The Annals of Probability
1998, Vol. 26, No. 3, 925–967

RANDOM PERTURBATIONS OF NONLINEAR OSCILLATORS

By Mark Freidlin[1] and Matthias Weber[2]

University of Maryland and Technische Universität Dresden

Degenerate white noise perturbations of Hamiltonian systems in R^2 are studied. In particular, perturbations of a nonlinear oscillator with 1 degree of freedom are considered. If the oscillator has more than one stable equilibrium, the long time behavior of the system is defined by a diffusion process on a graph. Inside the edges the process is defined by a standard averaging procedure, but to define the process for all $t > 0$, one should add gluing conditions at the vertices. Calculation of the gluing conditions is based on delicate Hörmander-type estimates.

1. Introduction. Consider an oscillator with 1 degree of freedom:

$$\ddot{X}_t + f(X_t) = 0, \qquad X_0 = x \in R^1, \qquad \dot{X}_0 = y \in R^1. \tag{1.1}$$

Let $f(x)$ be a smooth enough generic function such that

$$\limsup_{x\to-\infty} f(x) < 0, \qquad \liminf_{x\to\infty} f(x) > 0; \qquad F(x) = \int_0^x f(y)\,dy.$$

One can introduce the Hamiltonian $H(x, y) = \frac{1}{2}y^2 + F(x)$ of system (1.1) and rewrite (1.1) in the Hamiltonian form

$$\begin{aligned} &\dot{X}_t = Y_t \equiv \frac{\partial H}{\partial y}, \qquad \dot{Y}_t = -f(X_t) \equiv -\frac{\partial H}{\partial x}, \\ &X_0 = x, \qquad Y_0 = y. \end{aligned} \tag{1.2}$$

The phase picture of this system is given in Figure 1c. As is known, the Hamiltonian function $H(x, y)$ is a first integral of the system $H(X_t, Y_t) = H(x, y) =$ const. The flow in R^2 defined by system (1.2) preserves the area. The measure on each periodic trajectory with density const./$|\nabla H(x, y)|$ (with respect to the length element dl on the trajectory) is invariant.

Consider now random perturbations of system (1.1) by the white noise

$$\ddot{\tilde{X}}^\varepsilon_t + f(\tilde{X}^\varepsilon_t) = \sqrt{\varepsilon}\dot{W}_t, \qquad \tilde{X}^\varepsilon_0 = x, \qquad \dot{\tilde{X}}^\varepsilon_0 = y, \tag{1.3}$$

where W_t is the Wiener process in R^1 and ε is a small positive parameter. One can rewrite (1.3) as a system:

$$\dot{\tilde{X}}^\varepsilon_t = \tilde{Y}^\varepsilon_t, \qquad \dot{\tilde{Y}}^\varepsilon_t = -f(\tilde{X}^\varepsilon_t) + \sqrt{\varepsilon}\dot{W}_t, \qquad \tilde{X}^\varepsilon_0 = x, \qquad \tilde{Y}^\varepsilon_0 = y. \tag{1.4}$$

Received April 1997; revised January 1998.
[1]Supported by NSF Grant DMS-95-04177.
[2]Supported by a grant of Deutsche Forschungsgemeinschaft.
AMS 1991 *subject classifications*. Primary 60H10, 34C29, 35B20; secondary 35H05.
Key words and phrases. Averaging principle, random perturbations, Hamiltonian systems.

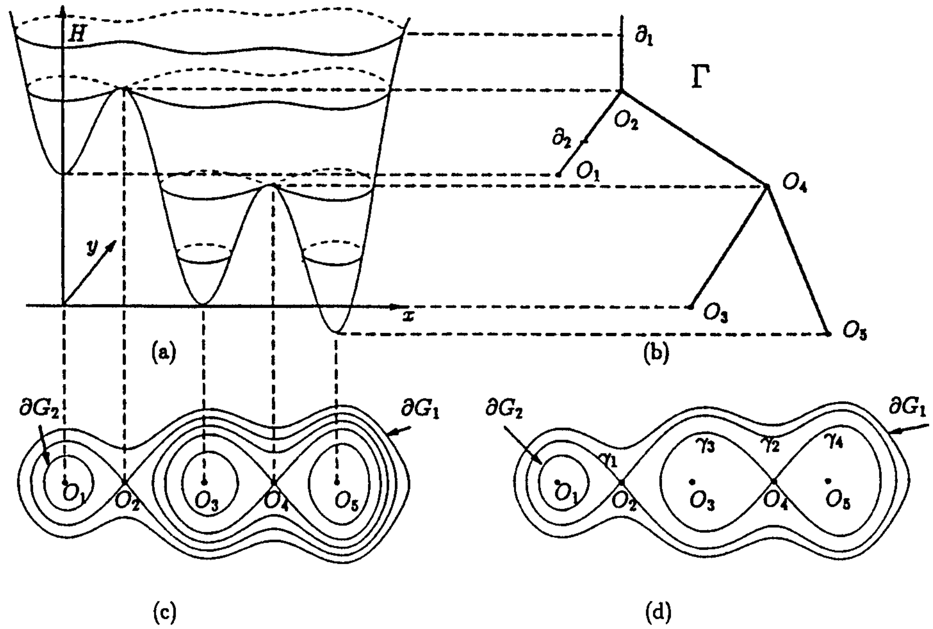

FIG. 1.

The trajectory $(\tilde{X}_t^\varepsilon, \tilde{Y}_t^\varepsilon)$ will be close to the trajectory of system (1.2) with the same initial conditions on any finite time interval if ε is small. More precisely,

$$\lim_{\varepsilon \downarrow 0} P_{x,y}\Big\{\max_{0 \le t \le T}\big(|\tilde{X}_t^\varepsilon - X_t| + |\tilde{Y}_t^\varepsilon - Y_t|\big) > \delta\Big\} = 0$$

for any δ, $T > 0$. Moreover, one can write down, under certain conditions, an asymptotic expansion $\tilde{X}_t^\varepsilon = X_t + \sqrt{\varepsilon} X_t^{(1)} + \varepsilon X_t^{(2)} + \cdots$ in the powers of $\sqrt{\varepsilon}$ valid on a finite time interval, but, as a rule, long time behavior of the perturbed system is of interest. A typical example of such a problem is the exit problem.

Let G denote the set of points (x, y) which are inside ∂G_1 and outside ∂G_2, where ∂G_1 and ∂G_2 are the trajectories of the nonperturbed system (components of the level sets of H) shown in Figure 1. Suppose the system described by (1.3) is working if $(\tilde{X}_t^\varepsilon, \tilde{Y}_t^\varepsilon) \in G$ and is out of service for

$$t \ge \tilde{\tau}^\varepsilon = \min\{t\colon (\tilde{X}_t^\varepsilon, \tilde{Y}_t^\varepsilon) \notin G\}.$$

The expected lifetime of the system $E_{x,y}\tilde{\tau}^\varepsilon$ is of interest [the subscript (x, y) in the expectation sign means the initial point]. Of course, one can write down a boundary problem for the function $\tilde{u}^\varepsilon(x, y) = E_{x,y}\tilde{\tau}^\varepsilon$:

$$\begin{aligned} &\frac{\varepsilon}{2}\frac{\partial^2 \tilde{u}^\varepsilon}{\partial y^2} + y\frac{\partial \tilde{u}^\varepsilon}{\partial x} - f(x)\frac{\partial \tilde{u}^\varepsilon}{\partial y} = -1, \qquad (x, y) \in G, \\ &\tilde{u}^\varepsilon(x, y)\big|_{(x,y)\in\partial G_1} = \tilde{u}^\varepsilon(x, y)\big|_{(x,y)\in\partial G_2} = 0, \end{aligned} \tag{1.5}$$

but it is not simple to solve this degenerate equation even numerically. One can see that $\tilde{u}^\varepsilon(x, y) \to \infty$ as $\varepsilon \downarrow 0$ for $(x, y) \in G$. It follows from the results of this paper that a nontrivial $\lim_{\varepsilon\downarrow 0} \varepsilon\tilde{u}^\varepsilon(x, y) = u(x, y)$ exists, and we calculate $u(x, y)$ explicitly.

To deal with finite time intervals, let us change the time in the process $(\tilde{X}_t^\varepsilon, \tilde{Y}_t^\varepsilon)$: put $X_t^\varepsilon = \tilde{X}_{t/\varepsilon}^\varepsilon$ and $Y_t^\varepsilon = \tilde{Y}_{t/\varepsilon}^\varepsilon$. The process $(X_t^\varepsilon, Y_t^\varepsilon)$ satisfies the equations

$$\dot{X}_t^\varepsilon = \frac{1}{\varepsilon} Y_t^\varepsilon, \qquad \dot{Y}_t^\varepsilon = -\frac{1}{\varepsilon} f(X_t^\varepsilon) + \dot{W}_t, \qquad X_0^\varepsilon = x, \qquad Y_0^\varepsilon = y. \tag{1.6}$$

Here W_t is a Wiener process which is different from W_t in (1.3) or (1.4). The displacement of $(X_t^\varepsilon, Y_t^\varepsilon)$ in a small but independent of ε time interval consists of the fast motion along the deterministic trajectories with "speed" of order ε^{-1} and the slow motion in the direction orthogonal to the deterministic trajectory with speed of order 1 as $\varepsilon \downarrow 0$. The fast component can be characterized by the invariant density const.$/|\nabla H(x, y)|$ on the corresponding nonperturbed trajectory. The slow component, at least locally, is described by the change of $H(X_t^\varepsilon, Y_t^\varepsilon)$. Let, first, the function $f(x)$ have just one zero (see Figure 2). This means that $F(x)$ has one minimum, say, at $x = 0$, $F(0) = 0$, as well as the function $H(x, y) = \frac{1}{2}y^2 + F(x)$. Then the value of $H(x, y)$ defines the deterministic trajectory in a unique way. Denote by $C(z)$, $z \geq 0$, the level set of $H(x, y)$: $C(z) = \{(x, y) \in R^2 \colon H(x, y) = z\}$. Applying the Itô formula to $H(X_t^\varepsilon, Y_t^\varepsilon)$ and taking into account that the gradient $\nabla H(x, y)$ is orthogonal to $\bar{\nabla} H(x, y) = (\partial H/\partial y, -\partial H/\partial x) = (y, -f(x))$, we have

$$H(X_t^\varepsilon, Y_t^\varepsilon) - H(x, y) = \int_0^t \frac{\partial H}{\partial y}(X_s^\varepsilon, Y_s^\varepsilon)\, dW_s + \frac{1}{2}\int_0^t \frac{\partial^2 H}{\partial y^2}(X_s^\varepsilon, Y_s^\varepsilon)\, ds. \tag{1.7}$$

Using the averaging procedure (with respect to the fast motion), it is easy to check that the second term in the right-hand side of (1.7) is equivalent to

$$\frac{t}{2}\int_{C(z)} \frac{H_{yy}(x, y)\, dl}{|\nabla H(x, y)|} \cdot \left(\int_{C(z)} \frac{dl}{|\nabla H(x, y)|}\right)^{-1}$$

for $0 < \varepsilon \ll t \ll 1$, $H(x, y) = z$. Here dl is the length element on $C(z)$.

Using the self-similarity of the Wiener process, the first integral in (1.7) can be written as

$$\tilde{W}_{\int_0^t |H_y(X_s^\varepsilon, Y_s^\varepsilon)|^2\, ds}$$

with a proper Wiener process $\tilde{W}_t$. Using the same averaging procedure, one can see that

$$\int_0^t \left|H_y(X_s^\varepsilon, Y_s^\varepsilon)\right|^2 ds \sim t\int_{C(z)} \frac{H_y^2(x, y)\, dl}{|\nabla H(x, y)|} \cdot \left(\int_{C(z)} \frac{dl}{|\nabla H(x, y)|}\right)^{-1},$$

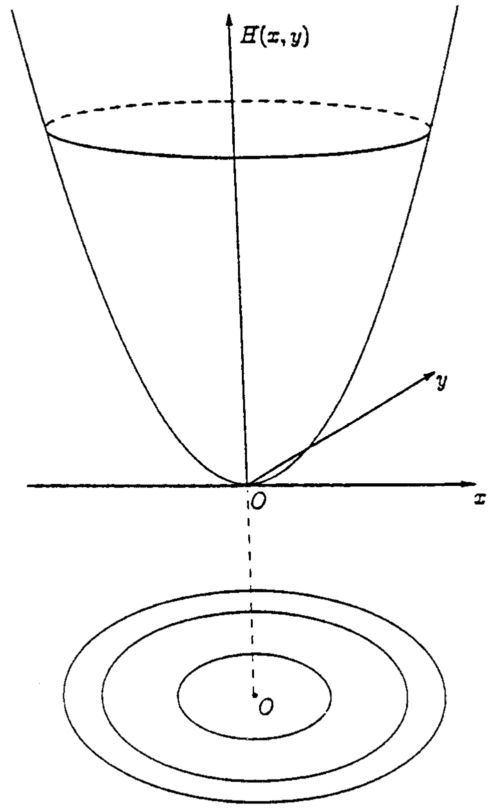

FIG. 2.

$0 < \varepsilon \ll t \ll 1$. This implies that the processes $Z_t^\varepsilon = H(X_t^\varepsilon, Y_t^\varepsilon)$ converge weakly on any finite time interval to the process Z_t governed by the operator

$$L = \frac{1}{2} A(z) \frac{d^2}{dz^2} + B(z) \frac{d}{dz},$$

$$A(z) = \lambda(z)^{-1} \int_{C(z)} \frac{H_y^2(x, y)\, dl}{|\nabla H(x, y)|},$$

$$B(z) = \lambda(z)^{-1} \int_{C(z)} \frac{H_{yy}(x, y)\, dl}{2|\nabla H(x, y)|},$$

$$\lambda(z) = \int_{C(z)} \frac{dl}{|\nabla H(x, y)|}.$$

The process Z_t changes in $R^+ = \{z \in R^1\colon z \geq 0\}$; the point $z = 0$ is inaccessible.

Since $C(z)$ is the level set of $H(x, y)$,

$$a(z) = \int_{C(z)} \frac{H_y^2(x, y)\,dl}{|\nabla H(x, y)|} = \int_{C(z)} (0, H_y) \cdot \frac{\nabla H\,dl}{|\nabla H|}$$

and the last integral is the flux of the vector field $(0, H_y(x, y))$ through the contour $C(z)$. Then, according to the Gauss theorem,

$$a(z) = \int_{G(z)} \operatorname{div}(0, H_y(x, y))\,dx\,dy = \int_{G(z)} H_{yy}(x, y)\,dx\,dy,$$

where $G(z)$ is the domain bounded by $C(z)$. Using this, one can easily derive that

$$\frac{d}{dz}a(z) = \frac{d}{dz}\int_{C(z)} \frac{H_y^2(x, y)\,dl}{|\nabla H(x, y)|} = \int_{C(z)} \frac{H_{yy}(x, y)\,dl}{|\nabla H(x, y)|}.$$

Thus the operator L corresponding to the limiting process Z_t can be written in the form

$$Lv(z) = \frac{1}{2\lambda(z)}\frac{d}{dz}\left(a(z)\frac{dv(z)}{dz}\right), \qquad z \geq 0. \tag{1.8}$$

We will see in Section 3 that explicit expressions for $\lambda(z)$ and for $a(z)$ through the function $f(x)$ can be given.

If, as before, $\tilde{\tau} = \min\{t\colon H(\tilde{X}_t^\varepsilon, \tilde{Y}_t^\varepsilon) \notin (M_1, M_2)\}$, with suitable $0 < M_1 < M_2$, we can conclude from the weak convergence of $H(X_t^\varepsilon, Y_t^\varepsilon)$ to Z_t that

$$\lim_{\varepsilon \downarrow 0} \varepsilon E_{x,y} \tilde{\tau}^\varepsilon = v(H(x, y)),$$

where $v(z)$ is the solution of the problem

$$\begin{aligned} Lv(z) &= -1, \qquad M_1 < z < M_2, \\ v(M_1) &= v(M_2) = 0 \end{aligned} \tag{1.9}$$

and is equal to zero for $z \notin (M_1, M_2)$. Problem (1.9), of course, can be solved explicitly.

Consider now the case of function $f(x)$ with more than one zero. This means that $F(x)$ and $H(x, y)$ have several critical points (see Figure 1). In this case, the set of trajectories can be divided into several families. Inside each family, $H(x, y)$ has different values on different periodic trajectories, but the values of $H(x, y)$ can be the same for trajectories from different families. For example, there are five families shown in Figure 1: trajectories inside γ_3, trajectories inside γ_4, trajectories inside γ_1, trajectories inside γ_2 but outside $\gamma_3 \cup \gamma_4$, trajectories around $\gamma_1 \cup \gamma_2$. The families are separated by the separatrices γ_1, γ_2, γ_3, γ_4. Each periodic trajectory is a connected component of $C(z)$ for some $z \in R$. The trajectory of the nonperturbed system, in the case of $H(x, y)$ with several critical points, is no longer defined by the value of $H(x, y)$ in a unique way. Thus the averaging procedure (along the fast motion) depends not only on the value of $H(x, y)$, but also on the index of the family. In other words, system (1.2) has an additional first integral $k(x, y)$ equal to the index of the family containing the trajectory starting at (x, y). This new discrete first integral $k(x, y)$ is independent of $H(x, y)$. This results in the fact that $H(X_t^\varepsilon, Y_t^\varepsilon)$ does not converge to a Markov process. To have, in the limit, a Markov process, one should extend the phase space by inclusion of the value of $k(x, y)$.

To realize this idea, consider the set of all connected components of the level sets of the Hamiltonian $H(x, y)$ provided with the natural topology. This set is homeomorphic to a graph Γ (see Figure 1). Each periodic trajectory corresponds to an interior point of one of the edges. The equilibrium points, where $H(x, y)$ has maximum or minimum, correspond to the vertices connected just with one edge. Such vertices are called exterior. Each saddle point O of system (1.2), together with two [we assume that $f(x)$ is a generic function] trajectories for which O is an attractor as $t \to \pm\infty$, corresponds to a vertex connected with three edges (interior vertex). For example, the vertices O_2 and O_4 in Figure 1b are interior vertices. The equilibrium point O_2 (O_4) together with the trajectories γ_1 and γ_2 (γ_3 and γ_4) corresponds to the vertex $O_2 \in \Gamma$ ($O_4 \in \Gamma$).

To introduce a coordinate system on Γ, let us index each edge of the graph Γ with a number $1, 2, \dots, n$. Then the value of $H(x, y)$ on the level set component corresponding to a point $P \in \Gamma$ together with the index $i = i(P)$ of the edge containing P forms a coordinate system on Γ. We write $O \sim I_k$ if the vertex O is an end of the edge I_k. If $O \sim I_{k_1}$, $O \sim I_{k_2}$, $O \sim I_{k_3}$ and H_0 is the value of $H(x, y)$ at the equilibrium point corresponding to O, then the coordinates (H_0, k_1), (H_0, k_2) and (H_0, k_3) correspond to the same point O. If a point (z, k) is not a vertex of Γ, it corresponds to a periodic trajectory $C_k(z)$. Each level set $C(z) = \{(x, y)\colon H(x, y) = z\}$ is a union of a finite number of connected components $C_k(z)$. One can define the discrete first integral $k(x, y)$ as the index of the edge $I_k \subset \Gamma$ containing the point corresponding to the periodic trajectory starting at (x, y).

Introduce a mapping $Y\colon R^2 \to \Gamma$ such that $Y(x, y) = (H(x, y), k(x, y)) \in \Gamma$. Consider processes $Y(X_t^\varepsilon, Y_t^\varepsilon) = (H(X_t^\varepsilon, Y_t^\varepsilon),\ k(X_t^\varepsilon, Y_t^\varepsilon))$ on Γ, $\varepsilon > 0$. We prove that these processes converge weakly in the space $C_{0T}(\Gamma)$ of continuous functions $\varphi\colon [0, T] \to \Gamma$ to a diffusion process Y_t on Γ.

A diffusion process on a graph Γ with edges $I_1, \dots, I_n$ and vertices $O_1, \dots, O_m$ is defined by a family of second order elliptic, maybe degenerate, operators $L_1, \dots, L_n$ and by gluing conditions at the vertices [6]. The operator L_k describes the process on I_k until it hits an end O_i of I_k. Then the gluing condition at O_i defines the process. We calculate the operators $L_1, \dots, L_n$ and the gluing conditions at the vertices for the limiting process Y_t on the graph corresponding to the Hamiltonian $H(x, y)$. The operator L_k on I_k, $k = 1, \dots, n$, is defined by formula (1.8), as in the case of one critical point, but one should replace the integration over $C(z)$ in the definition of $\lambda(z)$ and $a(z)$ by the integration over $C_k(z)$.

The gluing conditions for the limiting process are defined by the description of the space of functions on Γ belonging to the domain of definition of the generator of the process. At an interior vertex O_k, $O_k \sim I_{i_1}$, $O_k \sim I_{i_2}$ and $O_k \sim$

I_{i_3}, a smooth function $u(z, i)$ on Γ belongs to the domain of the generator iff

$$\alpha_{ki_1}\frac{du}{dz}(z, i_1)\bigg|_{(z, i_1)=O_k} + \alpha_{ki_2}\frac{du}{dz}(z, i_2)\bigg|_{(z, i_2)=O_k}$$
$$= S\alpha_{ki_3}\frac{du}{dz}(z, i_3)\bigg|_{(z, i_3)=O_k},$$

if $H(Y^{-1}(z, i))$ increases when z approaches O_k for $i = i_1$ and $i = i_2$ and decreases when $i = i_3$. To define the constants α_{ki_1}, α_{ki_2} and α_{ki_3}, one should consider the set $Y^{-1}(O_k)$. This set consists of two loops γ and γ', $\gamma \cap \gamma' = \{O_k\}$. For example, in Figure 1, loops γ_1 and γ_2 connected with O_2; γ_3 and γ_4 connected with O_4. Let γ and γ' be the loops which are the limits of the periodic trajectories corresponding to I_{i_1} and I_{i_2}, respectively. Then

$$\alpha_{ki_1} = \int_\gamma \frac{H_y^2(x, y)}{|\nabla H(x, y)|}\,dl, \qquad \alpha_{ki_2} = \int_{\gamma'} \frac{H_y^2(x, y)}{|\nabla H(x, y)|}\,dl, \qquad \alpha_{ki_3} = \alpha_{ki_1} + \alpha_{ki_2}.$$

No special conditions besides the boundness should be imposed at the exterior vertices. This corresponds to the fact that the process $(X_t^\varepsilon, Y_t^\varepsilon)$ with probability 1 never hits the point corresponding to the exterior vertex. The operators L_k and the gluing conditions define the limiting process in a unique way. The weak convergence of $Y(X_t^\varepsilon, Y_t^\varepsilon)$ to the process Y_t allows us, in particular, to calculate $\lim_{\varepsilon\downarrow 0} \varepsilon E_{x,y}\tilde{\tau}^\varepsilon$ explicitly (see Section 3).

Actually, here we study a slightly more general problem. Consider a Hamiltonian system with 1 degree of freedom:

$$\dot{X}_t = \frac{\partial H}{\partial y}(X_t, Y_t), \qquad \dot{Y}_t = -\frac{\partial H}{\partial x}(X_t, Y_t), \qquad X_0 = x \in R^1, \tag{1.10}$$
$$Y_0 = y \in R^1.$$

The Hamiltonian $H(x, y)$ is assumed to have continuous derivatives of any order, $\lim_{|x|+|y|\to\infty} H(x, y) = \infty$. Moreover, assume that $H(x, y)$ is a generic function. This means that $H(x, y)$ has a finite number of critical points and all of them are nondegenerate. In addition, let any critical value be accepted just at one critical point. Suppose the second of equation (1.10) is perturbed by a small white noise:

$$\dot{\tilde{X}}_t^\varepsilon = \frac{\partial H}{\partial y}(\tilde{X}_t^\varepsilon, \tilde{Y}_t^\varepsilon), \qquad \dot{\tilde{Y}}_t^\varepsilon = -\frac{\partial H}{\partial x}(\tilde{X}_t^\varepsilon, \tilde{Y}_t^\varepsilon) + \sqrt{\varepsilon}\dot{W}_t. \tag{1.11}$$

After the time change $X_t^\varepsilon = \tilde{X}_{t/\varepsilon}^\varepsilon$ and $Y_t^\varepsilon = \tilde{Y}_{t/\varepsilon}^\varepsilon$, we have the following equations for X_t^ε and Y_t^ε:

$$\dot{X}_t^\varepsilon = \frac{1}{\varepsilon}\frac{\partial H}{\partial y}(X_t^\varepsilon, Y_t^\varepsilon), \qquad \dot{Y}_t^\varepsilon = -\frac{1}{\varepsilon}\;\frac{\partial H}{\partial x}(X_t^\varepsilon, Y_t^\varepsilon) + \dot{W}_t, \tag{1.12}$$
$$X_0^\varepsilon = x, \qquad Y_0^\varepsilon = y.$$

Again, if $H(x, y)$ has just one critical point (minimum), $H(X_t^\varepsilon, Y_t^\varepsilon)$ converges as $\varepsilon \downarrow 0$ to a diffusion process which can be calculated using the averaging procedure. If $H(x, y)$ has several critical points, an additional first integral appears, and $H(X_t^\varepsilon, Y_t^\varepsilon)$ no longer converges to a Markov process. One can consider the graph Γ homeomorphic to the set of connected components of the level sets of the Hamiltonian and introduce the mapping $Y\colon R^2 \to \Gamma$ in the same way as above. We show that the processes $Y(X_t^\varepsilon, Y_t^\varepsilon)$ on Γ converge weakly to a diffusion process on Γ and calculate the characteristics of the limiting process. The result is:

THEOREM 1. *Let the Hamiltonian $H(\underset{\sim}{x})$, $\underset{\sim}{x} = (x, y) \in R^2$, be such that:*

(i) $H(\underset{\sim}{x}) \in C^\infty(R^2)$.

(ii) $H(\underset{\sim}{x}) \geq A^1|\underset{\sim}{x}|$, $|\nabla H(\underset{\sim}{x})| \geq A^2|\underset{\sim}{x}|$ *and* $\Delta H(\underset{\sim}{x}) \geq A^3$ *for sufficiently large* $|\underset{\sim}{x}|$, *where* A^1, A^2, *and* A^3 *are positive constants.*

(iii) $H(\underset{\sim}{x})$ *has a finite number of critical points* $\underset{\sim}{x}_1, \dots, \underset{\sim}{x}_N$, *at which the Hessian is nondegenerate.*

(iv) $H(\underset{\sim}{x}_i) \neq H(\underset{\sim}{x}_j)$, $i, j = 1, \dots, N, i \neq j$.

(v) $H_y(\underset{\sim}{x}) = 0 \Rightarrow H_{yy}(\underset{\sim}{x}) \neq 0$.

(vi) $0 < \lim_{\underset{\sim}{x} \in C(H(\underset{\sim}{x}_k)), \underset{\sim}{x} \to \underset{\sim}{x}_k} |H_x/H_y| < \infty$ *for any saddle point* $\underset{\sim}{x}_k$ *of* $H(\underset{\sim}{x})$.

Let $(X_t^\varepsilon, Y_t^\varepsilon; P_{\underset{\sim}{x}}^\varepsilon)$ *be the diffusion process on* R^2 *corresponding to the differential operator* $L^\varepsilon f(\underset{\sim}{x}) = (1/2) f_{yy}(\underset{\sim}{x}) + (1/\varepsilon)\bar{\nabla} H(\underset{\sim}{x}) \cdot \nabla f(\underset{\sim}{x})$, *where* $\bar{\nabla} H(x, y) = (\partial H/\partial y, -\partial H/\partial x)$. *Then the distributions of the processes* $Y(X_t^\varepsilon, Y_t^\varepsilon)$ *in the space of continuous functions with values in* $Y(R^2)$ *with respect to* $P_{\underset{\sim}{x}}^\varepsilon$ *converge weakly as* $\varepsilon \downarrow 0$ *to the probability measure* $P_{Y(\underset{\sim}{x})}$, *where* $(y(t), P_y)$ *is the process on the graph defined by operators* L_i:

$$L_i f_i(H) = \frac{1}{2} A_i(H) f_i''(H) + B_i(H) f_i'(H), \tag{1.13}$$

$$A_i(H) = \frac{\int_{C_i(H)} H_y^2(\underset{\sim}{x}) |\nabla H(\underset{\sim}{x})|^{-1}\, dl}{\int_{C_i(H)} |\nabla H(\underset{\sim}{x})|^{-1}\, dl}, \tag{1.14}$$

$$B_i(H) = \frac{\frac{1}{2}\int_{C_i(H)} H_{yy}(\underset{\sim}{x}) |\nabla H(\underset{\sim}{x})|^{-1}\, dl}{\int_{C_i(H)} |\nabla H(\underset{\sim}{x})|^{-1}\, dl} \tag{1.15}$$

on each edge I_i, *and gluing condition*

$$\sum_{i\colon I_i \sim O_k} \pm \beta_{ki} f_i'(H(\underset{\sim}{x}_k)) = 0, \qquad \beta_{ki} = \int_{C_{ki}} H_y^2(\underset{\sim}{x}) |\nabla H(\underset{\sim}{x})|^{-1}\, dl \tag{1.16}$$

at each interior vertex $O_k = Y(\underset{\sim}{x}_k)$. *The plus sign (+) should be taken in the* i*th term of (1.16) if the coordinate* H *on* I_i *is greater than* $H(\underset{\sim}{x}_k)$, *and the minus sign (–) otherwise. The function* $(H, i) \to f_i(H)$ *should be a continuous function on* Γ *as well as the function* $L_i f_i(H)$. *Further,* $f_i'(H)$ *denotes the derivative with respect to* H, *and* $f_i'(H(\underset{\sim}{x}_k)) = \lim_{H \to H(\underset{\sim}{x}_k), (H, i) \in I_i} f_i'(H)$.

The oscillator (1.1) is a special case of this result when $H(x, y) = \frac{1}{2}y^2 + F(x)$. The characteristics of the limiting process in this case can be calculated more explicitly and they have a simple geometric sense.

Random perturbations of a special equation of type (1.1) describing a phase synchronization model were briefly considered in [3]. Although there is no mathematical description of the limiting process there, the authors mentioned that the limiting process should be considered on a graph. The equation considered in [3] is not generic and the perturbations are a bit different from ours, so the small noise asymptotics for the phase synchronization model does not follow from the results of this paper. Actually, however, it can be calculated in a similar way. We will consider that model elsewhere.

Random perturbations of Hamiltonian systems with 1 degree of freedom in the case of several critical points were studied in [7]. Random perturbations of the vector field $(\partial H/\partial y, -\partial H/\partial x)$ by a nondegenerate white noise were considered there. One can generalize the results of this paper to the case of more general but nondegenerate perturbations.

The specificity of this paper is that just one component of the vector field is perturbed. Such kinds of perturbations are natural in many applied problems. The general approach in this paper is similar to the approach used in [7], although the limiting process is different, since the perturbations are different. The most important difference is that the perturbations now are degenerate. This leads to new serious difficulties in the proof of the Markov property for the limiting process. We overcome these difficulties using the Hörmander-type estimates for degenerate equations. The auxiliary a priori bounds are proved in the next subsection. Then we prove the weak convergence and calculate the characteristics of the limiting process for system (1.12). In the last section, we consider random perturbations of the oscillator (1.1).

We start with a lemma that explains the condition (v) of Theorem 1. Denote $\underline{x} = (x, y)$. Let H_1 and H_2, $H_1 < H_2$, belong to the set of values of $H(x, y)$ on $D_l = \{(x, y) \in R^2\colon Y(x, y) \in I_l\}$ and denote $D_l(H_1, H_2) = \{(x, y) \in D_l\colon H_1 < H(x, y) < H_2\}$.

LEMMA 1.1. *Let $D = D_l(H_1, H_2)$, $-\infty < H_1 < H_2 < \infty$. Then the set $\{\underline{x} \in D; H_y(\underline{x}) = 0\}$ consists of a finite number of mutually disjoint smooth curves $\partial_1, \dots, \partial_{n_l}$, each of which transversely intersects $C_l(H)$ at exactly one point for every $H_1 < H < H_2$.*

PROOF. Let $C_y(0) = \{\underline{x} \in R^2\colon H_y(\underline{x}) = 0\}$. By assumption (v) of Theorem 1 the set $C_y(0) \cap D$ contains no critical point of the function H_y. Thus, $C_y(0) \cap D$ consists of mutually disjoint smooth curves $\partial_1, \dots, \partial_{n_l}$. For $\underline{x} \in \partial_i$, $i \in \{1, \dots, n_l\}$, the vector $\bar{\nabla} H_y(\underline{x})$ is not zero, and it is a tangent vector to the curve ∂_i at $\underline{x}$. The vector $\nabla H(\underline{x})$ is orthogonal to the curve formed by the level set $C(H(\underline{x}))$ at $\underline{x}$. Now the statement of the lemma follows from

$$\bar{\nabla} H_y(\underline{x}) \cdot \nabla H(\underline{x}) \;=\; H_{yy}(\underline{x}) H_x(\underline{x}) \neq 0$$

as $\underline{x} \in C_y(0)$, and D contains no critical point of H. □

2. Proofs.

2.1. *An a priori estimate.* Let H_0, H_1 and H_2, $H_1 < H_0 < H_2$, belong to D_l. Similarly to [7], we introduce new orthogonal coordinates $(\tilde{h}, \theta)$ in $D_l(H_1, H_2) = \{(x, y) \in D_l \colon H_1 < H(x, y) < H_2\}$. The coordinate $\tilde{h}$ is given by

$$\tilde{h}(x, y) = H(x, y) - H_0.$$

The second coordinate θ for $(x, y) \in C_l(H_0)$ is defined after fixing a point $(x_0, y_0) \in C_l(H_0)$:

$$\theta(x, y) = \frac{2\pi \int_{(x_0, y_0)}^{(x, y)} |\nabla H(x, y)|\, dl}{\int_{C_l(H_0)} |\nabla H(x, y)|\, dl}.$$

The integration is taken along $C_l(H_0)$ with respect to the length dl, $0 \le \theta < 2\pi$. To define $\theta(x, y)$ for any point in $D_l(H_1, H_2)$, consider the family of curves orthogonal to $C_l(H)$, $H_1 < H < H_2$, and put $\theta(x, y) = \theta(x', y')$, where (x', y') is the point on $C_l(H_0)$ where the curve of the orthogonal family containing (x, y) crosses $C_l(H_0)$.

We observe that the equation

$$L^\varepsilon u = 0 \tag{2.1.1}$$

can be written in the new coordinates as

$$\left(\left(a_1(\tilde{h}, \theta)\frac{\partial}{\partial \tilde{h}} + a_2(\tilde{h}, \theta)\frac{\partial}{\partial \theta}\right)^2 + \frac{1}{\varepsilon}\, a_3(\tilde{h}, \theta)\frac{\partial}{\partial \theta}\right)u = 0, \tag{2.1.2}$$

where

$$a_1(\tilde{h}(x, y), \theta(x, y)) = \frac{1}{\sqrt{2}} H_y(x, y),$$

$$a_2(\tilde{h}(x, y), \theta(x, y)) = \frac{1}{\sqrt{2}} \theta_y(x, y),$$

$$a_3(\tilde{h}(x, y), \theta(x, y)) = (H_y\theta_x - H_x\theta_y)(x, y).$$

Now we change the coordinate $\tilde{h}$ to $(1/\sqrt{\varepsilon})\tilde{h}$ as follows. Define the operators L_1^ε and L_0^ε,

$$L_1^\varepsilon = a_1^\varepsilon(h, \theta)\frac{\partial}{\partial h} + \sqrt{\varepsilon}\, a_2^\varepsilon(h, \theta)\frac{\partial}{\partial \theta}, \tag{2.1.3}$$

$$L_0^\varepsilon = a_3^\varepsilon(h, \theta)\frac{\partial}{\partial \theta}, \tag{2.1.4}$$

where the coefficients are given by

$$a_i^\varepsilon(h, \theta) = a_i(\sqrt{\varepsilon}\, h, \theta), \qquad i = 1, 2, 3. \tag{2.1.5}$$

With these notations (2.1.2) becomes, after multiplication by ε,

$$\big((L_1^\varepsilon)^2 + L_0^\varepsilon\big)u^\varepsilon = 0, \tag{2.1.6}$$

where $u^\varepsilon(h,\theta)=u(\sqrt{\varepsilon}\,h,\theta)$. Note that there exists a $\bar b>0$ such that $|a_3^\varepsilon|>\bar b$ in $D_l(H_1,H_2)$.

In the following, we make use of the fact that the operator in (2.1.6) is hypoelliptic. Following the steps in Section 22.2 of [8], we derive an estimate as in Lemma 22.2.4 of [8]. This estimate is used to get an a priori estimate for $|u_\theta(0,\theta)|$ for solutions u of (2.1.2). We have to ensure that the estimates of [8] can be obtained independently of ε for small ε.

First, we state some facts about pseudodifferential operators. Let P be a pseudodifferential operator with symbol $p(h,\theta;\xi)$, $\xi=(\xi_1,\xi_2)$. That is,

$$(Pu)(h,\theta)=\frac{1}{(2\pi)^2}\int \exp(i(h\xi_1+\theta\xi_2))p(h,\theta;\xi)\hat u(\xi)\,d\xi,$$

where $\hat u$ denotes the Fourier transform of u. The operator P is of order $\le n$ if for any multi-indices $\alpha=(\alpha_1,\alpha_2)$, $\beta=(\beta_1,\beta_2)$ there exists a constant $A_{\alpha,\beta}$ (depending on α and β) such that

$$\big|(D_x^\beta D_\xi^\alpha\ p)(h,\theta;\xi)\big|\le A_{\alpha,\beta}(1+|\xi|^2)^{(n-|\alpha|)/2} \tag{2.1.7}$$

is satisfied for all $h,\theta,\xi_1,\xi_2\in R$. The smallest n in (2.1.7) is called order of P. Here D_x^β denotes $(-i(\partial/\partial h))^{\beta_1}(-i(\partial/\partial\theta))^{\beta_2}$ and D_ξ^α denotes $(-i(\partial/\partial\xi_1))^{\alpha_1}(-i(\partial/\partial\xi_2))^{\alpha_2}$. Note that for $s\in R$ the norm $\|\cdot\|_s$ is defined by

$$\|u\|_s^2=\frac{1}{(2\pi)^2}\int\big(1+|\xi|^2\big)^s|\hat u(\xi)|^2\,d\xi.$$

We frequently use the following lemma which can be found in Theorems 2.2.1–2.2.3 of [9].

LEMMA 2.1.1. *Let P_1 and P_2 be of order less than or equal to n_1, n_2 with symbols p_1 and p_2, respectively. Assume that the symbols are infinitely differentiable with respect to the variables h, θ, ξ_1, ξ_2, that they have the form $p_i(h,\theta;\xi)=p_i^0(\xi)+p_i^1(h,\theta;\xi)$, $i=1,2$, $\xi=(\xi_1,\xi_2)$, and that for a compact set K',*

$$p_i^1(h,\theta;\xi)=0,\qquad \xi\in R^2,\ (h,\theta)\in R^2\setminus K',\ i=1,2. \tag{2.1.8}$$

Then the following hold for $s\in R$:

(i) $\|P_iu\|_s\le A^i_{s,K'}\|u\|_{s+n_i}$, $i=1,2$.

(ii) *The operator P_1P_2 is of order less than or equal to n_1+n_2, has the property (2.1.8) and*

$$\|P_1P_2u\|_s\le A^{1,2}_{s,K'}\|u\|_{s+n_1+n_2}.$$

(iii) $P_1P_2=P_{(n)}+T_{(n)}$, $n\in N$, *where $P_{(n)}$ has symbol*

$$\sum_{|\alpha|\le n-1}\left(\frac{1}{\alpha!}\frac{\partial^{|\alpha|}}{\partial\xi^\alpha}p_1(h,\theta;\xi)\right)D_x^\alpha p_2(h,\theta;\xi)$$

and

$$\|T_{(n)}u\|_s \le A^{(n)}_{s,K'}\|u\|_{s+n_1+n_2-n}.$$

(iv) *The commutator* $[P_1, P_2] = P_1P_2 - P_2P_1$ *satisfies*

$$\big\|[P_1P_2]u\big\|_s \le A^{[1,2]}_{s,K'}\|u\|_{s+n_1+n_2-1}.$$

Here the constants $A^i_{s,K'}$, $A^{1,2}_{s,K'}$, $A^{(n)}_{s,K'}$ *and* $A^{[1,2]}_{s,K'}$ *depend only on* s, K' *and the estimates (2.1.7) for the operators* P_1 *and* P_2.

REMARK 2.1.2. If we apply a classical differential operator P to functions $u \in C_0^\infty(K)$, where K has positive distance to the boundary of a compact set $K' \supset K$, then we may assume that condition (2.1.8) is satisfied for P as $Pu = gPu$ for a function $g \in C_0^\infty(K')$ with $g = 1$ on K. Thus, we can consider the operator gP instead of P.

From now on, any reference to Lemma 2.1.1 is understood in the manner that the operators under consideration satisfy the assumptions of Lemma 2.1.1 and admit estimates of the form (2.1.7) independent of ε for small ε.

Now let $K' \subset R^2$ be a compact set containing $[-1,1]\times[-2\pi,2\pi]$ and let $K \subset [-1,1]\times[-\pi,\pi]$ be a compact set, such that K has a positive distance to the boundary of K'. From now on, let all ε be small enough to ensure that

$$H_0 + \sqrt{\varepsilon}h \in (H_1, H_2) \quad \text{for all } (h,\theta) \in K'.$$

As $(\partial/\partial h)a_i^\varepsilon(h,\theta) = \sqrt{\varepsilon}(\partial/\partial\tilde{h})a_i(\sqrt{\varepsilon}h,\theta)$, we have for small ε and any multi-index β with constants $A_{i,\beta}$ depending only on β,

$$|D_x^\beta a_i^\varepsilon(h,\theta)| \le \varepsilon^{\beta_1/2}A_{i,\beta}, \qquad i = 1,2,3,\ (h,\theta) \in K'. \tag{2.1.9}$$

Let E_s be the operator with symbol $(1+|\xi|^2)^{s/2}$ and $g \in C_0^\infty(K')$ such that $g = 1$ on K. We identify g with the operator of multiplication by g (having order 0). E_s and g satisfy the conditions of Lemma 2.1.1. By $\|\cdot\|$ and $(\cdot,\cdot)$ we denote the norm and the scalar product in $L_2(R^2)$. Define

$$P^\varepsilon = -(L_1^\varepsilon)^2 - L_0^\varepsilon$$

and $L_0^\varepsilon(h,\theta;\xi) = a_3^\varepsilon(h,\theta)\xi_2$ and $L_1^\varepsilon(h,\theta;\xi) = a_1^\varepsilon(h,\theta)\xi_1 + \sqrt{\varepsilon}a_2^\varepsilon(h,\theta)\xi_2$. Then the principal symbol of P^ε is $p_2^\varepsilon(h,\theta;\xi) = L_1^\varepsilon(h,\theta;\xi)^2$, and $P^\varepsilon = L_1^{\varepsilon*}L_1^\varepsilon + T^\varepsilon$ holds, where $L_1^{\varepsilon*} = -L_1^\varepsilon - w^\varepsilon$, $T^\varepsilon = -L_0^\varepsilon + w^\varepsilon L_1^\varepsilon$ and w^ε is the operator of multiplication by the function

$$w^\varepsilon = \frac{\partial}{\partial h}\,a_1^\varepsilon + \sqrt{\varepsilon}\,\frac{\partial}{\partial\theta}\,a_2^\varepsilon.$$

Note that by (2.1.9), w^ε is uniformly bounded for small ε. In what follows, all constants are independent of ε for small ε if not otherwise stated. Let $D_1 = -i(\partial/\partial h)$ and $D_2 = -i(\partial/\partial\theta)$.

LEMMA 2.1.3. *We have*

$$\|L_1^\varepsilon u\|^2 \le (P^\varepsilon u, u) + A_{K'}\|u\|^2, \qquad u \in C_0^\infty(K), \tag{2.1.10}$$

$$\begin{aligned}\sum_{j=1}^{2}\left\|\left(\frac{\partial}{\partial \xi_j} p_2^\varepsilon\right)(h,\theta;D)u\right\|_0^2 &+ \sum_{j=1}^{2}\left\|(D_j p_2^\varepsilon)(h,\theta;D)u\right\|_{-1}^2 \\ &\le A'_{K'}(P^\varepsilon u, u) + A''_{K'}\|u\|^2, \qquad u \in C_0^\infty(K),\end{aligned} \tag{2.1.11}$$

with constants $A_{K'}$, $A'_{K'}$ and $A''_{K'}$ depending only on K'.

PROOF. The identity

$$\begin{aligned}(P^\varepsilon u, u) &= (L_1^{\varepsilon *} L_1^\varepsilon u, u) + \big(((T^\varepsilon + T^{\varepsilon *})/2)u, u\big) \\ &= \left\|L_1^\varepsilon u\right\|^2 + \big(((T^\varepsilon + T^{\varepsilon *})/2)u, u\big)\end{aligned} \tag{2.1.12}$$

gives (2.1.10) since the operator $T^\varepsilon + T^{\varepsilon *}$ is equal to the operator of multiplication by the function

$$\tilde{w}^\varepsilon = \frac{\partial}{\partial \theta}\big(a_3^\varepsilon - \sqrt{\varepsilon}(w^\varepsilon a_2^\varepsilon)\big) - \frac{\partial}{\partial h}(w^\varepsilon a_1^\varepsilon),$$

which is an operator of order 0 satisfying the conditions of Lemma 2.1.1.

The symbol $(\partial/\partial\xi_1)p_2^\varepsilon$ is given by $2a_1^\varepsilon L_1^\varepsilon$. We have

$$\left\|2a_1^\varepsilon L_1^\varepsilon u\right\|_0^2 = \left\|g 2a_1^\varepsilon L_1^\varepsilon u\right\|_0^2 \le A_1\left\|L_1^\varepsilon u\right\|_0^2$$

by Lemma 2.1.1. So the estimate for $(\partial/\partial\xi_1)p_2^\varepsilon$ follows from (2.1.10). A similar estimate holds for $(\partial/\partial\xi_2)p_2^\varepsilon = 2\sqrt{\varepsilon}a_2^\varepsilon L_1^\varepsilon$. Finally, we get for $j = 1, 2$,

$$\begin{aligned}\left\|(D_j p_2^\varepsilon)(h,\theta;D)u\right\|_{-1}^2 &= \left\|2\big((D_j L_1^\varepsilon)L_1^\varepsilon\big)(h,\theta;D)u\right\|_{-1}^2 \\ &\le \left\|2\big((D_j L_1^\varepsilon)(h,\theta;D)L_1^\varepsilon\big)u\right\|_{-1}^2 \\ &\quad + 2\left\|\big((D_j L_1^\varepsilon)(h,\theta;D)L_1^\varepsilon - \big((D_j L_1^\varepsilon)L_1^\varepsilon\big)(h,\theta;D)\big)u\right\|_{-1}^2.\end{aligned}$$

The first term is equal to

$$2\left\|g\big((D_j L_1^\varepsilon)(h,\theta;D)L_1^\varepsilon\big)u\right\|_{-1}^2 \le A_2\|L_1^\varepsilon u\|_0^2$$

by Lemma 2.1.1(i). Lemma 2.1.1(iii) implies that the operator in the second term is of order 1 and that it admits a corresponding estimate. This completes the proof of (2.1.11). □

Now define the operators (of order 1)

$$Q_1^\varepsilon = L_1^\varepsilon, \qquad Q_2^\varepsilon = \frac{1}{2}(P^\varepsilon - P^{\varepsilon *}), \qquad Q_3^\varepsilon = [Q_1^\varepsilon, Q_2^\varepsilon],$$

$$P_j^\varepsilon = \big(E_{-1}(D_j p_2^\varepsilon)\big)(h,\theta;D), \qquad P^{\varepsilon j} = \left(\frac{\partial}{\partial \xi_j} p_2^\varepsilon\right)(h,\theta;D), \qquad j = 1, 2.$$

Note that $Q_2^\varepsilon = \frac{1}{2}(T^\varepsilon - T^{\varepsilon *}) = T^\varepsilon - \frac{1}{2}\tilde{w}^\varepsilon$.

LEMMA 2.1.4. *For $\delta \le 2^{1-k}$, $k = 1, 2, 3$, we get*

$$\|Q_k^\varepsilon u\|_{\delta-1} \le A_{K',K}\big(\|P^\varepsilon u\| + \|u\|\big), \qquad u \in C_0^\infty(K),$$

with a constant $A_{K',K}$ depending on K' and K.

PROOF. For $k = 1$, the statement follows from Lemma 2.1.3. For $k = 2$, we have to estimate

$$\begin{aligned}\|Q_2^\varepsilon u\|_{-1/2}^2 &= \|E_{-1/2}Q_2^\varepsilon u\|_0^2 = (Q_2^\varepsilon u, E_{-1/2}^2 Q_2^\varepsilon u)\\ &= \big(P^\varepsilon - \tfrac{1}{2}(P^\varepsilon + P^{\varepsilon *})u, E_{-1/2}^2 Q_2^\varepsilon u\big).\end{aligned}$$

Using Lemma 2.1.1(i) we get

$$\begin{aligned}\big|(P^\varepsilon u, E_{-1/2}^2 Q_2^\varepsilon u)\big| &\le A_3\|P^\varepsilon u\|\,\|Q_2^\varepsilon u\|_{-1} = A_3\|P^\varepsilon u\|\,\|gQ_2^\varepsilon u\|_{-1}\\ &\le A_4\|P^\varepsilon u\|\,\|u\|,\end{aligned}$$

as $E_{-1/2}^2$ is of order -1. It remains to consider

$$\begin{aligned}\tfrac{1}{2}\big((P^\varepsilon + P^{\varepsilon *})u, E_{-1/2}^2 Q_2^\varepsilon u\big) &= \big((L_1^{\varepsilon *}L_1^\varepsilon + \tilde{w}^\varepsilon/2)u, E_{-1/2}^2 Q_2^\varepsilon u\big)\\ &= \big(\tilde{P}^\varepsilon u, E_{-1/2}^2 Q_2^\varepsilon u\big) - A_5\big(u, E_{-1/2}^2 Q_2^\varepsilon u\big)\end{aligned}$$

by (2.1.12), where $\tilde{P}^\varepsilon = L_1^{\varepsilon *}L_1^\varepsilon + \tilde{w}^\varepsilon/2 + A_5 g$ with a constant A_5 large enough to ensure that $\tilde{P}^\varepsilon$ is a positive operator. We get

$$\big|(A_5 u, E_{-1/2}^2 Q_2^\varepsilon u)\big| \le A_6\|u\|^2$$

by Lemma 2.1.1, and

$$\begin{aligned}\big|\big(\tilde{P}^\varepsilon u, E_{-1/2}^2 Q_2^\varepsilon u\big)\big| &= \big|\big(\tilde{P}^\varepsilon u, g_1 E_{-1/2}^2 Q_2^\varepsilon u\big)\big|\\ &\le (\tilde{P}^\varepsilon u, u)^{1/2}\big(\tilde{P}^\varepsilon g_1 E_{-1/2}^2 Q_2^\varepsilon u, g_1 E_{-1/2}^2 Q_2^\varepsilon u\big)^{1/2}\\ &\le \tfrac{1}{2}\big((P^\varepsilon u, u) + \big(g_2 P^\varepsilon g_1 E_{-1/2}^2 Q_2^\varepsilon u, g_1 E_{-1/2}^2 Q_2^\varepsilon u\big)\big)\\ &\quad + A_5\|u\|^2 + A_5\big(\tilde{g}_1 E_{-1/2}^2 Q_2^\varepsilon u, g_1 E_{-1/2}^2 Q_2^\varepsilon u\big)\\ &\le A_7\big(\|P^\varepsilon u\| + \|u\|\big)^2 + \big(g_2 P^\varepsilon g_1 E_{-1/2}^2 Q_2^\varepsilon u, g_1 E_{-1/2}^2 Q_2^\varepsilon u\big).\end{aligned}$$

Here the functions are $g_1 \in C_0^\infty(K'')$ and $g_1 = 1$ on K, $g_2 \in C_0^\infty(K')$ and $g_2 = 1$ on K'' and $\tilde{g}_1 = gg_1g_2$ with a set K'', $K \subset K'' \subset K'$ such that the boundary of K'' has positive distance to the boundaries of K and K'. Define $Q^\varepsilon = g_1 E_{-1/2}^2 Q_2^\varepsilon$. By Lemma 2.1.1(iii), the operator $[g_2 P^\varepsilon, Q^\varepsilon]$ has symbol

$$\begin{aligned}&\sum_{j=1}^2 (D_j q^\varepsilon)(h,\theta;\xi) g_2(h,\theta)\frac{\partial}{\partial \xi_j} p_2^\varepsilon(h,\theta;\xi)\\ &\qquad - \sum_{j=1}^2 \frac{\partial}{\partial \xi_j} q^\varepsilon(h,\theta;\xi)(D_j g_2 p_2^\varepsilon)(h,\theta;\xi) + r_1^\varepsilon,\end{aligned}$$

where q^ε denotes the symbol of Q^ε and r_1^ε is the symbol of an operator of order 0 which can be estimated by Lemma 2.1.1(iii). By Lemma 2.1.1, the operator $[g_2P^\varepsilon, Q^\varepsilon]$ differs from the operator

$$\sum_{j=1}^{2}(D_jq^\varepsilon)(h,\theta;D)g_2(h,\theta)P^{\varepsilon j}(h,\theta;D)$$
$$-\sum_{j=1}^{2}\frac{\partial}{\partial\xi_j}q^\varepsilon(h,\theta;D)(D_jg_2p_2^\varepsilon)(h,\theta;D)$$

only by an operator of order 0 satisfying the conditions of Lemma 2.1.1. Denote $G^{\varepsilon j}=(D_jq^\varepsilon)(h,\theta;D)g_2(h,\theta)$ and $G_j^\varepsilon=(\partial/\partial\xi_j)q^\varepsilon(h,\theta;D)$, which are of order 0 and -1, respectively. We get

$$(g_2P^\varepsilon Q^\varepsilon u, Q^\varepsilon u)=(Q^\varepsilon g_2P^\varepsilon u, Q^\varepsilon u)+\big([g_2P^\varepsilon, Q^\varepsilon]u, Q^\varepsilon u\big). \tag{2.1.13}$$

From Lemma 2.1.1 it follows that

$$(Q^\varepsilon g_2P^\varepsilon u, Q^\varepsilon u)=\big(g_1E^2_{-1/2}Q_2^\varepsilon g_2P^\varepsilon u, g_1E^2_{-1/2}Q_2^\varepsilon u\big)\le A_8\|P^\varepsilon u\|\,\|u\|$$

and

$$\begin{aligned}
&\big([g_2P^\varepsilon, g_1E^2_{-1/2}Q_2^\varepsilon]u, g_1E^2_{-1/2}Q_2^\varepsilon u\big)\\
&\quad=\sum_{j=1}^{2}\Big\{\big(G^{\varepsilon j}g_2P^{\varepsilon j}u, g_1E^2_{-1/2}Q_2^\varepsilon u\big)\\
&\qquad\qquad-\big(G_j^\varepsilon D_j(g_2p_2^\varepsilon)(h,\theta;D)u, g_1E^2_{-1/2}Q_2^\varepsilon u\big)\Big\}\\
&\qquad+(R_2^\varepsilon u, g_1E^2_{-1/2}Q_2^\varepsilon u)\\
&\quad\le A_9\bigg(\sum_{j=1}^{2}\|P^{\varepsilon j}\|\,\|u\|+\|u\|^2\bigg)\\
&\qquad-\sum_{j=1}^{2}\big(G_j^\varepsilon D_j(g_2p_2^\varepsilon)(h,\theta;D)u, g_1E^2_{-1/2}Q_2^\varepsilon u\big),
\end{aligned}$$

where R_2^ε is of order 0 and is bounded by Lemma 2.1.1. As $g_2=1$ on $\operatorname{supp}u$, $D_j(g_2p_2^\varepsilon)(\cdot,\cdot;D)u=g_2(D_jp_2^\varepsilon)(\cdot,\cdot;D)u$ holds. This yields

$$\begin{aligned}
\big\|G_j^\varepsilon D_j(g_2p_2^\varepsilon)(h,\theta;D)u\big\|&=\big\|G_j^\varepsilon g_2E_1E_{-1}D_jp_2^\varepsilon(h,\theta;D)u\big\|\\
&=\big\|G_j^\varepsilon g_2E_1P_j^\varepsilon u\big\|\le A_{10}\big\|P_j^\varepsilon u\big\|
\end{aligned}$$

as G_j^ε has order -1 and g_2E_1 has order 1 and both satisfy the conditions of Lemma 2.1.1. This together with Lemma 2.1.3 completes the proof for $k=2$. For $k=3$ we have to estimate

$$\big\|[Q_1^\varepsilon, Q_2^\varepsilon]u\big\|^2_{-3/4}=\big([Q_1^\varepsilon, Q_2^\varepsilon]u, \tilde{Q}^\varepsilon u\big),$$

where $\tilde{Q}^\varepsilon = E^2_{-3/4}[Q^\varepsilon_1, Q^\varepsilon_2]$ has order $\frac{1}{2}$. Observe that $Q^{\varepsilon*}_1 = -Q^\varepsilon_1 - w^\varepsilon$ and $Q^{\varepsilon*}_2 = -Q^\varepsilon_2$. Hence

$$\begin{aligned}\big([Q^\varepsilon_1, Q^\varepsilon_2]u, \tilde{Q}^\varepsilon u\big) &= (Q^\varepsilon_1 Q^\varepsilon_2 u, g\tilde{Q}^\varepsilon u) - (Q^\varepsilon_2 Q^\varepsilon_1 u, g\tilde{Q}^\varepsilon u)\\ &= -(Q^\varepsilon_2 u, gQ^\varepsilon_1 g\tilde{Q}^\varepsilon u) - (Q^\varepsilon_2 u, gw^\varepsilon \tilde{Q}^\varepsilon u) + (Q^\varepsilon_1 u, gQ^\varepsilon_2 g\tilde{Q}^\varepsilon u).\end{aligned}$$

Using Lemma 2.1.1 and the cases $k = 1$ and 2, we can estimate

$$\begin{aligned}\big|(Q^\varepsilon_2 u, gQ^\varepsilon_1 g\tilde{Q}^\varepsilon u)\big| &\le \|Q^\varepsilon_2 u\|_{-1/2} \big\|gQ^\varepsilon_1 g\tilde{Q}^\varepsilon u\big\|_{1/2}\\ &\le A_{11}\big(\|P^\varepsilon u\| + \|u\|\big)\big(\big\|g\tilde{Q}^\varepsilon gQ^\varepsilon_1 u\big\|_{1/2} + \big\|[gQ^\varepsilon_1, g\tilde{Q}^\varepsilon]u\big\|_{1/2}\big)\\ &\le A_{12}\big(\|P^\varepsilon u\| + \|u\|\big)^2,\end{aligned}$$

as $g\tilde{Q}^\varepsilon g$ and $[gQ^\varepsilon_1, g\tilde{Q}^\varepsilon]$ are of order $-\frac{1}{2}$,

$$\big|(Q^\varepsilon_2 u, gw^\varepsilon \tilde{Q}^\varepsilon u)\big| \le \big\|Q^\varepsilon_2 u\big\|_{-1/2} \big\|gw^\varepsilon \tilde{Q}^\varepsilon u\big\|_{1/2} \le A_{13}\big(\|P^\varepsilon u\| + \|u\|\big)^2,$$

as $gw^\varepsilon \tilde{Q}^\varepsilon$ is of order $-\frac{1}{2}$, and

$$\big|(Q^\varepsilon_1 u, gQ^\varepsilon_2 g\tilde{Q}^\varepsilon u)\big| \le \|Q^\varepsilon_1 u\|_0 \|gQ^\varepsilon_2 g\tilde{Q}^\varepsilon u\|_0 \le A_{14}\big(\|P^\varepsilon u\| + \|u\|\big)^2,$$

as $\|gQ^\varepsilon_2 g\tilde{Q}^\varepsilon u\|_0 = \|g\tilde{Q}^\varepsilon gQ^\varepsilon_2 u\|_0 + \|[gQ^\varepsilon_2, g\tilde{Q}^\varepsilon]u\|_0 \le A_{15}(\|Q^\varepsilon_2 u\|_{-1/2} + \|u\|)$, and $g\tilde{Q}^\varepsilon g$ and $[gQ^\varepsilon_2, g\tilde{Q}^\varepsilon]$ have order $-\frac{1}{2}$. □

LEMMA 2.1.5. *We have*

$$\|u\|_{1/4} \le \tilde{A}_{K'}\big(\|P^\varepsilon u\| + \|u\|\big), \qquad u \in C^\infty_0(K), \tag{2.1.14}$$

with a constant $\tilde{A}_{K'}$ depending on K'.

PROOF. The statement follows from Lemma 2.1.4 if we show that there exists a constant $A_{K'}$ depending only on K' such that

$$\|u\|_{1/4} = \|E_1 u\|_{-3/4} \le A_{K'}\bigg(\sum_{j=1}^{3} \|Q^\varepsilon_j u\|_{-3/4} + \|u\|\bigg).$$

As

$$\|u\|^2_{1/4} = \|u\|^2_{-3/4} + \bigg\|\frac{\partial}{\partial h}u\bigg\|^2_{-3/4} + \bigg\|\frac{\partial}{\partial \theta}u\bigg\|^2_{-3/4},$$

we have to show that

$$\|D_i u\|_{-3/4} \le A_{16}\bigg(\sum_{j=1}^{3} \|Q^\varepsilon_j u\|_{-3/4} + \|u\|\bigg), \qquad i = 1, 2.$$

Using (2.1.9) we can write

$$Q_1^\varepsilon = a_1^\varepsilon \frac{\partial}{\partial h} + \sqrt{\varepsilon} a_2^\varepsilon \frac{\partial}{\partial \theta}, \tag{2.1.15}$$

$$Q_2^\varepsilon = \sqrt{\varepsilon} b_1^\varepsilon \frac{\partial}{\partial h} - (a_3^\varepsilon + \sqrt{\varepsilon} b_3^\varepsilon)\frac{\partial}{\partial \theta} - \frac{1}{2}\left(\frac{\partial}{\partial \theta} a_3^\varepsilon\right) - \sqrt{\varepsilon} b_2^\varepsilon, \tag{2.1.16}$$

$$Q_3^\varepsilon = \left(a_3^\varepsilon \left(\frac{\partial}{\partial \theta} a_1^\varepsilon\right) + \sqrt{\varepsilon} c_1^\varepsilon\right)\frac{\partial}{\partial h} + \sqrt{\varepsilon} c_2^\varepsilon \frac{\partial}{\partial \theta} + \sqrt{\varepsilon} c_3^\varepsilon \tag{2.1.17}$$

with functions b_i^ε, c_i^ε, $i = 1, 2, 3$, uniformly bounded on K' for small ε. Further,

$$\begin{aligned}\left(a_3^\varepsilon \frac{\partial}{\partial \theta} a_1^\varepsilon\right)(h, \theta) &= \left(a_3 \frac{\partial}{\partial \theta} a_1\right)\left(\tilde{h}(x, y), \theta(x, y)\right) \\ &= \frac{1}{\sqrt{2}}\left((H_y \theta_x - H_x \theta_y)\frac{\partial}{\partial \theta} H_y\right)(x, y) \\ &= \frac{1}{\sqrt{2}}\left(H_y H_{xy} - H_x H_{yy}\right)(x, y).\end{aligned} \tag{2.1.18}$$

Note that by our assumptions the right-hand side of (2.1.18) is not equal to 0 in a neighborhood of the zeros of H_y. So we can divide the set K' into two disjoint sets K_1 and K_2 such that there exist positive numbers δ_i, $i = 1, 2$, and a set $K_2' \supset K_2$ such that

$$|a_1^\varepsilon| > 2\delta_1 \quad \text{on } K_1, \tag{2.1.19}$$

$$|a_0^\varepsilon| = \left|\left(a_3^\varepsilon \frac{\partial}{\partial \theta} a_1^\varepsilon\right)\right| > 2\delta_2 \quad \text{on } K_2'. \tag{2.1.20}$$

Let $\varphi \in C_0^\infty(K' \setminus K_2)$ be a function with $\varphi = 1$ on $K \setminus K_2'$. On the set $K \setminus K_2'$ we can regard the identities (2.1.15) and (2.1.16) applied to $u \in C_0^\infty(K)$ as a linear system for u_h and u_θ with the (unique) solution

$$\begin{aligned}u_h = \frac{1}{D_1^\varepsilon}\bigg(&-\left(a_3^\varepsilon + \sqrt{\varepsilon} b_3^\varepsilon\right) Q_1^\varepsilon u + \sqrt{\varepsilon} a_2^\varepsilon Q_2^\varepsilon u \\ &+ \sqrt{\varepsilon} a_2^\varepsilon \left(\frac{1}{2}\left(\frac{\partial}{\partial \theta} a_3^\varepsilon\right) + \sqrt{\varepsilon} b_2^\varepsilon\right) u\bigg),\end{aligned} \tag{2.1.21}$$

$$u_\theta = \frac{1}{D_1^\varepsilon}\left(-\sqrt{\varepsilon} b_1^\varepsilon Q_1^\varepsilon u + a_1^\varepsilon Q_2^\varepsilon u + a_1^\varepsilon \left(\frac{1}{2}\left(\frac{\partial}{\partial \theta} a_3^\varepsilon\right) + \sqrt{\varepsilon} b_2^\varepsilon\right) u\right), \tag{2.1.22}$$

where $D_1^\varepsilon = -a_1^\varepsilon a_3^\varepsilon - \sqrt{\varepsilon}(a_1^\varepsilon b_3^\varepsilon + \sqrt{\varepsilon} a_2^\varepsilon b_1^\varepsilon)$.

Recall that there exists a constant $\bar{b} > 0$ such that

$$|a_3^\varepsilon(h, \theta)| \geq \bar{b}, \qquad (h, \theta) \in K'. \tag{2.1.23}$$

This implies

$$|D_1^\varepsilon(h, \theta)| \geq \bar{b}\ \delta_1, \qquad (h, \theta) \in K' \setminus K_2, \tag{2.1.24}$$

for ε small enough. Thus, the coefficients of $Q_i^\varepsilon u$, $i = 1, 2$, and u in (2.1.21) and (2.1.22) can be understood as operators of order 0 satisfying (after multiplication by φ and g) the conditions of Lemma 2.1.1 in $K' \setminus K_2$. This gives

$$\begin{aligned} \|\varphi D_j u\|_{-3/4} &= \|g\varphi D_j u\|_{-3/4} \\ &\le A_{17}\Big(\sum_{j=1}^{2} \|Q_j^\varepsilon u\|_{-3/4} + \|u\|\Big), \qquad j = 1, 2. \end{aligned} \tag{2.1.25}$$

We can make the analogous considerations on the set K_2' using the relations (2.1.16) and (2.1.17) instead of (2.1.15) and (2.1.16). The resulting formulas are similar to (2.1.21) and (2.1.22), but the determinant of the system is

$$D_2^\varepsilon = a_0^\varepsilon a_3^\varepsilon + \sqrt{\varepsilon}\big(\sqrt{\varepsilon} b_1^\varepsilon c_2^\varepsilon + c_1^\varepsilon\big(a_3^\varepsilon - \sqrt{\varepsilon} b_3^\varepsilon\big) - a_0^\varepsilon b_3^\varepsilon\big).$$

By (2.1.20) and (2.1.23) we get

$$|D_2^\varepsilon(h, \theta)| \ge \bar{b}\delta_2, \qquad (h, \theta) \in K_2', \tag{2.1.26}$$

for ε small enough, so that

$$\begin{aligned} \big\|(1-\varphi)D_j u\big\|_{-3/4} &= \big\|g(1-\varphi)D_j u\big\|_{-3/4} \\ &\le A_{18}\Big(\sum_{j=1}^{2} \big\|Q_j^\varepsilon u\big\|_{-3/4} + \|u\|\Big), \end{aligned} \tag{2.1.27}$$

$j = 1, 2$. Combining the relations (2.1.25) and (2.1.27), we get the desired result. $\square$

LEMMA 2.1.6. *For $s \in R$ we have*

$$\|u\|_{s+1/4} + \sum_{j=1}^{2} \|P_j^\varepsilon u\|_{s+1/8} + \sum_{j=1}^{2} \|P^{\varepsilon j} u\|_{s+1/8} \le A_{s, K', K}\big(\|P^\varepsilon u\|_s + \|u\|_s\big),$$

$u \in C_0^\infty(K)$, *with constants $A_{s, K', K}$ depending on s, K' and K.*

PROOF. By (2.1.11) we get

$$\big\|P_j^\varepsilon u\big\| \le A_{19}\big(\|P^\varepsilon u\|_{-1/8}^{1/2} \|u\|_{1/8}^{1/2} + \|u\|\big)$$

and the corresponding estimate for $P^{\varepsilon j}$, $j = 1, 2$. Thus we have for every $B > 0$,

$$\sum_{j=1}^{2} \big(\|P_j^\varepsilon u\| + \|P^{\varepsilon j} u\|\big) \le A_{20}\big(B\|P^\varepsilon u\|_{-1/8} + B^{-1}\|u\|_{1/8} + \|u\|\big). \tag{2.1.28}$$

Observe that by Lemma 2.1.1,

$$\begin{aligned} \|u\|_{s+1/4} &= \|E_s u\|_{1/4} = \|g_1 E_s u\|_{1/4} + \|[E_s, g_1]u\|_{1/4} \\ &\le \|g_1 E_s u\|_{1/4} + A_{21}\|u\|_s, \end{aligned} \tag{2.1.29}$$

$$
\begin{aligned}
\|P_j^\varepsilon u\|_{s+1/8} &\le \|g_1E_{-1}(D_jp_2^\varepsilon)(h,\theta;D)u\|_{s+1/8} \\
&\quad + \|[E_{-1},g_1]g_1(D_jp_2^\varepsilon)(h,\theta;D)u\|_{s+1/8} \\
&\le \|E_{s+1/8}g_1P_j^\varepsilon u\|_0 + A_{22}\|u\|_{s+1/8} \\
&\le \|g_1E_{s+1/8}g_1P_j^\varepsilon u\|_0 + \|[E_{s+1/8},g_1]g_1P_j^\varepsilon u\|_0 + A_{22}\|u\|_{s+1/8} \\
&\le \|g_1P_j^\varepsilon g_1E_{s+1/8}u\|_0 + \|[g_1E_{s+1/8},g_1P_j^\varepsilon]u\|_0 + A_{23}\|u\|_{s+1/8} \\
&\le \|P_j^\varepsilon g_1E_{s+1/8}u\|_0 + A_{24}\|u\|_{s+1/8},
\end{aligned}
\tag{2.1.30}
$$

$$
\begin{aligned}
\|P^{\varepsilon j}u\|_{s+1/8} &\le \|g_1E_{s+1/8}g_1P^{\varepsilon j}u\|_0 + \|[E_{s+1/8},g_1]g_1P^{\varepsilon j}u\|_0 \\
&\le \|g_1P^{\varepsilon j}g_1E_{s+1/8}u\|_0 + \|[g_1E_{s+1/8},g_1P^{\varepsilon j}]u\|_0 \\
&\quad + A_{25}\|u\|_{s+1/8} \\
&\le \|P^{\varepsilon j}g_1E_{s+1/8}u\|_0 + A_{26}\|u\|_{s+1/8}, \qquad j=1,2.
\end{aligned}
\tag{2.1.31}
$$

Let the functions g_1 and g_2 be as in the proof of Lemma 2.1.4. The relations (2.1.14) and (2.1.28) hold for any $K \subset K'$ having positive distance to the boundary of K' (with constants depending on K). So we may replace the function u in these relations by g_1E_su, $s \in R$. We need that for $t > 0$,

$$
\|P^\varepsilon g_1E_{s+t}u\|_{-t} \le \|P^\varepsilon u\|_s + A_{27}\left(\sum_{j=1}^{2}\|P_j^\varepsilon u\|_s + \|u\|_s\right).
\tag{2.1.32}
$$

This can be shown as follows: First note that

$$
\begin{aligned}
\|P^\varepsilon g_1E_{s+t}u\|_{-t} &= \|E_{-t}g_2P^\varepsilon g_1E_{s+t}u\|_0, \\
E_{-t}g_2P^\varepsilon g_1E_{s+t} &= E_{-t}E_{s+t}g_2P^\varepsilon g_1 + E_{-t}[g_2P^\varepsilon g_1, E_{s+t}],
\end{aligned}
$$

and $E_{-t}E_{s+t}g_2P^\varepsilon g_1u = E_sP^\varepsilon u$. Let p^ε denote the symbol of the operator $g_2P^\varepsilon g_1$. By Lemma 2.1.1(iii), the operator $[g_2P^\varepsilon g_1, E_{s+t}]$ has symbol

$$
\begin{aligned}
&-\sum_{j=1}^{2}\frac{\partial}{\partial\xi_j}(1+|\xi|^2)^{(s+t)/2}(D_jp^\varepsilon)(h,\theta;\xi) + r_1^\varepsilon \\
&\qquad = i(s+t)\sum_{j=1}^{2}\xi_j\big(1+|\xi|^2\big)^{(s+t-1)/2}(1+|\xi|^2)^{-1/2}(iD_jp^\varepsilon)(h,\theta;\xi) + r_1^\varepsilon,
\end{aligned}
$$

where r_1^ε is the symbol of an operator of order $s+t$. Thus, the operator $E_{-t}[g_2P^\varepsilon g_1, E_{s+t}]$ differs from the operator

$$
i(s+t)\sum_{j=1}^{2} g_1D_jE_{s-1}g_2P_j^\varepsilon
$$

only by an operator of order s having an uniform estimate for small ε and by an operator which maps the functions $u \in C_0^\infty(K)$ to 0. This proves (2.1.32).

Using formula (2.1.14) for the function $g_1 E_s u$ and formula (2.1.32) for $t = 0$, we get from formula (2.1.29) that

$$\begin{aligned} \|u\|_{s+1/4} &\le \|g_1 E_s u\|_{1/4} + A_{28}\|u\|_s \\ &\le A_{29}\big(\|P^\varepsilon g_1 E_s u\| + \|g_1 E_s u\|\big) + A_{28}\|u\|_s \\ &\le A_{29}\|P^\varepsilon u\|_s + A_{30}\left(\sum_{j=1}^{2} \|P_j^\varepsilon u\|_s + \|u\|_s\right) \\ &\le A_{31}\left(\|P^\varepsilon u\|_s + \sum_{j=1}^{2} \|P_j^\varepsilon u\|_s + \|u\|_{s+1/8}\right). \end{aligned} \tag{2.1.33}$$

Using formula (2.1.28) for $g_1 E_{s+1/8} u$ and formula (2.1.32) for $t = \frac{1}{8}$, we get from (2.1.30) and (2.1.31),

$$\begin{aligned} &\sum_{j=1}^{2} \|P_j^\varepsilon u\|_{s+1/8} + \sum_{j=1}^{2} \|P^{\varepsilon j} u\|_{s+1/8} \\ &\quad \le \sum_{j=1}^{2} \big\|P_j^\varepsilon g_1 E_{s+1/8} u\big\|_0 + \sum_{j=1}^{2} \big\|P^{\varepsilon j} g_1 E_{s+1/8} u\big\|_0 + A_{32}\|u\|_{s+1/8} \\ &\quad \le A_{33}\Big(B\big\|P^\varepsilon g_1 E_{s+1/8} u\big\|_{-1/8} + \frac{1}{B}\big\|g_1 E_{s+1/8} u\big\|_{1/8} + \big\|g_1 E_{s+1/8} u\big\|\Big) \\ &\qquad + A_{32}\|u\|_{s+1/8} \\ &\quad \le A_{33} B\left(\|P^\varepsilon u\|_s + A_{34}\left(\sum_{j=1}^{2} \|P_j^\varepsilon u\|_s + \|u\|_s\right)\right) + A_{35}\frac{1}{B}\|u\|_{s+1/4} \\ &\qquad + A_{36}\|u\|_{s+1/8}. \end{aligned} \tag{2.1.34}$$

Combining (2.1.33) and (2.1.34), we get for sufficiently large B,

$$\begin{aligned} &\|u\|_{s+1/4} + \sum_{j=1}^{2} \|P_j^\varepsilon u\|_{s+1/8} + \sum_{j=1}^{2} \|P^{\varepsilon j} u\|_{s+1/8} \\ &\quad \le A_{37}\left(\|P^\varepsilon u\|_s + \sum_{j=1}^{2} \|P_j^\varepsilon u\|_s + \sum_{j=1}^{2} \|P^{\varepsilon j} u\|_s + \|u\|_{s+1/8}\right). \end{aligned}$$

From this inequality, the statement of the lemma can be obtained by the same arguments as in the proof of the corresponding Lemma 22.2.4 of [8]. □

Next we derive from Lemma 2.1.6 an a priori estimate for solutions of (2.1.6).

LEMMA 2.1.7. *Let the set $K'' \subset K'$ be such that $a_1^\varepsilon \neq 0$ in K'' for small ε ($\varepsilon = 0$ included) and let $K \subset K''$ have positive distance to the boundary of K''. Let φ, $\tilde\varphi \in C_0^\infty(K'')$ with $\varphi = 1$ on K and $\tilde\varphi = 1$ on* $\operatorname{supp}\varphi$. *Then for any $t \in R$*

there exists a constant $A(K'', K, t, \varphi)$ *depending on* K'', K, φ *and* t *such that for any solution* $u^\varepsilon \in C^\infty$ *of*

$$P^\varepsilon u^\varepsilon = 0, \tag{2.1.35}$$

we have

$$\|\varphi u^\varepsilon\|_t \le A(K'', K, t, \varphi)\|\tilde{\varphi} u^\varepsilon\|. \tag{2.1.36}$$

PROOF. It is sufficient to show the statement for $t \in N$. Let $n = 8(t+1)+1$. We introduce a sequence of compact sets

$$K = K_0 \subset K_1 \subset \cdots \subset K_n = \operatorname{supp}\varphi,$$

such that K_i has positive distance to the boundary of K_{i+1}, $i = 1, \dots, n-1$, and a sequence of functions $\varphi_i \in C_0^\infty(K_i)$, with $\varphi_1 = \varphi$ and $\varphi_i = 1$ on K_{i-1}, $i = 1, \dots, n$. Then Lemma 2.1.6 implies

$$\|\varphi_i u^\varepsilon\|_{s+1/8} + \left\|P^{\varepsilon 1}(\varphi_i u^\varepsilon)\right\|_{s+1/8} \le A_{s,\,K'',\,K}\big(\|P^\varepsilon(\varphi_i u^\varepsilon)\|_s + \|\varphi_i u^\varepsilon\|_s\big). \tag{2.1.37}$$

It follows from (2.1.35) that

$$\begin{aligned} P^\varepsilon(\varphi_i u^\varepsilon) &= (P^\varepsilon \varphi_i)u^\varepsilon + 2(L_1^\varepsilon \varphi_i)(L_1^\varepsilon u^\varepsilon) \\ &= (P^\varepsilon \varphi_i)\varphi_{i+1}u^\varepsilon + 2(L_1^\varepsilon \varphi_i)(L_1^\varepsilon \varphi_{i+1}u^\varepsilon). \end{aligned} \tag{2.1.38}$$

Note that $P^{\varepsilon 1} = 2a_1^\varepsilon L_1^\varepsilon$. By (2.1.5), the assumption $a_1^\varepsilon \ne 0$ on K'' implies that $|a_1^\varepsilon| > \delta > 0$ on K'' for some δ independent of ε. We get from (2.1.37) and (2.1.38) that

$$\begin{aligned} \|\varphi u^\varepsilon\|_t \le A_{t-1/8,\,K'',\,K}\bigg(&\left\|(P^\varepsilon\varphi_1)\varphi_2 u^\varepsilon\right\|_{t-1/8} + \left\|2\frac{L_1^\varepsilon\varphi_1}{a_1^\varepsilon}a_1^\varepsilon\ L_1^\varepsilon(\varphi_2 u^\varepsilon)\right\|_{t-1/8} \\ &+ \left\|\varphi_1\varphi_2 u^\varepsilon\right\|_{t-1/8}\bigg). \end{aligned} \tag{2.1.39}$$

The factors $P^\varepsilon\varphi_1$, $L_1^\varepsilon\varphi_1/a_1^\varepsilon$ and φ_1 can be regarded as operators of order 0 satisfying the conditions of Lemma 2.1.1. Thus there exists a constant $A_1(K'', K, t, \varphi_1)$ such that

$$\begin{aligned} \|\varphi_1 u^\varepsilon\|_t &\le A_1(K'', K, t, \varphi_1)\big(\|\varphi_2 u^\varepsilon\|_{t-1/8} + \|2\ a_1^\varepsilon L_1^\varepsilon(\varphi_2 u^\varepsilon)\|_{t-1/8}\big) \\ &= A_1(K'', K, t, \varphi_1)\big(\|\varphi_2 u^\varepsilon\|_{t-1/8} + \left\|P^{\varepsilon 1}(\varphi_2 u^\varepsilon)\right\|_{t-1/8}\big). \end{aligned} \tag{2.1.40}$$

Applying (2.1.37) again, we get, using (2.1.38) and the same arguments as above,

$$
\begin{aligned}
\|\varphi u^\varepsilon\|_t &\le A_1(K'', K, t, \varphi_1) A_{t-2/8,\,K'',\,K}\big(\|P^\varepsilon(\varphi_2 u^\varepsilon)\|_{t-2/8} + \|\varphi_2 u^\varepsilon\|_{t-2/8}\big)\\
&= \tilde{A}_2(K'', K, t, \varphi_1)\\
&\quad \times \big(\|P^\varepsilon(\varphi_2 u^\varepsilon)\|_{t-2/8} + \|\varphi_2 u^\varepsilon\|_{t-2/8}\big)\\
&\le \tilde{A}_2(K'', K, t, \varphi_1)\bigg(\big\|(P^\varepsilon \varphi_2)\varphi_3 u^\varepsilon\big\|_{t-2/8} + \bigg\|2\frac{L_1^\varepsilon \varphi_2}{a_1^\varepsilon} a_1^\varepsilon L_1^\varepsilon(\varphi_3 u^\varepsilon)\bigg\|_{t-2/8} \qquad (2.1.41)\\
&\qquad\qquad + \|\varphi_2\varphi_3 u^\varepsilon\|_{t-2/8}\bigg)\\
&\le A_2(K'', K, t, \varphi_1, \varphi_2)\Big(\big\|\varphi_3 u^\varepsilon\big\|_{t-2/8} + \big\|P^{\varepsilon 1}(\varphi_3 u^\varepsilon)\big\|_{t-2/8}\Big).
\end{aligned}
$$

Continuing in this way, we finally obtain

$$
\begin{aligned}
\|\varphi u^\varepsilon\|_t &\le A_{n-1}(K'', K, t, \varphi_1, \varphi_2, \dots, \varphi_{n-1})\Big(\|\varphi_n u^\varepsilon\|_{-1} + \|\tilde{\varphi} P^{\varepsilon 1}(\varphi_n u^\varepsilon)\|_{-1}\Big)\\
&\le A(K'', K, t, \varphi)\|\tilde{\varphi} u^\varepsilon\|_0
\end{aligned}
$$

by Lemma 2.1.1, as the operator $\tilde{\varphi} P^{\varepsilon 1}$ has order 1. $\square$

LEMMA 2.1.8. *Let $\tilde{u}^\varepsilon$ be a solution of the differential equation (2.1.2) for $0 < \varepsilon < \varepsilon_0$ and let $S \subset [0, 2\pi]$ be a compact interval such that $a_1(0, \theta) \ne 0$ for $\theta \in S$. Let $\tilde{K}$ be an arbitrary neighborhood of the set $\{(0, \theta), \theta \in S\}$ such that $H_0 + \tilde{h} \in (H_1, H_2)$ for all $(\tilde{h}, \theta) \in \tilde{K}$ and assume that there exists a constant $A_1(\tilde{K})$ such that $|\tilde{u}^\varepsilon(h, \theta)| \le A_1(\tilde{K})$ for all $(h, \theta) \in \tilde{K}$ and $\varepsilon < \varepsilon_0$. Then there exist an $\varepsilon_1 > 0$ and a constant $A_2(\tilde{K})$ independent of ε such that*

$$
\left|\frac{\partial}{\partial\theta}\tilde{u}^\varepsilon(0, \theta)\right| \le A_2(\tilde{K})
$$

for all $0 < \varepsilon < \varepsilon_1$ and $\theta \in S$.

PROOF. Let ε_1 be such that for $0 < \varepsilon < \varepsilon_1$ we have $|\tilde{h}| < \varepsilon^{-1/2}$ if $(\tilde{h}, \theta) \in \tilde{K}$, and the functions u^ε defined by $u^\varepsilon(h, \theta) = \tilde{u}^\varepsilon(\varepsilon^{1/2} h, \theta)$ satisfy the conditions of Lemma 2.1.7. There exist compact sets $K,\ K''$, $\{0\} \times S \subset K \subset K''$, such that:

(i) each of these sets has positive distance to the boundary of the larger one.

(ii) $a_1^\varepsilon \ne 0$ on K''.

(iii) $(\sqrt{\varepsilon} h, \theta) \in \tilde{K}$ if $(h, \theta) \in K''$.

By Lemma 2.1.7 we get, for $t = 3$,

$$
\begin{aligned}
\|\varphi u^\varepsilon\|_3 &\le A(K'', K, 3, \varphi)\|\tilde{\varphi} u^\varepsilon\|\\
&\qquad \le A_3(\tilde{K}) \sup_{K''} |u^\varepsilon| \le A_3(\tilde{K}) \sup_{\tilde{K}} |\tilde{u}^\varepsilon| \le A_3(\tilde{K}) A_1(\tilde{K})
\end{aligned}
$$

with suitable functions φ and $\tilde{\varphi}$ and a constant $A_3(\tilde{K})$ depending only on $\tilde{K}$ (and on the sets K and K'' and the function φ chosen for $\tilde{K}$).

By Sobolev's lemma (see, e.g., [1]), this implies that there exists a constant $A_2(\tilde{K})$ such that

$$\left|\frac{\partial}{\partial\theta}u^\varepsilon(h,\theta)\right| \le A_2(\tilde{K}), \qquad (h,\theta)\in K.$$

As $(\partial/\partial\theta)u^\varepsilon(0,\theta) = (\partial/\partial\theta)\tilde{u}^\varepsilon(0,\theta)$, we get the statement of the lemma. □

2.2. *The Markov property.* Let G be a domain in R^2 bounded by components of the level sets of $H(x,y)$ and let $\tau = \tau^\varepsilon$ be the exit time of the process $(X_t^\varepsilon, Y_t^\varepsilon)$ from this domain. Using the a priori estimate obtained in the last subsection, we show in this subsection that the probability that $(X_\tau^\varepsilon, Y_\tau^\varepsilon)$ belongs to a certain connected component of the boundary of G depends for $\varepsilon \downarrow 0$ only on the value $H(x_0, y_0)$ at the initial point (x_0, y_0). We will use this result to prove the form of the gluing conditions for the limiting process on the graph in the next subsection. Therefore, we need this result especially for domains G containing a saddle point of $H(x,y)$. Actually, the result of this subsection implies the Markov property of the limiting process.

Let $\underline{x} = (x,y)$ and let $\underline{X}_t^\varepsilon = (X_t^\varepsilon, Y_t^\varepsilon)$ and $\tilde{\underline{X}}_t^\varepsilon = (\tilde{X}_t^\varepsilon, \tilde{Y}_t^\varepsilon)$ be the solutions of the systems (1.12) and (1.11), respectively. Further, let $\underline{X}_t(\underline{x}) = (X_t(\underline{x}), Y_t(\underline{x}))$ be the solution of the deterministic system (1.10) with the initial point $\underline{x} = (x,y)$. Itô's formula applied to $f(H(\underline{X}_t^\varepsilon))$, with a smooth function f, gives

$$\begin{aligned} f\big(H(\underline{X}_t^\varepsilon)\big) &= f(H(\underline{X}_0^\varepsilon)) + \int_0^t f'(H(\underline{X}_s^\varepsilon))H_y(\underline{X}_s^\varepsilon)\,dW_s \\ &\quad + \int_0^t \tfrac{1}{2}\big(f''(H(\underline{X}_s^\varepsilon))H_y^2(\underline{X}_s^\varepsilon) + f'(H(\underline{X}_s^\varepsilon))H_{yy}(\underline{X}_s^\varepsilon)\big)\,ds, \end{aligned} \tag{2.2.1}$$

$$\begin{aligned} E_{\underline{x}}^\varepsilon f(H(\underline{X}_\tau^\varepsilon)) &= f(H(\underline{x})) + E_{\underline{x}}^\varepsilon \int_0^\tau \tfrac{1}{2}\big(f''(H(\underline{X}_s^\varepsilon))H_y^2(\underline{X}_s^\varepsilon) \\ &\quad + f'(H(\underline{X}_s^\varepsilon))H_{yy}(\underline{X}_s^\varepsilon)\big)\,ds, \end{aligned} \tag{2.2.2}$$

for the time τ to exit any bounded region.

LEMMA 2.2.1. *For $k \in N$ there exists an $A(k) \ge 0$ such that for $T > 0$,*

$$P_{\underline{x}}^\varepsilon\Big\{\sup_{0<t<T}\big|\tilde{\underline{X}}_t^\varepsilon - \underline{X}_t(\underline{x})\big| \ge \eta\Big\} \le A(k)(e^{2LT}-1)^k\frac{\varepsilon^k}{\eta^{2k}},$$

where L is the Lipschitz constant of the function $\bar{\nabla}H$ and η is an arbitrary positive number.

This lemma is the analogue of Lemma 4.2 in [7] and can be proved by the same arguments.

For $\varepsilon > 0$ we consider the (fast) dynamical system

$$\dot{X}_t^\varepsilon(x) = \frac{1}{\varepsilon}\bar{\nabla} H\big(X_t^\varepsilon(x)\big), \qquad X_0^\varepsilon(x) = x. \tag{2.2.3}$$

COROLLARY 2.2.2. *With the notations of Lemma 2.2.1, we have*

$$P_x^\varepsilon\Big\{\sup_{0<t<\varepsilon T}\big|X_t^\varepsilon - X_t^\varepsilon(x)\big| \geq \eta\Big\} \leq A(k)\big(e^{2LT}-1\big)^k \frac{\varepsilon^k}{\eta^{2k}}.$$

Let μ denote the Lebesgue measure.

LEMMA 2.2.3. *Let $D = D_l(H_1, H_2)$, $H_1 \geq -\infty$, $H_2 \leq \infty$, $G \subset D$ and $T = \sup_{x\in D}\min\{t > 0\colon X_t(x) = x\}$. Assume that for $\eta > 0$ the set*

$$G_\eta = \big\{z \in D \text{ such that there exist } x \in \bar{G} \text{ and } s \in (-\eta,\eta) \text{ with } z = X_s(x)\big\}$$

has the property

$$\mu\big\{t \in (0,T)\colon X_t(x) \in G_\eta\big\} \leq T/B$$

for some $B \geq 4$ and for all $x \in D$. Then for any $T_0 > 0$,

$$P_x^\varepsilon\big\{\mu\{t \in (0,T_0)\colon X_t^\varepsilon \in G\} \geq 2T_0/B\big\} \to 0 \quad \text{if } \varepsilon \downarrow 0,$$

uniformly for all $x \in D$.

PROOF. By Corollary 2.2.2 there exists an $A_{38} > 0$ such that

$$P_x^\varepsilon\Big\{\sup_{0\leq t\leq\varepsilon T}\big|X_t^\varepsilon - X_t^\varepsilon(x)\big| \geq \tilde{\eta}\Big\} \leq A_{38}\frac{\varepsilon^2}{\tilde{\eta}^4} \tag{2.2.4}$$

for all $x \in D$, $\tilde{\eta} > 0$. Define

$$N_\varepsilon = \left[\frac{T_0}{\varepsilon T}\right], \qquad T_G^\varepsilon = \mu\big\{t \in (0,T_0)\colon X_t^\varepsilon \in G\big\} = \int_0^{T_0} \chi_G(X_t^\varepsilon)\,dt,$$

where $[t]$, $t \in R$, denotes the largest integer less than t and χ_G is the indicator function of the set G. Observe that

$$T_G^\varepsilon \leq \sum_{n=0}^{N_\varepsilon}\int_{n\varepsilon T}^{(n+1)\varepsilon T} \chi_G(X_t^\varepsilon)\,dt.$$

By the Markov property and (2.2.4), we get, for small ε and a suitable constant $\tilde{A} > 0$ depending on G and on the speed $\bar{\nabla} H(x)$ of $X_t(x)$,

$$\begin{aligned} P_x^\varepsilon\left\{T_G^\varepsilon \leq \frac{2T_0}{B}\right\} &\geq P_x^\varepsilon\left\{\sum_{n=0}^{N_\varepsilon}\int_{n\varepsilon T}^{(n+1)\varepsilon T} \chi_G(X_t^\varepsilon)\,dt \leq \frac{2T_0}{B}\right\} \\ &\geq P_x^\varepsilon\left\{\int_{n\varepsilon T}^{(n+1)\varepsilon T} \chi_G(X_t^\varepsilon)\,dt \leq \frac{2T_0}{B(N_\varepsilon+1)},\ n = 0,\ldots,N_\varepsilon\right\} \end{aligned}$$

$$\geq \left(\inf_{\underline{z}\in G} P^{\varepsilon}_{\underline{z}}\left\{\int_0^{\varepsilon T} \chi_G(\underline{X}^{\varepsilon}_t)\,dt \leq \frac{\varepsilon T}{B}\right\}\right)^{N_\varepsilon+1}$$

$$\geq \left(1 - \sup_{\underline{z}\in G} P^{\varepsilon}_{\underline{z}}\left\{\sup_{0\leq t\leq \varepsilon T}\left|\underline{X}^{\varepsilon}_t - \underline{X}^{\varepsilon}_t(\underline{z})\right| \geq \tilde{A}\,\eta\right\}\right)^{N_\varepsilon+1}$$

$$\geq \left(1 - \frac{A_{38}}{\tilde{A}^4}\,\frac{\varepsilon^2}{\eta^4}\right)^{T_0/\varepsilon T+1}.$$

Obviously, the right-hand side of this estimate tends to 1 if ε tends to 0. □

LEMMA 2.2.4. *Let $D = D_l(H_1, H_2)$, $|H_j| < \infty$, $j = 1, 2$, and $\tau^\varepsilon = \inf\{t\colon \underline{X}^\varepsilon_t \notin D\}$. Then for $\underline{x} \in D$ (and small ε) there exists an A_{39} such that $E^\varepsilon_{\underline{x}}\tau^\varepsilon < A_{39}$.*

PROOF. Let $\tilde{H}_1 < H_1$ and $\tilde{H}_2 > H_2$ such that $D \subset \tilde{D} = D_l(\tilde{H}_1, \tilde{H}_2)$. Choose functions c_1, $c_2 \in C^\infty(R^2)$ such that $c_1(\underline{x}) = H_y(\underline{x})$ and $c_2(\underline{x}) = H_{yy}(\underline{x})$ if $\underline{x} \in D$, and such that for an open set $G \subset \tilde{D}$ there exist positive numbers δ_1 and δ_2 such that

$$c_1^2(\underline{x}) \geq \delta_1, \qquad \underline{x} \in R^2 \setminus G, \tag{2.2.5}$$

$$c_1^2(\underline{x}) \leq \delta_1, \qquad \underline{x} \in G, \tag{2.2.6}$$

$$|c_2(\underline{x})| < \delta_2, \qquad \underline{x} \in D, \tag{2.2.7}$$

and such that the conditions of Lemma 2.2.3 are satisfied for $\tilde{D}$ and G with suitable numbers η and B (see also Lemma 1.1). For $\underline{x} \in D$, consider the process Z^ε_t defined by

$$Z^\varepsilon_t = H(\underline{x}) + \int_0^t c_1(\underline{X}^\varepsilon_s)\,dW_s + \tfrac{1}{2}\int_0^t c_2(\underline{X}^\varepsilon_s)\,ds. \tag{2.2.8}$$

One can find a Wiener process $\tilde{W}_t$ such that (see, e.g., [5], page 51)

$$\int_0^t c_1(\underline{X}^\varepsilon_s)\,dW_s = \tilde{W}_{\int_0^t c_1^2(\underline{X}^\varepsilon_s)\,ds}. \tag{2.2.9}$$

Note that the processes Z^ε_t and $H(\underline{X}^\varepsilon_t)$ coincide until τ^ε by formula (2.2.1), and

$$P^\varepsilon_{\underline{x}}\{\tau^\varepsilon < 1\} \geq P^\varepsilon_{\underline{x}}\left\{\sup_{0\leq t\leq 1} Z^\varepsilon_t > H_2\right\}, \qquad \underline{x} \in D.$$

Let $A_{40} = H_2 - H_1 + \delta_2$. By (2.2.8), (2.2.9) and Lemma 2.2.3, we get for $\underline{x} \in D$ and small ε,

$$P^\varepsilon_{\underline{x}}\{\tau^\varepsilon < 1\} \geq P^\varepsilon_{\underline{x}}\left\{\sup_{0\leq t\leq 1}\left\{H(\underline{x}) + \int_0^t c_1(\underline{X}^\varepsilon_s)\,dW_s + \frac{1}{2}\int_0^t c_2(\underline{X}^\varepsilon_s)\,ds\right\} \geq H_2\right\}$$

$$\geq P^\varepsilon_{\underline{x}}\left\{\sup_{0\leq t\leq 1}\int_0^t c_1(\underline{X}^\varepsilon_s)\,dW_s \geq A_{40}\right\}$$

$$\begin{aligned}&\ge P_{\underset{\sim}{x}}^{\varepsilon}\left\{\sup_{0\le t\le 1}\tilde{W}_{\int_0^t c_1^2(\underset{\sim}{X}_s^{\varepsilon})\,ds}\ge A_{40}\,\middle|\,T_G^{\varepsilon}<\frac{2}{B}\right\}P_{\underset{\sim}{x}}^{\varepsilon}\left\{T_G^{\varepsilon}<\frac{2}{B}\right\}\\&\ge\frac{1}{2}P_{\underset{\sim}{x}}^{\varepsilon}\left\{\sup_{0\le t\le\delta_1(1-2/B)}\tilde{W}_t\ge A_{40}\,\middle|\,T_G^{\varepsilon}<\frac{2}{B}\right\}\\&\ge\frac{1}{2}P_{\underset{\sim}{x}}^{\varepsilon}\{\tilde{W}_{\delta_1(1-2/B)}\ge A_{40}\}-\frac{1}{2}P_{\underset{\sim}{x}}^{\varepsilon}\left\{T_G^{\varepsilon}\ge\frac{2}{B}\right\}\ge\alpha\end{aligned}$$

for some $\alpha > 0$ independent of ε for small ε. By the strong Markov property we get

$$P_{\underset{\sim}{x}}^{\varepsilon}\{\tau^{\varepsilon}<n+1\,|\,\tau^{\varepsilon}\ge n\}\ge\alpha,\qquad n\in N,\underset{\sim}{x}\in D.$$

Thus $P_{\underset{\sim}{x}}^{\varepsilon}\{\tau^{\varepsilon}>n\}\le(1/\alpha)P_{\underset{\sim}{x}}^{\varepsilon}\{\tau^{\varepsilon}\in(n,n+1]\}$ implies $E_{\underset{\sim}{x}}^{\varepsilon}\tau^{\varepsilon}<1/\alpha$. □

As we are concerned only with the behavior of the processes until they leave a bounded domain in D_l, we can change the function H outside a large enough subset of D_l so that H becomes bounded, especially so that H is constant there. The distribution of $\underset{\sim}{X}_{\tau^{\varepsilon}}^{\varepsilon}$ is the same as the distribution of $\tilde{\underset{\sim}{X}}_{\tilde{\tau}^{\varepsilon}}^{\varepsilon}$, where $\tilde{\tau}^{\varepsilon}=\inf\{t\colon\tilde{\underset{\sim}{X}}_t^{\varepsilon}\notin D\}$. So we consider the (slow) process $\tilde{\underset{\sim}{X}}_t^{\varepsilon}$.

As in the beginning of Section 2.1, we introduce coordinates (h,θ) with the only difference that now $\tilde{h}(\underset{\sim}{x})=H(\underset{\sim}{x})-H_2$ or $\tilde{h}(\underset{\sim}{x})=H(\underset{\sim}{x})-H_1$ so that the points with coordinates $(0,\theta)$ correspond to the boundary $C_l(H_2)$ or $C_l(H_1)$, respectively. For brevity we restrict the considerations to the case of the boundary $C_l(H_2)$, this means $\tilde{h}(\underset{\sim}{x})=H(\underset{\sim}{x})-H_2$. It is easy to check that Lemmas 2.2.8 and 2.2.9 below hold also if H_2 is replaced by H_1.

With the coordinates (h,θ) the generator of the process $\tilde{\underset{\sim}{X}}_t^{\varepsilon}$ can be written

$$\tilde{L}^{\varepsilon}u=a_1^{\varepsilon 2}u_{hh}+a_3^{\varepsilon}u_{\theta}+\varepsilon a_2^{\varepsilon 2}u_{\theta\theta}+\varepsilon b_4^{\varepsilon}u_{\theta}+\sqrt{\varepsilon}2a_1^{\varepsilon}a_2^{\varepsilon}u_{\theta h}+\sqrt{\varepsilon}b_5^{\varepsilon}u_h,\tag{2.2.10}$$

where

$$\begin{aligned}b_4^{\varepsilon}\left(\frac{1}{\sqrt{\varepsilon}}h(x,y),\theta(x,y)\right)&=\frac{1}{2}\theta_{yy}(x,y),\\ b_5^{\varepsilon}\left(\frac{1}{\sqrt{\varepsilon}}h(x,y),\theta(x,y)\right)&=\frac{1}{2}H_{yy}(x,y)\end{aligned}$$

are functions bounded on compact sets.

Thus the process $\underset{\sim}{H}_t^{\varepsilon}=(h(\tilde{\underset{\sim}{X}}_t^{\varepsilon}),\theta(\tilde{\underset{\sim}{X}}_t^{\varepsilon}))$ solves the equation

$$d\underset{\sim}{H}_t^{\varepsilon}=\underset{\sim}{b}^{\varepsilon}(\underset{\sim}{H}_t^{\varepsilon})\,dt+\underset{\sim}{\sigma}^{\varepsilon}(\underset{\sim}{H}_t^{\varepsilon})\,d\underset{\sim}{W}_t\tag{2.2.11}$$

with

$$\begin{aligned}\underset{\sim}{b}^{\varepsilon}(h,\theta)&=\underset{\sim}{b}^0(\theta)+\sqrt{\varepsilon}\underset{\sim}{b}^{0,\varepsilon}(h,\theta)\\&=\begin{pmatrix}0\\b_2(\theta)\end{pmatrix}+\sqrt{\varepsilon}\begin{pmatrix}b_5^{\varepsilon}(h,\theta)\\b_3^{\varepsilon}(h,\theta)+\sqrt{\varepsilon}b_4^{\varepsilon}(h,\theta)\end{pmatrix},\end{aligned}$$

$$\sigma^\varepsilon(h,\theta) = \sigma^0(\theta) + \sqrt{\varepsilon}\sigma^{0,\varepsilon}(h,\theta)$$
$$= \begin{pmatrix} b_0(\theta) & 0 \\ 0 & 0 \end{pmatrix} + \sqrt{\varepsilon}\,\sqrt{2}\begin{pmatrix} b_1^\varepsilon(h,\theta) & 0 \\ a_2^\varepsilon(h,\theta) & 0 \end{pmatrix},$$

where by Taylor's expansion

$$a_1^\varepsilon(h,\theta) = \frac{1}{\sqrt{2}}\,b_0(\theta) + \sqrt{\varepsilon}\,b_1^\varepsilon(h,\theta), \qquad a_3^\varepsilon(h,\theta) = b_2(\theta) + \sqrt{\varepsilon}\,b_3^\varepsilon(h,\theta),$$

$b_0(\theta) = \sqrt{2}a_1^\varepsilon(0,\theta)$, $b_2(\theta) = a_3^\varepsilon(0,\theta)$ and $b_1^\varepsilon, b_2^\varepsilon$ are uniformly bounded on compact sets for small ε. Define the process $\underline{H}_t^0 = (H_t^0, \Theta_t^0)$ by

$$d\underline{H}_t^0 = \underline{b}^0(\underline{H}_t^0)\,dt + \sigma^0(\underline{H}_t^0)\,d\underline{W}_t \tag{2.2.12}$$

with the same Wiener process $\underline{W}_t$ as in (2.2.11).

LEMMA 2.2.5. *There exists a bounded function $A(t)$ such that for any $T > 0$, $\eta > 0$, $\varepsilon > 0$, $\underline{x} \in D$,*

$$P_{\underline{x}}^\varepsilon\Big\{\sup_{0\le t\le T}|\underline{H}_t^\varepsilon - \underline{H}_t^0| > \eta\Big\} \le A(T)\frac{\varepsilon}{\eta^2}. \tag{2.2.13}$$

The proof is standard (compare, e.g., Lemma 2.1.2 in [5]).

Define for $h_1, h_2 \in [-\infty,\infty]$, $h_1 < h_2$,

$$\tau^0(h_1,h_2) = \inf\{t\colon\ H_t^0 \notin (h_1,h_2)\}.$$

LEMMA 2.2.6. *We have for $h > 0$ and $B \subset [0,2\pi]$,*

$$P_{h,\theta}\{\Theta^0_{\tau^0(0,\infty)} \in B\} = \int_B f_{h,\theta}(\eta)\,d\eta$$

with a bounded function $f_{h,\theta}$.

PROOF. Let $h_0 > 0$. The process $\underline{H}_t^0$ corresponds to the operator

$$\frac{1}{2}\,b_0^2\,\frac{\partial^2}{\partial h^2} + b_2\frac{\partial}{\partial\theta}$$

For any $\delta > 0$ there exists $h_1 > h_0$ such that

$$P_{h_0,\theta}\{H^0_{\tau^0(0,h_2)} = h_2\} < \delta$$

for all $h_2 \ge h_1$. Thus

$$\big|P_{h_0,\theta}\{\Theta^0_{\tau^0(0,\infty)} \in B\} - P_{h_0,\theta}\{\underline{H}^0_{\tau^0(0,h_1)} \in \{0\}\times B\}\big| < \delta. \tag{2.2.14}$$

Let θ_t^θ be the bounded solution (modulo 2π) of the initial value problem

$$\tfrac{1}{2}b_0^2(\theta_t)\dot\theta_t = b_2(\theta_t), \qquad \theta_0 = \theta$$

(where $\dot\theta_t^\theta$ is unbounded at the zeros of b_0). The solution of this problem satisfies the equation

$$t = A_{41} + \int_0^{\theta_t^\theta} \frac{b_0^2(\xi)}{2b_2(\xi)}\, d\xi, \tag{2.2.15}$$

with a suitable constant A_{41}.

Let $\mathcal{H}_t^1 = (H_t^1, \theta_t^\theta)$ with a Wiener process H_t^1. Note that

$$P_{h_0,\,\theta}\{\Theta^0_{\tau^0(0,\,\infty)} \in B\} = P_{h_0,\,\theta}\{\theta^\theta_{\tau^1(0,\,\infty)} \in B\} = E_{h_0,\,\theta}\,\chi_B\big(\theta^\theta_{\tau^1(0,\,\infty)}\big)$$

as each of them solves the boundary value problem

$$\tfrac{1}{2} b_0^2(\theta) u_{hh}(h,\theta) + b_2(\theta) u_\theta(h,\theta) = 0, \qquad u(0,\theta) = \chi_B(\theta).$$

We consider now this boundary value problem with boundary conditions

$$u(0,\theta) = \varphi(\theta), \qquad u(h_1,\theta) = 0, \qquad u(h,\theta) = u(h,\theta+2\pi),$$

with a periodic function φ, $\varphi(\theta) = \varphi(\theta+2\pi)$. Evidently the solution u^{h_1} of this problem,

$$u^{h_1}(h,\theta) = E_{h,\,\theta}\tilde\varphi\big(\mathcal{H}^0_{\tau^0(0,\,h_1)}\big),$$

[with a function $\tilde\varphi$ such that $\tilde\varphi(0,\theta) = \varphi(\theta)$, $\tilde\varphi(h_1,\theta) = 0$], can be represented

$$u^{h_1}(h,\theta) = \int_0^\infty \varphi(\theta_t^\theta) f(h_1, h, t)\, dt,$$

where

$$f(h_1, h, dt) = P\big\{\tau(-h, h_1 - h) \in dt,\ W_{\tau(-h, h_1-h)} = -h\big\}$$

and $\tau(a,b) = \inf\{t > 0\colon W_t \notin (a,b)\}$. The explicit formulas are known for this probability (see, e.g., [2]). Let $T_0 = \inf\{t > 0\colon \theta_t^\theta = \theta\}$ be the period of θ_t^θ. As $f(h_1, h, t)$ tends to $f(h,t)$,

$$f(h,t) = P\big\{\tau(-\infty, h) \in dt\big\},$$

if $h_1 \to \infty$, we get for $u(h,\theta) = E_{h,\,\theta}\varphi(\theta^\theta_{\tau^0(0,\infty)})$ by (2.2.14) (using the substitution $\eta = \theta_t^\theta$),

$$\begin{aligned} u(h,\theta) &= \int_0^\infty \varphi(\theta_t^\theta) f(h,t)\, dt = \sum_{k=0}^\infty \int_{t=0}^{T_0} \varphi(\theta_t^\theta) f(h, t + kT_0)\, dt \\ &= \sum_{k=0}^\infty \int_{\eta=\theta}^{2\pi+\theta} \varphi(\eta) f\big(h, g^\theta(\eta) + kT_0\big) \frac{b_0^2(\eta)}{2b_2(\eta)}\, d\eta \\ &= \int_\theta^{\theta+2\pi} \varphi(\eta) \sum_{k=0}^\infty f\big(h, g^\theta(\eta) + kT_0\big) \frac{b_0^2(\eta)}{2b_2(\eta)}\, d\eta, \end{aligned} \tag{2.2.16}$$

where

$$g^\theta(\eta) = \int_\theta^\eta \frac{b_0^2(\xi)}{2b_2(\xi)}\, d\xi.$$

Thus the function in the integral at the right-hand side of (2.2.16) is the density $f_{h,\theta}$:

$$f_{h,\theta}(\eta) = \sum_{k=0}^{\infty} f(h, g^{\theta}(\eta) + kT_0)\frac{b_0^2(\eta)}{2b_2(\eta)}. \tag{2.2.17}$$ □

LEMMA 2.2.7. *For any $\delta > 0$ there exists $h_0 > 0$ such that for any interval $S = (\gamma_1, \gamma_2) \subset [0, 2\pi]$,*

$$\sup_{\theta_1,\,\theta_2\in[0,\,2\pi]}\Big| P_{h_0,\,\theta_1}\big\{H^0_{\tau^0(0,\,\infty)} \in \{0\}\times S\big\} - P_{h_0,\,\theta_2}\big\{H^0_{\tau^0(0,\,\infty)} \in \{0\}\times S\big\}\Big| < \delta. \tag{2.2.18}$$

PROOF. By (2.2.16) we have

$$\begin{aligned} &P_{h,\,\theta_1}\big\{H^0_{\tau^0(0,\,\infty)} \in \{0\}\times S\big\} - P_{h,\,\theta_2}\big\{H^0_{\tau^0(0,\,\infty)} \in \{0\}\times S\big\} \\ &\quad= \int_0^{\infty}\big(\chi_S(\theta_t^{\theta_1}) - \chi_S(\theta_t^{\theta_2})\big)f(h,t)\,dt. \end{aligned}$$

This can be written

$$\int_0^{\infty}(g(t) - g(t-\gamma))\,f(h,t)\,dt,$$

with a function g, $0 \le g \le 1$, and a suitable constant γ, $0 \le \gamma \le T_0$, where $T_0 = \inf\{t > 0\colon \theta_t^{\theta} = \theta\}$ is the period of θ_t^{θ}. The function f,

$$f(h,t) = \frac{1}{\sqrt{2\pi}}\,\frac{h}{t^{3/2}}\exp\left(-\frac{h^2}{2t}\right)$$

(see, e.g., [5]), for fixed h is a unimodular function with maximum at $t = h^2/3$ and

$$f\left(h, \frac{h^2}{3}\right) = \frac{3\sqrt{3}}{\sqrt{2\pi}}\,\frac{1}{h^2}e^{-3/2}.$$

Thus the function in the above integral tends to zero uniformly on intervals of finite length as $h \to \infty$. So it is sufficient to show that the expressions

$$\int_{\gamma}^{h^2/3}\big(f(h,t) - f(h,t-\gamma)\big)\,dt, \tag{2.2.19}$$

$$\int_{h^2/3+\gamma}^{\infty}\big(f(h,t) - f(h,t-\gamma)\big)\,dt \tag{2.2.20}$$

tend to 0 if $h \to \infty$. Note that

$$\int_0^z f(h,t)\,dt = 1 - \operatorname{erf}\left(\frac{h}{\sqrt{2z}}\right) \quad \text{where } \operatorname{erf}(z) = \frac{2}{\sqrt{\pi}}\int_0^z \exp(-s^2)\,ds.$$

So we get for the expression (2.2.19),

$$\operatorname{erf}\left(\frac{3}{\sqrt{2}}\frac{h}{\sqrt{3h^2-9\gamma}}\right)-\operatorname{erf}\left(\sqrt{\frac{3}{2}}\right)+1-\operatorname{erf}\left(\frac{h}{\sqrt{2\gamma}}\right),$$

which tends to 0 if $h\to\infty$. The same holds for the expression (2.2.20), as it is equal to

$$\operatorname{erf}\left(\frac{3}{\sqrt{2}}\frac{h}{\sqrt{3h^2+9\gamma}}\right)-\operatorname{erf}\left(\sqrt{\frac{3}{2}}\right). \qquad \square$$

LEMMA 2.2.8. *Let $D=D_l(H_1,H_2)$ and $H_1<H_0<H_2$. For any interval $S=(\gamma_1,\gamma_2)\subset[0,2\pi]$ and any $\delta>0$, there exists an $\varepsilon_0>0$ such that*

$$\sup_{x_1,x_2\in C_l(H_0)}\Big|P_{x_1}\{H(X^\varepsilon_{\tau^\varepsilon(D)})=H_2,\theta(X^\varepsilon_{\tau^\varepsilon(D)})\in S\} - P_{x_2}\{H(X^\varepsilon_{\tau^\varepsilon(D)})=H_2,\theta(X^\varepsilon_{\tau^\varepsilon(D)})\in S\}\Big|<\delta \tag{2.2.21}$$

for all $\varepsilon<\varepsilon_0$.

PROOF. We may assume without loss of generality that $\delta<1$. Let $\delta_0=\delta/25$. As will be shown in the next subsection, the processes $Z_t^\varepsilon=H(X_t^\varepsilon)$ and $X_0^\varepsilon=x\in D_l(H_1,H_2)$, stopped at the moment when they first leave (H_1,H_2), converge weakly on any time interval $[0,T_0]$ as $\varepsilon\to0$ to a nondegenerate diffusion process, the same for all $x\in D_l(H_1,H_2)$. Using this we can find $\varepsilon_1>0$ and $d_0>0$ such that

$$P_{x_3}\{H(X^\varepsilon_{\tau^\varepsilon(D)})=H_2\}<\delta_0 \quad \text{for all } x_3\in D_l(H_1,H_1+d_0),\ \varepsilon<\varepsilon_1, \tag{2.2.22}$$

and

$$\Big|P_{x_1}\{H(X^\varepsilon_{\tau^\varepsilon(D_l(H_1+d,H_2-d))})=H_j\pm d\} - P_{x_2}\{H(X^\varepsilon_{\tau^\varepsilon(D_l(H_1+d,H_2-d))})=H_j\pm d\}\Big|<\delta_0 \tag{2.2.23}$$

for all $x_1,x_2\in C_i(H_0)$, $\varepsilon<\varepsilon_1$ and $d<d_0$, $j=1,2$, where $+$ is taken if $j=1$ and $-$ is taken if $j=2$.

(i) Let h_0 be large enough such that by Lemma 2.2.7,

$$\sup_{\theta_1,\theta_2\in[0,2\pi]}\Big|P_{h_0,\theta_1}\{H^0_{\tau^0(0,\infty)}\in\{0\}\times B\}-P_{h_0,\theta_2}\{H^0_{\tau^0(0,\infty)}\in\{0\}\times B\}\Big|<\delta_0$$

for all intervals $B\subset[0,2\pi]$.

(ii) Let $\eta<(\gamma_2-\gamma_1)/4$ be small enough such that by Lemma 2.2.6,

$$\sup_{\theta\in[0,2\pi]}P_{h_0,\theta}\{H^0_{\tau^0(0,\infty)}\in\{0\}\times B_i\}<\delta_0,$$

where $B_i=[\gamma_i-\eta,\gamma_i+\eta]$.

(iii) Let $T > 0$ be large enough such that

$$P\Big\{\sup_{0\le t\le T} W_t < h_0 + 1\Big\} < \delta_0.$$

Fix $d < d_0$ and let $\varepsilon_0 < \varepsilon_1$ be small enough such that $h_0 < d\ {\varepsilon_0}^{-1/2}$ and
(iv) by Lemma 2.2.5,

$$\sup_{\theta\in[0,2\pi]} P_{h_0,\theta}\Big\{\sup_{0\le t\le T}\big|\underline{H}^{\varepsilon}_t - \underline{H}^0_t\big| > \eta\Big\} < \delta_0 \quad \text{for } 0<\varepsilon\le\varepsilon_0$$

and
(v) $$\sup_{\theta\in[0,2\pi]} P_{h_0,\theta}\big\{H^0_{\tau^0(-1,(H_2-H_1)/\sqrt{\varepsilon})} \neq -1\big\} < \delta_0, \qquad 0<\varepsilon<\varepsilon_0.$$

Define $D^{\varepsilon} = (0, (H_2 - H_1)/\sqrt{\varepsilon})$. For all $B \subset [0, 2\pi]$ we have

$$P_{h_0,\theta}\big\{\underline{H}^0_{\tau^0(D^{\varepsilon})} \in \{0\}\times B\big\} = P_{h_0,\theta}\big\{\underline{H}^0_{\tau^0(0,\infty)} \in \{0\}\times B,\ H^0_{\tau^0(D^{\varepsilon})} = 0\big\}.$$

So we get by (v),

$$P_{h_0,\theta}\big\{\underline{H}^0_{\tau^0(D^{\varepsilon})} \in \{0\}\times B\big\} - P_{h_0,\theta}\big\{\underline{H}^0_{\tau^0(0,\infty)} \in \{0\}\times B\big\} \le \delta_0. \tag{2.2.24}$$

Let $D_d = D_l(H_1 + d, H_2 - d)$ and $\underline{H}(\underline{X}^{\varepsilon}_t) = (H(\underline{X}^{\varepsilon}_t), \theta(\underline{X}^{\varepsilon}_t))$. Using the strong Markov property we get, by (2.2.22) and (2.2.23),

$$\begin{aligned}
&\Big|P_{\underline{X}_1}\big\{\underline{H}(\underline{X}^{\varepsilon}_{\tau^{\varepsilon}(D)}) \in \{H_2\}\times S\big\} - P_{\underline{X}_2}\big\{\underline{H}(\underline{X}^{\varepsilon}_{\tau^{\varepsilon}(D)}) \in \{H_2\}\times S\big\}\Big| \\
&\quad = \Big|P_{\underline{X}_1}\big\{\underline{H}(\underline{X}^{\varepsilon}_{\tau^{\varepsilon}(D)}) \in \{H_2\}\times S \,\big|\, H\big(\underline{X}^{\varepsilon}_{\tau^{\varepsilon}(D_d)}\big) = H_2 - d\big\} \\
&\qquad\quad \times \big(P_{\underline{X}_1}\big\{H(\underline{X}^{\varepsilon}_{\tau^{\varepsilon}(D_d)}) = H_2 - d\big\} - P_{\underline{X}_2}\big\{H(\underline{X}^{\varepsilon}_{\tau^{\varepsilon}(D_d)}) = H_2 - d\big\}\big) \\
&\qquad\quad + \big(P_{\underline{X}_1}\big\{\underline{H}(\underline{X}^{\varepsilon}_{\tau^{\varepsilon}(D)}) \in \{H_2\}\times S \,\big|\, H(\underline{X}^{\varepsilon}_{\tau^{\varepsilon}(D_d)}) = H_2 - d\big\} \\
&\qquad\qquad - P_{\underline{X}_2}\big\{\underline{H}(\underline{X}^{\varepsilon}_{\tau^{\varepsilon}(D)}) \in \{H_2\}\times S \,\big|\, H(\underline{X}^{\varepsilon}_{\tau^{\varepsilon}(D_d)}) = H_2 - d\big\}\big) \\
&\qquad\qquad \times P_{\underline{X}_2}\big\{H(\underline{X}^{\varepsilon}_{\tau^{\varepsilon}(D_d)}) = H_2 - d\big\} \\
&\qquad\quad + P_{\underline{X}_1}\big\{\underline{H}(\underline{X}^{\varepsilon}_{\tau^{\varepsilon}(D)}) \in \{H_2\}\times S \,\big|\, H(\underline{X}^{\varepsilon}_{\tau^{\varepsilon}(D_d)}) = H_1 + d\big\} \\
&\qquad\qquad \times P_{\underline{X}_1}\big\{H(\underline{X}^{\varepsilon}_{\tau^{\varepsilon}(D_d)}) = H_1 + d\big\} \\
&\qquad\quad - P_{\underline{X}_2}\big\{\underline{H}(\underline{X}^{\varepsilon}_{\tau^{\varepsilon}(D)}) \in \{H_2\}\times S \,\big|\, H(\underline{X}^{\varepsilon}_{\tau^{\varepsilon}(D_d)}) = H_1 + d\big\} \\
&\qquad\qquad\qquad\qquad \times P_{\underline{X}_2}\big\{H(\underline{X}^{\varepsilon}_{\tau^{\varepsilon}(D_d)}) = H_1 + d\big\}\Big| \\
&\quad \le \delta_0 + \Big|P_{\underline{X}_1}\big\{\underline{H}(\underline{X}^{\varepsilon}_{\tau^{\varepsilon}(D)}) \in \{H_2\}\times S \,\big|\, H(\underline{X}^{\varepsilon}_{\tau^{\varepsilon}(D_d)}) = H_2 - d\big\} \\
&\qquad\qquad - P_{\underline{X}_2}\big\{\underline{H}(\underline{X}^{\varepsilon}_{\tau^{\varepsilon}(D)}) \in \{H_2\}\times S \,\big|\, H(\underline{X}^{\varepsilon}_{\tau^{\varepsilon}(D_d)}) = H_2 - d\big\}\Big| \\
&\qquad + \sup_{\underline{X}\in C_l(H_1+d)} P_{\underline{X}}\big\{\underline{H}(\underline{X}^{\varepsilon}_{\tau^{\varepsilon}(D)}) \in \{H_2\}\times S\big\} \\
&\quad \le 2\delta_0 + \sup_{\underline{X}_3,\underline{X}_4\in C_l(H_2-d)} \Big|P_{\underline{X}_3}\big\{\underline{H}(\underline{X}^{\varepsilon}_{\tau^{\varepsilon}(D)}) \in \{H_2\}\times S\big\} \\
&\qquad\qquad - P_{\underline{X}_4}\big\{\underline{H}(\underline{X}^{\varepsilon}_{\tau^{\varepsilon}(D)}) \in \{H_2\}\times S\big\}\Big|.
\end{aligned} \tag{2.2.25}$$

Using the strong Markov property and the fact that the distributions of $\underline{X}^\varepsilon_\tau$ and $\tilde{\underline{X}}^\varepsilon_\tau$ are the same if τ is the time of first exit from a bounded domain, we can bound from above the supremum in the right-hand side of the above estimate by

$$\sup_{\theta_1,\,\theta_2\in[0,\,2\pi]}\Big|P_{h_0,\,\theta_1}\big\{\mathbb{H}^\varepsilon_{\tau^\varepsilon(D^\varepsilon)}\in\{0\}\times S\big\}-P_{h_0,\,\theta_2}\big\{\mathbb{H}^\varepsilon_{\tau^\varepsilon(D^\varepsilon)}\in\{0\}\times S\big\}\Big|$$

as $h_0<d\,\varepsilon^{-1/2}$. We get by (2.2.24) and (i),

$$\begin{aligned}
&\Big|P_{h_0,\,\theta_1}\big\{\mathbb{H}^\varepsilon_{\tau^\varepsilon(D^\varepsilon)}\in\{0\}\times S\big\}-P_{h_0,\,\theta_2}\big\{\mathbb{H}^\varepsilon_{\tau^\varepsilon(D^\varepsilon)}\in\{0\}\times S\big\}\Big|\\
&\quad\le 2\sup_{\theta\in[0,\,2\pi]}\Big|P_{h_0,\,\theta}\big\{\mathbb{H}^\varepsilon_{\tau^\varepsilon(D^\varepsilon)}\in\{0\}\times S\big\}-P_{h_0,\,\theta}\big\{\mathbb{H}^0_{\tau^0(D^\varepsilon)}\in\{0\}\times S_\eta\big\}\Big|\\
&\qquad+\sup_{\theta_1,\,\theta_2\in[0,\,2\pi]}\Big|P_{h_0,\,\theta_1}\big\{\mathbb{H}^0_{\tau^0(D^\varepsilon)}\in\{0\}\times S_\eta\big\}-P_{h_0,\,\theta_2}\big\{\mathbb{H}^0_{\tau^0(D^\varepsilon)}\in\{0\}\times S_\eta\big\}\Big|\\
&\quad\le 2\sup_{\theta\in[0,\,2\pi]}\Big|P_{h_0,\,\theta}\big\{\mathbb{H}^\varepsilon_{\tau^\varepsilon(D^\varepsilon)}\in\{0\}\times S\big\}-P_{h_0,\,\theta}\big\{\mathbb{H}^0_{\tau^0(D^\varepsilon)}\in\{0\}\times S_\eta\big\}\Big|+3\delta_0,
\end{aligned}$$

where $S_\eta=[\gamma_1+\eta,\gamma_2-\eta]$. Let $\tilde{S}_\eta=[\gamma_1-\eta,\gamma_2+\eta]$. Now

$$\begin{aligned}
&\Big|P_{h_0,\,\theta}\big\{\mathbb{H}^\varepsilon_{\tau^\varepsilon(D^\varepsilon)}\in\{0\}\times S\big\}-P_{h_0,\,\theta}\big\{\mathbb{H}^0_{\tau^0(D^\varepsilon)}\in\{0\}\times S_\eta\big\}\Big|\\
&\quad\le P_{h_0,\,\theta}\big\{\mathbb{H}^\varepsilon_{\tau^\varepsilon(D^\varepsilon)}\in\{0\}\times S,\,\mathbb{H}^0_{\tau^0(D^\varepsilon)}\notin\{0\}\times S_\eta\big\}\\
&\qquad+P_{h_0,\,\theta}\big\{\mathbb{H}^\varepsilon_{\tau^\varepsilon(D^\varepsilon)}\notin\{0\}\times S,\,\mathbb{H}^0_{\tau^0(D^\varepsilon)}\in\{0\}\times S_\eta\big\}\\
&\quad\le P_{h_0,\,\theta}\big\{\mathbb{H}^\varepsilon_{\tau^\varepsilon(D^\varepsilon)}\in\{0\}\times S,\,\mathbb{H}^0_{\tau^0(D^\varepsilon)}\notin\{0\}\times \tilde{S}_\eta\big\}\\
&\qquad+P_{h_0,\,\theta}\big\{\mathbb{H}^0_{\tau^0(D^\varepsilon)}\in\{0\}\times(\tilde{S}_\eta\setminus S_\eta)\big\}\\
&\qquad+P_{h_0,\,\theta}\big\{H^0_{\tau^0(-1,\ \varepsilon^{-1/2}(H_2-H_1))}\neq-1\big\}\\
&\qquad+P_{h_0,\,\theta}\Big\{\sup_{0\le t\le\tau^0(-1,\,\infty)}\big|\mathbb{H}^\varepsilon_t-\mathbb{H}^0_t\big|>\eta\Big|H^0_{\tau^0(-1,\ \varepsilon^{-1/2}(H_2-H_1))}=-1\Big\}.
\end{aligned}$$

By (ii) this is not greater than

$$\begin{aligned}
&2\delta_0+2\,P_{h_0,\,\theta}\Big\{\sup_{0\le t\le\tau^0(-1,\,\infty)}\big|\mathbb{H}^\varepsilon_t-\mathbb{H}^0_t\big|>\eta\Big|H^0_{\tau^0(-1,\,\varepsilon^{-1/2}(H_2-H_1))}=-1\big\}\\
&\qquad+2P_{h_0,\,\theta}\big\{H^0_{\tau^0(-1,\ \varepsilon^{-1/2}(H_2-H_1))}\neq-1\big\},
\end{aligned}$$

and by (iii), (iv) and (v) this can be estimated by

$$\begin{aligned}
&4\delta_0+3\,P_{h_0,\theta}\Big\{\sup_{0\le t\le\tau^0(-1,\,\infty)}\big|\mathbb{H}^\varepsilon_t-\mathbb{H}^0_t\big|>\eta,\,\tau^0(-1,\infty)<T\Big\}\\
&\qquad+3\,P_{h_0,\,\theta}\big\{\tau^0(-1,\infty)>T\big\}\le 10\delta_0.
\end{aligned}$$

Combining these estimates, we get the statement of the lemma. □

LEMMA 2.2.9. *Let* $D = D_l(H_1, H_2)$ *and* $H_1 < H_0 < H_2$. *For any* $\theta_0 \in [0, 2\pi]$ *and* $\delta > 0$ *there exists an open interval* S *containing* θ_0 *such that*

$$P^\varepsilon_{\underline{x}}\big\{H(X^\varepsilon_{\tau^\varepsilon(D)}) = H_2,\ \theta(X^\varepsilon_{\tau^\varepsilon(D)}) \in S\big\} < \delta$$

for all $\underline{x} \in C_l(H_0)$ *and small* ε.

As the exit distribution of the process $\underline{H}^0_t$ has a bounded density by Lemma 2.2.7, the statement can be proved by similar arguments as used in the proof of Lemma 2.2.8.

As mentioned above, it is easy to check that the statements of Lemmas 2.2.8 and 2.2.9 are also valid if H_2 is replaced by H_1.

Now we are able to prove a lemma which corresponds to the first part of the proof of Lemma 3.5 in [7]. Actually it gives the Markov property of the limiting process on the graph. Consider a vertex O_k. Without loss of generality we suppose that the segments meeting at O_k are I_1, I_2 and I_3, the region D_3 being the one adjoining the whole curve C_k. Suppose for definiteness that $H(\underline{x}) > H(\underline{x}_k)$ in D_3. Let $\gamma > 0$ be a small number, $D_k(\pm\gamma)$ be the connected component of $\{\underline{x}\colon H(O_k) - \gamma < H(\underline{x}) < H(O_k) + \gamma\}$ and $\tau^\varepsilon_k(\pm\gamma) = \inf\{t > 0\colon X^\varepsilon_t \notin D_k(\pm\gamma)\}$ and $C_{kj}(\gamma) = \{\underline{x} \in D_j\colon H(\underline{x}) = H(O_k) \pm \gamma\}$.

LEMMA 2.2.10. *For any* $\delta > 0$ *and* $0 < \gamma' < \gamma$ *there exists* $\varepsilon_0 > 0$ *such that for* $i,\ j = 1, 2, 3$, $0 < \varepsilon < \varepsilon_0$,

$$\sup_{\underline{x}_1, \underline{x}_2 \in C_{ki}(\gamma')} \Big| P^\varepsilon_{\underline{x}_1}\big\{X^\varepsilon_{\tau^\varepsilon_k(\pm\gamma)} \in C_{kj}(\gamma)\big\} - P^\varepsilon_{\underline{x}_2}\big\{X^\varepsilon_{\tau^\varepsilon_k(\pm\gamma)} \in C_{kj}(\gamma)\big\} \Big| < \delta.$$

PROOF. Let $C''_{ki} = C_{ki}(\gamma'/2) \cup C_{ki}((\gamma + \gamma')/2)$. By the strong Markov property, we get, for $\underline{x}_1, \underline{x}_2 \in C_{ki}(\gamma')$,

$$\begin{aligned} &\Big| P^\varepsilon_{\underline{x}_1}\big\{X^\varepsilon_{\tau^\varepsilon_k(\pm\gamma)} \in C_{kj}(\gamma)\big\} - P^\varepsilon_{\underline{x}_2}\big\{X^\varepsilon_{\tau^\varepsilon_k(\pm\gamma)} \in C_{kj}(\gamma)\big\} \Big| \\ &\qquad = \Big| \int_{C''_{ki}} g^\varepsilon(\xi)\big(p^\varepsilon(\underline{x}_1, d\xi) - p^\varepsilon(\underline{x}_2, d\xi)\big) \Big|, \end{aligned} \tag{2.2.26}$$

where $g^\varepsilon(\xi) = P_{\underline{x}} i^\varepsilon\{X^\varepsilon_{\tau^\varepsilon_k(\pm\gamma)} \in C_{kj}(\gamma)\}$ and $p^\varepsilon(\underline{x}, B) = P^\varepsilon_{\underline{x}}\{X^\varepsilon_{\tau^\varepsilon(D''_{ki})} \in B\}$. Here $B \subset C''_{ki}$ and D''_{ki} is the set containing the points between the two curves forming the set C''_{ki}.

Let $\underline{x}^0_1, \dots, \underline{x}^0_{n_1}$ denote all zeros of the function H_y in the set C''_{ki} (see Lemma 1.1). By Lemma 2.2.9 there exist open subsets $\tilde{S}^0_1, \dots, \tilde{S}^0_{n_1}$ of C''_{ki} such that $\underline{x}^0_i \in \tilde{S}^0_i$ and

$$\sum_{l=1}^{n_1} P^\varepsilon_{\underline{x}}\big\{X^\varepsilon_{\tau^\varepsilon(D''_{ki})} \in \tilde{S}^0_l\big\} < \delta/6 \tag{2.2.27}$$

for all $\underline{x} \in C_{ki}(\gamma')$ and small ε.

Further, the function g^ε solves the equation $L^\varepsilon g^\varepsilon = 0$. Thus by Lemma 2.1.8 there exists an $A_{42} > 0$ such that

$$|g^\varepsilon(\xi_1) - g^\varepsilon(\xi_2)| \le A_{42}|\xi_1 - \xi_2| \quad \text{for all } \xi_1, \xi_2 \in C''_{ki} \Big\backslash \bigcup_{l=1}^{n_1} \tilde{S}_l^0.$$

This implies the existence of a partition $\tilde{S}_1, \dots, \tilde{S}_{n_2}$ of the set $C''_{ki} \backslash \bigcup_{l=1}^{n_1} \tilde{S}_l^0$ such that with points $z_l \in \tilde{S}_l$ for small ε,

$$\left|g^\varepsilon(z_l) - g^\varepsilon(\xi)\right| < \delta/6, \qquad \xi \in \tilde{S}_l. \tag{2.2.28}$$

As $0 \le g^\varepsilon \le 1$, the right-hand side of (2.2.26) is not greater than

$$\sum_{l=1}^{n_1} \int_{\tilde{S}_l^0} \big(p^\varepsilon(x_1, d\xi) + p^\varepsilon(x_2, d\xi)\big) + \sum_{l=1}^{n_2} \left| \int_{\tilde{S}_l} \big(p^\varepsilon(x_1, d\xi) - p^\varepsilon(x_2, d\xi)\big) \right|$$
$$+ \sum_{l=1}^{n_2} \left\{ \int_{\tilde{S}_l} \left|g^\varepsilon(z_l) - g^\varepsilon(\xi)\right| p^\varepsilon(x_1, d\xi) + \int_{\tilde{S}_l} \left|g^\varepsilon(z_l) - g^\varepsilon(\xi)\right| p^\varepsilon(x_2, d\xi) \right\}.$$

We can find $\varepsilon_0 > 0$ such that for $0 < \varepsilon < \varepsilon_0$, the first term of this sum is less than $\delta/3$ by (2.2.27), the third term is less than $\delta/3$ by (2.2.28) and the second term is less than $\delta/3$ by Lemma 2.2.8. □

From the considerations of the last two subsections, especially from the proof of Lemma 2.2.10, we can deduce a result for a boundary value problem with the operator $(1/\varepsilon)\bar{\nabla} H \cdot \nabla + (\partial^2/\partial y^2)$:

COROLLARY 2.2.11. *Let $H(x)$ satisfy the conditions of Theorem 1 and let G be a bounded domain in R^2 containing a trajectory $C_l(H_0)$ of the dynamical system (1.10) for some l and a noncritical value H_0. Then, for any A_{43} and $\delta > 0$, there exists an $\varepsilon_0 > 0$, such that for any measurable function f with* $\operatorname{ess\,sup}|f| < A_{43}$ *and any* $0 < \varepsilon < \varepsilon_0$, *the solution u^ε of the boundary value problem,*

$$\frac{1}{\varepsilon}\bar{\nabla} H \cdot \nabla u^\varepsilon + u^\varepsilon_{yy} = 0, \qquad u^\varepsilon|_{\partial G} = f,$$

has the property

$$\sup_{x_1, x_2 \in C_l(H_0)} \left|u^\varepsilon(x_1) - u^\varepsilon(x_2)\right| < \delta.$$

2.3. *Proof of Theorem* 1. Theorem 2.2 in [7] that deals with the case of nondegenerate perturbation is proved using some results numbered Lemmas 3.1–3.5. The proofs of Lemmas 3.1–3.4 in [7] use a further series of results denoted Lemmas 4.1–4.10. All those can be used for the situation here after suitable changes and alterations in the proofs are made. Lemma 3.5 in [7] also holds here, but the first part of the proof changes essentially due to the fact that the perturbation of the Hamiltonian system is degenerate. This affects especially the proof of the Markov property of the limiting process on the graph. Here the results of the previous subsections replace Lemmas 5.1–5.5 in [7]. With modifications of Lemmas 3.1–3.5 and 4.1–4.10 in [7] that are discussed in this subsection, the proof of Theorem 1 is analogous to the proof of Theorem 2.2 in [7].

We start with the lemmas of Section 4 in [7]. Lemma 4.1 can be replaced by Lemma 2.2.4 of the present paper and Lemma 4.2 can be replaced by Lemma 2.2.1. Lemmas 4.3 and 4.4 and the resulting discussion of Particular Cases 1, 2, and 3 hold also in the case considered here and the proofs are the same. The statements of Lemmas 4.5–4.8 in [7] are also true for the processes considered in the present paper. The probability that a one-dimensional diffusion with positive (negative) drift leaves an interval at the left (right) end increases if the diffusion coefficient increases. By using this fact, the proof of Lemma 4.5 for the degenerate case is analogous to the proof in [7]. To extend the proof of Lemma 4.6 in [7] for the situation here we additionally have to make use of assumption (vi) of Theorem 1 to make sure that the coefficient denoted by a^{11} in [7] does not disappear. The rest of the proof is analogous.

The proof of Lemma 4.7 for the situation here is very similar to that in [7]. We replace ΔH by H_{yy} and ∇H by $(0, H_y)^*$. The only difficulty is that we cannot estimate H_y^2 from below by a positive constant as $|\nabla H|^2$ is estimated in [7]. This estimate has been used in [7] to get an estimate

$$P_x^\varepsilon\left\{\int_{\sigma_k}^{\tau_k} |\nabla H(X_s^\varepsilon)|^2\, ds \geq A_{44}\,\varepsilon/2 \text{ or } \sigma_k < \tau^\varepsilon \leq \tau_k \,\middle|\, X_s^\varepsilon,\ s \leq \sigma_k\right\} \geq 1 - \alpha(\varepsilon)$$

with $A_{44} > 0$ sufficiently small and $\alpha(\varepsilon) \to 0$ for $\varepsilon \to 0$; τ_k, τ^ε and σ_k are Markov times. To get a corresponding estimate for $|\nabla H(X_s^\varepsilon)|^2$ replaced by $H_y^2(X_s^\varepsilon)$ we divide the set of all trajectories into those for which $\sup_{\sigma_k \leq s \leq \tau_k} |X_s^\varepsilon - X_s^\varepsilon(X_{\sigma_k}^\varepsilon)| < \eta$ [cf. (2.2.3)] and into the complement of this set where η is sufficiently small. On the first set we get an estimate as above because the relative amount of time which the process X_t^ε spends near the zeros of H_y is "small." The arguments are the same as in the proof of Lemma 2.2.4 (see also Lemmas 1.1 and 2.2.3). The probability of the second set tends to 0 for $\varepsilon \to 0$.

The statement of Lemma 4.7′ in [7] holds for the situation here with $\Delta H(x_k)$ replaced by $H_{yy}(x_k)$. Note that assumption (v) of Theorem 1 guarantees that $H_{yy}(x_k) \neq 0$. The statement of Lemma 4.8 in [7] holds in the degenerate case with the same proof. The statement of Lemma 4.9 in [7] holds also in the situation here with the operators L_i from (1.13), (1.14) and (1.15). The proof is the same as in [7] after replacing ΔH by H_{yy} and ∇H by $(0, H_y)^*$. Lemma 4.10 in [7] contains some misprints, so we will give it here again for our situation.

LEMMA 2.3.1. *Let us consider the first time $\tau^\varepsilon = \tau_i^\varepsilon(H_1, H_2)$ of leaving the region $D_i(H_1, H_2)$. Let g be a continuous function on $[H_1, H_2]$ and let φ be a function defined only at the points H_1, H_2. Then*

$$\lim_{\varepsilon \to 0} E_x^\varepsilon\left[\varphi(H(X_{\tau^\varepsilon}^\varepsilon)) + \int_0^{\tau^\varepsilon} g(H(X_t^\varepsilon))\, dt\right] = f(H(x))$$

uniformly in $\underset{\sim}{x} \in D_i(H_1, H_2)$, *where*

$$f(H) = \frac{u_i(H_2) - u_i(H)}{u_i(H_2) - u_i(H_1)}\left[\varphi(H_1) + \int_{H_1}^{H} \big(u_i(h) - u_i(H_1)\big) g(h)\, dv_i(h)\right]$$
$$+ \frac{u_i(H) - u_i(H_1)}{u_i(H_2) - u_i(H_1)}\left[\varphi(H_2) + \int_{H}^{H_2} \big(u_i(H_2) - u_i(h)\big) g(h)\, dv_i(h)\right].$$

Here u_i and dv_i are the scale function and the speed measure of the diffusion process governed by the operator L_i. In general, there are no explicit formulas for u_i and v_i in this setting, but they are not needed in the proof, which can be copied from [7].

Now we are able to discuss Lemmas 3.1–3.5 in [7]. Lemma 3.1 can be used in the same form. In the proof of Lemma 3.2 the same changes have to be made as indicated above for the proof of Lemma 4.9. Lemma 3.3 holds also in the situation here and is a consequence of the corresponding Lemma 4.9 in this situation, too. Also the corresponding statement of Lemma 3.4 is true here and can be proved as in [7].

Thus we have the form of the operators governing the limiting diffusion in the interior of the edges of the graph and we can discuss the question of accessibility of the boundaries of the edges for these diffusions. Let $O_k \sim I_i$ be an interior vertex of the graph corresponding to a saddle point $\underset{\sim}{x}_k$ of $H(\underset{\sim}{x})$. The coefficients $A_i(H)$ and $B_i(H)$ in (1.13) [see (1.14) and (1.15)] have finite limits if H tends to $H_k = H(\underset{\sim}{x}_k)$. We have $0 < A_{45} < \int_{C_i(H)} H_y^2 |\nabla H|^{-1}\, dl < A_{46} < \infty$ for H close to $H_k = H(\underset{\sim}{x}_k)$, but $\lambda_i(H) = \int_{C_i(H)} |\nabla H|^{-1}\, dl$ tends to infinity as $H \to H_k$. If $S_i(H)$ denotes the area of the domain in R^2 bounded by $C_i(H)$, then $S_i'(H) = \lambda_i(H)$. Thus, the integral $\int_{H_0}^{H_k} \lambda_i(H)\, dH$ is finite for $(H_0, i) \in I_i$. This implies that

$$u_i(H_k) = \int_{H_0}^{H_k} \exp\left\{-\int_{H_0}^{z} 2B_i(u)A_i(u)\, du\right\} dz$$

and

$$v_i(H_k) = \int_{H_0}^{H_k} A_i(z)^{-1} \exp\left\{\int_{H_0}^{z} 2B_i(u)/A_i(u)\, du\right\} dz$$

are finite. Thus, the vertex O_k is accessible for all points of $I_i \sim O_k$ [4], and a gluing condition should be imposed for each interior vertex of the graph.

If O_k corresponds to an extremal point $\underset{\sim}{x}_k$ of the Hamiltonian $H(\underset{\sim}{x})$, and $I_i \sim O_k$, then near O_k the drift coefficient $B_i(H)$ is bounded, always has the sign to drive the diffusion away from the vertex and $|B_i(H)| > A_{47} > 0$. The diffusion coefficient $A_i(H)$ can be estimated $A_i(H) < A_{48}|H - H(\underset{\sim}{x}_k)|$, $A_{48} > 0$. The inaccessibility of the exterior vertex follows now from the respective property of the diffusion governed by the operator $\bar{L}_i f(H) = A_{48}|H - H(\underset{\sim}{x}_k)| f''(H) \pm A_{47} f'(H)$ with the sign so that the drift drives the diffusion away from the vertex.

Now we show that the statement of Lemma 3.5 in [7] holds also in the degenerate case. The proof starts as the corresponding proof in [7]. Formula (5.19) in [7] is already proved for the situation here by Lemma 2.2.10. Then, as in [7], we use the fact that the invariant measure μ for the processes $\underline{X}_t^\varepsilon$ (the Lebesgue measure) can be written as an integral with respect to the invariant measure of the embedded Markov chain. As in [7], let H_{k_1}, H_{k_2} be the limits between which the coordinate H on the edge I_j of the graph changes [if $H(I_j) = [H_{k_1}, \infty)$, introduce a new vertex with coordinates (H_{k_2}, j), where H_{k_2} is an arbitrary number greater than H_{k_1}]. For small $\delta > 0$, $I_j \sim O_k$, the set $C_{kj}(\delta) = \{\underline{x} \in D_j\colon H(\underline{x}) = H_{k_1} + \delta\}$ if $H(O_k) = H_{k_1}$, and $C_{kj}(\delta) = \{\underline{x} \in D_j\colon H(\underline{x}) = H_{k_2} - \delta\}$ if $H(O_k) = H_{k_2}$. Let $C(\delta) = \bigcup_{k,j} C_{kj}(\delta)$. By the same arguments as in [7] we get formula (5.23) in [7]:

$$\begin{aligned}(2.3.1)\qquad &\int_{R^2} g(H(\underline{x}))\chi_{D_j}(\underline{x})\,\mu(d\underline{x})\\ &\quad= \int_{\bigcup_{k:\ I_j\sim O_k} C_{kj}(\delta)} \nu^\varepsilon(d\underline{x}) E_{\underline{x}}^\varepsilon \int_0^{\tau_1} g\big(H(\underline{X}_t^\varepsilon)\big)\chi_{D_j}(\underline{X}_t^\varepsilon)\,dt\end{aligned}$$

for continuous functions g being different from zero only in $(H_{k_1} + \delta, H_{k_2} - \delta)$ and a Markov time τ_1 and measures ν^ε with the respective properties as in [7]. Let $d\tilde{v}_j$ denote the speed measure of the limiting diffusion on the graph obtained in [7] (denoted there by dv_j) and let dv_j and u_j be the speed measure and the scale function, respectively, of the limiting diffusion here. It follows from the well-known formulas for u_j and v_j that u_j' and v_j' exist and are positive and continuous, and that

$$(2.3.2)\qquad u_j'(H)v_j'(H) = \frac{2}{A_j} = \frac{2\int_{C_j(H)} |\nabla H(\underline{x})|^{-1}\,dl}{\int_{C_j(H)} H_y^2(\underline{x})|\nabla H(\underline{x})|^{-1}\,dl}.$$

Using Lemma 2.3.1 and the fact that $\tilde{v}_j(H)$ can be taken to be equal to the area enclosed by $C_j(H)$ (see [7]), the identity (2.3.1) can be written as

$$\begin{aligned}&\int_{H_{k_1}+\delta}^{H_{k_2}-\delta} g(h)\,d\tilde{v}_j(h)\\ &\quad= \int_{H_{k_1}+\delta}^{H_{k_2}-\delta} g(h)\frac{\tilde{v}_j'(h)}{v_j'(h)}\,dv_j(h)\\ &\quad= \nu^\varepsilon(C_{k_1 j}(\delta))\bigg[\frac{u_j(H_{k_1}+\delta) - u_j(H_{k_1}+\delta')}{u_j(H_{k_2}-\delta') - u_j(H_{k_1}+\delta')}\\ &\qquad\qquad \times \int_{H_{k_1}+\delta}^{H_{k_2}-\delta} \big(u_j(H_{k_2}-\delta') - u_j(h)\big) g(h)\,dv_j(h) + o(1)\bigg]\\ &\qquad + \nu^\varepsilon(C_{k_2 j}(\delta))\bigg[\frac{u_j(H_{k_2}-\delta') - u_j(H_{k_2}-\delta)}{u_j(H_{k_2}-\delta') - u_j(H_{k_1}+\delta')}\\ &\qquad\qquad \times \int_{H_{k_1}+\delta}^{H_{k_2}-\delta} \big(u_j(h) - u_j(H_{k_1}+\delta')\big) g(h)\,dv_j(h) + o(1)\bigg].\end{aligned}$$

Thus

$$
\begin{aligned}
\frac{\tilde v_j'(h)}{v_j'(h)} &= \nu^\varepsilon(C_{k_1 j}(\delta))\frac{u_j(H_{k_1}+\delta)-u_j(H_{k_1}+\delta')}{u_j(H_{k_2}-\delta')-u_j(H_{k_1}+\delta')} \\
&\quad\times\big(u_j(H_{k_2}-\delta')-u_j(h)\big) \\
&\quad+\nu^\varepsilon(C_{k_2 j}(\delta))\frac{u_j(H_{k_2}-\delta')-u_j(H_{k_2}-\delta)}{u_j(H_{k_2}-\delta')-u_j(H_{k_1}+\delta')} \\
&\quad\times\big(u_j(h)-u_j(H_{k_1}+\delta')\big)+o(1)
\end{aligned} \tag{2.3.3}
$$

for dv_j-almost all $h\in(H_{k_1}+\delta, H_{k_2}-\delta)$ and, as v_j' is strictly positive and all the functions are continuous, the formula (2.3.3) holds for all $h\in[H_{k_1}+\delta, H_{k_2}-\delta]$. If we take $h=H_{k_1}+\delta$ and $h=H_{k_2}-\delta$ we get a linear system for $\nu^\varepsilon(C_{k_1}(\delta))$ and $\nu^\varepsilon(C_{k_2}(\delta))$ from which we can easily deduce that

$$
\left|\nu^\varepsilon(C_{k_l}(\delta))-\frac{\tilde v_j'(H_{k_l})}{v_j'(H_{k_l})u_j'(H_{k_l})}\,\frac{1}{\delta-\delta'}\right|<\frac{\kappa}{\delta-\delta'},\qquad l=1,2,
$$

for some $\kappa>0$. As $(\tilde v_j'/v_j'\ u_j')=\frac12\int H_y^2\ |\nabla H|^{-1}\,dl$, we get the desired result by the same arguments as used at the end of the proof of Lemma 3.5 in [7].

3. The case of Hamiltonian $H(x,y)=\frac12 y^2+F(x)$. Equation (1.1) describes a nonlinear oscillator with 1 degree of freedom. Assume that the function $f(x)$ is in $C^\infty(R^1)$, $\liminf_{|x|\to\infty} f(x)\operatorname{sgn}(x)>0$, and let $f(x)$ have just a finite number of simple zeros, so that $f(x)$ and $f'(x)$ are not equal to zero simultaneously. Moreover, assume, for brevity, that all the local maxima of $F(x)=\int_0^x f(y)\,dy$ are different. Let $H(x,y)$ be the Hamilton function of the oscillator $H(x,y)=\frac12 y^2+F(x)$. Denote by Γ the graph corresponding to $H(x,y)$. Let Γ consist of n edges $I_1, I_2,\dots,I_n$ and m vertices $O_1,O_2,\dots,O_m$. Denote by $C_k(z)$ the component of $C(z)=\{(x,y)\colon H(x,y)=z\}$ corresponding to I_k. Of course, $C_k(z)$ is empty for some z and k. Let $S_k(z)$ be the area of the domain $G_k(z)\subset R^2$ bounded by $C_k(z)$ if the point $(z,k)\in\Gamma$ is not an end of I_k. If (z,k) is an end of I_k, put $S_k(z)=\lim_{z'\to z,\,(z',k)\in I_k} S_k(z')$. The function $S_k(z)$ is a function on the graph Γ. It is smooth inside the edges and can have discontinuities at the vertices.

Let $Y\colon R^2\to\Gamma$ be, as before, the mapping such that $Y(x,y)=(z,k)$, where $z=H(x,y)$, $k=k(x,y)$ is the index of the edge I_k containing the point corresponding to the level set component containing $(x,y)\in R^2$.

The function $H(x,y)=\frac12 y^2+F(x)$ satisfies the conditions of Theorem 1. Thus the processes $Y(X_t^\varepsilon, Y_t^\varepsilon)$, where $(X_t^\varepsilon, Y_t^\varepsilon)$ is defined by (1.6), converge to a diffusion process Y_t on Γ governed by the operators

$$
L_k v_k(z)=\frac{1}{2\lambda_k(z)}\frac{d}{dz}\left(a_k(z)\frac{dv_k(z)}{dz}\right),\qquad z\in(I_k),
$$

inside the edges and by the gluing conditions at the vertices. If O_i is an interior vertex and $I_{k_1} \sim O_i,\ I_{k_2} \sim O_i,\ I_{k_3} \sim O_i$, the gluing condition

$$
\begin{aligned}
&\alpha_{ik_1} \frac{dv_{k_1}(z)}{dz}\bigg|_{(z,\,k_1)=O_i} + \alpha_{ik_2} \frac{dv_{k_2}(z)}{dz}\bigg|_{(z,\,k_2)=O_i} \\
&\quad = \alpha_{ik_3} \frac{dv_{k_3}(z)}{dz}\bigg|_{(z,\,k_3)=O_i}
\end{aligned} \tag{3.1}
$$

should be imposed if the value of $H(x, y)$ is less than $H(Y^{-1}(O_i))$ for I_{k_1}, I_{k_2} and greater than $H(Y^{-1}(O_i))$ for I_{k_3}. The constants α_{ik_j} are defined as

$$
\alpha_{ik_j} = \int_{C_{k_j}(Y^{-1}(O_i))} \frac{H_y^2(x, y)}{|\nabla H(x, y)|}\, dl;
$$

$C_{k_j}(Y^{-1}(O_i))$ is the limit of $C_{k_j}(z')$ as $(z', k_j) \to O_i,\ j = 1, 2, 3$. The function $v_k(z)$ should be continuous on Γ.

In the case under consideration, one can give more explicit formulas for the coefficients of the operators and of the gluing conditions. Each set $C_k(z)$ is connected with two neighboring roots $\alpha_k(z)$ and $\beta_k(z)$ of the equation $F(x) = z$:

$$
C_k(z) = \left\{(x, y) \in R^2\colon \alpha_k(z) \le x \le \beta_k(z),\ y = \pm\sqrt{2(z - F(x))}\right\}.
$$

Then the coefficients $a_k(z)$, $\lambda_k(z)$ have the form

$$
\begin{aligned}
a_k(z) &= \int_{C_k(z)} \frac{H_y^2(x, y)\, dl}{|\nabla H(x, y)|} \\
&= 2\int_{\alpha_k(z)}^{\beta_k(z)} \sqrt{2(z - F(x))}\, dx = S_k(z), \\
\lambda_k(z) &= \int_{C_k(z)} \frac{dl}{|\nabla H(x, y)|} = 2\int_{\alpha_k(z)}^{\beta_k(z)} \frac{dx}{\sqrt{2(z - F(x))}} = \frac{d}{dz} S_k(z).
\end{aligned}
$$

We used here that $|\nabla H(x, y)| = \sqrt{f^2(x) + 2(z - F(x))},\ H_y^2 = 2(z - F(x))$,

$$
dl = \frac{|\nabla H(x, y)|}{|H_y(x, y)|}\, dx \quad \text{for } (x, y) \in C_k(z).
$$

Actually, the equality $\lambda_k(z) = S'_k(z)$ is true for any Hamiltonian $H(x, y)$. Thus the operators L_k can be written as

$$
L_k v_k(z) = \frac{1}{2S'_k(z)} \frac{d}{dz}\left(S_k(z) \frac{dv_k(z)}{dz}\right), \qquad (z, k) \in I_k, \tag{3.2}
$$

where $S_k(z)$ is the area of the domain bounded by $C_k(z)$.

The coefficients of the gluing conditions (3.1) can be expressed through the areas of the corresponding domains as well:

$$\alpha_{ik_j} = S_{k_j}(z), \qquad (z, k_j) = O_i, \qquad j = 1, 2, 3.$$

We have the following result.

THEOREM 2. *Assume that $f(x)$ satisfies the conditions formulated in the beginning of this section. Let $S_k(z)$ be the area of the domain bounded by $C_k(z)$ [for the critical values $z = H(x, y)$, the area is defined as the corresponding limit]. Then the processes $Y(X_t^\varepsilon, Y_t^\varepsilon)$, where $(X_t^\varepsilon, Y_t^\varepsilon)$ is defined by (1.6), converge weakly in the space of continuous functions φ: $[0, T] \to \Gamma$, for any $T > 0$, to the diffusion process on Γ governed by the operators (3.2) inside the edges and by the gluing conditions (3.1) with $\alpha_{ik_j} = S_{k_j}(z_i)$ at each interior vertex $O_i = (z_i, k_1) = (z_i, k_2) = (z_i, k_3)$.*

Using this result, one can calculate the main terms of the asymptotics for many interesting characteristics of the process $(\tilde{X}_t^\varepsilon, \tilde{Y}_t^\varepsilon)$ as $\varepsilon \downarrow 0$. Consider, for example, the asymptotics of the expectation of the exit time from a domain $G \subset R^2$: $u(x, y) = \lim_{\varepsilon\downarrow 0} \varepsilon E_{x, y}\tilde{\tau}^\varepsilon$, $\tilde{\tau}^\varepsilon = \min\{t: (\tilde{X}_t^\varepsilon, \tilde{Y}_t^\varepsilon) \notin G\}$. If the trajectory of the nonperturbed system, starting in $(x, y) \in G$, leaves G in a finite time, then $u(x, y) = 0$. Therefore just the domains bounded by the nonperturbed trajectories are of interest. Let G be bounded by the nonsingular trajectories $C_{k_1}(z_1), \ldots, C_{k_l}(z_l)$; (k_i, z_i) is the point of Γ corresponding to $C_{k_i}(z_i)$. In the example shown in Figure 1, $C_{k_1}(z_1) = \partial G_1$ and $C_{k_2}(z_2) = \partial G_2$.

LEMMA 3.1. *Let $\hat{\Gamma}$ be the domain in Γ bounded by the points (z_i, k_i), $i \in \{1, \ldots, l\}$, $\tau = \min\{t: Y_t \notin \hat{\Gamma}\}$ and $v_k(z) = E_{z,k}\tau$, $(z, k) \in \hat{\Gamma}$. Then $\lim_{\varepsilon\downarrow 0} \varepsilon E_{x, y}\tilde{\tau}^\varepsilon = v_k(z)$, where $Y(x, y) = (z, k) \in \hat{\Gamma}$.*

PROOF. Let $z_l = \max\{z_1, \ldots, z_l\}$ and G^* be the domain in R^2 bounded by $C_{k_l}(z_l)$, $\hat{\tau}^\varepsilon = \min\{t: (X_t^\varepsilon, Y_t^\varepsilon) \notin G^*\}$. It is clear that $P_{x, y}^\varepsilon\{\tau^\varepsilon \le \hat{\tau}^\varepsilon\} = 1$, $(x, y) \in G^*$, where $\tau^\varepsilon = \varepsilon\tilde{\tau}^\varepsilon = \min\{t: (X_t^\varepsilon, Y_t^\varepsilon) \notin G\}$. Applying the Itô formula to $H(X_t^\varepsilon, Y_t^\varepsilon)$, we have

$$H(X_{\hat{\tau}^\varepsilon}^\varepsilon, Y_{\hat{\tau}^\varepsilon}^\varepsilon) - H(x, y) = \int_0^{\hat{\tau}^\varepsilon} H_y(X_s^\varepsilon, Y_s^\varepsilon)\, dW_s + \tfrac{1}{2}\int_0^{\hat{\tau}^\varepsilon} H_{yy}\, ds.$$

Taking into account that $H_{yy} \equiv 1$, we conclude that

$$\varepsilon E_{x, y}^\varepsilon \tilde{\tau}^\varepsilon = E_{x, y}^\varepsilon \tau^\varepsilon \le E_{x, y}^\varepsilon \hat{\tau}^\varepsilon \le z - \min_{(x, y)\in G^*} H(x, y) < \infty.$$

Since the last bound holds uniformly for all $(x, y) \in G^*$, we derive, using the Markov property, that all the moments of τ^ε are bounded uniformly in $\varepsilon > 0$, $(x, y) \in G$. In particular, $E_{x, y}^\varepsilon(\tau^\varepsilon)^2 \le B < \infty$.

Let $\chi_{\tau^\varepsilon \le T}$ be the indicator function of $\{\tau^\varepsilon \le T\}$. For any $T > 0$, we have

$$\begin{aligned} 0 &\le E_{x, y}^\varepsilon \tau^\varepsilon - E_{x, y}^\varepsilon \tau^\varepsilon \chi_{\tau^\varepsilon \le T} = E_{x,y}^\varepsilon \tau^\varepsilon \chi_{\tau^\varepsilon > T} \\ &\le \sqrt{E_{x,y}^\varepsilon (\tau^\varepsilon)^2\ P_{x, y}^\varepsilon\{\tau^\varepsilon < T\}} \le \frac{B}{T}. \end{aligned} \tag{3.3}$$

Now, one can consider $\tau^\varepsilon \chi_{\tau^\varepsilon \le T}$ as a functional on the trajectories of the process $Y_t^\varepsilon = Y(X_t^\varepsilon, Y_t^\varepsilon)$ on Γ. These processes converge weakly in C_{0T} to the process Y_t on Γ as $\varepsilon \downarrow 0$. The functional $\tau^\varepsilon \chi_{\tau^\varepsilon \le T}$ is not continuous in C_{0T}, but the set where it is discontinuous has probability zero for the limiting process Y_t, since the diffusion coefficient of Y_t at the boundary of $\hat{\Gamma}$ is not zero. Thus we can conclude from the weak convergence that

$$\lim_{\varepsilon \downarrow 0} E_{x,y}^\varepsilon \tau^\varepsilon \chi_{\tau^\varepsilon \le T} = E_{Y(x,y)} \tau \chi_{\tau \le T}.$$

This equality together with (3.3) implies the statement of the lemma. □

The function $v_k(z) = E_{z,k}\tau$, as it follows from the theory of Markov processes, is the solution of the boundary problem

$$L_k v_k(z) = -1, \qquad (z,k) \in \hat{\Gamma}; (z,k) \text{ is not a vertex,}$$
$$v_{k_i}(z_i) = 0, \qquad i = 1, \dots, l. \tag{3.4}$$

One should add the gluing conditions at the interior vertices and the continuity on $\hat{\Gamma}$.

Problem (3.4) can be solved, in a sense, explicitly. Equations (3.4) are linear, and the general solution is the sum of a solution of the nonhomogeneous problem (satisfying the gluing condition, of course) and the general solution of the equations with zero in the right-hand side. It is clear that the function $v_k(z) \equiv -2z$ satisfies the nonhomogeneous equations. The solutions of the homogeneous equations (satisfying the gluing conditions) can be constructed in the following way. Single out one of the boundary points of $\hat{\Gamma}$, say (z_l, k_l). Consider the edges of $\hat{\Gamma}$ which contain a boundary point different from (z_l, k_l) (these are the edges $I_{k_1}, \dots, I_{k_{l-1}}$) and write a constant c_j on each I_{k_j}. Write zero on any edge of $\hat{\Gamma}$ which has an exterior vertex and no boundary points, besides, maybe, the point (z_l, k_l). Now, define constants c_j for the rest of the edges of $\hat{\Gamma}$ so that if a vertex $O = (H_0, \nu)$ is the common point of I_{ν_1}, I_{ν_2}, I_{ν_3} and the coordinate z on I_{ν_1} and I_{ν_2} is smaller than H_0 (thus, z is greater than H_0 on I_{ν_3}), then $c_{\nu_3} = c_{\nu_1} + c_{\nu_2}$. This condition allows us to extend the sequence $c_1, \dots, c_{l-1}$ to all the edges included in $\hat{\Gamma}$ in a unique way.

For any point $(z,k) \in \hat{\Gamma}$, there exists a unique path leading from (z,k) to (z_l, k_l). Let $C(t) = c_j$ and $S(t) = S_{k_j}(z)$ if $t = (z, k_j) \in \Gamma$. Put

$$w^{c_0, c_1, \dots, c_{l-1}}(z,k) = \int_{(z,k)}^{(z_l,k_l)} \frac{C(t)\,dt}{S(t)} + c_0.$$

It is easy to see that the function $w^{c_0, c_1, \dots, c_{l-1}}(z,k)$ on Γ satisfies the equations $L_k w^{c_0, c_1, \dots, c_{l-1}}(z,k) = 0$, if (z,k) is not a vertex, and satisfies the gluing conditions at the vertices. Choose the constants $c_0, c_1, \dots, c_{l-1}$ from the boundary conditions at the points (z_i, k_i):

$$w^{c_0, c_1, \dots, c_l}(z_i, k_i) = 2z_i, \qquad i = 1, \dots, l. \tag{3.5}$$

This is a system of linear algebraic equations with respect to $c_0, \dots, c_l$. One can check that system (3.5) defines the constants $c_0, \dots, c_l$ in a unique way. Then the function

$$v_k(z) = -2z + w^{c_0, c_1, \dots, c_{l-1}}(z, k)$$

is the solution of the problem (3.4) and

$$u(x, y) = \lim_{\varepsilon \downarrow 0} \varepsilon E^{\varepsilon}_{x, y} \tilde{\tau}^{\varepsilon} = v_{k(x, y)}(\tfrac{1}{2} y^2 + F(x));$$

$k(x, y)$ is the index of the edge containing the point $Y(x, y)$.

Let, for example, the function $f(x)$ and the domain G be as in Figure 1. Then $\hat{\Gamma}$ consists of edges $\tilde{I}_1 = (\partial_1, O_2)$, $\tilde{I}_2 = (O_2, \partial_2)$, $I_3 = (O_2, O_4)$, $I_4 = (O_3, O_4)$, $I_5 = (O_5, O_4)$; $\partial_1 = Y(\partial G_1)$, $\partial_2 = Y(\partial G_2)$. We prescribe 0 to I_4, I_5 and prescribe c_1 to $\tilde{I}_1$. The rule of extension of these constants to the other edges gives us $c_3 = 0$, $c_1 = c_2$. Let H_1 and H_2 be the values of $H(x, y)$ on ∂G_1 and ∂G_2, respectively, and let $H(O_2)$ be the value of $H(x, y)$ at O_2. The constants c_1, c_0 satisfy the equations

$$\begin{aligned} v_1(H_1) &= c_0 - 2H_1 = 0, \\ v_2(H_2) &= -2H_2 + c_0 + c_1 \int_{(H_2, 2)}^{(H_1, 1)} \frac{dt}{S(t)} = 0. \end{aligned}$$

Solving this system, we have

$$c_0 = 2H_1, \qquad c_1 = 2(H_2 - H_1)\left(\int_{(H_2,2)}^{(H_1,1)} \frac{dt}{S(t)}\right)^{-1},$$

and thus

$$\begin{aligned} u(x, y) &= -2H(x, y) + 2H_1 \\ &\quad + 2(H_2 - H_1)\left(\int_{(H_2, 2)}^{(H_1, 1)} \frac{dt}{S(t)}\right)^{-1} \int_{(H(x, y), k(x, y))}^{(H_1, 1)} \frac{dt}{S(t)}, \\ &\qquad\qquad (H(x, y), k(x, y)) \in \tilde{I}_1 \cup \tilde{I}_2; \\ u(x, y) &= u(O_2) + 2H(O_2) - 2H(x, y); \qquad Y(x, y) \in I_3 \cup I_4 \cup I_5. \end{aligned}$$

REFERENCES

[1] ADAMS, R. A. (1975). *Sobolev Spaces.* Academic Press, New York.

[2] BASKAKOVA, P. E. and BORODIN, A. N. (1995). Tables of distributions of functionals of Brownian motion. *J. Math. Sci.* 75 1873–1883.

[3] BURDZEIKO, B., IGNAT'EV, YU., KHASMINSKII, R. and B. SHAHGILDIAM (1979). Statistical analysis of a model of phase synchronization. In *Proceedings of the International Conference on Information Theory, Tbilisi*, 64–67 (in Russian).

[4] FELLER, W. (1954). Diffusion processes in one dimension. *Trans. Amer. Math. Soc.* 97 1–31.

[5] FREIDLIN, M. (1985). *Functional Integration and Partial Differential Equations.* Princeton Univ. Press.

[6] FREIDLIN, M. I. and WENTZELL, A. D. (1993). Diffusion processes on graphs and the averaging principle. *Ann. Probab.* 21 2215–2245.

[7] FREIDLIN, M. I. and WENTZELL, A. D. (1994). Random perturbations of Hamiltonian systems. *Mem. Amer. Math. Soc.* 109 (523).

[8] HÖRMANDER, L. (1985). *The Analysis of Linear Partial Differential Operators* 3. Springer, Berlin.

[9] OLEINIK, A. O. and RADKEVIC, E. V. (1973). *Second Order Equations with Nonnegative Characteristic Form.* Plenum, New York.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND 20742
E-MAIL: mif@math.umd.edu

INSTITUTE OF MATHEMATICAL STOCHASTICS
TECHNICAL UNIVERSITY OF DRESDEN
D-01062 DRESDEN
GERMANY
E-MAIL: matthias.weber@math.tu-dresden.de