Width of stochastic layers in near-integrable two-dimensional symplectic maps

D. Treschev
Moscow State University

Abstract

Width of stochastic layers around separatrix "eight" is studied for two-dimensional symplectic maps close to integrable. We prove that under rather general assumptions the relation $w/d \sim 1/\log \mu$ holds, where w is the width of the stochastic layer, d is the width of a lobe domain and μ is a multiplier at the corresponding hyperbolic point.

1 An informal introduction

Let T be an analytic symplectic self-map of the two-dimensional domain N. We assume that T is close to an integrable map. (The precise meaning of this assumption will be clarified below). We assume also that T has a hyperbolic fixed point $\hat{z}$ whose asymptotic curves (separatrices) look as shown on Fig. 1. Three invariant curves $\gamma_\pm$ and γ_0 closest to the separatrices form the boundary of the stochastic layer. The width w of the stochastic layer is one of important quantities characterizing chaotic properties of T in the vicinity of the separatrices.

In the present paper we show that under some natural assumptions the following relation holds:

$$w/d \sim 1/\lambda. \tag{1.1}$$

Here d is the width of a lobe domain D bounded by segments of the separatrices and $\lambda > 0$ is logarithm of the larger multiplier at the hyperbolic fixed point $\hat{z}$. The symbol $\sim$ means that if we have a smooth family T_ε, of analytic symplectic maps, where T_0 is integrable then for sufficiently small ε

$$C_1/\lambda(\varepsilon) < w(\varepsilon)/d(\varepsilon) < C_2/\lambda(\varepsilon),$$

where C_1 and C_2 are positive constants. Here the family T_ε must satisfy certain regularity conditions (see Section 4). We believe that these conditions are generic. In some cases the genericity is proved. The constants C_1 and C_2 can be estimated (see Section 4).

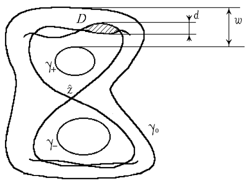

Figure 1: Separatrices of the hyperbolic fixed point $\hat{z}$ and the stochastic layer.

If $\lambda \sim 1$, we see that the quantities w and d are of the same order. However, if the stochastic layer appears when a resonant invariant curve of an integrable map disintegrates, the quantity λ is close to zero. Hence, in this case w is much greater than d. This situation is typical for exponentially small separatrix splitting. In particular, if T satisfies a certain symmetry condition, we expect that the estimate

$$\lim_{\varepsilon \to 0} \frac{w(\varepsilon)\lambda(\varepsilon)}{d(\varepsilon)} = \frac{4\pi}{k_0}, \tag{1.2}$$

holds, where

$$\begin{aligned} k_0 \quad = \quad & \inf\{k' : \text{ for all } k > k' \text{ the standard map} \\ & \begin{pmatrix} I \\ \varphi \end{pmatrix} \mapsto \begin{pmatrix} J \\ \psi \end{pmatrix} = \begin{pmatrix} I + \frac{k}{2\pi}\sin(2\pi\varphi) \\ \varphi + J \end{pmatrix} \\ & \text{has no invariant curve homotopic to the circle } I = 0\}. \end{aligned}$$

The constant $k_0 = 0.971635\ldots$ was evaluated numerically in [13, 19, 23].

It is possible to present an invariant version of the relation (1.1). Let $\mathcal{A}$ be the symplectic area of the stochastic layer and $\mathcal{A}_D$ the area of a lobe. Then we have:

$$\mathcal{A} \sim \frac{\mathcal{A}_D \log \mathcal{A}_D^{-1}}{\lambda^2}. \tag{1.3}$$

The corresponding analog of the formula (1.2) (under the same assumptions) is

$$\lim_{\varepsilon \to 0} \frac{\mathcal{A}\lambda^2}{\mathcal{A}_D \log \mathcal{A}_D^{-1}} = \frac{8\pi^2}{k_0}. \tag{1.4}$$

In the general (non-symmetric) case the fractions

$$\frac{w(\varepsilon)\lambda(\varepsilon)}{d(\varepsilon)}, \qquad \frac{\mathcal{A}\lambda^2}{\mathcal{A}_D \log \mathcal{A}_D^{-1}}$$

do not have limits as $\varepsilon \to 0$ but oscillate between two positive constants.

Formulas (1.1)–(1.4) can be regarded as relations between w, $\mathcal{A}$ and the quantities $\lambda, d, \mathcal{A}_D$. In the perturbative situation the last ones are standard to evaluate: to obtain the functions $d(\varepsilon)$ and $\mathcal{A}_D(\varepsilon)$ one can use the Poincaré-Melnikov theory or its generalizations to the case of exponentially small splitting; $\lambda(\varepsilon)$ in the main approximation is usually evaluated easily.

Particular cases of the relation (1.1) were discovered in [5, 9, 28], but no rigorous proofs were presented there. In these papers such useful object as separatrix map was introduced. Analyzing this map, Dovbysh [7] proved the estimate $w/d \leq \text{const}$ provided $\lambda \sim 1$. In paper [1] the estimate $w/d \geq \text{const}$ is established in the same case. Lazutkin [16] has obtained the estimate (1.1) for separatrices of the standard map.

Because of some technical reasons we have to postpone precise formulation of our main result to Section 4. The plan of the paper is as follows. In Sections 2–3 and 6 we introduce our main tools: normal coordinates, gluing maps and the separatrix map. The main results (Theorems 1–3) are presented in Section 4. In particular, relations (1.1) and (1.3) follow from Theorem 1. Relations (1.2) and (1.4) are consequences of Corollaries 4.3–4.4 provided Conjecture 2 is true. In Section 5 we discuss applications of Theorems 1–3. Proof of the theorems is contained in Sections 7–8. In Section 9 we prove some technical statements from Section 5.

2 Normal coordinates

Below we always assume that the map T is analytic in z.

In the vicinity of the point $\hat{z}$ the map T is a hyperbolic rotation. More precisely, the following lemma holds.

Lemma 2.1 *It is possible to choose in a neighborhood of the point $\hat{z}$ on N real-analytic symplectic coordinates (x, y) such that the following assertions hold.*

(1) Coordinates of $\hat{z}$ vanish.

(2) The map T has the form

$$(x, y) \to L(x, y) = (x\mathcal{M}, y/\mathcal{M}), \qquad \mathcal{M} = \mathcal{M}(xy). \tag{2.1}$$

where the function $\mathcal{M}$ is analytic and $\mathcal{M}(0) = \mu > 1$.

On a formal level existence of such coordinates was established by Birkhoff. They are called normal ones. Convergence of the procedure introducing the normal coordinates was established in [22]. The map (2.1) has the first integral xy.

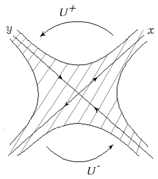

Figure 2: Continued normal coordinates and the gluing maps

Remark 2.1 *The normal coordinates* (x, y) *are not defined uniquely. For any real-analytic at zero function* $r(z)$ ($r(0) \neq 0$) *the coordinates* $x' = x/r(xy)$, $y' = y\, r(xy)$ *are normal as well.*

Remark 2.2 *Below we assume that inside the higher loop of the "eight" (see Fig. 1) we have* $x > 0$, $y > 0$.

Remark 2.3 *If the map* T *depends smoothly on a parameter (say,* ε*), the normal coordinates also can be chosen depending smoothly on* ε *while the point* $\hat{z} = \hat{z}(\varepsilon)$ *remains hyperbolic.*

The normal coordinates (x, y) can be continued to a neighborhood of the separatrices according to the following inductive procedure. Suppose that the point $z \in N$ has the coordinates (x, y). Then we define coordinates of the points $T(z)$ and $T^{-1}(z)$ by $L(x, y)$ and $L^{-1}(x, y)$ respectively. Far from the point $\hat{z}$ we have at least two different continuations of the coordinates (by T^k and by T^{-k}, see Fig. 2). Hence, we have two gluing transformations:

$$\begin{aligned} U^+ &: \{(x, y) : y \text{ is small}, \ x > 0, x \sim 1\} \to \{(x, y) : y > 0, y \sim 1, \ x \text{ is small}\}, \\ U^- &: \{(x, y) : y \text{ is small}, \ x < 0, x \sim 1\} \to \{(x, y) : y < 0, y \sim 1, \ x \text{ is small}\}. \end{aligned} \tag{2.2}$$

The map U^+ corresponds to the higher loop of the "eight" and U^- to the lower one. The maps L and $U^\pm$ obviously commute:

$$U^\pm \circ L = L \circ U^\pm. \tag{2.3}$$

3 Gluing maps

The following two assertions can be regarded as a motivation for one of our basic assumptions (see assumption **(B)** from Section 4).

Proposition 3.1 *Suppose that the map T has an analytic integral $F : N \to N$ (i.e. $F = F \circ T$). Then in normal coordinates $F = F(xy)$. Moreover, if the critical point $x = y = 0$ of the function F is non-degenerate, the gluing maps preserve the product xy and*

$$U^{\pm}(x, y) = \big(xy(\alpha_{\pm}^2/x + y\kappa_{\pm}(xy))^{-1}, \alpha_{\pm}^2/x + y\kappa_{\pm}(xy)\big)$$

Here $\alpha_{\pm}$ are positive constants and $\kappa_{\pm}(r)$ are functions analytic at the point $r = 0$.

Proof. The relation $F = F(xy)$ follows from the Tailor expansion of the equality $F = F \circ T$ at the point $x = y = 0$. Non-degeneracy of the critical point $x = y = 0$ means that there exists an inverse function $xy = xy(F)$. It is also a first integral. Putting $U^+(x, y) = (X, Y)$ we have:

$$xy = XY, \qquad dx \wedge dy = dX \wedge dY. \tag{3.1}$$

Putting $Y = Y(x, y)$, we see from the first equality (3.1) that $X = xy/Y(x, y)$. Then after simple transformations the second equality (3.1) implies the equation

$$y\partial Y/\partial y - x\partial Y/\partial x = Y.$$

The general solution real-analytic at $y = 0$ has the form $Y = x^{-1}a(xy)$. It remains to show that $a(0) > 0$. This inequality follows from our choice of normal coordinates (see relations (2.2)). Proposition 3.1 is proved.

If the map $T = T_\varepsilon$ is a perturbation of an integrable one, its gluing maps have the form

$$U_\varepsilon^{\pm} = \big(xy(\alpha_{\pm}^2/x + y\kappa_{\pm}(xy))^{-1}, \alpha_{\pm}^2/x + y\kappa_{\pm}(xy)\big) + \big(\hat{f}(x, y, \varepsilon), \hat{g}(x, y, \varepsilon)\big).$$

In general they do not preserve the function xy. According to the relation (2.3),

$$\hat{f}_{\pm}(x\mathcal{M}, y/\mathcal{M}, \varepsilon) = \mathcal{M}\hat{f}(x, y, \varepsilon), \quad \hat{g}_{\pm}(x\mathcal{M}, y/\mathcal{M}, \varepsilon) = \mathcal{M}^{-1}\hat{g}(x, y, \varepsilon). \tag{3.2}$$

Proposition 3.2 *If the maps $U^{\pm}$ are analytic in x and y in the domains*

$$\{x_0 < \pm x < X_0, \quad |y| < Y_0\}, \tag{3.3}$$

for some positive x_0, X_0 and Y_0, the following relations hold:

$$\hat{f}_{\pm}(x, y, \varepsilon) = xf_{\pm}(x, y, \varepsilon), \quad \hat{g}_{\pm}(x, y, \varepsilon) = yg_{\pm}(x, y, \varepsilon).$$

The functions $f_{\pm}$, $g_{\pm}$ are analytic in the domains (3.3) and

$$f_{\pm}(x\mathcal{M}, y/\mathcal{M}, \varepsilon) = f_{\pm}(x, y, \varepsilon), \quad g_{\pm}(x\mathcal{M}, y/\mathcal{M}, \varepsilon) = g_{\pm}(x, y, \varepsilon). \tag{3.4}$$

Proof of Proposition 3.2 is simple. It can be obtained with the help of (3.2) by expanding the functions $\hat{f}_{\pm}$, $\hat{g}_{\pm}$ into Tailor series in y.

4 Main results

From now on we assume that the map T depends smoothly on the parameter $\varepsilon \in [0, \varepsilon_0]$. In other words we will deal with the smooth family T_ε of analytic symplectic maps. Sometimes the "unperturbed" map T_0 turns out to be the identity. Because of this the zero value of the parameter ε is in a sense exceptional. We assume that for positive ε the family T_ε possesses the following properties.

(A) *The maps T_ε have hyperbolic fixed points $\hat{z}(\varepsilon)$ smoothly depending on ε. Separatrices of each point $\hat{z}(\varepsilon)$ form figure eight structure, see Fig. 1.*

(B) *The gluing maps $U_\varepsilon^\pm$ have the form*

$$\begin{aligned} U_\varepsilon^\pm(x, y) &= \big(xy(\alpha_\pm^2/x + y\kappa_\pm(xy, \varepsilon))^{-1},\, \alpha_\pm^2/x + y\kappa_\pm(xy, \varepsilon)\big) \\ &\qquad + \delta(\varepsilon)\big(xf_\pm(x, y, \varepsilon), yg_\pm(x, y, \varepsilon)\big), \end{aligned} \tag{4.1}$$

where $\kappa_\pm, f_\pm$ and $g_\pm$ are analytic in x, y and smooth in $\varepsilon \in (0, \varepsilon_0)$. The function $\delta(\varepsilon)$, $0 < \varepsilon \le \varepsilon_0$ is smooth, positive and $\lim_{\varepsilon\to 0} \delta(\varepsilon) = 0$.

We put

$$f_\pm(\xi, 0, \varepsilon) = \alpha_\pm^{-2}\hat{\nu}_\pm(\log|\xi|/\lambda(\varepsilon), \varepsilon), \qquad \lambda(\varepsilon) = \log \mathcal{M}(0, \varepsilon). \tag{4.2}$$

The functions $\hat{\nu}_\pm$ are obviously analytic. Due to (3.4) they are 1-periodic in the first argument. We assume that

(C) *The functions $\hat{\nu}_\pm$ have the form*

$$\hat{\nu}_\pm(\xi, \varepsilon) = \nu_\pm(\xi) + o(1) \quad \text{as} \quad \varepsilon \to 0, \tag{4.3}$$

where the functions $\nu_\pm$ do not vanish identically.

Below sometimes we need the following symmetry condition

(S) *The gluing maps $U_\varepsilon^\pm$ coincide, i.e. $U_\varepsilon^+(x, y) = -U_\varepsilon^-(-x, -y)$.*

Remark 4.1 *The term smoothness above means C^∞-smoothness. However, our results hold also in finitely smooth case.*

Remark 4.2 *Conditions* **(A)**–**(B)** *and relation (4.3) hold for any smooth family T_ε of analytic maps such that the fixed point $\hat{z}(0)$ of the map T_0 is hyperbolic, T_0 is integrable and separatrices of T_0 form a figure eight curve. Condition in this case is generic. We believe that conditions* **(A)**–**(B)** *are generic in the case $\lim_{\varepsilon\to 0} \lambda(\varepsilon) = 0$ as well.*

We define stochastic layer as the domain $\mathcal{SL}(\varepsilon) \subset N$ bounded by continuous invariant curves $\gamma_\pm$ and γ_0 (see Fig. 1), closest to the hyperbolic point $\hat{z}(\varepsilon)$. We are going to measure the width of $\mathcal{SL}(\varepsilon)$ in terms of xy. More precisely, we introduce the following quantities:

$$\underline{w}_* = \min_{(x,y)\in\gamma_*} |xy|, \quad \overline{w}_* = \max_{(x,y)\in\gamma_*} |xy|,$$

where $*$ denotes any of the symbols $+$, $-$ or 0. It is natural to say that the width w of the stochastic layer varies in the interval $(\underline{w}_\bullet + \underline{w}_0, \overline{w}_\bullet + \overline{w}_0)$, where

$$\underline{w}_\bullet = \min\{\underline{w}_+, \underline{w}_-\}, \quad \overline{w}_\bullet = \max\{\overline{w}_+, \overline{w}_-\}.$$

We define width of the lobe D as

$$d = \max_D xy - \min_D xy. \tag{4.4}$$

Since the function xy does not depend on choice of normal coordinates, the quantities $\underline{w}_*, \overline{w}_*$ and d are well-defined.

Lemma 4.1 *The following relations hold.*
(1) $d \sim \delta$, $\mathcal{A}_D \sim \lambda\delta$.
(2) If D is a lobe on the upper (lower) loop and $\nu_+(\xi) = a\sin(2\pi(\xi - \xi_+))$ (respectively $\nu_-(\xi) = a\sin(2\pi(\xi - \xi_-))$) with some constants a and $\xi_\pm$ then $d = \delta(a + o(1))$ and $\mathcal{A}_D = \lambda\delta(a/\pi + o(1))$ as $\varepsilon \to 0$.
(3) $\underline{\mathcal{A}} \le \mathcal{A} \le \overline{\mathcal{A}}$, where

$$\underline{\mathcal{A}} = 2|\underline{w}_0 \log \underline{w}_0^{-1} + O(\underline{w}_0)| + \sum_{*\in\{+,-\}} |\underline{w}_* \log \underline{w}_*^{-1} + O(\underline{w}_*)|,$$
$$\overline{\mathcal{A}} = 2|\overline{w}_0 \log \overline{w}_0^{-1} + O(\overline{w}_0)| + \sum_{*\in\{+,-\}} |\overline{w}_* \log \overline{w}_*^{-1} + O(\overline{w}_*)|.$$

A proof of this lemma is presented in Section 6.

Remark 4.3 *Assertion (2) can be easily generalized to the case of arbitrary 1-periodic functions $\nu_\pm$. Suppose for example that D is a lobe on the upper loop. The domain D is determined by two subsequent zeros $\xi_1 < \xi_2$ of ν. Then*

$$d = \delta\big(\max_{\xi_1\le\xi\le\xi_2} |\nu_+(\xi)| + o(1)\big), \quad \mathcal{A}_D = \lambda\delta\Big(\int_{\xi_1}^{\xi_2} |\nu_+(\xi)|\, d\xi + o(1)\Big) \quad \text{as } \varepsilon \to 0.$$

We omit the proof of this remark since it is almost the same as the proof of assertion (2) of Lemma lem:dA.

Theorem 1 *Suppose that the maps T_ε satisfy assumptions* **(A)**–**(C)**, *$\varepsilon > 0$ is sufficiently small, $0 < \lambda(\varepsilon) < \lambda_0$, and $\lim_{\varepsilon\to 0} \delta(\varepsilon)/\lambda^2(\varepsilon) = 0$. Then the following estimates hold:*

$$\begin{aligned} c_1\delta(\varepsilon)/\lambda(\varepsilon) \quad &\le \quad \underline{w}_\pm \le \overline{w}_\pm \le c_2\delta(\varepsilon)/\lambda(\varepsilon), \\ & \qquad \underline{w}_0 \le \overline{w}_0 \le c_2\delta(\varepsilon)/\lambda(\varepsilon) \end{aligned}$$

with some positive constants c_1 and c_2.

Moreover, if the family T_ε satisfies condition **(S)** *then*

$$\underline{w}_0 \geq c_1 \delta(\varepsilon)/\lambda(\varepsilon). \tag{4.5}$$

Theorem 1 is a corollary from Theorem 4 (Section 6).

Remark 4.4 *If the conditions* **(A)**–**(C)** *hold, we see that width of the stochastic layer measured by the function xy varies in the interval $[c_1\delta/\lambda, 2c_2\delta/\lambda]$. This implies the estimate (1.1) since according to Lemma 4.1, $d \sim \delta$.*

Conjecture 1 *The estimate (4.5) is valid without assumption* **(S)**.

Remark 4.5 *The area $\mathcal{A}$ of $\mathcal{SL}(\varepsilon)$ is a symplectic invariant. According to Lemma 4.1 and Theorem 1 for some positive constants $\tilde{c}_1, \tilde{c}_2$*

$$\frac{\tilde{c}_1\delta(\varepsilon)}{\lambda(\varepsilon)} \log \frac{\lambda(\varepsilon)}{\delta(\varepsilon)} \leq \mathcal{A} \leq \frac{\tilde{c}_2\delta(\varepsilon)}{\lambda(\varepsilon)} \log \frac{\lambda(\varepsilon)}{\delta(\varepsilon)}$$

Now consider the case $\lim_{\varepsilon\to 0} \lambda(\varepsilon) = 0$ in more detail. This case appears when separatrix splitting is exponentially small in ε. The following two specific properties of the corresponding separatrix map help us to analyze the situation. First, we have one more small parameter: λ. Secondly, studying of concrete problems (the standard map, pendulum with rapidly oscillating suspension point and many others, [6, 10, 11, 12, 15, 17, 18, 25, 26, 27]) shows that in this case in general the 1-periodic functions $\nu_\pm$ contain only the first harmonics, i.e.

$$\nu_\pm(\xi) = a_\pm \sin(2\pi(\xi - \xi_\pm)), \tag{4.6}$$

where $a_\pm \in \mathbf{R}$ and $\xi_\pm \in \mathbf{R}/\mathbf{Z}$ are constants. There are also some general theoretical argument which show that in the case in question one should expect that the higher Fourier harmonics of ν_+ can be neglected and equalities (4.6) hold. However, this has not been proved yet. If the separatrix splitting is not exponentially small, the functions $\nu_\pm$ can contain infinitely many Fourier harmonics (see Eq. (5.4) below).

Theorem 2 *Suppose that assumptions* **(A)**–**(C)** *hold, and*

$$\lim_{\varepsilon\to 0} \lambda(\varepsilon) = 0, \qquad \lim_{\varepsilon\to 0} \delta(\varepsilon)/\lambda^2(\varepsilon) = 0.$$

Then $\overline{w}_* - \underline{w}_* = O(\delta)$ for any $* \in \{+, -\}$.

Moreover, if condition **(S)** *holds, $\overline{w}_0 - \underline{w}_0 = O(\delta)$.*

Corollary 4.1 *Let all assumptions (including* **(S)***) of Theorem 2 be satisfied. Then according to Theorem 1, $\overline{w}_* \geq \underline{w}_* = O(\delta/\lambda)$, and we have the equalities $\overline{w}_*/\underline{w}_* = 1 + O(\lambda)$, $* \in \{+, -, 0\}$.*

To formulate another theorem we need some definition. Let $C_{a,\omega,\rho}$ be the following symplectic self-map of the cylinder $(\varphi \bmod 1, I)$:

$$C_{a,\omega,\rho}(\varphi, I) = (\psi, J) = \big(\varphi + \omega + J,\ I + \rho^{-1} a \sin(2\pi\varphi)\big).$$

The map $C_{1,0,\rho}$ is called standard. Usually instead of ρ the notation $k = 2\pi/\rho$ is used.

We call a closed curve on the cylinder rotational invariant (RI-curve) if it is homotopic to the one $\{I = 0\}$. We call an invariant with respect to $C_{a,\omega,\rho}$ closed curve hard if any close to $C_{a,\omega,\rho}$ analytic exact symplectic map has a close invariant curve. Now we define three functions

$$\overline{m}(a,\omega), \quad \underline{m}(a,\omega), \quad \overline{m}_0(a', a'', \omega', \omega''),$$

as follows:

$$\begin{aligned}
\overline{m} = \inf\{\rho_0 > 0 :\ & \text{for any } \rho \geq \rho_0 \\
& \text{the map } C_{a,\omega,\rho} \ \text{ has an invariant hard RI curve}\},\\
\underline{m} = \sup\{\rho_0 > 0 :\ & \text{for any } 0 < \rho < \rho_0 \\
& \text{the map } C_{a,\omega,\rho} \ \text{ has no invariant RI curve}\},\\
\overline{m}_0 = \inf\{\rho_0 > 0 :\ & \text{for any } \rho \geq \rho_0 \text{ the map} \\
& C_{a',\omega',\rho} \circ C_{a'',\omega'',\rho} \ \text{ has an invariant hard RI curve}\}.
\end{aligned}$$

Now let us present simple properties of these functions. We begin with $\underline{m}$ and $\overline{m}$. Since ω can be eliminated from $C_{a,\omega,\rho}$ by a change of the action, the functions $\underline{m}$ and $\overline{m}$ do not depend on ω. Moreover, obviously

$$\underline{m} = \underline{c}a, \qquad \overline{m} = \overline{c}a$$

for some constants $0 < \underline{c} \leq \overline{c}$.

The constant $\underline{c}$ was evaluated in a series of papers (see for example, [3, 13, 19, 20, 21, 23]). The sharpest estimates were obtained numerically [13, 19, 23]. In our notations the result is as follows:

$$\underline{c} = 2\pi/0.971635\ldots$$

Conjecture 2 *The constants $\underline{c}$ and $\overline{c}$ coincide.*

Numerical simulations confirm this conjecture.

The function $\overline{m}_0$ also depends on its arguments in some specific way.

Proposition 4.1 *(1)* $0 < \overline{m}_0 < \infty$.

(2) $\overline{m}_0(a, a, \omega, \omega) = \overline{c}a$.

(3) The function $\overline{m}_0$ can be presented in the form

$$\overline{m}_0(a', a'', \omega', \omega'') = a\overline{f}(\eta, \omega' - \omega''),$$
$$a' = a\cos(2\pi\eta), \quad a'' = a\sin(2\pi\eta),$$

where the function $\overline{f}$ is 1-periodic in both arguments and satisfies the identities

$$\overline{f}(\eta, \omega) = \overline{f}(-\eta, \omega) = \overline{f}(1/2 - \eta, \omega) = \overline{f}(1/4 - \eta, -\omega), \qquad \eta, \omega \in \mathbf{R}/\mathbf{Z}.$$

Proof of Proposition 4.1 as well as proofs of Theorems 2–3 (see below) are presented in Section sec:8.

Theorem 3 *Suppose that assumptions* **(A)**–**(C)** *hold,*

$$\lim_{\varepsilon\to 0} \lambda(\varepsilon) = 0, \qquad \lim_{\varepsilon\to 0} \delta(\varepsilon)/\lambda^2(\varepsilon) = 0,$$

and the equalities (4.6) are satisfied. Then the following assertions are valid.

a) The quantities $\underline{w}_\pm$ and $\overline{w}_\pm$ satisfy the inequalities

$$a_\pm\, \underline{c}(1 + o(1)) \le \lambda \cdot \underline{w}_\pm/\delta \le \lambda \cdot \overline{w}_\pm/\delta \le a_\pm\, \overline{c}(1 + o(1)). \tag{4.7}$$

b). In the case $\alpha_+ \ne \alpha_-$

$$\lambda \cdot \underline{w}_0/\delta \le \lambda \cdot \overline{w}_0/\delta \le \sup_{\omega', \omega''} \overline{m}_0(a_+, a_-, \omega', \omega'')(1 + o(1)).$$

c). If $\alpha_+ = \alpha_-$ then

$$\lambda \cdot \underline{w}_0/\delta \le \lambda \cdot \overline{w}_0/\delta \le \overline{m}_0(a_+, a_-, 2\xi_+, 2\xi_-)(1 + o(1)).$$

d). If condition **(S)** *holds then*

$$a_+\, \underline{c}(1 + o(1)) \le \lambda \cdot \underline{w}_0/\delta \le \lambda \cdot \overline{w}_0/\delta \le a_+\, \overline{c}(1 + o(1)). \tag{4.8}$$

Corollary 4.2 *If all assumptions (including* **(S)***) of Theorem 3 hold, upper $\overline{w}$ and lower $\underline{w}$ estimates for width of the stochastic layer*

$$\overline{w} = \overline{w}_+ + \overline{w}_0 = \overline{w}_- + \overline{w}_0, \qquad \underline{w} = \underline{w}_+ + \underline{w}_0 = \underline{w}_- + \underline{w}_0$$

satisfy the inequalities

$$2a_+\delta\underline{c}(1 + o(1)) \le \lambda\underline{w} \le \lambda\overline{w} \le 2a_+\delta\overline{c}(1 + o(1)). \tag{4.9}$$

Corollary 4.3 *Suppose that all assumptions (including* **(S)***) of Theorem 3 hold and Conjecture 2 is true. Then (4.7) and (4.8) give us asymptotic formulas for the quantities $\underline{w}_*, \overline{w}_*$, $* \in \{+, -, 0\}$:*

$$\lim_{\varepsilon\to 0} \left(\lambda w_*/\delta\right) = a_+\underline{c} = a_+\overline{c}. \tag{4.10}$$

The estimates (4.9)–(4.10) imply (1.2). Indeed, the quantities $w(\varepsilon)$ and $d(\varepsilon)$ measured in terms of xy are as follows:

$$w(\varepsilon) \approx \underline{w} \approx \overline{w}, \qquad d(\varepsilon) = \delta a_+ = \delta a_-.$$

The last two equalities follow from Lemma 4.1. Lemma 4.1 implies also the estimates presented in the following

Corollary 4.4 *Suppose that all assumptions (including* **(S)***) of Theorem 3 hold and Conjecture 2 is true. Then the area of the stochastic layer satisfies the relation*

$$\mathcal{A} = \frac{4\underline{c}a_+\delta(\varepsilon)}{\lambda(\varepsilon)} \log \delta^{-1}(\varepsilon)(1 + o(1)). \tag{4.11}$$

The area of a lobe equals

$$\mathcal{A}_D = \frac{a_+\delta(\varepsilon)\lambda(\varepsilon)}{\pi}(1 + o(1)) \tag{4.12}$$

The estimates (4.11)–(4.12) imply (1.4).

In the case $\alpha_+ \neq \alpha_-$ the quantities $\lambda\underline{w}_0/\delta$ and $\lambda\overline{w}_0/\delta$ oscillate between two distinct positive constants as $\varepsilon \to 0$.

5 Examples.

5.1 Polynomial splitting

Suppose that the family T_ε is smooth in ε for $|\varepsilon| \leq \varepsilon_0$ and $\lambda(0) > 0$. According to assumption **(B)** the map T_0 is integrable and separatrices of the hyperbolic fixed point $\hat{z}(0)$ form a figure eight curve. For small $\varepsilon \neq 0$ they in general split. We put $\delta(\varepsilon) = \varepsilon$. The functions $\lambda(\varepsilon), \kappa_\pm, f_\pm, g_\pm$ are smooth in ε and analytic in x, y for small values of ε and xy. Thus, results of Section 4 are applicable.

Now we show that the parameters $\alpha_\pm$ and the functions $\nu_\pm$ usually can be evaluated explicitly for concrete families T_ε. Indeed, let F be a first integral of the map T_0, $F(\hat{z}(0)) = 0$. We assume that the critical point $\hat{z}(0)$ of the function F is non-degenerate.

Let $\Gamma^\pm : \mathbf{R} \to \{z \in N : F(z) = 0\}$ be the higher and lower separatrix loops parameterized in a natural way:

$$T_0 \circ \Gamma^\pm(t) = \Gamma^\pm(t+1).$$

Such parameterizations are not defined uniquely. Actually, we use the natural parametrization defined below in Remark 5.1.

It is possible to introduce a symplectic coordinate system (q, p) in the vicinity of $\hat{z}(0)$ such that the curves $\Gamma^\pm(t)$ satisfy the relations

$$\begin{aligned} (q,p) &= (c_u^\pm \mu^t + O(\mu^{2t}), O(\mu^{2t})) \quad &\text{for} \quad t \to -\infty, \\ (q,p) &= (O(\mu^{-2t}), c_s^\pm \mu^{-t} + O(\mu^{-2t})) \quad &\text{for} \quad t \to +\infty, \end{aligned} \tag{5.1}$$

where $c_{s,u}^+ > 0$, $c_{s,u}^- < 0$, and $\mu = \mathcal{M}(0,0) = e^{\lambda(0)}$. These coordinates can be obtained by a linear change from any symplectic ones defined in a neighborhood of $\hat{z}(0)$. The constants $c_{s,u}^{\pm}$ are not defined uniquely but only one of them is independent.

Proposition 5.1 *(a). The quantities $c_s^+ c_u^+$, $c_s^- c_u^-$, c_s^+/c_s^- do not depend on choice of the coordinates (q,p).*

(b). The quantities $c_s^+ c_u^+$, $c_s^- c_u^-$ preserve after changes $t \to t + \hat{c}^{\pm}$ of the parameterizations of $\Gamma^{\pm}$.

(c). $\alpha_{\pm}^2 = c_s^{\pm} c_u^{\pm}$.

Proof of Proposition 5.1 and of the ones 5.2–5.3 (see below) is presented in Section 9.

Remark 5.1 *Given normal coordinates (x,y) there exist unique natural parameterizations on $\Gamma^{\pm}$ such that these curves are determined by the relations*

$$\begin{array}{lll} (i) & \Gamma^{\pm} & = (\pm\alpha_{\pm}\mu^t, 0), \\ (ii) & \Gamma^{\pm} & = (0, \pm\alpha_{\pm}\mu^{-t}). \end{array} \tag{5.2}$$

Below we use these parameterizations of the separatrices.

The functions $\nu_{\pm}$ can be obtained by using a discrete analog of the Poincaré–Melnikov integral. We put

$$w(T_0(z)) = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} T_\varepsilon(z), \qquad \nu_{\pm}^*(t) = \sum_{n=-\infty}^{\infty} dF \cdot w(\Gamma^{\pm}(t+n)),$$

where the expression $dF \cdot w(A)$ should be regarded as the application of the covector dF to the vector w at the point $A \in N$.

Since the critical point $\hat{z}(0)$ is non-degenerate, we can write

$$F(q,p) = F_0 qp + \text{ terms of at least third order.}$$

Proposition 5.2 *The following relations hold:*

$$\nu_{\pm}(t) = F_0^{-1} \nu_{\pm}^*(t - \lambda^{-1} \log \alpha_{\pm}).$$

An important particular case is the one when the maps T_ε in question are the Poincaré maps for a certain periodic in time Hamiltonian system with one and a half degrees of freedom. More precisely, let $H : N \times \mathbf{R} \times (-\varepsilon_0, \varepsilon_0) \to \mathbf{R}$ be a Hamiltonian function periodic in t:

$$H(z, t+\tau, \varepsilon) = H(z, t, \varepsilon)$$

for some $\tau > 0$. Assume that T_ε is the time τ shift along solutions of the corresponding Hamiltonian system, where initial conditions are taken at the time moment $t = 0$. Assume also that the function $H_0(z) = H(z, t, 0)$ does not depend on time. Then obviously, the functions $\Gamma^{\pm}(t/\tau)$ are separatrix solutions of the unperturbed system (the system with Hamiltonian H_0).

Proposition 5.3 *The functions $\nu^*_\pm$ can be presented in the form of Poincaré-Melnikov integral:*

$$\nu^*_\pm(t) = \int_{-\infty}^{+\infty} \{H', F\}(\Gamma^\pm(t + s/\tau), s)\, ds,$$

where $\{\,,\}$ is the ordinary Poisson bracket and

$$H'(x, y, t) = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} H(x, y, t, \varepsilon).$$

5.2 Stochastic layer for a pendulum

Consider as an example a pendulum with vertically periodically oscillating suspension point. Hamiltonian of the system is as follows:

$$H(\hat{q}, \hat{p}, t, \varepsilon) = \hat{p}^2/2 + \Omega^2 \cos\hat{q} + \varepsilon\theta(\omega t)\cos\hat{q}. \tag{5.3}$$

Here $\hat{q} = \hat{q} \bmod 2\pi$ is the angle between the pendulum and the vertical, $\hat{p}$ is the corresponding momentum, $\Omega > 0$ is the "interior frequency" of the system (Ω^2 equals gravity acceleration divided by the length of the pendulum), ω is the frequency of the suspension point oscillations and the parameter ε is proportional to the amplitude of the oscillations multiplied by ω^2. The form of the oscillations is determined by the 2π-periodic function θ. In this subsection we assume that ε is small and the other parameters in the system are of order 1.

The Poincaré map in this system has the hyperbolic fixed point $\hat{q} = \hat{p} = 0$. It is easy to calculate:

$$\mu(\varepsilon) = e^{2\pi\Omega/\omega} + O(\varepsilon), \quad \lambda(\varepsilon) = 2\pi\Omega/\omega + O(\varepsilon).$$

The separatrices $\Gamma^\pm$ have the form

$$(\hat{q}, \hat{p}) = (4 \arctan e^{\pm 2\pi\Omega t/\omega}, \pm 2\Omega/\cosh(2\pi\Omega t/\omega)).$$

The variables (q, p) (see (5.1)) are as follows:

$$q = \hat{q}\sqrt{\Omega/2} + \hat{p}/\sqrt{2\Omega}, \quad p = -\hat{q}\sqrt{\Omega/2} + \hat{p}/\sqrt{2\Omega}.$$

We see that $c_s^+ = c_u^+ = 4\sqrt{2\Omega}$ and $c_s^- = c_u^- = -4\sqrt{2\Omega}$. Therefore, according to Proposition 5.1, $\alpha_\pm^2 = 32\Omega$.

The first integral F has the form

$$F = H_0 = \hat{p}^2/2 + \Omega^2 \cos\hat{q} = \Omega qp + \text{ terms of third order.}$$

We obtain: $F_0 = \Omega$.

By using Proposition 5.3 we have:

$$\nu^*_\pm(t) = \int_{-\infty}^{+\infty} \theta(\omega s)\big(\hat{p}\sin\hat{q}\big)\big|_{\Gamma^\pm(t+\omega s/2\pi)}\, ds.$$

Direct calculation yields:

$$\nu_{\pm}^{*}(t) = -\sum_{m\in\mathbf{Z}\setminus\{0\}} \frac{2\pi i\omega^2}{\Omega^2} \frac{m^2\theta_m}{\sinh(\pi m\omega/(2\Omega))} e^{-2\pi imt}, \tag{5.4}$$

where the Fourier expansion of the function θ has the form

$$\theta(s) = \sum_{m\in\mathbf{Z}} \theta_m e^{ims}.$$

Note that the separatrix map in question satisfies condition **(S)** because the involution $(\hat{q},\hat{p}) \to (-\hat{q},-\hat{p})$ preserves the Hamiltonian (5.3). According to Theorem 1, width of the stochastic layer around "eight-like" separatrix structure of the point $\hat{q} = \hat{p} = 0$ is of order ε provided $\theta, \Omega, \omega \sim 1$ and θ is not constant. Under the same assumptions the area $\mathcal{A}$ of the stochastic layer is of order $\varepsilon \log \varepsilon^{-1}$.

5.3 Exponentially small splitting

Equality (5.4) is valid provided the frequency ratio ω/Ω is moderate. In the case $\omega/\Omega \sim \varepsilon^{-1}$ the standard Poincaré-Melnikov theory fails and it is necessary to apply a theory which can handle exponentially small effects. (Note, that formally (5.4) predicts exponentially small in ω/Ω value for $\nu_{\pm}$.)

The phenomenon of exponentially small separatrix splitting was discovered by Poincaré [24]. Later it was studied by Arnold [2] as a mechanism generating (Arnold) diffusion in Hamiltonian systems close to integrable ones. In physical papers [28, 5] this phenomenon was analyzed by means of the separatrix map. Intensive quantitative studying of exponentially small separatrix splitting was initiated by the papers: [15, 17] (see also [6]). Now there are essentially two parallel theories one can use to analyze the situation for broad variety of systems. The first one was constructed by Lazutkin with collaborators [10, 11, 12, 17, 18]. The second theory begins to form in [25, 26, 27]. The main idea of [25, 26, 27] is to reduce considerably the perturbation by a carefully chosen coordinate change. Actually, in certain coordinates the Hamiltonian (5.3) is autonomous up to an exponentially small in ω/Ω term, playing the role of a perturbation.

More precisely, let us consider the system with Hamiltonian (5.3) under the following assumptions:

$$\Omega = 1, \quad \omega = 1/\varepsilon, \quad \theta(s) = 2\varepsilon^{-1} B \cos s.$$

The larger multiplier of the hyperbolic fixed point $\hat{q} = \hat{p} = 0$ of the corresponding Poincaré map is $\mu = e^{2\pi\varepsilon}$.

The separatrix map, generated by this Poincaré map, satisfies condition **(S)**. In [26] (see also [25]) we have proved that there exists a symplectic change of coordinates $\hat{q}, \hat{p} \to \tilde{q}, \tilde{p}$ close to the identity, periodic in time and analytic in a complex neighborhood of the separatrices $\Gamma^\pm$ of the system with the Hamiltonian $H_0 = \hat{p}^2/2 + \cos\hat{q}$ such that the new Hamiltonian function takes the form

$$\tilde{H}(\tilde{q},\tilde{p},t,\varepsilon) = H_0(\tilde{q},\tilde{p}) + \varepsilon H_1(\tilde{q},\tilde{p},\varepsilon) + \exp(-\tilde{c}/\varepsilon)\tilde{H}_2(\tilde{q},\tilde{p},t,\varepsilon),$$

where $\tilde{c} \in [0,\pi/2)$ is arbitrary, the functions H_1, H_2 are analytic in $\tilde{q},\tilde{p}$ in the vicinity of $\Gamma^\pm$, smooth in $\varepsilon > 0$ and the function $\tilde{H}_2$ is analytic and 2π-periodic in t.

The ordinary Poincaré-Melnikov theory applied to this system for positive $\tilde{c}$ gives a correct asymptotics of the separatrix splitting. (Actually we are able to prove this only for $\tilde{c} > \pi/4$, [25, 26].) In particular we obtain:

$$\nu^*_\pm(t) = -2\pi\varepsilon^{-2}e^{-\pi\varepsilon^{-1}/2}Bf(B^2)\sin(2\pi t),$$

where f is an entire real-analytic function, $f(0) = 2$. In [25, 26, 27] this function was investigated numerically. It turned out that $f(z) = \sum_0^\infty f_n z^n$, where $f_0 = 2$ and

$$\begin{array}{lllllll} f_1 & = & 0.65856738\ldots, & & f_2 & = & 6.651741\ldots\cdot 10^{-2}, \\ f_3 & = & 3.21010\ldots\cdot 10^{-3}, & & f_4 & = & 9.03367\ldots\cdot 10^{-5}, \\ f_5 & = & 1.6620\ldots\cdot 10^{-6}, & & f_6 & = & 2.1534\ldots\cdot 10^{-8}, \\ f_7 & = & 2.070\ldots\cdot 10^{-10}, & & f_8 & = & 1.53\ldots\cdot 10^{-12}. \end{array}$$

According to Theorem 3 width of the stochastic layer near the separatrices $\Gamma^\pm$ satisfies the estimate (4.9) with

$$a_+\delta = 2\pi\varepsilon^{-2}e^{-\pi\varepsilon^{-1}/2}Bf(B^2), \quad \lambda = 2\pi\varepsilon.$$

Area of this layer can be estimated as follows:

$$\frac{2\pi\underline{c}}{\varepsilon^4}e^{-\pi\varepsilon^{-1}/2}Bf(B^2)(1+o(1)) \le \mathcal{A} \le \frac{2\pi\overline{c}}{\varepsilon^4}e^{-\pi\varepsilon^{-1}/2}Bf(B^2)(1+o(1)).$$

Recall that according to Conjecture 2 we have $\underline{c} = \overline{c}$.

6 Separatrix map

In the vicinity of the separatrices the maps T_ε generate the so-called separatrix maps S_ε. To define the last ones we identify any two points z_1 and z_2 of N whose normal coordinates (x_1, y_1) and (x_2, y_2) satisfy the relation

$$x_2 = \mathcal{M}^n(x_1y_1,\varepsilon)x_1, \quad y_2 = \mathcal{M}^{-n}(x_1y_1,\varepsilon)y_1, \qquad n \in \mathbf{Z}. \tag{6.1}$$

(The parameter ε is assumed fixed.) After this identification the maps S_ε are determined by the gluing maps. Below we assume that the argument y in the gluing maps is small and x is of order one. By using assumptions **(A)–(C)** and relations (4.1)–(4.2), we can write:

$$U_\varepsilon^\pm(x,y) = \Big(\alpha_\pm^{-2}x(xy+\delta\hat\nu_\pm(\log|x|/\log\mathcal{M}(0,\varepsilon),\varepsilon)),\ \alpha_\pm^2/x\Big) + \big(O(|y|+\delta)^2, O(|y|+\delta)\big). \tag{6.2}$$

Putting formally $\delta = O(y) = 0$ we obtain an integrable map (the function xy is an analytic first integral). Let us introduce the action-angle variables for the corresponding separatrix map. We put

$$I = \lambda\cdot\delta^{-1}xy, \quad \varphi = \lambda^{-1}\log|x|. \tag{6.3}$$

It is easy to check that $dI\wedge d\varphi = \delta^{-1}dy\wedge dx$. In the variables (I,φ) the identification (6.1) takes the form $\varphi = \varphi \bmod 1$.

We restore the information about sign of the variable x (lost in (6.3)) by adding the sign $+$ or $-$ to the coordinate system (I,φ) according to the value of $\mathrm{sign}(x)$. In the variables (I,φ,σ), $\sigma\in\{+,-\}$ the maps S_ε have the form

$$\begin{aligned} &S_\varepsilon(I,\varphi,\sigma) = (J,\psi,\rho),\\ J &= I + \lambda\cdot\hat\nu_\sigma(\varphi,\varepsilon) + \delta\lambda\cdot O(1+|I|/\lambda)^2,\\ \psi &= \varphi + \frac{1}{\lambda}\Big[\log\frac{\delta}{\alpha_\sigma^2\lambda} + \log|J| + \delta\cdot O(1+|I|/\lambda)^2\Big],\\ \rho &= \sigma\,\mathrm{sign}(J). \end{aligned} \tag{6.4}$$

Indeed, let $U_\varepsilon^\sigma(x,y) = (X,Y)$. According to (6.2), we get:

$$\begin{aligned} XY &= xy + \delta\hat\nu_\sigma(\varphi,\varepsilon) + O(|y|+\delta)^2,\\ \frac{\log|X|}{\lambda} &= \frac{1}{\lambda}\Big[-\log|Y| + \log\frac{\delta}{\lambda} + \log|\lambda\delta^{-1}XY|\Big]\\ &= \frac{1}{\lambda}\Big[\log|x| + \log\frac{\delta}{\lambda\alpha_\sigma^2} + \log|\lambda\delta^{-1}XY| + O(|y|+\delta)\Big]. \end{aligned}$$

These equalities together with (6.3) imply (6.4). Here we have used also the estimate $O(y) = O(\delta I/\lambda)$. Note that formulas similar to (6.4) are contained in [8].

In the main approximation the dependence of the separatrix map on $\log\delta$ is periodic. Indeed, for any integer n the change $\delta\to e^{n\lambda}\delta$ preserves the map

$$\begin{aligned} J &= I + \lambda\cdot\hat\nu_\sigma(\varphi,\varepsilon),\\ \psi &= \varphi + \frac{1}{\lambda}\Big[\log\frac{\delta}{\alpha_\sigma^2\lambda} + \log|J|\Big],\\ \rho &= \sigma\,\mathrm{sign}(J). \end{aligned}$$

Earlier this property was pointed out in [29]

Note that for large values of the action variable the separatrix map is close to integrable. Indeed, let us put $I = I_0(1+\lambda u)$ and $J = I_0(1+\lambda v)$, where $|I_0|$ is large. The separatrix map turns out to be close to the following one:

$$v = u, \quad \psi = \varphi + \frac{1}{\lambda}\Big[\log\Big|\frac{I_0\delta}{\alpha_\sigma^2\lambda}\Big| + \log(1+\lambda v)\Big], \quad \rho = \sigma\,\mathrm{sign}(I_0).$$

Proof of Lemma 4.1. The separatrix map is not defined on the stable separatrix and its inverse is not defined on the unstable one. According to equalities (6.4) the unstable separatrix has the form $\{I = 0\}$ and the stable one the form $\{J = J(I,\varphi) = 0\}$. If δ is small, the last equation is reduced to the one

$$I = -\lambda(\hat\nu_\sigma(\varphi,\varepsilon) + O(\delta)) = -\lambda(\nu_\sigma(\varphi) + o(1)) \quad \text{as } \varepsilon \to 0.$$

Therefore, homoclinic points (i.e. the points which belong to both stable and unstable separatrix) correspond to zeros of the functions $\nu_\pm(\varphi)$. A lobe can be determined by two neighboring zeros φ_1, φ_2 of ν_σ.

Let φ_1 and φ_2 define the lobe D. Sign of ν_σ is constant on the interval (φ_1, φ_2). According to (4.4) and (6.3) we obtain the formula for the width of the lobe D:

$$d = \delta(\max_{\varphi_1\le\varphi\le\varphi_2} \nu_\sigma(\varphi) + o(1)) \quad \text{as } \varepsilon \to 0. \tag{6.5}$$

The area $\mathcal{A}_D$ is defined as absolute value of $\int_D dy\wedge dx = \delta\int_D dI\wedge d\varphi$. Therefore we have:

$$\mathcal{A}_D = \delta\lambda\Big|\int_{\varphi_1}^{\varphi_2} \nu_\sigma(\varphi)\,d\varphi + o(1)\Big|. \tag{6.6}$$

Equalities (6.5)–(6.6) imply assertions (1)–(2) of the lemma.

Now we prove assertion (3). In normal coordinates x, y the stochastic layer occupies the domain $\mathcal{SL}$ such that $\underline{\mathcal{D}} \subset \mathcal{SL} \subset \overline{\mathcal{D}}$, where

$$\underline{\mathcal{D}} \;=\; \Big\{(x,y) : x^2+y^2 < \underline{r}^2, \quad |xy| < \begin{cases} \underline{w}_+, & \text{if } x>0 \text{ and } y>0,\\ \underline{w}_-, & \text{if } x<0 \text{ and } y<0,\\ \underline{w}_0, & \text{if } xy<0, \end{cases} \Big\},$$

$$\overline{\mathcal{D}} \;=\; \Big\{(x,y) : x^2+y^2 < \overline{r}^2\}, \quad |xy| < \begin{cases} \overline{w}_+, & \text{if } x>0 \text{ and } y>0,\\ \overline{w}_-, & \text{if } x<0 \text{ and } y<0,\\ \overline{w}_0, & \text{if } xy<0. \end{cases} \Big\},$$

and $0 < \underline{r} \le \overline{r}$ are constants. By using the inequalities

$$\mathrm{area}(\underline{\mathcal{D}}) \ge \underline{\mathcal{A}}, \quad \mathrm{area}(\overline{\mathcal{D}}) \le \overline{\mathcal{A}},$$

we obtain the required estimate.

7 The boundary of the stochastic layer

Theorem 4 *Suppose that $\varepsilon > 0$ is sufficiently small, $0 < \lambda < \lambda_0$, and*

$$\lim_{\varepsilon \to 0} \delta(\varepsilon)/\lambda^2(\varepsilon) = 0.$$

Then there are two constants $0 < I_{\min} < I_{\max}$ (depending only on the functions $\nu_\pm$) such that the following conditions hold.

(1) The inequality $|I| < I_{\max}$ is satisfied on the invariant curves $\gamma_\pm, \gamma_0$.

(2) The curves γ_+ and γ_- do not intersect with the domain $|I| \le I_{\min}$.

(3) The estimate $I_{\max} - I_{\min} < 2I_{\min}\lambda e^{2\lambda}$ holds.

(4) If condition **(S)** *is valid, the curve γ_0 also does not intersect with the domain $|I| \le I_{\min}$.*

Remark 7.1 *By using equalities (6.3) we can rewrite the estimates presented in Theorem 4 in terms of xy instead of I. These estimates imply Theorems 1 and 2.*

Proof of Theorem 4.

1. Recall that the curves γ_+, γ_- and γ_0 lie in the domains $I > 0$, $\sigma > 0$; $I > 0$, $\sigma < 0$; and $I < 0$ respectively.

Introducing new variables

$$I = I_0(1 + u\lambda), \quad J = I_0(1 + v\lambda), \qquad I_0 = \text{const.}$$

we rewrite the equations (6.4) as follows:

$$\begin{aligned} v &= u + I_0^{-1}\hat{\nu}_\sigma(\varphi, \varepsilon) + I_0^{-1}\delta O(1 + |I|/\lambda)^2, \\ \psi &= \varphi + \omega_\sigma + \lambda^{-1}\log|1 + v\lambda| + \delta\lambda^{-1}\, O(1 + |I|/\lambda)^2, \\ \rho &= \sigma \cdot \operatorname{sign} I_0 \cdot \operatorname{sign}(1 + v\lambda), \end{aligned} \tag{7.1}$$

where the parameters ω_σ have the form

$$\omega_\sigma = \frac{1}{\lambda}\log\frac{|I_0|\delta}{\alpha_\sigma^2\lambda}.$$

The most important parameters in the system (7.1) are δ and I_0^{-1}. The first one is small. Its smallness is equivalent to smallness of ε. The parameter I_0 (the "unperturbed" action) is responsible for type of the dynamics (regular or chaotic).

Putting formally $I_0^{-1} = \delta\lambda \cdot O(1 + I/\lambda) = 0$ we obtain the integrable map

$$\begin{aligned} v &= u, \\ \psi &= \varphi + \omega_\sigma + \lambda^{-1}\log|1 + v\lambda|, \\ \rho &= \sigma \operatorname{sign} I_0. \end{aligned}$$

This map is non-degenerate because $\partial\psi(u,\varphi)/\partial u = (1+u\lambda)^{-1} \neq 0$.

Suppose that the following inequalities hold:

$$\hat{c}_0^{-1} \leq |I_0^{-1}| \leq 1, \quad \delta \leq \hat{c}_2^{-1} \ll \frac{1}{\lambda(1+\hat{c}_0/\lambda)^2}, \quad c_0 \gg 1.$$

These relations mean that in the map (7.1) the terms proportional to $\delta O(1+I/\lambda)^2$ are negligible. If the parameter $|I_0^{-1}|$ is small enough (being nevertheless greater than $\hat{c}_0^{-1}$), the map is close to integrable. According to the KAM theory the separatrix map (7.1) has an invariant curve $u = u(\varphi) = u_0 + O(I_0^{-1})$, where the value u_0 is such that the "unperturbed" frequency $\omega_\sigma + \lambda^{-1}\log|1+u\lambda|$ is Diophantine. (Obviously, we can take $|u_0| < 1/\lambda$.) Recall that for $I_0 > 0$ we obtain invariant curves homotopic to γ_σ and for $I_0 < 0$ the invariant curve is homotopic to γ_0. This proves assertion (1) of Theorem 4 since we can put $I_{\max} = \max_\varphi |I_0|(1+u(\varphi)\lambda)$.

2. We prove assertion (2) only for the curve γ_+. Arguments in the case of γ_- are analogous. First, we note that (6.4) is a twist map for small values of $\varepsilon > 0$. This means that $\partial\psi(I,\varphi)/\partial I > 0$. Following ideas of Lazutkin [16], we use one Birkhoff's result [4] (see also [14]):

Lemma 7.1 *Let γ be a continuous invariant curve of a twist area-preserving map. If γ is homotopic to a curve $\{I = \mathrm{const}\}$ then γ is a graph of a Lipschitz continuous function $I = r(\varphi)$.*

Lipschitz constant of the function $r(\varphi)$ can be estimated as follows. Let $Q = \partial(J,\psi)/\partial(I,\varphi)$ be the Jacobi matrix of the map T_ε. Since T_ε is a twist map, there exist constants $\eta_1, \eta_2 > 0$ such that for any point (I,φ) of the curve the inequalities

$$(0,1)Q\begin{pmatrix} \eta_1 \\ -1 \end{pmatrix} > 0, \quad (0,1)Q^{-1}\begin{pmatrix} \eta_2 \\ 1 \end{pmatrix} < 0$$

hold. (Note that any $\eta_1' \geq \eta_1$ and $\eta_2' \geq \eta_2$ share properties of η_1, η_2.) The following lemma in fact is contained in [14].

Lemma 7.2 *For any two points φ_1, φ_2 such that $0 \leq \varphi_2 - \varphi_1 \leq 1/2$ the function $r(\varphi)$ satisfies the inequality*

$$-\eta_1(\varphi_2 - \varphi_1) < r(\varphi_2) - r(\varphi_1) < \eta_2(\varphi_2 - \varphi_1). \tag{7.2}$$

Remark 7.2 *Inequalities of type $0 \leq \varphi_2 - \varphi_1 \leq 1/2$, $\varphi_1, \varphi_2 \in \mathbf{R}/\mathbf{Z}$ should be understood in the sense that it is possible to find two points $\tilde{\varphi}_1$, $\tilde{\varphi}_2$ on the universal cover $\mathbf{R}$ of the circle $\mathbf{R}/\mathbf{Z}$ such that the natural projection maps them into φ_1, φ_2 respectively and $0 \leq \tilde{\varphi}_2 - \tilde{\varphi}_1 \leq 1/2$. Below we do not refer to the universal cover and hope that there will be no misunderstanding.*

Lemma 7.2 is a corollary of Lemma 7.1. Indeed, assuming that at least one of inequalities (7.2) does not hold for some pair $\varphi_1 < \varphi_2$, where φ_1 and φ_2 are close to one another, we obtain that for the images $(r(\psi_1), \psi_1)$, $(r(\psi_2), \psi_2)$ of the points $(r(\varphi_1), \varphi_1)$, $(r(\varphi_2), \varphi_2)$ with respect to the map T_ε (or T_ε^{-1}) the inequality $\psi_2 < \psi_1$ holds.

Let the curve γ_+ have the parameterization $(I, \varphi) = (r_+(\varphi), \varphi)$. Below we show that

$$r_{\max} = \max_{\varphi} r_+(\varphi)$$

is estimated from below by a positive constant (Proposition 7.1) and then that $r_+(\varphi) > r_{\max} e^{-2\lambda}$ (Proposition 7.2). Propositions 7.1–7.2 obviously imply assertion (2) of Theorem 4 for the curve γ_+. Assertion (4) of Theorem 4 can be proved in the same way due to condition **(S)**. Assertion (3) also follows from Proposition 7.2. Indeed, we have the inequality $I_{\max} < I_{\min} e^{2\lambda}$. By using the simple relation $e^{2\lambda} < 1 + 2\lambda e^{2\lambda}$, we obtain the desired estimate.

The function r_+ is positive, 1-periodic and Lipschitz continuous. We estimate the Lipschitz constant for small values of ε by using Lemma 7.2. First, we put $\varepsilon = 0$ (this obviously implies the equality $\delta = 0$). The Jacobi matrix Q takes the form

$$Q = \begin{pmatrix} 1, & \lambda \nu_+'(\varphi) \\ (J\lambda)^{-1}, & 1 + J^{-1}\nu_+'(\varphi) \end{pmatrix},$$

where $J = r_+(\varphi) + \lambda\, \nu_+'(\varphi) = r_+(\psi)$ and $\nu_+' = d\nu_+/d\varphi$. Direct calculations yield the following estimate: $\eta_2 \ge r_{\max}\lambda$.

For small nonzero values of δ the estimate is slightly weaker. We can put:

$$\eta_2 = 2r_{\max}\lambda. \tag{7.3}$$

According to assumption **(C)** (see Section 4) the function ν_+ is analytic and non-constant. We put $\nu'_{\min} = \min_\varphi d\nu_+/d\varphi$. Obviously, $\nu'_{\min} < 0$.

Proposition 7.1 *Suppose that $\varepsilon > 0$ is sufficiently small. Then*

$$r_{\max} \ge -\nu'_{\min}/7. \tag{7.4}$$

Proof of Proposition 7.1. We put again for simplicity $\varepsilon = 0$. Let the map (6.4) transform the points $(r_+(\varphi_j), \varphi_j)$, $j = 1, 2$ to the ones $(r_+(\psi_j), \psi_j)$.

We assume that the points φ_1, φ_2 are close to one another and $\varphi_1 < \varphi_2$. The points ψ_1, ψ_2 are also close to one another and $\psi_1 < \psi_2$. By using the equalities (6.4), we have:

$$\Delta\psi = \Delta\varphi + \frac{1}{\lambda} \log\Big(1 + \frac{\Delta r_+ + \lambda \cdot \Delta\nu_+}{r_+(\varphi_1) + \lambda \cdot \nu_+(\varphi_1)}\Big), \tag{7.5}$$

where

$$\Delta\psi = \psi_2 - \psi_1, \qquad \Delta\varphi = \varphi_2 - \varphi_1 > 0,$$
$$\Delta r_+ = r_+(\varphi_2) - r_+(\varphi_1), \qquad \Delta\nu_+ = \nu_+(\varphi_2) - \nu_+(\varphi_1).$$

We take φ_1 such that $\nu'_+(\varphi_1) = \nu'_{\min}$. For small positive $\Delta\varphi$ we have the estimates:

$$\Delta r_+ \le \eta_2 \Delta\varphi < 2r_{\max}\lambda \cdot \Delta\varphi, \qquad \Delta\nu_+ \le \nu'_{\min}\Delta\varphi/2 < 0.$$

Here we have used relations (7.2)–(7.3). Now suppose that $r_{\max} < -\nu'_{\min}/6$. Then the equality (7.5) implies the estimate

$$\begin{aligned}\Delta\psi &\le \Delta\varphi + \frac{1}{\lambda}\log\Big(1 + \frac{2r_{\max} + \nu'_{\min}/2}{r_{\max}}\lambda\Delta\varphi\Big)\\ &< \Delta\varphi + \frac{(2r_{\max} + \nu'_{\min}/2)\Delta\varphi}{r_{\max}} = (3r_{\max} + \nu'_{\min}/2)\frac{\Delta\varphi}{r_{\max}} < 0.\end{aligned}$$

This relation contradicts to the inequality $\Delta\psi > 0$. For small values of ε the estimates preserve (with possibly slightly modified constants). Because of this we set in (6.8) $\nu'_{\min}/7$ instead of $\nu'_{\min}/6$. Proposition 7.1 is proved.

Proposition 7.2 *For any $\varphi \in \mathbf{R}/\mathbf{Z}$ the inequality $r_{\max}/r_+(\varphi) < e^{2\lambda}$ holds.*

Proof of Proposition 7.2. We put $\varepsilon = 0$ and assume that there exists a point $\hat\varphi \in \mathbf{R}/\mathbf{Z}$ such that $r_{\max} \ge r_+(\hat\varphi)e^{\lambda}$. It is possible to assume that $0 < \hat\varphi - \varphi_0 < 1$, where the point φ_0 satisfies the equality $r_+(\varphi_0) = r_{\max}$. Take a pair $\theta_0, \hat\theta \in \mathbf{R}/\mathbf{Z}$ such that the map $(6.4)|_{\varepsilon=0}$ transforms the points $(r_+(\theta_0), \theta_0)$, $(r_+(\hat\theta), \hat\theta)$ respectively into the ones $(r_+(\varphi_0), \varphi_0)$, $(r_+(\hat\varphi), \hat\varphi)$.

Now we are going to show that $\hat\theta - \theta_0 > 1$ in the sense that image of the curve

$$\{(I, \varphi) : I = r_+(s), \quad \varphi = s, \quad \hat\theta \le s \le \theta_0\}$$

covers all the circle being projected to $\mathbf{R}/\mathbf{Z}$ in the natural way: $(I, \varphi) \to \varphi$. By proving this fact we obtain a contradiction with Lemma 7.1 and consequently, prove Proposition 7.2.

We estimate the quantity $\hat\theta - \theta_0$ by using the equalities $(6.4)|_{\varepsilon=0}$:

$$\hat\varphi - \varphi_0 = \hat\theta - \theta_0 + \frac{1}{\lambda}\log\frac{r_+(\hat\varphi)}{r_+(\varphi_0)}.$$

Therefore, we have:

$$\hat\theta - \theta_0 = \hat\varphi - \varphi_0 - \frac{1}{\lambda}\log\frac{r_+(\hat\varphi)}{r_{\max}} > \hat\varphi - \varphi_0 + 1 > 1$$

according to our assumption. In the case of small nonzero ε the estimates are weaker. Because of this we assume in Proposition 7.2 that $r_{\max}/r_+(\varphi) < e^{2\lambda}$. Proposition 7.2 is proved.

8 The case of small $\lambda(\varepsilon)$.

8.1 Proof of Proposition 4.1

(1) The inequality $\overline{m}_0 > 0$ follows from KAM-theory. Finiteness of the function $\overline{m}_0$ is also a standard fact (see for example [20]).

(2) This assertion follows from the equalities $\overline{m}_0(a, a, \omega, \omega) = \overline{m}(a, \omega)$.

(3) The following identities hold for any real $\beta > 0, a', a'', \omega', \omega'', \omega$,

$$\begin{aligned} \overline{m}_0(a', a'', \omega', \omega'') &= \overline{m}_0(a', a'', \omega' + \omega, \omega'' + \omega), && (8.1) \\ \overline{m}_0(\beta a', \beta a'', \omega', \omega'') &= \beta \overline{m}_0(a', a'', \omega', \omega''), && (8.2) \\ \overline{m}_0(-a', a'', \omega' + \pi, \omega'' - \pi) &= \overline{m}_0(a', a'', \omega', \omega'') \\ &= \overline{m}_0(a', -a'', \omega' - \pi, \omega'' + \pi), && (8.3) \\ \overline{m}_0(a'', a', \omega'', \omega') &= \overline{m}_0(a', a'', \omega', \omega''). && (8.4) \end{aligned}$$

Indeed, the map $C_{a',\omega',\rho} \circ C_{a'',\omega'',\rho}$ has the form $(u, \varphi) \to (w, \theta)$, where

$$\left\{ \begin{array}{l} w = v + \rho^{-1} a' \sin\psi, \\ \theta = \psi + \omega' + w, \end{array} \right. \qquad \left\{ \begin{array}{l} v = u + \rho^{-1} a'' \sin\varphi, \\ \psi = \varphi + \omega'' + v. \end{array} \right.$$

Since $C_{a',\omega'+\omega,\rho} \circ C_{a'',\omega''+\omega,\rho}$ can be obtained from $C_{a',\omega',\rho} \circ C_{a'',\omega'',\rho}$ by the change $u, v, w \to u + \omega, v + \omega, w + \omega$, we get (8.1). The identity (8.2) follows from the equality

$$C_{\beta a',\omega',\rho} \circ C_{\beta a'',\omega'',\rho} = C_{a',\omega',\beta^{-1}\rho} \circ C_{a'',\omega'',\beta^{-1}\rho}.$$

Since $C_{-a',\omega'+\pi,\rho} \circ C_{a'',\omega''-\pi,\rho}$ can be obtained from $C_{a',\omega',\rho} \circ C_{a'',\omega'',\rho}$ by the change $\varphi, \theta \to \varphi - \pi, \theta - \pi$, we get (8.3). The equality (8.4) follows from conjugacy of the maps $C_{a',\omega',\rho} \circ C_{a'',\omega'',\rho}$ and $C_{a'',\omega'',\rho} \circ C_{a',\omega',\rho}$.

Identities (8.1)–(8.4) obviously imply assertion (3) of Proposition 4.1.

8.2 Proof of Theorem 3

According to the equalities (4.6) the map (7.1) takes the form

$$\begin{aligned} v &= u + I_0^{-1} a_\sigma \sin(2\pi\tilde{\varphi}) + I_0^{-1} \delta O(1 + \delta/\lambda)^2, \\ \tilde{\psi} &= \tilde{\varphi} + \omega_\sigma + \xi_\sigma - \xi_\rho + v + I_0^{-1} \delta O(1 + \delta/\lambda)^2, \qquad (8.5) \\ \rho &= \sigma \cdot \operatorname{sign} I_0, \qquad \tilde{\varphi} = \varphi - \xi_\sigma, \quad \tilde{\psi} = \psi - \xi_\sigma. \end{aligned}$$

Properties of the separatrix map presented in Theorem 3 can be easily obtained with the help of equalities (8.5). Assertions a) and b) follow from definition of the functions $\underline{m}, \overline{m}, \overline{m}_0$. Assertion c) is based on Proposition 4.1 and the equality $\omega_+ = \omega_-$ for $\alpha_+ = \alpha_-$. Assertion d) is reduced to a) due to the symmetry condition **(S)**.

9 Appendix

9.1 Proof of Proposition 5.1

Let (q', p') be another coordinate system satisfying relations (5.1) (in general with other constants $c^{\pm}_{s,u}$). It easy to show that

$$q' = qc' + O(q^2 + p^2), \quad p' = p/c' + O(q^2 + p^2)$$

with some positive c'. These equalities imply assertion (a) of Proposition 5.1.

Changing parameterization of Γ^+ (or Γ^-) we multiply c^+_s (respectively, c^-_s) by a positive constant and divide c^+_u (respectively, c^-_u) by the same constant. This implies assertion (b).

To prove assertion (c) we note that normal coordinates (x, y) also satisfy relations of type (5.1). Moreover, the curves $\Gamma^{\pm}$ have the parameterizations

$$\begin{array}{lccl} (i) & \Gamma^{\pm} & = & (c^{\pm}_u \mu^t, 0), \\ (ii) & \Gamma^{\pm} & = & (0,\, c^{\pm}_s \mu^{-t}). \end{array}$$

The gluing maps $U^{\pm}_0$ transform parameterizations (i) into the ones (ii):

$$(0,\, c^{\pm}_s \mu^{-t}) = U^{\pm}_0(c^{\pm}_u \mu^t, 0) = \big(0,\, \alpha^2_{\pm}/(c^{\pm}_u \mu^t)\big).$$

These equalities imply assertion (c) of Proposition 5.1.

9.2 Proof of Proposition 5.2

Let A_ε be a smooth curve on N. Then the following useful equality holds:

$$dF \cdot \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} T^n_\varepsilon(A_\varepsilon) - dF \cdot \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} A_\varepsilon = \operatorname{sign} n \cdot \sum_k dF \cdot w(T^k_0(A_0)), \tag{9.1}$$

where $k \in \{0, 1, \ldots, n-1\}$ if $n > 0$, and $k \in \{-1, -2, \ldots, n\}$ if $n < 0$. This formula can be obtained by a direct calculation with the help of the relation

$$dF \cdot (dT^m_0 v) = dF \cdot v, \qquad v \text{ is a vector field on } N.$$

The last relation can be obtained as a result of differentiation of the identity $F \circ T^m_0 = F$ along the vector field v.

Now suppose that the point A^{s+}_ε (respectively, A^{u+}_ε) lies on the higher stable separatrix Γ^{s+}_ε (respectively, on the higher unstable one Γ^{u+}_ε) for any $|\varepsilon| < \varepsilon_0$ and $A^{u+}_0 = A^{s+}_0 = A_0$. Note that $\Gamma^{s+}_0 = \Gamma^{u+}_0 = \Gamma^+$. By using relation (9.1), we have for any natural n:

$$\begin{aligned} dF \cdot \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} A^{s+}_\varepsilon &= dF \cdot \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} T^n_\varepsilon(A^{s+}_\varepsilon) - \sum_{k=0}^{n-1} dF \cdot w(T^k_0(A_0)), \\ dF \cdot \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} A^{u+}_\varepsilon &= dF \cdot \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} T^{-n}_\varepsilon(A^{u+}_\varepsilon) + \sum_{k=-n}^{-1} dF \cdot w(T^k_0(A_0)). \end{aligned}$$

Hence,

$$dF \cdot \Big(\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} A_\varepsilon^{u+} - \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} (A_\varepsilon^{s+})\Big) = \sum_{k=-\infty}^{+\infty} dF \cdot w(T_0^k(A_0)). \tag{9.2}$$

Here we have used the relation

$$\lim_{n\to\infty} \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} T_\varepsilon^n(A_\varepsilon^{s+}) = \lim_{n\to\infty} \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} T_\varepsilon^n(A_\varepsilon^{u+}).$$

Now let us calculate the left hand side of (9.2) in another way. In the normal coordinates $(x, y) = (x(\varepsilon), y(\varepsilon))$ the curves A_ε^{s+} and A_ε^{u+} have the following form:

$$\begin{array}{rcll} A_\varepsilon^{u+} &=& (j_u(\varepsilon), 0), & \quad j_u(\varepsilon) > 0, \\ A_\varepsilon^{s+} &=& (0, j_s(\varepsilon)), & \quad j_s(\varepsilon) > 0. \end{array}$$

We can present the curve A_ε^{u+} in another way. By using explicit form of U_ε^+, we get:

$$A_\varepsilon^{u+} = U_\varepsilon^+(j_u(\varepsilon), 0) = \Big(\varepsilon j_u(0) f_+(j_u(0), 0, 0) + O(\varepsilon^2),\ \alpha_+^2/j_u(0) + O(\varepsilon)\Big).$$

Since $A_0^{s+} = A_0^{u+}$, we have: $j_u(0) j_s(0) = \alpha_+^2$.

The normal coordinates (x, y) depend on ε. We must take into account this fact when evaluating in these coordinates any of the vectors $dA_\varepsilon^{u,s+}/d\varepsilon$. Nevertheless, we can forget about this when evaluating the difference $\frac{d}{d\varepsilon}\big|_{\varepsilon=0} A_\varepsilon^{u+} - \frac{d}{d\varepsilon}\big|_{\varepsilon=0} A_\varepsilon^{s+}$. The difference equals

$$\begin{aligned} &\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} \big(\varepsilon j_u(0) f_+(j_u(0), 0, 0) + O(\varepsilon^2), \quad \alpha_+^2/j_u(0) + O(\varepsilon) - j_s(\varepsilon)\big) \\ &\qquad\qquad = (j_u(0) f_+(j_u(0), 0, 0),\ \hat{j}), \end{aligned} \tag{9.3}$$

where $\hat{j}$ is some constant.

The first integral F satisfies the relation

$$F(x, y) = F_0 xy + O(\varepsilon) + O(x^2 y^2).$$

Hence, according to (9.3),

$$\begin{aligned} dF \cdot \Big(\frac{d}{d\varepsilon}\Big|_{\varepsilon=0} A_\varepsilon^{u+} - \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} A_\varepsilon^{s+}\Big) &= F_0 \cdot (j_s(0), 0) \cdot (j_u(0) f_+(j_u(0), 0, 0),\ \hat{j}) \\ &= F_0\, j_s(0)\, j_u(0)\, f_+(j_u(0), 0, 0) \\ &= F_0\, \nu_+(\log j_u(0)/\lambda). \end{aligned} \tag{9.4}$$

We take $A_0 = \Gamma^+(t)$. Then

$$T_0^m(A_0) = \Gamma^+(t + m), \qquad m \in \mathbf{Z}. \tag{9.5}$$

According to (5.2) we have:

$$j_u(0) = \alpha_+ \mu^t, \quad j_s(0) = \alpha_+ \mu^{-t}. \tag{9.6}$$

By using the relations (9.2)–(9.6) we get:

$$F_0\, \nu_+(\log \alpha_+/\lambda + t) = \sum_{k=-\infty}^{+\infty} dF \cdot w(\Gamma^+(t+k)) = \nu_+^*(t).$$

Analogously we obtain the equality $F_0\, \nu_-(\log \alpha_-/\lambda + t) = \nu_-^*(t)$.

9.3 Proof of Proposition 5.3

Let $z_\varepsilon(t) = z_0(t) + \varepsilon z'(t) + O(\varepsilon^2)$ be the solution of the system with Hamiltonian H and with the initial conditions $z_\varepsilon(0) = (x, y)$. Then, obviously,

$$T_\varepsilon(x, y) = z_\varepsilon(\tau). \tag{9.7}$$

Below we assume that the point (x, y) does not depend on ε. Then differentiating (9.7) with respect to ε at the point $\varepsilon = 0$, we get:

$$w(T_0(x, y)) = z'(\tau). \tag{9.8}$$

The first integral F satisfies the equality $\{H_0, F\} = 0$. Differentiating the identity $\frac{d}{dt} F(z_\varepsilon(t)) = \{H, F\}(z_\varepsilon(t), t, \varepsilon)$ with respect to ε and putting $\varepsilon = 0$, we obtain:

$$\frac{d}{dt}\Big(dF(z_0(t)) \cdot z'(t)\Big) = \{H', F\}(z_0(t)). \tag{9.9}$$

Now let us put $z_0(t) = \Gamma^+(t_0 + t/\tau)$. Integrating (9.9), we get:

$$\begin{aligned} &dF(\Gamma^+(t_0+1)) \cdot z'(\tau) - dF(\Gamma^+(t_0)) \cdot z'(0) \\ &= \int_0^\tau \{H', F\}(\Gamma^+(t_0 + t/\tau), t)\, dt. \end{aligned}$$

By using the equality (9.8) and the one $z'(0) = 0$, we obtain:

$$dF \cdot w(\Gamma^+(t_0+1)) = \int_0^\tau \{H', F\}(\Gamma^+(t_0 + s/\tau), s)\, ds.$$

Analogously, it is possible to get the corresponding equality for Γ^-. These equalities together with definition of the function ν_+^* imply Proposition 5.3.

Acknowledgments. We thank Ya. G. Sinai who attracted our attention to the problem considered in the paper. We are grateful to S. V. Bolotin, V. V. Kozlov and A. I. Neishtadt for fruitful discussions. We estimate highly the work of the referees whose criticism helped to improve readability of the paper and to correct misprints.

The work was partially supported by Russian Foundation of Basic Research (Grant 96-01-00747) and by INTAS (Grant 93-339-ext).

References

[1] T. Ahn, G. Kim and S. Kim, Analysis of the separatrix map in Hamiltonian systems, Physica **89D**, 1996, p. 315–328

[2] V. Arnold Instability of dynamical systems with several degrees of freedom, Dokl. Acad. Nauk SSSR, v. 156, n. 1, pp. 9–12, (1964) (in Russian), English transl. in Sov. Math. Dokl., v. 5, pp. 581–589, (1964)

[3] S. Aubry, The twist map, the extended Frenkel-Kontorova model and the Devil's staircase, Physica **7D**, 1983, p. 240–258

[4] G. D. Birkhoff, Surface transformations and their dynamical applications, Acta Mathematicae, 1920, v. 43

[5] B. V. Chirikov, A universal instability of many-dimensional oscillation systems, Phys. Rep. 1979, v. 52, no. 5, p. 263–379

[6] A. Delshams and T. M. Seara, An asymptotic Expressions for the Splitting of Separatrices of the Rapidly Forced Pendulum, Comm. Math. Phys. 1992, v. 150, p. 433–463

[7] S. A. Dovbysh, Structure of the Kolmogorov set near separatrices of a plane map, Math. Zametki, 1989, v. 46, no. 4, p. 112–114 (in Russian)

[8] S. A. Dovbysh, Intersection of asymptotic surfaces in the perturbed Euler-Poinsot problem, Prikl. Mat. Mekh. 1987, v. 51, no. 3, p. 363–370 (in Russian)

[9] N. N. Filonenko, R. Z. Sagdeev and G. M. Zaslavsky, Destruction of magnetic surfaces by magnetic field irregularities, Nuclear Fusion 1967, v. 7, p. 253–266

[10] V. G. Gelfreich, Reference systems for splitting of separatrices, Nonlinearity, V. 10 (1997) P. 175–193.

[11] V. G. Gelfreich, V. F. Lazutkin and N. V. Svanidze, Refined formula to separatrix splitting for the standard map, Physica D, 1994, v. 71, no. 2, p. 101

[12] V. G. Gelfreich, V. F. Lazutkin and M. B. Tabanov, Exponentially small splitting in Hamiltonian systems, Chaos 1991, v. 1, no. 2, p. 137–142

[13] J. M. Greene, A method for determining a stochastic transition, J. Math. Phys., 1979, v. 20, p. 1183–1201

[14] M. R. Hermann, Sur les courbes invariantes par les difféomorphismes de l'anneau Astérisque: 1983, v. 103–104

[15] P. Holmes, J. Marsden and J. Scheurle, Exponentially small splittings of separatrices with applications to KAM theory and degenerate bifurcations, Contemp. Math. 1988, v. 81, p. 213–244

[16] V. F. Lazutkin, On the width of a stability zone near separatrices of the standard map, Dokl. Akad. Nauk SSSR 1990, v. 313, no. 2, p. 268–272 (in Russian)

[17] V. F. Lazutkin, Splitting of separatrices for standard Chirikov's mapping, VINITI no. 6372-84, 24 Sept. 1984 (in Russian)

[18] V. F. Lazutkin, I. G. Schachmannski and M. B. Tabanov, Splitting of separatrices for standard and semistandard mappings, Physica D 1989, v. 40, p. 235–248

[19] R. S. MacKay, A renormalization approach to invariant circles in area-preserving maps Physica **7D** 1983, p. 283–300

[20] R. S. MacKay and I. C. Percival, Converse KAM: Theory and practice. Commun. Math. Phys. 1985, v. 98, p. 469–512

[21] J. N. Mather, Non-existence of invariant circles, Erg. Theory Dyn. Sys. 1984, v. 4, p. 301–311

[22] J. Moser, The analytic invariants of an area preserving mapping near hyperbolic fixed point, Commun. Pure Appl. Math. 1956, v. 9, p. 673–692

[23] A. Olvera and C. Simó, An obstruction method for the destruction of invariant curves Physica **26D**, 1987, p. 181–192

[24] H. Poincaré Les méthodes nouvelles de la mécanique celeste, V. 1–3,: Gauthier-Villars, Paris, 1892, 1893, 1899.

[25] D. Treschev, Separatrix Splitting for a Pendulum with Rapidly Oscillating Suspension Point, Preprint no. 273, Centre de Recerca Matematica, Institute D'Estudis Catalans, Barcelona, Spain, 1994

[26] D. Treschev, Separatrix Splitting for a Pendulum with Rapidly Oscillating Suspension Point, Russian J. of Math. Phys. , 1997, V. 5, No. 1, P. 63–98.

[27] D. Treschev, An averaging method for Hamiltonian systems, exponentially close to integrable ones, Chaos 1996, v. 6, no. 1, p. 6–14

[28] G. M. Zaslavsky and N. N. Filonenko, Stochastic instability of trapped particles and the conditions of applicability of the quasilinear approximation, Zh. Eksp. Teor. Fiz. 1968, v. 54, p. 1590–1602, (in Russian)

[29] G. M. Zaslavsky and S. S. Abdullaev, Scaling properties and anomalous transport of particles inside the stochastic layer. Phys. Review E, 1995, V. 51, No. 5, P. 3901–3910.