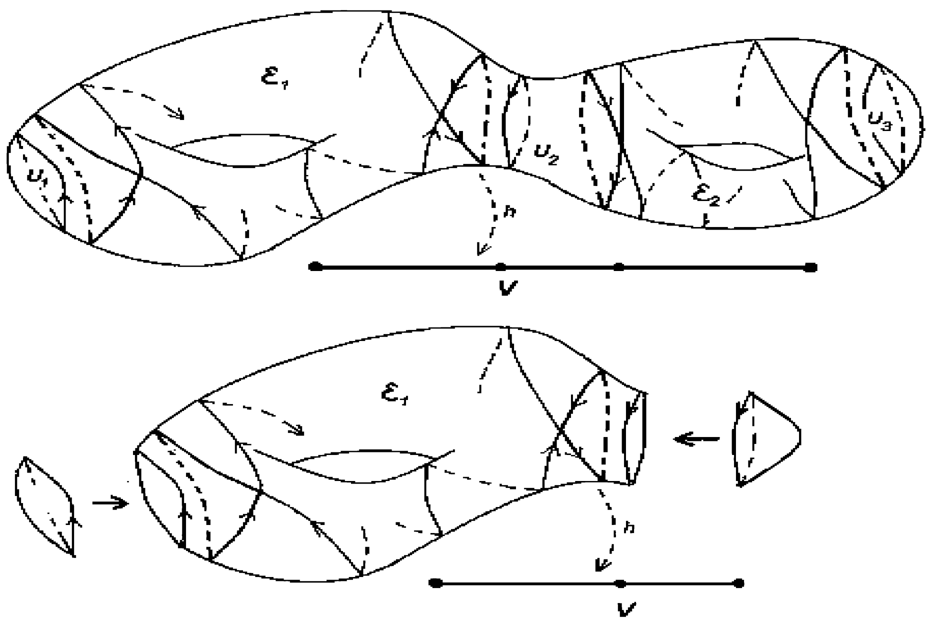

FIGURE 5. Transition from the surface depicted in Figure 3 to a star-shaped one.

Assume that (2.1) holds for the process on M'. Then the process $h(X_t^\varepsilon)$ converges weakly to the limiting diffusion process on the star-shaped graph $\mathbb{G}'$ corresponding to M'.

Let $\tau = \min(\eta_1, T)$ and let $\overline{\tau}$ be the corresponding stopping time on $C([0,T], \mathbb{G}')$ (the space of continuous functions with values in $\mathbb{G}'$). In other words, $\overline{\tau}(h(X_t^\varepsilon)) = \tau$. Note that the function

$$g(\omega) = f(\omega(\overline{\tau})) - f(\omega(0)) - \int_0^{\overline{\tau}} \mathcal{L} f(\omega(s))ds, \quad \omega \in C([0,T], \mathbb{G}'),$$

is bounded and continuous almost surely with respect to the measure induced by the limiting process and the measures induced by the processes $h(X_t^\varepsilon)$. The integral of g with respect to the limiting measure is equal to zero, since $\mathcal{L}$ is the generator of the limiting process. Therefore, the weak convergence implies that (2.2) holds. Thus it remains to prove the following lemma.

Lemma 2.2. *Suppose that the graph $\mathbb{G}$ is star-shaped. Then for any function $f \in D(\mathcal{L})$ and any $T > 0$ we have*

$$\mathbb{E}_x[f(h(X_T^\varepsilon)) - f(h(X_0^\varepsilon)) - \int_0^T \mathcal{L} f(h(X_s^\varepsilon))ds] \to 0 \quad \text{as } \varepsilon \to 0, \tag{2.3}$$

uniformly in $x \in M$.

3. PROOF OF THE MAIN RESULT

In this section we prove Lemma 2.2. In order to fix our ideas we consider the case where V corresponds to a component $\mathcal{E}$ of positive measure. The modifications needed in the case when V corresponds to a saddle point are discussed in Section 5.

The proof of Lemma 2.2 will rely on several other lemmas. Below we shall introduce a number of processes, stopping times, and sets, which will depend on ε. However, we shall not always incorporate this dependence on ε into our notation, so one must be careful to distinguish between the objects which do not depend on ε and those which do.

Let $\overline{\gamma}_k = \gamma_k(-\varepsilon^{1/2})$ if $m_k < 0$ and let $\overline{\gamma}_k = \gamma_k(\varepsilon^{1/2})$ if $m_k > 0$. Let $\overline{\gamma} = \bigcup_{k=1}^r \overline{\gamma}_k$. Recall that $\gamma = \bigcup_{k=1}^r \gamma_k$ is the boundary of U.

Let σ be the first time when the process X_t^ε reaches γ and let τ be the first time when the process reaches $\overline{\gamma}$.

Lemma 3.1. *For each function $f \in D(\mathcal{L})$ we have*

$$\sup_{x\in M} |\mathbb{E}_x[f(h(X_\sigma^\varepsilon)) - f(h(X_0^\varepsilon)) - \int_0^\sigma \mathcal{L}f(h(X_s^\varepsilon))ds]| \to 0 \quad \text{as } \varepsilon \to 0. \tag{3.1}$$

Proof. If the supremum is restricted to the set $M \setminus \mathrm{Cl}(\mathcal{E})$, then the statement follows from the averaging principle inside a periodic component (see [12], [10]). The statement with the supremum taken over $\mathrm{Cl}(\mathcal{E})$ immediately follows from Lemma 4.3 proved in Section 4 if one takes into account that $f(h(x)) = \text{const}$ for $x \in \mathrm{Cl}(\mathcal{E})$. (The proof of Lemma 4.3 does not rely on (3.1).) □

We also need the following lemma that concerns the behavior of the process inside the periodic component.

Lemma 3.2. *There is a constant K independent of ε such that*

$$\sup_{x\in U} \mathbb{E}_x \sigma < K. \tag{3.2}$$

Moreover, for each $k = 1, \dots, r$,

$$\lim_{\mathrm{dist}(x,\mathcal{E})\downarrow 0, x\in U} \mathbb{E}_x \sigma = 0 \tag{3.3}$$

uniformly in ε.

Proof. Both statements follow from [10]. (A more precise statement on the asymptotics of $\mathbb{E}_x\sigma$ when x is a function of ε and $H(x) - H(A_k) \to 0$ can be found in Lemma 4.4 of [14].) □

We inductively define the following two sequences of stopping times. Let $\sigma_0 = \sigma$. For $n \geq 0$ let τ_n be the first time following σ_n when the process reaches $\overline{\gamma}$. For $n \geq 1$ let σ_n be the first time following τ_{n-1} when the process reaches γ.

We can consider the following discrete time Markov chains $\xi_n^1 = X_{\sigma_n}^\varepsilon$ and $\xi_n^2 = X_{\tau_n}^\varepsilon$ with the state spaces γ and $\overline{\gamma}$, respectively. Since the generator of X_t^ε is nondegenerate, both chains satisfy the Doeblin condition. Therefore there are unique invariant measures ν and μ (which depend on ε) on γ and $\overline{\gamma}$ for the chains ξ_n^1 and ξ_n^2, respectively.

Given a stopping time σ^* such that $\sigma^*(\omega) \in \{\sigma_0(\omega), \sigma_1(\omega), \dots\}$ for each ω, we denote by $m_{\sigma^*}^x$ the measure on γ induced by $X_{\sigma^*}^\varepsilon$ starting at x, that is,

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

The following lemma will be proved in Section 4.

Lemma 3.3. *For each $\delta > 0$ and all sufficiently small ε there is a stopping time $\sigma^* = \sigma^*(\delta, \varepsilon, x)$ defined on an extension of the probability space, such that $\mathbb{E}_x \sigma^* \leq \delta$ and*

$$\sup_{x \in \gamma} \mathrm{Var}(m^x_{\sigma^*}(dy) - \nu(dy)) \leq \delta, \tag{3.4}$$

where Var *is the total variation of the signed measure.*

Proof of Lemma 2.2. Let $f \in D(\mathcal{L})$, $T > 0$, and let $\eta > 0$ be fixed. We would like to show that the absolute value of the left-hand side of (2.3) is less than η for all sufficiently small positive ε.

First, we replace the time interval $[0, T]$ by a larger one, $[0, \widetilde{\sigma}]$, where $\widetilde{\sigma}$ is the first of the stopping times σ_n, which is greater than or equal to T, that is,

$$\widetilde{\sigma} = \min_{n : \sigma_n \geq T} \sigma_n.$$

Using the Markov property of the process, the difference can be rewritten as

$$\begin{aligned} &|\mathbb{E}_x[f(h(X^\varepsilon_{\widetilde{\sigma}})) - f(h(X^\varepsilon_0)) - \int_0^{\widetilde{\sigma}} \mathcal{L}f(h(X^\varepsilon_s))ds] \\ &\quad - \mathbb{E}_x[f(h(X^\varepsilon_T)) - f(h(X^\varepsilon_0)) - \int_0^T \mathcal{L}f(h(X^\varepsilon_s))ds]| \\ &\quad\quad = |\mathbb{E}_x \mathbb{E}_{X^\varepsilon_T}[f(h(X^\varepsilon_\sigma)) - f(h(X^\varepsilon_0)) - \int_0^\sigma \mathcal{L}f(h(X^\varepsilon_s))ds]|. \end{aligned}$$

The latter expression can be made smaller than $\eta/3$ for all sufficiently small ε due to (3.1). Therefore, it remains to show that

$$|\mathbb{E}_x[f(h(X^\varepsilon_{\widetilde{\sigma}})) - f(h(X^\varepsilon_0)) - \int_0^{\widetilde{\sigma}} \mathcal{L}f(h(X^\varepsilon_s))ds]| < \frac{2\eta}{3} \tag{3.5}$$

for all sufficiently small ε.

Let $\mathcal{F}_k$ be the σ-algebra of events determined by the time σ_k, and define

$$\alpha_k = \int_{\sigma_k}^{\sigma_{k+1}} \mathcal{L}f(h(X^\varepsilon_s))ds.$$

Since $f(h(x)) = \mathrm{const}$ for $x \in \gamma$, the expectation in the left-hand side of (3.5) can be rewritten as

$$\begin{aligned} &\mathbb{E}_x[f(h(X^\varepsilon_{\widetilde{\sigma}})) - f(h(X^\varepsilon_0)) - \int_0^{\widetilde{\sigma}} \mathcal{L}f(h(X^\varepsilon_s))ds] \\ &\quad = \mathbb{E}_x[f(h(X^\varepsilon_\sigma)) - f(h(X^\varepsilon_0)) - \int_0^\sigma \mathcal{L}f(h(X^\varepsilon_s))ds] + \mathbb{E}_x \sum_{k=0}^{\widetilde{n}} \alpha_k, \end{aligned}$$

where $\widetilde{n}$ is a certain stopping time with respect to the filtration $\{\mathcal{F}_k\}$. The absolute value of the first term on the right-hand side can be made smaller than $\eta/3$ due to (3.1). Define

$$\beta_k = \sum_{n=0}^\infty \mathbb{E}_x(\alpha_{k+n} | \mathcal{F}_k).$$

Below we will prove that the series converges almost surely and for each $\delta > 0$ and all sufficiently small ε

$$\sup_{k \geq 0} |\beta_k| \leq \delta \quad \mathbb{P}_x\text{-almost surely.} \tag{3.6}$$

Assume first that (3.6) holds. Note that $\mathbb{E}_x(\alpha_k - \beta_k + \beta_{k+1}|\mathcal{F}_k) = 0$, and thus

$$\left(\sum_{k=0}^{n}(\alpha_k - \beta_k + \beta_{k+1}), \mathcal{F}_{n+1}\right), \quad n \geq 0,$$

is a martingale. Note that

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

The first term on the right-hand side is equal to zero by the optional sampling theorem, while the absolute value of the second term can be made smaller than $\eta/3$ by choosing $\delta = \eta/6$ in (3.6). It remains to prove (3.6).

Applying the strong Markov property with respect to the stopping time σ, we see that (3.6) will follow if we prove that

$$\sup_{x\in\gamma}\left|\sum_{n=0}^{\infty}\mathbb{E}_x\alpha_n\right| \leq \delta \tag{3.7}$$

for all sufficiently small ε. We need the fact that

$$\mathbb{E}_\nu \alpha_0 = 0, \tag{3.8}$$

which is proved below. First let us show that (3.7) holds if we assume that (3.8) is true. Observe that $\sup_{x\in\gamma}\mathbb{E}_x\alpha_n$ converges to zero exponentially fast as $n \to \infty$ (with the exponent that may depend on ε) since the measure on γ induced by $X^\varepsilon_{\sigma_n}$ converges exponentially to the invariant measure. Hence the series in (3.7) converges absolutely uniformly in x.

By the strong Markov property with respect to the stopping time σ^*, taking (3.8) into account,

$$\left|\sum_{n=0}^{\infty}\mathbb{E}_x\alpha_n\right| \leq ||\mathcal{L}f||_{C(M)}|\mathbb{E}_x\sigma^*| + \mathrm{Var}(m^x_{\sigma^*} - \nu)\sup_{y\in\gamma}\left|\sum_{n=0}^{\infty}\mathbb{E}_y\alpha_n\right|,$$

where σ^* is the same as in Lemma 3.3. Choose a small number $\delta' \ll \delta$. Taking the supremum in x on both sides and applying Lemma 3.3 with δ' in the right-hand side of (3.4), we obtain

$$\sup_{x\in\gamma}\left|\sum_{n=0}^{\infty}\mathbb{E}_x\alpha_n\right| \leq \delta'||\mathcal{L}f||_{C(M)} + \delta'\sup_{x\in\gamma}\left|\sum_{n=0}^{\infty}\mathbb{E}_x\alpha_n\right|.$$

This immediately implies (3.7) since an arbitrarily small δ' can be taken. It remains to prove (3.8), which is the subject of Lemma 3.4 below. □

Lemma 3.4. *For each $\varepsilon > 0$, we have $\mathbb{E}_\nu \int_0^{\sigma_1}\mathcal{L}f(h(X^\varepsilon_s))ds = 0$.*

Proof. Since the process X^ε_t is ergodic and the measure λ on M is the invariant measure, by the law of large numbers

$$\lim_{t\to\infty}\mathbb{E}_\nu\frac{\int_0^t\mathcal{L}f(h(X^\varepsilon_s))ds}{t} = \int_M \mathcal{L}f(h(x))d\lambda(x). \tag{3.9}$$

At the same time, since the Markov chain $X^\varepsilon_{\sigma_n}$, $n \geq 0$, on γ is ergodic and has invariant measure ν, by the law of large numbers,

$$\lim_{n\to\infty}\frac{\sigma_n}{n} = \mathbb{E}_\nu\sigma_1$$

almost surely with respect to the initial distribution ν, and therefore the limit in the left-hand side of (3.9) is equal to

$$\frac{\mathbb{E}_\nu \int_0^{\sigma_1} \mathcal{L}f(h(X_s^\varepsilon))ds}{\mathbb{E}_\nu \sigma_1}.$$

It remains therefore to show that the right-hand side of (3.9) is equal to zero.

For $h_k \in I_k$, define

$$g_k(h_k) = \int_{\gamma_k(h_k)} \frac{1}{|\nabla H|} dl.$$

Then representing $M = \mathcal{E} \cup \mathrm{Cl}(\bigcup_{k=1}^r U_k)$ and replacing the integral over each U_k by an integral over I_k, we obtain

$$\int_M \mathcal{L}f(h(x))d\lambda(x) = \lambda(\mathcal{E})\mathcal{L}f(0) + \sum_{k=1}^r \int_{I_k} g_k(h_k)L_k f(h_k)dh_k.$$

Note that from (1.3) and (1.4) and the Stokes formula it follows that $(g_k a_k)' = g_k b_k$. Therefore

$$\int_{I_k} g_k L_k f dh_k = \int_{I_k} (g_k a_k f'' + g_k b_k f')dh_k = \int_{I_k} (g_k a_k f')' dh_k = -p_k^V \lim_{h_k \to 0} f'(h_k),$$

where the last equality follows from (1.5). The result now follows from (1.6). $\square$

4. Exit from the ergodic component. Lemma on approaching the invariant measure

4.1. A general perturbation result. We make use of the following result proved in [20].

Let Γ be a selfadjoint nonnegative unbounded operator with discrete spectrum on a separable Hilbert space $\mathcal{H}$. Let λ_j be the eigenvalues of Γ and let e_j be the associated orthonormal eigenfunctions. Let $H^m(\Gamma)$ be the space of vectors of the form $\psi = \sum_j c_j e_j$ such that

$$||\psi||^2_{H^m(\Gamma)} = \sum_j |c_j|^2 (\lambda_j + 1)^m < \infty.$$

Let L be a selfadjoint operator on $\mathcal{H}$ such that

$$||L\psi||_{\mathcal{H}} \le \mathrm{Const}||\psi||_{H^1(\Gamma)} \quad \text{and} \quad ||e^{iLt}\psi||_{H^1(\Gamma)} \le B(t)||\psi||_{H^1(\Gamma)}$$

for some continuous function $B(t)$. For each ε, let the function $\varphi_\varepsilon(t)$ solve the equation

$$\frac{d\varphi_\varepsilon}{dt} = \frac{i}{\varepsilon}L\varphi_\varepsilon - \Gamma\varphi_\varepsilon, \quad \varphi_\varepsilon(0) = \psi.$$

Such a solution exists and is unique in the class of functions $\varphi(t)$ satisfying

$$\varphi(t) \in L^2([0,T], H^1(\Gamma)) \cap C([0,T], \mathcal{H}), \qquad \varphi'(t) \in L^2([0,T], H^{-1}(\Gamma))$$

(see Remark 2 on page 649 of [3]). (In our application, below, the equation is parabolic and the existence of solutions and their smoothness away from $t = 0$ does not pose any problem.) Let V be the closed subspace of $\mathcal{H}$ generated by the eigenfunctions of L which lie in $H^1(\Gamma)$. Let $P : \mathcal{H} \to \mathcal{H}$ denote the orthonormal projection to V.

Proposition 4.1 ([20, Theorem 2.4]). *For each T, τ, $\eta > 0$ there exists $\varepsilon_0 = \varepsilon_0(T, \tau, \eta)$ such that for each $\psi \in \mathcal{H}$ with $||\psi||_{\mathcal{H}} \le 1$ we have*

$$\mathrm{mes}\{t \le T : ||(1-P)\varphi_\varepsilon(t)||^2_{\mathcal{H}} > \eta\} \le \tau$$

when $\varepsilon \le \varepsilon_0$.

We will use the following corollary of this proposition. Let M be a compact manifold with a smooth measure λ. Let G^v_t be the flow generated by an incompressible vector field v. Let $E \subset M$ be an open connected invariant set whose boundary consists of a finite number of piecewise smooth hypersurfaces S_k, $k = 1, \ldots, r$. We assume that there are finitely many nonintersecting invariant sets $E_1, \ldots, E_m \subseteq E$ such that $\lambda(E \setminus (E_1 \cup \cdots \cup E_m)) = 0$ and the flow (E_i, G^v_t, λ) is ergodic for each i. Consider the process X^ε_t with the generator $\frac{1}{\varepsilon} L_v + L_D$ where L_D is a selfadjoint nondegenerate elliptic operator of the second order. Let us fix $k \in \{1, \ldots, r\}$. Let $u^\varepsilon(t, y)$ be the probability that the process starting at $y \in M$ does not reach S_k before time t.

Lemma 4.2. *Under the above assumptions, for each k, the $L^1(E)$-norm of $u^\varepsilon(t, \cdot)$ tends to zero as $\varepsilon \downarrow 0$ for each $t > 0$.*

Proof. Let us take an arbitrary Riemannian metric on M. This allows us to introduce the corresponding Sobolev spaces $H^s(M)$, $s \in \mathbb{R}$. Note that $u^\varepsilon(t, y)$ is the solution of the parabolic equation

$$\frac{\partial u^\varepsilon(t,y)}{\partial t} = \left(L_D + \frac{1}{\varepsilon} L_v\right) u^\varepsilon, \quad y \in M \setminus S_k,$$

$$u^\varepsilon(0, y) = 1, \quad y \in M \setminus S_k, \qquad u^\varepsilon(t, y) = 0, \quad t > 0,\ y \in S_k.$$

We apply Proposition 4.1 with $T = t$, $\mathcal{H} = L^2(M)$, $\Gamma = -L_D$, $L = -iL_v$, where Γ and L are the operators on $\mathcal{H}$ with the domains $H^2_0(M \setminus S_k)$ and $H^1_0(M \setminus S_k)$, respectively.

Then $H^1(\Gamma) = H^1_0(M \setminus S_k)$ is the standard Sobolev space and the norms in these two spaces are equivalent. Let us show that each eigenfunction φ of L in $H^1_0(M \setminus S_k)$ is equal to zero on E.

Since φ is an eigenfunction, $\varphi(G^v_t x) = e^{ikt}\varphi(x)$ for some k and all t. Therefore $|\varphi|(G^v_t x) = |\varphi(x)|$, and by ergodicity there are constants a_i, such that $|\varphi(x)| = a_i$ for $x \in E_i$, $i = 1, \ldots, m$. Let us combine the sets E_i into the groups $F_1, \ldots, F_r$ ($r \le m$) so that the values of $|\varphi(x)|$ are constant on each F_i and distinct for different $i \in \{1, \ldots, r\}$. The value of $|\varphi(x)|$ on F_i will be denoted by b_i.

We thus have the function φ that belongs to $H^1_0(M \setminus S_k)$, such that $|\varphi|$ takes a finite number of values on E. We claim that $|\varphi|$ takes only one value on E.

For $h_0 > 0$, let $E^{h_0} = \{x \in E,\ \mathrm{dist}(x, \partial E) > h_0\}$. Let $g_i(x) = \chi_{F_i}(x)$. Let w be a smooth vector field on M with the length of $w(x)$ less than or equal to one for each x. Let $G^w_h(x)$ be the flow generated by the vector field w.

Then for some positive constant K that depends on $b_1, \ldots, b_r$, a positive constant K_1, and each $h \in (0, h_0)$,

$$\int_{E^{h_0}} |g_i(G^w_h x) - g_i(x)| d\lambda(x) \le K \int_{E^{h_0}} ||\varphi(G^w_h x)| - |\varphi(x)||^2 d\lambda(x)$$
$$\le K \int_{E^{h_0}} |\varphi(G^w_h x) - \varphi(x)|^2 d\lambda(x) \le K K_1 h^2 ||\varphi||^2_{H^1(E)}.$$

Similarly, if j, N are positive integers such that $1 \le j \le N$, then

$$\int_{E^{h_0}} |g_i(G^w_{\frac{jh}{N}} x) - g_i(G^w_{\frac{(j-1)h}{N}} x)| d\lambda(x) \le KK_1 \frac{h^2}{N^2} ||\varphi||^2_{H^1(E)}.$$

Therefore,

$$\int_{E^{h_0}} |g_i(G^w_h x) - g_i(x)| d\lambda(x) \le \sum_{j=1}^{N} \int_{E^{h_0}} |g_i(G^w_{\frac{jh}{N}} x) - g_i(G^w_{\frac{(j-1)h}{N}} x)| d\lambda(x)$$
$$\le KK_1 \frac{h^2}{N} ||\varphi||^2_{H^1(E)}.$$

Taking the limit as $N \to \infty$, we see that $\int_{E^{h_0}} |g_i(G^w_h x) - g_i(x)| d\lambda(x) = 0$. Since w and h were arbitrary, this easily implies that g_i is a constant on E^{h_0}. Since h_0 can be taken to be arbitrarily small, this implies that g_i is constant on E for each i, and therefore $|\varphi|$ takes only one value on E.

We now can proceed with the proof of the fact that φ is zero on E. There is a sequence of $C^\infty(M)$ functions f_n, such that the zero set of f_n is a co-dimension 2 submanifold and $f_n \to \varphi$ in $H^1(M)$. Indeed, since $\varphi \in H^1_0(M \setminus S_k)$, it can be approximated by C^∞ functions since C^∞ is dense in $H^1_0(M \setminus S_k)$. Furthermore, by the Sard Lemma, an arbitrarily small complex constant can be added to a C^∞ function to make sure that its zero set becomes a co-dimension 2 submanifold, thus proving the existence of the sequence f_n.

Note that $||(|f_n|)||_{H^1(M)} \le ||f_n||_{H^1(M)}$ since the zero set of f_n has co-dimension 2. Since the embedding of $H^1(M)$ into $H^{3/4}(M)$ is compact, there is a subsequence f_{n_m} and a function $g \in H^{3/4}(M)$ such that $|f_{n_m}| \to g$ in $H^{3/4}(M)$. Since $|f_{n_m}| \to |\varphi|$ in $L^2(M)$, we conclude that $g = |\varphi|$.

Let $\mathrm{tr} : H^{3/4}(M) \to L^2(S_k)$ be the restriction of a function onto the part of the boundary. Since tr is a continuous operator, we have $\mathrm{tr}|f_{n_m}| \to \mathrm{tr}|\varphi|$ in $L^2(S_k)$. On the other hand, $\mathrm{tr}|f_{n_m}| = |\mathrm{tr}\, f_{n_m}| \to |\mathrm{tr}\, \varphi|$ in $L^2(S_k)$ since $|\cdot|$ (taking the absolute value) is a continuous mapping of $L^2(S_k)$ into itself. Therefore $\mathrm{tr}|\varphi| = |\mathrm{tr}\, \varphi|$. Since $\mathrm{tr}\, \varphi = 0$, we have $\mathrm{tr}|\varphi| = 0$, and since $|\varphi|$ is constant on E, the constant must equal zero. Therefore, $\varphi = |\varphi| = 0$ on E.

We have thus demonstrated that all the eigenfunctions vanish on E. Therefore Proposition 4.1 implies that $\int_E (u^\varepsilon(t, y))^2 d\lambda(y) < \eta$ for a large set of times. Since u^ε is a decreasing function of time, it follows that $||u^\varepsilon(t, \cdot)||_{L^2(E)}$ tends to 0 and hence $||u^\varepsilon(t, \cdot)||_{L^1(E)}$ tends to 0. □

Remark. It is proved in [19] that for a typical vector field v on a surface the associated operator L_v has no nonconstant L^1-eigenfunctions on E. In order to apply Proposition 4.1 one needs to check a weaker condition that there are no H^1-eigenfunctions. The proof of Lemma 4.2 shows that this weaker condition holds for *all* vector fields satisfying the assumptions of Theorem 1.2 (in fact, for all the vector fields without saddle connections).

4.2. **Analysis of the exit time.** We start this section by proving several lemmas that describe the behavior of the process in and near the ergodic component. The last one of these lemmas gives a bound on the time it takes for the process that starts in $\mathrm{Cl}(\mathcal{E})$ to reach $\overline{\gamma}_k$ that will be needed for the proof of Lemma 3.3.

Let us show that if the process starts at $x \in \mathrm{Cl}(\mathcal{E})$, then it reaches γ sufficiently fast.

Lemma 4.3. *We have the limit*

$$\lim_{\varepsilon\downarrow 0} \sup_{x\in \mathrm{Cl}(\mathcal{E})} \mathbb{E}_x \sigma = 0.$$

Proof. For $t > 0$ and $x, y \in M$, let m_t^x be the measure induced by X_t^ε starting at x and let $p_x(t, y)$ be the density of m_t^x with respect to the measure λ. The main result of [7] implies that $p_x(t, y)$ is uniformly bounded in ε, x, y for each $t > 0$. In particular, for each $\delta > 0$ there is $k(\delta)$ such that $p_x(\delta/2, y) \le k(\delta)$. Let $u^\varepsilon(t, y)$ be the probability that the process starting at $y \in \mathcal{E}$ does not leave $\mathcal{E}$ before time t. By Lemma 4.2, the $L^1(\mathcal{E})$-norm of $u^\varepsilon(\delta/2, \cdot)$ tends to 0 as $\varepsilon \downarrow 0$. Therefore,

$$\mathbb{P}_x(\sigma > \delta) \le \int_{\mathcal{E}} p_x(\delta/2, y) u^\varepsilon(\delta/2, y) d\lambda(y) \le k(\delta) ||u^\varepsilon(\delta/2, \cdot)||_{L^1(\mathcal{E})} \to 0 \quad \text{as } \varepsilon \downarrow 0.$$

Since δ was arbitrary, the Markov property now implies the statement of the lemma. □

Let $\overline{\lambda}$ be the normalized measure λ on $\mathcal{E}$. Let σ^k be the first time when the process reaches γ_k. We next show that if the process starts with the distribution $\overline{\lambda}$, then it reaches γ_k sufficiently fast.

Lemma 4.4. *For each $k = 1, \ldots, r$ and $\delta > 0$ we have the limit*

$$\lim_{\varepsilon\downarrow 0} \mathbb{P}_{\overline{\lambda}}(\sigma^k > \delta) = 0. \tag{4.1}$$

Moreover, for each $k = 1, \ldots, r$ and $\delta > 0$ there are $\delta' \in (0, \delta)$ and a stopping time $\eta = \eta(\varepsilon) \in [\delta', \delta]$ such that

$$\mathbb{P}_{\overline{\lambda}}(X_\eta^\varepsilon \in \gamma_k) \ge 1 - \delta \tag{4.2}$$

for all sufficiently small ε.

Proof. Let $u^\varepsilon(t, y)$ be the probability that the process starting at y does not reach γ_k before time t. Thus the probability that the process starting with $\overline{\lambda}$ does not reach γ_k before time δ is

$$\mathbb{P}_{\overline{\lambda}}(\sigma^k > \delta) = \int_{\mathcal{E}} u^\varepsilon(\delta, y) d\overline{\lambda}(y).$$

We can apply Lemma 4.2 to conclude that the last integral tends to zero. This proves the first part of the lemma.

The second part follows from the first part and two additional facts:

(a) For each δ', the distribution of $X_{\delta'}^\varepsilon$ starting with distribution $\overline{\lambda}$ has density with respect to λ that is bounded by $1/\lambda(\mathcal{E})$.

(b) Let $\delta > 0$. Then for all sufficiently small δ' and ε we have

$$\mathbb{P}_{\overline{\lambda}}(X_{\delta'}^\varepsilon \notin \mathcal{E}) \le \delta.$$

Now (a) immediately follows from the invariance of λ for the process X_t^ε. Observe that for each $\varkappa > 0$ we have $\mathbb{P}_{\overline{\lambda}}(\mathrm{dist}(X_{\delta'}^\varepsilon, \mathcal{E}) > \varkappa) \to 0$ uniformly in ε as $\delta' \downarrow 0$. This, the first statement of the lemma, and the invariance of λ immediately imply (b). □

Next, we consider the process starting at $x \in \mathrm{Cl}(\mathcal{E})$ rather than starting with the initial distribution $\overline{\lambda}$. First, we need the following simple facts:

Lemma 4.5. (a) *For each $\delta > 0$, there is a $\delta' \in (0, \delta)$ and a stopping time $\eta = \eta(\varepsilon, x) \in [\delta', \delta]$ such that*

$$\sup_{x \in \mathrm{Cl}(\mathcal{E})} \mathbb{P}_x(X^\varepsilon_\eta \notin \mathrm{Cl}(\mathcal{E})) \le \delta$$

for all sufficiently small ε.

(b) *There is a constant $c > 0$ such that for each $\delta > 0$ we have*

$$\sup_{x \in \gamma_k} \mathbb{P}_x(\inf\{t \ge 0 : X^\varepsilon_t \in \overline{\gamma}_k\} > \delta) \le 1 - c$$

for all sufficiently small ε.

(c) *There are a positive constant c and nontrivial subcurves $J_k \subset \gamma_k$, $k = 1, \ldots, r$, such that for each k and each $x \in \overline{\gamma}_k$ the measure induced by X^ε_σ on γ has a component with density bounded from below by c on J_k.*

(d) *We have the limit $\lim_{\varepsilon \downarrow 0} \sup_{x \in \overline{\gamma}_k} \mathbb{E}_x \sigma = 0$.*

(e) *There is a $c > 0$ such that for each $\delta > 0$, each k, all sufficiently small ε (depending on δ), and each $x \in \gamma_k$, there is a stopping time $\eta = \eta(\varepsilon, x) \in [0, \delta]$ with the property that X^ε_η belongs to γ_k with positive probability and the measure induced by X^ε_η on γ_k has a component with density bounded from below by c on J_k.*

Proof. Observe that $\sup_{x \in \mathrm{Cl}(\mathcal{E})} \mathbb{P}_x(\mathrm{dist}(X^\varepsilon_{\delta'}, \mathcal{E}) > \varkappa) \to 0$ uniformly in ε as $\delta' \downarrow 0$. Therefore (a) follows from (3.3). Statement (d) also immediately follows from (3.3).

Statements (b) and (c) follow from the proof of Lemma 2.3 of [6]. The main idea in that lemma was to introduce new coordinates (roughly speaking, dividing H by $\sqrt{\varepsilon}$ and subtracting the fast rotation from the angle variable) so that the generator becomes uniformly elliptic in ε away from the critical points of H. The boundedness of density from below is then known (see [15]).

Finally, (e) immediately follows from (b), (c), (d), and the Markov property of the flow. $\square$

Lemma 4.6. *For each $k = 1, \ldots, r$ and $\bar{\delta} > 0$ we have the limit*

$$\lim_{\varepsilon \downarrow 0} \sup_{x \in \mathrm{Cl}(\mathcal{E})} \mathbb{P}_x(\sigma^k > \bar{\delta}) = 0.$$

Proof. The Markov property together with part (a) of Lemma 4.5 allows us to reduce the statement of the lemma to the following: there is a $c > 0$ such that for each $\hat{\delta} > 0$ we have

$$\sup_{x \in \mathrm{Cl}(\mathcal{E})} \mathbb{P}_x(\sigma^k > \hat{\delta}) \le 1 - c \tag{4.3}$$

for all sufficiently small ε.

Indeed, suppose that (4.3) holds. Fix $\kappa > 0$. Take l such that $(1-c)^l < \kappa/2$. Take $\delta < \min(\kappa/(2l), \bar{\delta}/l)$. Let δ' be given by Lemma 4.5(a). Define stopping times $\eta_1, \ldots, \eta_j, \ldots$ as follows. If $\sigma^k < \delta'$, let $\eta_1 = \sigma^k$; otherwise let η_1 be the first time after δ' such that $X_{\eta_1} \in \mathrm{Cl}(\mathcal{E})$. Next, if η_j is already defined, let $\eta_{j+1} = \sigma^k$ if $\sigma^k < \eta_j + \delta'$ and let η_{j+1} be the first time after $\eta_j + \delta'$ such that $X_{\eta_{j+1}} \in \mathrm{Cl}(\mathcal{E})$. Lemma 4.5(a) implies that

$$\mathbb{P}(\eta_l > \bar{\delta}) < l\delta < \frac{\kappa}{2}.$$

On the other hand (4.3) with $\hat{\delta} = \delta'$ gives

$$\mathbb{P}(\eta_l < \sigma^k) < (1-c)^l < \frac{\kappa}{2}.$$

Hence $\mathbb{P}(\sigma^k > \bar{\delta}) < \kappa$. Since κ is arbitrary, the lemma follows. It remains to prove (4.3).

Let J_j be the subcurves from Lemma 4.5. Denote by λ_j the uniform distribution on J_j. We claim that there is a constant $\bar{c} > 0$ such that for each j, k and all sufficiently small ε we have

$$\mathbb{P}_{\lambda_j}\left(\sigma^k < \frac{\hat{\delta}}{2}\right) > \bar{c}. \tag{4.4}$$

Indeed, by (4.2) there is a $\delta' \in (0, \hat{\delta}/2)$ and a stopping time $\eta = \eta(\varepsilon) \in [\delta', \hat{\delta}/2]$ such that

$$\mathbb{P}_{\overline{\lambda}}(X^\varepsilon_\eta \in \gamma_k) \geq 1 - \hat{\delta}/2. \tag{4.5}$$

From (4.1) and part (e) of Lemma 4.5 it follows that there is a stopping time $\eta' \in [0, \delta']$ such that the measure induced by $X^\varepsilon_{\eta'}$ (starting with $\overline{\lambda}$) has a component with density on J_j with respect to λ_j that is bounded from below by $c > 0$. Since c can be chosen independently of δ', by taking a sufficiently small $\hat{\delta}$ we can ensure that $\hat{\delta}/2 < c$, and therefore $\bar{c} := 1 - \hat{\delta}/(2c) > 0$. By the strong Markov property of the process X^ε_t with respect to η',

$$\mathbb{P}_{\overline{\lambda}}(X^\varepsilon_\eta \in \gamma_k) \leq 1 - c\mathbb{P}_{\lambda_j}\left(\sigma^k \geq \frac{\hat{\delta}}{2}\right),$$

and therefore (4.4) follows from (4.5).

From Lemma 4.3 and part (e) of Lemma 4.5 it follows that there is a stopping time $\eta \in [0, \hat{\delta}/2]$ with the property that the measure induced by X^ε_η (starting from x) has a component with density with respect to λ_j bounded from below by a positive constant on J_j for some j. This, together with (4.4), implies (4.3). □

Let τ^k be the first time when the process reaches $\overline{\gamma}_k$ (observe that $\tau^k > \sigma^k$ when the process starts in $\mathrm{Cl}(\mathcal{E})$).

Lemma 4.7. *For each $k = 1, \ldots, r$ and $\delta > 0$ we have the limit*

$$\lim_{\varepsilon \downarrow 0} \sup_{x \in \mathrm{Cl}(\mathcal{E})} \mathbb{P}_x(\tau^k > \delta) = 0.$$

Proof. As in the proof of Lemma 4.6, the statement will follow if we show that there is a $c > 0$ such that for $\delta > 0$ we have

$$\sup_{x \in \mathrm{Cl}(\mathcal{E})} \mathbb{P}_x(\tau^k > \delta) \leq 1 - c$$

for all sufficiently small ε. This follows from Lemma 4.6, part (b) of Lemma 4.5, and the Markov property of the process. □

Proof of Lemma 3.3. Due to (3.2) and the Markov property of the process, it is sufficient to prove that for each $\delta > 0$ there is a stopping time $\overline{\sigma}$ such that

$$\mathbb{P}_x(\overline{\sigma} \leq \delta) \geq 1 - \delta \tag{4.6}$$

and (3.4) holds with $\overline{\sigma}$ instead of σ^*.

For $k = 1, \dots, r$, let $\sigma_0^k, \sigma_1^k, \dots$ be the subsequence of $\sigma_0, \sigma_1, \dots$ defined by the condition that $X^\varepsilon_{\sigma_n^k} \in \gamma_k$ for each n. Let $q_k = \nu(\gamma_k)$. Note that for each k

$$\nu_k(A) = \frac{\nu(A)}{q_k}, \quad A \in \mathcal{B}(\gamma_k), \tag{4.7}$$

is the invariant measure for $X^\varepsilon_{\sigma_n^k}$ (see [4]).

Let m be a random variable independent of X_t^ε taking values $1, \dots, r$ with probabilities q_k, $k = 1, \dots, r$. Consider the Markov chain $Y_n^k = X^\varepsilon_{\sigma_n^k}$, $n \geq 0$, with the state space γ_k. We have exponentially fast in n convergence of the distributions of Y_n^k to the invariant measure since part (c) of Lemma 4.5 implies that Y_n^k satisfies the Doeblin condition uniformly in ε. Let N be such that the distribution of Y_n^k is δ-close to the invariant measure for each k. Note that N can be chosen independently of ε. Let $\overline{\sigma} = \sigma_N^m$, and thus $X^\varepsilon_{\overline{\sigma}} = Y_N^m$. The fact that $\overline{\sigma}$ satisfies (4.6) now follows from Lemma 4.7 and (3.3). □

Remark. In the companion paper [5] we need a slightly stronger version of Lemma 4.7. Namely, let $\widetilde{\gamma}_k$ be the set of points in U_k whose distance from γ_k is equal to ε^α, where $\alpha \in [1/4, 1/2]$. Let $\widetilde{\tau}^k$ be the first time when the process reaches $\widetilde{\gamma}_k$. We claim that Lemma 4.7 holds with τ^k replaced by $\widetilde{\tau}^k$.

Let us sketch the proof of this fact. Similarly to the arguments above, it is sufficient to show that there is a $c > 0$ such that for each $\delta > 0$ we have

$$\sup_{x \in \mathrm{Cl}(\mathcal{E})} \mathbb{P}_x(\widetilde{\tau}^k > \delta) \leq 1 - c \tag{4.8}$$

for all sufficiently small ε. We claim that

$$\mathbb{E}_{\nu_k} \widetilde{\tau}^k \to 0 \quad \text{as } \varepsilon \downarrow 0 \tag{4.9}$$

where ν_k is the measure on γ_k defined in (4.7). Indeed let $\tilde{\mu}_k$ be the invariant measure for $X_{\tilde{\tau}_m^k}$ where the $\tilde{\tau}_m^k$ are consecutive visits to $\tilde{\gamma}_k$. Equation (15) of [6] reads

$$\lim_{\varepsilon \downarrow 0} \frac{\mathbb{E}_{\nu_k} \tilde{\tau}^k}{\mathbb{E}_{\tilde{\mu}_k} \sigma^k} = \frac{\lambda(M - U_k)}{\lambda(U_k)}.$$

(In [6] this result is proved for $M = \mathbb{T}^2$, but the same argument works for arbitrary surfaces since it only relies on the ergodicity of the process X_t^ε, the asymptotic closeness of $\tilde{\gamma}_k$ and γ_k, and the fact that each of these curves separates the manifold into two disjoint domains.) Hence (4.9) follows from part (d) of Lemma 4.5.

In view of (4.9), in order to see that (4.8) holds, it is sufficient to note the following:

(a) ν_k has a component with density (with respect to the Lebesgue measure) bounded from below by a positive constant on J_k.

(b) By Lemma 4.6 and part (e) of Lemma 4.5, for each $\delta > 0$ and $x \in \mathrm{Cl}(\mathcal{E})$ there is a stopping time $\eta \leq \delta$ such that the distribution of X_η^ε has a component with density that is bounded from below, uniformly in x, by a positive constant on J_k.

5. Saddle points

In the case when V corresponds to a saddle point the result has already been proved in [10]. We also note that the proof presented in Section 3 also works in this case with significant simplifications since the results of Section 4 become trivial.

This simplifies the original argument of [10] by avoiding some of the more technical steps in their proof.

6. Extensions and applications

In this section we describe several applications of Theorem 1.2 following [5].

First, we note that the assumption that L_D is selfadjoint can be removed. Namely, if we do not assume selfadjointness, then L_D can be written in the following form in local coordinates (x, y) on U_k such that $\boldsymbol{\omega} = dxdy$:

$$L_D(f) = L_\beta(f) + \frac{1}{2}\mathrm{div}(\alpha \nabla f).$$

In this case (1.4) has to be modified as

$$b_k(h_k) = \frac{1}{2}\left(\int_{\gamma_k(h_k)} \frac{1}{|\nabla H|}dl\right)^{-1} \int_{\gamma_k(h_k)} \frac{2\langle \beta + u, \nabla H\rangle + \alpha \cdot H''}{|\nabla H|}dl, \tag{6.1}$$

while (1.3) and (1.5) remain unchanged.

Theorem 6.1. *Theorem* 1.2 *remains valid in the dissipative case provided that* (1.4) *is replaced by* (6.1).

Proof. The derivation of Theorem 6.1 from Theorem 1.2 was done in [5] for the case of the torus, but the proof relied only on the facts that the result was known in the conservative case ($\beta = 0$) and that $\mathbb{G}$ had only one accessible vertex. As was explained in Section 2, the general graphs can be reduced to the star-shaped ones, so we can derive Theorem 6.1 from Theorem 1.2 following the argument of [5]. □

Next we consider the random processes $X_t^{\varkappa,\varepsilon}$ with the generator

$$\frac{1}{\varepsilon}L_v(f) + L_\beta(f) + \frac{\varkappa}{2}\mathrm{div}(\alpha\nabla f)$$

where $\varkappa$ is a small parameter. We assume that the initial distribution of the process does not depend on ε and $\varkappa$. Let $Y_t^\varkappa$ denote the limiting process on $\mathbb{G}$ (in the limit $\varepsilon \downarrow 0$). We claim that $Y_t^\varkappa$ has a limit as $\varkappa \downarrow 0$. To describe this limit, define for each edge I_k containing an accessible vertex V,

$$\overline{\psi}_{V,k} = \int_{\gamma_k} \frac{\langle \beta, \nabla H\rangle}{|\nabla H|}dl$$

and

$$\psi_k(h_k) = \left(\int_{\gamma_k(h_k)} \frac{1}{|\nabla H|}dl\right)^{-1} \int_{\gamma_k(h_k)} \frac{\langle \beta, \nabla H\rangle}{|\nabla H|}dl.$$

We assume that $\overline{\psi}_{V,k} \neq 0$ for each accessible vertex and each k.

Given an edge I_k and a vertex V on this edge, let $s(V,k) = 1$ if $\overline{\psi}_{V,k} > 0$ and $h_k > 0$ on I_k, and let $s(V,k) = 0$ otherwise. Denote $q(V,k) = s(V,k)|\overline{\psi}_{V,k}|$.

Let Y_t be the strong Markov process on $\mathbb{G}$ such that:

(a) Inside an edge I_k the process moves deterministically according to the equation

$$\dot{Y}_t = \psi_k(Y_t).$$

(b) If Y_t reaches a vertex V and $s(V,k) = 0$ for all the edges containing V, then Y_t stays at V indefinitely.

(c) If there exist edges such that $s(V,k)=1$, then Y_t leaves V instantaneously if V corresponds to a saddle. If V corresponds to a transitivity component $\mathcal{E}_i$, then Y_t stays at V for a time η which has exponential distribution with parameter

$$\frac{\sum_k q(V,k)}{\lambda(\mathcal{E}_i)},$$

where the sum is over the edges containing V.

(d) When Y_t leaves V, it chooses one of the edges $I_{\overline{k}}$, where $s(V,\overline{k})=1$, with probability

$$\frac{q(V,\overline{k})}{\sum_k q(V,k)}$$

independently of η.

Theorem 6.2. *The measure on $C([0,\infty),\mathbb{G})$ induced by the process $Y_t^{\varkappa}$ converges weakly to the measure induced by the process Y_t with the initial distribution $h(X_0)$.*

Proof. Arguing as in Section 2, we can reduce the statement to the case where $\mathbb{G}$ has one only vertex, in which case the result was proved in [5]. $\square$

Theorem 6.2 says that if we perturb an incompressible vector field v on M by a small dissipative vector field $\varepsilon\beta$, then (after a viscosity type regularization), the system exhibits intermittent behavior. Namely, the limiting motion is deterministic on a finite or infinite number of time segments separated by random delays. Moreover, the consecutive deterministic segments are chosen at random. In the case of M being the torus considered in [5], the graph $\mathbb{G}$ is a tree, and hence the process Y_t visits each point at most once and converges to a random limit as $t\to\infty$. In contrast, if $\mathbb{G}$ has loops, then Y_t may visit some parts of the graph infinitely many times allowing for an infinite sequence of transitions between the intervals with deterministic and random behavior.

Acknowledgements

The authors are grateful to Mark Freidlin for introducing them to this problem and to both Mark Freidlin and Andrej Zlatos for insightful discussions. The authors are also grateful to the referees for a number of valuable suggestions.

References

[1] V. I. Arnol′d, *Topological and ergodic properties of closed* 1*-forms with incommensurable periods*, Funktsional. Anal. i Prilozhen. **25** (1991), no. 2, 1–12, 96, DOI 10.1007/BF01079587 (Russian); English transl., Funct. Anal. Appl. **25** (1991), no. 2, 81–90. MR1142204 (93e:58104)

[2] Henri Berestycki, François Hamel, and Nikolai Nadirashvili, *Elliptic eigenvalue problems with large drift and applications to nonlinear propagation phenomena*, Comm. Math. Phys. **253** (2005), no. 2, 451–480, DOI 10.1007/s00220-004-1201-9. MR2140256 (2006b:35057)

[3] P. Constantin, A. Kiselev, L. Ryzhik, and A. Zlatoš, *Diffusion and mixing in fluid flow*, Ann. of Math. (2) **168** (2008), no. 2, 643–674, DOI 10.4007/annals.2008.168.643. MR2434887 (2009e:58045)

[4] I. P. Cornfeld, S. V. Fomin, and Ya. G. Sinaĭ, *Ergodic theory*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 245, Springer-Verlag, New York, 1982. Translated from the Russian by A. B. Sosinskiĭ. MR832433 (87f:28019)

[5] D. Dolgopyat, M. Freidlin, and L. Koralov, *Deterministic and stochastic perturbations of area preserving flows on a two-dimensional torus,* Erg. Th. Dyn. Sys. **32**, no. 3 (2012), 899–918.

[6] Dmitry Dolgopyat and Leonid Koralov, *Averaging of Hamiltonian flows with an ergodic component*, Ann. Probab. **36** (2008), no. 6, 1999–2049, DOI 10.1214/07-AOP372. MR2478675 (2010d:37105)

[7] Brice Franke, *Integral inequalities for the fundamental solutions of diffusions on manifolds with divergence-free drift*, Math. Z. **246** (2004), no. 1-2, 373–403, DOI 10.1007/s00209-003-0604-1. MR2031461 (2005e:58060)

[8] Mark Freidlin, *Markov processes and differential equations: asymptotic problems*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 1996. MR1399081 (97f:60150)

[9] Mark I. Freidlin and Alexander D. Wentzell, *Diffusion processes on graphs and the averaging principle*, Ann. Probab. **21** (1993), no. 4, 2215–2245. MR1245308 (94j:60116)

[10] Mark I. Freidlin and Alexander D. Wentzell, *Random perturbations of Hamiltonian systems*, Mem. Amer. Math. Soc. **109** (1994), no. 523, viii+82. MR1201269 (94j:35064)

[11] Ya. G. Sinaĭ and K. M. Khanin, *Mixing of some classes of special flows over rotations of the circle*, Funktsional. Anal. i Prilozhen. **26** (1992), no. 3, 1–21, DOI 10.1007/BF01075628 (Russian); English transl., Funct. Anal. Appl. **26** (1992), no. 3, 155–169. MR1189019 (93j:58079)

[12] R. Z. Has′minskiĭ, *On the principle of averaging the Itô's stochastic differential equations*, Kybernetika (Prague) **4** (1968), 260–279 (Russian, with Czech summary). MR0260052 (41 #4681)

[13] Anatole Katok and Boris Hasselblatt, *Introduction to the modern theory of dynamical systems*, Encyclopedia of Mathematics and its Applications, vol. 54, Cambridge University Press, Cambridge, 1995. With a supplementary chapter by Katok and Leonardo Mendoza. MR1326374 (96c:58055)

[14] L. Koralov, *Random perturbations of 2-dimensional Hamiltonian flows*, Probab. Theory Related Fields **129** (2004), no. 1, 37–62, DOI 10.1007/s00440-003-0320-0. MR2052862 (2005e:60179)

[15] Carlo Miranda, *Partial differential equations of elliptic type*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 2, Springer-Verlag, New York, 1970. Second revised edition. Translated from the Italian by Zane C. Motteler. MR0284700 (44 #1924)

[16] Igor Nikolaev and Evgeny Zhuzhoma, *Flows on 2-dimensional manifolds*, Lecture Notes in Mathematics, vol. 1705, Springer-Verlag, Berlin, 1999. An overview. MR1707298 (2001b:37065)

[17] S. P. Novikov, *The semiclassical electron in a magnetic field and lattice. Some problems of low-dimensional "periodic" topology*, Geom. Funct. Anal. **5** (1995), no. 2, 434–444, DOI 10.1007/BF01895674. MR1334874 (96f:58124)

[18] Richard B. Sowers, *Random perturbations of two-dimensional pseudoperiodic flows*, Illinois J. Math. **50** (2006), no. 1-4, 853–959 (electronic). MR2247849 (2008a:60084)

[19] Corinna Ulcigrai, *Weak mixing for logarithmic flows over interval exchange transformations*, J. Mod. Dyn. **3** (2009), no. 1, 35–49, DOI 10.3934/jmd.2009.3.35. MR2481331 (2010b:37005)

[20] Andrej Zlatoš, *Diffusion in fluid flow: dissipation enhancement by flows in 2D*, Comm. Partial Differential Equations **35** (2010), no. 3, 496–534, DOI 10.1080/03605300903362546. MR2748635 (2012c:35178)

Department of Mathematics, University of Maryland, College Park, Maryland 20742
E-mail address: dmitry@math.umd.edu

Department of Mathematics, University of Maryland, College Park, Maryland 20742
E-mail address: koralov@math.umd.edu