Journal of Statistical Physics, Vol. 123, No. 6, June 2006 (© 2006)
DOI: 10.1007/s10955-006-9133-8

Long-Time Behavior of Weakly Coupled Oscillators

M.I. Freidlin[1] and A.D. Wentzell[2]

Received January 4, 2006; accepted May 17, 2006
Published Online: July 21, 2006

We consider small perturbations of a simple completely integrable system with many degrees of freedom: a collection of independent one-degree-of-freedom oscillators (in the perturbed system the individual oscillators are no longer independent). We show that the long-time behavior of such a system, even in the case of purely deterministic perturbations, should, in general, be described as a stochastic process. The limiting stochastic process is a Markov process on an open book space corresponding to the collection of first integrals of the non-perturbed system.

KEY WORDS: Averaging principle, weakly coupled oscillators

1. INTRODUCTION

Let us consider a system of n independent one-degree-of-freedom oscillators described by

$$\ddot{q}_i = -V_i(q_i), \qquad q_i \in \mathbb{R}^1, \quad i = 1, \ldots, n. \tag{1.1}$$

Denote by $H_i(q_i, p_i)$ the Hamiltonian of the i-th oscillator: $H_i(q_i, p_i) = \frac{p_i^2}{2} + V_i(q_i)$. We assume that the Hamiltonians $H_i(q, p)$ are generic: each of them has a finite number of critical points, which are non-degenerate, not more than one for each connected component of a level set $\{(q, p) : H_i(q, p) = \text{const}\}$; and $\lim_{|q|\to\infty} V_i(q) = \infty$. Let $x_i = (q_i, p_i) \in \mathbb{R}^2$, $\boldsymbol{x} = (x_1, \ldots, x_n)$, $\boldsymbol{H}(\boldsymbol{x}) = \boldsymbol{H}(x_1, \ldots, x_n) = \sum_{i=1}^n H_i(x_i)$. The system (1.1) can be written in the form

$$\dot{\boldsymbol{X}}(t) = \overline{\nabla} \boldsymbol{H}\,(\boldsymbol{X}(t))\,, \tag{1.2}$$

[1] University of Maryland, College Park; e-mail: mif@math.umd.edu.
[2] Tulane University, New Orleans; e-mail: wentzell@math.tulane.edu.

where $\overline{\nabla}\boldsymbol{H}(\boldsymbol{x})$ is the skew gradient of $\boldsymbol{H}(\boldsymbol{x})$. The Hamiltonian system (1.2) is, of course, a completely integrable one with n first integrals $H_1(q_1, p_1), \dots, H_n(q_n, p_n)$.

Consider now a deterministic perturbation of system (1.2):

$$\dot{\widetilde{\boldsymbol{X}}}^{\varepsilon}(t) = \overline{\nabla}\boldsymbol{H}\big(\widetilde{\boldsymbol{X}}^{\varepsilon}(t)\big) + \varepsilon\boldsymbol{\beta}\big(\widetilde{\boldsymbol{X}}^{\varepsilon}(t)\big).$$

A system described by such an equation can be called a system of weakly coupled oscillators. We don't assume that the perturbed system is again a Hamiltonian one. For instance, a small friction in the oscillators is a typical example of perturbations which we are interested in. See example in Sec. 7. It is clear that significant deviations of $\widetilde{\boldsymbol{X}}^{\varepsilon}(t)$ from $\boldsymbol{X}(t)$ occur in time intervals of order of ε^{-1}, so it is convenient to change the time scale and consider $\boldsymbol{X}^{\varepsilon}(t) = \widetilde{\boldsymbol{X}}^{\varepsilon}(t/\varepsilon)$ satisfying the equation

$$\dot{\boldsymbol{X}}^{\varepsilon}(t) = \frac{1}{\varepsilon}\overline{\nabla}\boldsymbol{H}\left(\boldsymbol{X}^{\varepsilon}(t)\right) + \boldsymbol{\beta}\left(\boldsymbol{X}^{\varepsilon}(t)\right). \tag{1.3}$$

The dynamics described by (1.3) has two components: the fast one which is, actually, the motion along the non-perturbed trajectories, and slow motion "across" the trajectories. The slow component, in the case of the Hamiltonians $H_i(x_i)$ having one minimum each, can be described by the values of $H_i(X_i^{\varepsilon}(t))$, $i = 1, \dots, n$. In the general situation of the Hamiltonians having several critical points, to describe the slow component we consider, for every $i = 1, \dots, n$, the graph Γ_i obtained by identifying all points of the plane $\mathbb{R}^2$ belonging to the same connected component of level sets $\{x : H_i(x) = \text{const}\}$ (see Refs. 2–4). It was shown in these papers that in the case of $n = 1$ such a graph is the natural phase space on which we should consider the slow motion. For $n > 1$, if the two-dimensional components β_i of the perturbation vector $\boldsymbol{\beta}(\boldsymbol{x})$ depend on the whole $\boldsymbol{x} = (x_1, \dots, x_n)$ rather than on x_i only, we cannot consider the oscillators separately. The phase space for the slow motion should be $\boldsymbol{\Gamma} = \Gamma_1 \times \Gamma_2 \times \cdots \times \Gamma_n$. This $\boldsymbol{\Gamma}$ is what is called an *open book*, having finitely many n-dimensional *pages* $\boldsymbol{\gamma}_j$, and a *binding* $\boldsymbol{B}$ consisting of pieces of smaller dimensions $\boldsymbol{B}_{n-1}, \boldsymbol{B}_{n-2}, \dots, \boldsymbol{B}_1, \boldsymbol{B}_0$.

Let $\mathfrak{Y}_i(x)$ be the identification mapping $\mathbb{R}^2 \mapsto \Gamma_i$ corresponding to the Hamiltonian $H_i(x)$; let us define $\mathfrak{Y}(\boldsymbol{x}) = \mathfrak{Y}(x_1, \dots, x_n) = (\mathfrak{Y}_1(x_1), \dots, \mathfrak{Y}_n(x_n)) \in \boldsymbol{\Gamma}$. The slow component of the solution of (1.3) is, by definition, $\boldsymbol{Y}^{\varepsilon}(t) = \mathfrak{Y}\left(\boldsymbol{X}^{\varepsilon}(t)\right)$.

As an example, we consider in Sec. 7 a system of two oscillators (Fig. 1). There, the open book consists of three two-dimensional pages and the binding (Fig. 2). One can see in that example that the limiting slow motion, which is a stochastic process on the open book with stochasticity concentrated on the one-dimensional binding, can be described in a rather explicit form.

The main goal of this paper is to study the limiting behavior of the slow motion $\boldsymbol{Y}^{\varepsilon}(t)$ as $\varepsilon \downarrow 0$. In the simplest situation of one degree of freedom and the Hamiltonian without saddle points $Y^\varepsilon(t) = H(X^\varepsilon(t))$ converges, uniformly in every finite time interval, to an averaged motion $\overline{Y}(t)$ being the solution of a certain equation $\dot{\overline{Y}}(t) = \overline{b}(\overline{Y}(t))$, $\overline{Y}(0) = H(X^\varepsilon(0))$. In the case of many degrees of freedom, even in a region where the Hamiltonians have no critical points, this is not the case. One should, as it is well known, introduce the condition that the set of resonance tori is small in some sense (see Ref. 7 and references therein), and even then convergence will take place only in the sense of convergence in measure (with respect to the Lebesgue measure in the space of initial conditions). If the Hamiltonians have saddle points, it can be seen even in the case of one degree of freedom that the slow component $\boldsymbol{Y}^\varepsilon(t) = \mathfrak{Y}(X^\varepsilon(t))$ may have no limit as $\varepsilon \downarrow 0$ (see, for example, Ref. 2). As it follows from Refs. 1, 8, the problem, in the case of no critical points, can be "regularized" by adding small random perturbation to the initial point. In the case of saddle points present, such a regularization, generally, is not sufficient even for systems with one degree of freedom (see Ref. 2): if the number of saddle points of the Hamiltonian is greater than 1, the limit may not exist. In more detail: Let $X^{\varepsilon,\delta}(t)$ be the solution of $\dot{X}^{\varepsilon,\delta}(t) = \frac{1}{\varepsilon}\overline{\nabla} H(X^{\varepsilon,\delta}(t)) + \beta(X^{\varepsilon,\delta}(t))$ with the randomly perturbed initial condition $X^{\varepsilon,\delta}(0) = x + \delta \cdot \xi$, where the two-dimensional random variable ξ is, say, uniformly distributed in the unit circle. Then the slow component $Y^{\varepsilon,\delta}(t)$ may have no limit as first ε and then δ go to zero (such a double limit does exist for systems with at most one saddle point).

It was shown in Ref. 2 that for systems with one degree of freedom one can regularize the problem by adding a stochastic perturbation not to the initial point but rather to the right-hand side of the equation. We are going to do this for n-degrees-of-freedom systems. Let us replace (1.3) by the equation

$$\dot{\boldsymbol{X}}^{\varepsilon,\varkappa}(t) = \frac{1}{\varepsilon}\overline{\nabla}\boldsymbol{H}(\boldsymbol{X}^{\varepsilon,\varkappa}(t)) + \boldsymbol{\beta}(\boldsymbol{X}^{\varepsilon,\varkappa}(t)) + \sqrt{\varkappa}\boldsymbol{\sigma}\dot{\boldsymbol{W}}(t), \tag{1.4}$$

where $\boldsymbol{W}(t)$ is a $2n$-dimensional Wiener process, and $\boldsymbol{\sigma}$ a $2n \times 2n$ matrix having 2×2 nonzero matrices $\sigma_i = (\sigma_{i;jk})^2_{j,k=1}$ on the diagonal, and 0 elsewhere (for simplicity's sake, we don't consider matrices depending on $\boldsymbol{x}$; if $\boldsymbol{\beta}(\boldsymbol{x}) \equiv 0$, the components $X_i^{\varepsilon,\varkappa}(t)$ of $\boldsymbol{X}^{\varepsilon,\varkappa}(t)$ are independent). The slow component $\boldsymbol{Y}^{\varepsilon,\varkappa}(t) = \mathfrak{Y}(\boldsymbol{X}^{\varepsilon,\varkappa}(t))$ of the diffusion process described by the equation (1.4) is a continuous stochastic process (generally, not a Markov one) on the open book $\boldsymbol{\Gamma}$. Under some natural additional assumptions, we show that, as $\varepsilon \downarrow 0$, the process $\boldsymbol{Y}^{\varepsilon,\varkappa}(t)$ converges weakly to a diffusion process $\overline{\boldsymbol{Y}}^{\varkappa}(t)$ on $\boldsymbol{\Gamma}$. We evaluate the generator of $\overline{\boldsymbol{Y}}^{\varkappa}(t)$, which is described as certain differential operators on n-dimensional pages of our open book, plus some appropriate gluing conditions at its binding. Then we find the weak limit $\overline{\boldsymbol{Y}}^{0}(t)$ of the processes $\overline{\boldsymbol{Y}}^{\varkappa}(t)$ as $\varkappa \downarrow 0$. It turns out that this limit exists; being the limit of Markov processes $\overline{Y}^{\varkappa}(t)$, it is quite natural for $\overline{\boldsymbol{Y}}^{0}(t)$ to be also a Markov one. Inside each n-dimensional page $\boldsymbol{\gamma}_j$ of $\boldsymbol{\Gamma}$, $\overline{\boldsymbol{Y}}^{0}(t)$ is a deterministic motion described as the solution of a differential equation with coefficients $\overline{\boldsymbol{\beta}}(\boldsymbol{y})$ obtained by averaging $\beta_i(x_1, \ldots, x_n) \cdot \overline{\nabla} H_i(x_i)$ over the set $\mathfrak{Y}^{-1}(\boldsymbol{y})$ (which is a connected component of a level set $\{\boldsymbol{x} = (x_1, \ldots, x_n) : H_1(x_1) = \text{const}_1, \ldots, H_n(x_n) = \text{const}_n\}$). At the binding $\boldsymbol{B}$ of the open book $\boldsymbol{\Gamma}$, the process $\overline{\boldsymbol{Y}}^0(t)$ displays a stochastic behavior: if $\overline{\boldsymbol{Y}}^0(t)$ comes to a point $\boldsymbol{y}$ belonging to the binding, it goes without any delay to one of the pages meeting at the part of the binding containing this point, to each page with its own probability (depending on the point $\boldsymbol{y} \in \boldsymbol{B}$). (The limiting process $\overline{\boldsymbol{Y}}^0(t)$ cannot go back to the page $\boldsymbol{\gamma}_j$ from which it came to the binding at or near the point $\boldsymbol{y}$ at which it has come to the binding, since the velocity $\overline{\boldsymbol{\beta}}(\boldsymbol{y})$ is directed *from* the page $\boldsymbol{\gamma}_j$ to the binding).

It turns out that the limiting process $\overline{\boldsymbol{Y}}^0(t)$ is independent of the choice of the matrix $\boldsymbol{\sigma}$ (taken within the class we described above). This means that the stochasticity of the limiting slow motion is, actually, an intrinsic property of the deterministic system (1.3) with $\varkappa = 0$. The stochastic term $\sqrt{\varkappa}\boldsymbol{\sigma}\dot{\boldsymbol{W}}(t)$ is used just for regularization of the deterministic problem. The real reason for this stochasticity is instability of saddle points of the Hamiltonian.

Our main results will be formulated: that about the limit as $\varepsilon \downarrow 0$ for positive $\varkappa$, in Sec. 5 (the preparatory work being done in Secs. 3 and 4); that about the limit as $\varkappa \downarrow 0$, at the end of Sec. 6. An example will be considered in Sec. 7.

2. NOTATIONS

What was written above was to explain what we are planning to do in the simplest way; in this section, we are going to describe a slightly more general class of systems that can be handled the same way, and introduce notations.

Instead of denoting the two coordinates of a point x in the plane with q, p, we are going to denote them ξ_1, ξ_2 (this is more convenient because then we can write sums $\sum_{i=1}^{2}$); we'll no longer assume that the Hamiltonian $H_i(x)$ has the form $\frac{p^2}{2} + V_k(q)$: it will be an arbitrary smooth function such that $\lim_{|x|\to\infty} H_i(x) = \infty$. The skew gradient has the form $\overline{\nabla} H_i(x) = (\frac{\partial H_i}{\partial \xi_2}, -\frac{\partial H_i}{\partial \xi_1})$. We'll assume that, for sufficiently large $|x|$, $H_i(x) \geq A_1|x|^2$, $A_2|x| \leq |\nabla H_i(x)| \leq A_3|x|$, and the matrix of the second derivatives $(\frac{\partial^2 H_i(x)}{\partial \xi_j \partial \xi_r})_{j,r=1,2}$ is bounded and uniformly positive definite for large $|x|$.

In the notations already introduced, $\mathfrak{Y}_i : \mathbb{R}^2 \mapsto \Gamma_i$ is the identification mapping corresponding to the i-th Hamiltonian. The graph Γ_i consists of vertices O_{ik} (each vertex is the image under the mapping $\mathfrak{Y}_i$ of a critical point x_{ik} of the Hamiltonian H_i) and edges I_{il}. The fact of an edge I_{il} having a vertex O_{ik} as one of its ends will be noted as $I_{il} \sim O_{ik}$. A point $y \in \Gamma_i$ will be characterized by two coordinates: $y = (l, H)$, where l is the number of the edge I_l of the graph containing y, and $H = H_i(y)$. For $y = (l, H) \in \Gamma_i$, the notation $y \to \infty$ will mean that $H \to \infty$ (this may happen only along one edge I_{il} of the graph).

A vertex O_{ik} of the graph Γ_i will be called *internal* if the critical point x_{ik} belonging to $\mathfrak{Y}_i^{-1}(O_{ik})$ is a saddle; and *external* if this critical point is a local extremum. Exactly one edge I_{il} enters every external vertex; and exactly three edges meet at every internal one: $I_{il_1}, I_{il_2}, I_{il_3} \sim O_{ik}$. So an internal vertex has three sets of coordinates: $O_{ik} = (l_1, H_{ik}) = (l_2, H_{ik}) = (l_3, H_{ik})$, where $H_{ik} = H_i(O_{ik})$.

For shortness, $C_i(y)$, $y \in \Gamma_i$, will denote the inverse image $\mathfrak{Y}_i^{-1}(y)$. The same kind of notation, $C_i(x)$, will be used for $x \in \mathbb{R}^2$: $C_i(x) = \mathfrak{Y}_i^{-1}(\mathfrak{Y}_i(x))$ is the connected component of a level set of the Hamiltonian H_i containing the point x.

For $1 \le i \le n$ and $y \in \Gamma_i$ not being a vertex of this graph we take

$$T_i(y) = \oint\limits_{C_i(y)} \frac{1}{|\nabla H_i(x)|} \ell(dx), \tag{2.1}$$

where $\ell(dx)$ denotes integration with respect to the curve length. This is the period of the rotation of the system $\dot{X}(t) = \overline{\nabla} H_i(X(t))$ along the curve $C_i(y)$. The same notation $T_i(x) = T_i(\mathfrak{Y}_i(x))$ will be used for points $x \in \mathbb{R}^2$.

For $y \in \Gamma_i$ not being a vertex we define a measure μ_y^i concentrated on $C_i(y)$ by

$$\mu_y^i(A) = T_i(y)^{-1} \oint\limits_{C_i(y)} \frac{I_A(x)}{|\nabla H_i(x)|} \ell(dx),$$

where $I_A(x)$ is the indicator function of $A \subset C_i(y)$; and if y is a vertex $O_{ik} \in \Gamma_i$, we define μ_y^i as a unit mass concentrated at the equilibrium point $x_{ik} \in \mathfrak{Y}_i^{-1}(O_{ik})$.

The measure μ_y^i clearly depends on $y \in \Gamma_i$ in a weakly continuous way.

Now we define

$$a_{i;jr} = \sum_{t=1}^{2} \sigma_{i;jt}\sigma_{i;rt}, \quad 1 \le j, r \le 2;$$

$$\overline{a}_i(y) = \oint\limits_{C_i(y)} \sum_{j,r=1}^{2} a_{i;jr} \cdot \frac{\partial H_i}{\partial \xi_j} \frac{\partial H_i}{\partial \xi_r} \mu_y^i(dx),$$

$$\overline{b}_{i0}^{\varkappa}(y) = \frac{\varkappa}{2} \oint\limits_{C_i(y)} \sum_{j,r=1}^{2} a_{i;jr} \cdot \frac{\partial^2 H_i(x)}{\partial \xi_j \partial \xi_r} \mu_{y_i}^i(dx)$$

for $y \in \Gamma_i$; and for $\boldsymbol{y} = (y_1, \ldots, y_n) \in \boldsymbol{\Gamma}$ we take $\overline{b}_i^{\varkappa}(\boldsymbol{y}) = \overline{b}_{i0}^{\varkappa}(y_i) + \overline{\beta}_i(\boldsymbol{y})$, where

$$\overline{\beta}_i(\boldsymbol{y}) = \oint_{C_1(y_1)} \cdots \oint_{C_n(y_n)} \beta_i(x_1, \ldots, x_n) \cdot \nabla H_i(x_i) \mu_{y_1}^1(dx_1) \ldots \mu_{y_n}^n(dx_n). \quad (2.2)$$

Let O_{ik} be an interior vertex of the graph Γ_i; we have $O_{ik} = (l_1, H_{ik}) = (l_2, H_{ik}) = (l_3, H_{ik})$, where $H_{ik} = H_i(O_{ik})$, and l_1, l_2, l_3 are the numbers of the edges I_{il_s}, $s = 1, 2, 3$, meeting at O_{ik}. Let us define $C_{i;kl_s}$ as the part of the curve $C_i(O_{ik})$ that forms a part of the boundary of $\mathfrak{Y}_i^{-1}(I_{il_s} \setminus \{O_{ik}\})$. One of $C_{i;kl_s}$ coincides with the whole curve $C_i(O_{ik})$, consisting of two "loops"; and the other two $C_{i;kl_s}$ are these loops taken separately.

We define

$$\alpha_{i;kl_s} = \oint_{C_{i;kl_s}} \sum_{j,r=1}^{2} a_{i;jr} \cdot \frac{\frac{\partial H_i}{\partial \xi_j} \frac{\partial H_i}{\partial \xi_r}}{|\nabla H_i(x)|} \ell(dx), \quad s = 1, 2, 3. \quad (2.3)$$

Now for $y = (l, H)$ being an interior point of an edge $I_{il} \subseteq \Gamma_i$, and for every function $f(y) = f(l, H)$ that is twice continuously differentiable in H, we take

$$\overline{L}_{i0}^{\varkappa} f(y) = \frac{\varkappa}{2} \overline{a}_i(y) \cdot \frac{d^2 f(l, H)}{dH^2} + \overline{b}_{i0}^{\varkappa}(y) \cdot \frac{df(l, H)}{dH}.$$

A function $f(y) = f(l, H)$ on Γ_i is said to belong to $D_i = D_i^{\varkappa}$ if the following requirements are satisfied: it is continuous on Γ_i and has a finite limit as $y \to \infty$ (in other words: it is continuous on $\Gamma_i \cup \{\infty\}$); it is twice continuously differentiable (with respect to H) on the interior parts of the edges I_{il} of the graph; for every vertex $O_{ik} = (l, H_{ik})$ of the graph and every edge $I_{il} \sim O_{ik}$ a finite limit $\lim_{H \to H_{ik}} \frac{df(l,H)}{dH}$ exists; finite limits

$$\lim_{y \to O_{ik}} \overline{L}_{i0}^{\varkappa} f(y) \quad (2.4)$$

exist for all vertices O_{ik}, and a finite limit

$$\lim_{y \to \infty} \overline{L}_{i0}^{\varkappa} f(y) \quad (2.5)$$

too; and f satisfies the gluing conditions

$$\sum_{s=1}^{3} (\pm \alpha_{i;kl_s}) \cdot \lim_{H \to H_{ik}} \frac{df(l_s, H)}{dH} = 0 \quad (2.6)$$

at every interior vertex $O_{ik} = (l_1, H_{ik}) = (l_2, H_{ik}) = (l_3, H_{ik})$, where the sign "+" is taken if the edge $I_{i;l_s}$ consists of points (l_s, H) with $H \geq H_{ik}$, and "−" if $H \leq H_{ik}$ for $(l_s, H) \in I_{il}$.

For $f \in D_i$, we define the value of the function $\overline{L}_{i0}^{\varkappa} f$ at a vertex O_{ik} as the limit (2.4).

Proposition 2.1. (see Refs. 3 and 4) *For every function u_0 on Γ_i belonging to D_i there exists a unique solution $u(t, y)$ of the differential equation $\frac{\partial u(t,y)}{\partial t} = \overline{L}_{i0} u(t, y)$ with initial condition $u(0, y) = u_0(y)$ such that $u(t, \bullet) \in D_i$ for every $t > 0$.*

The solution $\overline{P}_{i0;y}^{\varkappa}$ of the martingale problem corresponding to the operator $\overline{L}_{i0}^{\varkappa}$ with domain D_i with initial distribution concentrated at an arbitrary point $y \in \Gamma_i$ exists and is unique.

As a matter of fact, the probability measure solving this problem, with $y = \mathfrak{Y}_i(x)$, is the weak limit as $\varepsilon \downarrow 0$ of the function-space distribution of the random function $Y_{i0}^{\varepsilon,\varkappa}(\bullet) = \mathfrak{Y}_i(X_{i0}^{\varepsilon,\varkappa}(\bullet))$, where $X_{i0}^{\varepsilon,\varkappa}$ is the solution of the equation $\dot{X}_{i0}^{\varepsilon,\varkappa}(t) = \frac{1}{\varepsilon}\overline{\nabla} H_i(X_{i0}^{\varepsilon,\varkappa}(t)) + \sqrt{\varkappa}\sigma_i \dot{W}_i(t)$, $X_{i0}^{\varepsilon,\varkappa}(0) = x$, with the drift coefficient $\beta_i \equiv 0$ (see Refs. 3 and 4).

Now let us define the differential operator $\overline{\boldsymbol{L}}^{\varkappa}$ on the open book $\boldsymbol{\Gamma}$.

By definition, a function $\boldsymbol{f}(\boldsymbol{y}) = \boldsymbol{f}(y_1, \ldots, y_n) = \boldsymbol{f}(l_1, H_1, \ldots, l_n, H_n)$ on $\boldsymbol{\Gamma} = \Gamma_1 \times \cdots \times \Gamma_n$ belongs to $\boldsymbol{D} = \boldsymbol{D}^{\varkappa}$ if the following conditions are satisfied:

$\boldsymbol{f}$ is continuous on $(\Gamma_1 \cup \{\infty\}) \times \cdots \times (\Gamma_n \cup \{\infty\})$;

$\boldsymbol{f}$ has first and second continuous partial derivatives in H_i for $y_i = (l_i, H_i)$ in the interior parts of edges I_{il_i} of the graph Γ_i, $i = 1, \ldots, n$ (so that we can apply the operator $\overline{L}_{i0}$ to $\boldsymbol{f}$ in its argument y_i);

For every vertex $O_{ik} = (l, H_{ik}) \in \Gamma_i$, every edge of this graph $I_{il} \sim O_{ik}$, and arbitrary points $y_{j0} \in \Gamma_j \cup \{\infty\}$, $j \neq i$, a finite limit

$$\lim_{H_i \to H_{ik}, y_j \to y_{j0}, j \neq i} \frac{\partial \boldsymbol{f}(y_1, \ldots, l, H_i, \ldots, y_n)}{\partial H_i}$$

(with y_i approaching O_{ik} along the edge I_{il}) exists;

finite limits

$$\lim_{y_i \to O_{ik}, y_j \to y_{j0}, j \neq i} \overline{L}_{i0}^{\varkappa} \boldsymbol{f}(y_1, \ldots, y_n)$$

exist for all vertices $O_{ik} \in \Gamma_i$ and all $y_{j0} \in \Gamma_j \cup \{\infty\}$, $j \neq i$, where the operator $\overline{L}_{i0}^{\varkappa}$ is applied to the function $\boldsymbol{f}$ in its i-th argument;

$\boldsymbol{f}$ satisfies the gluing conditions

$$\sum_{s=1}^{3} (\pm \alpha_{i;ks}) \cdot \lim_{H_i \to H_{ik}, y_j \to y_{j0}, j \neq i} \frac{\partial \boldsymbol{f}(y_1, \ldots, l_s, H_i, \ldots, y_n)}{\partial H_i} = 0 \qquad (2.7)$$

for every interior vertex $O_{ik} = (l_1, H_{ik}) = (l_2, H_{ik}) = (l_3, H_{ik})$ of the graph Γ_i, $i = 1, \ldots, n$, and all points $y_j \in \Gamma_j$, $j \neq i$.

We'll be also considering a smaller domain $\boldsymbol{D}_0 = \boldsymbol{D}_0^{\varkappa}$ consisting of all linear combinations of functions $\boldsymbol{f}$ having the form $\boldsymbol{f}(y_1, \ldots, y_n) = f_1(y_1) \cdot \ldots \cdot f_n(y_n)$, where $f_i \in D_i$, $i = 1, \ldots, n$.

Now we define the operator $\overline{\boldsymbol{L}}^{\varkappa}$.

For a function $\boldsymbol{f} \in \boldsymbol{D}$ we define the function $\overline{\boldsymbol{L}}^{\varkappa} \boldsymbol{f}(\boldsymbol{y})$, $\boldsymbol{y} \in \boldsymbol{\Gamma}$, by

$$\overline{\boldsymbol{L}}^{\varkappa} \boldsymbol{f}(\boldsymbol{y}) = \overline{\boldsymbol{L}}^{\varkappa} \boldsymbol{f}(y_1, \ldots, y_n) = \sum_{i=1}^{n} \overline{b}_i^{\varkappa}(\boldsymbol{y}) \cdot \frac{\partial \boldsymbol{f}}{\partial H_i} + \frac{\varkappa}{2} \sum_{i=1} \overline{a}_i(y_i) \cdot \frac{\partial^2 \boldsymbol{f}}{\partial H_i^2},$$

where the summands are replaced by the corresponding limits for y_i being vertices of Γ_i. The limits at vertices $O_{ik} \in \Gamma_i$ exist because they do for the operators $\overline{L}_{i0}^{\varkappa}$, and the coefficients $\overline{\beta}_i(\boldsymbol{y})$ (defined by (2.2)) in the difference of the operators $\overline{\boldsymbol{L}}^{\varkappa} - \sum_{i=1}^{n} \overline{L}_{i0}^{\varkappa}$ have zero limits at vertices, this being because the measure $\mu_{O_{ik}}^{i}$ is concentrated at the critical point $x_{ik} \in \mathfrak{Y}_i^{-1}(O_{ik})$, and $\nabla H_i(x_{ik}) = 0$.

Considering the averaged operator $\overline{\boldsymbol{L}}^{\varkappa}$ on the domain $\boldsymbol{D}$ is more natural; however for our purposes and methods of proof we need it only on a much smaller domain $\boldsymbol{D}_0$.

3. EXISTENCE AND UNIQUENESS OF THE SOLUTION OF THE MARTINGALE PROBLEM

In this section and Secs. 4 and 5, we are going to speak about limit passage as $\varepsilon \downarrow 0$ and about the limiting stochastic process; the parameter $\varkappa$ is fixed. So we are going to drop the mention of this parameter in our notations, and write the diffusion coefficients without the factor $\varkappa$.

As it is explained in Refs. 5, 6, several things are needed to establish weak convergence of the function-space distribution $\boldsymbol{P}_{\boldsymbol{x}}^{\varepsilon}$ of the random function $\boldsymbol{Y}^{\varepsilon}(\bullet) = \mathfrak{Y}(\boldsymbol{X}^{\varepsilon}(\bullet))$ with respect to the probability measure $\boldsymbol{P}_{\boldsymbol{x}}^{\varepsilon}$ to that of the Markov process $\overline{\boldsymbol{Y}}(\bullet)$ with generating operator $\overline{\boldsymbol{L}}$ with respect to the probability measure $\overline{\boldsymbol{P}}_{\mathfrak{Y}(\boldsymbol{x})}$:

- "tightness" of the family of measures $\mathbf{P}_{\boldsymbol{x}}^{\varepsilon}$;
- convergence

$$\boldsymbol{E}_{\boldsymbol{x}}^{\varepsilon} \int_0^{\infty} e^{-\lambda t} \left[\lambda \boldsymbol{f}(\boldsymbol{Y}^{\varepsilon}(t)) - \overline{\boldsymbol{L}} \boldsymbol{f}(\boldsymbol{Y}^{\varepsilon}(t))\right] dt \to \boldsymbol{f}(\mathfrak{Y}(\boldsymbol{x})) \tag{3.1}$$

  as $\varepsilon \downarrow 0$ for every function $\boldsymbol{f}$ belonging to a set $\boldsymbol{D}_{\overline{L}}$ and for every positive λ; and
- uniqueness of the solution of the martingale problem corresponding to the operator $\overline{\boldsymbol{L}}$ considered on the domain $\boldsymbol{D}_{\overline{L}}$.

The tightness is pretty easy to establish—we are not going to stop at it; the problems of uniqueness and of convergence (3.1) are considered independently from one another. We are going to consider the uniqueness first.

This is done in pretty much the same way as in Ref. 6.

Probably the most natural way of proving uniqueness of the solution of the martingale problem corresponding to the operator $\overline{\boldsymbol{L}}$ on $\boldsymbol{\Gamma}$ with domain $\boldsymbol{D}_{\overline{\boldsymbol{L}}}$ is proving existence of the solution of the problem

$$\frac{\partial \boldsymbol{u}(t, \boldsymbol{y})}{\partial t} = \overline{\boldsymbol{L}}\boldsymbol{u}(t, \boldsymbol{y}), \qquad \boldsymbol{u}(0, \boldsymbol{y}) = \boldsymbol{u}_0(\boldsymbol{y}), \qquad \boldsymbol{u}(t, \bullet) \in \boldsymbol{D}_{\overline{\boldsymbol{L}}}, \quad t > 0, \tag{3.2}$$

for a dense set of initial values $\boldsymbol{u}_0$. This is essentially a version of Theorem 6.3.2 of Ref. 9 freed from the reference to the space of infinitely differentiable function. But this is an equation on an open book, not on merely a multidimensional region, and not one on a one-dimensional structure (graph); in addition the coefficients of the operator $\overline{\boldsymbol{L}}$ degenerate at the binding of the open book; and we don't know how to find solutions of this equation, even if we take $\boldsymbol{D}_{\overline{\boldsymbol{L}}} = \boldsymbol{D}$, the larger of the two versions of domain. So we apply an oblique method.

If we take $\boldsymbol{\beta}(\boldsymbol{x}) \equiv 0$ and consider the corresponding operator $\overline{\boldsymbol{L}}_0$, we have: $\overline{\boldsymbol{L}}_0 \boldsymbol{f}(\boldsymbol{y}) = \sum_{i=1}^{n} \overline{L}_{i0} \boldsymbol{f}(y_1, \dots, y_n)$, where the operator $\overline{L}_{i0}$ is applied to the function in its i-th argument. As for solutions of the martingale problem, clearly the probability measure $\overline{\mathbf{P}}_{0;\boldsymbol{y}} = \overline{P}_{10;y_1} \times \dots \times \overline{P}_{n0;y_n}$ (the joint distribution of n independent solutions of the corresponding martingale problems) is a solution of the martingale problem corresponding to the operator $\overline{\boldsymbol{L}}_0$ (with the domain $\boldsymbol{D}_{\overline{\boldsymbol{L}}_0} = \boldsymbol{D}_0$); and this establishes the existence.

Is this solution unique?

Let us take $\boldsymbol{u}_0(\boldsymbol{y}) = \boldsymbol{u}_0(y_1, \dots, y_n) = \prod_{i=1}^{n} u_{i0}(y_i)$, $u_{0i} \in D_i$. Let $u_i(t, y)$, $t \geq 0$, $y \in \Gamma_i$, be the solution of the problem $\frac{\partial u_i(t,y)}{\partial t} = \overline{L}_{i0} u_i(t, y)$, $u_i(0, y) = u_{i0}(y)$, $u_i(t, \bullet) \in D_i$, $t > 0$. Then the function $\boldsymbol{u}(t, \boldsymbol{y}) = \prod_{i=1}^{n} u_i(t, y_i)$ clearly solves the problem (3.2) with $\overline{\boldsymbol{L}} = \overline{\boldsymbol{L}}_0$ and $\boldsymbol{D}_{\overline{\boldsymbol{L}}} = \boldsymbol{D}_0$. Since linear combinations of functions belonging to $\boldsymbol{D}_0$ form a dense set in the space of continuous functions, the existence and uniqueness problem is solved for $\boldsymbol{\beta}(\boldsymbol{x}) \equiv 0$, $\overline{\boldsymbol{\beta}}(\boldsymbol{y}) = \left(\overline{\beta}_1(\boldsymbol{y}), \dots, \overline{\beta}_n(\boldsymbol{y})\right) \equiv 0$.

It turns out that we can change the drift coefficients in the generating operator by making an absolutely continuous change of the probability measure.

Proposition 3.1. *Let $\boldsymbol{e}(\boldsymbol{y}) = (e_1(\boldsymbol{y}), \dots, e_n(\boldsymbol{y}))$ be a measurable function on $\boldsymbol{\Gamma}$ such that $e_i(\boldsymbol{y}) = 0$ for y_i being a vertex of Γ_i, and the functions $\overline{a}_i(y_i) \cdot e_i(\boldsymbol{y})^2$ are bounded. Let $\mathbf{C}[0, \infty)$ be the space of continuous functions $\mathfrak{y}(t) = (\mathfrak{y}_1(t), \dots, \mathfrak{y}_n(t))$, $0 \leq t < \infty$, $\mathfrak{y}_i(t) \in \Gamma_i$. Let us introduce the random functions*

$$m_i(t) = H_i\left(\mathfrak{y}_i(t)\right) - \int_0^t \overline{b}_i\left(\mathfrak{y}(s)\right)\, ds. \tag{3.3}$$

Suppose $\mathbf{P}$ *is a solution of the martingale problem associated with the operator* $\overline{\boldsymbol{L}}$ *with domain* $\boldsymbol{D}$.

Then the random functions (3.3) are square-integrable martingales with respect to $\mathbf{P}$*, and stochastic integrals* $\int_0^t e_i\,(\mathfrak{y}(s))\,dm_i(s)$ *are defined. Take*

$$\pi[0,t] = \exp\left\{\sum_{i=1}^{n}\int_0^t e_i\,(\mathfrak{y}(s))\;dm_i(s) - \frac{1}{2}\sum_{i=1}^{n}\int_0^t \overline{a}_i\,(\mathfrak{y}_i(s))\,e_i\,(\mathfrak{y}(s))^2\,ds\right\},$$

and define the probability measure $\hat{\mathbf{P}}$ *by* $\hat{\mathbf{P}}(B) = \mathbf{E}(B;\pi[0,t])$ $(\mathbf{E}(B;\ )$ *being the expectation corresponding to the probability measure* $\mathbf{P}$ *taken over the set* B*) for events* B *belonging to the algebra* $\bigcup_{0\le t<\infty}\sigma\{\mathfrak{y}(s),\,0\le s\le t\}$ *and by extension on the* σ*-algebra generated by all random variables* $\mathfrak{y}(t)$*,* $0\le t<\infty$.

Then $\hat{\mathbf{P}}$ *is a solution of the martingale problem corresponding to the linear operator* $\hat{\overline{\boldsymbol{L}}}$ *defined the same way as* $\overline{\boldsymbol{L}}$*, with the same coefficients* $\overline{a}_i(y_i)$ *as* $\overline{\boldsymbol{L}}$*, with* $\hat{\overline{b}}_i(\boldsymbol{y}) = \overline{b}_i(\boldsymbol{y}) + \overline{a}_i(y_i)\cdot e_i(\boldsymbol{y})$*, and the same domain as* $\overline{\boldsymbol{L}}$.

The *proof* is similar to that of Propositions 5.3 and 6.1 in Ref. 6.

Now we can take $\mathbf{P} = \overline{\mathbf{P}}_{0;y}$ (the probability measure solving the martingale problem associated with the operator $\overline{\boldsymbol{L}}_0$) and $e_i(\boldsymbol{y}) = \overline{\beta}_i(\boldsymbol{y})/\overline{a}_i(y_i)$ (replacing it with 0 when y is a vertex of the graph Γ_i). If the perturbing drift function $\boldsymbol{\beta}(\boldsymbol{x})$ is bounded, the functions $\overline{a}_i(y_i)\cdot e_i(\boldsymbol{y})^2 = \overline{\beta}_i(\boldsymbol{y})^2/\overline{a}_i(y_i)$, are bounded, because for $y_i = (l, H)$ close to an interior vertice $O_{ik} = (l, H_{ik})$, $\boldsymbol{y} = (y_1,\dots,y_n)$ we have $\overline{\beta}_i(\boldsymbol{y}) = O(\frac{1}{|\ln|H-H_{ik}||})$, $\overline{a}_i(y_i)\sim\frac{\text{const}}{|\ln|H-H_{ik}||}$; near exterior vertices $\overline{\beta}_i(\boldsymbol{y}) = O(|H-H_{ik}|^{1/2})$, $\overline{a}_i(y_i)\sim\text{const}\cdot|H-H_{ik}|$; and as $H\to\infty$, we have $\overline{\beta}_i(\boldsymbol{y}) = O(\sqrt{H})$, $\overline{a}_i(y_i)\ge\text{const}\cdot H$. So we can apply Proposition 3.1 and get the probability measure $\hat{\mathbf{P}}$ solving the martingale problem associated with the operator $\overline{\boldsymbol{L}}$. This takes care of the existence problem.

For *uniqueness*, we apply the same Proposition with $\mathbf{P} = \overline{\mathbf{P}}_{\boldsymbol{y}}$ (a measure solving the martingale problem associated with $\overline{\boldsymbol{L}}$, with the initial distribution concentrated at the point $\boldsymbol{y}$) and $e_i(\boldsymbol{y}) = -\overline{\beta}_i(\boldsymbol{y})/\overline{a}_i(y_i)$; the probability measure $\hat{\mathbf{P}}$ solves the martingale problem associated with the operator $\overline{\boldsymbol{L}}_0$. From the uniqueness of such a measure we deduce uniqueness for the martingale problem associated with $\overline{\boldsymbol{L}}$.

4. CONVERGENCE AS $\varepsilon\downarrow 0$. CASE OF REGIONS WITH THE $(n-2)$-DIMENSIONAL PART OF THE BINDING CUT OUT

We have introduced the notation $T_i(x)$ for the period of the solution of $\dot{X}(t) = \overline{H}_i\,(X(t))$ starting from the point x (see the paragraph containing formula (2.1)). The corresponding frequency will be $\omega_i(x) = 1/T_i(x)$.

Let us introduce our main restriction on these frequencies (cf. Ref. 5):

Condition ⋆: The set of points $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^{2n}$ for which the frequencies $\omega_1(x_1), \ldots, \omega_n(x_n)$ are rationally dependent has zero Lebesgue measure.

Let $\boldsymbol{\tau}$ be the time at which the process $\boldsymbol{Y}^\varepsilon(t)$ leaves a region $\boldsymbol{\Gamma}_0 \subset \boldsymbol{\Gamma}$ (we are not showing the dependence of $\boldsymbol{\tau}$ on ε to avoid cumbersome notations). We are going to prove that

$$\boldsymbol{E}_{\boldsymbol{x}}^\varepsilon\Big[e^{-\lambda\boldsymbol{\tau}} \boldsymbol{f}(\boldsymbol{Y}^\varepsilon(\boldsymbol{\tau})) + \int_0^{\boldsymbol{\tau}} e^{-\lambda t}\big[\lambda \boldsymbol{f}(\boldsymbol{Y}^\varepsilon(t)) - \overline{\boldsymbol{L}}\boldsymbol{f}(\boldsymbol{Y}^\varepsilon(t))\big]dt\Big] \to \boldsymbol{f}(\mathfrak{Y}(\boldsymbol{x})) \tag{4.1}$$

as $\varepsilon \downarrow 0$ for different classes of regions $\boldsymbol{\Gamma}_0$. Because of the uniqueness result that we already have, (4.1) means that the distribution of the random function $\boldsymbol{Y}^\varepsilon(\bullet)$ stopped at the time $\boldsymbol{\tau}$ converges weakly to that of the process $\overline{\boldsymbol{Y}}(\bullet)$ stopped at the time when it leaves $\boldsymbol{\Gamma}_0$: to express it shorter but less precisely, weak convergence takes place before the time of leaving $\boldsymbol{\Gamma}_0$.

Proposition 4.1. *Suppose that the Hamiltonians H_i are four times continuously differentiable; for sufficiently large $|x|$, $H_i(x) \geq A_1|x|^2$, $A_2|x| \leq |\nabla H_i(x)| \leq A_3|x|$, where A_i are positive constants, and the matrix of the second derivatives $(\frac{\partial^2 H_i(x)}{\partial\xi_j\partial\xi_r})_{j,r=1,2}$ is bounded and uniformly positive definite for large $|x|$; that* **Condition ⋆** *is satisfied; and that $\boldsymbol{\beta}(\boldsymbol{x})$ is bounded.*

For every $i = 1, \ldots, n$, let Γ_{i0} be a subedge of an edge I_{il_i} of the graph Γ_i: a subedge whose ends are interior points of I_{il_i}; and let $\boldsymbol{\Gamma}_0 = \Gamma_{10} \times \cdots \times \Gamma_{n0}$.

Then for every $\boldsymbol{f} \in \boldsymbol{D}$ and every $\lambda > 0$ (4.1) is satisfied, uniformly in $\boldsymbol{x} \in \mathfrak{Y}^{-1}(\boldsymbol{\Gamma}_0)$.

Proof. In the region $\mathfrak{Y}_i^{-1}(\Gamma_{i0}) \subset \mathbb{R}^2$ (which is homeomorphic to an anulus) we introduce action-angle coordinates: $H_i(x)$ and $\varphi_i(x)$, the last one changing in the unit circle, so that for $\boldsymbol{X}^\varepsilon(t) = (X_1^\varepsilon(t), \ldots, X_n^\varepsilon(t)) = (H_1^\varepsilon(t), \varphi_1^\varepsilon(t), \ldots, H_n^\varepsilon(t), \varphi_n^\varepsilon(t))$ we have $\frac{d\varphi_i^\varepsilon(t)}{dt} = \frac{1}{\varepsilon}\omega_i(H_i^\varepsilon(t)) + c_i(\boldsymbol{X}^\varepsilon(t)) + \tilde{\sigma}_i(X_i^\varepsilon(t))\dot{W}_i(t)$. Then we apply Theorem 3 of Ref. 5 (the uniformity in $\boldsymbol{x}$ was not included in the formulation of that theorem, but the proof is essentially the same).

The following result is some sort of cross between the main result of Ref. 6 and that of Ref. 5:

Proposition 4.2. *Let the conditions imposed on the Hamiltonians H_i and on $\boldsymbol{\beta}(\boldsymbol{x})$ in* Proposition 4.1 *be satisfied. Let all Γ_{i0} but one be subedges as described in* Proposition 4.1*, and the remaining one Γ_{i0} is just a compact subgraph of Γ_i.*

Then for every $\boldsymbol{f} \in \boldsymbol{D}$ and every $\lambda > 0$ (4.1) is satisfied, uniformly in $\boldsymbol{x} \in \mathfrak{Y}^{-1}(\boldsymbol{\Gamma}_0)$, $\boldsymbol{\Gamma}_0 = \Gamma_{10} \times \cdots \times \Gamma_{n0}$.

The proof is a combination of those in Ref. 6 and in Ref. 5: the sequence of stopping times $0 = \tau_0 \leq \sigma_1 \leq \tau_1 \leq \sigma_2 \leq \cdots$ is constructed similarly to Ref. 6; we handle the time intervals from τ_i to σ_{i+1} in a way similar to that of Ref. 6 (taking the gluing conditions into account), and for the time intervals from σ_i to τ_i, we apply Proposition 4.1.

Proposition 4.3. *Let the conditions imposed on the Hamiltonians H_i and on $\boldsymbol{\beta}(\boldsymbol{x})$ in* Proposition 4.1 *be satisfied. For small $d > 0$, let $\Gamma_i(\leq d)$ be the union of closed d-neighborhoods of all vertices $O_{ik} \in \Gamma_i$:*

$$\Gamma_i(\leq d) = \bigcup_{O_{ik}\in\Gamma_i, I_{il}\sim O_{ik}} \{(l, H) : |H - H_i(O_{ik})| \leq d\};$$

let us define $\boldsymbol{\Gamma}(d)$ by

$$\boldsymbol{\Gamma}(d) = \boldsymbol{\Gamma} \setminus \bigcup_{1\leq i<j\leq n} \Gamma_1 \times \cdots \times \Gamma_{i-1} \times \Gamma_i(\leq d) \times \Gamma_{i+1} \times \cdots \times \Gamma_{j-1} \times \Gamma_j(\leq d) \times \Gamma_{j+1} \times \cdots \times \Gamma_n \tag{4.2}$$

(the open book $\boldsymbol{\Gamma}$ from which some neighborhood of the $(n-2)$-dimensional part of the binding is deleted);

$$\boldsymbol{\Gamma}_0 = \{\boldsymbol{y} = (y_1, \ldots, y_n) \in \boldsymbol{\Gamma}(d) : H_i(y_i) < R, 1 \leq i \leq n\}.$$

Then for every $\boldsymbol{f} \in \boldsymbol{D}$ and every $\lambda > 0$ (4.1) is satisfied, uniformly in $\boldsymbol{x} \in \mathfrak{Y}^{-1}(\boldsymbol{\Gamma}_0)$.

Limit passage as $R \to \infty$ yields the same for $\boldsymbol{\Gamma}_0 = \boldsymbol{\Gamma}(d)$, uniformly in $\boldsymbol{x}$ changing in every compact subset of $\mathfrak{Y}^{-1}(\boldsymbol{\Gamma}(d))$.

Proof of Proposition 4.3. Choose a positive $\delta < d$. Let us define $\boldsymbol{\tau}_0 = 0$;

$$\boldsymbol{\tau}_1 = \min\{t \geq 0 : \text{ one of } Y_j^\varepsilon(t) \in \Gamma_j(\leq d), \text{ or } \boldsymbol{Y}^\varepsilon(t) \notin \boldsymbol{\Gamma}_0\};$$

and for $i > 0$ we define $\boldsymbol{\tau}_{i+1}$ as being equal to $\boldsymbol{\tau}_i$ if $\boldsymbol{Y}^\varepsilon(\boldsymbol{\tau}_i) \notin \boldsymbol{\Gamma}_0$, and by

$$\boldsymbol{\tau}_{i+1} = \min\{t \geq \boldsymbol{\tau}_i : \text{ one of } Y_j^\varepsilon(t) \in \Gamma_j(\leq \delta), 1 \leq j \leq n, j \neq k, \text{ or } \boldsymbol{Y}^\varepsilon(t) \notin \boldsymbol{\Gamma}_0\}$$

if $Y_k^\varepsilon(\boldsymbol{\tau}_i) \in \Gamma_k(\leq \delta)$.

Note that there can be only one k such that $Y_k^\varepsilon(\boldsymbol{\tau}_i) \in \Gamma_k(\leq \delta)$, because if both this and $Y_j^\varepsilon(\boldsymbol{\tau}_i) \in \Gamma_j(\leq \delta)$, $j \neq k$, held, then $\boldsymbol{Y}^\varepsilon(\boldsymbol{\tau}_i)$ would have been deep inside $\Gamma_1 \times \cdots \times \Gamma_{k-1} \times \Gamma_k(\leq d) \times \Gamma_{k+1} \times \cdots \times \Gamma_{j-1} \times \Gamma_j(\leq d) \times \Gamma_{j+1} \times \cdots \times \Gamma_n$, and the time $\boldsymbol{\tau}_i$ would be *after* leaving $\boldsymbol{\Gamma}_0$.

It is clear that all $\boldsymbol{\tau}_i$, starting with some i, are equal to $\boldsymbol{\tau}$.

We have:

$$\boldsymbol{E}_{\boldsymbol{x}}^\varepsilon\Big[e^{-\lambda\boldsymbol{\tau}} \boldsymbol{f}(\boldsymbol{Y}^\varepsilon(\boldsymbol{\tau})) + \int_0^{\boldsymbol{\tau}} e^{-\lambda t}\big[\lambda \boldsymbol{f}(\boldsymbol{Y}^\varepsilon(t)) + \overline{\boldsymbol{L}} \boldsymbol{f}(\boldsymbol{Y}^\varepsilon(t))\big]dt\Big] - \boldsymbol{f}(\mathfrak{Y}(\boldsymbol{x}))$$

$$= \sum_{i=0}^{\infty} \boldsymbol{E}_{\boldsymbol{x}}^{\varepsilon}\Big[e^{-\lambda \boldsymbol{\tau}_{i+1}} \boldsymbol{f}(\boldsymbol{Y}^{\varepsilon}(\boldsymbol{\tau}_{i+1})) + \int_{\boldsymbol{\tau}_i}^{\boldsymbol{\tau}_{i+1}} e^{-\lambda t}\big[\lambda \boldsymbol{f}(\boldsymbol{Y}^{\varepsilon}(t)) + \overline{\boldsymbol{L}} \boldsymbol{f}(\boldsymbol{Y}^{\varepsilon}(t))\big] dt$$
$$- e^{-\lambda \boldsymbol{\tau}_i} \boldsymbol{f}(\boldsymbol{Y}^{\varepsilon}(\boldsymbol{\tau}_i))\Big]. \tag{4.3}$$

The zeroth summand converges to 0, uniformly in $\boldsymbol{x} = (x_1, \ldots, x_n)$, by Proposition 4.1, where we take as Γ_{i0} the edge of the graph Γ_i containing the point $\mathfrak{Y}_i(x_i)$ with δ-neighborhoods of its ends deleted. To the i-th summand in (4.3), $i > 1$, we apply the strong Markov property with respect to $\boldsymbol{\tau}_i$; and we get that it is equal to $\boldsymbol{E}_{\boldsymbol{x}}^{\varepsilon} \psi^{\varepsilon}(\boldsymbol{X}^{\varepsilon}(\boldsymbol{\tau}_i))$, where $\psi^{\varepsilon}(\boldsymbol{x}') = \psi^{\varepsilon}(x_1', \ldots, x_n') = 0$ if $\mathfrak{Y}(\boldsymbol{x}') \notin \boldsymbol{\Gamma}_0$, and

$$\psi^{\varepsilon}(\boldsymbol{x}') = \boldsymbol{E}_{\boldsymbol{x}'}^{\varepsilon}\Big[e^{-\lambda \boldsymbol{\sigma}_k} \boldsymbol{f}(\boldsymbol{Y}^{\varepsilon}(\boldsymbol{\sigma}_k)) + \int_0^{\boldsymbol{\sigma}_k} \big[\lambda \boldsymbol{f}(\boldsymbol{Y}^{\varepsilon}(t)) + \overline{\boldsymbol{L}} \boldsymbol{f}(\boldsymbol{Y}^{\varepsilon}(t))\big] dt\Big] - \boldsymbol{f}\big(\mathfrak{Y}(\boldsymbol{x}')\big)$$

if $\mathfrak{Y}_k(x_k') \in \Gamma_k(\le \delta)$, where

$$\boldsymbol{\sigma}_k = \min\{t \ge 0 : \text{ one of } Y_j^{\varepsilon}(t) \in \Gamma_j(\le \delta),\ j \ne k, \text{ or } \boldsymbol{Y}^{\varepsilon}(t) \notin \boldsymbol{\Gamma}_0\}.$$

By Proposition 4.2, we have that $\psi^{\varepsilon}(\boldsymbol{x}') \to 0$ as $\varepsilon \downarrow 0$, uniformly in $\boldsymbol{x}'$ (as Γ_{k0} we take $\{y \in \Gamma_k : H_k(y) \le R\}$, and as Γ_{j0}, $j \ne k$, the edge of the graph Γ_j containing the point $\mathfrak{Y}_j(x_j')$ with δ-neighborhoods of its ends deleted).

To conclude the proof, we show that the expectation $\boldsymbol{E}_{\boldsymbol{x}}^{\varepsilon} \sum_{i:\boldsymbol{\tau}_i < \boldsymbol{\tau}} e^{-\lambda \boldsymbol{\tau}_i}$ is uniformly bounded for small positive ε; and this is done taking into account that between the times $\boldsymbol{\tau}_i$ and $\boldsymbol{\tau}_{i+1} < \boldsymbol{\tau}$ the process $\boldsymbol{Y}^{\varepsilon}(t)$ has to travel at least the positive distance $d - \delta$.

So we have proved what we needed before the time of coming too close to the $(n-2)$-dimensional part of the binding. What remains to be checked is that we come too close to it only with a very small probability. This will be done in the next section.

5. INACCESSIBILITY OF THE $(n-2)$-DIMENSIONAL PART OF THE BINDING

If O_{ik} is an exterior vertex of the graph Γ_i, it is clear that the set $\{\boldsymbol{y} = (y_1, \ldots, y_n) : y_i = O_{ik}\}$ is inaccessible for the limiting process $\overline{\boldsymbol{Y}}(t)$ starting from any point that does not belong to this set. An *interior* vertex *can* be reached in finite time; but if O_{ik} is an interior vertex of the graph Γ_i, and O_{jr} of Γ_j, $j \ne i$, it turns out that the set $\{\boldsymbol{y} = (y_1, \ldots, y_n) : y_i = O_{ik}, y_j = O_{jr}\}$ (an $(n-2)$-dimensional part of the binding of the open book $\boldsymbol{\Gamma}$) is inaccessible for $\overline{\boldsymbol{Y}}(t)$ if this process starts outside this set.

This statement is similar to that of inaccessibility of a single point for a two-dimensional Wiener process starting at a different point; and to prove it we need to find the asymptotics of the diffusion coefficients $\overline{a}_i(y_i)$ of the process $\overline{\boldsymbol{Y}}(t)$ as y_i approaches a vertex.

Proposition 5.1. *Let the Hamiltonian H_i be three times continuously differentiable and generic as described in the beginning of the paper. Let $O_{ik} = (l, H_{ik})$ be an interior vertex of the graph Γ_i, and let I_{il} be an edge of Γ_i whose one end is O_{ik}.*

Then there exist constants $A^i_{kl} > 0$ and B^i_{kl} such that

$$\overline{a}_i(l, H) = \frac{A^i_{kl}}{\left|\ln|H - H_{ik}|\right|} + \frac{B^i_{kl}}{\left|\ln|H - H_{ik}|\right|^2} + O\left(\frac{1}{\left|\ln|H - H_{ik}|\right|^2 \cdot \sqrt{|H - H_{ik}|}}\right) \tag{5.1}$$

as $H \to H_{ik}$.

Proof. Let us denote $H - H_{ik} = d$. By the Morse Lemma, we can introduce new coordinates α_1, α_2 instead of ξ_1, ξ_2 in a neighborhood of the saddle point $x_{ik} \in \mathfrak{Y}_i^{-1}(O_{ik})$ so that in this neighborhood $H_i(x) = H_{ik} + \alpha_1 \cdot \alpha_2$. In this neighborhood the curve $C_i(l, H_{ik} + d)$ is described as $\{(\alpha_1, \alpha_2) : \alpha_1\alpha_2 = d\}$, $d \neq 0$, and we have, for the length measure on the curve:

$$\frac{\ell(dx)}{|\nabla H_i(x)|} = A(\alpha_1, d/\alpha_1)\frac{|d\alpha_1|}{|\alpha_1|} = A(d/\alpha_2, \alpha_2)\frac{|d\alpha_2|}{|\alpha_2|},$$

where $A(\alpha_1, \alpha_2)$ is a positive continuously differentiable function.

For $y = (l, H) \in I_{il}$ we have either $H \geq H_{ik}$, or $H \leq H_{ik}$. For definiteness, let it be $H \geq H_{ik}$ (and so $d > 0$).

For small $a > 0$ and $d > 0$, let us denote $C_i(l, H_{ik} + d; < a)$ the part of the curve $C_i(l, H_{ik} + d)$ lying in the neighborhood mentioned above, with $|\alpha_1|$, $|\alpha_2| < a$; $C_i(l, H_{ik} + d; > a)$ will denote the remaining part of $C_i(l, H_{ik} + d)$.

Let us prove that there exist a positive constant C^i_{kl} and a constant D^i_{kl} such that

$$T_i(l, H_{ik} + d) = C^i_{kl} \cdot \left|\ln|d|\right| + D^i_{kl} + O(\sqrt{|d|}) \tag{5.2}$$

as $d \to 0$.

The integral (2.1) defining $T_i(l, H)$ is equal to the integral over $C_i(l, H_{ik} + d; > a)$ plus that over $C_i(l, H_{ik} + d; < a)$.

The first integral is taken over the part of the curve that is far from the equilibrium point, the integrand is smooth, the curve and its ends depend smoothly on d; so it is a continuously differentiable function of d, and

$$\int\limits_{C_i(l,H_{ik}+d;>a)} \frac{1}{|\nabla H_i(x)|}\,\ell(dx) = \int\limits_{C_{i;kl}(>a)} \frac{1}{|\nabla H_i(x)|}\,\ell(dx) + O(d),$$

where $C_{i;kl}(>a)$ is the part of $C_{i;kl}$ with $|\alpha_1|, |\alpha_2| > a$.

The integral over $C_i(l, H_{ik}+d; < a)$, $d > 0$, is equal either to $\int_{d/a}^{a} A(\alpha_1, d/\alpha_1)\frac{d\alpha_1}{\alpha_1}$, or to $\int_{-a}^{-d/a} A(\alpha_1, d/\alpha_1)\frac{d\alpha_1}{|\alpha_1|}$, or to the sum of these two integrals (the first one for one edge of the graph adjacent to O_{ik}, the second integral for another one, and the sum for the third one). Let us find the asymptotics of $\int_{d/a}^{a} A(\alpha_1, d/\alpha_1)\frac{d\alpha_1}{\alpha_1}$.

Changing the variable in some part of the integration range, we write this integral as

$$\int_{\sqrt{d}}^{a} \frac{A(\alpha_1, d/\alpha_1)}{\alpha_1}\,d\alpha_1 + \int_{\sqrt{d}}^{a} \frac{A(d/\alpha_2, \alpha_2)}{\alpha_2}\,d\alpha_2. \tag{5.3}$$

The first integral here is equal to

$$\int_{\sqrt{d}}^{a} \frac{A(0,0)}{\alpha_1}\,d\alpha_1 + \int_0^{a} \frac{A(\alpha_1,0)-A(0,0)}{\alpha_1}\,d\alpha_1 - \int_0^{\sqrt{d}} \frac{A(\alpha_1,0)-A(0,0)}{\alpha_1}\,d\alpha_1 + \int_{\sqrt{d}}^{a} \frac{A(\alpha_1, d/\alpha_1)-A(\alpha_1,0)}{\alpha_1}\,d\alpha_1.$$

The first integral here is equal to $A(0,0)\cdot(\ln a - \frac{1}{2}\ln d)$; the second one converges because the function A is smooth, and it is equal to some constant not depending on d; the third integral is $O(\sqrt{d})$; and the fourth does not exceed in absolute value

$$\int_{\sqrt{d}}^{a} \frac{\mathrm{const}\cdot d/\alpha_1}{\alpha_1}\,d\alpha_1 = \mathrm{const}\cdot d\cdot\left(\frac{1}{\sqrt{d}} - \frac{1}{a}\right) = O\left(\sqrt{d}\right).$$

So the first integral in (5.3) is equal to $\frac{1}{2}A(0,0)\cdot|\ln d| + A(0,0)\cdot\ln a + \int_0^a \frac{A(\alpha_1,0)-A(0,0)}{\alpha_1}\,d\alpha_1 + O(\sqrt{d})$; and we deal with the second integral in (5.3) in a similar way.

If the integral over $C_i(l, H_{ik}+d; < a)$, is equal to $\int_{d/a}^{a} A(\alpha_1, d/\alpha_1)\frac{d\alpha_1}{\alpha_1}$, we get formula (5.2) with $C^i_{kl} = A(0,0)$ and $D^i_{kl} = \int_{C_i(l,H_{ik};>a)} \frac{1}{|\nabla H_i(x)|}\,\ell(dx) + 2\ln a + \int_0^a \frac{A(\alpha_1,0)-A(0,0)}{\alpha_1}\,d\alpha_1 + \int_0^a \frac{A(0,\alpha_2)-A(0,0)}{\alpha_2}\,d\alpha_2$; for the two other cases (and for $d < 0$), the expressions for the constants are different.

In a similar way we prove that

$$\oint\limits_{C_i(l,H_{ik}+d)} \frac{\sum_{r,s=1}^{2} a_{i;rs}(x)\cdot\frac{\partial H_i}{\partial\xi_r}\frac{\partial H_i}{\partial\xi_s}}{|\nabla H_i(x)|}\,\ell(dx) = \oint\limits_{C_{i;kl}} \frac{\sum_{r,s=1}^{2} a_{i;rs}(x)\cdot\frac{\partial H_i}{\partial\xi_r}\frac{\partial H_i}{\partial\xi_s}}{|\nabla H_i(x)|}\,\ell(dx) + O(|d|\ln|d|)$$

as $d \to 0$ (the integral in the right-hand side converges because $H_i(x)$ has a critical point at x_{ik}, and the value of this integral is positive). Together with (5.2), this yields (5.1).

For simplicity of notations, we formulate our next Proposition for the first two coordinates, denoting two interior vertices of Γ_1, Γ_2 with O_{1k}, O_{2k} with the same k, and the edges meeting at them with I_{i1}, I_{i2}, I_{i3}, $i = 1, 2$.

Proposition 5.2. *Suppose that*

$$\overline{a}_i(l, H) = \frac{A^i_{kl}}{\left|\ln|H - H_{ik}|\right|} + \frac{B^i_{kl}}{\left|\ln|H - H_{ik}|\right|^2} + o\left(\frac{1}{\left|\ln|H - H_{ik}|\right|^2 \cdot \ln\left|\ln|H - H_{ik}|\right|}\right)$$

as $H \to H_{ik}$, $i = 1, 2$, $l = 1, 2, 3$; and $\overline{b}_i(\boldsymbol{y})$, $i = 1, \ldots, n$, are bounded near the points $\boldsymbol{y}$ with $y_1 = O_{1k}$, $y_2 = O_{2k}$.

Then the set $\{\boldsymbol{y} : y_1 = O_{1k}, y_2 = O_{2k}\}$ is inaccessible for the process $\overline{\boldsymbol{Y}}(t)$ corresponding to the operator $\overline{\boldsymbol{L}}$ starting from points outside this set.

Proof. For further simplicity of notations, let us introduce new coordinates, denoted with the letter z, along the edges $I_{il} \subseteq \Gamma_i$, $i = 1, 2$, $l = 1, 2, 3$, changing between 0 and some $r_{i;kl}$, so that

$$\overline{a}_i(l, z) = \frac{1}{\left|\ln|z|\right|} + \frac{C^i_{kl}}{\left|\ln|z|\right|^2} + o\left(\frac{1}{\left|\ln|z|\right|^2 \cdot \ln\left|\ln|z|\right|}\right)$$

as $z \downarrow 0$.

Let us define the following functions: for $i = 1, 2$, $y = (l, z)$, $l = 1, 2, 3$,

$$h_i(y) = h_i(l, z) = \begin{cases} z^2 \cdot |\ln z| + D^i_{kl} z^2 & \text{for } 0 < z \le r_{i;kl}, \\ 0 & \text{for } z = 0, \end{cases}$$

where $D^i_{kl} = 1 - C^i_{kl}$; for y_1, y_2 in some neighborhoods of O_{1k}, O_{2k},

$$u = u(y_1, y_2) = h_1(y_1) + h_2(y_2);$$

for $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$ we take

$$\boldsymbol{F}(\boldsymbol{y}) = f(u),$$

where $f(u)$ is a twice continuously differentiable function.

Let us evaluate $\overline{\boldsymbol{L}}\boldsymbol{f}(\boldsymbol{y})$.

We have for $y_i = (l_i, z_i)$, $i = 1, 2$, $1 \le l_i \le 3$:

$$\overline{\boldsymbol{L}}\boldsymbol{f}(\boldsymbol{y}) = f''(u) \cdot \sum_{i=1}^{2} \frac{\overline{a}_i(y_i)}{2} \cdot h'_i(l_i, z_i)^2 + f'(u) \cdot \sum_{i=1}^{2} [\frac{\overline{a}_i(y_i)}{2} \cdot h''_i(l_i, z_i)$$

$$+\overline{b}_i(\boldsymbol{y}) \cdot h_i'(l_i, z_i)\big].$$

At the points at which $f'(u) \neq 0$ this can be written as

$$\overline{\boldsymbol{L}} \boldsymbol{f}(\boldsymbol{y}) = -f'(u) \cdot \left\{ -\frac{d}{du} \ln |f'(u)| \cdot \sum_{i=1}^{2} \frac{\overline{a}_i}{2} \cdot (h_i')^2 - \sum_{i=1}^{2} \left[\frac{\overline{a}_i}{2} \cdot h_i'' + \overline{b}^i \cdot h_i' \right] \right\}.$$

The function $\boldsymbol{F}$ does not belong to $\boldsymbol{D}_0$; however, if we approximate the function f, together with its first and second derivatives, with polynomials f_m, $m = 1, 2, 3, \ldots$, and take $\boldsymbol{F}_m(\boldsymbol{y}) = f_m(u)$, we get the functions $\boldsymbol{F}_m$ that do belong to $\boldsymbol{D}_0$. This is because $\boldsymbol{F}_m(\boldsymbol{y})$ is a linear combination of products of functions of y_i, $i = 1, \ldots, n$; $\frac{\partial f_m(u)}{\partial y_i}\Big|_{y_i=0} = 0$, $i = 1, \ldots, n$, and the gluing conditions are satisfied for every choice of coefficients $\alpha_{i;kl}$.

Using this, we prove easily that $\boldsymbol{F}\left(\overline{\boldsymbol{Y}}(t)\right) - \int_0^t \overline{\boldsymbol{L}} F\left(\overline{\boldsymbol{Y}}(s)\right) ds$, $t \geq 0$, is a martingale with respect to each of the probabilities $\overline{\mathbf{P}}_{\boldsymbol{y}}$, $\boldsymbol{y} \in \boldsymbol{\Gamma}$.

Now let $f(u) = \ln \ln |\ln u|$ for $\rho \leq u \leq R$, where R and $\rho \in (0, R)$ are sufficiently small. Let us check that $\overline{L} \boldsymbol{F}(\boldsymbol{y}) < 0$ for $\rho \leq u \leq R$.

Differentiating, we get:

$$f'(u) = -\frac{1}{u \cdot |\ln u| \cdot \ln |\ln u|}, \quad \frac{d}{du} \ln |f'(u)| = -\frac{1}{u} + \frac{1}{u \cdot |\ln u|} + \frac{1}{u \cdot |\ln u| \cdot \ln |\ln u|}$$

for $\rho < u < R$,

$$h_i'(l_i, z_i) = 2z_i \cdot |\ln z_i| + \left(2D_{kl_i}^i - 1\right) z_i, \qquad h_i''(l_i, z_i) = 2|\ln z_i| + (2D_{kl_i}^i - 3),$$

$$\begin{aligned} \overline{\boldsymbol{L}}\mathrm{F}(\boldsymbol{y}) = -f'(u) \cdot \Bigg\{ &- \frac{d}{du} \ln |f'(u)| \cdot \frac{1}{2} \sum_{i=1}^{2} \left[\frac{1}{|\ln z_i|} + \frac{C_{kl_i}^i}{|\ln z_i|^2} \right. \\ &\left. + o\left(\frac{1}{|\ln z_i|^2 \cdot \ln |\ln z_i|} \right) \right] \times \left[4z_i^2 \cdot |\ln z_i|^2 + 4\left(2D_{kl_i}^i - 1\right) z_i^2 |\ln z_i| + O(z_i^2) \right] \\ &- \frac{1}{2} \sum_{i=1}^{2} \left[\frac{1}{|\ln z_i|} + \frac{C_{kl_i}^i}{|\ln z_i|^2} + o\left(\frac{1}{|\ln z_i|^2 \cdot \ln |\ln z_i|} \right) \right] \cdot [2|\ln z_i| \\ &+ (2D_{kl_i}^i - 3)] - \sum_{i=1}^{2} \overline{b}_i(\boldsymbol{y}) \cdot O(z_i |\ln z_i|) \Bigg\}. \end{aligned} \tag{5.4}$$

Opening the brackets, we obtain that the first sum, together with the factor $1/2$, is equal to $2\sum_{i=1}^{2} [z_i^2 |\ln z_i| + (C_{kl_i}^i + 2D_{kl_i}^i - 1) z_i^2 + o((z^i)^2 / \ln |\ln z^i|)]$. Because of our choice of $D_{kl_i}^i$, we have $C_{kl_i}^i + 2D_{kl_i}^i - 1 = D_{kl_i}^i$, and this sum can be replaced with $2u + \sum_{i=1}^{2} o(z_i^2 / \ln |\ln z_i|)$. Opening the brackets in the second sum in (5.4), we get that this sum, with the factor $(-1/2)$, is equal to $-2 + \sum_{i=1}^{2} 1/2|\ln z_i| + \sum_{i=1}^{2} o(\frac{1}{|\ln z_i| \cdot \ln |\ln z_i|})$; and the last sum in (5.4) also can be included in this $o(\ )$. So the quantity between the braces in formula (5.4) is equal to

$$-\frac{2}{|\ln u|}+\sum_{i=1}^{2}\frac{1}{2|\ln z_i|}-\frac{2}{|\ln u|\cdot\ln|\ln u|}+\frac{1}{u}\cdot\sum_{i=1}^{2}o\left(\frac{z_i^2}{\ln|\ln z_i|}\right)$$
$$+\sum_{i=1}^{2}o\left(\frac{1}{|\ln z_i|\cdot\ln|\ln z_i|}\right). \tag{5.5}$$

Since $h_i(l_i,z_i)\sim z_i^2\cdot|\ln z_i|$, we have:

$$(1+o(1))\cdot(\max_i z_i)^2\cdot|\ln(\max_i z_i)|\le u=\sum_{i=1}^{2}h_i(l_i,z_i)$$
$$\le(2+o(1))\cdot(\max_i z_i)^2\cdot|\ln(\max_i z_i)|,$$

$|\ln u|=2|\ln(\max_i z_i)|-\ln|\ln(\max_i z_i)|+O(1)$, $2/|\ln u|>1/|\ln(\max_i z_i)|$ for sufficiently small u (or z_1, z_2); so $-2/|\ln u|+\sum_{i=1}^{2}1/2|\ln z_i|\le-2/|\ln u|+1/|\ln(\max_i z_i)|<0$. We have $-\frac{2}{|\ln u|\cdot\ln|\ln u|}\sim-\frac{1}{|\ln(\max_i z_i)|\cdot\ln|\ln(\max_i z_i)|}$, and the last two terms in the sum (5.5) are $o(\frac{1}{|\ln(\max_i z_i)|\cdot\ln|\ln(\max_i z_i)|})$.

So the expression (5.5), and with it $\overline{\boldsymbol{L}}\mathbf{F}(\boldsymbol{y})$, is negative for $\boldsymbol{y}$ such that $\rho\le u\le R$.

Let $\boldsymbol{\tau}_{\rho R}$ be the first time that the process $\overline{\boldsymbol{Y}}(t)$ leaves the set $\{\boldsymbol{y}:\rho<u<R\}$. Because $\overline{\boldsymbol{L}}F(\boldsymbol{y})<0$ for $\boldsymbol{y}$ in this set, we get for such $\boldsymbol{y}$

$$\overline{\boldsymbol{E}}_{\boldsymbol{y}}\mathbf{F}\left(\overline{\boldsymbol{Y}}(\boldsymbol{\tau}_{\rho R})\right)\le\mathbf{F}(\boldsymbol{y}),$$

from which

$$\overline{\boldsymbol{P}}_{\boldsymbol{y}}\big\{\overline{\boldsymbol{Y}}(\boldsymbol{\tau}_{\rho R})\in\{\boldsymbol{y}:u=\rho\}\big\}\le\frac{\mathbf{F}(\boldsymbol{y})-f(R)}{f(\rho)-f(R)}=\frac{\mathbf{F}(\boldsymbol{y})-\ln\ln|\ln R|}{\ln\ln|\ln\rho|-\ln\ln|\ln R|}.$$

If we take (for fixed R) ρ small enough, we see that the probability for the process $\overline{\boldsymbol{Y}}(t)$ to reach a small neighborhood of the set $\{\boldsymbol{y}:y_1=O_{1k},y_2=O_{2k}\}$ before reaching the set $\{\boldsymbol{y}:u=R\}$ can be made arbitrarily small. Letting the parameter ρ characterising the smallness of this neighborhood go to 0, we get that the probability to reach the set $\{\boldsymbol{y}:y_1=O_{1k},y_2=O_{2k}\}$ before the set $\{\boldsymbol{y}:u=R\}$ is equal to 0.

After this, only little effort is needed to prove that the process $\overline{\boldsymbol{Y}}(t)$ cannot reach the $(n-2)$-dimensional part $\bigcup_{i\ne j,O_{ik}\in\Gamma_i,O_{jr}\in\Gamma_j}\{\boldsymbol{y}:y_i=O_{ik},y_j=O_{jr}\}$ of the binding with positive probability (if the starting point $\overline{\boldsymbol{Y}}(0)$ is not in this set).

Since the techniques we are using require such things as expectations of $e^{\lambda\boldsymbol{\tau}}\boldsymbol{f}(\boldsymbol{Y}^\varepsilon(\boldsymbol{\tau}))$ and the like, let us formulate the results in such terms:

Proposition 5.3. *For every point $\boldsymbol{y} \in \boldsymbol{\Gamma}$ not belonging to the $(n-2)$-dimensional part of the binding $\boldsymbol{B}_{n-2}$, for every $\lambda > 0$, and for every positive γ there exists a positive d such that*

$$\overline{\boldsymbol{E}}_{\boldsymbol{y}} e^{-\lambda \boldsymbol{\tau}} < \gamma,$$

where $\boldsymbol{\tau}$ is the time at which the process $\overline{\boldsymbol{Y}}(t)$ leaves the region $\boldsymbol{\Gamma}(d)$ defined by (4.2)*; and that, for sufficiently small ε, for every $\boldsymbol{x} \notin \mathfrak{Y}^{-1}(\boldsymbol{B}_{n-2})$, for every $\lambda > 0$, and for every positive γ there exists a positive d such that*

$$\boldsymbol{E}^{\varepsilon}_{\boldsymbol{y}} e^{-\lambda \boldsymbol{\tau}^{\varepsilon}} < \gamma,$$

where $\boldsymbol{\tau}^{\varepsilon}$ is the time at which the process $\boldsymbol{Y}^{\varepsilon}(t)$ leaves $\boldsymbol{\Gamma}(d)$.

Combining this with the results of Secs. 3 and 4, we get

Theorem 1. *Let the conditions imposed on the Hamiltonians H_i and on $\boldsymbol{\beta}(\boldsymbol{x})$ in* Proposition 4.1 *be satisfied. Then for every point $\boldsymbol{x}$ such that $\mathfrak{Y}(\boldsymbol{x})$ does not belong to the $(n-2)$-dimensional part of the binding $\boldsymbol{B}_{n-2}$ (i.e. such that no two coordinates of the point $\mathfrak{Y}(\boldsymbol{x}) \in \boldsymbol{\Gamma}$ are vertices of the corresponding graphs) the function-space distribution of $\boldsymbol{Y}^{\varepsilon}(\bullet)$ with respect to the probability $\boldsymbol{P}^{\varepsilon}_{\boldsymbol{x}}$ converges weakly as $\varepsilon \downarrow 0$ to that of $\overline{\boldsymbol{Y}}(\bullet)$ with respect to the probability $\overline{\boldsymbol{P}}_{\mathfrak{Y}(\boldsymbol{x})}$.*

Remark: It can be proved that the same is true for $\mathfrak{Y}(\boldsymbol{x}) \in \boldsymbol{B}_{n-2}$. The proof is based on the facts that the process $\overline{\boldsymbol{Y}}(t)$ starting from $\boldsymbol{B}_{n-2}$ almost surely leaves this set immediately, and does not return to it (see Propositions 5.1 and 5.2); and for small ε, the process $\boldsymbol{Y}^{\varepsilon}(t)$ starting in a small neigborhood of $\mathfrak{Y}^{-1}(\boldsymbol{B}_{n-2})$ leaves this neighborhood very soon with probability very close to 1, and with probability very close to 1 does not return to a smaller neighborhood for a very long time (see Proposition 5.3).

6. VANISHING-NOISE ASYMPTOTICS

Now we resume our initial notations, with explicit mention of the noise parameter $\varkappa$.

The diffusion process $\overline{\boldsymbol{Y}}^{\varkappa}(t)$ on the open book $\boldsymbol{\Gamma}$, the weak limit of the process $\boldsymbol{Y}^{\varepsilon,\varkappa}(t)$ as $\varepsilon \downarrow 0$, is governed, inside its pages, by the differential operator $\overline{\boldsymbol{L}}^{\varkappa}$ introduced in Sec. 2, which can be rewritten in the form

$$\overline{\boldsymbol{L}}^{\varkappa} \boldsymbol{f}(\boldsymbol{y}) = \overline{\boldsymbol{L}}^{\varkappa} \boldsymbol{f}(y_1, \ldots, y_n) = \sum_{i=1}^{n} \frac{1}{T_i(y_i)} \left[\overline{B}_i(\boldsymbol{y}) \frac{\partial \boldsymbol{f}}{\partial H_i} + \frac{\varkappa}{2} \frac{\partial}{\partial H_i} \left(\overline{A}_i(y_i) \frac{\partial \boldsymbol{f}}{\partial H_i} \right) \right],$$

where

$$\overline{B}_i(\boldsymbol{y}) = \frac{1}{\prod\limits_{1\le j\le n, j\ne i} T_j(y_j)} \oint\limits_{C_1(y_1)} \cdots \oint\limits_{C_n(y_n)} \frac{\beta_i(\boldsymbol{x}) \cdot \nabla H_i(x_i)}{\prod_{i=1}^n |\nabla H_i(x_i)|}\, \ell(dx_1)\cdots\ell(dx_n), \tag{6.1}$$

$$\overline{A}_i(y_i) = \left| \int\limits_{G_i(y_i)} \operatorname{div}[a_i \nabla H_i(x)]\, dx \right|, \quad a_i = \sigma_i\sigma_i^*,$$

and $G_i(y_i)$ is the region in $\mathbb{R}^2$ bounded by the contour $C_i(y_i)$.

The coefficients $\alpha_{i;kl_s}$ in the gluing conditions (2.6), (2.7), defined by formula (2.3), also can be rewritten in the form

$$\alpha_{i;kl_s} = \left| \int\limits_{G_{i;kl_s}} \operatorname{div}[a_i \nabla H_i(x)]\, dx \right|,$$

where $G_{i;kl_s}$ are the regions enclosed by $C_{i;kl_s}$. If the values of $H - H_{ik}$ for $y = (l_3, H) \in I_{il_3}$ have the sign opposite to that for y in the edges I_{il_1}, I_{il_2}, we have $\alpha_{i;kl_3} = \alpha_{i;kl_1} + \alpha_{i;kl_2}$.

Consider one more Markov process $\overline{\boldsymbol{Y}}^0(t)$ on $\boldsymbol{\Gamma}$. Inside each n-dimensional page $\boldsymbol{\gamma}_j \subset \boldsymbol{\Gamma}$, this process is deterministic, being the solution of the differential equation

$$\dot{\overline{\boldsymbol{H}}}^0(t) = \overline{\boldsymbol{\beta}}\left(\overline{\boldsymbol{Y}}^0(t)\right) = \left(\overline{\beta}_1\left(\overline{\boldsymbol{Y}}^0(t)\right), \ldots, \overline{\beta}_n\left(\overline{\boldsymbol{Y}}^0(t)\right)\right), \tag{6.2}$$

where $\overline{\boldsymbol{H}}^0(t) = (\overline{H}_1^0(t), \ldots, \overline{H}_n^0(t))$ is the vector of second coordinates of the components $\overline{Y}_i^0(t)$ of $\overline{\boldsymbol{Y}}^0(t)$: $\overline{\boldsymbol{Y}}^0(t) = (\overline{Y}_1^0(t), \ldots, \overline{Y}_n^0(t)) = (\overline{l}_1^0(t), \overline{H}_1^0(t), \ldots, \overline{l}_n^0(t), \overline{H}_n^0(t))$ (note that the numbers $\overline{l}_1^0(t), \ldots, \overline{l}_n^0(t)$, characterizing the page on which the process $\overline{\boldsymbol{Y}}^0$ is at the time t, remain constant as long as it is in the same page). The coefficients $\overline{\beta}_i$, defined by (2.2), can be rewritten in the form $\overline{\beta}_i(\boldsymbol{y}) = \frac{1}{T_i(y_i)}\overline{B}_i(\boldsymbol{y})$, where $\overline{B}_i$ is given by (6.1).

Now we are going to describe what happens with $\overline{\boldsymbol{Y}}^0(t)$ at the binding. The coefficient $\overline{\beta}_i(y_1, \ldots, y_n)$ is equal to 0 for y_i being a vertex O_{ik} of the graph Γ_i, so there is a solution of (6.2) with $\overline{Y}_i^0(t) \equiv O_{ik}$; but it turns out that sometimes the solution is not unique.

Let us introduce the following functions $\pi_{il}(y_1, \ldots, y_n)$, defined only for y_i being a vertex $O_{ik} = (l, H_{ik})$ of the graph Γ_i, and the edge $I_{il} \sim O_{ik}$:

$$\pi_{il}(\boldsymbol{y}) = \pi_{il}(y_1, \ldots, y_n)$$

$$= \oint\limits_{C_1(y_1)} \cdots \oint\limits_{C_{i-1}(y_{i-1})} \oint\limits_{C_{i;kl}} \oint\limits_{C_{i+1}(y_{i+1})} \cdots \oint\limits_{C_n(y_n)} \frac{\beta_i(x_1, \ldots, x_n) \cdot \nabla H_i(x_i)}{|\nabla H_i(x_i)|}$$
$$\times \mu^1_{y_1}(dx_1) \ldots \mu^{i-1}_{y_{i-1}}(dx_{i-1}) \ell(dx_i) \mu^{i+1}_{y_{i+1}}(dx_{i+1}) \ldots \mu^n_{y_n}(dx_n).$$

For $y_i = O_{ik}$ being an exterior vertex, $C_{i;kl}$ consists of one point x_{ik} (an extremum of the Hamiltonian H_i), and we take $\pi_{il}(y_1, \ldots, y_n) = 0$; at an interior vertex, three $\pi_{il}(y_1, \ldots, y_n)$ are defined, corresponding to the three edges I_{il_s}, $s = 1, 2, 3$, meeting at O_{ik}.

The coefficients π_{il_s} can be rewritten in the form

$$\pi_{il_s}(\boldsymbol{y}) = \pm \oint\limits_{C_1(y_1)} \cdots \oint\limits_{C_{i-1}(y_{i-1})} \iint\limits_{G_{i;kl_s}} \oint\limits_{C_{i+1}(y_{i+1})} \cdots \oint\limits_{C_n(y_n)} \operatorname{div}_i \beta_i(x_1, \ldots, x_n)\, \mu^1_{y_1}(dx_1)$$
$$\times \cdots \mu^{i-1}_{y_{i-1}}(dx_{i-1}) dx_i \mu^{i+1}_{y_{i+1}}(dx_{i+1}) \cdots \mu^n_{y_n}(dx_n), \tag{6.3}$$

where $G_{i;kl_s}$ is the region encircled by the curve $C_{i;kl_s}$, and div_i denotes the divergence taken with respect to the two-dimensional variable x_i (the sign $+$ is taken if the gradient ∇H_i is directed from the region $G_{i;kl_s}$ outside, and $-$ in the opposite case).

The coefficient $\overline{\beta}_i(y_1, \ldots, y_n)$, defined by formula (2.2), for y_i not being a vertex of the graph Γ_i can be written in the form

$$\overline{\beta}_i(y_1, \ldots, y_n)$$
$$= T_i(y_i)^{-1} \oint\limits_{C_1(y_1)} \cdots \oint\limits_{C_{i-1}(y_{i-1})} \oint\limits_{C_i(y_i)} \oint\limits_{C_{i+1}(y_{i+1})} \cdots \oint\limits_{C_n(y_n)} \frac{\beta_i(x_1, \ldots, x_n) \cdot \nabla H_i(x_i)}{|\nabla H_i(x_i)|}$$
$$\times \mu^1_{y_1}(dx_1) \times \cdots \mu^{i-1}_{y_{i-1}}(dx_{i-1}) \ell(dx_i) \mu^{i+1}_{y_{i+1}}(dx_{i+1}) \cdots \mu^n_{y_n}(dx_n). \tag{6.4}$$

If a point $y_i = (l, H)$ approaches a vertex $O_{ik} \in \Gamma_i$ along the edge I_{il} (that is, $H \to H_{ik} = H_i(O_{ik})$), we have:

$$\overline{\beta}_i(y_1, \ldots, y_n) = \frac{\pi_{il}(y_1, \ldots, O_{ik}, \ldots, y_n) + o(1)}{T_i(y_i)}. \tag{6.5}$$

So, if the edge $I_{il} \sim O_{ik}$ consists of points (l, H) with $H \geq H_{ik}$, and a point $\boldsymbol{y} \in \boldsymbol{\Gamma}$ with $y_i \in I_{il}$ is close to a point of the binding with the i-th coordinate being O_{ik}, and $\pi_{il}(y_1, \ldots, O_{ik}, \ldots, y_n)$ is positive, the vector $\overline{\boldsymbol{\beta}}(\boldsymbol{y})$ is directed *towards* the page corresponding to edge $I_{il} \subseteq \Gamma_i$ (away from the binding); if π_{il} is negative, $\overline{\boldsymbol{\beta}}(\boldsymbol{y})$ is directed from the page towards the binding. If $H \leq H_{ik}$ for $(l, H) \in I_{il}$, the directions are opposite.

If not all coefficients $\pi_{il_s}(\boldsymbol{y})$ are 0, then at least one of them is such that the vector $\overline{\boldsymbol{\beta}}(\boldsymbol{y})$ is directed towards the page $\boldsymbol{\gamma}_{il_s}$ corresponding to the edge I_{il_s}, and it turns out that there is a solution of the equation (6.1) with initial condition

$\overline{\boldsymbol{Y}}^0(t_0) = \boldsymbol{y}$ lying strictly inside this page for some time interval (t_0, t_1) (there can be two such pages, but not three, because one of the coefficients $\pi_{il_s}(\boldsymbol{y})$ is equal to the sum of two others).

So, if the process $\overline{\boldsymbol{Y}}^0(t)$ reaches, at time $t = t_0 < \infty$, a point $\boldsymbol{y}$ of the binding with y_i being an internal vertex of Γ_i, and y_j, $j \neq i$, not being vertices of the corresponding graphs, and there is exactly one edge $I_{il_s} \subset \Gamma_i$ such that $\pi_{il_s}(\boldsymbol{y})$ is non-zero and $\overline{\boldsymbol{\beta}}(\boldsymbol{y})$ is directed towards the page $\boldsymbol{\gamma}_{il_s}$ corresponding to this edge, the process goes after the time t_0, without any delay, to this page. If there are *two* such edges, say, I_{il_1} and I_{il_2}, the process goes, without any delay, either to the page $\boldsymbol{\gamma}_{il_1}$ corresponding to the edge I_{il_1}, with probability

$$P_{il_1}(\boldsymbol{y}) = \frac{|\pi_{i;l_1}(\boldsymbol{y})|}{|\pi_{i;l_1}(\boldsymbol{y})| + |\pi_{i;l_2}(\boldsymbol{y})|}, \tag{6.6}$$

or to that corresponding to I_{il_2}, with probability

$$P_{il_2}(\boldsymbol{y}) = \frac{|\pi_{i;l_2}(\boldsymbol{y})|}{|\pi_{i;l_1}(\boldsymbol{y})| + |\pi_{i;l_2}(\boldsymbol{y})|}; \tag{6.7}$$

and this independently from what happened before the time t_0. Note that (6.5), (6.6) can also be rewritten as $|\pi_{i;l_1}(\boldsymbol{y})|/|\pi_{i;l_3}(\boldsymbol{y})|$, $|\pi_{i;l_2}(\boldsymbol{y})|/|\pi_{i;l_3}(\boldsymbol{y})|$.

Let us introduce the set

$$\boldsymbol{E} = \{\boldsymbol{y} = (y_1, \ldots, y_n) \in \boldsymbol{\Gamma} : y_i \text{ is a vertex } O_{ik} \text{ of } \Gamma_i, \text{ and } \pi_{i;kl}(\boldsymbol{y}) = 0$$
$$\text{for at least one edge } I_{il} \sim O_{ik}, \text{ or } y_i = O_{ik}, y_j = O_{jr} \text{ for some } i \neq j\}.$$

The Markov process $\overline{\boldsymbol{Y}}^0(t)$ is determined by the above description before it reaches the exceptional set $\boldsymbol{E}$. We can define it after this time too, say, letting it stop at reaching $\boldsymbol{E}$; but anyway we are not going to make any statement about $\overline{\boldsymbol{Y}}^0(t)$ after reaching this set.

Theorem 2. *Suppose that for some $T > 0$ and some $\boldsymbol{y} \in \boldsymbol{\Gamma}$ the process $\overline{\boldsymbol{Y}}^0(t)$ does not reach the exceptional set $\boldsymbol{E}$ with any positive probability for $t \leq T$. Then the distribution of the random function $\overline{\boldsymbol{Y}}^\varkappa(t)$, $0 \leq t \leq T$, with starting point $\overline{\boldsymbol{Y}}^\varkappa(0) = \boldsymbol{y}$ in the space of continuous functions on the interval $[0, T]$ converges weakly as $\varkappa \downarrow 0$ to that of $\overline{\boldsymbol{Y}}^0(t)$, $0 \leq t \leq T$.*

To prove this theorem, we can check that

(i) the family of the distributions of $\overline{\boldsymbol{Y}}^\varkappa(t)$, $0 \leq t \leq T$, is tight in weak topology;
(ii) if $\{\overline{\boldsymbol{Y}}^{\varkappa_m}(t),\ m = 1, 2, 3, \ldots\}$, $\varkappa^m \to 0$ is a sequence of processes whose distributions converge weakly, then the limiting process coincides with $\overline{\boldsymbol{Y}}^0(t)$.

The tightness takes no more effort than the same for the family of the processes $\boldsymbol{Y}^{\varepsilon,\varkappa}(t)$, and we skip the proof.

The property (ii) can be checked as follows. First, outside a δ-neighborhood $\boldsymbol{B}_\delta$, $\delta > 0$, of the binding the coefficients $\overline{\boldsymbol{\beta}}(\boldsymbol{y})$ are Lipschitz continuous; taking this into account, we prove that $\overline{\boldsymbol{Y}}^{\varkappa}(t)$ converges in probability as $\varkappa \downarrow 0$ to the solution $\overline{\boldsymbol{Y}}^{0}(t)$ of (6.1) with the same initial condition, uniformly on the time interval $[0, T \wedge \min\{t : \overline{\boldsymbol{Y}}^{0}(t) \in \boldsymbol{B}_\delta\}]$.

Second: Suppose the pages $\boldsymbol{\gamma}_{il_1}$, $\boldsymbol{\gamma}_{il_2}$, $\boldsymbol{\gamma}_{il_3}$ are attached to the binding near the point $\boldsymbol{y}_0$ at which the process $\overline{\boldsymbol{Y}}^{0}(t)$ touches the binding from the page $\boldsymbol{\gamma}_{il_3}$; and suppose that the point $\boldsymbol{y}_0$ does not belong to the exceptional set $\boldsymbol{E}$. Let $\boldsymbol{\tau}_\delta^\varkappa = \min\{t : \overline{\boldsymbol{Y}}^{\varkappa}(t) \notin \boldsymbol{B}_\delta\}$. Then the expectation of $\boldsymbol{\tau}_\delta^\varkappa$ for the process $\overline{\boldsymbol{Y}}^{\varkappa}(t)$ starting at the points $\boldsymbol{y} \in \boldsymbol{B}$ that are close to $\boldsymbol{y}_0$ can be estimated: it is bounded by $A\delta|\ln\delta|$, where A is a constant that does not depend on the parameter $\varkappa$.

Third: Let $\boldsymbol{y}_0 \notin \boldsymbol{E}$, $y_{0i} = O_{ik}$, be such that there are two edges I_{il_1}, I_{il_2} with $\pi_{il_1}(\boldsymbol{y}_0)$, $\pi_{il_2}(\boldsymbol{y}_0)$ such that the vector $\overline{\boldsymbol{\beta}}(\boldsymbol{y})$ is directed towards the pages $\boldsymbol{\gamma}_{il_1}$, $\boldsymbol{\gamma}_{il_2}$ for $\boldsymbol{y}$ in these pages being close to $\boldsymbol{y}_0$. Then for every positive γ we can find $\delta_0 > 0$ so small that for every $\delta \in (0, \delta_0)$ there exists a positive $\varkappa_0$ such that for $0 < \varkappa < \varkappa_0$ for $\boldsymbol{y} \in \boldsymbol{B}$ at a distance less than δ_0 from $\boldsymbol{y}_0$

$$|\overline{\boldsymbol{P}}_{\boldsymbol{y}}^{\varkappa}\{\overline{\boldsymbol{Y}}^{\varkappa}(\boldsymbol{\tau}_\delta^\varkappa) \in \boldsymbol{\gamma}_{il_s}\} - P_{il_s}(\boldsymbol{y}_0)| < \gamma, \quad s = 1, 2,$$

$$\overline{\boldsymbol{P}}_{\boldsymbol{y}}^{\varkappa}\{\overline{\boldsymbol{Y}}^{\varkappa}(\boldsymbol{\tau}_\delta^\varkappa) \in \boldsymbol{\gamma}_{il_3}\} < \gamma.$$

The last two statements are obtained from Lemmas 2.2 and 2.3 of Ref. 2 using the (sub)martingale technique.

Theorems 1 and 2 together yield the following

Theorem 3. *Let the conditions imposed on the Hamiltonians H_i and on $\boldsymbol{\beta}(\boldsymbol{x})$ in Proposition 4.1 be satisfied; and let a point $\boldsymbol{x} \in \mathbb{R}^{2n}$ be such that for a positive T the process $\overline{\boldsymbol{Y}}^{0}(t)$ with $\overline{\boldsymbol{Y}}^{0}(0) = \mathfrak{Y}(\boldsymbol{x})$ does not reach the exceptional set $\boldsymbol{E}$ with any positive probability for $0 \le t \le T$.*

Then the slow component $\boldsymbol{Y}^{\varepsilon,\varkappa}(t)$ of the process $\boldsymbol{X}^{\varepsilon,\varkappa}(t)$ converges weakly to $\overline{\boldsymbol{Y}}^{0}(t)$ in the space of continuous functions on $[0, T]$ with values in $\boldsymbol{\Gamma}$ as, first, $\varepsilon \downarrow 0$ and then $\varkappa \downarrow 0$.

7. EXAMPLE

Consider a system of two weakly coupled oscillators which, after an appropriate time change, has the form

$$\dot{q}_1^{\varepsilon,\varkappa} = \frac{1}{\varepsilon} p_1^{\varepsilon,\varkappa}, \quad \dot{p}_1^{\varepsilon,\varkappa} = -\frac{1}{\varepsilon} V_1\big(q_1^{\varepsilon,\varkappa}\big) - p_1^{\varepsilon,\varkappa} + \varkappa \dot{W}_1,$$

$$\dot{q}_2^{\varepsilon,\varkappa} = \frac{1}{\varepsilon} p_2^{\varepsilon,\varkappa}, \quad \dot{p}_2^{\varepsilon,\varkappa} = -\frac{1}{\varepsilon} V_2\big(q_2^{\varepsilon,\varkappa}\big) - \alpha\big(q_1^{\varepsilon,\varkappa}\big) p_2^{\varepsilon,\varkappa} + \varkappa \dot{W}_2, \tag{7.1}$$

where W_1, W_2 are independent Wiener processes. We assume that $\alpha(q) > 0$, $V_1(q)$ has two local minima, and V_2 has just one minimum. The Hamiltonians $H_i(q, p) = \frac{p^2}{2} + V_i(q)$ as well as the corresponding phase pictures and the graphs Γ_i, $i = 1, 2$, are shown in Fig. 1.

The open book $\mathbf{\Gamma} = \Gamma_1 \times \Gamma_2$ is shown in Fig. 2. It consists of three pages $\boldsymbol{\gamma}_l = I_{1l} \times \Gamma_2, l = 1, 2, 3,$ and the binding $\boldsymbol{B} = (\{O_{11}, O_{12}, O_{13}\} \times \Gamma_2) \cup (\Gamma_1 \times \{O_{21}\})$. The only part of the binding that is accessible for the process $\boldsymbol{Y}^{\varepsilon,\varkappa}$ or for the limiting process (and so the only one that we have to take into account) is $\{O_{12}\} \times \Gamma_2$.

The coordinates on $\mathbf{\Gamma}$ are $(l_1, H_1; l_2, H_2)$; since Γ_2 consists of just one edge, l_2 is always the same and can be omitted.

For $y = (l, H) \in \Gamma_i$ not being a vertex, let $S_i(y)$ denote the area of the region in $\mathbb{R}^2$ bounded by the curve $C_i(y)$ corresponding to this point (here $i = 1, 2$). For a fixed edge $I_l \subseteq \Gamma_i$, $S_i(y)$ is in a one-to-one correspondence with $H = H_i(y)$. Note that the function $S_1(y)$ has different limits as y approaches the interior vertex O_{12} along different edges: the limits along I_{11} and I_{12} are the areas S_1, S_2 of the regions G_1, G_2 (see Fig. 1c); that along the edge I_{13} is equal to $S_1 + S_2$.

Let us find the double limit $\overline{\boldsymbol{Y}}^0(t)$ of the slow component $\boldsymbol{Y}^{\varepsilon,\varkappa}(t)$ of the process described by the system (7.1) as first $\varepsilon \downarrow 0$ and then $\varkappa \downarrow 0$.

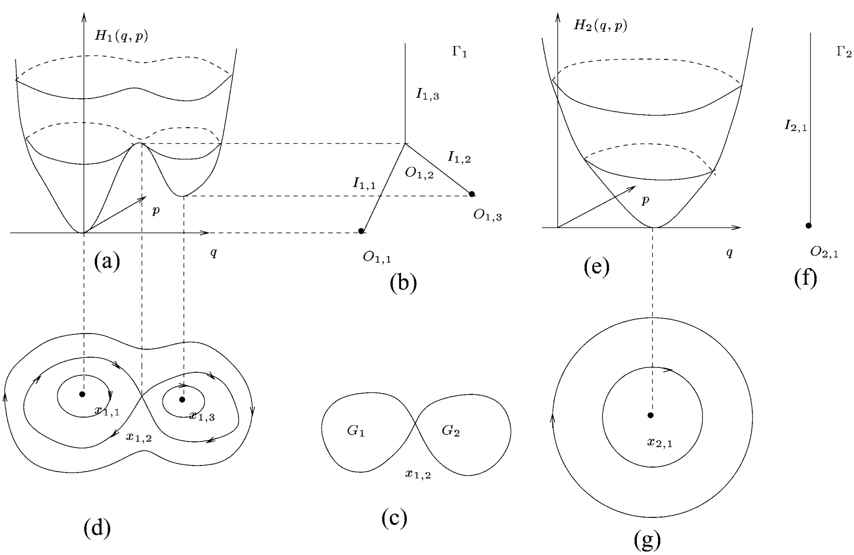

Fig. 1.

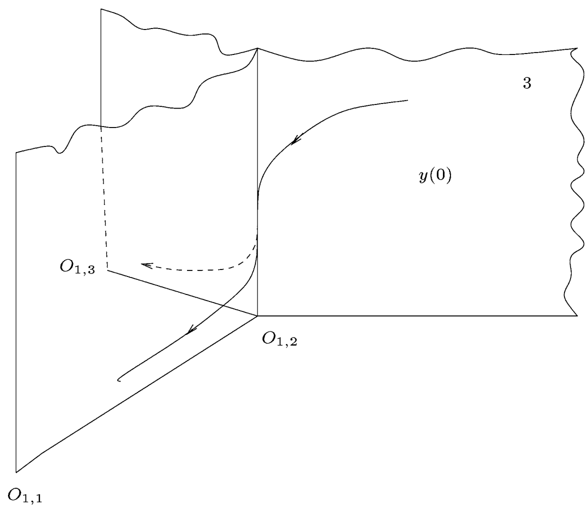

Fig. 2.

Applying formula (6.4), we find, for y_1 being not a vertex:

$$\overline{\beta}_1(y_1, y_2) = T_1(y_1)^{-1} \oint\limits_{C_1(y_1)} \frac{(0, -p) \cdot \nabla H_1(x)}{|\nabla H_1(x)|}\, \ell(dx),$$

where p is the second coordinate of a point $x = (q, p) \in \mathbb{R}^2$. The contour integral here is equal to the double integral of the divergence of the vector field $(0, -p)$ over the region enclosed by $C_1(y_1)$, i.e. to $-S_1(y_1) = -S_1(l_1, H_1)$. The integral (2.1) defining $T_1(y_1)$ is equal to $\frac{dS_1(l_1, H_1)}{dH_1}$; so the differential equation governing the first coordinate $\overline{Y}^0_1(t) = (\overline{l}^0_1(t), \overline{H}^0_1(t))$ of the limiting process $\overline{\boldsymbol{Y}}^0(t)$ while it is within the same page ($\overline{l}^0_1(t) = \text{const}$) is

$$\dot{\overline{H}}{}^0_1 = -\frac{S_1\left(\overline{l}^0_1, \overline{H}^0_1\right)}{S_1'\left(\overline{l}^0_1, \overline{H}^0_1\right)}. \tag{7.2}$$

Similarly, for y_2 not being a vertex of Γ_2,

$$\overline{\beta}_2(y_1, y_2) = \overline{\beta}_2(y_1, 1, H_2) = -\overline{\alpha}(y_1) \cdot \frac{S_2(1, H_2)}{S_2'(1, H_2)},$$

where

$$\overline{\alpha}(y_1) = T_1(y_1)^{-1} \oint\limits_{C_1(y_1)} \frac{\alpha(q)}{|\nabla H_1(x)|}\, \ell(dx);$$

and

$$\dot{\overline{H}}{}_2^0 = -\overline{\alpha}\left(\overline{Y}_1^0\right) \cdot \frac{S_2\left(1, \overline{H}_2^0\right)}{S_2'\left(1, \overline{H}_1^0\right)}. \tag{7.3}$$

Solving the differential equation (7.2), we obtain:

$$S_1\left(\overline{Y}_1^0(t)\right) = S_1\left(\overline{Y}_1^0(0)\right)e^{-t}, \tag{7.4}$$

this formula holding as long as $\overline{Y}_1^0(t)$ is in the same page as the initial value $\overline{Y}_1^0(0)$. For the initial point $\overline{\boldsymbol{Y}}(0)$ in the pages $\boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_2$ (i.e., for $\overline{Y}_1^0(0) \in \Gamma_{11}$ or Γ_{12}), this holds for all $t \in [0, \infty)$; for $\overline{\boldsymbol{Y}}(0) \in \boldsymbol{\gamma}_3$, formula (7.4) holds for $0 \le t \le t_0 = \ln \frac{S_1(\overline{Y}_1^0(0))}{S_1+S_2}$. After this time, $\overline{\boldsymbol{Y}}^0(t)$ goes to one of the pages $\boldsymbol{\gamma}_l$, $l = 1$, 2, and for $t > t_0$ we have:

$$S_1\left(\overline{Y}_1^0(t)\right) = S_l e^{-(t-t_0)}. \tag{7.5}$$

Since within one edge we have a one-to-one correspondence between $S_1(y) = S_1(l, H)$ and y (or H), formulas (7.4), (7.5) allow us to find $\overline{Y}_1^0(t) = (\overline{l}_1^0(t), \overline{H}_1^0(t))$ if we know to which of the pages $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2$ the process $\overline{\boldsymbol{Y}}^0(t)$ turns after the time t_0.

Solving the equation (7.3), we get:

$$S_2\left(\overline{Y}_2^0(t)\right) = S_2\left(1, \overline{H}_2^0(t)\right) = S_2\left(\overline{Y}_2^0(0)\right) \cdot \exp\left\{-\int_0^t \overline{\alpha}\left(\overline{Y}_1^0(s)\right) ds\right\}, \tag{7.6}$$

which allows us to find $\overline{Y}_2^0(t) = (1, \overline{H}_2^0(t))$ (again, if we know to which of the pages $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2$ the process turns after the time t_0).

As for the probabilities with which the process $\overline{\boldsymbol{Y}}^0(t)$ starting in the page $\boldsymbol{\gamma}_3$ turns to $\boldsymbol{\gamma}_1$ or $\boldsymbol{\gamma}_2$, formula (6.3) yields: $\pi_{11}(O_{12}, y_2) = -S_1$, $\pi_{13}(O_{12}, y_2) = -S_2$, $\pi_{12}(O_{12}, y_2) = -S_1 - S_2$, so the probabilities are $P_1 = \frac{S_1}{S_1+S_2}$, $P_2 = \frac{S_2}{S_1+S_2}$ (these probabilities do not depend on the second coordinate y_2 of the point at which $\overline{\boldsymbol{Y}}^0(t)$ reaches the binding).

Equalities (7.4), (7.5), (7.6) together with the probabilities P_1, P_2 describe the limiting slow motion on the open book $\boldsymbol{\Gamma}$ for the coupled system (7.1)—see Fig. 2.

ACKNOWLEDGMENT

The work of the first author was supported in part by the NSF grant DMS0503950.

REFERENCES

1. D.V. Anosov, Averaging in systems of ordinary differential equations with rapidly oscillating solutions. *Izv. Akad. Nauk SSSR, Math.* **24**:721–742 (1960).
2. M. Brin and M. Freidlin, On stochastic behavior of perturbed Hamiltonian systems. *Ergodic Theory Dynam. Syst.* **20**:55–76 (2000).
3. M. Freidlin and A. Wentzell, Random perturbations of Hamiltonian systems. *Mem. Amer. Math. Soc.* **523** (1994).
4. M. Freidlin and A. Wentzell, *Random Perturbations of Dynamical Systems*, 2nd edition (Springer, 1998).
5. M. Freidlin and A. Wentzell, Averaging principle for stochastic perturbations of multifrequency systems. *Stochastics Dynam.* **3**:393–408 (2003).
6. M. Freidlin and A. Wentzell, Averaging principle and stochastic processes on an open book. *Stochastic Process. Appl.* **113**:101–126 (2004).
7. P. Lochak and C. Meunier, Multiphase Averaging for Classical Systems (Springer, 1988).
8. A. Neishtadt, Averaging principle in multi-frequency systems. *Dokl. Akad. Nauk SSSR* **223**:314–317 (1975); Averaging in multi-frequency systems II. *ibid* **226**:1295–1298 (1976).
9. D.W. Stroock and S.R.S.Varadhan, Multidimensional Diffusion Processes (Springer, 1979).