ENOC-2005, Eindhoven, Netherlands, 7-12 August 2005

NEW RESULTS ON RANDOM PERTURBATIONS OF PSEUDOPERIODIC FLOWS

Richard B. Sowers
Department of Mathematics
University of Illinois at Urbana–Champaign
Urbana, IL 61801
r-sowers@math.uiuc.edu

Abstract
Arnol'd in 1991 characterized "pseudoperiodic" flows on the 2-dimensional torus as having both ergodic components, periodic orbits, and homoclinic orbits. We consider small random perturbations of such flows. Under appropriate scaling of time, we search for an averaged picture which describes the evolution of local "energies". Under certain circumstances, we identify a limiting Markov process with glueing conditions (as suggested by Freidlin in 1996) which characterizes energy evolution.

Key words
Stochastic Averaging, Pseudoperiodic Flows, Glueing Conditions

1 Introduction

This is a précis of the results of [Sowersc] (with proofs omitted). An important technique in the analysis of many physical systems is the circle of ideas known as *model reduction*; i.e., the development of rigorous methods to replace, often in some limiting regime, a complicated system by a simpler, or lower-dimensional one. More particularly, the ideas known as *averaging* take advantage of regular behavior along a *fast* motion to identify lower-dimensional dynamics for a *slow* motion.

If the fast motion is deterministic, averaging requires, naturally, the existence of long-term averages. This is definitely possible if the fast motion is periodic. It is also possible if the fast motion is ergodic. Arnol'd [Arnol'd, 1991] (see also [Sinaĭ and Khanin, 1992]) identified a simple system which contains both ergodic and periodic orbits; a pseudoperiodic system on the two-dimensional torus. For such a flow, there is a partition of the torus into an *ergodic* region, and a collection of *traps*. Inside each of the traps, the flow can be described via a local Hamiltonian. Our interest is how small diffusive perturbations cause transitions between the traps and the ergodic class.

Since our interest is the effect of *small* noise, we have a separation of scales. The fast variable is the position within orbits of the dominant dynamical system; an *angle*. The slow variable distinguishes between orbits; an *action*. The theory of averaging (in this case, stochastic averaging) suggests that we look for closed dynamics of the action variable. The effective coefficients of these closed dynamics are given, informally, by fixing the slow variable and taking long-time averages in the fast variable. In the simplest cases, when all orbits are periodic, the space of action variables is usually diffeomorphic to a line, and is formally given by taking the quotient with respect to the action of the fast orbits. When there are bifurcations in the topology of the orbits, the notion of *chain equivalence* is the correct way to include the effect of small perturbations; then the action variable in general takes values in a graph, or more generally, a stratified space (see [Freidlin and Wentzell, 1994], [Freidlin and Weber, 1998], [Freidlin and Weber, 1999], [Novikov, Papanicolaou, and Ryzhik], [Sowers, 2002]). The formal asymptotic goal is to show that when the trajectories of the original randomly perturbed dynamical system are projected onto the space of action variables, i.e., the space of chain-equivalent classes, they asymptotically (under an appropriate change of time scale), tend to a Markov process (on the space of chain-equivalent classes). The interesting part is the effect of bifurcations, which create different strata; at the chain-equivalent set representing these strata, glueing conditions must be imposed.

Returning to the pseudoperiodic problem, the focus of this paper is the effect of the ergodic class. Sooner or later, the diffusive perturbations will push the randomly-perturbed trajectory into the ergodic class. The ergodicity will then take the particle everywhere in the ergodic class, and eventually it will exit back into a trap. A quantification of this effect was was conjectured in [Freidlin, 1996, p. 74]. Chain equivalence collapses the whole ergodic class to a single point, and the conjecture is that the limiting process is *sticky* at this point, with a computable stickiness coefficient. Our goal is to show that in a certain weak sense, this is true.

We admit that the arguments of this paper are perhaps a bit more complex than the audience might anticipate. We hope that the need for a more general understanding of the interplay of noise and dynamics will compensate for the complexity.

2 Problem Statement and Main Result

Let's start with a *pseudoperiodic Hamiltonian* on $\mathbb{R}^2$. Let $\langle\cdot,\cdot\rangle_{\mathbb{R}^2}$ be the standard inner product on $\mathbb{R}^2$.

Assumption 2.1. *Let* H *in* $C^\infty(\mathbb{R}^2)$ *and* $\boldsymbol{\omega}=(\omega_1,\omega_2)$ *in* $\mathbb{R}^2$ *be such that: firstly,* H *is Morse, secondly,* ω_1 *and* ω_2 *are incommensurable (i.e.,* $\langle\boldsymbol{\omega},K\rangle_{\mathbb{R}^2}\neq 0$ *for all nonzero* K *in* $\mathbb{Z}^2$*), and thirdly,* $\mathsf{H}(x+K)=\mathsf{H}(x)+\langle\boldsymbol{\omega},K\rangle_{\mathbb{R}^2}$ *for all* $x\in\mathbb{R}^2$ *and* $K\in\mathbb{Z}^2\subset\mathbb{R}^2$.

Define[1] the vector field

$$(\mathsf{Y}_e\varphi)(x)\overset{\text{def}}{=}\left(\frac{\partial\mathsf{H}}{\partial x_2}\frac{\partial\varphi}{\partial x_1}-\frac{\partial\mathsf{H}}{\partial x_1}\frac{\partial\varphi}{\partial x_2}\right)(x)\qquad(1)$$

for all $\varphi\in C^\infty(\mathbb{R}^2)$ and $x=(x_1,x_2)\in\mathbb{R}^2$ (i.e., Y_e is the symplectic or skew gradient of H).

We now want to add diffusivity, albeit in a periodic way. Let $C^\infty_p(\mathbb{R}^2)$ be the set of those $f\in C^\infty(\mathbb{R}^2)$ such that $f(x+K)=f(x)$ for all $x\in\mathbb{R}^2$ and $K\in\mathbb{Z}^2$. Note that $\partial\mathsf{H}/\partial x_1$ and $\partial\mathsf{H}/\partial x_2$ are both in $C^\infty_p(\mathbb{R}^2)$. Thus $\mathsf{Y}C^\infty_p(\mathbb{R}^2)\subset C^\infty_p(\mathbb{R}^2)$. We now define a *diffusion generator and bracket*.

Assumption 2.2. *Let* $\mathscr{L}_e$ *be a second-order partial differential operator of the form*

$$(\mathscr{L}_e f)(x)\overset{def}{=}\frac{1}{2}\sum_{i,j\in\{1,2\}}a^{(2)}_{i,j}(x)\frac{\partial^2 f}{\partial x_i\partial x_j}(x)+\sum_{i\in\{1,2\}}a^{(1)}_i(x)\frac{\partial f}{\partial x_i}(x)$$

for all $f\in C^2(\mathbb{R}^2)$ *and* $x=(x_1,x_2)\in\mathbb{R}^2$*, where the* $a^{(2)}_{i,j}$*'s and* $a^{(1)}_i$*'s are in* $C^\infty_p(\mathbb{R}^2)$*. We require for simplicity that* $\mathscr{L}_e$ *be strongly elliptic; i.e., that*

$$\sum_{i,j\in\{1,2\}}a^{(2)}_{i,j}(x)\frac{\partial f}{\partial x_i}(x)\frac{\partial f}{\partial x_2}(x)>0$$

for all $x=(x_1,x_2)\in\mathbb{R}^2$ *and all* $f\in C^1(\mathbb{R}^2)$ *with* $df(x)\neq 0$.

Then $\mathscr{L}C^\infty_p(\mathbb{R}^2)\subset C^\infty_p(\mathbb{R}^2)$.

Since the $\partial\mathsf{H}/\partial x_i$'s, $a^{(1)}_i$'s, and $a^{(2)}_{i,j}$'s are all in $C^\infty_p(\mathbb{R}^2)$, we can now move to the two dimensional torus $\mathbb{T}\overset{\text{def}}{=}\mathbb{R}^2/\mathbb{Z}^2$. Let $\mathfrak{t}:\mathbb{R}^2\to\mathbb{T}$ the standard covering map; i.e., $\mathfrak{t}(x)=x+\mathbb{Z}^2$ for all $x\in\mathbb{R}^2$. We then define the vector field Y and the second-order operator $\mathscr{L}$ by requiring that $(\mathsf{Y}\varphi)(\mathfrak{t}(x))=(\mathsf{Y}_e(\varphi\circ\mathfrak{t}))(x)$ and $(\mathscr{L}\varphi)(\mathfrak{t}(x))=(\mathscr{L}_e(\varphi\circ\mathfrak{t}))(x)$ for all $\varphi\in C^\infty(\mathbb{T})$ and all $x\in\mathbb{R}^2$.

We will consider the Markov process on $\mathbb{T}$ whose generator is $\mathscr{L}^\varepsilon\overset{\text{def}}{=}\frac{1}{\varepsilon^2}\mathsf{Y}+\mathscr{L}$ (with domain $\mathscr{D}(\mathscr{L}^\varepsilon)\supset C^2(\mathbb{T})$). We will construct this Markov process in a canonical way, via the martingale problem (see [Ethier and Kurtz, 1986] and [Stroock and Varadhan, 1979]). Define the event space $\Omega\overset{\text{def}}{=}C([0,\infty);\mathbb{T})$. Define the coordinate functions $X_t(\omega)\overset{\text{def}}{=}\omega(t)$ for all $t\geq 0$ and all $\omega\in\Omega$. For each $t\geq 0$, define $\mathscr{F}_t\overset{\text{def}}{=}\sigma\{X_s;\,0\leq s\leq t\}$ and define a sigma-algebra on Ω by $\mathscr{F}\overset{\text{def}}{=}\bigvee_{t\geq 0}\mathscr{F}_t$. We can now define our principal objects of interest, starting with the *original martingale problem*.

Definition 2.3. *Fix* $x_\circ\in\mathbb{T}$*. For each* $\varepsilon>0$*, let* $\mathbb{P}^\varepsilon\in\mathscr{P}(C([0,\infty);\mathbb{T}))$ *be a solution to the martingale problem with generator* $\mathscr{L}^\varepsilon$ *whose domain contains* $C^2(\mathbb{T})$ *(as a dense subset), and initial condition* $\delta_{x_\circ}$*. Let* $\mathbb{E}^\varepsilon$ *be the corresponding expectation operator. This means the following. Firstly, that* $\mathbb{P}^\varepsilon\{X_0=x_\circ\}=1$*. Secondly, that if we fix* $f\in C^2(\mathbb{T})$*,* $0\leq r_1<r_2\cdots<r_n\leq s<t$ *and* $\{\varphi_j;\,j=1,2\dots n\}\subset C_b(\mathbb{T})$*, then*

$$\mathbb{E}^\varepsilon\left[\left\{f(X_t)-f(X_s)-\int_s^t(\mathscr{L}^\varepsilon f)(X_u)du\right\}\prod_{j=1}^n\varphi_j(X_{r_j})\right]=0.$$

In other words, $\mathbb{P}^\varepsilon$ is the law of the stochastic differential equation

$$dY^\varepsilon_t=\frac{1}{\varepsilon^2}\mathsf{Y}(Y^\varepsilon_t)dt+\check{a}_0(Y^\varepsilon_t)dt+\sum_{i\in\{1,2\}}\check{a}_i(Y^\varepsilon_t)\circ dW^i_t\qquad t>0\quad(2)$$

where W^1 and W^2 are two independent standard Wiener processes, and where $\check{a}_0$, $\check{a}_1$, and $\check{a}_2$ are smooth vector fields on $\mathbb{T}$ such that (in Hörmander form)

$$\frac{1}{2}\sum_{i\in\{1,2\}}\check{a}_i^2+\check{a}_0=\mathscr{L}.$$

The generator $\mathscr{L}^\varepsilon$ is, of course, a speeded-up version of the operator $\mathsf{Y}+\varepsilon^2\mathscr{L}$ (to get the corresponding stochastic differential equation, change (2) as follows: remove the $1/\varepsilon^2$ from the Y term, put ε^2 in front of $\check{a}_0$, and ε in front of $\check{a}_1$ and $\check{a}_2$). The operator $\mathsf{Y}+\varepsilon^2\mathscr{L}$ represents a combination of motion along

[1] We shall attach the superscript e when referrring to the Euclidean space $\mathbb{R}^2$, which we endow with the standard metric and symplectic form.

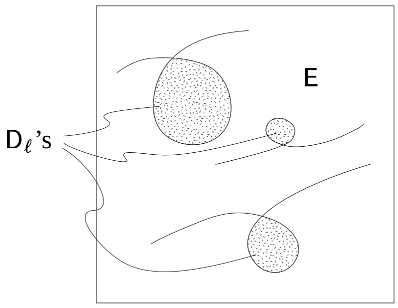

Figure 1. Pseudoperiodic Flow

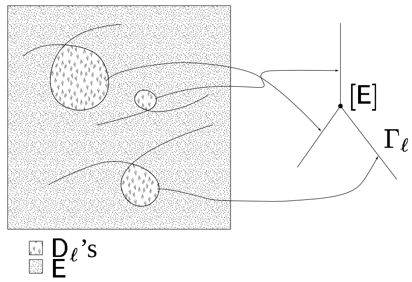

Figure 2. Chain Equivalence Reduction

the integral curves of Ч and small random perturbations. The change in time scale stems from a desire to see how diffusive perturbations cause motion *across* the orbits of Ч.

Let now $\mathfrak{z}$ be the flow of diffeomorphisms of $\mathbb{T}$ defined by

$$\begin{aligned} \dot{\mathfrak{z}}_t(x) &\stackrel{\text{def}}{=} Ч(\mathfrak{z}_t(x)) \qquad t \geq 0 \\ \mathfrak{z}_0(x) &= x. \end{aligned} \qquad x \in \mathbb{T}$$

The novelty of our problem comes from the structure of $\mathfrak{z}$, which Arnol'd [Arnol'd, 1991] identified. There is a partition of $\mathbb{T}$ into a finite collection $\{\mathsf{D}_\ell;\ \ell \in \Lambda\}$ of closed *traps* (Λ is simply the index set) and an open *ergodic* set E. Both E and each of the D_ℓ's is invariant under $\mathfrak{z}$. The interior of each trap is diffeomorphic to the open unit disk in $\mathbb{R}^2$, and $\partial \mathsf{D}_\ell$ is a homoclinic orbit of $\mathfrak{z}$ with fixed point $\mathfrak{x}_\ell$. Furthermore, for each trap D_ℓ, there is an $\mathsf{H}_{T,\ell} \in C^\infty(\mathbb{T})$ (a Hamiltonian) such that $\mathsf{H}_{T,\ell} \equiv 0$ on $\partial \mathsf{D}_\ell$ and such that Ч is the symplectic gradient of $\mathsf{H}_{T,\ell}$ on D_ℓ (i.e., $Ч(\mathfrak{t}(x)) = T\mathfrak{t}(\bar{\nabla}_e(\mathsf{H}_{T,\ell} \circ \mathfrak{t}))(x)$ for all $x \in \mathfrak{t}^{-1}(\mathsf{D}_\ell)$) (here $\bar{\nabla}_e$ is the symplectic gradient with respect to the Euclidean symplectic form on $\mathbb{R}^2$; see (1)). In E, the orbits of $\mathfrak{z}$ are dense. See Figure 1. For simplicity, we impose an extra assumption.

Assumption 2.4. *We assume that for each $\ell \in \Lambda$, there is a unique critical point $\mathfrak{p}_\ell \in \mathsf{D}_\ell^\circ$ of $\mathsf{H}_{T,\ell}$.*

We could actually remove this assumption (see Remark 2.5). The requirement that H is Morse implies that the $\mathfrak{p}_\ell$'s are elliptic fixed points.

A general tool in considering small diffusive perturbations of conservative systems is the notion of *chain equivalence* relative to $\mathfrak{z}$ (see [Conley, 1978] and [Robinson, 1999]). For a positive integer N, $\delta \in (0,1)$ and $T \in (0,\infty)$, we say that there is an (N,δ)-chain of time T from $x \in \mathbb{T}$ to $y \in \mathbb{T}$ if there is a sequence $(z_j;\ j = 1,2\dots N)$ of points in $\mathbb{T}$ and a sequence $0 = t_0 < t_1 \cdots < t_N = T$ of times such that $z_0 = x$ and $z_N = y$ and such that $\|\mathfrak{z}_{t_j - t_{j-1}}(z_{j-1}) - z_j\| < \delta$ for all $1 \leq j \leq N$. We say that $x \Rightarrow y$, where x and y are in $\mathbb{T}$, if there is a positive integer N such that for each $\delta \in (0,1)$ and $T \in (0,\infty)$, there is an (N,δ)-chain of time $T' \in (T,\infty)$ from x to y. We say that $x \sim y$ if $x \Rightarrow y$ and $y \Rightarrow x$. We note that $x \sim x$ for each $x \in \mathbb{T}$, and that $\sim$ is an equivalence relation on $\mathbb{T}$. Define $\mathbb{M} \stackrel{\text{def}}{=} \mathbb{T}/\sim$ and endow $\mathbb{M}$ with the quotient topology defined by $\sim$. For $x \in \mathbb{T}$, we let $[x] \stackrel{\text{def}}{=} \{y \in \mathbb{T} : y \sim x\}$ be the equivalence class of x (the chain components of $\mathbb{T}$). Then

$$\mathbb{M} = \bigcup_{\ell \in \Lambda} \Gamma_\ell \cup \bar{\mathsf{E}} \cup \bigcup_{\ell \in \Lambda} \{\mathfrak{p}_\ell\} \tag{3}$$

where $\Gamma_\ell = \{[x] : x \in \mathsf{D}_\ell^\circ \setminus \{\mathfrak{p}_\ell\}\}$ for all $\ell \in \Lambda$. We note that $\bar{\mathsf{E}}$ and the $\{\mathfrak{p}_\ell\}$'s are all single points in $\mathbb{M}$. It is easy to see that for each $\ell \in \Lambda$, Γ_ℓ is a one-dimensional open C^∞ manifold (diffeomorphic to $(0,1)$), and the points $\bar{\mathsf{E}}$ and $\{\mathfrak{p}_\ell\}$ are the limits of points in Γ_ℓ. This makes $\mathbb{M}$ into a *stratified space* [Goresky and MacPherson, 1988] if we enforce the ordering $\bar{\mathsf{E}} \prec \Gamma_\ell$ and $\{\mathfrak{p}_\ell\} \prec \Gamma_\ell$ for all $\ell \in \Lambda$ (see also [Sowers, 2002] for another example of a Markov process on a stratified space resulting from averaging). We note that (3) represents $\mathbb{M}$ as a disjoint union of open manifolds and a collection of boundary points; the set $\Gamma_\Lambda \stackrel{\text{def}}{=} \bigcup_{\ell \in \Lambda} \Gamma_\ell$ consists of the open manifolds.

We also note that there is a homeomorphism between Γ and a "wye" (see Figure 2). Let $\{\mathbf{v}_\ell;\ \ell \in \Lambda\}$ be a collection of unit vectors in $\mathbb{R}^2$ such that for each distinct pair ℓ and ℓ' of elements of Λ, $\mathbf{v}_\ell$ and $\mathbf{v}_{\ell'}$ are linearly independent. Define $\daleth : \mathbb{M} \to \mathbb{R}^2$ by

$$\daleth([x]) \stackrel{\text{def}}{=} \begin{cases} |\mathsf{H}_{T,\ell}(x)|\mathbf{v}_\ell & \text{if } x \in \mathsf{D}_\ell \text{ for } \ell \in \Lambda \\ \mathbf{0}_e & \text{if } x \in \mathsf{E} \end{cases}$$

where $\mathbf{0}_e$ is the origin of $\mathbb{R}^2$. The image of $\daleth$ is the "spider" $\{\mathbf{0}_e\} \cup \bigcup_{\ell \in \Lambda} (0,\hbar]\mathbf{v}_\ell$. Since $\daleth$ is a homeomorphism into $\mathbb{R}^2$, we can define $d^\dagger([x],[y]) \stackrel{\text{def}}{=} \|\daleth(x) - \daleth(y)\|_e$ for all x and y in $\mathbb{T}$, where $\|\cdot\|_e$ is the standard metric on $\mathbb{R}^2$. Then $d^\dagger$ is a metric on $\mathbb{M}$, and we can see that $\mathbb{M}$ is in fact Polish.

Remark 2.5. *If we remove the assumption of Assumption 2.4, then the topology of $\mathbb{M}$ will be more complicated. Inside each D_ℓ, we should use an analysis like that of [Freidlin and Wentzell, 1994].*

Set $X_t^M \stackrel{\text{def}}{=} [X_t]$ for all $t \ge 0$ and define $\mathbb{P}^{\varepsilon,\dagger}(A) \stackrel{\text{def}}{=} \mathbb{P}^\varepsilon\{X^M \in A\}$ for all A in $\mathscr{B}(C([0,\infty),\mathbb{M}))$; i.e., $\mathbb{P}^{\varepsilon,\dagger}$ is the law of the projection of the process $t \mapsto X_t$ onto $C([0,\infty),\mathbb{M})$.

It will not be hard to show *tightness*.

Proposition 2.6. *The $\mathbb{P}^{\varepsilon,\dagger}$'s are tight in the Prohorov topology on $\mathscr{P}(C([0,\infty);\mathbb{M}))$.*

Thus it is appropriate to investigate the existence and uniqueness of the limit $\lim_{\varepsilon\to 0}\mathbb{P}^{\varepsilon,\dagger}$, this limit being in the Prohorov topology. We want to show that in certain cases, this limit exists along a subsequence, and can be identified as a certain Markov process. We note that since $[X]$ records only part of the location of X, $\mathbb{P}^{\varepsilon,\dagger}$ is not Markovian for $\varepsilon > 0$. Our goal is to show that as $\varepsilon \to 0$, the limit *is* Markovian (and thus that the discarded information can be replaced via *effective* coefficients).

As long as X^M stays in Γ_Λ, it should tend to a process with *averaged* coefficients. Let $\mathfrak{m}(x) \stackrel{\text{def}}{=} [x]$ for all $x \in \mathbb{T}$ and note that $\mathfrak{m}^{-1}(\Gamma_\Lambda) = \bigcup_{\ell\in\Lambda}(\mathsf{D}_\ell^\circ\setminus\{\mathfrak{p}_\ell\})$. Define a linear *averaging* operator $\mathcal{A} : C(\mathfrak{m}^{-1}(\Gamma_\Lambda)) \to C(\Gamma_\Lambda)$. For $\varphi \in C(\mathfrak{m}^{-1}(\Gamma_\Lambda))$, define

$$(\mathcal{A}\varphi)([x]) = \frac{\int_{z\in[x]}\varphi(z)\|\mathsf{Ч}(z)\|^{-1}\mathscr{H}^1(dz)}{\int_{z\in[x]}\|\mathsf{Ч}(z)\|^{-1}\mathscr{H}^1(dz)} \qquad (4)$$

for all $[x] \in \Gamma_\Lambda$, where $\|\cdot\|$ is the standard metric on $T\mathbb{T}$ and where $\mathscr{H}^1$ is standard 1-dimensional Hausdorff measure on $\mathbb{T}$. If $\varphi \in C(\mathfrak{m}^{-1}(\Gamma_\Lambda))$, then

$$(\mathcal{A}\varphi)([x]) = \lim_{T\to\infty}\frac{1}{T}\int_0^T \varphi(\mathfrak{z}_s(x))ds$$

for all $x \in \mathfrak{m}^{-1}(\Gamma_\Lambda)$. Define next an averaged diffusive operator

$$(\mathscr{L}_{\text{ave}}f)([x]) \stackrel{\text{def}}{=} (\mathcal{A}(\mathscr{L}(f\circ\mathfrak{m})))([x])$$

for all $x \in \mathfrak{m}^{-1}(\Gamma_\Lambda)$ and $f \in C^2(\Gamma_\Lambda)$ (since Γ_Λ is a C^∞ manifold, $f\circ\mathfrak{m} \in C^2(\mathfrak{m}^{-1}(\Gamma_\Lambda))$. We then expect that the limiting dynamics of X^M will given by the generator $\mathscr{L}_{\text{ave}}$ as long as it remains in Γ_Λ.

The remaining, and most interesting, question, is the limiting behavior at $[\mathsf{E}]$. The following was conjectured in [Freidlin, 1996, p. 74]. Since $\mathscr{L}_{\text{ave}}$ is a nondegenerate elliptic operator on $C^2(\Gamma_\Lambda)$ we can consequently define the nonnegative bilinear form $\langle\cdot,\cdot\rangle_{\text{ave}}$ on $T^*\Gamma_\Lambda$ by

$$\begin{aligned}\langle df, dg\rangle_{\text{ave}}([x]) &= (\mathscr{L}_{\text{ave}}(fg))([x]) \\ &\quad - f([x])(\mathscr{L}_{\text{ave}}g)([x]) - g([x])(\mathscr{L}_{\text{ave}}f)([x])\end{aligned}$$

for all f and g in $C^2(\Gamma_\Lambda)$ and all $[x] \in \Gamma_\Lambda$. We next define *area* functions for each Γ_ℓ. Let $\mathscr{H}^2$ be standard 2-dimensional Hausdorff measure on $\mathbb{T}$. For each $\ell \in \Lambda$ and each $[x] \in \Gamma_\ell$, define

$$\eth_\ell([x]) \stackrel{\text{def}}{=} \mathscr{H}^2\{z \in \mathsf{D}_\ell : |\mathsf{H}_{T,\ell}(z)| \le |\mathsf{H}_{T,\ell}(x)|\}. \qquad (5)$$

We then define the glueing operator

$$\mathscr{G}_\ell f \stackrel{\text{def}}{=} \lim_{\substack{[x]\to[\mathsf{E}]\\ [x]\in\Gamma_\ell}} \langle df, d\eth_\ell\rangle_{\mathscr{L},\text{ave}}([x])$$

if this limit exists.

Lemma 2.7. *Fix $f \in C^2(\Gamma_\Lambda)$ such that*

$$\lim_{\substack{[x]\to[\mathsf{E}]\\ [x]\in\Gamma_\Lambda}} (\mathscr{L}_{ave}f)([x])$$

exists. Then $\mathscr{G}_\ell f$ is well-defined for each $\ell \in \Lambda$.

The proof of this is the same as that of [Sowers, 2003, Lemma 1.5]. We can write $\mathscr{G}_\ell$ in a better way. Let $\boldsymbol{\sigma}$ in $C^\infty(\mathbb{T})$ be such that $\boldsymbol{\sigma}(\mathfrak{t}(x)) = \langle d\mathsf{H}, d\mathsf{H}\rangle(x)$ for all $x \in \mathbb{R}^2$. We then define

$$\mathrm{G}_\ell \stackrel{\text{def}}{=} \int_{z\in\partial\mathsf{D}_\ell}\frac{\boldsymbol{\sigma}(z)}{\|\mathsf{Ч}(z)\|}\mathscr{H}^1(dz). \qquad (6)$$

Next, let Λ_P be the collection of $\ell \in \Lambda$ such that $\mathsf{H}_{T,\ell} > 0$ on D_ℓ°, and let Λ_W be the collection of $\ell \in \Lambda$ such that $\mathsf{H}_{T,\ell} < 0$ on D_ℓ°. Define $\mathrm{S}_\ell \stackrel{\text{def}}{=} 1$ if $\ell \in \Lambda_P$, and define $\mathrm{S}_\ell \stackrel{\text{def}}{=} -1$ if $\ell \in \Lambda_W$. Fix $f \in C^1(\Gamma_\Lambda)$ and for each $\ell \in \Lambda$ let $f_\ell : \mathbb{R} \to \mathbb{R}$ be such that $f_\ell(\mathsf{H}_{T,\ell}(x)) = f(\mathfrak{m}(x))$ for all $x \in \mathsf{D}_\ell$, then $\mathscr{G}_\ell f$ exists if and only if $\dot{f}_\ell(0)$ exists (as a limit from $\mathsf{H}_{T,\ell}(\mathfrak{m}^{-1}(\mathsf{D}_\ell^\circ))$), and then $\mathscr{G}_\ell f = \mathrm{S}_\ell \mathrm{G}_\ell \dot{f}_\ell(0)$.

We can now define our *limiting domain and generator*

Definition 2.8. *Define*

$$\begin{aligned}\mathscr{A}^\dagger \stackrel{def}{=} \Big\{(f,g) \in C(\mathbb{M})\times C(\mathbb{M}) :& \\ f \in C^2(\Gamma_\Lambda),\ g = \mathscr{L}_{ave}f \text{ on } \Gamma_\Lambda,& \\ \text{and } 2g([\mathsf{E}])\mathscr{H}^2(\mathsf{E}) = \sum_{\ell\in\Lambda}\mathscr{G}_\ell f&\Big\}\end{aligned}$$

The requirement that $2g([\mathsf{E}])\mathscr{H}^2(\mathsf{E}) = \sum_{\ell\in\Lambda}\mathscr{G}_\ell f$ is call the **glueing condition**. It means that $[\mathsf{E}]$ is "sticky" (see [Harrison and Lemoine, 1981]); i.e., that, asymptotically, X^M spends positive time at $[\mathsf{E}]$. See also Remark 2.13 of [Sowers, 2002] for some motivational comments. Finally, we don't need to explicitly specify any boundary conditions at the $\{\mathfrak{p}_\ell\}$'s, since these points will be inaccessible for the limiting diffusion.

Remark 2.9. *It is not surprising that the glueing coefficient at* $[\mathsf{E}]$ *involves* $\mathscr{H}^2$ *near* E *(see* (5)*). Assume, for the sake of discussion, that the* $\mathbb{P}^\varepsilon$*-law of* X_t *has a density with respect to 2-dimensional Lebesgue measure (i.e.,* $\mathscr{H}^2$*) for each* $t \ge 0$*. This density is described by a PDE both in* E *and outside of* E*. To uniquely specify this density, we need only specify some boundary conditions which relate the densities inside and outside* $\partial\mathsf{E}$*. Of course the* $\mathbb{P}^\varepsilon$ *dynamics of* X *have no singularities at* $\partial\mathsf{E}$*, so it should be enough to specify that the density is continuous at* $\partial\mathsf{E}$*, and that the flux across* $\partial\mathsf{E}$ *should sum up to zero. These boundary conditions at* $\partial\mathsf{E}$ *implicitly use the fact that the densities are with respect to* standard Lebesgue measure *on both sides of* $\partial\mathsf{E}$*. The glueing conditions give us continuity and conservation of flux when we are* $\mathscr{H}^2$ *on both sides of* E*. Note that the definition of the* $\eth_\ell$*'s would not change if we would replace* $\mathsf{H}_{T,\ell}$ *in* (5) *by some homeomorphism of* $\mathsf{H}_{T,\ell}$*; the important issue is that we are using 2-dimensional Lebesgue measure.*

To formalize the theory surrounding $\mathscr{A}^\dagger$, we make the usual setup on the event space $\Omega^\dagger \overset{\text{def}}{=} C([0,\infty);\mathbb{M})$. Define the coordinate functions $X^\dagger_t(\omega) \overset{\text{def}}{=} \omega(t)$ for all $t \ge 0$ and all $\omega \in \Omega^\dagger$. For each $t \ge 0$, define $\mathscr{F}^\dagger_t \overset{\text{def}}{=} \sigma\{X^\dagger_s;\, 0 \le s \le t\}$ and define a sigma-algebra on $\Omega^\dagger$ by $\mathscr{F}^\dagger \overset{\text{def}}{=} \bigvee_{t\ge 0}\mathscr{F}^\dagger_t$.

Proposition 2.10. $\mathscr{A}^\dagger$ *generates a strongly continuous, positive, contraction semigroup on* $C(\mathbb{M})$*; i.e., there is a unique* $\mathbb{P}^\dagger \in \mathscr{P}(\Omega^\dagger)$ *which solves the martingale problem with generator* $\mathscr{A}^\dagger$ *and initial distribution* $\delta_{[x_\circ]}$*. In other words, there is a unique* $\mathbb{P}^\dagger \in \mathscr{P}(\Omega^\dagger)$ *such that* $\mathbb{P}^\dagger\{X^\dagger_0 = [x_\circ]\} = 1$ *and such that for* $(f,g) \in \mathscr{A}^\dagger$*,* $0 \le r_1 < r_2 \dots < r_n \le s < t$*, and* $\{\varphi^\dagger_j;\, j = 1,2\dots n\} \subset C(\mathbb{M})$*,*

$$\mathbb{E}^\dagger\left[\left\{f(X^\dagger_t) - f(X^\dagger_s) - \int_s^t g(X^\dagger_u)du\right\}\prod_{j=1}^n \varphi^\dagger_j(X^\dagger_{r_j})\right] = 0,$$

where $\mathbb{E}^\dagger$ *denotes the expectation operator associated with* $\mathbb{P}^\dagger$*.*

To properly state our results, we need some restrictions on the ratio $\varrho \overset{\text{def}}{=} \frac{\omega_1}{\omega_2}$, which is by assumption irrational. First, let's construct the continued-fraction expansion of ϱ. For all $z \in \mathbb{R}$, define $\lfloor z\rfloor \overset{\text{def}}{=} \max\{j \in \mathbb{Z} : j \le z\}$ and $\iota(z) \overset{\text{def}}{=} z - \lfloor z\rfloor$. Define $\Bbbk_1 \overset{\text{def}}{=} \varrho$, and for $n \in \mathbb{N}$ (as usual $\mathbb{N} \overset{\text{def}}{=} \{1,2\dots\}$), recursively define $\Bbbk_{n+1} \overset{\text{def}}{=} \frac{1}{\iota(\Bbbk_n)}$. For each $n \in \mathbb{N}$, we then define $\bar{\Bbbk}_n \overset{\text{def}}{=} \lfloor \Bbbk_n\rfloor$. For each $j \in \mathbb{N}$, define $[[j]] \overset{\text{def}}{=} j$, and if we have defined $[[j_1, j_2 \dots j_N]] \in (0,\infty)$ for all $(j_1, j_2 \dots j_N) \in \mathbb{N}^N$ for some $N \in \mathbb{N}$, we then define

$$[[j_1, j_2 \dots j_{N+1}]] \overset{\text{def}}{=} j_1 + \frac{1}{[[j_2,\dots j_{N+1}]]}$$

for all $(j_1, j_2 \dots j_{N+1}) \in \mathbb{N}^{N+1}$; this will of course be positive. We then define $\varrho_N \overset{\text{def}}{=} [\bar{\Bbbk}_1, \bar{\Bbbk}_2 \dots \bar{\Bbbk}_N]$ for each nonnegative integer N. This is the continued-fraction expansion of ϱ. We can write ϱ_N as $\varrho_N = \mathsf{a}^{(n)}_N/\mathsf{a}^{(d)}_N$ where $\mathsf{a}^{(n)}_N$ and $\mathsf{a}^{(d)}_N$ are relatively prime integers; then $\mathsf{a}^{(d)}_N \nearrow \infty$.

Our main theorem is

Theorem 2.11. *Fix* $\gamma > 0$*. Assume that* ϱ *is such that*

$$\lim_{N\to\infty}\frac{\left(\mathsf{a}^{(d)}_N\right)^{721/14+\gamma}}{\mathsf{a}^{(d)}_{N+1}} = 0. \tag{7}$$

Define

$$\varepsilon_N \overset{def}{=} \left(\frac{1}{\mathsf{a}^{(d)}_N}\right)^{105/4+\gamma/2} \tag{8}$$

for all $N \in \mathbb{N}$*. Then* $\lim_{N\to\infty}\mathbb{P}^{\varepsilon_N,\dagger} = \mathbb{P}^\dagger$*.*

The reason why complicated exponents appear is given in [Sowersc]. We can in fact consider sequences other than that given by (8); again, see [Sowersc].

The crucial idea is the following. Fix $(f,g) \in \mathscr{A}^\dagger$. Define

$$f_{\text{outer}}(x) \overset{\text{def}}{=} f([\mathsf{E}]) + \sum_{\ell\in\Lambda}\dot{f}_\ell(0)\mathsf{H}_{T,\ell}(x)\chi_{\mathsf{D}_\ell}(x)$$

for all $x \in \mathbb{T}$. As we will see, this is a good approximation of $f \circ \mathfrak{m}$ near E. The main technical result is the following

Proposition 2.12. *Assume that* (7) *and* (8) *hold. Then there is a sequence* $(\Psi^{\varepsilon_N};\, N \in \mathbb{N})$ *of functions such that for each* $N \in \mathbb{N}$*,* $\Psi^{\varepsilon_N} + f_{outer}$ *is in* $C^1(\mathbb{T})$ *and is* C^2 *except on a codimension-one subset of* $\mathbb{T}$*, and such that* $\lim_{N\to\infty}\|\Psi^{\varepsilon_N}\|_{C(\mathbb{T})} = 0$ *and*

$$\lim_{N\to\infty}\mathbb{E}^{\varepsilon_N}\left[\left\{\int_{u=s}^t \{(\mathscr{L}^{\varepsilon_N}\Psi^{\varepsilon_N})(X_u) - g([\mathsf{E}])\chi_{\mathsf{E}}(X_u)\}\,du\right\}^-\right] = 0$$

for all $0 \le s < t$*.*

The point of this result is that we can find a small *corrector* which compensates for the loss of smoothness of f_{outer}; as we shall see, we can find such a corrector precisely when the glueing conditions are satisfied.

This is connected with the collection of ideas known as *perturbed test functions*; see Remark 7.4 of [Sowers, 2003] for an explanation. In PDE terminology, we need to be able to construct a boundary layer with the correct smoothness.

Neglecting complications, we should then roughly have that

$$\begin{aligned}M_t^{(f,g)} &\overset{\text{def}}{=} f(X_t^M) - \int_0^t g(X_s^M)ds \\ &= \{(f\circ\mathfrak{m})(X_t) + \Psi^\varepsilon(X_t) \\ &\quad - \int_0^t (\mathscr{L}^\varepsilon(f\circ\mathfrak{m}+\Psi^\varepsilon))(X_u)du\Big\} - \Psi^\varepsilon(X_t) \\ &\quad + \int_0^t \{(\mathscr{L}^\varepsilon(f\circ\mathfrak{m}))(X_u) \\ &\quad -(g\circ\mathfrak{m})(X_u)\}\,\chi_{\mathbb{T}\setminus\mathsf{E}}(X_u)du \\ &\quad + \int_0^t \{(\mathscr{L}^\varepsilon\Psi^\varepsilon)(X_u) - g([\mathsf{E}])\chi_{\mathsf{E}}(X_u)\}\,du.\end{aligned}$$

Since $f\circ\mathfrak{m}$ is constant on E, $\mathscr{L}^\varepsilon(f\circ\mathfrak{m})$ vanishes there. The martingale problem ensures that the term in braces is a martingale. Proposition 2.12 implies that Ψ^ε should be asymptotically negligible. Stochastic averaging should show that the two penultimate lines asymptotically cancel each other. Proposition 2.12 also implies that the last line is asymptotically nonnegative. Thus $M^{(f,g)}$ is asymptotically a submartingale. Since $\mathscr{A}^\dagger$ is a vector space, it contains $(-f,-g)$. Thus the exact same arguments show that $M^{(-f,-g)} = -M^{(f,g)}$ is a submartingale, so $M^{(f,g)}$ is asymptotically a martingale (cf. the submartingale problem of [Stroock and Varadhan, 1971]). This implies the convergence of Theorem 2.11.

Inside the D_ℓ's, standard stochastic averaging calculations give us the corrector function (see [Sowersa]). In E and near the $\partial\mathsf{D}_\ell$'s, we need to *glue*; this is the goal of the next two sections.

3 Relaxation of the Hamiltonian

We now construct a sequence of approximate Hamiltonians on $\mathbb{R}^2$. The salient features of these will be that they generate a flow on $\mathbb{T}$ which *i*) agrees with Ч on the D_ℓ's, for which *ii*) all of the $\mathfrak{x}_\ell$'s are on the same heteroclinic cycle, and *iii*) the flow is periodic on E except on this heteroclinic cycle.

It is not hard to show that there is a disjoint collection $\{\mathcal{D}_\ell,\ \ell\in\Lambda\}$ of open subsets of $\mathbb{T}$ such that for each $\ell\in\Lambda$, $\mathsf{D}_\ell \subset\subset \mathcal{D}_\ell$ and such that $\mathfrak{t}$ is evenly covered over $\mathcal{D}_\ell$ (see [Fulton, 1995] for the definition of even coverings). For each $\ell\in\Lambda$, let $\mathfrak{x}^e_\ell\in\mathbb{R}^2$ be the unique point in $[0,1)^2$ such that $\mathfrak{t}(\mathfrak{x}^e_\ell)=\mathfrak{x}_\ell$, and let $\mathcal{D}^e_\ell$ be the connected component of $\mathfrak{t}^{-1}(\mathcal{D}_\ell)$ which contains $\mathfrak{x}^e_\ell$. Then $\mathfrak{t}|_{\mathcal{D}^e_\ell}$ is a diffeomorphism; we let $\check{\mathfrak{t}}_\ell$ be its inverse. For each $\ell\in\Lambda$, we can find an open subset $\mathcal{D}'_\ell$ of $\mathbb{T}$ such that $\mathsf{D}_\ell\subset\subset\mathcal{D}'_\ell\subset\subset\mathcal{D}_\ell$. Define $\mathsf{D}^e_\ell \overset{\text{def}}{=} \check{\mathfrak{t}}(\mathsf{D}_\ell)$ and $\mathcal{D}'^{,e}_\ell \overset{\text{def}}{=} \check{\mathfrak{t}}(\mathcal{D}'_\ell)$ for all $\ell\in\Lambda$. For each $\ell\in\Lambda$, let $\varpi_\ell\in C^\infty(\mathbb{R}^2;[0,1])$ be such that $\varpi_\ell=1$ on $\mathcal{D}'^{,e}_\ell$ and $\operatorname{supp}\varpi_\ell\subset\mathcal{D}^e_\ell$.

Next, fix $\ell\in\Lambda$. Since ϱ is irrational, there are integers $J^{(1)}_{N,\ell}$ and $J^{(2)}_{N,\ell}$ such that

$$\left|J^{(1)}_{N,\ell}\varrho + J^{(2)}_{N,\ell} - \frac{\mathsf{H}(\mathfrak{x}^e_\ell)}{\omega_2}\right| \le |\boldsymbol{\nu}_N|^2,$$

where $\boldsymbol{\nu}_N \overset{\text{def}}{=} \varrho-\varrho_N$. Set $\mathbf{J}_{N,\ell}\overset{\text{def}}{=}\left(J^{(1)}_{N,\ell}, J^{(2)}_{N,\ell}\right)$ and $\textsc{h}_{N,\ell}\overset{\text{def}}{=}\mathsf{H}\left(\mathfrak{x}^e_\ell-\mathbf{J}_{N,\ell}\right)$. We note that thus

$$\begin{aligned}\textsc{h}_{N,\ell} &= \mathsf{H}(\mathfrak{x}^e_\ell) - J^{(1)}_{N,\ell}\omega_1 - J^{(2)}_{N,\ell}\omega_2 \\ &= \omega_2\left\{\frac{\mathsf{H}(\mathfrak{x}^e_\ell)}{\omega_2} - J^{(1)}_{N,\ell}\varrho - J^{(2)}_{N,\ell}\right\}\end{aligned}$$

so $|\textsc{h}_{N,\ell}|\le|\omega_2||\boldsymbol{\nu}_N|^2$.

Defining $\mathbf{e}_1\overset{\text{def}}{=}(1,0)\in\mathbb{R}^2$, set

$$\begin{aligned}\hat{\mathsf{H}}_N(x) \overset{\text{def}}{=} \omega_2\Big\{&-\langle x,\mathbf{e}_1\rangle_{\mathbb{R}^2} \\ &+\sum_{\substack{K\in\mathbb{Z}^2\\ \ell\in\Lambda}}\varpi_\ell(x-K)\,\langle x-K-\mathbf{J}_{N,\ell},\mathbf{e}_1\rangle_{\mathbb{R}^2}\Big\} \\ &-\sum_{\substack{K\in\mathbb{Z}^2\\ \ell\in\Lambda}}\varpi_\ell(x-K)\frac{\textsc{h}_{N,\ell}}{\boldsymbol{\nu}_N}. \qquad x\in\mathbb{R}^2\end{aligned}$$

We then define the perturbed Hamiltonian $\mathsf{H}_N\overset{\text{def}}{=}\mathsf{H}+\boldsymbol{\nu}_N\hat{\mathsf{H}}_N$ and the perturbed frequency vector $\boldsymbol{\omega}_N\overset{\text{def}}{=}\boldsymbol{\omega}-\omega_2\mathbf{e}_1\boldsymbol{\nu}_N=(\omega_2\varrho_N,\omega_2)$.

Lemma 3.1. *Fix a positive integer N. For any $x\in\mathbb{R}^2$ and $K\in\mathbb{Z}^2$, $\mathsf{H}_N(x+K)=\mathsf{H}_N(x)+\langle\boldsymbol{\omega}_N,K\rangle_{\mathbb{R}^2}$. Secondly, $\mathsf{H}_N-\mathsf{H}$ is locally constant on $\mathfrak{t}^{-1}(\mathsf{D}_\ell)$ and furthermore $\mathsf{H}_N(\mathfrak{x}^e_\ell-\mathbf{J}_{N,\ell})=0$.*

Analogously to (1), we now define the vector fields $\text{Ч}_{e,N}$ and $\hat{\text{Ч}}_{e,N}$ on $\mathbb{R}^2$ by

$$\begin{aligned}(\text{Ч}_{e,N}\varphi)(x) &\overset{\text{def}}{=} \left(\frac{\partial\mathsf{H}_N}{\partial x_2}\frac{\partial\varphi}{\partial x_1} - \frac{\partial\mathsf{H}_N}{\partial x_1}\frac{\partial\varphi}{\partial x_2}\right)(x) \\ (\hat{\text{Ч}}_{e,N}\varphi)(x) &\overset{\text{def}}{=} \left(\frac{\partial\hat{\mathsf{H}}_N}{\partial x_2}\frac{\partial\varphi}{\partial x_1} - \frac{\partial\hat{\mathsf{H}}_N}{\partial x_1}\frac{\partial\varphi}{\partial x_2}\right)(x)\end{aligned}$$

for all $\varphi\in C^\infty(\mathbb{R}^2)$ and $x=(x_1,x_2)\in\mathbb{R}^2$, and similarly to the definition of Ч preceding Definition 2.3, we then define vector fields Ч_N and $\hat{\text{Ч}}_N$ on $\mathbb{T}$ by requiring that $(\text{Ч}_N\varphi)(\mathfrak{t}(x))=(\text{Ч}_{e,N}(\varphi\circ\mathfrak{t}))(x)$ and $(\hat{\text{Ч}}_N\varphi)(\mathfrak{t}(x))=(\hat{\text{Ч}}_{e,N}(\varphi\circ\mathfrak{t}))(x)$ for all $\varphi\in C^\infty(\mathbb{T})$ and $x\in\mathbb{R}^2$. Then $\text{Ч}_N=\text{Ч}+\boldsymbol{\nu}_N\hat{\text{Ч}}_N$ for all $N\in\mathbb{N}$.

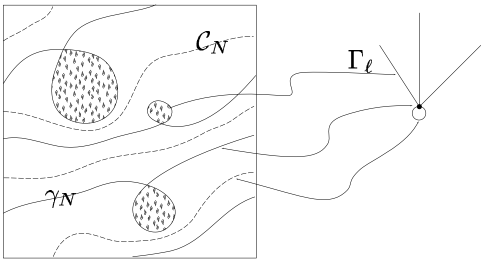

Figure 3. γ_N and $\mathcal{C}_N$

Lemma 3.2. *For each $\ell \in \Lambda$ and $N \in \mathbb{N}$, $Ч_N = Ч$ on $\bigcup_{\ell\in\Lambda} \mathsf{D}_\ell$. Secondly, for N large enough, $\{x \in \mathbb{T} : Ч_N(x) = 0\} = \{x \in \mathbb{T} : Ч(x) = 0\}$.*

Let $\mathfrak{z}^N$ be the flow of diffeomorphisms of $\mathbb{T}$ defined by

$$\begin{aligned} \dot{\mathfrak{z}}^N_t(x) &\overset{\text{def}}{=} Ч_N(\mathfrak{z}^N_t(x)) \quad t \ge 0 \\ \mathfrak{z}^N_0(x) &= x. \end{aligned} \qquad x \in \mathbb{T}$$

By Lemma 3.2, $\mathfrak{z}^N_t = \mathfrak{z}_t$ on $\bigcup_{\ell\in\Lambda} \mathsf{D}_\ell$ for all $t \in \mathbb{R}$. Let now $\sim_N$ denote chain equivalence with respect to $\mathfrak{z}^N$, and let $[x]_N \overset{\text{def}}{=} \{x' \in \mathbb{T} : x' \sim_N x\}$ for all $x \in \mathbb{T}$. Since $\mathfrak{z}^N = \mathfrak{z}$ in the D_ℓ's and $\mathfrak{z}$ is periodic in D_ℓ, all points in all of the D_ℓ's are chain-recurrent and $\{[x]_N : x \in \mathsf{D}^\circ_\ell\} \simeq (0, \mathsf{H}_{T,\ell}(\mathfrak{p}_\ell)]$ for $\ell \in \Lambda_P$ and $\{[x]_N : x \in \mathsf{D}^\circ_\ell\} \simeq [-\mathsf{H}_{T,\ell}(\mathfrak{p}_\ell), 0)$ for $\ell \in \Lambda_W$. We next claim that $\{[x]_N;\ x \in \mathsf{E}\}$ is a *circle*. Define $\mathsf{r}_N \overset{\text{def}}{=} \frac{\omega_2}{\mathsf{a}^{(d)}_N}$ for all $N \in \mathbb{N}$. For all $K = (k_1, k_2) \in \mathbb{Z}^2$,

$$\begin{aligned} \mathsf{H}_N(\mathfrak{x}^e_\ell + K) &= \mathsf{H}_N(\mathfrak{x}^e_\ell - \mathbf{J}_{N,\ell}) \\ &\quad + \omega_2 \varrho_N (J^{(1)}_{N,\ell} + k_1) + \omega_2 (J^{(2)}_{N,\ell} + k_2) \\ &= \mathsf{r}_N \left\{ \mathsf{a}^{(n)}_N (J^{(1)}_{N,\ell} + k_1) + \mathsf{a}^{(d)}_N (J^{(2)}_{N,\ell} + k_2) \right\}; \end{aligned}$$

hence $\mathsf{H}_N(\mathfrak{x}^e_\ell + K) \in \mathsf{r}_N\mathbb{Z}$.

For each $N \in \mathbb{N}$, next define the codimension-one sets

$$\begin{aligned} \gamma_N &\overset{\text{def}}{=} \mathsf{t}\left\{x \in \mathsf{t}^{-1}(\overline{\mathsf{E}}) : \mathsf{H}_N(x) \in \mathbb{Z}\mathsf{r}_N\right\} \\ \mathcal{C}_N &\overset{\text{def}}{=} \mathsf{t}\left\{x \in \mathsf{t}^{-1}(\mathsf{E}) : \mathsf{H}_N(x) \in \left(\mathbb{Z} + \frac{1}{2}\right) \mathsf{r}_N\right\}; \end{aligned}$$

see Figure 3.

Lemma 3.3. *For large enough $N \in \mathbb{N}$, $\gamma_N \cap (\mathsf{E} \cup \mathfrak{X})$ and γ_N are path-connected and all orbits of $\mathfrak{z}^N$ in $\mathsf{E} \setminus \gamma_N$ are periodic.*

Thus γ_N is the unique heteroclinic cycle of $\mathfrak{z}^N$ in $\mathbb{T}$, and furthermore all points in E are chain-recurrent under $\mathfrak{z}^N$.

We next claim that there is a local Hamiltonian in $\mathbb{T} \setminus \mathcal{C}_N$.

Lemma 3.4. *For each $N \in \mathbb{N}$, there is an $\mathsf{H}^{loc}_N \in B(\mathbb{T})$ such that H^{loc}_N is C^∞ on $\mathbb{T} \setminus \mathcal{C}_N$, and such that*

$$\mathsf{H}^{loc}_N(\mathsf{t}(x)) = \mathsf{H}_N(x) - \left\lfloor \frac{\mathsf{H}_N(x)}{r_N} + \frac{1}{2} \right\rfloor r_N$$

for all $x \in \mathsf{t}^{-1}(\overline{\mathsf{E}})$ and such that $\mathsf{H}^{loc}_N(x) = \mathsf{H}_{T,\ell}(x)$ for all $x \in \mathsf{D}^\circ_\ell$ for all $\ell \in \Lambda$.

This tells us that $\{[x]_N;\ x \in \mathsf{E}\}$ is the circle of circumference r_N. Since γ_N contains the $\partial\mathsf{D}_\ell$'s, which are limit points of the D_ℓ's, we have that $\{[x]_N;\ x \in \mathbb{T}\}$ is a *whiskered circle*, where one whisker is attached for each trap. Define $\mathcal{W} \overset{\text{def}}{=} (-1/2, 1/2)$ we thus have that $\mathsf{H}^{\text{loc}}_N : \mathsf{E} \setminus \mathcal{C}_N = \mathsf{r}_N\mathcal{W}$.

Remark 3.5. *By removing $\mathcal{C}_N$, we have approximated E by a long and thin ribbon. More precisely, the width of $\mathsf{E} \setminus \mathcal{C}_N$, as parametrized by H^{loc}_N, is r_N (the width of $r_N\mathcal{W}$). Since the area of $\mathsf{E}\setminus\mathcal{C}_N$ is equal to that of E (we have removed only the one-dimensional manifold $\mathcal{C}_N$), the "length" of the ribbon should be of order $1/r_N$.*

4 Dominant Analysis

We now want to use the machinery of [Sowers, 2005] to "glue" at γ_N. We have two questions to answer. First of all, if we start on γ_N, what is the relative likelihood of going into each of the D_ℓ's, versus the relative likelihood of going back into E? Secondly, if we start in E, how long does it (on average) take to get to γ_N, where we are back to the first question. The first question is one of *glueing*, and the second question should involve some sort of *Poisson* equation (recall that one usually studies occupation times by solving Poisson equations). The combination of the likelihood of going back into E and then returning to γ_N at a later time should give the stickiness coefficient, and the relative likelihood of going into the different D_ℓ's should give the $\textsc{g}_\ell$'s. Lemma 4.4 is the result which brings all of this into focus.

The long-term behavior of $\mathfrak{z}$ in E should (and will) be important in our calculations. For $\varphi \in C(\mathsf{E})$, define

$$(\mathcal{A}\varphi)([\mathsf{E}]) \overset{\text{def}}{=} \frac{1}{\mathscr{H}^2(\mathsf{E})} \int_{z\in\mathsf{E}} \varphi(z)\mathscr{H}^2(dz);$$

then for any $x \in \mathsf{E}$,

$$\lim_{T\to\infty} \frac{1}{T} \int_{s=0}^{T} \varphi(\mathfrak{z}_s(x))ds = (\mathcal{A}\varphi)([\mathsf{E}]);$$

this is an extension of (4).

Let's first focus on γ_N. Recall that [Sowers, 2005] gives us a solvability condition for correcting for a *smoothness defect* in certain types of test functions near a homoclinic orbit. Let's write down the function we wish to correct. Define $\boldsymbol{\sigma}_N \in C^\infty(\mathbb{T})$ by requiring that

$\boldsymbol{\sigma}_N(\mathsf{t}(x)) = \langle d\mathsf{H}_N, d\mathsf{H}_N\rangle(x)$ for all $x \in \mathbb{R}^2$. Recall the $\textsc{g}_\ell$'s of (6), and define

$$\tilde{\textsc{g}}_N \overset{\text{def}}{=} \int_{z\in\gamma_N\cap\mathsf{E}} \frac{\boldsymbol{\sigma}_N(z)}{\|\text{Ч}_N(z)\|}\mathscr{H}^1(dz). \tag{9}$$

Local calculations show that $\lim_{N\to\infty}\mathsf{r}_N\tilde{\textsc{g}}_N = \daleth$, where $\daleth \overset{\text{def}}{=} (\mathcal{A}\boldsymbol{\sigma})([\mathsf{E}])\mathscr{H}^2(\mathsf{E})$.

Now define $F_N : \mathbb{T}\to\mathbb{R}$ as follows. If $x\in\mathsf{E}$ and $\mathsf{H}^{\text{loc}}_N(x) > 0$, define

$$F_N(x) \overset{\text{def}}{=} \frac{\mathsf{r}_N}{\daleth}\left\{\sum_{\ell\in\Lambda_W}\textsc{g}_\ell\dot{f}_\ell(0)\right\}\mathsf{H}^{\text{loc}}_N(x),$$

while if $x\in\mathsf{E}$ and $\mathsf{H}^{\text{loc}}_N(x) < 0$, define

$$F_N(x) \overset{\text{def}}{=} \frac{\mathsf{r}_N}{\daleth}\left\{\sum_{\ell\in\Lambda_P}\textsc{g}_\ell\dot{f}_\ell(0)\right\}\mathsf{H}^{\text{loc}}_N(x);$$

if $x\in\mathsf{D}_\ell$, define $F_N(x) \overset{\text{def}}{=} \dot{f}_\ell(0)\mathsf{H}^{\text{loc}}_N(x)$. This captures, to first order, the behavior of $f\circ\mathfrak{m}$ on the D_ℓ's near $\partial\mathsf{E}$. We claim that we can find a small corrector function which compensates for the loss of smoothness of F_N across γ_N. To see this, we first decompose $\mathbb{T}\setminus\gamma_N$ into connected components (see [Sowers, 2005]). Note that

$$\begin{aligned}
&\{x\in\mathbb{T} : \mathsf{H}^{\text{loc}}_N(x) > 0\}\\
&\qquad = \{x\in\mathsf{E} : \mathsf{H}^{\text{loc}}_N(x) > 0\} \cup \bigcup_{\ell\in\Lambda_P}\mathsf{D}_\ell\\
&\{x\in\mathbb{T} : \mathsf{H}^{\text{loc}}_N(x) < 0\}\\
&\qquad = \{x\in\mathsf{E} : \mathsf{H}^{\text{loc}}_N(x) < 0\} \cup \bigcup_{\ell\in\Lambda_W}\mathsf{D}_\ell,
\end{aligned}$$

both of these being disjoint unions. We then have that

$$\begin{aligned}
&\gamma_N\cap\partial\{x\in\mathbb{T} : \mathsf{H}^{\text{loc}}_N(x) > 0\}\\
&\qquad = \left\{(\gamma_N\cap\mathsf{E}) \cup \bigcup_{\ell\in\Lambda_W}\partial\mathsf{D}_\ell\right\} \cup \bigcup_{\ell\in\Lambda_P}\partial\mathsf{D}_\ell\\
&\gamma_N\cap\partial\{x\in\mathbb{T} : \mathsf{H}^{\text{loc}}_N(x) < 0\}\\
&\qquad = \left\{(\gamma_N\cap\mathsf{E}) \cup \bigcup_{\ell\in\Lambda_P}\partial\mathsf{D}_\ell\right\} \cup \bigcup_{\ell\in\Lambda_W}\partial\mathsf{D}_\ell.
\end{aligned}$$

We thus have that

$$\begin{aligned}
&\sum_{\ell\in\Lambda_P}\textsc{g}_\ell\dot{f}_\ell(0)\\
&\quad + \left\{\tilde{\textsc{g}}_N + \sum_{\ell\in\Lambda_W}\textsc{g}_\ell\right\}\left\{\frac{\mathsf{r}_N}{\daleth}\sum_{\ell\in\Lambda_W}\textsc{g}_\ell\dot{f}_\ell(0)\right\}\\
&\quad \approx \sum_{\ell\in\Lambda_P}\textsc{g}_\ell\dot{f}_\ell(0) + \sum_{\ell\in\Lambda_W}\textsc{g}_\ell\dot{f}_\ell(0)\\
&\quad \approx \sum_{\ell\in\Lambda_W}\textsc{g}_\ell\dot{f}_\ell(0)\\
&\quad + \left\{\tilde{\textsc{g}}_N + \sum_{\ell\in\Lambda_P}\textsc{g}_\ell\right\}\left\{\frac{\mathsf{r}_N}{\daleth}\sum_{\ell\in\Lambda_P}\textsc{g}_\ell\dot{f}_\ell(0)\right\}.
\end{aligned}$$

This should allow us to glue as in [Sowers, 2005]. To make this precise, define

$$\begin{aligned}
\mho^+_N &\overset{\text{def}}{=} \frac{\sum_{\ell\in\Lambda_W}\textsc{g}_\ell\dot{f}_\ell(0)}{\sum_{\ell\in\Lambda_P}\textsc{g}_\ell}\left\{1 - \frac{\mathsf{r}_N\tilde{\textsc{g}}_N}{\gimel}\right.\\
&\qquad\left. - \frac{\mathsf{r}_N}{\gimel}\sum_{\ell\in\Lambda_W}\textsc{g}_\ell\right\}\\
\mho^-_N &\overset{\text{def}}{=} \frac{\sum_{\ell\in\Lambda_P}\textsc{g}_\ell\dot{f}_\ell(0)}{\sum_{\ell\in\Lambda_W}\textsc{g}_\ell}\left\{1 - \frac{\mathsf{r}_N\tilde{\textsc{g}}_N}{\gimel}\right.\\
&\qquad\left. - \frac{\mathsf{r}_N}{\gimel}\sum_{\ell\in\Lambda_P}\textsc{g}_\ell\right\}\\
\hat{F}_N(x) &\overset{\text{def}}{=} F_N(x) + \mho^+_N\sum_{\ell\in\Lambda_P}\mathsf{H}^{\text{loc}}_N(x)\chi_{\mathsf{D}_\ell}(x)\\
&\qquad + \mho^-_N\sum_{\ell\in\Lambda_W}\mathsf{H}^{\text{loc}}_N(x)\chi_{\mathsf{D}_\ell}(x).
\end{aligned}$$

Then we exactly have that

$$\begin{aligned}
&\sum_{\ell\in\Lambda_P}\textsc{g}_\ell\{\dot{f}_\ell(0) + \mho^+_N\}\\
&\quad + \left\{\tilde{\textsc{g}}_N + \sum_{\ell\in\Lambda_W}\textsc{g}_\ell\right\}\left\{\frac{\mathsf{r}_N}{\daleth}\sum_{\ell\in\Lambda_W}\textsc{g}_\ell\dot{f}_\ell(0)\right\}\\
&\quad = \sum_{\ell\in\Lambda_P}\textsc{g}_\ell\dot{f}_\ell(0) + \sum_{\ell\in\Lambda_W}\textsc{g}_\ell\dot{f}_\ell(0)\\
&\quad = \sum_{\ell\in\Lambda_W}\textsc{g}_\ell\{\dot{f}_\ell(0) + \mho^-_N\}\\
&\quad + \left\{\tilde{\textsc{g}}_N + \sum_{\ell\in\Lambda_P}\textsc{g}_\ell\right\}\left\{\frac{\mathsf{r}_N}{\daleth}\sum_{\ell\in\Lambda_P}\textsc{g}_\ell\dot{f}_\ell(0)\right\}.
\end{aligned}$$

Using the asymptotics of $\tilde{\textsc{g}}_N$ (given right after (9)) and the fact that $\lim_{N\to\infty}\mathsf{r}_N = 0$, we have that $\lim_{N\to\infty}|\mho^+_N| = \lim_{N\to\infty}|\mho^+_N| = 0$. We now can carry out the main task of *glueing*.

Proposition 4.1. *There is a collection* $\{\tilde{\Psi}^{\varepsilon_N}; N\in\mathbb{N}\}$ *of functions such that for all* $N\in\mathbb{N}$, $\tilde{\Psi}^{\varepsilon_N} + \hat{F}_N$ *is in* $C^1(\mathbb{T}\setminus\mathcal{C}_N)$ *and is* C^2 *except on* $\mathcal{C}_N\cup\gamma_N$, *and such that* $\lim_{N\to\infty}\|\tilde{\Psi}^{\varepsilon_N}\|_{C(\mathbb{T})} = 0$ *and*

$$\varliminf_{N\to\infty}\inf_{x\in\mathbb{T}\setminus(\mathcal{C}_N\cup\gamma_N)}(\mathscr{L}^{\varepsilon_N}\tilde{\Psi}^{\varepsilon_N})(x) \geq 0$$

The proof of this is fairly detailed, and relies upon the ideas of [Sowersb], which in turn depends upon the ideas of [Sowers, 2005]; see [Sowersc].

Let's now look inside E, i.e., at the issue of the Poisson equation. Here we want to solve the PDE $\mathscr{L}^\varepsilon u \approx g([\mathsf{E}])$ such that $\hat{F}_N$ captures the nonsmooth behavior of u near γ_N. Since γ_N and $\hat{F}_N$ are both given in terms of $\mathsf{H}^{\text{loc}}_N$, let's look at the effect of $\mathscr{L}^\varepsilon$ on functions of $\mathsf{H}^{\text{loc}}_N$. Define $\xi_N \in C^\infty_p(\mathbb{T})$ by requiring that

$$\xi_N(\mathfrak{t}(x)) \stackrel{\text{def}}{=} \frac{1}{\boldsymbol{\nu}_N}(\bar{\nabla}_e \mathsf{H}, \nabla_e \mathsf{H}_N)_e(x)$$

for all $x \in \mathbb{R}^2$. If $u(x) = U(\mathsf{H}^{\text{loc}}_N(x)/\mathsf{r}_N)$ (this scaling turns $\mathsf{E} \setminus \mathcal{C}_N$ into a reference strip of *unit* width) on some open subset $\mathcal{O}$ of $\mathsf{E} \setminus \mathcal{C}_N$, where $U \in C^2(\mathbb{R})$, then (note that $\xi_N \circ \mathfrak{t} = (\bar{\nabla}_e \mathsf{H}, \nabla_e \hat{\mathsf{H}}_N)_e$)

$$\begin{aligned}(\mathscr{L}^\varepsilon u)(x) = \frac{\boldsymbol{\nu}_N \xi_N(x)}{\varepsilon^2 \mathsf{r}_N}\dot{U}\left(\frac{\mathsf{H}^{\text{loc}}_N(x)}{\mathsf{r}_N}\right) \\ + \frac{\boldsymbol{\sigma}_N(x)}{2\mathsf{r}_N^2}\ddot{U}\left(\frac{\mathsf{H}^{\text{loc}}_N(x)}{\mathsf{r}_N}\right) \\ + \frac{\boldsymbol{\beta}_N(x)}{\mathsf{r}_N}\dot{U}\left(\frac{\mathsf{H}^{\text{loc}}_N(x)}{\mathsf{r}_N}\right)\end{aligned} \quad (10)$$

for all $x \in \mathcal{O}$, where $\boldsymbol{\beta}_N \in C^\infty(\mathbb{T})$ is defined by requiring that $\boldsymbol{\beta}_N(\mathfrak{t}(x)) = (\mathscr{L}\mathsf{H}_N)(x)$ for all $x \in \mathbb{R}^2$. The first two terms are the dominant ones.

The theory of *averaging* tells us that we should replace the coefficients in (10) by *constants*. The operator $\mathscr{L}^\varepsilon$ generates a drift of size $1/\varepsilon^2$ in the direction of Ч; keeping in mind that we are using $Ч_N$ as an approximation of Ч, we have a drift along γ_N of size $1/\varepsilon^2$. Comparing this to the drift and diffusion in the $\nabla \mathsf{H}^{\text{loc}}_N$-direction, we should have a separation of scales, and be able to replace ξ_N and $\boldsymbol{\sigma}_N$ by their averages over the orbits of $\mathfrak{z}^N$ (which is an approximation of the average with respect to $\mathfrak{z}$).

Lemma 4.2. *There is a $K > 0$ such that*

$$\left|(\mathcal{A}\xi_N)([\mathsf{E}]) - \frac{\omega_2}{\mathscr{H}^2(\mathsf{E})}\right| \le K|\boldsymbol{\nu}_N|$$

for all $N \in \mathbb{N}$.

For all $N \in \mathbb{N}$ and $\varepsilon \in (0,1)$, define now

$$\mu_{N,\varepsilon} \stackrel{\text{def}}{=} \frac{(\omega_2/\mathscr{H}^2(\mathsf{E}))}{(\mathcal{A}\boldsymbol{\sigma})([\mathsf{E}])}\frac{\nu_N \mathsf{r}_N}{\varepsilon^2} = \frac{\omega_2}{\daleth}\frac{\nu_N \mathsf{r}_N}{\varepsilon^2}.$$

Then $\mu_{N,\varepsilon}$ is the ratio of the asymptotic averaged drift to asymptotic averaged diffusion coefficients in (10).

Let's now get back to our Poisson equation in E. For $N \in \mathbb{N}$ and $\varepsilon \in (0,1)$, define

$$\mathsf{u}^{N,\varepsilon}_P(h) \stackrel{\text{def}}{=} \frac{1}{\mu_{N,\varepsilon}}\left\{h - \frac{1-\exp[-2\mu_{N,\varepsilon}h]}{1-\exp[-2\mu_{N,\varepsilon}]}\right\}$$

for all $h \in \mathbb{R}$. The importance of $\mathsf{u}^{N,\varepsilon}_P$ is the following.

Lemma 4.3. *For each $\varepsilon \in (0,1)$ and $N \in \mathbb{N}$, $u^{N,\varepsilon}_P \in C^2([0,1])$ and*

$$\mu_{N,\varepsilon}\dot{u}^{N,\varepsilon}_P(h) + \frac{1}{2}\ddot{u}^{N,\varepsilon}_P(h) = 1, \qquad h \in (0,1)$$

$u^{N,\varepsilon}_P(1) = u^{N,\varepsilon}_P(0) = 0$, and $\dot{u}^{N,\varepsilon}_P(1) - \dot{u}^{N,\varepsilon}_P(0) = 2$. Furthermore, there is a $K > 0$ such that $\left|u^{N,\varepsilon}_P(h)\right| \le K$ and $\left|\dot{u}^{N,\varepsilon}_P(h)\right| \le K$ for all $h \in [0,1]$, $\varepsilon \in (0,1)$, and $N \in \mathbb{N}$.

Let's now define two constants; set

$$\begin{aligned}\hat{\mathsf{u}}^{N,\varepsilon}_+ &\stackrel{\text{def}}{=} \left\{\frac{g([\mathsf{E}])\dot{\mathsf{u}}^{N,\varepsilon}_P(1)}{(\mathcal{A}\sigma)([\mathsf{E}])} - \frac{1}{\beth}\sum_{\ell \in \Lambda_P} {}_{\mathrm{G}}\ell \dot{f}_\ell(0)\right\}\mathsf{r}_N \\ &\quad + \mho^+_N \\ \hat{\mathsf{u}}^{N,\varepsilon}_- &\stackrel{\text{def}}{=} \left\{\frac{g([\mathsf{E}])\dot{\mathsf{u}}^{N,\varepsilon}_P(0)}{(\mathcal{A}\sigma)([\mathsf{E}])} - \frac{1}{\beth}\sum_{\ell \in \Lambda_W} {}_{\mathrm{G}}\ell \dot{f}_\ell(0)\right\}\mathsf{r}_N \\ &\quad + \mho^-_N\end{aligned}$$

for all $\varepsilon \in (0,1)$ and $N \in \mathbb{N}$. In light of the asymptotics of $\mho^\pm_N$ (given just prior to Proposition 4.1) and the bound on $\dot{\mathsf{u}}^{N,\varepsilon}_P$ in Lemma 4.3, we have that $\lim_{\substack{N\to\infty \\ \varepsilon\to 0}}\left|\hat{\mathsf{u}}^{N,\varepsilon}_+\right| = \lim_{\substack{N\to\infty \\ \varepsilon\to 0}}\left|\hat{\mathsf{u}}^{N,\varepsilon}_-\right| = 0$.

Finally, define

$$\begin{aligned}\mathsf{U}^{N,\varepsilon}_P(x) &\stackrel{\text{def}}{=} \sum_{\ell \in \Lambda_P}\hat{\mathsf{u}}^{N,\varepsilon}_+ \mathsf{H}_{T,\ell}(x)\chi_{\mathsf{D}_\ell}(x) \\ &\quad + \sum_{\ell \in \Lambda_W}\hat{\mathsf{u}}^{N,\varepsilon}_- \mathsf{H}_{T,\ell}(x)\chi_{\mathsf{D}_\ell}(x) \\ &\quad + \frac{g([\mathsf{E}])}{(\mathcal{A}\sigma)([\mathsf{E}])}\mathsf{r}_N^2 \mathsf{u}^{N,\varepsilon}_P\left(\iota\left(\frac{\mathsf{H}^{\text{loc}}_N(x)}{\mathsf{r}_N}\right)\right)\chi_{\mathsf{E}}(x)\end{aligned}$$

for all $x \in \mathbb{T}$, $\varepsilon \in (0,1)$, and $N \in \mathbb{N}$ (where ι is as was defined after Proposition 2.10).

It is fairly easy to see that for each $\varepsilon \in (0,1)$ and $N \in \mathbb{N}$, $\mathsf{U}^{N,\varepsilon}_P$ is in $C(\mathbb{T})$ and is C^∞ on $\mathbb{T} \setminus \gamma_N$. We also have

Lemma 4.4. *For each $\varepsilon \in (0,1)$ and $N \in \mathbb{N}$, $U^{N,\varepsilon}_P - \hat{F}_N + f_{outer}$ is C^1 at γ_N.*

Proof. The important question is differentiability at points of $\gamma_N \cap \mathsf{E}$. We need to show that

$$\begin{aligned}\frac{g([\mathsf{E}])}{(\mathcal{A}\sigma)([\mathsf{E}])}\mathsf{r}_N \dot{u}^{N,\varepsilon}_P(1) - \frac{\mathsf{r}_N}{\daleth}\sum_{\ell \in \Lambda_P} {}_{\mathrm{G}}\ell \dot{f}_\ell(0) \\ = \frac{g([\mathsf{E}])}{(\mathcal{A}\sigma)([\mathsf{E}])}\mathsf{r}_N \dot{u}^{N,\varepsilon}_P(0) - \frac{\mathsf{r}_N}{\daleth}\sum_{\ell \in \Lambda_W} {}_{\mathrm{G}}\ell \dot{f}_\ell(0);\end{aligned}$$

the first term is the transversal derivative of $\mathsf{U}_P^{N,\varepsilon} - \hat{F}_N + f_{\text{outer}}$ when we approach $\gamma_N \cap \mathsf{E}$ from the direction where $\mathsf{H}_N^{\text{loc}} < 0$, and the second term is the transversal derivative of $\mathsf{U}_P^{N,\varepsilon} - \hat{F}_N + f_{\text{outer}}$ when we approach $\gamma_N \cap \mathsf{E}$ from the direction where $\mathsf{H}_N^{\text{loc}} > 0$. In light of the definition of $\urcorner$ (given right after (9)), this is equivalent to showing that

$$g([\mathsf{E}]) \left\{ \dot{u}_P^{N,\varepsilon}(1) - \dot{u}_P^{N,\varepsilon}(0) \right\} = \frac{1}{\mathscr{H}^2(\mathsf{E})} \left\{ \sum_{\ell \in \Lambda_P} \mathrm{G}_\ell \dot{f}_\ell(0) - \sum_{\ell \in \Lambda_W} \mathrm{G}_\ell \dot{f}_\ell(0) \right\},$$

and substituting the claim of Lemma 4.3 involving $\dot{\mathsf{u}}_P^{N,\varepsilon}(0)$ and $\mathsf{u}_P^{N,\varepsilon}(1)$, this is in turn equivalent to showing that $2g([\mathsf{E}])\mathscr{H}^2(\mathsf{E}) = \sum_{\ell \in \Lambda_P} \mathrm{G}_\ell \dot{f}_\ell(0) - \sum_{\ell \in \Lambda_W} \mathrm{G}_\ell \dot{f}_\ell(0)$. **This is exactly the glueing condition**. □

Of course to use this, we need to show that it is good enough that $\mathsf{U}_P^{N,\varepsilon}$ satisfies the appropriate *averaged* ODE in E; we can use [Sowersa] to bound these averaging errors. Since $\mathsf{U}_P^{N,\varepsilon_N} + \tilde{\Psi}^{\varepsilon_N} + f_{\text{outer}} = \{\mathsf{U}_P^{N,\varepsilon_N} - \hat{F}_N + f_{\text{outer}}\} + \{\tilde{\Psi}^{\varepsilon_N} + \hat{F}_N\}$, the above lemma means that $\mathsf{U}_P^{N,\varepsilon} + \tilde{\Psi}^{\varepsilon_N} + f_{\text{outer}}$ is C^1 at γ_N. This is the main idea in the construction of the Ψ^{ε_N}'s of Proposition 2.12 near E.

5 Conclusions

We have studied the behavior of a randomly-perturbed 2-dimensional pseudoperiodic system. The time scale of interest is of the order such that we can see transitions between the different parts of the phase space identified by Arnold; i.e., traps and an ergodic class. In a weak sense, we have shown that this system converges to a diffusion on a graph with a "delay" (stickiness) at the vertex. We hope that these techniques will be of use in the general ongoing investigation into the effect of small noise upon dynamical systems.

6 Acknowledgements

The author would like to thank Mark Freidlin for recommending this problem. This research was supported by NSF DMS 0305925.

References

Arnol'd, V. I. [1991] *Topological and ergodic properties of closed* 1*-forms with incommensurable periods*, Funktsional. Anal. i Prilozhen. **25**, no. 2, 1–12, 96.

Conley, C. [1978] *Isolated invariant sets and the Morse index*, CBMS Regional Conference Series in Mathematics, vol. 38, American Mathematical Society, Providence, R.I.

Ethier, S. N. and Kurtz, T. G. [1986] *Markov Processes: Characterization and Convergence*, John Wiley & Sons Inc., New York. 1986.

Freidlin, M. [1996] *Markov processes and differential equations: asymptotic problems*, Birkhäuser Verlag, Basel.

Freidlin, M. and Weber, M. [1998] *Random perturbations of nonlinear oscillators*, Ann. Probab. **26**, no. 3, 925–967.

Freidlin, M. and Weber, M. [1999] *A remark on random perturbations of the nonlinear pendulum*, Ann. Appl. Probab. **9**, no. 3, 611–628.

Freidlin, M. I. and Wentzell, A. D. [1994] *Random perturbations of Hamiltonian systems*, Mem. Amer. Math. Soc. **109**, no. 523, viii+82.

Fulton, W. [1995] *Algebraic topology*, Springer-Verlag, New York.

Goresky, M. and MacPherson, R. [1998] *Stratified Morse Theory*, Springer-Verlag, Berlin.

Harrison, J. M. and Lemoine, A. J. [1981] *Sticky Brownian motion as the limit of storage processes*, J. Appl. Probab. **18**, no. 1, 216–226.

Novikov, A., Papanicolaou, G., and Ryzhik, L. *Boundary layers for cellular flows at high péclet numbers*, Comm. Pure Appl. Math., to appear.

Robinson, C. [1999] *Dynamical systems*, second ed., Studies in Advanced Mathematics, CRC Press, Boca Raton, FL.

Sinaĭ, Y. G. and Khanin, K. M. [1992] *Mixing of some classes of special flows over rotations of the circle*, Funktsional. Anal. i Prilozhen. **26**, no. 3, 1–21.

Sowers, R. B. [A] *On the breakdown and regularity of an approximate solution of the cohomological equation (of averaging) in 2-dimensional systems*, Submitted.

Sowers, R. B. [B] *Bounds on corrector functions for diffusively-perturbed transport equations near a (long) heteroclinic cycle*, Submitted.

Sowers, R. B. [C] *Random perturbations of two-dimensional pseudoperiodic flows*, submitted.

Sowers, R. B. [2002], *Stochastic averaging with a flattened Hamiltonian: a Markov process on a stratified space* (*a whiskered sphere*), Trans. Amer. Math. Soc. **354**, no. 3, 853–900.

Sowers, R. B. [2003] *Stochastic averaging near a homoclinic orbit with multiplicative noise*, Stoch. Dyn. **3**, no. 3, 299–391.

Sowers, R. B. [2005] *A boundary layer theory for diffusively perturbed transport around a heteroclinic cycle*, Comm. Pure Appl. Math. **58**, no. 1, 30–84.

Stroock, D. W. and Varadhan, S. R. S. [1971] *Diffusion processes with boundary conditions*, Comm. Pure Appl. Math. **24**, 147–225.

Stroock, D. W. and Varadhan, S. R. S. [1979] *Multidimensional Diffusion Processes*, Springer-Verlag, Berlin.