On the phase change for perturbations of Hamiltonian systems (non-parametric case) *

Anatoly Neishtadt and Alexey Okunev

Abstract

We consider perturbations of Hamiltonian systems with one degree of freedom such that the evolution leads to separatrix crossings. Such crossings are described by a parameter called the pseudo-phase. We prove a formula for the dependence of the pseudo-phase on the initial conditions.

1 Introduction

We start with a Hamiltonian system with one degree of freedom

$$\dot{q} = \frac{\partial H}{\partial p}, \ \dot{p} = -\frac{\partial H}{\partial q} \tag{1.1}$$

with the Hamiltonian $H(p,q)$. We assume that H has a saddle C with two separatix loops l_1 and l_2 forming a figure eight. We also assume $H(C) = 0$, $H > 0$ outside the loops and $H < 0$ inside each loop.

Then we add a small perturbation εf:

$$\begin{aligned} \dot{q} &= \frac{\partial H}{\partial p} + \varepsilon f_q(p,q,\varepsilon), \\ \dot{p} &= -\frac{\partial H}{\partial q} + \varepsilon f_p(p,q,\varepsilon). \end{aligned} \tag{1.2}$$

We assume that H is analytic and f is C^2. We use that H is analytic to apply the local normal form [3] in a neighborhood of C. Denote by $f_h(p,q,\varepsilon) = f_q \frac{\partial H}{\partial q} + f_p \frac{\partial H}{\partial p}$ the rate of change of H divided by ε. For $i = 1,\ 2$ denote $\Theta_i = -\oint_{l_i} f_h(p(t), q(t), 0) dt$ (here t is the time for the unperturbed system). Let $\Theta_3 = \Theta_1 + \Theta_2$. We assume that $\Theta_1, \Theta_2 > 0$, then $\Theta_3 > 0$.

As $\Theta_3 > 0$, the trajectories of the perturbed system starting close to the figure eight outside it eventually approach the separatrices of the unperturbed system. We study the phase change for such trajectories. Formulas describing such phase change were obtained (using the averaging method) in [2] for Hamiltonian systems with one degree of freedom and slow time dependence; in [6] for slow-fast Hamiltonian systems with one degree of freedom corresponding to slow motion. In [1] the authors use the averaging method to compute the phase change for perturbed strongly nonlinear oscillators. Unlike [2] and [6], they do not provide an estimate for the accuracy of using the averaging method, but instead check that the result compares well with numerical experiments.

A parameter called the pseudo-phase (we use the terminology from [6]) describes the phase at the moment of separatrix crossing. We show that a formula for the pseudo-phase similar to the formula from [6] also holds for our case, this is done in Section 6.

The general plan of the proof is close to the one in [6]. However, instead of the improved adiabatic invariant considered in [6] we consider the averaged system of order 2. An important part of our paper is obtaining estimates for the coefficients of this system (in particular, proving that solutions of this system cross the separatrices of the unperturbed Hamiltonian system).

2 Energy-angle variables

Let us consider the action-angle variables I, φ; $\varphi \in [0, 2\pi)$ We will assume that $\varphi = 0$ corresponds to a specific transversal Γ that is chosen in Section 9.1. It will

*The work was supported by the Leverhulme Trust (Grant No. RPG-2018-143).

	Estimates	Obtained in
T	$T = O(ln(h)),\ \frac{\partial T}{\partial h} = O(h^{-1}),\ \frac{\partial^2 T}{\partial h^2} = O(h^{-2})$	Section 9.3
ω	$\omega = O(ln^{-1} h),\ \frac{\partial \omega}{\partial h} = O(h^{-1}\ln^{-2} h),\ \frac{\partial^2 \omega}{\partial h^2} = O(h^{-2}\ln^{-2} h)$	Section 9.3
f_h	$f_h = O(1),\ \frac{\partial f_h}{\partial h} = O(h^{-1}),\ \frac{\partial f_h}{\partial \varphi} = O(\ln h),$ $\frac{\partial^2 f_h}{\partial h^2} = O(h^{-2}),\ \frac{\partial^2 f_h}{\partial h \partial \varphi} = O(h^{-1}\ln h),\ \frac{\partial^2 f_h}{\partial \varphi^2} = O(\ln^2 h)$	Section 9.5
f_φ	$f_\varphi, \frac{\partial f_\varphi}{\partial \varphi} = O(h^{-1});\ f_\varphi(h, 0) = O(h^{-1/2}\ln^{-1} h);$ $\frac{\partial f_\varphi}{\partial h} = O(h^{-2})$	Section 9.5
$u_{h,1}$	$u_{h,1} = O(1); \frac{\partial u_{h,1}}{\partial \varphi} = O(\ln h);$ $\frac{\partial u_{h,1}}{\partial h}, \frac{\partial^2 u_{h,1}}{\partial h \partial \varphi} = O(h^{-1}\ln h); \frac{\partial^2 u_{h,1}}{\partial h^2} = O(h^{-2}\ln h)$	Section 10
$u_{\varphi,1}$	$u_{\varphi,1}, \frac{\partial u_{\varphi,1}}{\partial \varphi} = O(h^{-1}\ln h); \frac{\partial u_{\varphi,1}}{\partial h} = O(h^{-2}\ln h)$	Section 10
$\overline{f}_{h,1}$	$\overline{f}_{h,1} = O(\ln^{-1} h),\ \frac{\partial \overline{f}_{h,1}}{\partial h} = O(h^{-1}\ln^{-2} h)$	Section 10
$\overline{f}_{\varphi,1}$	$\overline{f}_{\varphi,1} = O(h^{-1}\ln^{-3} h),\ \frac{\partial \overline{f}_{\varphi,1}}{\partial h} = O(h^{-2})$	Sections 9.6, 10
$u_{h,2}$	$u_{h,2}, \frac{\partial u_{h,2}}{\partial \varphi} = O(h^{-1}\ln^3 h); \frac{\partial u_{h,2}}{\partial h} = O(h^{-2}\ln^3 h)$	Section 10
$\overline{f}_{h,2}$	$\overline{f}_{h,2} = O(\ln^{-1} h),\ \frac{\partial \overline{f}_{h,2}}{\partial h} = O(h^{-1})$	Section 10
$\overline{f}_{h,3}$	$\overline{f}_{h,3} = O(h^{-2}\ln^4 h)$ for $h > \varepsilon \ln^{10} \varepsilon$.	Section 10
$\overline{f}_{\varphi,2}$	$\overline{f}_{\varphi,2} = O(h^{-2}\ln h)$ for $h > \varepsilon \ln^{10} \varepsilon$.	Section 10
$\hat{f}_{*,*}$	For $\hat{f}_{h,1}, \hat{f}_{h,2}, \hat{f}_{h,3}, \hat{f}_{\varphi,1}, \hat{f}_{\varphi,2}$ and derivatives: same as for the corresponding expression with overline instead of hat.	Section 10

Table 1: estimates. Note that the estimates for functions that depend on ε, e.g. f_*, $u_{*,*}$ and $\overline{f}_{*,*}$, are uniform in ε.

be tangent to the bisector of the angle between the separatrices. We will assume that the separatrices are numbered in such way that for a trajectory outside the figure eight close to the separatrices it is close to l_2 for $0 < \varphi < \pi$ and to l_1 for $\pi < \varphi < 2\pi$.

Denote by h the value of the Hamiltonian. We will always assume $h > 0$, as we study a trajectory approaching the separatrices from the outside. We will use the "energy-angle" variables h, φ. In these variables the unperturbed system (1.1) is written as $\dot{h} = 0, \dot{\varphi} = \omega(h)$. Denote by $T(h) = \frac{2\pi}{\omega(h)}$ the period of the unperturbed system. We will sometimes use the time t passed from the last crossing of the transversal $\varphi = 0$ instead of φ. We have $t = \frac{\varphi T}{2\pi}$.

Denote by f_h, f_φ the components of f in the energy-angle variables: $f_y = f_q \frac{\partial y}{\partial q} + f_p \frac{\partial y}{\partial p}$ for $y = h, \varphi,\ y = y(q, p)$. Then the perturbed system (1.2) is written as

$$\begin{aligned} \dot{h} &= \varepsilon f_h(h, \varphi, \varepsilon), \\ \dot{\varphi} &= \omega(h) + \varepsilon f_\varphi(h, \varphi, \varepsilon). \end{aligned} \tag{2.1}$$

Let us state a useful relation between the derivatives of the components of f.

Lemma 2.1.

$$\frac{\partial f_h}{\partial h} + \frac{\partial f_\varphi}{\partial \varphi} + \frac{1}{T}\frac{dT}{dh} f_h = \operatorname{div}(f), \textit{ where } \operatorname{div}(f) = \frac{\partial f_q}{\partial q} + \frac{\partial f_p}{\partial p} \tag{2.2}$$

Proof. Let us first prove that $\frac{\partial f_q}{\partial q} + \frac{\partial f_p}{\partial p} = \frac{\partial f_I}{\partial I} + \frac{\partial f_\varphi}{\partial \varphi}$. Here f_I, f_φ is the vector field f written in the action-angle variables.

Recall that the divergence of a vector field v with respect to a volume form α is a function $\operatorname{div}_\alpha(v)$ such that $\mathcal{L}_v(\alpha) = \operatorname{div}_\alpha(v) \cdot \alpha$ (here $\mathcal{L}$ denotes the Lie derivative). In the coordinates x, y for the euclidean volume form $dx \wedge dy$ we have $\operatorname{div}_{dx\wedge dy}(v) = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y}$.

Hence the equality rewrites as $\operatorname{div}_{dp\wedge dq}(f) = \operatorname{div}_{dI\wedge d\varphi}(f)$. But since I, φ are the action-angle variables, $dp \wedge dq = dI \wedge d\varphi$.

Finally, using that $\frac{\partial h}{\partial I} = \omega(h)$ (this follows from the Hamiltonian equations in the action-angles variables) and $f_I = \frac{\partial I}{\partial h} f_h$, we can compute that $\frac{\partial f_I}{\partial I} = \frac{\partial f_h}{\partial h} + \frac{1}{T}\frac{\partial T}{\partial h} f_h$. □

3 Averaging chart

Remark 3.1. *The formulas in this section are also valid for the parametric case where H depends on a parameter z. For the parametric case we can set h to be the column vector (h, z).*

We start with the system (2.1). Let us find a change of variables

$$\begin{aligned} h &= \overline{h} + \varepsilon u_{h,1}(\overline{h}, \overline{\varphi}, \varepsilon) + \varepsilon^2 u_{h,2}(\overline{h}, \overline{\varphi}, \varepsilon), \\ \varphi &= \overline{\varphi} + \varepsilon u_{\varphi,1}(\overline{h}, \overline{\varphi}, \varepsilon) \end{aligned} \tag{3.1}$$

that transforms (2.1) to the following form:

$$\begin{aligned} \dot{\overline{h}} &= \varepsilon \overline{f}_{h,1}(\overline{h}, \varepsilon) + \varepsilon^2 \overline{f}_{h,2}(\overline{h}, \varepsilon) + \varepsilon^3 \overline{f}_{h,3}(\overline{h}, \overline{\varphi}, \varepsilon), \\ \dot{\overline{\varphi}} &= \omega(\overline{h}) + \varepsilon \overline{f}_{\varphi,1}(\overline{h}, \varepsilon) + \varepsilon^2 \overline{f}_{\varphi,2}(\overline{h}, \overline{\varphi}, \varepsilon). \end{aligned} \tag{3.2}$$

Let us call the new chart $\overline{h}, \overline{\varphi}$ the *averaging chart.* For brevity we will often omit the dependence of the functions f_*, $\overline{f}_{*,*}$ and $u_{*,*}$ on ε.

Lemma 3.1. *For $k = h,\ i = 1, 2$ and for $k = \varphi,\ i = 1$ we have*

$$\overline{f}_{k,i}(h) = \langle Y_{k,i}(h, \varphi) \rangle_\varphi, \tag{3.3}$$

$$\overline{f}_{k,i}(h) + \omega(h) \frac{\partial u_{k,i}}{\partial \varphi}(h, \varphi) = Y_{k,i}(h, \varphi), \tag{3.4}$$

where

$$\begin{aligned} Y_{h,1} &= f_h, \\ Y_{\varphi,1} &= f_\varphi + \frac{\partial \omega}{\partial h} u_{h,1}, \\ Y_{h,2} &= \frac{\partial f_h}{\partial h} u_{h,1} + \frac{\partial f_h}{\partial \varphi} u_{\varphi,1} - \frac{\partial u_{h,1}}{\partial h} \overline{f}_{h,1} - \frac{\partial u_{h,1}}{\partial \varphi} \overline{f}_{\varphi,1}. \end{aligned} \tag{3.5}$$

The formulas for $\overline{f}_{h,3}$ and $\overline{f}_{\varphi,2}$ are stated in Lemma 8.1 below.

We will prove this lemma in Section 8. The formulas above uniquely define $\overline{f}_{k,i}$ and $u_{k,i}$ under an additional assumption that for $k = h,\ i = 1, 2$ and for $k = \varphi,\ i = 1$ we have

$$\langle u_{k,i} \rangle_\varphi = 0.$$

We will always assume this to hold.

For $h \to 0$ many expressions introduced above tend to infinity. We will use the estimates given in Table 1, these estimates will be proved below.

Remark 3.2. *For h large compared with ε (i.e. $h > \varepsilon \ln^{10} \varepsilon$) the derivative of the coordinate change given by (3.1) is close to identity (by (3.1) and Table 1). This means that for such h this coordinate change is invertible.*

Since $\langle \frac{\partial u_{k,i}}{\partial \varphi} \rangle_\varphi = 0$, we have

$$\overline{f}_{k,i} = \langle Y_{k,i} \rangle_\varphi.$$

Using that $\langle \frac{\partial u_{k,i}}{\partial \varphi} \rangle_\varphi = 0$, $\langle \frac{\partial u_{k,i}}{\partial h} \rangle_\varphi = \frac{\partial}{\partial h} \langle u_{k,i} \rangle_\varphi = 0$, we can simplify this for $\overline{f}_{h,2}$:

$$\overline{f}_{h,2} = \langle \frac{\partial f_h}{\partial h} u_{h,1} + \frac{\partial f_h}{\partial \varphi} u_{\varphi,1} \rangle_\varphi. \tag{3.6}$$

Using Lemma 2.1, we can prove another formula for $\overline{f}_{h,2}$ that gives a better estimate for $h \to 0$.

Lemma 3.2.

$$\overline{f}_{h,2} = \langle \operatorname{div}(f) u_{h,1} \rangle_\varphi, \text{ where } \operatorname{div}(f) = \frac{\partial f_q}{\partial q} + \frac{\partial f_p}{\partial p}. \tag{3.7}$$

This lemma is proved in Section 7.

The following formula is similar to formula 2 from [5]. We will prove it in Section 7.

Lemma 3.3.

$$u_{h,1}(h, t_0) = \frac{1}{T} \int_0^T \left(t - \frac{T}{2}\right) f_h(h, t + t_0) dt. \tag{3.8}$$

Here the second argument in f_h is not φ, as usual, but the time t. We use the notation $f_h(h, t) = f_h(h, \varphi(h, t))$.

This can also be rewritten as follows:

$$u_{h,1}(h, t_0) = \frac{1}{2\pi} \int_0^T (\varphi(t) - \pi) f_h(h, t + t_0) dt. \tag{3.9}$$

4 Averaged system of order 2

The coefficients of the the initial system (3.2) in the averaging chart depend on ε. We would like the coefficients of the averaged system that we define in this section to be independent of ε. To this end, let us introduce some notation. First, let us expand

$$f(p, q, \varepsilon) = f^0(p, q) + \varepsilon f^1(p, q) + \varepsilon^2 f^2(p, q, \varepsilon), \tag{4.1}$$

where $f^0(p, q) = f(p, q, 0)$ and $f^1(p, q) = \frac{\partial f}{\partial \varepsilon}(p, q, 0)$. Clearly, f^0_q, f^0_p, f^1_q and f^1_p are smooth functions of p and q. The functions f^2_p and f^2_q are uniformly bounded by some constant independent of ε (by Taylor's theorem with the Lagrange form of remainder). Let us also consider the perturbed system (2.1) with the petrurbation $\varepsilon f^0(h, \varphi)$ instead of $\varepsilon f(h, \varphi, \varepsilon)$. For such system we may also consider a coordinate change of form (3.1) that transforms it to the form (3.2). Let us add an upper index 0 to the coefficients of these equations (e.g. $u^0_{h,1}$, $\overline{f}^0_{\varphi,1}$) to show that we started with the perturbation εf^0. The coefficients $u^0_{*,*}$ and $\overline{f}^0_{*,*}$ are determined by the same formulas as $u_{*,*}$ and $\overline{f}_{*,*}$, but we should plug f^0 instead of f into those formulas.

Now let us rewrite (3.2) in such way that only the coefficients next to the largest powers of ε depend on ε. This is done simply by expanding the coefficients similarly to (4.1). The resulting system will be

$$\begin{aligned} \dot{\overline{h}} &= \varepsilon \hat{f}_{h,1}(\overline{h}) + \varepsilon^2 \hat{f}_{h,2}(\overline{h}) + \varepsilon^3 \hat{f}_{h,3}(\overline{h}, \overline{\varphi}, \varepsilon), \\ \dot{\overline{\varphi}} &= \omega(\overline{h}) + \varepsilon \hat{f}_{\varphi,1}(\overline{h}) + \varepsilon^2 \hat{f}_{\varphi,2}(\overline{h}, \overline{\varphi}, \varepsilon), \end{aligned} \tag{4.2}$$

where

$$\hat{f}_{h,1} = \overline{f}^0_{h,1}, \ \hat{f}_{\varphi,1} = \overline{f}^0_{\varphi,1}, \ \hat{f}_{h,2} = \overline{f}^0_{h,2} + \langle f^1_h(h, \varphi) \rangle_\varphi \tag{4.3}$$

(here f^1_h is the h-component of f^1 written in (h, φ) coordinates), and $\hat{f}_{\varphi,2}$ and $\hat{f}_{h,3}$ satisfy the estimates in Table 9.3. The estimates for $\hat{f}_{\star,\star}$ will be proved in Lemma 10.2 below, one can also find formulas for $\hat{f}_{\varphi,2}, \hat{f}_{h,3}$ there. Also note that by [4, Corollary 3.1] we have $\int_0^T f^0_h dt = -\Theta_3 + O(h \ln h)$, so we have

$$\hat{f}_{h,1} = \frac{-\Theta_3 + O(h \ln h)}{T}. \tag{4.4}$$

The *averaged system of order* 2 is obtained from the system (4.2) by removing all terms on the right hand side that depend on φ:

$$\begin{aligned} \dot{\hat{h}} &= \varepsilon \hat{f}_{h,1}(\hat{h}) + \varepsilon^2 \hat{f}_{h,2}(\hat{h}) \\ \dot{\hat{\varphi}} &= \omega(\hat{h}) + \varepsilon \omega_1(\hat{h}). \end{aligned} \tag{4.5}$$

Here we denote $\omega_1(\hat{h}) = \hat{f}_{\varphi,1}(\hat{h})$ in order to match with [6]. We will sometimes call this system simply the *averaged system*.

Let us introduce the slow time $\tau = \varepsilon t$. Then the first equation in (4.5) can be written as follows:

$$\frac{\partial \hat{h}}{\partial \tau} = \hat{f}_{h,1}(\hat{h}) + \varepsilon \hat{f}_{h,2}(\hat{h}). \tag{4.6}$$

By the estimates on $\hat{f}_{h,1}$ and $\hat{f}_{h,2}$ from Table 1 we get that

$$\frac{\partial \hat{h}}{\partial \tau} = \frac{-\Theta_3 + O(\hat{h} \ln \hat{h}) + O(\varepsilon)}{T}. \tag{4.7}$$

As $\Theta_3 > 0$, this means that for small ε any solution $\hat{h}(\tau), \hat{\varphi}(\tau)$ of the averaged system of order 2 starting close to the separatrices crosses the separatrix of the initial unperturbed Hamiltonian system. Denote by τ_* the slow time at the moment of crossing, $\hat{h}(\tau_*) = 0$. From (4.7) we also see that for small ε, h and $\tau < \tau_*$ the function $\hat{h}(\tau)$ is decreasing. By (4.7) we also have

$$\frac{\partial \tau}{\partial \hat{h}} = -\frac{T}{\Theta_3}(1 + O(\hat{h} \ln \hat{h}) + O(\varepsilon)). \tag{4.8}$$

Let us prove that the solution of the averaged system (4.5) approximates the solution of (3.2) for h separated from zero:

Lemma 4.1. *Consider a solution $\overline{h}(\tau), \overline{\varphi}(\tau)$ of (3.2) with initial conditions $\overline{h}(0), \overline{\varphi}(0)$. Consider also a solution $\hat{h}(\tau)$ of (4.6) with initial condition $\hat{h}(0)$ such that $|\overline{h}(0) - \hat{h}(0)| \le C\varepsilon^2$ for some C. Then for small enough ε for any τ such that*

$$\hat{h}(\tau) > \varepsilon \ln^{10} \varepsilon \tag{4.9}$$

we have the following estimates (in the error terms below we write h for $\hat{h}(\tau)$, e.g. $O(h)$ instead of $O(\hat{h}(\tau))$):

$$\overline{h}(\tau) - \hat{h}(\tau) = O(\varepsilon^2 h^{-1} \ln^5 h),$$
$$\overline{\varphi}(\tau) - \overline{\varphi}(0) = \varepsilon^{-1} \int_0^\tau \Big(\omega(\hat{h}(\tau')) + \varepsilon\omega_1(\hat{h}(\tau'))\Big) d\tau' + O(\varepsilon h^{-1} \ln^4 h).$$

Proof. Let us start with the estimate for $\overline{h}(\tau) - \hat{h}(\tau)$. We will first only consider the system up to some moment τ_{fin} such that for all $\tau < \tau_{fin}$ we have

$$0.5\hat{h}(\tau) < \overline{h}(\tau) \le 2\hat{h}(\tau). \tag{4.10}$$

In order to receive a better estimate, let us switch from h to the action I. From the Hamiltonian equations in I, φ-chart we have $\frac{\partial h}{\partial I} = \omega$. Denote $\overline{I} = I(\overline{h})$, $\hat{I} = I(\hat{h})$. Denote $\hat{f}_{I,j} = \omega^{-1} \hat{f}_{h,j}$, $j = 1, 2, 3$. As $\overline{h}, \overline{\varphi}$ is a solution of (3.2), it is also a solution of (4.2). From (4.2), (4.5) we have

$$\begin{aligned} \dot{\overline{I}} &= \varepsilon \hat{f}_{I,1}(\overline{h}) + \varepsilon^2 \hat{f}_{I,2}(\overline{h}) + \varepsilon^3 \hat{f}_{I,3}(\overline{h}, \overline{\varphi}, \varepsilon), \\ \dot{\hat{I}} &= \varepsilon \hat{f}_{I,1}(\hat{h}) + \varepsilon^2 \hat{f}_{I,2}(\hat{h}). \end{aligned} \tag{4.11}$$

As $\hat{f}_{I,1} = (2\pi)^{-1} \int_0^T f_h^0 dt$, the estimate $\frac{\partial}{\partial h}(\int_0^T f_h^0 dt) = O(\ln h)$ ([4, Lemma 3.2]) yields $\frac{\partial \hat{f}_{I,1}}{\partial h} = O(\ln h)$.

Denote $\Delta = |\overline{I}(\tau) - \hat{I}(\tau)|$. From (4.11) we have the following differential inequality for Δ:

$$\dot{\Delta} \le \varepsilon a(\tau)\Delta + \varepsilon^3 b(\tau), \tag{4.12}$$

where $a(\tau) = |\frac{\partial \hat{f}_{I,1}}{\partial I}(\xi) + \varepsilon \frac{\partial \hat{f}_{I,2}}{\partial I}(\xi)|$ for some $\xi \in [\overline{I}(\tau), \hat{I}(\tau)]$ and $b(\tau) = |\hat{f}_{I,3}(\overline{I}(\tau), \overline{\varphi}(\tau))|$. We have $\varepsilon \frac{\partial \hat{f}_{I,2}}{\partial I} = \varepsilon\omega \frac{\partial}{\partial h}(\omega^{-1} \hat{f}_{I,2}) = o(1)$ by (4.10), (4.9) and Table 1. Hence, we have the estimate $a(\tau) = O(1)$. By Table 1 we have $b(\tau) = O(h^{-2} \ln^5 h)$.

As in [6], we use the following estimate for Δ obtained by solving (4.12):

$$\Delta(\tau) \le \exp\left(\int_0^\tau a(\tau')d\tau'\right)\left(\Delta(0) + \varepsilon^2 \int_0^\tau b(\tau')d\tau'\right).$$

Using (4.8) and the estimates for a and b, we can make a change of variable and compute the integrals above as integrals $d\hat{h}$:

$$\int_0^\tau a(\tau')d\tau' = O(1); \qquad \int_0^\tau b(\tau')d\tau' = O(h^{-1} \ln^6 h).$$

As $\Delta(0) = O(\varepsilon^2)$, this gives the estimate $\Delta(\tau) = O(\varepsilon^2 h^{-1} \ln^6 h)$. As $\frac{\partial h}{\partial I} = \omega$, by (4.10) we have

$$|\overline{h}(\tau) - \hat{h}(\tau)| = O(\varepsilon^2 h^{-1} \ln^5 h). \tag{4.13}$$

From the estimate on $\overline{h}(\tau) - \hat{h}(\tau)$ we have just proved and (4.9) we get that $\overline{h}(\tau) - \hat{h}(\tau) = O(\varepsilon \ln^{-5} \varepsilon)$ and so for small ε we have $|\overline{h}(\tau) - \hat{h}(\tau)| < 0.5\varepsilon \ln^{10} \varepsilon < 0.5\hat{h}(\tau)$, so the condition (4.10) actually holds for all τ considered in this lemma.

Let us now prove the estimate for φ. Denote $\omega_{0,1}(h) = \omega(h) + \varepsilon\omega_1(h)$. Then from (4.2) we have

$$\overline{\varphi}(\tau) - \overline{\varphi}(0) = \varepsilon^{-1} \int_0^\tau \Big(\omega_{0,1}(\overline{h}(\tau')) + \varepsilon^2 \hat{f}_{\varphi,2}(\overline{h}(\tau'), \overline{\varphi}(\tau'))\Big) d\tau'.$$

From Table 1 and (4.9) we have $\frac{\partial \omega_{0,1}}{\partial h} = O(h^{-1} \ln^{-2} h)$. Thus from (4.13) we have $\omega_{0,1}(\overline{h}(\tau)) - \omega_{0,1}(\hat{h}(\tau)) = O(\varepsilon^2 h^{-2} \ln^3 h)$. From Table 1 we have $\hat{f}_{\varphi,2}(\overline{h}(\tau), \overline{\varphi}(\tau)) = O(h^{-2} \ln h)$. So

$$\overline{\varphi}(\tau) - \overline{\varphi}(0) = \varepsilon^{-1} \int_0^\tau \omega_{0,1}(\hat{h}(\tau'))d\tau' + \varepsilon \int_0^\tau O((\hat{h}(\tau'))^{-2} \ln^3 \hat{h}(\tau'))d\tau'.$$

Again, we make a change of variable and compute the error term as an integral $d\hat{h}$: $\varepsilon \int_0^\tau O((\hat{h}(\tau'))^{-2} \ln^3 \hat{h}(\tau'))d\tau = O(\varepsilon h^{-1} \ln^4 h)$. This proves the formula for φ.

□

5 Cancellation lemma

In this section we prove the following lemma. It will be useful when we prove the formula for the phase, because due to this lemma two terms will cancel out. Recall the notation $\omega_1(h) = \hat{f}_{\varphi,1} = \overline{f}^0_{\varphi,1}$.

Lemma 5.1. *Consider a solution $\hat{h}$ of the averaged system (4.5). Take any $\tau_1 < \tau_2 < \tau_*$ and denote $h_1 = \hat{h}(\tau_1)$, $h_2 = \hat{h}(\tau_2)$. We have $0 < h_2 < h_1$. Then*

$$\int_{\tau_1}^{\tau_2} \omega_1(\hat{h}(\tau))d\tau = -\frac{2\pi}{\Theta_3}\Big|_{h_1}^{h_2} u^0_{h,1}(\hat{h}, 0) + O(h_1^{1/2}) + O(\varepsilon \ln^{-1} h). \quad (5.1)$$

Let us first estimate ω_1. Denote $\mathcal{I}(h) = \int_0^{2\pi} t(\varphi) f^0_h(\varphi)d\varphi$.

Lemma 5.2.

$$\omega_1(h) = \frac{1}{T}\frac{d\mathcal{I}}{dh} + O(h^{-1/2}\ln^{-1} h). \quad (5.2)$$

Proof. Integrating by parts, we can write

$$2\pi\omega_1 = \int_0^{2\pi} f^0_\varphi d\varphi = \Big|_0^{2\pi} \varphi f^0_\varphi - \int_0^{2\pi} \varphi \frac{\partial f^0_\varphi}{\partial \varphi} d\varphi.$$

Using (2.2) and the equality $\frac{1}{T}\frac{\partial}{\partial h}(T f^0_h) = \frac{\partial f^0_h}{\partial h} + \frac{1}{T}\frac{dT}{dh} f^0_h$, this rewrites as

$$2\pi\omega_1 = 2\pi f^0_\varphi(h, 0) - \int_0^{2\pi} \varphi \operatorname{div}(f^0) d\varphi + \int_0^{2\pi} \varphi \frac{1}{T}\frac{\partial}{\partial h}(T f^0_h) d\varphi.$$

By Table 1 the first term is $O(h^{-1/2}\ln^{-1} h)$. The second term is $O(1)$ as $\operatorname{div}(f^0)$ is bounded. As $\frac{\partial}{\partial h}$ commutes with integrating by φ, we can rewrite the last term as $\frac{1}{T}\frac{d}{dh}\int_0^{2\pi} \varphi T f^0_h d\varphi = \frac{2\pi}{T}\frac{d\mathcal{I}}{dh}$. We have obtained (5.2). □

Proof of Lemma 5.1. By Table 1 we have $\omega_1 = \overline{f}^0_{\varphi,1} = O(h^{-1}\ln^{-3} h)$. This means that $\int_0^{h'} T\omega_1 dh$ converges, so from (4.8) we have

$$\int_{\tau_1}^{\tau_2} \omega_1(\hat{h}(\tau))d\tau = -\frac{1}{\Theta_3}\int_{h_1}^{h_2} T\omega_1 d\hat{h} + O(h_1 \ln^{-1} h_1) + O(\varepsilon \ln^{-1} h).$$

By Lemma 5.2 we can rewrite this as

$$\int_{\tau_1}^{\tau_2} \omega_1(\hat{h}(\tau))d\tau = -\frac{1}{\Theta_3}\Big|_{h_1}^{h_2} \mathcal{I}(\hat{h}) + O(h_1^{1/2}) + O(\varepsilon \ln^{-1} h). \quad (5.3)$$

As $dt = \frac{T d\varphi}{2\pi}$, by (3.8) we have

$$u^0_{h,1}(h, 0) = \frac{1}{T}\int_0^T \left(t - \frac{T}{2}\right) f^0_h(t)dt = \frac{1}{2\pi}\int_0^{2\pi} t f^0_h(t) d\varphi - \frac{1}{2}\int_0^T f^0_h dt.$$

By [4, Corollary 3.1] $\int_0^T f^0_h(t)dt = -\Theta_3 + O(h \ln h)$. Hence,

$$\Big|_{h_1}^{h_2} u^0_{h,1}(\hat{h}, 0) = \frac{1}{2\pi}\Big|_{h_1}^{h_2} \mathcal{I} + O(h_1 \ln h_1).$$

Comparing this with (5.3), we get (5.1). □

6 Formula for the pseudo-phase

Consider a solution $h(\tau), \varphi(\tau)$ of the perturbed equation (2.1) that approaches the separatrices. Let the initial conditions be $h(0) = h_0, \varphi(0) = \varphi_0$. Set

$$\hat{h}_0 = h_0 - \varepsilon u^0_{h,1}(h_0, \varphi_0).$$

By (3.1) and Lemma 10.2 this approximates the value of h in the averaging chart corresponding to h_0, φ_0 with error $O(\varepsilon^2)$. Let $\hat{h}(\tau)$ be the solution of (4.6) with this initial condition. Let τ_* be the first time such that $\hat{h}(\tau_*) = 0$. In Section 4 we showed that τ_* exists. Let h_{-1} be the value of $h(\tau)$ at the first crossing of the transversal $\varphi = 0$ with $h(\tau) < \varepsilon\Theta_3 + \varepsilon^{4/3}$. For small ε it exists by Lemma 6.3 below. As h decreases by approximately $\varepsilon\Theta_3$ during one turn, for most points this will be the last crossing of this transversal. However, this choice of h_{-1} allows us to dismiss crossings of the transversal with $h < O(\varepsilon^{4/3})$. If we wished to consider crossings with small h, we would have to consider points being captured into the saddle of the perturbed system. Let h_{-2} be the value of h at the crossing before h_{-1}, h_{-3} be the previous crossing, and so on. Let us prove a formula for the pseudo-phase $\frac{h_{-1}}{\varepsilon\Theta_3}$. This formula (6.1) is similar to the one from [6], see also Section 1 for more references.

$$\frac{h_{-1}}{\varepsilon\Theta_3} = \left\{\frac{1}{2\pi}\left(\varphi_0 + \frac{1}{\varepsilon}\int_{\tau=0}^{\tau_*} \left(\omega(\hat{h}(\tau)) + \varepsilon\omega_1(\hat{h}(\tau))d\tau\right)\right) + \frac{u_*}{\Theta_3} + O(\varepsilon^{1/3}\ln^{11/3}\varepsilon)\right\} + O(\varepsilon^{1/3}), \quad (6.1)$$

where $u_* = \frac{1}{4}(\Theta_1 - \Theta_2)$. Note that $u_* = \lim_{h\to+0} u^0_{h,1}(h, 0)$ by Lemma 6.1 below. Let us also recall the notation $\omega_1 = \hat{f}_{\varphi,1} = \overline{f}^0_{\varphi,1}$ and that Θ_2 corresponds to $0 < \varphi < \pi$ and Θ_2 to $\pi < \varphi < 2\pi$.

Remark 6.1. *$\varphi = 0$ corresponds to the transversal Γ defined in Section 9.1. However, one may easily show that* (6.1) *also holds if we take as Γ any transversal tangent to the bisector of the angle between the separatrices.*

Remark 6.2. *We have assumed earlier that $\Theta_1, \Theta_2 > 0$. If they have different signs with $\Theta_3 = \Theta_1 + \Theta_2 > 0$, the last transversal crossing can happen for $h > \varepsilon\Theta_3$. In this case h_{-1} should be defined in such way that $h > \varepsilon^{4/3}$ during all the time before the moment corresponging to h_{-1}. Then the right-hand side of* (6.1) *would give the fractional part of $\frac{h_{-1}}{\varepsilon\Theta_3}$.*

First let us prove some auxiliary statements.

Lemma 6.1.

$$\lim_{h\to+0} u^0_{h,1}(h,0) = \frac{\Theta_1 - \Theta_2}{4},$$

Here Θ_2 corresponds to $0 < \varphi < \pi$ and Θ_1 to $\pi < \varphi < 2\pi$.

Proof. Let us split the integral expression (3.9) (with f replaced by f^0) for $u^0_{h,1}(h,0)$ into the integrals over the part of the trajectory near l_1 and near l_2. For the first part the value of $\varphi(t) - \pi$ is close to $\pi/2$ far away from the saddle C. But close to C we have $f_h \approx 0$, so the integral near l_1 is close to $\Theta_1/4$. Similarly, the integral near l_2 is close to $-\Theta_2/4$. □

Lemma 6.2. *Take $\tau_1 < \tau_*$, denote $h_1 = \hat h(\tau_1)$. Then we have*

$$\int_{\tau_1}^{\tau_*} \omega(\hat h(\tau))d\tau = \frac{2\pi}{\Theta_3} h_1 + O(\varepsilon h_1) + O(h_1^2 \ln h_1). \tag{6.2}$$

Proof. As $\omega T = 2\pi$, (4.8) impies that

$$\int_{\tau_1}^{\tau_*} \omega(\hat h(\tau))d\tau = -\frac{2\pi}{\Theta_3} \int_{h_1}^{0} \Big(1 + O(\hat h \ln \hat h) + O(\varepsilon)\Big) d\hat h,$$

which gives the required estimate. □

Lemma 6.3. *Assume $\Theta_1, \Theta_2 > 0$. Then there is $c_1 > 0$ such that for all small enough ε the following holds. Take a point $(h_0, 0)$ on the transversal $\varphi = 0$ with $\varepsilon\Theta_3 + \varepsilon^{4/3} \le h_0 < c_1$. Then the orbit of this point intersects the transversal $\varphi = 0$ once more with*

$$h = h_0 - \varepsilon\Theta_3 + O(\varepsilon h_0 \ln h_0) + O(\varepsilon^2 h_0^{-1/2}).$$

Proof. By [4, Lemma 3.5, Corollary 3.4] there is $c_2 > 0$ such that for $c_2\varepsilon \le h_0 \le c_1$ the orbit crosses the transversal and we have $h = h_0 - \varepsilon \oint_{h=h_0} f_h dt + O(\varepsilon^2 h_0^{-1/2})$. By (4.1) we have $f_h - f_h^0 = \varepsilon\psi(p,q,\varepsilon)$ for some smooth ψ. By [4, Lemma 3.2] $\oint_{h=h_0} \psi dt = O(1)$, so $\oint_{h=h_0} f_h dt = \oint_{h=h_0} f_h^0 dt + O(\varepsilon) = -\Theta_3 + O(h_0 \ln h_0) + O(\varepsilon)$ by (4.4). As $\varepsilon = O(h)$, this gives the required estimate.

For $\varepsilon\Theta_3 + \varepsilon^{4/3} \le h_0 < c_2\varepsilon$ by [4, Proposition 5.1] the orbit of our point intersects the transversal $\varphi = 0$ once more (the condition $\Theta_1, \Theta_2 > 0$ is used here). Moreover, arguing as in the proof of [4, Proposition 5.1], we can get that for this new intersection $h = h_0 - \varepsilon\Theta_3 + O(\varepsilon^{3/2})$. As $\Theta_3\varepsilon \le h_0 < c_2\varepsilon$, this estimate is equivalent to the one claimed in the lemma. □

Lemma 6.4. *There is a constant c_1 such that for $h_{-n} > c_1$ we have*

$$h_{-n} = h_{-1} + \varepsilon\Theta_3(n-1) + O(h_{-n}^2 \ln h_{-n}) + \varepsilon O((h_{-n})^{1/2}). \tag{6.3}$$

Proof. This follows from Lemma 6.3 by summation. □

Let us return to the proof of the formula for pseudo-phase. We denote by $\overline{h}(\tau), \overline{\varphi}(\tau)$ the solution $h(\tau), \varphi(\tau)$ of the initial equation, written in the averaged chart (3.1). Denote $\overline{h}_0 = \overline{h}(0)$, $\overline{\varphi}_0 = \overline{\varphi}(0)$. We have $\hat h_0 - \overline{h}_0 = O(\varepsilon^2)$, so we may use Lemma 4.1

Lemma 6.5. *For τ such that*

$$\hat h(\tau) > \varepsilon \ln^{10} \varepsilon \tag{6.4}$$

$h(\tau)$, $\overline{h}(\tau)$ and $\hat h(\tau)$ are close:

$$\begin{aligned} \hat h - \overline{h} &= O(\varepsilon^2 \hat h^{-1} \ln^5 \hat h) = o(\hat h), \\ \overline{h} - h &= O(\varepsilon) = o(\hat h). \end{aligned} \tag{6.5}$$

Proof. The first estimate is given by Lemma 4.1. To obtain the second one, we just plug the estimates from Table 1 into the equation $h = \overline{h} + \varepsilon u_{h,1} + \varepsilon^2 u_{h,2}$ from (3.1). $\square$

Consider a moment τ_1 such that $\varphi(\tau_1) = 0$ and $\hat{h}(\tau_1)$ is as close as possible to $\varepsilon^{2/3} \ln^{4/3} \varepsilon$. Note that we have (6.4) for $\tau = \tau_1$. We may check that under the condition (6.4) the difference between $\hat{h}(\tau)$ for consequtive times τ with $\varphi(\tau) = 0$ is $O(\varepsilon)$. Indeed, the time between consequtive fast times of crossing the transversal $\varphi = 0$ is $O(T)$ and $\dot{\hat{h}}$ is $O(T^{-1})$. Hence,

$$\hat{h}(\tau_1) = (1 + o(1))\varepsilon^{2/3} \ln^{4/3} \varepsilon. \tag{6.6}$$

Denote by $h_1, \varphi_1, \overline{h}_1, \overline{\varphi}_1, \hat{h}_1$ the values of $h, \varphi, \overline{h}, \overline{\varphi}, \hat{h}$ at the slow time τ_1. As justified by (6.5), we may write h_1 instead of $\hat{h}_1$ and $\overline{h}_1$ in the error terms. For brevity let us even denote $h = h_1$ for the error terms and write simply $O(h)$.

We will split the integral in (6.1) into integrals from 0 to τ_1 and from τ_1 to τ_*. First, let us check that

$$\varphi_0 + \frac{1}{\varepsilon} \int_{\tau=0}^{\tau_1} \Big(\omega(\hat{h}(\tau)) + \varepsilon \omega_1(\hat{h}(\tau)) \Big) d\tau = 2\pi m + O(\varepsilon h^{-1} \ln^4 h), \tag{6.7}$$

where $m \in \mathbb{Z}$. By Lemma 4.1 we have

$$\frac{1}{\varepsilon} \int_{\tau=0}^{\tau_1} \Big(\omega(\hat{h}(\tau)) + \varepsilon \omega_1(\hat{h}(\tau)) \Big) d\tau = \overline{\varphi}_1 - \overline{\varphi}_0 + O(\varepsilon h^{-1} \ln^4 h).$$

We also have $\varphi = \overline{\varphi} + \varepsilon u_{\varphi,1}$. By Table 1 $u_{\varphi,1} = O(h^{-1} \ln h)$, so $\overline{\varphi}_1 - \overline{\varphi}_0 = \varphi_1 - \varphi_0 + O(\varepsilon h^{-1} \ln h)$. As $\varphi_1 = 2\pi m$, this gives the required equality (6.7).

Now let us use (6.2), (5.1) to compute the remaining terms in (6.1). We have

$$\begin{aligned} \frac{1}{2\pi\varepsilon} \bigg(\int_{\tau=\tau_1}^{\tau_*} (\omega(\hat{h}(\tau)) + \varepsilon\omega_1(\hat{h}(\tau))) d\tau \bigg) + \frac{u_*}{\Theta_3} &= \\ = \frac{1}{\varepsilon\Theta_3} \Big(\hat{h}_1 + \varepsilon u^0_{h,1}(\hat{h}_1, 0) \Big) &+ O(h^{1/2}) + O(\varepsilon^{-1} h^2 \ln h). \end{aligned} \tag{6.8}$$

Note that the term $O(\varepsilon \ln^{-1} h)$ from (5.1) is absorbed into $O(h^{1/2})$ by (6.4). By Table 1 and (6.4) we have $\frac{\partial}{\partial h} \varepsilon u^0_{h,1} = o(1)$. Hence, by (6.5) we have

$$\hat{h}_1 + \varepsilon u^0_{h,1}(\hat{h}_1, 0) = \overline{h}_1 + \varepsilon u^0_{h,1}(\overline{h}_1, 0) + O(\varepsilon^2 h^{-1} \ln^5 h) = \overline{h}_1 + \varepsilon u_{h,1}(\overline{h}_1, 0, \varepsilon) + O(\varepsilon^2 h^{-1} \ln^5 h).$$

The last equality is justified by Lemma 10.2. The error term $O(\varepsilon^2)$ appears, but it is absorbed into $O(\varepsilon^2 h^{-1} \ln^5 h)$. As $0 = \varphi_1 = \overline{\varphi}_1 + \varepsilon u_{\varphi,1}$, by Table 1 we have $\overline{\varphi}_1 = O(\varepsilon h^{-1} \ln h)$. Hence, by the estimate $\frac{\partial u_{h,1}}{\partial \varphi} = O(\ln h)$ from Table 1 we get

$$\varepsilon u_{h,1}(\overline{h}_1, 0, \varepsilon) = \varepsilon u_{h,1}(\overline{h}_1, \overline{\varphi}_1, \varepsilon) + O(\varepsilon^2 h^{-1} \ln^2 h)$$

and

$$\hat{h}_1 + \varepsilon u^0_{h,1}(\hat{h}_1, 0) = \overline{h}_1 + \varepsilon u_{h,1}(\overline{h}_1, \overline{\varphi}_1, \varepsilon) + O(\varepsilon^2 h^{-1} \ln^5 h).$$

As

$$h_1 = \overline{h}_1 + \varepsilon u_{h,1}(\overline{h}_1, \overline{\varphi}_1, \varepsilon) + \varepsilon^2 u_{h,2}(\overline{h}_1, \overline{\varphi}_1, \varepsilon)$$

and by (3.1) and $\varepsilon^2 u_{h,2}$ is small by Table 1, we obtain

$$\hat{h}_1 + \varepsilon u^0_{h,1}(\hat{h}_1, 0) = h_1 + O(\varepsilon^2 h^{-1} \ln^5 h).$$

Combining this with (6.8), we get

$$\frac{1}{2\pi\varepsilon} \bigg(\int_{\tau=\tau_1}^{\tau_*} (\omega(\hat{h}(\tau)) + \varepsilon\omega_1(\hat{h}(\tau))) d\tau \bigg) + \frac{u_*}{\Theta_3} = \frac{h_1}{\varepsilon\Theta_3} - R(h_1)$$

with the error term

$$R = O(h^{1/2}) + O(\varepsilon h^{-1} \ln^5 h) + O(\varepsilon^{-1} h^2 \ln h).$$

After taking a sum with (6.7), we get

$$\frac{h_1}{\varepsilon\Theta_3} = \frac{1}{2\pi} \bigg(\varphi_0 + \frac{1}{\varepsilon} \int_{\tau=\tau_0}^{\tau_*} (\omega(\hat{h}(\tau)) + \varepsilon\omega_1(\hat{h}(\tau)) d\tau) \bigg) + \frac{u_*}{\Theta_3} - m + R(h_1).$$

Note that R absorbs the error term in (6.7). Applying (6.3), we get the required formula (6.1), but with the error term $R(h_1)$ depending on h_1. Note that the error term in (6.3) is not greater than R. Then we just plug in the expression (6.6) for h_1 and obtain $R = O(\varepsilon^{1/3} \ln^{11/3} \varepsilon)$. One may check that (6.6) minimizes the error term. Indeed, first we check that up to some power of $\ln \varepsilon$ the value of R is minimal for $h \approx \varepsilon^{2/3}$. Then $\ln h \approx (2/3) \ln \varepsilon$, and from this we see that R is minimal for h given by (6.6). This completes the proof of formula (6.1).

7 Proofs

Proof of Lemma 3.2. First let us check that

$$\langle \frac{\partial f_h}{\partial \varphi} u_{\varphi,1} \rangle_\varphi = \langle \frac{\partial f_\varphi}{\partial \varphi} u_{h,1} \rangle_\varphi. \tag{7.1}$$

We shall use the following equalities (see Lemma 3.1):

$$\frac{\partial u_{h,1}}{\partial \varphi} = \frac{1}{\omega}(f_h - \overline{f}_{h,1}), \qquad \frac{\partial u_{\varphi,1}}{\partial \varphi} = \frac{1}{\omega}(f_\varphi + \frac{\partial \omega}{\partial h} u_{h,1} - \overline{f}_{\varphi,1}), \tag{7.2}$$
$$\overline{f}_{h,1} = \langle f_h \rangle_\varphi, \qquad \overline{f}_{\varphi,1} = \langle f_\varphi \rangle_\varphi.$$

Integrating by parts, we get

$$2\pi \langle \frac{\partial f_h}{\partial \varphi} u_{\varphi,1} \rangle_\varphi = \int_{\varphi=0}^{2\pi} \frac{\partial f_h}{\partial \varphi} u_{\varphi,1} d\varphi = -\int_{\varphi=0}^{2\pi} f_h \frac{\partial u_{\varphi,1}}{\partial \varphi} d\varphi =$$
$$= -\frac{1}{\omega} \int_{\varphi=0}^{2\pi} f_h f_\varphi d\varphi \; - \; \frac{1}{\omega}\frac{\partial \omega}{\partial h} \int_{\varphi=0}^{2\pi} f_h u_{h,1} d\varphi \; + \; \frac{1}{\omega} \overline{f}_{\varphi,1} \int_{\varphi=0}^{2\pi} f_h d\varphi.$$

Using (7.2), we can rewrite the integral in the second term:

$$\frac{1}{\omega} \int_{\varphi=0}^{2\pi} f_h u_{h,1} d\varphi \; = \; \int_{\varphi=0}^{2\pi} \frac{\partial u_{h,1}}{\partial \varphi} u_{h,1} d\varphi + \frac{1}{\omega} \overline{f}_{h,1} \int_{\varphi=0}^{2\pi} u_{h,1} d\varphi \; = \; 0. \tag{7.3}$$

Indeed, the first term is $\int_{\varphi=0}^{2\pi} u_{h,1} du_{h,1} = 0$, and the second term is also zero because $\langle u_{h,1} \rangle_\varphi = 0$. Hence, we have

$$2\pi \langle \frac{\partial f_h}{\partial \varphi} u_{\varphi,1} \rangle_\varphi = -\frac{1}{\omega} \int_{\varphi=0}^{2\pi} f_h f_\varphi d\varphi \; + \; \frac{2\pi}{\omega} \overline{f}_{h,1} \overline{f}_{\varphi,1}.$$

Similarly,

$$2\pi \langle \frac{\partial f_\varphi}{\partial \varphi} u_{h,1} \rangle_\varphi = \int_{\varphi=0}^{2\pi} \frac{\partial f_\varphi}{\partial \varphi} u_{h,1} d\varphi = -\int_{\varphi=0}^{2\pi} f_\varphi \frac{\partial u_{h,1}}{\partial \varphi} d\varphi =$$
$$= -\frac{1}{\omega} \int_{\varphi=0}^{2\pi} f_h f_\varphi d\varphi \; + \; \frac{1}{\omega} \overline{f}_{h,1} \int_{\varphi=0}^{2\pi} f_\varphi d\varphi.$$

We obtained (7.1). This means that (3.6) can be rewritten as $\overline{f}_{h,2} = \langle (\frac{\partial f_h}{\partial h} + \frac{\partial f_\varphi}{\partial \varphi}) u_{h,1} \rangle_\varphi$. By Lemma 2.1 this equals $\langle (\frac{\partial f_q}{\partial q} + \frac{\partial f_p}{\partial p}) u_{h,1} \rangle_\varphi - \frac{1}{T}\frac{\partial T}{\partial h} \langle f_h u_{h,1} \rangle_\varphi$. By (7.3) the last term is equal to zero. This means

$$\overline{f}_{h,2} = \langle (\frac{\partial f_q}{\partial q} + \frac{\partial f_p}{\partial p}) u_{h,1} \rangle_\varphi.$$

□

Proof of Lemma 3.3. The function $u_{h,1}$ is uniquely determined by two properties. The first one is that $\frac{\partial u_{h,1}}{\partial t} = f_h(t) - \langle f_h \rangle_t$ (this follows from (3.4), (3.5)). Denote by U the expression on the right hand side of (3.8). We have

$$\frac{\partial U}{\partial t_0} = \frac{1}{T} \int_0^T \Big(t - \frac{T}{2}\Big) \frac{\partial f_h}{\partial t}(t + t_0) dt.$$

Integrating by parts, this can be rewritten as

$$\frac{\partial U}{\partial t_0} = \frac{1}{T} \Big|_{t=0}^{T} f_h(t + t_0)\Big(t - \frac{T}{2}\Big) - \frac{1}{T} \int_0^T f_h(t + t_0) dt = f_h(t_0) - \langle f_h \rangle_t.$$

Hence the first property of $u_{h,1}$ holds for U.

The second property is that $\langle u_{h,1} \rangle_t = 0$. This also holds for U, it is checked by writing $\int U(t_0) dt_0$ as a double integral and changing the order of integration. □

8 Formulas for the averaging chart

In this section we present formulas for $\overline{f}_{\varphi,2}$ and $\overline{f}_{h,3}$ from Lemma 3.1 and prove this lemma. We use the notation introduced in Section 3.

Lemma 8.1.

- *Denote by x the column vector (h, φ) and by $\overline{x}$ the column vector $(\overline{h}, \overline{\varphi})$. Let $\overline{f}_{x,i} = (\overline{f}_{h,i}, \overline{f}_{\varphi,i})$, $u_{x,i} = (u_{h,i}, u_{\varphi,i})$.*

- *For $k = x, h, \varphi$ denote $u_{k,1,2} = u_{k,1} + \varepsilon u_{k,2}$, $\overline{f}_{k,1,2} = \overline{f}_{k,1} + \varepsilon \overline{f}_{k,2}$, $\overline{f}_{k,2,3} = \overline{f}_{k,2} + \varepsilon \overline{f}_{k,3}$, $\overline{f}_{k,1,2,3} = \overline{f}_{k,1} + \varepsilon \overline{f}_{k,2} + \varepsilon^2 \overline{f}_{k,3}$. For $k = x$ the terms $u_{\varphi,2}$, $\overline{f}_{\varphi,3}$ appear, we set $u_{\varphi,2} = \overline{f}_{\varphi,3} = 0$.*
- *For a vector-function $g(x) = (g_1, \dots, g_l)$ denote $\left(\frac{\partial g}{\partial x}\right)_{int} = (\frac{\partial g_1}{\partial x}(\xi_1), \dots, \frac{\partial g_l}{\partial x}(\xi_l))$, $\left(\frac{\partial^2 g}{\partial x^2}\right)_{int} = (\frac{\partial^2 g_1}{\partial x^2}(\eta_1), \dots, \frac{\partial^2 g_l}{\partial x^2}(\eta_l))$, where ξ_i, η_i are some intermediate points on the segment $[x, \overline{x}]$.*

Then we have the following system of linear equations determining $\overline{f}_{\varphi,2}$ and $\overline{f}_{h,3}$:

$$
\begin{aligned}
&(1 + \varepsilon \frac{\partial u_{\varphi,1}}{\partial \varphi})\overline{f}_{\varphi,2} + \varepsilon^2 \frac{\partial u_{\varphi,1}}{\partial h}\overline{f}_{h,3} = \\
&\quad = \frac{\partial \omega}{\partial h} u_{h,2} + \frac{1}{2} u^T_{h,1,2} \Big(\frac{\partial^2 \omega}{\partial h^2}\Big)_{int} u_{h,1,2} + \Big(\frac{\partial f_\varphi}{\partial x}\Big)_{int} u_{x,1,2} - \\
&\quad - \frac{\partial u_{\varphi,1}}{\partial h}\overline{f}_{h,1,2} - \frac{\partial u_{\varphi,1}}{\partial \varphi}\overline{f}_{\varphi,1}, \\
&(1 + \varepsilon \frac{\partial u_{h,1,2}}{\partial h})\overline{f}_{h,3} + \frac{\partial u_{h,1,2}}{\partial \varphi}\overline{f}_{\varphi,2} = \\
&\quad = \frac{\partial f_h}{\partial h} u_{h,2} + \frac{1}{2} u^T_{x,1,2} \Big(\frac{\partial^2 f_h}{\partial x^2}\Big)_{int} u_{x,1,2} - \\
&\quad - \frac{\partial u_{h,1}}{\partial h}\overline{f}_{h,2} - \frac{\partial u_{h,2}}{\partial h}\overline{f}_{h,1,2} - \frac{\partial u_{h,2}}{\partial \varphi}\overline{f}_{\varphi,1}.
\end{aligned}
\tag{8.1}
$$

Proof of lemmas 3.1 and 8.1. We shall differentiate the coordinate change (3.1) with respect to the time and rewrite all emerging terms as functions of $\overline{x}$. The derivatives of the left hand sides of (3.1) are given by (2.1). They are functions of x, let us write Taylor's expansions at the point $\overline{x}$. We group together the terms of order at least 3 for the coordinate change in h and 2 for the change in φ

$$
\begin{aligned}
\dot h &= \varepsilon f_h(x) = \varepsilon f_h(\overline{x} + \varepsilon u_{x,1} + \varepsilon^2 u_{x,2}) = \\
&= \varepsilon f_h(\overline{x}) + \varepsilon^2 \frac{\partial f_h}{\partial x} u_{x,1} + \varepsilon^3 \Big(\frac{\partial f_h}{\partial h} u_{h,2} + \frac{1}{2} u^T_{x,1,2} \Big(\frac{\partial^2 f_h}{\partial x^2}\Big)_{int} u_{x,1,2} \Big), \\
\dot\varphi &= \omega(h) + \varepsilon f_\varphi(x) = \omega(\overline{h} + \varepsilon u_{h,1} + \varepsilon^2 u_{h,2}) + \varepsilon f_\varphi(\overline{x} + \varepsilon u_{x,1} + \varepsilon^2 u_{x,2}) = \\
&= \omega(\overline{h}) + \varepsilon \Big(\frac{\partial \omega}{\partial h} u_{h,1} + f_\varphi(\overline{x}) \Big) + \\
&\qquad + \varepsilon^2 \Big(\frac{\partial \omega}{\partial h} u_{h,2} + \frac{1}{2} u^T_{h,1,2} \Big(\frac{\partial^2 \omega}{\partial h^2}\Big)_{int} u_{h,1,2} + \Big(\frac{\partial f_\varphi}{\partial x}\Big)_{int} u_{x,1,2} \Big).
\end{aligned}
$$

Now we write the terms containing the derivatives of $u_{k,i}$.

$$
\begin{aligned}
&\varepsilon \dot u_{h,1}(\overline{x}) + \varepsilon^2 \dot u_{h,2}(\overline{x}) = \\
&= \varepsilon \frac{\partial u_{h,1}}{\partial \varphi}\omega + \varepsilon^2 \Big(\frac{\partial u_{h,2}}{\partial \varphi}\omega + \frac{\partial u_{h,1}}{\partial x}\overline{f}_{x,1} \Big) + \varepsilon^3 \Big(\frac{\partial u_{h,1}}{\partial x}\overline{f}_{x,2,3} + \frac{\partial u_{h,2}}{\partial x}\overline{f}_{x,1,2,3} \Big), \\
&\varepsilon \dot u_{\varphi,1}(\overline{x}) = \varepsilon \frac{\partial u_{\varphi,1}}{\partial \varphi}\omega + \varepsilon^2 \frac{\partial u_{\varphi,1}}{\partial x}\overline{f}_{x,1,2,3}.
\end{aligned}
$$

Let us plug these expressions together with (3.2) into the time derivative of (3.1). Equating the terms of the same order in ε (grouping together the terms with order at least 3 for the equation on h and 2 for the equation on φ), we get (3.4) and (3.5), as well as the following equations:

$$
\begin{aligned}
\overline{f}_{\varphi,2} &= \frac{\partial \omega}{\partial h} u_{h,2} + \frac{1}{2} u^T_{h,1,2} \Big(\frac{\partial^2 \omega}{\partial h^2}\Big)_{int} u_{h,1,2} + \Big(\frac{\partial f_\varphi}{\partial x}\Big)_{int} u_{x,1,2} - \frac{\partial u_{\varphi,1}}{\partial x}\overline{f}_{x,1,2,3}, \\
\overline{f}_{h,3} &= \frac{\partial f_h}{\partial h} u_{h,2} + \frac{1}{2} u^T_{x,1,2} \Big(\frac{\partial^2 f_h}{\partial x^2}\Big)_{int} u_{x,1,2} - \frac{\partial u_{h,1}}{\partial x}\overline{f}_{x,2,3} - \frac{\partial u_{h,2}}{\partial x}\overline{f}_{x,1,2,3},
\end{aligned}
$$

which are equivalent to (8.1), we just expand some terms like $\overline{f}_{x,1,2,3}$ in order to move the terms containing $\overline{f}_{\varphi,2}$ and $\overline{f}_{h,3}$ to the left hand side. □

9 Estimates related to the energy-angle variables

9.1 The coordinates $\tilde h$, $\tilde t_i$

Our goal in this section is to estimate how q, p depend on h, φ for $h \to 0$. To do so, we introduce new coordinates $\tilde h, \tilde t_i$. The subscript i is here because there will be different coordinate systems in different parts of the phase space. Then we will estimate how q, p depend on $\tilde{h}, \tilde{t}_i$ and how $\tilde{h}, \tilde{t}_i$ depend on h, φ. Combining these estimates, we will get the required estimates of the dependence of q, p on h, φ.

For simplicity we will assume that the Hamiltonian H is analytic. Then by [3] one can find a new coordinate system x, y in the neighborhood of the saddle C such that this coordinate change is analytic and volume preserving, and the unperturbed system in the new coordinates is determined by a Hamiltonian $H_{x,y} = H_{x,y}(xy)$.[1] Let $\tilde{h} = xy$, denote $a(\tilde{h}) = \frac{dH_{x,y}}{d\tilde{h}}$. Then in the new chart the unperturbed system rewrites as

$$\dot{x} = a(\tilde{h})x, \ \dot{y} = -a(\tilde{h})y. \tag{9.1}$$

Note that $\tilde{h}$ is a first integral of this system. Also note that as $H_{x,y}(x, y) = H(p, q)$, we have that $\tilde{h}$ is a smooth function of h. This also means that $\tilde{h}$ is defined on the whole phase space, even far from C. We will assume $a > 0$ (else we may swap x and y). Rescaling x and y if needed, we may assume that the neighborhood of C where the new coordinates are defined contains the square $\mathcal{S} = \{-1 \le x, y \le 1\}$.

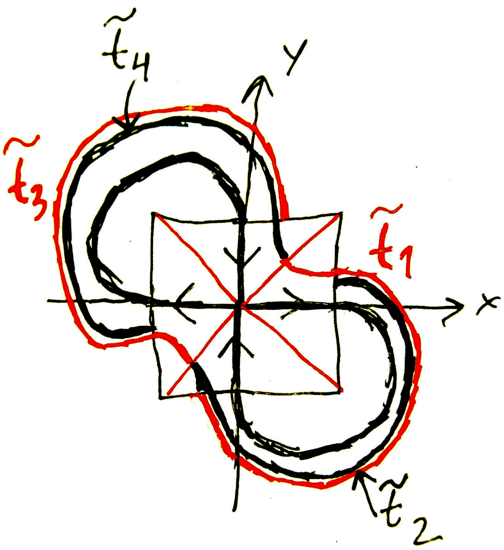

Figure 1: Domains where $\tilde{t}_i$ are defined.

The diagonals $x = \pm y$ split $\mathcal{S}$ into four triangles adjanced to each of its sides. In each such triangle let us introduce the time $\tilde{t}_i$ (it can be positive or negative) that passes after the trajectory intersects the adjanced side of $\mathcal{S}$. The time $\tilde{t}_i$ can also be continued outside the square to the neighborhood of the separatix crossing the transversal $\tilde{t}_i = 0$ (it is a side of $\mathcal{S}$). Domains where each $\tilde{t}_i$ is defined are drawn in figure 1. Note that the coordinate systems $\tilde{h}, \tilde{t}_i$ cover the whole phase space (we only consider h close to zero here).

We will assume that $\varphi = 0$ corresponds to the transversal Γ given by $x = y \ge 0$.

9.2 Estimates on how q, p depend on $\tilde{h}, \tilde{t}_i$

Outside of $\mathcal{S}$ each point of the phase space is covered by two coordinate systems $\tilde{h}, \tilde{t}_i$. For both of them the coordinate change $p, q \leftrightarrow \tilde{h}, \tilde{t}_i$ is defined and is smooth without singularities. So we only need to consider what happens inside $\mathcal{S}$. For definiteness, let us restrict ourselves to the triangle $\{x \ge y, \ x \ge -y\} \cap \mathcal{S}$. For brevity we will write just $\tilde{t}$ for the coordinate $\tilde{t}_i$ defined in this triangle. This means that $\tilde{t}$ is the time after the trajectory intersects the line $x = 1$. Moreover, we will only consider the upper half of this triangle ($x \ge y, \ y \ge 0$). Then $\tilde{t}$ will be negative. We have

$$\begin{aligned} &x = e^{a(\tilde{h})\tilde{t}}, \ y = \tilde{h}e^{-a(\tilde{h})\tilde{t}}, \\ &\tilde{h} = xy, \ \tilde{t} = \frac{\ln x}{a(xy)}; \end{aligned} \tag{9.2}$$

$$\frac{\partial x}{\partial \tilde{h}} = a'(\tilde{h})\tilde{t}x, \ \frac{\partial x}{\partial \tilde{t}} = a(\tilde{h})x, \ \frac{\partial y}{\partial \tilde{h}} = -a'(\tilde{h})\tilde{t}y + \frac{1}{x}, \ \frac{\partial y}{\partial \tilde{t}} = -a(\tilde{h})y. \tag{9.3}$$

[1]The result of [3] is for the case when H periodically depends on the time, but one may check that when this dependence is constant the coordinate change constructed in [3] does not depend on the time.

Note that $\tilde{t}x = \tilde{t}e^{a(\tilde{h})\tilde{t}} = O(1)$, as $a(\tilde{h})\tilde{t} < 0$. It follows that

$$\frac{\partial y}{\partial \tilde{h}} = O(h^{-1/2}); \ \frac{\partial x}{\partial \tilde{h}}, \ \frac{\partial y}{\partial \tilde{t}}, \ \frac{\partial x}{\partial \tilde{t}} = O(1). \tag{9.4}$$

Here we use that $x \geq \tilde{h}^{1/2}$. Also note that $\frac{\tilde{h}}{h} \to c \neq 0$ as $h \to 0$, so we may write $O(h^k)$ instead of $O(\tilde{h}^k)$. It also follows from (9.3) that

$$\begin{aligned} \frac{\partial^2 y}{\partial \tilde{h}^2} &= -\frac{2a'(\tilde{h})\tilde{t}}{x} + \cdots = O(h^{-1/2}\ln h); \\ \frac{\partial^2 y}{\partial \tilde{t}\partial \tilde{h}} &= O(h^{-1/2}); \ \frac{\partial^2 y}{\partial \tilde{t}^2}, \frac{\partial^2 x}{\partial \tilde{h}^2}, \frac{\partial^2 x}{\partial \tilde{t}^2}, \frac{\partial^2 x}{\partial \tilde{t}\partial \tilde{h}} = O(1). \end{aligned} \tag{9.5}$$

Now let us return from (x, y) to (q, p). Let us consider a smooth function $\psi(p, q)$ without singularities (e.g. p or q), inside $\mathcal{S}$ it is a smooth function of x, y. Hence, all partial derivatives of orders 1 and 2 of ψ with respect to x, y are $O(1)$. We will use the following formula (a_i, b_i are some coordinate systems)

$$\frac{\partial^2}{\partial a_i \partial a_j} = \sum_l \frac{\partial^2 b_l}{\partial a_i \partial a_j}\frac{\partial}{\partial b_l} + \sum_{k,l} \frac{\partial b_l}{\partial a_j}\frac{\partial b_k}{\partial a_i}\frac{\partial^2}{\partial b_k \partial b_l}. \tag{9.6}$$

Using this formula and (9.4), (9.5), we get the following estimates.

$$\begin{aligned} &\frac{\partial \psi}{\partial \tilde{h}} = O(h^{-1/2}), \frac{\partial \psi}{\partial \tilde{t}_i} = O(1); \\ &\frac{\partial^2 \psi}{\partial \tilde{h}^2} = O(h^{-1}), \frac{\partial^2 \psi}{\partial \tilde{t}\partial \tilde{h}} = O(h^{-1/2}), \ \frac{\partial^2 \psi}{\partial \tilde{t}^2} = O(1). \end{aligned} \tag{9.7}$$

These estimates are valid everywhere: we obtained them in a part of $\mathcal{S}$, in other parts of $\mathcal{S}$ they can be obtained similarly, and outside of S we even have $O(1)$ on all right hand sides as the considered coordinate change is smooth.

9.3 Estimates on how $\tilde{h}$, $\tilde{t}_i$ depend on h, φ

First, recall that $\tilde{h}$ is a smooth function of h without singularities and $\frac{\tilde{h}}{h} \to c \neq 0$ as $h \to 0$.

Denote by $S(\tilde{h})$ the time that the solution with given $\tilde{h}$ takes to get from the diagonal of the square $\mathcal{S}$ to its side. Then the total time spent inside the square during each period is $4S(\tilde{h})$. From (9.2) we have $S(\tilde{h}) = -\frac{\ln \tilde{h}}{2a(\tilde{h})}$. Hence,

$$S = O(\ln h), \ \frac{dS}{dh} = O(h^{-1}), \ \frac{d^2 S}{dh^2} = O(h^{-2}). \tag{9.8}$$

Denote by $T_{reg,1}(h)$ and $T_{reg,2}(h)$ the times that the solution spends outside $\mathcal{S}$ near each of the separatix loops during each period. These are smooth functions of h. Then

$$T = 4S + T_{reg,1} + T_{reg,2}.$$

From (9.8) we get the estimates on T, ω from Table 1.

For each $\tilde{t}_i$ we have $\tilde{t}_i = t - t_{0,i}$, where $t_{0,i}$ is the value of t corresponding to $\tilde{t}_i = 0$. We have $t_{0,i} = kS + k_1 T_{reg,1} + k_2 T_{reg,2}$ with $k \in \{1, 2, 3\}$; $k_1, k_2 \in \{0, 1\}$ (see figure 1). Hence, we have

$$\tilde{t}_i = (4S(\tilde{h}) + T_{reg,1}(h) + T_{reg,2}(h))\varphi - kS(\tilde{h}) - k_1 T_{reg,1}(h) - k_2 T_{reg,2}(h).$$

From this (and smooth dependence of $\tilde{h}$ on h) we get

$$\begin{aligned} &\frac{\partial \tilde{t}_i}{\partial \varphi} = O(\ln h), \ \frac{\partial \tilde{t}_i}{\partial h} = O(h^{-1}), \\ &\frac{\partial^2 \tilde{t}_i}{\partial \varphi^2} = 0, \ \frac{\partial^2 \tilde{t}_i}{\partial h \partial \varphi} = O(h^{-1}), \ \frac{\partial^2 \tilde{t}_i}{\partial h^2} = O(h^{-2}), \\ &\frac{\partial \tilde{h}}{\partial h} = O(1), \ \frac{\partial^2 \tilde{h}}{\partial h^2} = O(1), \ \frac{\partial \tilde{h}}{\partial \varphi} = 0. \end{aligned} \tag{9.9}$$

9.4 Estimates on how q, p depend on h, φ

As above, let $\psi(p, q)$ be a smooth function without singularities. Applying the formula (9.6) to (9.7) and (9.9), we get the following estimates:

$$\begin{aligned} &\frac{\partial \psi}{\partial h} = O(h^{-1}), \ \frac{\partial \psi}{\partial \varphi} = O(\ln h), \\ &\frac{\partial^2 \psi}{\partial h^2} = O(h^{-2}), \ \frac{\partial^2 \psi}{\partial h \partial \varphi} = O(h^{-1}\ln h), \ \frac{\partial^2 \psi}{\partial \varphi^2} = O(\ln^2 h). \end{aligned} \tag{9.10}$$

9.5 Estimates on f

Here we obtain the estimates on f_h and f_φ from Table 1. The estimates on f_h follow from (9.10) as $f_h = f_q \frac{\partial h}{\partial q} + f_p \frac{\partial h}{\partial p}$ is smooth without singularities.

From (2.2) and Table 1 we have $\frac{\partial f_\varphi}{\partial \varphi} = O(h^{-1})$. Let us apply $\frac{\partial}{\partial h}$ to (2.2). As $\mathrm{div}(f)$ is smooth, by (9.10) we have $\frac{\partial}{\partial h}(\mathrm{div}(f)) = O(h^{-1})$. So we obtain an estimate $\frac{\partial^2 f_\varphi}{\partial \varphi \partial h} = O(h^{-2})$. The values of f_φ and $\frac{\partial f_\varphi}{\partial h}$ are determined only by the local behavior of f. If we estimate these functions near one separatrix loop, we may assume that $f = 0$ far from this separatrix loop. We have just estimated $\frac{\partial}{\partial \varphi}$ of these functions, so $f_\varphi = O(h^{-1})$, $\frac{\partial f_\varphi}{\partial h} = O(h^{-2})$, as in Table 1.

Finally, let us estimate $f_\varphi(h, 0)$. For $x, y > 0$ we have $t = \frac{1}{2a(\tilde{h})}(\ln x - \ln y)$, this is obtained by solving (9.1) with initial conditions $x = y = \tilde{h}^{1/2}$ for $t = 0$. For $\varphi = 0$ (and therefore $t = 0$, $x = y = \tilde{h}^{1/2}$) we have

$$\frac{\partial \varphi}{\partial x} = \frac{\partial}{\partial x}(\omega t) = \omega \frac{\partial t}{\partial x} = \frac{\omega}{2a(\tilde{h})x} = O(h^{-1/2} \ln^{-1} h).$$

Similarly, $\frac{\partial \varphi}{\partial y} = O(h^{-1/2} \ln^{-1} h)$, and $f_\varphi = f_x \frac{\partial \varphi}{\partial x} + f_y \frac{\partial \varphi}{\partial y} = O(h^{-1/2} \ln^{-1} h)$. Here f_x, f_y are the components of the vector field f written in the x, y chart, they are $O(1)$.

9.6 Estimate on $\overline{f}_{\varphi,1}$

It will be convenient to prove the estimate $\overline{f}_{\varphi,1} = O(h^{-1} \ln^{-3} h)$ here. As $\langle u_{h,1} \rangle_\varphi = 0$, by Lemma 3.1 we have

$$\overline{f}_{\varphi,1} = \langle f_\varphi \rangle_\varphi = \frac{1}{T} \int_0^T f_\varphi dt. \tag{9.11}$$

We split this integral into four disjoint integrals over the domains where the charts $\tilde{h}, \tilde{t}_i$ are defined. For definiteness we will consider only one such chart $\tilde{h}, \tilde{t}$ corresponding to the triangle $0 < y < x < 1 \subset \mathcal{S}$ (and continued outside $\mathcal{S}$ as discussed above). Denote by γ the part of our trajectory that is covered by this chart. We can rewrite the part of (9.11) covered by our chart as

$$\frac{1}{T} \int_\gamma \Big(\frac{\partial \varphi}{\partial \tilde{t}} f_{\tilde{t}} + \frac{\partial \varphi}{\partial \tilde{h}} f_{\tilde{h}} \Big) d\tilde{t}, \tag{9.12}$$

where $f_{\tilde{h}}, f_{\tilde{t}}$ are the components of f in the coordinates $\tilde{h}, \tilde{t}$. Note that $dt = d\tilde{t}$.

Denote by $\varphi_0(\tilde{h})$ the value of φ at the moment $\tilde{t} = 0$. We have $\varphi = \varphi_0 + \omega \tilde{t}$ and $\frac{\partial \varphi}{\partial \tilde{h}} = \frac{\partial \varphi_0}{\partial \tilde{h}} + \tilde{t} \frac{\partial \omega}{\partial \tilde{h}}$. We have $\frac{\varphi_0}{2\pi} = \frac{kS + s_1}{4S + s_2} = \frac{k}{4} + \frac{s_1 - k s_2/4}{4S + s_2}$. Here $k \in \mathbb{Z}$, S was defined in Section 9.3, and s_1, s_2 are smooth functions of $\tilde{h}$ corresponding to the time spend outside $\mathcal{S}$. Hence, $\frac{\partial \varphi_0}{\partial \tilde{h}} = O(h^{-1} \ln^{-2} h)$. We also have $\frac{\partial \varphi}{\partial \tilde{t}} = \omega = O(\ln^{-1} h)$, $\frac{\partial \omega}{\partial \tilde{h}} = O(h^{-1} \ln^{-2} h)$, $T = O(\ln h)$. Hence, (9.12) rewrites as

$$O(\ln^{-2} h) \int_\gamma |f_{\tilde{t}}| d\tilde{t} + O(h^{-1} \ln^{-3} h) \int_\gamma O(1 + |\tilde{t}|) |f_{\tilde{h}}| d\tilde{t}. \tag{9.13}$$

Split γ into parts that lie inside and outside $\mathcal{S}$: $\gamma = \gamma_{in} \cup \gamma_{out}$. Outside $\mathcal{S}$ the functions $\tilde{t}, \tilde{h}$ are smooth functions of p, q without singularities, so $f_{\tilde{h}}, f_{\tilde{t}} = O(1)$. Also, outside $\mathcal{S}$ we have $\tilde{t} = O(1)$, so the integral (9.12) over γ_{in} is $O(h^{-1} \ln^{-3} h)$.

Inside $\mathcal{S}$ we have

$$f_{\tilde{h}} = f_x \frac{\partial \tilde{h}}{\partial x} + f_y \frac{\partial \tilde{h}}{\partial y}, f_{\tilde{t}} = f_x \frac{\partial \tilde{t}}{\partial x} + f_y \frac{\partial \tilde{t}}{\partial y}.$$

Recall that inside $\mathcal{S}$ by (9.2) we have $\tilde{t} = O(\ln x)$, $\frac{\partial \tilde{t}}{\partial x} = O(x^{-1})$, $\frac{\partial \tilde{t}}{\partial y} = O(\ln x)$, $\tilde{h} = xy$. We also have $y = O(x)$, $f_x = O(1)$, $f_y = O(1)$. This gives the estimates

$$f_{\tilde{h}} = O(x), f_{\tilde{t}} = O(x^{-1}).$$

We have

$$\int_{\gamma_{in}} |f_{\tilde{t}}| d\tilde{t} = \int_{\gamma_{in}} O(x^{-1}) d\tilde{t} = \int_{\gamma_{in}} O(x^{-2}) dx = O(h^{-1/2}).$$

Here we used that on γ_{in} we have $x \geq \tilde{h}^{1/2}$. We also have

$$\int_{\gamma_{in}} O(1 + |\tilde{t}|) |f_{\tilde{h}}| d\tilde{t} = \int_{\gamma_{in}} O(1 + |\ln x|) dx = O(1).$$

Plugging this estimates in (9.13), we obtain that the integral (9.12) over γ_{in} is also $O(h^{-1} \ln^{-3} h)$. Hence, the part of integral (9.11) corresponding to the chart $\tilde{h}, \tilde{t}$ is $O(h^{-1} \ln^{-3} h)$. For other charts $\tilde{h}, \tilde{t}_i$ we have the same estimate, so $\overline{f}_{\varphi,1} = O(h^{-1} \ln^{-3} h)$.

10 Estimates related to the averaging chart

In this section we prove the estimates from Table 1 for the functions $u_{k,i}$ and $\overline{f}_{k,i}$. We will also prove the estimates for the functions $\hat{f}_{k,i}$. The following lemma allows to mass-produce estimates for $u_{k,i}$ and $\overline{f}_{k,i}$. However, these estimates are not always good, so we will estimate some of these functions differently.

Lemma 10.1. *Let functions $\overline{f}$, u be determined by the equations (3.3), (3.4) and the condition $\langle u \rangle_\varphi = 0$. Suppose that we have estimates for the function Y when $h \to 0$:*

$$Y = O(Y_0(h)), \ \frac{\partial Y}{\partial h} = O(Y_1(h)), \ \frac{\partial^2 Y}{\partial h^2} = O(Y_2(h)).$$

Then we have

1. $\overline{f} = O(Y_0)$, $\frac{\partial \overline{f}}{\partial h} = O(Y_1)$, $\frac{\partial^2 \overline{f}}{\partial h^2} = O(Y_2)$
2. $\frac{\partial u}{\partial \varphi}, u = O(Y_0 \ln h)$
3. $\frac{\partial^2 u}{\partial \varphi \partial h}, \ \frac{\partial u}{\partial h} = O(Y_1 \ln h) + O(Y_0 \cdot h^{-1})$
4. $\frac{\partial^3 u}{\partial \varphi \partial h^2}, \frac{\partial^2 u}{\partial h^2} = O(Y_2 \ln h) + O(Y_1 \cdot h^{-1}) + O(Y_0 \cdot h^{-2})$.

Proof. Item 1 follows from (3.3). Let us rewrite (3.4) as

$$\frac{\partial u}{\partial \varphi} = \frac{T}{2\pi}(Y - \overline{f}). \tag{10.1}$$

The first part of item 2 immediately follows from (10.1) and the estimates on T from Table 1, while the second part follows from the first part and the condition $\langle u \rangle_\varphi = 0$. The second parts of items 3 and 4 follow from the first parts in the same way. To get the first parts we differentiate (10.1) with respect to h once or twice, respectively, and use the estimates on T from Table 1. □

Let us now start with the function $\overline{f}_{h,1}$. We have $\overline{f}_{h,1} = \frac{\int_0^T f_h dt}{T}$. By [4, Corollary 3.1] we have $\int_0^T f_h dt = O(1)$, this gives the estimate for $\overline{f}_{h,1}$ itself. To estimate $\frac{\partial \overline{f}_{h,1}}{\partial h}$, we use the estimate $\frac{\partial}{\partial h}(\int_0^T f_h dt) = O(\ln h)$ ([4, Lemma 3.2]).

Let us estimate the function $u_{h,1}$, we will estimate the derivatives later. Since the saddle C is a critical point of H, the function $f_h = \frac{\partial h}{\partial q} f_q + \frac{\partial h}{\partial p} f_p$ vanishes at C. By [4, Lemma 3.2], it follows that $\oint_{H=h} |f_h| dt = O(1)$. This and (3.9) means that $u_{h,1} = O(1)$.

We use the formulas for $Y_{k,i}$ from (3.5). As $Y_{h,1} = f_h$, by Lemma 10.1 and the estimates for f_h from Table 1 we obtain the estimates for the derivatives of $u_{h,1}$. Next, we have $Y_{\varphi,1} = f_\varphi + \frac{\partial \omega}{\partial h} u_{h,1}$. From the estimates we have just obtained and the estimates on f_φ in Table 1 we get $Y_{\varphi,1} = O(h^{-1})$, $\frac{\partial Y_{\varphi,1}}{\partial h} = O(h^{-2})$. This gives us the estimates on the functions $\overline{f}_{\varphi,1}$ and $u_{\varphi,1}$ and their derivatives. However, for the function $\overline{f}_{\varphi,1}$ itself we have obtained a better estimate in Section 9.6.

To prove the estimate for $f_{h,2}$, we use (3.7). As $\langle u_{h,1} \rangle_\varphi = 0$, we can replace there $\operatorname{div}(f)$ with $\operatorname{div}(f) - \operatorname{div}(f)(C)$:

$$\overline{f}_{2,h} = \langle (\operatorname{div}(f) - \operatorname{div}(f)(C)) u_{h,1} \rangle_\varphi. \tag{10.2}$$

We have $\langle |\operatorname{div}(f) - \operatorname{div}(f)(C)| \rangle_\varphi = \frac{1}{T} \oint_{H=h} |\operatorname{div}(f) - \operatorname{div}(f)(C)| dt = \frac{O(1)}{T}$ by [4, Lemma 3.2]. As $u_{h,1} = O(1)$, by (10.2) this implies $\overline{f}_{2,h} = \frac{O(1)}{T} = O(\ln^{-1} h)$. By (9.10) we have $\frac{\partial \operatorname{div}(f)}{\partial h} = O(h^{-1})$. We also use the estimate $\frac{\partial u_{h,1}}{\partial h} = O(h^{-1} \ln h)$ obtained above. Then taking $\frac{\partial}{\partial h}$ of (10.2) we get $\frac{\partial \overline{f}_{2,h}}{\partial h} = O(h^{-1})$.

For the function $u_{h,2}$ we use Lemma 10.1 and estimate $Y_{h,2}$ given in (3.5) by using Table 1: $Y_{h,2} = O(h^{-1} \ln^2 h)$, $\frac{\partial Y_{h,2}}{\partial h} = O(h^{-2} \ln^2 h)$.

To estimate the functions $\overline{f}_{h,3}$ and $\overline{f}_{\varphi,2}$, we need to assume that

$$h > \varepsilon \ln^{10} \varepsilon. \tag{10.3}$$

By (8.1) we have the following system of equations:

$$(1 + \varepsilon \frac{\partial u_{\varphi,1}}{\partial \varphi}) \overline{f}_{\varphi,2} + \varepsilon^2 \frac{\partial u_{\varphi,1}}{\partial h} \overline{f}_{h,3} = A,$$
$$(1 + \varepsilon \frac{\partial u_{h,1,2}}{\partial h}) \overline{f}_{h,3} + \frac{\partial u_{h,1,2}}{\partial \varphi} \overline{f}_{\varphi,2} = B.$$

From (10.3), (3.1) and the estimates on $u_{h,1}$ and $u_{h,2}$ we have

$$\overline{h} = h(1 + o(1)).$$

This allows us to estimate the intermediate values from (8.1) as if they were at the point h. Using Table 1, we have $A = O(h^{-2} \ln h)$, $B = O(h^{-2} \ln^4 h)$. We can substitute the expression for $\overline{f}_{h,3}$ from the second equation into the first one. This yields

$$\overline{f}_{\varphi,2}\left(1 + \varepsilon \frac{\partial u_{\varphi,1}}{\partial \varphi} - \varepsilon^2 \frac{\partial u_{\varphi,1}}{\partial h} \frac{\partial u_{h,1,2}}{\partial \varphi} (1 + \varepsilon \frac{\partial u_{h,1,2}}{\partial h})^{-1}\right) = \\ = A - \varepsilon^2 \frac{\partial u_{\varphi,1}}{\partial h} (1 + \varepsilon \frac{\partial u_{h,1,2}}{\partial h})^{-1} B.$$

From (10.3) and Table 1 we see that $\frac{\partial u_{h,1,2}}{\partial \varphi} = O(\ln h)$ and $\varepsilon \frac{\partial u_{h,1,2}}{\partial h} = o(1)$. Hence, we have $\overline{f}_{\varphi,2} = O(h^{-2} \ln h)$. Then from the second equation we obtain $\overline{f}_{h,3} = O(h^{-2} \ln^4 h)$.

Lemma 10.2. *The estimates for the functions $\overline{f}_{h,1}, \overline{f}_{h,2}, \overline{f}_{h,3}, \overline{f}_{\varphi,1}, \overline{f}_{\varphi,2}$ and their derivatives stated in Table 1 also hold for the corresponding functions $\hat{f}_{h,1}, \hat{f}_{h,2}, \hat{f}_{h,3}, \hat{f}_{\varphi,1}, \hat{f}_{\varphi,2}$ and their derivatives. Moreover, we have $|u_{h,1}(h, \varphi, \varepsilon) - u^0_{h,1}(h, \varphi)| = O(\varepsilon)$.*

Proof. Recall that the expressions $\overline{f}^0_{*,*}$ are computed by the same formulas as $\overline{f}_{*,*}$, with the perturbation f replaced by f^0. This means that the estimates we have for $\overline{f}_{*,*}$ (they are valid for any smooth perturbation f) also hold for $\overline{f}^0_{*,*}$. By (4.3) we have $\hat{f}_{h,1} = \overline{f}^0_{h,1}$ and $\hat{f}_{\varphi,1} = \overline{f}^0_{\varphi,1}$, so for these expressions and their derivatives the lemma holds.

By (4.3) we also have $\hat{f}_{h,2} = \overline{f}^0_{h,2} + \langle f^1_h(h, \varphi) \rangle_\varphi$. Denote $\psi = \langle f^1_h(h, \varphi) \rangle_\varphi$. Similarly to the estimate on $\overline{f}_{h,1}$ above, we have $\psi = O(\ln^{-1} h)$ and $\frac{\partial \psi}{\partial h} = O(h^{-1} \ln^{-2} h)$. Therefore, the estimates for $\overline{f}_{h,2}$ and $\frac{\partial \overline{f}_{h,2}}{\partial h}$ from Table 1 also hold for $\hat{f}_{h,2}$.

We have $\hat{f}_{\varphi,2} = \overline{f}_{\varphi,2} + \varepsilon^{-1}(\overline{f}_{\varphi,1} - \overline{f}^0_{\varphi,1})$. Using (4.1), we get $\varepsilon^{-1}(\overline{f}_{\varphi,1} - \overline{f}^0_{\varphi,1}) = \langle f^1_\varphi + \varepsilon f^2_\varphi \rangle_\varphi$, where f^i_φ is the φ-component of f^i written in the energy-angle coordinates. As the estimate for $\overline{f}_{\varphi,1} = \langle f \rangle_\varphi$ holds for any smooth f, we can plug in $f^1_\varphi + \varepsilon f^2_\varphi$ instead of f and get the estimate $\langle f^1_\varphi + \varepsilon f^2_\varphi \rangle_\varphi = O(h^{-1} \ln^{-3} h)$. As f^2_p and f^2_q are uniformly bounded by a constant independent of ε, one may check that this estimate is uniform in ε. Therefore, the estimate for $\overline{f}_{\varphi,2}$ also holds for $\hat{f}_{\varphi,2}$.

Before estimating $\hat{f}_{h,3}$ we need to prove the second statement of the lemma. The map $\mathcal{U} : f \to u_{h,1}$ is linear by (3.8). Hence, $u_{h,1}(h, \varphi, \varepsilon) - u^0_{h,1}(h, \varphi) = \mathcal{U}(f(p, q, \varepsilon) - f^0(p, q)) = \varepsilon \mathcal{U}(f^1(p, q) + \varepsilon f^2(p, q, \varepsilon)) = O(\varepsilon)$. The last equality holds, because the estimate for $u_{h,1}$ gives that $\mathcal{U}(g) = O(1)$ for all smooth g. Again, this estimate is uniform in ε because f^2_p and f^2_q are uniformly bounded.

We have $\hat{f}_{h,3} = \overline{f}_{h,3} + \langle f^2_h \rangle_\varphi + \varepsilon^{-1}(\overline{f}_{h,2} - \overline{f}^0_{h,2})$. Clearly, $\langle f^2_h \rangle_\varphi = O(1)$. By (3.7) we have $\varepsilon^{-1}(\overline{f}_{h,2} - \overline{f}^0_{h,2}) = \varepsilon^{-1} \langle \operatorname{div} f \; u_{h,1} - \operatorname{div} f^0 \; u^0_{h,1} \rangle_\varphi$. As $u_{h,1} = u^0_{h,1} + O(\varepsilon)$ and $\operatorname{div} f = \operatorname{div} f^0 + O(\varepsilon)$, we get $\varepsilon^{-1}(\overline{f}_{h,2} - \overline{f}^0_{h,2}) = O(1)$, thus proving the estimate for $\hat{f}_{h,3}$. □

References

[1] F Jay Bourland, Richard Haberman, and William L Kath. "Averaging methods for the phase shift of arbitrarily perturbed strongly nonlinear oscillators with an application to capture". *SIAM Journal on Applied Mathematics* 51.4 (1991), pp. 1150–1167.

[2] John R Cary and Rex T Skodje. "Phase change between separatrix crossings". *Physica D: Nonlinear Phenomena* 36.3 (1989), pp. 287–316.

[3] Jürgen Moser. "The analytic invariants of an area-preserving mapping near a hyperbolic fixed point". *Communications on Pure and Applied Mathematics* 9.4 (1956), pp. 673–692.

[4] Anatoly Neishtadt. "Averaging method for systems with separatrix crossing". *Nonlinearity* 30.7 (2017), p. 2871.

[5] Anatoly Neishtadt. "Change of an adiabatic invariant at a separatrix". *Fizika plazmy* 12.8 (1986), p. 992. [Engl. transl.: Sov. J. Plasma Phys. 12, 568-573 (1986)].

[6] Anatoly Neishtadt and Alexei Vasiliev. "Phase change between separatrix crossings in slow–fast Hamiltonian systems". *Nonlinearity* 18.3 (2005), p. 1393.

Anatoly Neishtadt,
Department of Mathematical Sciences,
Loughborough University, Loughborough LE11 3TU, United Kingdom;
Space Research Institute, Moscow 117342, Russia
E-mail : a.neishtadt@lboro.ac.uk

Alexey Okunev,
Department of Mathematical Sciences,
Loughborough University, Loughborough LE11 3TU, United Kingdom
E-mail : a.okunev@lboro.ac.uk