ON STOCHASTIC BEHAVIOR OF PERTURBED HAMILTONIAN SYSTEMS

MICHAEL BRIN AND MARK FREIDLIN

ABSTRACT. We consider deterministic perturbations $\ddot{q}^\varepsilon(t)+F'(q^\varepsilon(t)) = \varepsilon b(\dot{q}^\varepsilon(t), q^\varepsilon(t))$ of an oscillator $\ddot{q}+F'(q) = 0$, $q \in \mathbb{R}^1$. Assume that $\lim_{|q|\to\infty} F(q) = \infty$ and that $F'(q)$ has a finite number of nondegenerate zeros. For a generic F, if $\operatorname{div} v(x) < 0$ (an analogue of friction), then typical orbits are attracted to points where F has a local minimum. For $0 < \varepsilon \ll 1$, the equilibrium to which the trajectory is attracted is "random". To study this randomness which is caused by the sensitive behavior of trajectories near the saddle points we consider the graph Γ homeomorphic to the space of connected components of the level sets of the Hamiltonian $H(p, q) = p^2/2 + F(q)$. We show that, as $\varepsilon \to 0$, the slow component of $(p^\epsilon(t/\epsilon), q^\epsilon(t/\epsilon))$ tends to a certain stochastic process on Γ which is deterministic inside the edges and branches at the interior vertices into adjacent edges with probabilities which can be calculated through the Hamiltonian H and the perturbation b.

1. INTRODUCTION

An oscillator with one degree of freedom is given by the following equation

$$\ddot{q}(t) + f(q(t)) = 0,\ q(0) = q_0,\ \dot{q}(0) = p_0\,, \tag{1.1}$$

where $q \in \mathbb{R}^1$. The Hamiltonian of the system is

$$H(p, q) = \frac{1}{2}p^2 + F(q)\,,$$

where $F(q) = \int_0^q f(u)\,du$ is the potential and $p = \dot{q}$. Although the results of this paper can be generalized for a wider class of potentials (e.g., periodic potentials) we impose for brevity the following restrictions on the potential. We assume that $\lim_{|q|\to\infty} F(q) = \infty$ and that $f(q)$ is a generic smooth function, so that $f(q)$ has a finite number of nondegenerate zeros (if $f(q) = 0$, then $f'(q) \neq 0$) and the critical values of F are pairwise distinct.

1991 *Mathematics Subject Classification.* 34F05,34C37,58F05,60H10,60J60.

The authors were supported in part by NSF grants DMS-9504135, DMS-9504177 and by a BSF grant 9200087.

We consider perturbations of (1.1) and show that the long time behavior of the perturbed system is in a certain sense stochastic.

Let $b(p,q)$, $p,q\in\mathbb{R}^2$, be a smooth function with bounded first and second derivatives. Consider the perturbed equation

$$\ddot{q}^\varepsilon(t)+f(q^\varepsilon(t))=\varepsilon b(\dot{q}^\varepsilon(t),q^\varepsilon(t)),\ 0\le\varepsilon\ll 1, \tag{1.2}$$

and assume first that $f(q)$ has only one zero, at $q=0$ with $f'(q)>0$ (see Figure 1).

Of course, one can consider an expansion of $q^\varepsilon(t)$ in powers of ε. This describes the solutions of the perturbed system (1.2) on any finite time interval. However, this approach cannot answer more interesting questions concerning the behavior of the perturbed system on large time intervals of the order of ε^{-1}. A typical example of such a question is the exit problem: how long does it take for the trajectory $(p^\varepsilon(t),q^\varepsilon(t))$ of (1.2) to exit a domain $G\subset\mathbb{R}^2$, say, bounded by two trajectories $C(z_i)=\{x\in\mathbb{R}^2 : H(x)=z_i\}$, $i=1,2$, of (1.1)?

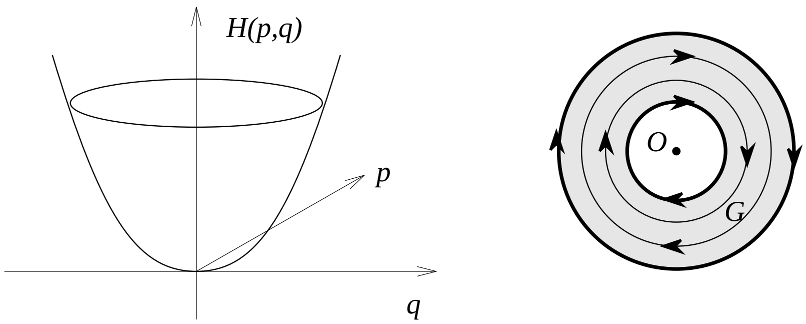

FIGURE 1

The long time behavior of $X^\varepsilon(t)=(p^\varepsilon(t),q^\varepsilon(t))$ is of course related to the averaging principle. Since $H(p,q)$ is a first integral for (1.1), for a small ε, $H(p^\varepsilon(t),q^\varepsilon(t))$ changes slowly so that the flow of (1.2) in $\mathbb{R}^2$ has a fast component and a slow component. The fast component is basically the rotation along the corresponding nonperturbed trajectory $(p(t),q(t))$. Near an energy level z of (1.1), asymptotically as $\varepsilon\downarrow 0$, the fast component can be characterized by the invariant measure of the nonperturbed system on the level set $C(z)$ of H. The density of the normalized invariant measure on $C(z)$ with respect to the length element dl is

$$\left(|\nabla H(x)|\oint_{C(z)}\frac{dl}{|\nabla H(x)|}\right)^{-1},\quad x=(p,q).$$

The slow motion is described by the evolution of $H(p^\varepsilon(t), q^\varepsilon(t)) = \frac{1}{2}(p^\varepsilon(t))^2 + F(q^\varepsilon(t))$ which has the rate of order ε. To deal with finite time intervals one rescales the time by considering $\tilde{p}^\varepsilon(t) = p^\varepsilon(t/\varepsilon)$, $\tilde{q}^\varepsilon(t) = q^\varepsilon(t/\varepsilon)$, $X^\varepsilon(t) = (p^\varepsilon(t), q^\varepsilon(t))$, $\widetilde{X}^\varepsilon(t) = X^\varepsilon(t/\varepsilon)$. Then

$$\ddot{\tilde{q}}^\varepsilon(t) + \frac{1}{\varepsilon} f(\tilde{q}^\varepsilon(t)) = b\left(\dot{\tilde{q}}^\varepsilon(t), \tilde{q}^\varepsilon(t)\right) \tag{1.3}$$

and it follows that

$$H\left(\widetilde{X}^\varepsilon(t)\right) - H\left(\widetilde{X}^\varepsilon(0)\right) = \int_0^t \tilde{p}^\varepsilon(s) b\left(\tilde{p}^\varepsilon(s), \tilde{q}^\varepsilon(s)\right) ds\,. \tag{1.4}$$

On a small but independent of ε time interval $[t, t+\Delta]$, $\Delta \ll 1$, the energy $H\left(\widetilde{X}^\varepsilon(t)\right)$ changes by an amount of order Δ, uniformly in $\varepsilon \in (0,1]$. The number of revolutions that the fast motion makes along the corresponding level set during the same time is of order $\Delta\varepsilon^{-1}$. Therefore, applying the classical averaging principle (see, e.g., [FWen2], Chapter 7), one gets from (1.4) that asymptotically as $\varepsilon \downarrow 0$, the slow component $H\left(\widetilde{X}^\varepsilon(t)\right)$ converges uniformly on any finite time interval $[0, t_0]$ to the averaged motion $H(t)$:

$$\dot{H}(t) = \frac{1}{T(H(t))} \oint_{C(H(t))} \frac{pb(p,q)\, dl}{\sqrt{p^2 + f^2(q)}}\,, \tag{1.5}$$

where $T(z) = \oint_{C(z)} \frac{dl}{\sqrt{p^2 + f^2(q)}}$ is the period of the oscillations with energy $H(p,q) = z$, $C(z) = \{(p,q) \in \mathbb{R}^2 : H(p,q) = z\}$. By the divergence theorem, the integral in the right hand side of (1.5) equals

$$B(z) = \int_{G(z)} \frac{\partial b(p,q)}{\partial p}\, dp\, dq\,,\ z = H(t), \tag{1.6}$$

where $G(z)$ is the domain in $\mathbb{R}^2$ bounded by $C(z)$. Note that if $S(z) = \text{area}(G(z))$, then $T(z) = S'(z)$. Thus (1.5) takes the form

$$\dot{H}(t) = \frac{B(H(t))}{S'(H(t))}\,. \tag{1.7}$$

In particular, if $B(z)$ does not change sign, the slow component is monotone. If $B(z_0) = 0$ and $B'(z_0) < 0$, then the perturbed equation has a stable limit cycle near the curve $C(z)$. The point $H = 0$ is inaccessible in finite time for the trajectory $H(t)$ with $H(0) > 0$.

Consider now the case when f has several zeros. Suppose for example, that $H(p,q) = \frac{p^2}{2} + F(q)$ has a saddle at $\mathcal{O} = (0,0)$ (i.e.,

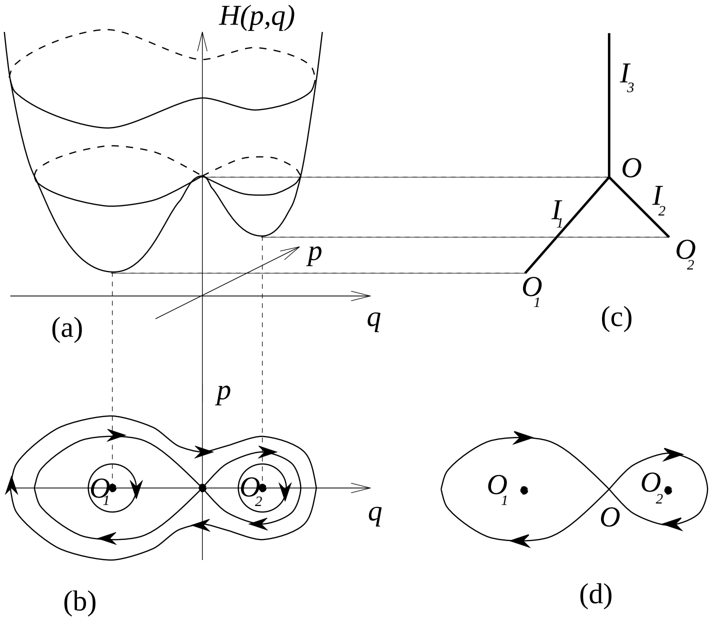

FIGURE 2

the potential $F(q)$ has a local maximum at $q = 0$ and two minima see Figure 2). For brevity assume that $\dfrac{\partial b(p,q)}{\partial p} < 0$, $(p,q) \in \mathbb{R}^2$, (a kind of friction for the oscillator) and let $H(p_0, q_0) > H(\mathcal{O})$. Then, by (1.7), $H(t)$ decreases as t increases. Since $|B(H(\mathcal{O}))|$ is finite and $T(z) \to \infty$ as $z \downarrow H(\mathcal{O})$, the right hand side $B(z)/T(z)$ of (1.7) equals 0 at $z = H(\mathcal{O})$. It is easy to check that $T(z) \sim |\ln(z - H(\mathcal{O}))|^{-1}$ as $z \downarrow H(\mathcal{O})$. Therefore, in spite of the fact that $B(z)/T(z) = 0$ at $z = H(\mathcal{O})$, the trajectory $H(t)$ will reach the point $z = H(\mathcal{O})$ in a finite time $T_0 = T_0(x_0)$, $\widetilde{X}^\varepsilon(0) = x_0 = (p_0, q_0)$. The phase portrait of the unperturbed system is shown in Figure 2b. Since $\dfrac{\partial b(p,q)}{\partial p} < 0$, as a result of the perturbation, the centers will become stable spiral points but the saddle persists, see Figure 3. The trajectories $X^\varepsilon(t)$ of the perturbed system, except the two separatrices of $\mathcal{O}'$, tend to one of the attractors $\mathcal{O}'_1$ or $\mathcal{O}'_2$ depending on the initial point and ε. As we explain below, it is natural to view the limiting slow motion on the graph corresponding to the Hamiltonian H (see [FWen1], [FWen2]).

The set of the connected components of the level sets $C(z) = H^{-1}(z)$, with the natural topology, is homeomorphic to a tree Γ with vertices $\mathcal{O}_i$ and edges I_k. Each level set $C(z)$ consists of a finite number of components $C_k(z)$. By the genericity assumption on F, each interior

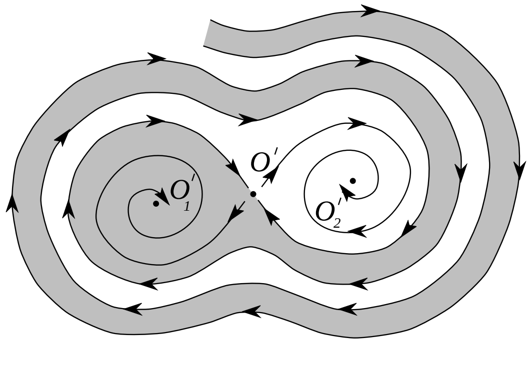

FIGURE 3

vertex is adjacent to exactly three edges; it corresponds to a saddle of H and represents an ∞-shaped component of the saddle's level set. The vertices of Γ of degree 1 (i.e., with only one adjacent edge) correspond to the minima and maxima of H, they represent the components of the level sets consisting of single points. The points lying inside the edges of Γ correspond to the periodic orbits of (1.1). A point $u \in I_k$ is characterized by the edge number k and the corresponding value of the Hamiltonian. The pair (H, k) defines global coordinates on Γ; more than one pair may correspond to a vertex. Let $Q : \mathbb{R}^2 \to \Gamma$ be the projection $x \mapsto (H(x), k(x)) \in \Gamma$ so that $Q(C_k(z)) \in I_k$. Note that $H(x)$ and $k(x)$ are first integrals of the unperturbed system. If H has more than one local minimum, k is independent of H.

In general, when the averaging principle is applied to a perturbed system, the slow component describes the evolution of the first integrals of the unperturbed system along the perturbed trajectories. In our case there is a smooth first integral H and an integer valued first integral k. Their joint evolution naturally happens on Γ.

In Section 2 we construct a stochastic process $Q(t)$ on Γ. Inside each edge e of Γ the process $Q(t)$ is deterministic and is governed by Equation (1.7) for e. If $\dfrac{\partial b(p,q)}{\partial p} < 0$, then $H(\widetilde{X}^\varepsilon(t))$ decreases (i.e., $Q(t)$ moves down along Γ, see Figure 3) and $Q(t)$ branches at each interior vertex $\mathcal{O}$ of Γ with probabilities proportional to the integrals of $\dfrac{\partial b(p,q)}{\partial p}$ over the regions in $\mathbb{R}^2$ bounded by the separatrices of the saddle of (1.1) corresponding to $\mathcal{O}$.

In this paper we show that for a potential with several local minima (and in general, for Hamiltonian systems with one degree of freedom and several critical points), as $\varepsilon \downarrow 0$, the slow component of $\widetilde{X}^\varepsilon(t)$ tends in a certain sense to the stochastic process $Q(t)$.

There are several ways to give the convergence a precise meaning. In Section 2 we consider random perturbations of the right hand side of (1.2). We replace $\varepsilon b(p,q)$ with $\varepsilon b(p,q)+\sqrt{\varepsilon\kappa}\sigma(q)\dot{W}_t$, where $\dot{W}_t$ is the standard one-dimensional white noise, $\kappa>0$ and $\sigma(q)$ is a smooth positive function. We prove that, as first ε and then κ tend to 0, the process $Q\left(\widetilde{X}^{\varepsilon,\kappa}_{t/\varepsilon}\right)$ on Γ converges weakly to the process $Q(t)$ on Γ. We emphasize that the limit process $Q(t)$ does not depend on the choice of $\sigma(q)$.

In Section 3 we consider random perturbations of the initial point $X^\varepsilon(0)$ of the deterministic trajectory. Let $X^\varepsilon(0)=x+\xi_\delta=x_\delta$, where ξ_δ is random variable in $\mathbb{R}^2$ with a continuous density, which converges to 0 as $\delta\to 0$. Let $\widetilde{X}^{\varepsilon,\delta}(t)=\left(\tilde{p}^{\varepsilon,\delta}(t),\tilde{q}^{\varepsilon,\delta}(t)\right)$ be the solution of (1.3) with initial condition $x_\delta=(p_\delta,q_\delta)$. The slow motion is the projection $Q\left(\widetilde{X}^{\varepsilon,\delta}(t)\right)$ of $\widetilde{X}^{\varepsilon,\delta}(t)$ to Γ. We prove in Section 3 that for any $T>0$, as first ε and then δ tend to 0, the stochastic process $Q\left(\widetilde{X}^{\varepsilon,\delta}(t)\right)$ on Γ converges weakly in the space of continuous functions $\phi:[0,T]\to\Gamma$ to the stochastic process $Q(t)$ on Γ.

The fact that in both cases and for different σ the limit process $Q(t)$ on Γ for the slow motion is the same, shows that $Q(t)$ is determined by the intrinsic properties of the Hamiltonian system and its deterministic perturbation $\varepsilon b(p,q)$ and not by the random noise which we use to regularize the problem.

We also consider a more general situation when the oscillator (1.1) is replaced by a general Hamiltonian system with one degree of freedom.

2. Perturbation of the vector field

Let W_t, $t\geq 0$, be the one-dimensional Wiener process, $\sigma(x)$, $x\in\mathbb{R}$, be a positive C^∞-function and let $0<\kappa\ll 1$. In this section we consider a random perturbation of the right hand side of equation (1.2) of the following form:

$$\ddot{q}^{\varepsilon,\kappa}(t)+f\left(q^{\varepsilon,\kappa}(t)\right)=\varepsilon b\left(\dot{q}^{\varepsilon,\kappa}(t),q^{\varepsilon,\kappa}(t)\right)+\sqrt{\varepsilon\kappa}\sigma\left(q^{\varepsilon,\kappa}(t)\right)\dot{W}_t\,. \tag{2.1}$$

Now $q^{\varepsilon,\kappa}(t)$ is a stochastic process in $\mathbb{R}$. To make it a Markov process we consider the pair $p^{\varepsilon,\kappa}(t)=\dot{q}^{\varepsilon,\kappa}(t)$ and $q^{\varepsilon,\kappa}(t)$. After rescaling time $\tilde{p}^{\varepsilon,\kappa}(t)=p^{\varepsilon,\kappa}(t/\varepsilon)$, $\tilde{q}^{\varepsilon,\kappa}(t)=q^{\varepsilon,\kappa}(t/\varepsilon)$ we obtain the following equations for $\tilde{p}^{\varepsilon,\kappa}(t)$ and $\tilde{q}^{\varepsilon,\kappa}(t)$:

$$\begin{aligned}\dot{\tilde{p}}^{\varepsilon,\kappa}(t)&=\frac{1}{\varepsilon}f\left(\tilde{q}^{\varepsilon,\kappa}(t)\right)+b\left(\tilde{p}^{\varepsilon,\kappa}(t),\tilde{q}^{\varepsilon,\kappa}(t)\right)+\sqrt{\kappa}\sigma\left(\tilde{q}^{\varepsilon,\kappa}(t)\right)\dot{W}_t\\ \dot{\tilde{q}}^{\varepsilon,\kappa}(t)&=\frac{1}{\varepsilon}\tilde{p}^{\varepsilon,\kappa}(t)\,.\end{aligned} \tag{2.2}$$

The generator of the diffusion process $\widetilde{X}^{\varepsilon,\kappa}(t) = (\tilde{p}^{\varepsilon,\kappa}(t), \tilde{q}^{\varepsilon,\kappa}(t))$ is

$$\mathcal{L}^{\varepsilon,\kappa} = \frac{\kappa}{2}\sigma^2(q)\frac{\partial^2}{\partial p^2} + b(p,q)\frac{\partial}{\partial p} - \frac{1}{\varepsilon}f(q)\frac{\partial}{\partial p} + \frac{1}{\varepsilon}p\frac{\partial}{\partial q}\,.$$

Let $\Gamma = \{I_1, \ldots, I_n; \mathcal{O}_1, \ldots, \mathcal{O}_m\}$ be the graph corresponding to $H(p,q) = \frac{1}{2}p^2 + F(q)$ and $Q : \mathbb{R}^2 \to \Gamma$ be the corresponding projection, $Q(p,q) = (H(p,q), i(p,q)) \in \Gamma$, $(p,q) \in \mathbb{R}^2$. If an edge I_j is adjacent to a vertex $\mathcal{O}_k$, we write $I_j \sim \mathcal{O}_k$.

Let $\mathcal{O}_k$ be an interior vertex of Γ with adjacent edges $I_{i_1}, I_{i_2}, I_{i_3} \sim \mathcal{O}_k$. The set $Q^{-1}(\mathcal{O}_k)$ consists of the equilibrium point $\mathcal{O}_k$ and two separatrices γ_k^1 and γ_k^2 of $\mathcal{O}_k$ for the unperturbed system (1.1). Denote by $G_j(\mathcal{O}_k)$ the domain bounded by γ_k^j and assume that $Q^{-1}(I_{i_j}) \subseteq G_j(\mathcal{O}_k)$, $j = 1, 2$. Define

$$\begin{aligned} \beta_{kj} &= \int_{G_j(\mathcal{O}_k)} \sigma^2(q)\, dp\, dq\,, \quad j = 1, 2, \\ \beta_{k3} &= -(\beta_{k1} + \beta_{k2})\,. \end{aligned} \tag{2.3}$$

Consider the process $Q^{\varepsilon,\kappa}(t) = Q\left(\tilde{p}^{\varepsilon,\kappa}(t), \tilde{q}^{\varepsilon,\kappa}(t)\right)$ on Γ, which is the slow component of $(\tilde{p}^{\varepsilon,\kappa}(t), \tilde{q}^{\varepsilon,\kappa}(t))$ as $\varepsilon \downarrow 0$. It follows from [FWeb1], [FWeb2] that for any fixed $\kappa > 0$, as $\varepsilon \downarrow 0$, the process $Q^{\varepsilon,\kappa}(t)$ converges weakly in the space of continuous functions to a diffusion process Q_t^κ on Γ. Inside any edge $I_i \subset \Gamma$, the generator $\mathcal{A}$ of Q_t^κ (on smooth functions) is

$$L_i = \frac{\kappa}{2S_i'(z)}\frac{d}{dz}\left(A_i(z)\frac{d}{dz}\right) + \frac{1}{S_i'(z)}B_i(z)\frac{d}{dz}\,, \tag{2.4}$$

where $S_i(z)$ is the area of the domain in $\mathbb{R}^2$ bounded by $C_i(z) = Q^{-1}(z,i)$, $S_i'(z) = \frac{dS_i(z)}{dz}$ and

$$A_i(z) = \int_{G_i(z)} \sigma^2(q)\, dp\, dq\,, \quad B_i(z) = \int_{G_i(z)} \frac{\partial b(p,q)}{\partial p}\, dp\, dq\,.$$

Let $g : \Gamma \to \mathbb{R}$, be a continuous function such that $\mathcal{A}g$ is continuous and g is smooth inside the edges of Γ. Then g belongs to the domain $D_{\mathcal{A}}$ of the generator $\mathcal{A}$ of Q_t^κ if and only if the following gluing condition is satisfied at any interior vertex $\mathcal{O}_k \in \Gamma$:

$$\sum_{j=1}^{3} \beta_{kj} D_j g(\mathcal{O}_k) = 0\,, \tag{2.5}$$

where the β's are defined by (2.3), D_j denotes the z-derivative along I_{i_j} and $I_{i_1}, I_{i_2}, I_{i_3} \sim \mathcal{O}_k$. The limiting process Q_t^κ is uniquely defined by the operators L_i and the gluing conditions (2.5).

Note that, if an edge $I_{i_j} \sim \mathcal{O}_k$, then the limit $\lim_{(z,i_j)\to\mathcal{O}_k} B_{i_j}(z) = B_j(\mathcal{O}_k)$ exists. If $I_{i_1}, I_{i_2}, I_{i_3} \sim \mathcal{O}_k$ and z increases along I_{i_1} and I_{i_2} and decreases along I_{i_3} as the point approaches $\mathcal{O}_k$, then

$$B_3(\mathcal{O}_k) = B_1(\mathcal{O}_k) + B_2(\mathcal{O}_k)$$

for any interior vertex $\mathcal{O}_k \in \Gamma$. We assume for brevity that

$$B_i(\mathcal{O}_k) \neq 0, \quad i = 1, 2, 3, \tag{2.6}$$

for any interior vertex $\mathcal{O}_k \in \Gamma$.

Let I_i be an edge adjacent to an interior vertex $\mathcal{O}_k$. We call I_i an *exit* edge if $H(Q^{-1}(u))$ increases (decreases) along I_i as $u \in \Gamma$ approaches $\mathcal{O}_k$ and if $B_i(z) < 0$ ($B_i(z) > 0$) for z close to $H(Q^{-1}(\mathcal{O}_k))$. If $I_i \sim \mathcal{O}_k$ and is not an exit edge, it is called an *entrance* edge for $\mathcal{O}_k$. By (2.6), every interior vertex is adjacent to at least one exit edge and at least one entrance edge.

Consider the stochastic process $Q(t)$ on Γ defined by the following properties:

1. Inside any edge $I_i \subset \Gamma$, the process $Q(t)$ is deterministic motion with speed $B_i(z)/S_i'(z)$:
$$\frac{dQ(t)}{dt} = \frac{B_i(Q(t))}{S_i'(Q(t))}, \quad Q(t) \in I_i\,.$$
2. If an interior vertex $\mathcal{O}_k$ is adjacent to only one exit edge I_i, then $Q(t)$ leaves $\mathcal{O}_k$ without delay along I_i.
3. If $\mathcal{O}_k$ is an interior vertex and I_{i_1}, $I_{i_2} \sim \mathcal{O}_k$ are exit edges, then $Q(t)$ leaves $\mathcal{O}_k$ without delay along I_{i_1} or I_{i_2} with probabilities
$$P_1 = \frac{|B_1(\mathcal{O}_k)|}{|B_1(\mathcal{O}_k)| + |B_2(\mathcal{O}_k)|}, \quad P_2 = \frac{|B_2(\mathcal{O}_k)|}{|B_1(\mathcal{O}_k)| + |B_2(\mathcal{O}_k)|},$$
respectively, independently of the past.

Note that the speed $\dfrac{dQ(t)}{dt}$ is equal to 0 at any vertex of Γ but, as simple calculations show, near an interior vertex $\mathcal{O}_k$ with $H\left(Q^{-1}(\mathcal{O}_k)\right) = H_0$, this 0 is of order $(|\ln|z - H_0||)^{-1}$ as $|z - H_0| \to 0$. Therefore the interior vertices are accessible for $Q(t)$ in finite time and the trajectory $Q(t)$ can leave such a vertex without delay. The exterior vertices are inaccessible for $Q(t)$.

Let γ_k^1 and γ_k^2 be the separatrices of $\mathcal{O}_k$ for the unperturbed system (1.1) and let G_k^j be the domain bounded by γ_k^j, $j = 1, 2$. Then

$$B_j(\mathcal{O}_k) = \int_{G_k^j} \frac{\partial b(p,q)}{\partial p}\, dp\, dq\,, \;\; j = 1, 2; \;\; B_3(\mathcal{O}_k) = B_1(\mathcal{O}_k) + B_2(\mathcal{O}_k)\,.$$

We will show that, for any $T \in [0, \infty)$, as $\kappa \downarrow 0$, the process $Q^\kappa(t)$, $t \in [0, T]$, on Γ converges weakly in the space of continuous functions $\phi : [0, T] \to \Gamma$ to the process $Q(t)$. Note that the "randomness" of $Q(t)$ is localized at the interior vertices with two exit edges. Since the measure in the space of trajectories of $Q(t)$ is independent of the random perturbation (i.e., independent of $\sigma(q)$), the process $Q(t)$ describes intrinsic properties of the Hamiltonian system (1.1) and its deterministic perturbations.

The Hamiltonian H can be viewed as a function on Γ. Let ρ be the metric on Γ induced by the distance $|H(u_1) - H(u_2)|$ inside the edges of Γ.

Denote by $\mathcal{C}_{0T}(\Gamma)$ the space of continuous functions on $[0, T]$ with values in Γ.

LEMMA 2.1. *For any compact subset $K \subset \Gamma$, the family of stochastic processes $Q^\kappa(t)$, $0 < \kappa \leq 1$, $Q^\kappa(0) \in K$, is tight in the topology of weak convergence in $\mathcal{C}_{0T}(\Gamma)$.*

Proof. The proof of this lemma is similar to the proof of Lemma 3.2 from [FWen1] and we omit it. □

Set $\mathcal{E}_h(u) = \{v \in \Gamma : \rho(u, v) < h\}$ for $u \in \Gamma$ and $\tau_h^\kappa = \min\{t : Q^\kappa(t) \notin \mathcal{E}_h(u)\}$. As usual, we use the subscript for probability and expectation to indicate the initial point.

LEMMA 2.2. *Let $\mathcal{O}_k$ be an interior vertex of Γ and let $I_{i_1}, I_{i_2}, I_{i_3} \sim \mathcal{O}_k$.*

If I_{i_1} is the only exit edge for $\mathcal{O}_k$, then for a small enough h

$$\lim_{\kappa \downarrow 0} P_{\mathcal{O}_k} \{Q^\kappa(\tau_h^\kappa) \in I_{i_1}\} = 1 \,. \tag{2.7}$$

If there are two exit edges I_{i_1} and I_{i_2}, then for a small enough h

$$\lim_{\kappa \downarrow 0} P_{\mathcal{O}_k} \{Q^\kappa(\tau_h^\kappa) \in I_{i_3}\} = 0 \,, \tag{2.8}$$

$$\lim_{\kappa \downarrow 0} P_{\mathcal{O}_k} \left\{Q^\kappa(\tau_h^\kappa) \in I_{i_j}\right\} = P_j \,, \ j = 1, 2 \,. \tag{2.9}$$

Proof. For $u = (z, i) \in \mathcal{E}_h(\mathcal{O}_k)$, set $v_j^\kappa(u) = v_j^\kappa(z, i) = P_{z,i} \left\{Q^\kappa(\tau_h^\kappa) \in I_{i_j}\right\}$. The function $v_j^\kappa(z, i)$ is the unique continuous solution of the following problem

$$
\begin{aligned}
L_{i_m} v_j^\kappa(z, i_m) = 0,\ (z, i_m) \in \mathcal{E}_h(\mathcal{O}_k) \setminus \{\mathcal{O}_k\},\ m = 1, 2, 3;\\
v_j^\kappa(z, i_m)\Big|_{(z,i_m)\in\partial\mathcal{E}_h(\mathcal{O}_k)\cap I_{i_m}} = 0 \text{ for } m \neq j,\\
v_j^\kappa(z, i_j)\Big|_{(z,i_j)\in\partial\mathcal{E}_h(\mathcal{O}_k)} = 1;\\
\sum_{m=1}^{3} \beta_{km} D_m v_j^\kappa(\mathcal{O}_k) = 0\,.
\end{aligned}
\tag{2.10}
$$

Operators L_i and constants β_{km} where introduced in (2.4) and (2.3).

Problem (2.10) can be solved explicitly and then equalities (2.7), (2.8) and (2.9) can be checked. Without loss of generality assume that $H\left(Q^{-1}(\mathcal{O}_k)\right) = 0$, $H(Q^{-1}(z, i_3)) > 0$ and $H(Q^{-1}(z, i_1)), H(Q^{-1}(z, i_2)) < 0$. Define

$$
U_m^\kappa(z) = \int_0^z \frac{2B_{i_m}(y) + \kappa A'_{i_m}(y)}{A_{i_m}(y)}\, dy\,,
$$

$$
V_m^\kappa(x) = \int_0^x e^{-\frac{1}{\kappa} U_m^\kappa(z)} dz\,, \quad m = 1, 2, 3.
$$

Note that $A_{i_m}(y)$ is strictly positive and bounded; $A'_{i_m}(y) \to \infty$ as $y \to 0$ but this singularity is integrable.

The solution of (2.10) has the form:

$$
v_j^\kappa(z, i) = \begin{cases} \dfrac{(1 - d_j^\kappa)V_i^\kappa(z) + d_j^\kappa V_i^\kappa(-h)}{V_i^\kappa(-h)}, & \text{if } i = 1, 2,\ j = i\,;\\[2ex] \dfrac{d_j^\kappa V_i^\kappa(-h) - V_i^\kappa(z)}{V_i^\kappa(-h)}, & \text{if } i = 1, 2,\ j \neq i\,; \end{cases}
$$

$$
v_j^\kappa(z, 3) = \begin{cases} \dfrac{(1 - d_j^\kappa)V_3^\kappa(z) - d_j^\kappa V_3^\kappa(h)}{V_3^\kappa(h)}, & \text{if } j = 3\,;\\[2ex] \dfrac{d_j^\kappa V_3^\kappa(h) - V_3^\kappa(z)}{V_3^\kappa(h)}, & \text{if } j \neq 3\,. \end{cases}
$$

The constants d_j^κ are defined by the following gluing conditions

$$
\sum_{m=1}^{3} \beta_{km} D_m v_j^\kappa(\mathcal{O}_k) = 0\,, \quad j = 1, 2, 3\,. \tag{2.11}
$$

It follows from the definition of $v_j^\kappa(z, i)$ that

$$
P_{\mathcal{O}_k}\left\{Q^\kappa(\tau_h^\kappa) \in I_{i_j}\right\} = v_j^\kappa(\mathcal{O}_k) = d_j^\kappa\,.
$$

Choose h so small that $B_{i_m}(z) \neq 0$ in $\mathcal{E}_h(\mathcal{O}_k)$. Since $V_m^\kappa(x) \to \infty$ as $\kappa \downarrow 0$ if (κ, m) lies in an entrance edge I_m for $\mathcal{O}_k$, one obtains easily from (2.11) that $d_m^\kappa \to 0$ as $\kappa \downarrow 0$. This yields (2.7) and (2.8).

To prove (2.9) observe that since I_{i_3} is an exit edge for $\mathcal{O}_k$, (2.11) implies that

$$\frac{(1-d_1^\kappa)\beta_{k1}}{V_1^\kappa(-h)} - \frac{d_1^\kappa\beta_{k2}}{V_2^\kappa(-h)} o_\kappa(1)\,, \text{ as } \kappa\downarrow 0\,,$$
$$\frac{(1-d_2^\kappa)\beta_{k2}}{V_2^\kappa(-h)} - \frac{d_2^\kappa\beta_{k1}}{V_1^\kappa(-h)} o_\kappa(1)\,, \text{ as } \kappa\downarrow 0\,.$$

Therefore

$$d_1^\kappa + d_2^\kappa = 1 + o_\kappa(1)\,, \quad \frac{d_1^\kappa}{d_2^\kappa} = \frac{\beta_{k1}V_2^\kappa(-h)}{\beta_{k2}V_1^\kappa(-h)}\,. \tag{2.12}$$

From the last equality we obtain that

$$\lim_{\kappa\downarrow 0}\frac{d_1^\kappa}{d_2^\kappa} = \lim_{\kappa\downarrow 0}\frac{\beta_{k1}\int_0^{-h}\exp\left\{-\int_0^z \frac{2B_{i_1}(y)}{\kappa A_{i_1}(y)}dy - \int_0^z \frac{A'_{i_1}(y)}{A_{i_1}(y)}dy\right\}dz}{\beta_{k2}\int_0^{-h}\exp\left\{-\int_0^z \frac{2B_{i_2}(y)}{\kappa A_{i_2}(y)}dy - \int_0^z \frac{A'_{i_2}(y)}{A_{i_2}(y)}dy\right\}dz} = \tag{2.13}$$
$$\lim_{\kappa\downarrow 0}\frac{\beta_{k1}\int_0^{-h}\exp\left\{-\int_0^z \frac{2B_{i_1}(y)}{\kappa A_{i_1}(y)}dy\right\}dz}{\beta_{k2}\int_0^{-h}\exp\left\{-\int_0^z \frac{2B_{i_2}(y)}{\kappa A_{i_2}(y)}dy\right\}dz} =$$
$$\lim_{\kappa\downarrow 0}\frac{\beta_{k1}\int_0^{-h}\exp\left\{-\frac{z|B_{i_1}(0)|}{\kappa A_{i_1}(0)}\right\}dz}{\beta_{k2}\int_0^{-h}\exp\left\{-\frac{z|B_{i_2}(0)|}{\kappa A_{i_2}(0)}\right\}dz} = \frac{|B_{i_1}(0)|}{|B_{i_2}(0)|}\,.$$

We used in the last equality that $\beta_{kj} = \lim_{z\to 0} A_{i_j}(z) = A_{i_j}(0)$, $j = 1, 2$. Formulas (2.12) and (2.13) imply (2.9). $\square$

LEMMA 2.3. *Assume that $B_i(\mathcal{O}_k) \neq 0$, $i = 1, 2, 3$, for an interior vertex $\mathcal{O}_k$. Then there is $h_0 > 0$ and $A > 0$ such that for any $h \in (0, h_0]$, any $(z, i) \in \mathcal{E}_h(\mathcal{O}_k)$ and any small enough κ*

$$E_{z,i}\tau_h^\kappa(\mathcal{O}_k) \le Ah|\ln h|\,.$$

Proof. The lemma can be proved by writing down and solving explicitly a boundary value problem in $\mathcal{E}_h(\mathcal{O}_k) \subset \Gamma$ for $u^\kappa(z, i) = E_{z,i}\tau_h^\kappa(\mathcal{O}_k)$. To reduce the calculations we use comparison arguments (the maximum principle).

As before we assume without loss of generality that $\mathcal{O}_k$ is an interior vertex, I_{i_1}, I_{i_2}, $I_{i_3} \sim \mathcal{O}_k$ and that $H(\mathcal{O}_k) = 0$. Let D_{in} (D_{out}) be the subset of $\{i_1, i_2, i_3\}$ consisting of the indices of the entrance (exit) edges of $\mathcal{O}_k$. Let $h_0 > 0$ be so small that the functions $B_{i_j}(z)$ have no zeros in $\mathcal{E}(\mathcal{O}_k)$. Consider the domain

$$M_h(\mathcal{O}_k) = \{\cup_{j\in D_{\text{in}}} I_j \cup \{\mathcal{O}_k\} \cup [(\cup_{j\in D_{\text{out}}} I_j) \cap \mathcal{E}_h(\mathcal{O}_k)]\,.$$

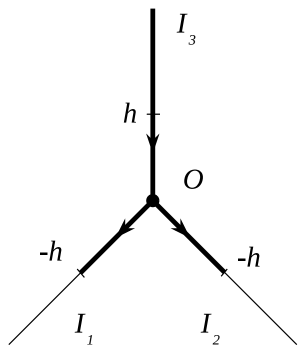

FIGURE 4

For example, in Figure 4 $D_{\text{in}} = \{3\}$, $D_{\text{out}} = \{1,2\}$, the set $M_h(\mathcal{O}_k)$ is shown by a thick line. Since we are interested in the h-neighborhood of $\mathcal{O}_k$, $0 < h \le h_0$, we can assume that the functions $B_{i_j}(z)$ are uniformly bounded away from 0 in the edge I_{i_j}, $i_j \in D_{\text{in}}$. Moreover, we replace the edges I_{i_j}, $j \in D_{\text{in}}$, by infinite rays attached to $\mathcal{O}_k$ which increases the exit time from $M_h(\mathcal{O}_k)$. The signs of the functions $B_{i_j}(z)$ are such that the drift is directed to $\mathcal{O}_k$ on the entrance edge(s) and away from $\mathcal{O}_k$ on the exit edge(s). The exit time from $M_h(\mathcal{O}_k)$ increases if the drift in the entrance edge(s) is replaced by a weaker drift with the same sign and independent of $i \in D_{\text{in}}$ and the diffusion coefficients are replaced by a smaller positive and independent of $i \in D_{\text{in}}$ constant. Similarly the drift and diffusion in the exit edge(s) can be replaced by quantities independent of the edge, increasing the exit time. We denote the modified process by $\overline{Q}^\kappa$.

Since the coefficients are now the same for the entrance edge(s) and the same for the exit edge(s), the exit time of $\overline{Q}^\kappa$ from $M_h(\mathcal{O}_k)$ is the same as the exit time from $\overline{M}_h = \{x \in \mathbb{R} : x > -h\}$ for the one-dimensional process $\overline{X}_t$ whose diffusion and drift coefficients on $[-h, 0]$ are the same as those of $\overline{Q}^\kappa$ on the exit edge(s) and whose diffusion and drift coefficients on $[0, \infty)$ are the same as those of $\overline{Q}^\kappa$ on the entrance edges. Since $S_i'(z)$ has a singularity of order $|\ln|z||$ at $\mathcal{O}_k$, we conclude that the function $u(z,i)$ is bounded from above for $|z| \le h$ by the function $v^\kappa(z) = E_z \overline{\tau}^\kappa$, where $\overline{\tau}^\kappa = \min\{t : \overline{X}_t \le -h\}$. The function $v^\kappa(z)$ is the minimal positive solution of the problem

$$
\begin{aligned}
&\kappa v''(z) - (\beta - \kappa\gamma\,|\ln^*|z||)\,v'(z) = -\alpha\,|\ln^*|z||\,,\\
&z > -h,\ z \ne 0,\ v(-h) = 0, \qquad\qquad (2.14)\\
&v(z) \text{ and } v'(z) \text{ are continuous at } z = 0\,,
\end{aligned}
$$

where α, β, γ are appropriate positive constants, $\ln^* |z| = \ln |z|$ if $|z| < e$ and $\ln^* |z| = 1$ if $|z| > e$. Since the coefficients have a singularity at $z = 0$, equation (2.14) is satisfied only for $z \neq 0$. The solution of (2.14) is

$$v(z) =$$

$$\frac{\alpha}{\kappa}\int_{-h}^{z} e^{\frac{1}{\kappa}\int_0^y (\beta - \kappa\gamma\,|\ln^*|z||)\,dz}\,dy \int_y^{\infty} e^{-\frac{1}{\kappa}\int_0^z (\beta - \kappa\gamma\,|\ln^*|v||)\,dv}\,|\ln^*|z||\;dz\,.$$

Using the integrability of $|\ln |z||$ at $z = 0$ we conclude that for $|z| < h$ and some positive constants A_i

$$(2.15)\quad v(z) \leq \frac{A_1}{\kappa}\int_{-h}^{h} e^{\frac{\beta y}{\kappa}} dy \int_y^{\infty} e^{-\frac{\beta v}{\kappa}}\,|\ln^*|v||\,dv =$$

$$A_2\int_{-h}^{h} de^{\frac{\beta y}{\kappa}} \cdot \int_y^{\infty} e^{-\frac{\beta v}{\kappa}}\,|\ln^*|v||\,dv =$$

$$A_2 e^{\frac{\beta y}{\kappa}} \int_y^{\infty} e^{-\frac{\beta v}{\kappa}}\,|\ln^*|v||\,dv\Big|_{-h}^{h} + A_2\int_{-h}^{h} |\ln^*|v||\,dv$$

$$\leq A_3 h\,|\ln h| + A_2 e^{\frac{\beta h}{\kappa}}\int_h^{\infty} e^{-\frac{\beta v}{\kappa}}\,|\ln^* v|\,dv - A_2 e^{-\frac{\beta h}{\kappa}}\int_{-h}^{\infty} e^{-\frac{\beta v}{\kappa}}\,|\ln^*|v||\,dv$$

Since

$$e^{\frac{\beta h}{\kappa}}\int_h^{\infty} e^{-\frac{\beta v}{\kappa}}\,|\ln^* v|\,dv \leq |\ln h|\int_h^{\infty} e^{-\frac{\beta(v-h)}{\kappa}} dv =$$

$$= |\ln h|\kappa\int_0^{\infty} e^{-\beta v} dv = A_4\kappa\ln h\,,$$

$$e^{-\frac{\beta h}{\kappa}}\int_{-h}^{\infty} e^{-\frac{\beta v}{\kappa}}\,|\ln^*|v||\,dv =$$

$$e^{-\frac{\beta h}{\kappa}}\int_{-h}^{h} e^{-\frac{\beta v}{\kappa}}\,|\ln|v||\,dv + e^{-\beta h}\int_h^{\infty} e^{-\frac{\beta v}{\kappa}}\,|\ln^* v|\,dv \leq$$

$$\int_{-h}^{h} |\ln|v||\,dv + A_4\kappa\ln h \leq A_5\left(h\,|\ln h| + \kappa\ln h\right)\,,$$

we conclude from (2.15) that, for $|x| \leq h$ and κ small enough,

$$v(x) \leq A_6 h\,|\ln h|$$

and the lemma follows. $\square$

The argument in the proof of Lemma (2.3) implies the following fact.

COROLLARY 2.4. *For any* $\delta > 0$

$$\lim_{h\downarrow 0} \max_{x\in\mathcal{E}_h(\mathcal{O}_k)} P_x\left\{\hat{\tau}_h^\kappa > \delta\right\} = 0 \text{ uniformly in } \kappa \in (0, \kappa_0]\,, \tag{2.16}$$

where $\hat{\tau}_h^\kappa = \min\{t : Q^\kappa(t) \notin M_h(\mathcal{O}_k)\}$.

For an edge $I_i \subset \Gamma$ with ends $\mathcal{O}_{k_1}$ and $\mathcal{O}_{k_2}$, set

$$T_i^{\kappa,h} = \min\{t : Q^\kappa(t) \notin I_i \setminus (\mathcal{E}_h(\mathcal{O}_{k_1}) \cup \mathcal{E}_h(\mathcal{O}_{k_2}))\}\ ,\ \kappa, h \geq 0\,;$$

$$T_i^{\kappa,h}(t) = T_i^{\kappa,h} \wedge t\,,\ t > 0\,;\ T_i^{0,0}(t) = T_i(t)\,.$$

LEMMA 2.5. *For any* $\delta > 0$, $t > 0$

$$\lim_{\kappa\downarrow 0} P_{z,i}\left\{\max_{0\leq s\leq T_i(t)} |Q^\kappa(s) - Q(s)| > \delta\right\} = 0\,, \tag{2.17}$$

where the subindex (z, i) *means that* $Q^\kappa(0) = (z, i) = Q(0)$.

Proof. The coefficients of the equations for $Q^\kappa(t)$ and $Q(t)$ are Lipschitz before time $T_i^{\kappa,h} \wedge T^{0,h}(t)$. Therefore (2.17) with $T_i(t)$ replaced by $T_i^{\kappa,h} \wedge T^{0,h}(t)$ follows by standard Gronwall–Bellman inequality arguments (cf. Section 4.2 of [F]). This together with (2.16) implies the lemma. ☐

By Lemmas 2.2, 2.3 and 2.5, as $\kappa \downarrow 0$, the finite dimensional distributions of $Q^\kappa(t)$, $0 \leq t \leq T$, converge to the finite dimensional distributions of $Q(t)$. This convergence together with the weak tightness of the process $Q^\kappa(t)$ guaranteed by Lemma 2.1, implies the weak convergence of $Q^\kappa(t)$ to $Q(t)$ in the space $\mathcal{C}_{0T}(\Gamma)$. Theorem 2.6 below follows from the weak convergence of $Q^{\varepsilon,\kappa}(t)$, $0 \leq t \leq T$, to $Q^\kappa(t)$ as $\varepsilon \downarrow 0$ in $\mathcal{C}_{0T}(\Gamma)$ and from the weak convergence of $Q^\kappa(t)$ to $Q(t)$ as $\kappa \downarrow 0$.

THEOREM 2.6. *Assume that condition (2.6) is satisfied. Then for any* $T > 0$ *the process* $Q^{\varepsilon,\kappa}(t) = Q\left(X_{t/\varepsilon}^{\varepsilon,\kappa}\right)$, $t \in [0, T]$, *converges weakly in* $\mathcal{C}_{0T}(\Gamma)$, *as first* $\varepsilon \downarrow 0$ *and then* $\kappa \downarrow 0$, *to the process* $Q(t)$ *on* Γ.

Consider now a general Hamiltonian system with one degree of freedom

$$\dot{X}(t) = \overline{\nabla} H(X(t))\,,\ X(0) = (p, q) = x \in \mathbb{R}^2\,,$$

$$\overline{\nabla} H(p, q) = \left(-\frac{\partial H}{\partial q}, \frac{\partial H}{\partial p}\right)\,.$$

We assume that $H(x)$ is a smooth generic function with $\lim_{|x|\to\infty} H(x) = \infty$. For a smooth bounded vector field $b(x)$ in $\mathbb{R}^2$ consider the perturbed system

$$\dot{X}^\varepsilon(t) = \overline{\nabla} H\left(X^\varepsilon(t)\right) + \varepsilon b\left(X^\varepsilon(t)\right)\,,\ 0 < \varepsilon \ll 1\,. \tag{2.18}$$

To study the slow component of this system we rescale the time by setting $\widetilde{X}^{\varepsilon}(t) = X^{\varepsilon}(t/\varepsilon)$. Similarly to the system (1.1), the slow component $Q^{\varepsilon}(t)$ of $\widetilde{X}^{\varepsilon}(t)$ is the image of $\widetilde{X}^{\varepsilon}(t)$ under the projection $Q : \mathbb{R}^2 \to \Gamma$, $Q^{\varepsilon}(t) = Q\left(\widetilde{X}^{\varepsilon}(t)\right)$ (see also [FWen1]). It turns out that similarly to the system (1.1), as $\varepsilon \downarrow 0$, the deterministic process $Q^{\varepsilon}(t)$ converges, to a stochastic process $Q(t)$ on the graph Γ.

To give this statement a meaning, we consider again random perturbations of $\widetilde{X}^{\varepsilon}(t)$. Let $a(x)$, $x \in \mathcal{R}^2$, be a uniformly positive definite 2×2 matrix with bounded smooth coefficients. For $\kappa > 0$, define the random perturbation $\widetilde{X}^{\varepsilon,\kappa}(t)$ of $\widetilde{X}^{\varepsilon}(t)$ as the diffusion process in $\mathcal{R}^2$ governed by the operator

$$\mathcal{L}u(x) = \frac{\kappa}{2}\mathrm{div}\ (a(x)\nabla u(x)) + b(x)\cdot\nabla u(x) + \frac{1}{\varepsilon}\overline{\nabla} H(x)\cdot\nabla u(x)\,.$$

Consider the diffusion process $Q^{\kappa}(t)$ on Γ which is defined as follows. Inside an edge $I_i \subset \Gamma$, the generator $\mathcal{A}$ of the process $Q^{\kappa}(t)$ (on smooth functions) is

$$L_i u_i(z) = \frac{\kappa}{2T_i(z)}\frac{d}{dz}\left(A_i(z)\frac{du_i(z)}{dz}\right) + \frac{B_i(z)}{T_i(z)}\frac{du_i(z)}{dz}\,,$$

where

$$A_i(z) = \oint_{C_i(z)} \frac{a(x)\nabla H(x)\cdot\nabla H(x)}{|\nabla H(x)|}dl\,,$$

$$T_i(z) = \oint_{C_i(z)} \frac{dl}{|\nabla H(x)|}\,, \quad B_i(z) = \oint_{C_i(z)} b(x)\cdot n_i(x)\,dl\,,$$

here $C_i(z)$ is the component of the level set $C(z) = \{x : H(x) = z\}$ corresponding to a point $(z,i) \in I_i$ and $n_i(x)$ is the exterior normal to $C_i(z)$. By the divergence theorem

$$B_i(z) = \int_{G_i(z)} \mathrm{div}\, b(x)\,dx\,,$$

where $G_i(z)$ is the domain bounded by $C_i(z)$.

Let $\mathcal{O}_k$ be an interior vertex of Γ and let $I_j \sim \mathcal{O}_k$, $j = 1,2,3$. As $(z,j) \in I_j$ approaches $\mathcal{O}_k$, the set $Q^{-1}(z,j)$ tends to a closed curve γ_k^j which is a subset of $Q^{-1}(\mathcal{O}_k)$. For two of the edges, the curves γ_k^j are the separatrices of $\mathcal{O}_k$ and for the third one γ_k^j is their union.

A function $u(x)$, $x \in \Gamma$, smooth inside the edges of Γ, belongs to the domain of definition of the generator $\mathcal{A}$ of $Q^{\kappa}(t)$ if it is continuous together with $\mathcal{A}u(x)$ on Γ and if the following gluing conditions are satisfied at every interior vertex $\mathcal{O}_k$:

$$\sum_{j:\, I_j \sim \mathcal{O}_k} \pm \beta_{kj} D_j u(\mathcal{O}_k) = 0\,.$$

where

$$\beta_{kj} = \oint_{\gamma_k^j} \frac{a(x)\nabla H(x) \cdot \nabla H(x)}{|\nabla H(x)|} dl\,.$$

The "+" ("-") sign in front of β_{kj} corresponds to H growing (decreasing) as the point approaches $\mathcal{O}_k$ along I_j. The operators L_i and the gluing conditions uniquely determine the process $Q^\kappa(t)$.

By a slight generalization of the arguments in [FWen1], [FWen2] and by the absolute continuity arguments from [FWeb2] we obtain that for any $T > 0$ and $\kappa > 0$, the process $Q^{\varepsilon,\kappa}(t) = Q\left(\widetilde{X}^{\varepsilon,\kappa}(t)\right)$, $0 \le t \le T$, converges, as $\varepsilon \downarrow 0$, to $Q^\kappa(t)$.

For an interior vertex $\mathcal{O}_k$ and an edge $I_j \sim \mathcal{O}_k$, set

$$B_j(\mathcal{O}_k) = \int_{G_j(\mathcal{O}_k)} \operatorname{div} b(x)\, dx\,,$$

where, as before, $G_j(\mathcal{O}_k)$ is the domain bounded by the curve γ_k^j corresponding to I_j. Define a process $Q(t)$ on Γ as follows:

1. If $Q(t)$ belongs to an open edge $I_i \subset \Gamma$, then
$$\frac{dQ(t)}{dt} = \frac{1}{T_i\left(Q(t)\right)} \int_{G_i(Q(t))} \operatorname{div} b(x)\, dx\,.$$
2. If there is only one exit edge I_i for an interior vertex $\mathcal{O}_k$, the process $Q(t)$ leaves $\mathcal{O}_k$ without delay along I_i.
3. If there are two exit edges I_{i_1} and I_{i_2} for an interior vertex $\mathcal{O}_k$, the process $Q(t)$ leaves $\mathcal{O}_k$ without delay along I_{i_j}, $j = 1, 2$, with probability
$$P_j = \frac{|B_j(\mathcal{O}_k)|}{|B_1(\mathcal{O}_k)| + |B_2(\mathcal{O}_k)|}$$
independently of the past.

Using the same arguments as in Theorem (2.6) we obtain the following result.

THEOREM 2.7. *Suppose $B_j(\mathcal{O}_k) \neq 0$ for any interior vertex $\mathcal{O}_k$ and edge $I_j \sim \mathcal{O}_k$. Then, as first $\varepsilon \downarrow 0$ and then $\kappa \downarrow 0$, the process $Q^{\varepsilon,\kappa}(t) = Q\left(\widetilde{X}^{\varepsilon,\kappa}(t)\right)$ converges weakly in $\mathcal{C}_{0T}(\Gamma)$ to the process $Q(t)$.*

Note that Theorem (2.6) is not a special case of Theorem (2.7) because of the degeneracy of the random perturbation in Theorem (2.6). A special case of Theorem (2.7) when $a(x)$ is the unit matrix was studied in [W]. If $H(p,q) = \frac{1}{2}p^2 + f(q),$, the process $Q(t)$ in Theorem (2.7) is the same as in Theorem (2.6). We emphasize that the limiting process is independent of the type of the random perturbations. If equation (2.18) is perturbed by a stationary process $\sqrt{\kappa\varepsilon}\xi(t)$ with strong enough mixing properties and $\xi(t)$ is not degenerate in a certain sense, then we expect the same process $Q(t)$ in the limit, as first $\varepsilon \downarrow 0$ and then $\kappa \downarrow 0$.

3. Perturbation of the initial condition

In this section we consider perturbations of initial conditions for the system

$$\dot{X}^\varepsilon(t) = \overline{\nabla} H(X^\varepsilon(t)) + \varepsilon b(X^\varepsilon(t)), \tag{3.1}$$

with a generic Hamiltonian $H = \frac{1}{2}p^2 + F(q)$, i.e., $F''(q) \neq 0$ for any critical poit q of F and all critical values of F are distinct. We assume that $\lim_{|q|\to\infty} F(q) = \infty$ and that $\operatorname{div} b(x) < 0$ for all x. Under these assumptions, the set of connected components of the level sets $C(z)$ of H, considered with the natural topology, is a tree Γ with projection $Q : \mathbb{R}^2 \to \Gamma$, $x \mapsto Q(x) = (H(x), k(x))$, where $Q(x) \in I_{k(x)}$. Level sets $C(z)$ with large values of z correspond to the only semi-infinite edge I_0 of Γ. The first coordinate increases if the point in Γ approaches I_0. Since $\operatorname{div} b(x) < 0$, each interior vertex $\mathcal{O}$ of Γ corresponds to a saddle of (1.1) (which we also denote $\mathcal{O}$) and has degree three with one entrance edge I_e and two exit edges I_l and I_r which we call *left* and *right*. The Hamiltonian H decreases as the point in Γ approaches the vertex along the entrance edge and increases as the point approaches along the exit edges. If $g \to \mathcal{O}$ along I_l (respectively I_r), then the closed orbit $Q^{-1}(g)$ of (1.1) tends to a separatrix γ_l (respectively γ_r) of the corresponding saddle $\mathcal{O}$. If $g \to \mathcal{O}$ along I_e, then the closed orbit $Q^{-1}(g)$ of (1.1) tends to the ∞-shaped figure $\gamma_l \cup \gamma_r$. Let G_l and G_r be the domains bounded by γ_l and γ_r. Set

$$J(I_l) = \int_{G_l} \operatorname{div} b(x)\, dx\,, \qquad J(I_r) = \int_{G_r} \operatorname{div} b(x)\, dx$$

$$P(I_l) = \frac{J_l}{J_l + J_r}\,, \qquad P(I_r) = \frac{J_r}{J_l + J_r}\,.$$

If I is an edge of Γ, we refer to the end with the larger value of H as the *top* of I and the end with the smaller value of H as the *bottom* of I. The edge I_0 has no top. For two edges $I, I' \subset \Gamma$, we write $I \geq I'$ if there is a path $\alpha(I, I')$ in Γ from the bottom of I to the top of I' along which H decreases. If $I \geq I'$, then such a path is unique; it is a finite collection of edges.

THEOREM 3.1. *Suppose I and I' are edges of Γ with $I \geq I'$ and $\alpha(I, I') = (I_1, I_2, \dots, I_n)$, $I_1 = I$, $I_n = I'$. Let $\mathcal{U} \subset Q^{-1}(I)$ be open and let $\mathcal{U}^\varepsilon \subset \mathcal{U}$ be the set of initial points in $\mathcal{U}$ whose orbits under (1.2) eventually enter $Q^{-1}(I')$. Then*

$$\lim_{\varepsilon \to 0} \frac{m(\mathcal{U}^\varepsilon)}{m(\mathcal{U})} = \prod_{j=2}^{n} p(I_j)\,,$$

where m is the Lebesgue measure in $\mathbb{R}^2$.

Theorem 3.1 follows by an inductive application of Proposition 3.3.

After this paper was written, Ya. Sinai pointed out that a statement similar to Theorem 3.1 appears (without a proof) in V. Arnold's paper [A] (see Section 4.3); the authors are not aware of any published proof.

As before, the slow motion is described by the evolution of the projection of $H(p^\varepsilon(t), q^\varepsilon(t))$ to Γ which has the rate of order ε. To deal with finite time intervals we rescale the time by considering $\tilde{p}^\varepsilon(t) = p^\varepsilon(t/\varepsilon)$, $\tilde{q}^\varepsilon(t) = q^\varepsilon(t/\varepsilon)$, $X^\varepsilon(t) = (p^\varepsilon(t), q^\varepsilon(t))$, $\widetilde{X}^\varepsilon(t) = X^\varepsilon(t/\varepsilon)$. Then

$$\ddot{\tilde{q}}^\varepsilon(t) + \frac{1}{\varepsilon} f(\tilde{q}^\varepsilon(t)) = b\left(\dot{\tilde{q}}^\varepsilon(t), \tilde{q}^\varepsilon(t)\right)$$

and

$$\dot{\widetilde{X}}^\varepsilon(t) = \frac{1}{\varepsilon} \overline{\nabla} H\left(\widetilde{X}^\epsilon(t)\right) + b\left(\widetilde{X}^\epsilon(t)\right)\,. \tag{3.2}$$

It follows that

$$H\left(\widetilde{X}^\varepsilon(t)\right) - H\left(\widetilde{X}^\varepsilon(0)\right) = \int_0^t \tilde{p}^\varepsilon(s) b\left(\tilde{p}^\varepsilon(s), \tilde{q}^\varepsilon(s)\right)\, ds\,. \tag{3.3}$$

Let $\widetilde{X}^\varepsilon(t, x_0)$ be the solution of (3.2) with initial condition $\widetilde{X}^\varepsilon(0, x_0) = x_0$.

Since for every interior vertex $\mathcal{O}$ of Γ, the time during which $\widetilde{X}^\varepsilon(t, x)$ passes through a neighborhood of $Q^{-1}(\mathcal{O})$ is uniformly bounded in ε, the following theorem follows immediately from Theorem 3.1.

THEOREM 3.2. *Let ξ_δ, $\delta > 0$, be a two-dimensional random variable with a continuous density and such that $\xi_\delta \to 0$ as $\delta \downarrow 0$.*

Then, for every $T \geq 0$, as first ε and then δ tends to 0, the stochastic process $Q\left(\widetilde{X}^{\varepsilon}(t, x+\xi_\delta)\right)$ converges weakly to the process Q_t, $Q_0 = x$.

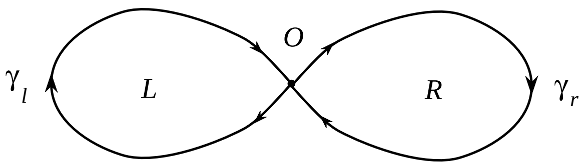

FIGURE 5

We begin by considering the behavior of (3.1) in the special case when H has one local maximum and two minima. The local maximum corresponds to a saddle critical point $\mathcal{O}^{\varepsilon}$ of (3.1). As $\varepsilon \downarrow 0$, the saddle $\mathcal{O}^{\varepsilon}$ tends to a saddle $\mathcal{O}$ of (1.1). The two separatrices γ_l, γ_r of $\mathcal{O}$ bound a figure ∞-shaped region consisting of two domains L and R shown in Figure 5. Since $\operatorname{div} b(x) < 0$, for a small enough $\varepsilon > 0$, the separatrices spiral into the corresponding domains to the left and right of $\mathcal{O}^{\varepsilon}$. Let $\Psi^t(\cdot, \varepsilon)$ denote the time-t map of (3.1), G^t denote the time-t map of the gradient flow

$$\dot{X} = \nabla H(X),$$

and $G(z)$ the trajectory of a point z under the gradient flow. Let $LB(\varepsilon)$ and $RB(\varepsilon)$ be the basins of attraction of L and B for (3.1), respectively:

$$LB^{\varepsilon} = \{x \in \mathbb{R}^2 : \Psi^t(x,\varepsilon) \in L, \text{ for } t \geq T(x) \geq 0\},$$
$$RB^{\varepsilon} = \{x \in \mathbb{R}^2 : \Psi^t(x,\varepsilon) \in R, \text{ for } t \geq T(x) \geq 0\}.$$

The sets LB^{ε} and RB^{ε} consist of the central parts that are close to L and B, respectively, and thin ribbons which we refer to as *flow ribbons*. The boundaries of the flow ribbons are the stable separatrices γ_l^{ε} and γ_r^{ε} of $\mathcal{O}^{\varepsilon}$ (see Figure 6).

Set

$$J_l = \int_L \operatorname{div} b(\xi)\, d\xi\,, \quad J_r = \int_R \operatorname{div} b(\xi)\, d\xi\,.$$

Using the terminology of Section 2, the vertex $\mathcal{O}$ of Γ has one entrance edge I and two exit edges I_l and I_r corresponding to L and B, respectively. Let $\mathcal{B}$ be the preimage of I under the projection $Q : \mathbb{R}^2 \to \Gamma$.

PROPOSITION 3.3. *Let $\mathcal{O}^{\varepsilon}$ be a saddle of (3.1). Then, for every open set $\mathcal{U} \subset \mathcal{B} \setminus (L \cup R)$,*

$$\lim_{\varepsilon \to 0} \frac{m(\mathcal{U} \cap LB^{\varepsilon})}{m(\mathcal{U} \cap RB^{\varepsilon})} = \frac{J_l}{J_r}\,.$$

Proof. We will estimate the H-thickness of the flow ribbons, i.e., the difference between the values of H at the boundaries of the flow ribbons. For a point x from γ_r^ε or γ_l^ε, let $LB^\varepsilon(x)$ be the region obtained from LB^ε by cutting off the tail flow ribbon along the connected component of $G(x)\cap\overline{LB^\varepsilon}$ containing x and let $RB^\varepsilon(x)$ be the region obtained from RB^ε by cutting off the tail flow ribbon along the connected component of $G(x)\cap\overline{RB^\varepsilon}$ containing x. Assume that $y,z\in G(x)$, $x\in\gamma_r^\varepsilon$, $y\in\gamma_l^\varepsilon$, $z\in\gamma_r^\varepsilon$ and the segment of $G(x)$ between x and z intersects the separatrices only at x, y and z. Then $LB^\varepsilon(x)=LB^\varepsilon(y)$ is bounded by the segment of γ_r^ε from $\mathcal{O}^\varepsilon$ to x, the segment of γ_l^ε from $\mathcal{O}^\varepsilon$ to y and the segment of $G(x)$ from x to y. Similarly $RB^\varepsilon(y)=RB^\varepsilon(z)$ is bounded by the segment of γ_l^ε from $\mathcal{O}^\varepsilon$ to y, the segment of γ_r^ε from $\mathcal{O}^\varepsilon$ to z and the segment of $G(z)=G(y)$ from y to z. For $y=G^t(x)$, let $F(x,y)$ denote the flux of $\overline{\nabla}H+\varepsilon b$ through the segment of T between y and z:

$$F(x,y)=\int_0^t\left(\overline{\nabla}H(G^\tau(x))+\varepsilon b(G^\tau(x))\right)\cdot\overline{\nabla}H(G^\tau(x))\,d\tau\,. \tag{3.4}$$

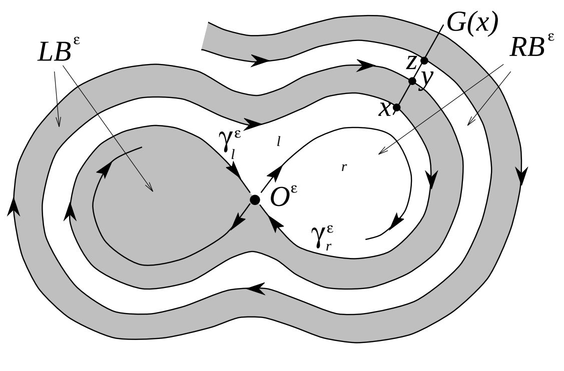

FIGURE 6

LEMMA 3.4. *Let $H(\mathcal{O}^\varepsilon)=a$, let $b>a$ be such that $(a,b]$ does not contain critical values of H. Then there is $C>0$ with the following property. Suppose that $y,z\in G(x)$, $x\in\gamma_r^\varepsilon$, $y\in\gamma_l^\varepsilon$, $z\in\gamma_r^\varepsilon$ and the segment of $G(x)$ between x and z intersects the separatrices only at x, y and z. Assume also that $b>H(z)>H(y)>H(x)=H(\mathcal{O}^\varepsilon)+\delta$ with $\delta>0$. Then*

$$\lim_{\varepsilon\to 0}\left|\frac{H(z)-H(y)}{H(y)-H(x)}-\frac{J_R}{J_L}\right|<C\sqrt{\delta}\,.$$

Proof. Since $\mathcal{O}^\varepsilon$ is a nondegenerate critical point, the area of $H^{-1}(a-\delta, a+\delta)$ does not exceed $C_1\sqrt{\delta}$ for some constant $C_1 > 0$. Therefore

$$\left|\int_{H^{-1}(a-\delta,a+\delta)} \operatorname{div} b(\xi)\, d\xi\right| \le C_2(H,b)\sqrt{\delta}\,. \tag{3.5}$$

Since $\mathcal{O}^\varepsilon$ is a nondegenerate critical point, $|\nabla H(\xi)| > C_2(H)\sqrt{\delta}$ for $\xi \notin H^{-1}(a-\delta, a+\delta)$. Therefore, assuming that $y = G^t(x)$

$$\begin{aligned} |H(y) - H(x) - F(x,y)| &= \left|\int_0^t |\nabla H(G^\tau(x))|^2\, d\tau - F(x,y)\right| \\ &= \varepsilon\left|\int_0^t b(G^\tau(x))\cdot\overline{\nabla} H(G^\tau(x))\, d\tau\right| \le \frac{\varepsilon}{\delta}\left(H(y) - H(x)\right)\,. \end{aligned} \tag{3.6}$$

Similarly

$$|H(z) - H(y) - F(y,z)| \le \frac{\varepsilon}{\delta}\left(H(z) - H(y)\right)\,. \tag{3.7}$$

By the divergence theorem,

$$F(x,y) = \varepsilon\int_{LB^\varepsilon(x,y)} \operatorname{div} b(\xi)\, d\xi\,, \quad F(y,z) = \varepsilon\int_{RB^\varepsilon(y,z)} \operatorname{div} b(\xi)\, d\xi\,.$$

Therefore, by (3.5),

$$\begin{aligned} |F(x,y) - J_r| &\le C_3\sqrt{\delta}\,|J_r| \\ |F(y,z) - J_l| &\le C_4\sqrt{\delta}\,|J_l|\,, \end{aligned} \tag{3.8}$$

where C_3 and C_4 are constants. The lemma follows from (3.6), (3.7) and (3.8).

$\square$

To keep track of the relative areas of different basins of attraction away from the equilibrium $\mathcal{O}$ we construct a convenient coordinate system away from the critical points in which we apply the averaging principle to the equation in variations for (3.1).

Let $-\infty < a < b < \infty$ and let $C(a,b)$ be a component of $H^{-1}([a,b])$ which does not contain any critical points of H. We will use the solutions of (1.1) and the gradient flow of H to construct a convenient coordinate system. Let G^t be the time-t map of the differential equation

$$\dot{X} = \nabla H(X) \tag{3.9}$$

and $\Psi^t(\cdot,\varepsilon)$ be the time-t map of (3.1). Call a solution of (1.1) or (3.1) *regular* if it is not an equilibrium and does not tend to an equilibrium in either direction. The regular solution of (1.1) that starts at x is periodic with period $T(x)$. Let $y \in C(a,b)$ be a point with $H(y) = a$ and let $\phi = 2\pi t/T(y)$ be the rescaled to length 2π time parameter on the solution $S(y)$ of (3.1) starting at y. Let $\Pi : C(a,b) \to S(y)$ be the projection along the gradient flow (3.9). Then $\phi(\Pi(\Psi^t(x,\varepsilon)))$ is a smooth function and $(H(x), \phi(\Pi(x)))$ is a smooth coordinate system in $C(a,b)$. Let

$$h(x,\varepsilon) = \left\{ \left\| dG^\tau(x) \left(\overline{\nabla}(X^\varepsilon(t)) + \varepsilon b(X^\varepsilon(t)) \right) \right\| l(S(y)) \right\}^{-1} ,$$

where $l(S(y))$ is the length of $S(y)$, $dG^t(x)$ is the derivative of $G^t(x)$ with respect to x and τ is such that $G^\tau(x) = \Pi(x)$. Then in coordinates (H, ϕ)

$$\dot{X}^\varepsilon(t) = \left(\overline{\nabla}(X^\varepsilon(t)) + \varepsilon b(X^\varepsilon(t)) \right) h(X^\varepsilon(t), \varepsilon) \tag{3.10}$$

has the form

$$\begin{cases} \dot{H} &= \varepsilon u(\varepsilon, H, \phi) \\ \dot{\phi} &= 1 \, , \end{cases} \tag{3.11}$$

where u is a smooth uniformly bounded function whose derivatives are uniformly bounded for $x \in F$ and $\varepsilon \in (0, \varepsilon_0)$.

Observe that by changing the velocity by a factor of h in (3.10) we do not change the trajectories of (3.1) and therefore do not change the basins LB^ε and RB^ε.

Consider now the following system of differential equations for the derivative of the solution of (3.11) with respect to the initial point in the direction of the variable H:

$$\begin{cases} \dfrac{d}{dt}\Delta H &= \varepsilon \dfrac{\partial u(\varepsilon, H, \phi)}{\partial H} \Delta H \\ \dot{H} &= \varepsilon u(\varepsilon, H, \phi) \\ \dot{\phi} &= 1 \end{cases}$$

After the change of time $t \mapsto t/\varepsilon$ we obtain the following system:

$$\begin{cases} \dfrac{d}{dt}\Delta H &= \dfrac{\partial u(\varepsilon, H, \phi)}{\partial H} \Delta H \\ \dot{H} &= u(\varepsilon, H, \phi) \\ \dot{\phi} &= \varepsilon^{-1} \end{cases} \tag{3.12}$$

LEMMA 3.5. *As $\varepsilon \to 0$, the solutions of (3.12) converge uniformly in $\mathcal{B} \setminus (L \cup R)$, with first derivatives with respect to H, on bounded time* *intervals, to the solutions of the averaged system*

$$\begin{cases} \dfrac{d}{dt}\Delta H &= \left(\displaystyle\int_0^{2\pi} \dfrac{\partial u(0,H,\phi)}{\partial H} d\phi\right) \Delta H \\ \dot H &= \displaystyle\int_0^{2\pi} u(0,H,\phi) d\phi\,. \end{cases}$$

Proof. The statement follows immediately by the averaging principle (see [FWen2], Chapter 7) applied to (3.12). $\square$

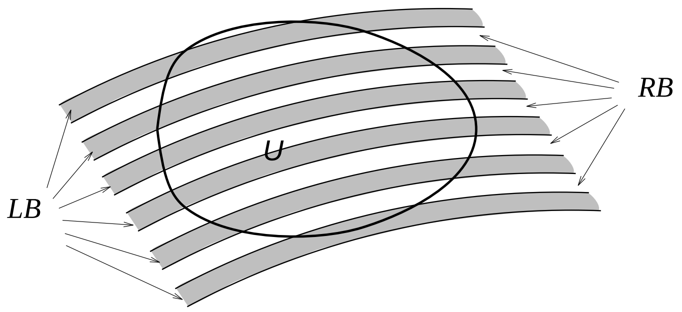

FIGURE 7

To complete the proof of Proposition 3.3 observe that the flow ribbons of the basins LB^ε and RB^ε partition the neighborhood $\mathcal{U}$ into narrow curvilinear rectangles (see Figure 7) of almost constant width. Since $\mathcal{U}$ does not contain the critical points of H, the proportions of LB^ε and RB^ε in $\mathcal{U}$ can be measured by the H-widths of the flow ribbons. By Lemma 3.4, for a small δ, the ratio of the H-widths of the LB^ε flow ribbons and the RB^ε flow ribbons is close to J_L/J_R in the δ-neighborhood of $\mathcal{O}^\varepsilon$. By Lemma 3.5, as $\varepsilon \to 0$, this ratio tends to a constant in $\mathcal{B}$. This finishes the proof of Proposition 3.3. $\square$

Remark 3.1. The issue of independence from the past does not arise in Section 3 since we consider the oscillator system and assume that the divergence of the perturbation is negative. Together with the genericity assumptions this implies that the tree Γ is binary, one edge is an infinite ray corresponding to large values of H, all interior vertices have one entrance and two exit edges; H decreases from left to right in Figure 8 and decreases along the solutions of 1.2. In particular, all trajectories have the same past and the limit process is Markovian.

If one does not assume that $\operatorname{div} b < 0$, then Γ may have an edge I with ends $\mathcal{O}_1$ and $\mathcal{O}_2$ such that $\mathcal{O}_1$ has two entrance edges and one exit edge and $\mathcal{O}_2$ has one entrance and two exit edges, see Figure 9. The preimage of I in the phase plane under the projection Q is homeomorphic to a

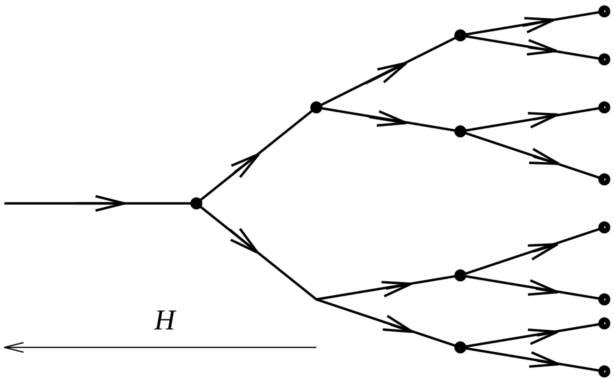

FIGURE 8

cylinder shown in Figure 9 on the right. There is a vertical cut in the cylinder with the flow ribbons from $\mathcal{O}_1$ shaded to the left and the flow ribbons from $\mathcal{O}_2$ shaded to the right; the past behavior of the trajectory is determined by the left shading and the future by the right. The slope and width of the flow ribbons are of order ε and the number of revolutions they make around the cylinder is of order ε^{-1}. As $\varepsilon \downarrow 0$, the flow ribbons to the left of the cut are sliding up and the flow ribbons to the right of the cut are sliding down at a rate of order $\Delta\varepsilon/\varepsilon$. Therefore, for any nontrivial ratios of the integrals of $\operatorname{div} b$ over the areas enclosed by the separatrices of the saddles, the conditional measures of initial conditions corresponding to different pasts and futures do not have limits as $\varepsilon \downarrow 0$. It seems though that if one averages the conditional measures with respect to ε, then the limits exist and the average (in ε) future behavior is asymptotically independent of the past.

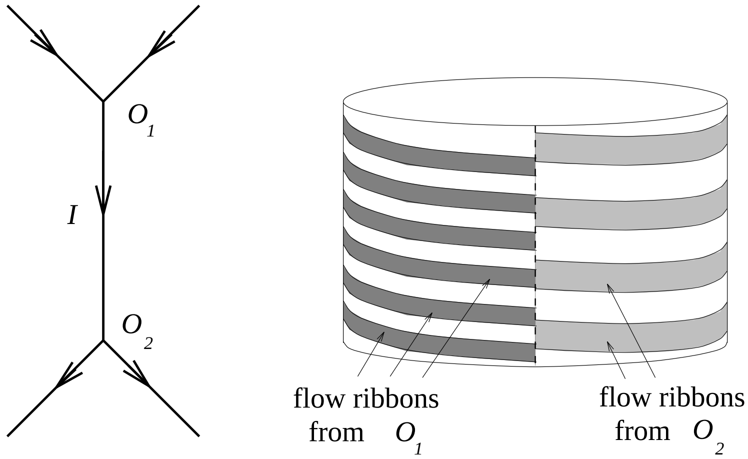

FIGURE 9

EXAMPLE 3.2. Consider the oscillator with $H(p,q) = \dfrac{p^2}{2} + F(q)$, where the potential F has only one local maximum at q^*, $H(q^*) = H^*$, separating two local minima and $\lim_{|q|\to\infty} F(q) = \infty$ (see Figure 2). For $(z,k) \in I_k$, the component $C_k(z)$ of the level set $H^{-1}(z)$ is a periodic trajectory of the oscillator. Let $T_k(z)$ be the period of $C_k(z)$ and $S_k(z)$ be the area of the region $G_k(z)$ bounded by $C_k(z)$, $k = 1,2,3$. Then $S_k'(H) = T_k(z)$.

Assume that the oscillator is perturbed by small friction proportional to the velocity:

$$\ddot{q}^\varepsilon(t) + F\left(q^\varepsilon(t)\right) = -\varepsilon\beta\dot{q}^\varepsilon(t)\,.$$

Applying the results of this paper we conclude that the slow component, after the rescaling of time $t \mapsto t/\varepsilon$, converges, as $\varepsilon \downarrow 0$, to a process $Q(t)$ on Γ whose description follows. Inside I_k, $k = 1,2,3$, the process $Q(t)$ is governed by the following differential equation:

$$\frac{dH}{dt} = \int_{G_k(z)} \operatorname{div}\left(-\beta p, 0\right) = -\beta \frac{S_k(H)}{S_k'(H)}\,, \tag{3.13}$$

so that

$$S_k(H(t)) = S_k(H(0))e^{-\beta t}\,.$$

If the initial energy level $H(0)$ is high enough, the slow motion starts at a point $(H(0),3) \in I_3$ and is described by $S_3(H(t)) = S_3(H(0))e^{-\beta t}$ until time t^* when H decreases to H^*, $t^* = -\dfrac{1}{\beta}\ln\dfrac{S_3(H^*)}{S_3(H(0))}$. At time t^* the process $Q(t)$ branches into I_1 or I_2 with probabilities

$$P_i = \frac{S_i(H^*)}{S_1(H^*) + S_2(H^*)}, \quad i = 1,2,$$

where $S_i(H^*)$, are the areas of the regions bounded by the two separatrices of the saddle $(q^*,0)$. Inside I_1 and I_2 the evolution of the slow component is given by $S_i(H(t)) = S_3(H^*)e^{-\beta(t-t^*)}$.

References

[A] Arnold V.I., Small denominators and problems of stability of motion in classical and celestial mechanics, Russian Mathematical Surveys, vol. 18 (1963), #6, pp. 86-191.

[F] Freidlin M.I., Functional Integration and Partial Differential Equations, Princeton University Press, 1985.

[FWeb1] Freidlin M.I., Weber M., Random perturbations of nonlinear oscillators, accepted by the Annals of Probability.

[FWeb2] Freidlin M.I., Weber M., A remark on random perturbations of a nonlinear pendulum, submitted to the Annals of Applied Probability.

[FWen1] Freidlin M.I., Wentzell A.D., Random perturbations of Hamiltonian systems, Memoirs of the AMS, vol. 523 (1994), pp. 1-82.

[FWen2] Freidlin M.I., Wentzell A.D., Random Perturbations of Dynamical Systems, 2nd edition, Springer, 1998.

[W] Wolansky G., Limit theorems for a dynamical system in the presence of resonances and homoclinic orbits, Journal of Differential Equations, vol. 83(1990), pp. 300-335.

Department of Mathematics, University of Maryland, College Park, MD 20742, mib@math.umd.edu, mif@math.umd.edu