Averaging method for systems with separatrix crossing

Anatoly Neishtadt[1,2]

[1] Loughborough University, Loughborough, LE11 3TU, UK
[2] Space Research Institute, Moscow, 117997, Russia

Abstract

The averaging method provides a powerful tool for studying evolution in near-integrable systems. Existence of separatrices in the phase space of the underlying integrable system is an obstacle for application of standard results that justify using of averaging. We establish estimates that allow to use averaging method when the underlying integrable system is a system with one rotating phase, and the evolution leads to separatrix crossings.

1 Introduction

An averaging method (see, e.g., [5]) is a powerful tool for study a long-term evolution in systems which are small perturbations of integrable systems. Many applications of this method are for one-frequency systems (also called systems with one fast rotating phase). In these cases in the phase space of the corresponding unperturbed system there is a domain foliated by closed trajectories - invariant circles. Averaging of perturbations over these circles provides a closed system for an approximate description of perturbed dynamics in this domain. Typically, in systems under consideration there are several such domains. These domains are bounded by surfaces on which this foliation has singularities. Classical results justifying averaging method [5] guarantee its applicability for description of evolution not too close to these separating surfaces. However, it is rather typical that evolution leads
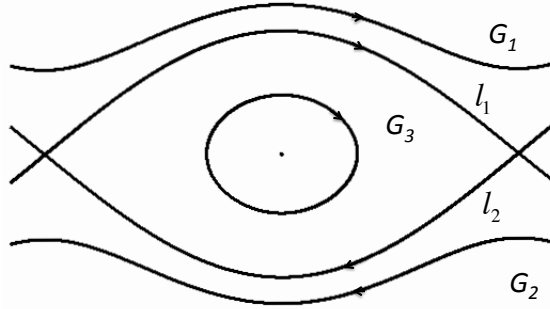
1

Figure 1: Phase portrait of a pendulum.

to crossing of these surfaces. The goal of this paper is to provide justification of a modified version of the averaging method for description of such an evolution.

A paradigmatic example of problems considered in this paper is a pendulum under the action of perturbations, e.g., of a small friction, a small constant torque, and a slow change of its length. An unperturbed pendulum could be in one of three regimes of motion: it could rotate in one or other direction, or oscillate. In the phase portrait of the pendulum these three regimes are demarcated by separatrices (Fig. 1). Motion of the pendulum evolves slowly under the action of perturbations. In the process of this evolution the pendulum can change the regime of its motion. In the phase plane the phase point crosses an instant separatrix of the unperturbed pendulum. Evolution of energy far from the separatrix can be described by the averaging method. Classical results justifying this method [5] are not applicable in the case of crossings of a separatrix. Moreover, this crossing leads to a remarkable probabilistic scattering. Initial data for different outcomes of separatrix crossings are mixed, and it is reasonable to consider each outcome as a random event with a definite probability. This probabilistic approach was first described in a similar problem in [21] and then independently in [14].

A natural way to describe evolution in the considered system is to use averaging method up to arrival to the separatrix, to calculate probabilities of capture into different domains at the separatrix, and to use the averaging method starting from the separatrix in the domain in which the system continues its evolution. In the current paper we justify such an approach for a rather wide class of one-frequency systems that change a qualitative

2

character of their motion in the process of evolution. The obtained estimates of the accuracy of the averaging method are sharp.

Part of results of this paper was announced (without proofs) in [26] (see also [4], Subsection 6.1.10) on the basis of estimates in [25].

2 Averaging method and averaging theorem for the separatrix crossing

In this section an averaging theorem is formulated that justifies the averaging method for description of the separatrix crossing. The proof is based on propositions given in Sections 3, 4, 5. There are probability phenomena due to separatrix crossing. Hence the recipe of the averaging method here includes calculations of the corresponding probabilities, and the averaging theorem justifies these calculations. All considerations are for systems of the form (2.1) below. We explain in the Appendix relation of this form to the general form of one-frequency systems with separatrix crossings.

2.1 Outline of the problem

We consider systems described by differential equations of the form

$$\dot{q} = \frac{\partial E}{\partial p} + \varepsilon f_1, \ \dot{p} = -\frac{\partial E}{\partial q} + \varepsilon f_2, \ \dot{z} = \varepsilon f_3 \,, \tag{2.1}$$

$$E = E(p, q, z), \ f_i = f_i(p, q, z, \varepsilon), i = 1, 2, 3, \ (p, q) \in \mathbb{R}^2, z \in \mathbb{R}^{l-2} \,.$$

Here $\varepsilon > 0$ is a small parameter characterising the rate of evolution. For $\varepsilon = 0$, $z = \text{const}$ we have *an unperturbed system* for p, q, which is a Hamiltonian system with one degree of freedom. The function E is an unperturbed Hamiltonian, and the functions εf_i are the perturbations. It is supposed, that there are separatrices in the phase portrait of the unperturbed system Fig. 2. In the course of evolution the projection of the phase point onto the plane (p, q) crosses a separatrix.

Far from the separatrices instead of (p, q) it is possible to use the variables $h = E$ and φ, where φ is "the angle" (from the pair "action-angle" variables [2] of the unperturbed system). Then for h, z, φ we get the perturbed system having the standard form of system with one rotating phase [5]: in this system h, z are called *slow variables*, φ is *the rotating phase*. It is a classical
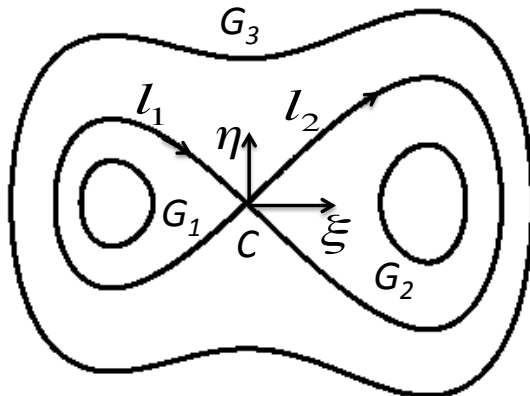
3

Figure 2: Phase portrait of the unperturbed system.

result that the averaged with respect to φ system describes the evolution of h, z far from separatrices with accuracy $O(\varepsilon)$ during the time interval of order $1/\varepsilon$. At the separatrices the frequency of the unperturbed motion vanishes, and also the equations in variables h, z, φ have singularities. In a region that includes a separatrix, the conditions of the classical theorem about accuracy of the averaging [5] fail and the applicability of the averaging method for the description of the evolution near the separatrices requires a justification.

Separatrix crossing leads to probability phenomena [1, 14, 15, 21]. As a simple example let as consider the motion of a particle in one dimension in double-well potential, Fig. 3a, perturbed by a small, of order ε, dissipation [1]. Phase portrait of the perturbed system is shown in Fig. 3b, where the initial conditions for the capture into the region, surrounded by the right separatrix loop, are shaded. The shaded strips far from the saddle have width of order ε and form a spiral with a step of order ε. Therefore small, of order ε, change of the initial conditions can change the result of evolution. As the initial conditions are always known only with some finite accuracy, the deterministic approach to the problem fails when $\varepsilon \to 0$. But it is possible to define in some natural way and to calculate the probabilities of capture into different regions after the separatrix crossing [1].

For systems of such types, a procedure of an approximate description of the evolution consists of using the averaged system up to the separatrix and calculation of the probability of capture into one or another region on the separatrix. It will be seen, that for majority of initial conditions this pro-
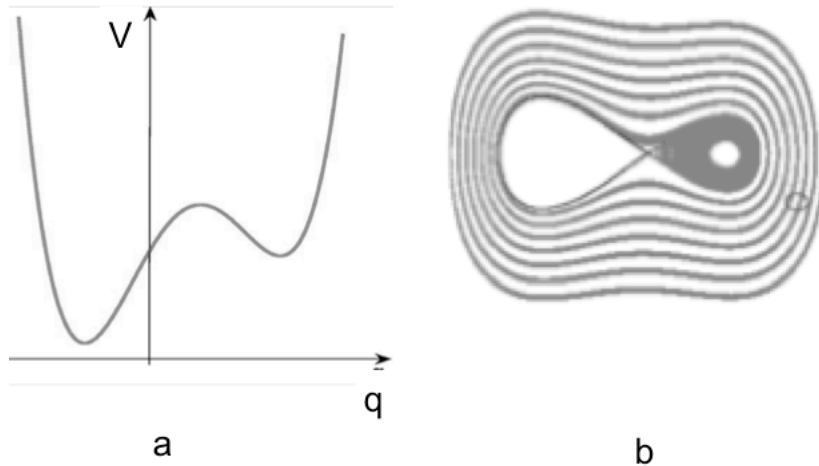
Figure 3: a) Double-well potential. b) Effect of a small friction.

cedure describes the behaviour of the slow variables with accuracy $O(\varepsilon \ln \varepsilon)$ during time of order $1/\varepsilon$. The measure of the "bad" set of initial conditions, for which this description is not valid, tends to 0 faster, than any given power of ε as $\varepsilon \to 0$. The general formula for the probability of the capture into one or another region (in the sense of definition in [1]) also will be proved.

2.2 Formulation of Hypotheses

System (2.1) is considered for $(p, q, z) \in D, |\varepsilon| < \varepsilon_1$, where D is a bounded domain in $\mathbb{R}^l$, $\varepsilon_1 = $ const. We denote B the projection of D onto z-space, and $G = G(z)$ the section of D by the two-dimensional plane $z = $ const. It is supposed that each domain $G(z) \subset \mathbb{R}^2$ is composed of the whole trajectories of the unperturbed system. It is supposed that the following assumptions are satisfied.

A. The function E is of smoothness C^3, and the functions f_i are of smoothness C^2 with respect to p, q, z. Functions f_i have one continuous derivative with respect to ε.

B. For $z \in B$ the phase portrait of the unperturbed system in the domain $G(z)$ has the form shown in Fig. 2. The unstable stationary point C is a non-degenerate saddle point. The separatrices $l_1 = l_1(z)$ and $l_2 = l_2(z)$ divide the unperturbed phase portrait into three regions $G_\nu = G_\nu(z)$, $\nu = 1, 2, 3$.

5

In what follows we assume that the Hamiltonian E is normalised in such a way, that $E = 0$ at the saddle point C, and therefore, on the separatrices. Then $E > 0$ in the region G_3, $E < 0$ in the regions G_1 and G_2. We denote $l_3 = l_1 \cup l_2$.

C. Introduce quantities

$$
\begin{aligned}
\Theta_\nu(z) &= -\oint_{l_\nu} \left(\frac{\partial E}{\partial q} f_1^0 + \frac{\partial E}{\partial p} f_2^0 + \frac{\partial E}{\partial z} f_3^0 \right) dt\,, \quad \nu = 1, 2, \qquad (2.2) \\
\Theta_3(z) &= \Theta_1(z) + \Theta_2(z)\,.
\end{aligned}
$$

Integrals in (2.2) are calculated along the unperturbed separatrices parametrized by the time t of the unperturbed motion; $f_i^0 = f_i(p, q, z, 0)$, $i = 1, 2, 3$. Integrals (2.2) are improper, because the motion along a separatrix takes infinite time. Our normalisation of E guarantees the convergence of the integrals as it is proved at the end of this section. We assume that the values Θ_ν, $\nu = 1, 2, 3$ are different from zero. In what follows, for certainty, the values Θ_ν, $\nu = 1, 2, 3$ are supposed to be positive.

Let us explain the meaning of the condition **C**. In the region G_ν for small $|E| > \varepsilon$, in the perturbed motion, a phase point makes rounds that are close to the unperturbed separatrix l_ν. The change of the value of E during one such round is close to the value $-\varepsilon\Theta_\nu$. Therefore, for phase points with small $|E|$, condition **C** ensures an approach the separatrix in the region G_3 and a departure from the separatrix in the regions G_1 and G_2. The convergence of integrals (2.2) is a corollary of the following assertion.

Lemma 2.1 *The first derivatives of the function E with respect to p, q, z vanish at the point C.*

Proof. The derivatives with respect to p, q vanish at the point C because the point C is an equilibrium position of the unperturbed system. Let $p_C(z), q_C(z)$ be coordinates of the point C. The Hamiltonian is normalised by the condition $E(p_C(z), q_C(z), z) \equiv 0$. Calculating the derivative of this equality with respect to z and taking into account that $\frac{\partial E}{\partial p} = \frac{\partial E}{\partial q} = 0$ at the point C, we get that $\frac{\partial E}{\partial z} = 0$ at the point C. $\qquad \square$

Lemma 2.1 implies that

$$
(\partial E/\partial \alpha)/\sqrt{(\partial E/\partial p)^2 + (\partial E/\partial q)^2}, \alpha = p, q, z,
$$

tend to finite limits as a point (p, q) tends to the point C along a separatrix. Let us use in the integrals (2.2) the arc length along the separatrix as a new independent variable. Then integrands do not have singularities, and therefore integrals (2.2) converge. Moreover, Θ_ν are smooth functions of z.

2.3   Averaged system

Let us define the averaged system separately for each region G_ν first. Let

$$\Sigma_\nu = \{(h, z) \colon z \in B, h = E(p, q, z), (p, q) \in G_\nu(z)\}, \ \nu = 1, 2, 3 \,.$$

The averaged in the region G_ν system is, by definition, the following system of differential equations in Σ_ν:

$$
\begin{aligned}
\dot{h} &= \frac{\varepsilon}{T} \oint_{E=h} \left(\frac{\partial E}{\partial q} f_1^0 + \frac{\partial E}{\partial p} f_2^0 + \frac{\partial E}{\partial z} f_3^0\right) dt \,, & (2.3)\\
\dot{z} &= \frac{\varepsilon}{T} \oint_{E=h} f_3^0 dt \,.
\end{aligned}
$$

Here integrals are calculated along the level line $E = h$ of the Hamiltonian situated in the domain $G_\nu(z)$. This level line is parametrized by the time t of the unperturbed motion along it, and $T = T(h, z)$ is the period of this motion. To write down this averaged system we calculate the rate of changing of $h = E(p, q, z)$ and z in the perturbed system, and then average the obtained expressions over t along the level line $E = h$ (for $\varepsilon = 0$ in the arguments of f_i). This averaging is equivalent to the averaging over the angular variable φ discussed in Subsection 2.1.

The period T grows proportionally to $-\ln|h|$ in the principal approximation as $h \to 0$ (see Lemma 3.3). When $h = 0$ it is reasonable to extent the definition of the right hand sides of (2.3) by continuity, putting

$$\dot{h} \mid_{h=0} = 0, \ \dot{z} \mid_{h=0} = \varepsilon f_{3C}^0 \,,$$

where f_{3C}^0 is the value of the function f_3^0 at the point C. Now we can combine three averaged systems in different regions into one "whole" averaged system. The phase space of the "whole" averaged system is a singular manifold, glued of three parts Σ_1, Σ_2 and Σ_3 along the set $\{h = 0\}$, Fig. 4. We will call the set $\{h = 0\}$ the separatrix for the averaged system.

According to condition **C**, $\dot{h} < 0$ in the averaged system for small $h \neq 0$ in all regions.
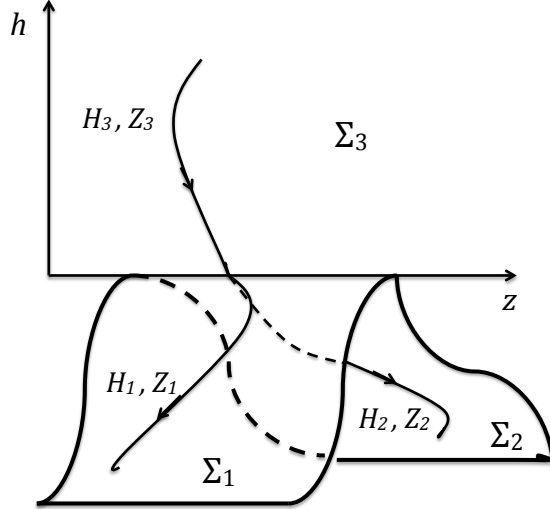
Figure 4: The phase space of the averaged system.

Definition. A solution $H(\tau)$, $Z(\tau)$, $\tau = \varepsilon t$, of the averaged system in G_ν such that $H(\tau) \to 0$, $Z(\tau) \to z_*$ as $\tau \to \tau_* - 0$ (for the region G_3) or $\tau \to \tau_* + 0$ (for the regions $G_{1,2}$) is called a *solution crossing the separatrix at the point z_* at the moment τ_**.

The following lemma is proved in Subsection 3.6.

Lemma 2.2 *a) For any $z_* \in B$, τ_*, and $\nu = 1, 2, 3$ there exists a unique solution of the averaged system in G_ν crossing the separatrix at the point z_* at the moment τ_*.*
b). If $z_0 \in B$, then for small enough $|h_0|$ the solution $H(\tau)$, $Z(\tau)$ of the averaged system with the initial conditions $H(\tau_0) = h_0, Z(\tau_0) = z_0$ crosses the separatrix at some moment $\tau = \tau_*$ ($\tau_* > \tau_0$, if $h_0 > 0$ and $\tau_* < \tau_0$ if $h_0 < 0$).

Take any point $(p_0, q_0, z_0) \in D$ such that $(p_0, q_0) \in G_3(z_0)$. Denote $h_0 = E(p_0, q_0, z_0)$. Consider the solution $(H_3(\tau), Z_3(\tau))$ of the averaged system in G_3 with the initial condition (h_0, z_0) at $\tau = 0$ (Fig. 4). Suppose that this solution crosses the separatrix at some $\tau = \tau_*$, i.e. $H_3(\tau_*) = 0$, $Z_3(\tau_*) = z_*$. According to Lemma 2.2, we can consider the solution $(H_\nu(\tau), Z_\nu(\tau))$ of the system averaged in G_ν, $\nu = 1, 2$ with the initial condition $(0, z_*)$ at $\tau = \tau_*$. This solution is well defined for τ close enough to τ_*, $\tau > \tau_*$. For an

approximate description of the behaviour of the values E, z in the perturbed system (2.1) we use the solution (H_3, Z_3) for $0 \le \varepsilon t \le \tau_*$, and the solution (H_ν, Z_ν) for $\varepsilon t > \tau_*$ with $\nu = 1$ or 2, if the phase point has been captured into the region G_ν after the separatrix crossing. We define for $\tau \le \tau_*$ the functions H_ν, Z_ν, $\nu = 1, 2$, by the relations $H_\nu(\tau) = H_3(\tau)$, $Z_\nu(\tau) = Z_3(\tau)$. These functions (H_ν, Z_ν) are called the solutions of the averaged system with the initial condition (h_0, z_0) at $t = 0$.

We attribute the probability $\Theta_\nu(z_*)/\Theta_3(z_*)$ to the capture into G_ν of the initial point (p_0, q_0, z_0). The meaning of this definition of the probability will be clear from the results of Subsection 2.4. The function P_ν defined by the formula

$$P_\nu(z) = \Theta_\nu(z)/\Theta_3(z) \tag{2.4}$$

will be called the probability of the capture into G_ν, $\nu = 1, 2$, at the separatrix.

We will need a lemma which allows to estimate a distance between two solutions of the averaged system with initial conditions near the separatrix.

Lemma 2.3 *Let two solutions of the averaged system, $(H_\nu(\tau), Z_\nu(\tau))$ and $(H'_\nu(\tau), Z'_\nu(\tau))$, $0 \le \tau \le K$, be given. Suppose that for some $\tau_0 \in [0, K]$ and some $\delta > 0$ these solutions satisfy the following condition:*

$$|H_\nu(\tau_0)| + |H'_\nu(\tau_0)| + |Z_\nu(\tau_0) - Z'_\nu(\tau_0)| < \delta.$$

If δ is small enough, then for $0 \le \tau \le K$ the following estimate is valid:

$$|H_\nu(\tau) - H'_\nu(\tau)| + |Z_\nu(\tau) - Z'_\nu(\tau)| < O\left(\delta + \frac{\delta|\ln \delta|}{1 + |\ln |H_\nu(\tau)||}\right).$$

The proof is given in Subsection 3.6.

In what follows, the action I of the unperturbed system will be important. The action $I = I(h, z)$ of the unperturbed trajectory $E = h$ in the region G_ν is the area enclosed by this trajectory, divided by 2π. We have [19]

$$\frac{\partial I}{\partial h} = \frac{1}{2\pi} T(h, z), \tag{2.5}$$

$$\frac{\partial I}{\partial z} = -\frac{1}{2\pi} \oint_{E=h} \frac{\partial E}{\partial z} dt. \tag{2.6}$$

9

With the aid of the formulas (2.5) and (2.6) the rate of change of I along a trajectory of the averaged system is found to be:

$$2\pi\frac{dI}{dt} = \varepsilon\oint_{E=h}(\frac{\partial E}{\partial q}f_1^0 + \frac{\partial E}{\partial p}f_2^0 + \frac{\partial E}{\partial z}f_3^0)dt$$
$$- \frac{\varepsilon}{T}\oint_{E=h}\frac{\partial E}{\partial z}dt\oint_{E=h}f_3^0 dt. \qquad (2.7)$$

A corollary of the formula (2.6) when $h \to 0$ is the following useful formula for the areas $S_{1,2} = S_{1,2}(z)$ of the regions $G_{1,2}$ and the area $S_3 = S_3(z)$ of the region $G_1 \cup G_2$:

$$\frac{\partial S_\nu}{\partial z} = -\oint_{l_\nu}\frac{\partial E}{\partial z}dt, \nu = 1, 2, 3. \qquad (2.8)$$

From (2.2) for the averaged system in the region G_ν we get:

$$\lim_{h\to 0}\frac{d}{dt}(2\pi I - S_\nu) = -\varepsilon\Theta_\nu(z), \ \nu = 1, 2, 3. \qquad (2.9)$$

A consequence of this equation is the above-mentioned property (see Lemma 2.2) that for solutions of the averaged system with small $|h|$ the arrival at the separatrix takes a finite time (in the region G_3 this time is positive, in the regions $G_{1,2}$ it is negative).

2.4 Estimates in the averaging method

Let a point $\hat{M}_0 = (\hat{p}_0, \hat{q}_0, \hat{z}_0)$ belong to the region D, and let $\hat{I}_0, \hat{\varphi}_0$ be the values of the action-angle variables I, φ at this point. The following sets are well defined and lie in D for small enough δ:

$$U^\delta = \{p, q, z : |z - \hat{z}_0| < \delta, |I - \hat{I}_0| < \delta, |\varphi - \hat{\varphi}_0| < \delta\}, \qquad (2.10)$$
$$W^\delta = \{p, q, z : |z - \hat{z}_0| < \delta, |I - \hat{I}_0| < \delta\}.$$

Denote $\hat{h}_0 = E(\hat{p}_0, \hat{q}_0, \hat{z}_0)$. We assume that solutions of the averaged system with initial data $(\hat{h}_0, \hat{z}_0)$ are well defined for $0 \le \tau \le K$ and cross the separatrix at some $\tau = \hat{\tau}_*$, $z = \hat{z}_*$. Denote:

- $(p(t), q(t), z(t))$ the solution of the perturbed system (2.1) with initial data $(p_0, q_0, z_0) \in U^\delta$ at $t = 0$,

- $h(t) = E(p(t), q(t), z(t))$ the value of the Hamiltonian along this solution,

10

- $h_0 = h(0)$,

- (H_ν, Z_ν), $\nu = 1, 2$, the solutions of the averaged system with the initial condition (h_0, z_0) at $\tau = 0$,

- τ_* the moment of separatrix crossing for the solutions (H_ν, Z_ν).

The value δ is supposed to be small enough so that the solutions (H_ν, Z_ν), $\nu = 1, 2$, are well defined for $0 \leq \tau \leq K$, and $\tau_* < K$. Fix any natural number $r \geq 2$. In what follows K_i (and afterwards k_i, c_i, d_i, ν_i) are positive constants, i.e. values independent of ε, δ and initial conditions $(p_0, q_0, z_0) \in W^\delta$. The appearance of K_i in some relation is equivalent to the assertion that there exists K_i satisfying this relation for small enough $\varepsilon > 0, \delta > 0, \varepsilon < \delta^2$ (and similarly for other constants). The following theorem summarises the principal features of the averaging method for separatrix-crossing orbits:

Theorem 1 *There exists a representation $U^\delta = U_1^\delta \cup U_2^\delta \cup v$ with the following properties.*
I. If $(p_0, q_0, z_0) \in U_\nu^\delta$, $\nu = 1, 2$, then the behaviour of E, z in the perturbed system is described approximately by the solution (H_ν, Z_ν) of the averaged system, and the following estimates hold

$$|h(t) - H_\nu(\varepsilon t)| + |z(t) - Z_\nu(\varepsilon t)| \; < \; K_1 \varepsilon \quad for \quad 0 \leq \varepsilon t \leq \tau_*, \tag{2.11}$$

$$|h(t) - H_\nu(\varepsilon t)| + |z(t) - Z_\nu(\varepsilon t)| \; < \; K_1 \varepsilon + \frac{K_2 \varepsilon |\ln \varepsilon|}{1 + |\ln |H_\nu(\varepsilon t)||} \quad for \quad \tau_* \leq \varepsilon t \leq K.$$

For $0 \leq \varepsilon t \leq \tau_* - K_3 \varepsilon |\ln \varepsilon|$ the point $(p(t), q(t))$ moves in the region $G_3(z(t))$, while for $\tau_* + K_3 \varepsilon |\ln \varepsilon| \leq \varepsilon t \leq K$ it moves in the region $G_\nu(z(t))$.

II.
$$\left| \frac{\mathrm{mes}\, U_\nu^\delta}{\mathrm{mes}\, U^\delta} - \frac{\Theta_\nu(\hat{z}_*)}{\Theta_3(\hat{z}_*)} \right| < K_4 \left(\delta + \frac{\varepsilon |\ln \varepsilon|}{\delta} \right).$$

III. $\mathrm{mes}\, v < k_5 \varepsilon^r \delta^{-1} \mathrm{mes}\, U^\delta$.

Here $\mathrm{mes}(\cdot)$ is the standard phase volume in $\mathbb{R}^l$.

This theorem will be proved by means of a series of propositions established in the following three Sections (Propositions 2.1, 2.2, 2.3).

It is natural to consider the relative measure of the set of points from a small neighbourhood of $\hat{M}_0$ that will be captured into the region $G_\nu, \nu = 1, 2$

11

for small ε as the value at the point $\hat{M}_0$ of the probability density of capture into G_ν. This approach is formalised as follows (cf. [1]).

Definition The value at the point $\hat{M}_0$ of *the probability density of capture into G_ν, $\nu = 1, 2$,* (or, for brevity, *the probability of capture of $\hat{M}_0$ into G_ν*) is

$$Q_\nu(\hat{M}_0) = \lim_{\delta \to 0} \lim_{\varepsilon \to 0} \frac{\mathrm{mes}\, U_\nu^\delta}{\mathrm{mes}\, U^\delta}\,. \tag{2.12}$$

Corollary 2.1 *The probability of capture of the point $\hat{M}_0$ into the region G_ν is given by the formula*

$$Q_\nu(\hat{M}_0) = \frac{\Theta_\nu(\hat{z}_*)}{\Theta_3(\hat{z}_*)},\ \ \nu = 1, 2\,. \tag{2.13}$$

By means of the function P_ν defined by equation (2.4) – the probability of the capture into G_ν at the separatrix – the last formula can be rewritten in the form

$$Q_\nu(\hat{M}_0) = P_\nu(\hat{z}_*),\ \ \nu = 1, 2\,.$$

Remarks

1. The formulation of the problem of separatrix-crossing has been described in the context of an eight-figure separatrix that is being approached by orbits of the perturbed and averaged systems. It will be clear from the nature of the estimates in the succeeding sections, that the basic results embodied in Theorem 1 can be carried over to the other geometric pictures of separatrix-crossing that are possible in $\mathbb{R}^2$. In fact, the phase space need not be $\mathbb{R}^2$. It could, for example, be a cylinder or a sphere. The important hypotheses are those demanding the non-degeneracy of the saddle equilibrium and the non-vanishing of the numbers Θ_ν, $\nu = 1, 2, 3$ (see Appendix). These conditions may be relaxed: they are needed only when the orbit approaches the separatrix. During the evolution prior to that time, these conditions are not needed. Metamorphoses of the phase portrait may take place as long as the phase point is far from separatrices when this happens. Situations wherein the crossing of a separatrix occurs for values of z for which $\Theta_\nu = 0$ or at which the non-degeneracy condition for the saddle fails are viewed as degenerate, and the estimates of the averaging method will in general be poorer in these cases. Examples of separatrix-crossing near a "newborn" saddle are considered in [11, 16, 18, 20].

2. The conclusion that the error in the use of the averaging method is $O(\varepsilon \ln \varepsilon)$, which follows from Theorem 1, cannot be improved. This follows from asymptotic formulas for change of the adiabatic invariant at a separatrix in Hamiltonian systems ([27], [9], [23], [24]) and also for systems perturbed by a weak dissipation [7].

3. The assertion of the Theorem 1 remains valid if the standard volume $\mathrm{mes}\,(\cdot)$ is replaced by any measure in $\mathbb{R}^l$ that has a smooth density, independent of ε, with respect to the standard volume. The formula for the probability given by (2.12) and (2.13) remains valid if U^δ is any domain with piecewise-smooth boundary having diameter δ (but in this case the estimate in the right hand side in part II of the Theorem 1 may not be satisfied).

4. For the example of motion of a particle in one dimension in double-well potential perturbed by a small dissipation the formula for probability is given in [1]. The proof is contained in [8].

5. A different approach for introducing probability was suggested in [13, 28]. White noise of order $\varepsilon\delta$ was added to the right-hand side of equations (2.1), in the case when the parameter z is absent from the problem. In this problem, capture into one or another region becomes a genuinely random event. Again taking the limit of the probability of capture as $\varepsilon \to 0$ (first) and $\delta \to 0$, one recovers the same formula for the probability as that found in Subsection 2.4 above.

6. D.V.Anosov has suggested yet another approach for introducing probability in the considered problem[1]. Denote $\varkappa_\nu(\varepsilon_0)$ the measure of the set of values $\varepsilon \in (0, \varepsilon_0]$ such that $\hat{M}_0 \in U_\nu^\delta$ for these values of ε, $\nu = 1, 2$. Then we define the probability of capture of $\hat{M}_0$ into G_ν as $\lim_{\varepsilon_0 \to 0} \varkappa_\nu(\varepsilon_0)/\varepsilon_0$. This definition was not discussed in a literature. One can show that, e.g. for the case when $f_1 = f_2 = 0$, $f_3 = \mathrm{const}$ in equations (2.1) this probability is again given by formula (2.13). It looks plausible that this is the case for the general form of system (2.1) as well.

7. An important open question is when results of consecutive crossings of separatrices can be considered as statistically independent. This is not the case, for example, if (2.1) is a Hamiltonian system with slowly varying parameter, and $\Theta_1 \equiv \Theta_2$ [10]. However, it looks as a reasonable hypothesis that typically there is such an independence.

---

[1]This was a comment in a meeting of the Moscow Mathematical Society.

2.5 Derivation of the estimates in the averaging method

In this subsection Theorem 1 is derived from several principal Propositions describing approach the separatrix, passage through its narrow neighbourhood, and departure from the separatrix. These Propositions are proved in Sections 3, 4, 5. Approach the separatrix is described by the following assertion.

Proposition 2.1 *For all initial conditions $(p_0, q_0, z_0) \in W^\delta$ and for $t > 0$ while $H_3(\varepsilon t) \geq k_1 \varepsilon$ the following estimates are valid:*

$$|h(t) - H_3(\varepsilon t)| + |z(t) - Z_3(\varepsilon t)| = O(\varepsilon), \quad \frac{1}{2} H_3(\varepsilon t) \leq h(t) \leq 2 H_3(\varepsilon t).$$

Departure from the separatrix in the regions $G_{1,2}$ is described by an analogous assertion.

Proposition 2.2 *Let a moment of time $t' \in [0, K/\varepsilon]$ exist such that $h(t') = -k_2 \varepsilon$, $(p(t'), q(t')) \in G_\nu(z(t'))$, $\nu = 1, 2$, $|\varepsilon t' - \tau_*| + |Z(t') - Z_3(\tau_*)| < k_3^{-1}$. Then for $t' \leq t \leq K/\varepsilon$ the behaviour of E, z is approximately described by the solution (H'_ν, Z'_ν) of the averaged in G_ν system with the initial condition $h(t'), z(t')$ at $\tau = \varepsilon t'$ as follows:*

$$|h(t) - H'_\nu(\varepsilon t)| + |z(t) - Z'_\nu(\varepsilon t)| = O\left(\frac{\varepsilon \ln \varepsilon}{1 + |\ln |H'_\nu(\varepsilon t)||}\right), \quad 2H'_\nu(\varepsilon t) \leq h(t) \leq \frac{1}{2} H'_\nu(\varepsilon t)$$

It follows from Proposition 2.1 that there exists a moment of time $t_- = t_-(p_0, q_0, z_0, \varepsilon)$ such that at this moment $h(t) = 2k_1 \varepsilon$ for the first time. Let $t_+ = t_+(p_0, q_0, z_0, \varepsilon)$ be the moment of time such that at this moment $h(t) = -k_2 \varepsilon$ for the first time at the segment $[0, K/\varepsilon]$. The moment of time t_+ is defined, in general, not for all initial conditions. The following assertion describes passage through a narrow neighbourhood of the separatrices during the interval of time $t_- \leq t \leq t_+$.

Proposition 2.3 *There exists a representation $W^\delta = W_3^\delta \cup w$ such that $\mathrm{mes}\, w = O(\varepsilon^r) \mathrm{mes}\, W^\delta$ and for $(p_0, q_0, z_0) \in W_3^\delta$ the moment of time t_+ is well defined and $t_+ = t_- + O(\ln \varepsilon)$. For $t_- \leq t \leq t_+$ and for $\nu = 1, 2$ the following estimate holds:*

$$|h(t)| + |H_\nu(\varepsilon t)| + |z(t) - Z_\nu(\varepsilon t)| = O(\varepsilon).$$

14

Let W_ν^δ, $\nu = 1, 2$, be the sets of points $(p_0, q_0, z_0) \in W_3^\delta$ such that $(p(t_+), q(t_+)) \in G_\nu(z(t_+))$.

Proposition 2.4 *The measure of the set W_ν^δ is estimated as*

$$\text{mes } W_\nu^\delta = \int_{W^\delta} P_\nu(Z_3(\tau_*)) dp_0 dq_0 dz_0 + O(\varepsilon \ln \varepsilon \, \delta^{-1} \text{ mes } W^\delta), \quad \nu = 1, 2 \,.$$

Here $(H_3(\tau), Z_3(\tau))$ is the solution of the averaged system with the initial condition $H_3(0) = E(p_0, q_0, z_0), Z_3(0) = z_0$, and τ_* is the moment of separatrix crossing for this solution, dz_0 is the standart volume element in $\mathbb{R}^{l-2}$.

Denote $U_\nu^\delta = U^\delta \cap W_\nu^\delta$, $v = U^\delta \cap w$.

Proposition 2.5 *The measure of the set U_ν^δ meets the estimate of part II in Theorem 1.*

For $0 \le t \le t_+$ the estimates in Theorem 1 hold due to the estimates of Propositions 2.2 and 2.3. From these estimates we get also that it is possible to choose t_+ as t' in Propositions 2.2. Then at $t = t_+$ the distance between solutions (H_ν', Z_ν') and (H_ν, Z_ν) is $O(\varepsilon)$. Therefore from Lemma 2.3 and Proposition 2.2 we get the estimates in Theorem 1 for $t_+ \le t \le K/\varepsilon$. This completes the proof of Theorem 1.

To proof Corollary 2.1 it is enough to consider a cover of U^δ by the union of sets

$$\begin{aligned} U_{j_1, j_2, j_3}^\kappa &= \{p, q, z \colon |z - \hat{z}_{j_1}| \le \kappa, |I - \hat{I}_{j_2}| \le \kappa, |\varphi - \hat{\varphi}_{j_3}| \le \kappa\}, \quad (2.14) \\ \hat{z}_{j_1} &= \hat{z}_0 + 2\kappa j_1, \; \hat{I}_{j_2} = \hat{I}_0 + 2\kappa j_2, \; \hat{\varphi}_{j_3} = \hat{\varphi}_0 + 2\kappa j_3 \,. \end{aligned}$$

Here j_2, j_3 are integer numbers, and j_1 is an integer $(l-2)$-dimensional vector. Then one should apply Theorem 1 to those of sets U_{j_1, j_2, j_3}^κ, which intersect U^δ. Now proceed to the limit first as $\varepsilon \to 0$ and then as $\kappa \to 0$. We get the formula

$$\lim_{\varepsilon \to 0} \text{mes } U_\nu^\delta = \int_{U^\delta} P_\nu(Z_3(\tau_*)) dp_0 dq_0 dz_0$$

which implies Corollary 2.1.

3 Estimates of the accuracy of the averaging method up to separatrix

In this section the proof of Proposition 2.1 is given. The proof of Proposition 2.2 is entirely analogous, and only a sketch of it is given. We use several lemmas, which are formulated in Subsections 3.1, 3.2. This lemmas are proven in Subsections 3.4, 3.5.

Only motion in the region G_3 is considered in the principal part of this section. Thus we will omit index $\nu = 3$ at solution (H_ν, Z_ν). The following notation is used: $I(h, z)$ - the value of the "action" variable for the trajectory $E = h$ of the unperturbed system, $j(t) = I(h(t), z(t)), J(\varepsilon t) = I(H(\varepsilon t), Z(\varepsilon t))$.

Below $z \in B - c_1^{-1}$, $(p, q, z) \in D - c_1^{-1}$ in all estimates, and the constant c_1 is chosen in such a way that $3c_1^{-1}$-neighbourhood of the set $\{z \colon z = Z_\nu(\tau), \nu = 1, 2; \tau \in [0, K]\}$ belongs to B, and $3c_1^{-1}$-neighbourhood of the set

$$\{(p, q, z) \colon z = Z_\nu(\tau), E(p, q, z) = H_\nu(\tau), \ \nu = 1, 2; \ \tau \in [0, K]\}$$

belongs to D.

3.1 Lemmas on unperturbed motion

Let $\varphi = \varphi(p, q, z, \varepsilon)$ and $\psi = \psi(p, q, z, \varepsilon)$ be smooth functions, and φ vanish at the saddle point C identically with respect to z, ε. Let $0 < h < 1/2, z \in B - c_1^{-1}$.

Lemma 3.1
$$\oint_{E=h} \varphi dt = \oint_{l_3} \varphi dt + O(h \ln h). \tag{3.1}$$

Corollary 3.1 *For $0 < h < c_2^{-1}$ the following estimate is valid:*

$$\oint_{E=h} \left(\frac{\partial E}{\partial q} f_1^0 + \frac{\partial E}{\partial p} f_2^0 + \frac{\partial E}{\partial z} f_3^0 \right) dt = -\Theta_3 + O(h \ln h) < -c_3^{-1}. \tag{3.2}$$

16

Lemma 3.2

$$(1) \oint_{E=h} |\varphi| dt = O(1); \qquad (2) \oint_{E=h} |\psi| dt = O(\ln h); \qquad (3.3)$$

$$(3) \frac{\partial}{\partial h} \oint_{E=h} \psi dt = O(h^{-1}); \quad (4) \frac{\partial}{\partial z} \oint_{E=h} \psi dt = O(\ln h);$$

$$(5) \frac{\partial}{\partial h} \oint_{E=h} \varphi dt = O(\ln h); \quad (6) \frac{\partial}{\partial z} \oint_{E=h} \varphi dt = O(1).$$

Corollary 3.2

$$\frac{\partial I}{\partial h} = \frac{1}{2\pi} T = O(\ln h), \quad \frac{\partial I}{\partial z} = -\frac{1}{2\pi} \oint_{E=h} \frac{\partial E}{\partial z} = O(1), \qquad (3.4)$$

$$\frac{\partial^2 I}{\partial h^2} = O(h^{-1}), \quad \frac{\partial^2 I}{\partial h \partial z} = O(\ln h), \quad \frac{\partial^2 I}{\partial z^2} = O(1).$$

Lemma 3.3 $T = -2a \ln h + b_3 + O(h \ln h)$, where $a = a(z), b_3 = b_3(z), a = 1/\omega_0$, and $\omega_0 > 0$ is the eigenvalue of the saddle point C.

Corollary 3.3 $T > c_4^{-1} |\ln h|$.

Remark. For the period T_i of the unperturbed trajectory $E = h < 0$ in the region $G_i, i = 1, 2$, the following expansion is valid:

$$T_i = -a \ln |h| + b_i + O(h \ln |h|),$$

where $b_3 = b_1 + b_2$.

Lemma 3.4

$$\frac{\partial}{\partial h} \frac{1}{T} \oint_{E=h} \psi dt = O(h^{-1} \ln^{-2} h). \qquad (3.5)$$

3.2 Lemmas on perturbed motion

Let $C\xi\eta = C\xi\eta(z)$ be the system of principal axes for the saddle point C, oriented as it is shown in Fig. 2.

Lemma 3.5 *Let at a moment of time t' the point $(p(t'), q(t'))$ lie on the axis $C\eta(z(t'))$ in c_5^{-1}-neighbourhood of the point C, and $c_6\varepsilon \le h(t') \le c_7^{-1} \le c_2^{-1}$, $z(t') \in B - 2c_1^{-1}$. Then there exists a moment of time $t'' = t' + O(\ln h(t'))$ such that a) for $t' \le t \le t''$ the solution $(p(t), q(t), z(t))$ is well defined, and*

17

the point $(p(t''), q(t''))$ lies on the axis $C\eta(z(t''))$ in c_5^{-1}-neighbourhood of the point C; b) for $t' \le t \le t''$ the following estimates are satisfied:

$$h(t) = h(t') + O(\varepsilon), \quad \frac{1}{2}h(t') \le h(t) \le 2h(t'),$$

$$|z(t) - z(t')| + |j(t) - j(t')| = O(\varepsilon \ln h(t')); \tag{3.6}$$

c) integrals of functions ψ, φ (see Section 3.1) along the trajectory are estimated as follows:

$$\int_{t'}^{t''} \psi dt = \oint_{E=h(t')} \psi dt + \varepsilon O(h^{-1}(t')),$$

$$\int_{t'}^{t''} \varphi dt = \oint_{E=h(t')} \varphi dt + \varepsilon O(h^{-1/2}(t')). \tag{3.7}$$

Here integrals in the right hand side are calculated along the unperturbed trajectory for $z = z(t')$, $\varepsilon = 0$.

Corollary 3.4

$$t'' - t' = \int_{t'}^{t''} dt = T(h(t'), z(t')) + \varepsilon O(h^{-1}(t')), \tag{3.8}$$

$$h(t') - h(t'') = -\varepsilon \int_{t'}^{t''} \left(\frac{\partial E}{\partial q} f_1 + \frac{\partial E}{\partial p} f_2 + \frac{\partial E}{\partial z} f_3 \right) dt$$

$$= -\varepsilon \oint_{E=h(t')} \left(\frac{\partial E}{\partial q} f_1^0 + \frac{\partial E}{\partial p} f_2^0 + \frac{\partial E}{\partial z} f_3^0 \right) dt + \varepsilon^2 O(h^{-1/2}(t')) > \frac{1}{2} c_3^{-1} \varepsilon.$$

Lemma 3.6 *Let at a moment of time t' the conditions $c_6\varepsilon \le h(t') \le c_7^{-1}$, $z(t') \in B - 2c_1^{-1}$ be satisfied. Then there exists a moment of time $t'' = t' + O(\ln h(t'))$ such that for $t' \le t \le t''$ the solution $(p(t), q(t), z(t))$ is well defined, it meets estimates (3.6), and the point $(p(t''), q(t''))$ lies on the axis $C\eta(z(t''))$ in c_5^{-1}-neighbourhood of the point C.*

3.3 Proof of Proposition 2.1

I. Let t_* be the maximal moment of time at the interval $[0, K/\varepsilon]$ such that for $0 \le t \le t_*$ the following estimates hold:

$$(p(t), q(t), z(t)) \in D - c_1^{-1}, \quad h(t) \ge \frac{1}{2} c_7^{-1}.$$

For $0 \leq t \leq t_*$ the frequency of the unperturbed motion is separated from 0 by a positive constant. Therefore, usual estimates of the averaging method for one-frequency systems [5] are valid:

$$|h(t) - H(\varepsilon t)| + |z(t) - Z(\varepsilon t)| + |j(t) - J(\varepsilon t)| = O(\varepsilon). \qquad (3.9)$$

Thus, for $0 \leq t \leq t^*$ estimates of Proposition 2.1 hold, and $h(t_*) = \frac{1}{2}c_7^{-1}$.

The moment t_* meets conditions of Lemma 3.6. Due to this lemma a moment of time $t_1 = t_* + O(1)$ is defined such that for $t_* \leq t \leq t_1$ the estimate (3.9) holds, and the point $(p(t_1), q(t_1))$ lies on the axis $C\eta(z(t_1))$ in c_5^{-1}-neighbourhood of the saddle point C.

II. Let t_{**} be the maximal moment of time in the interval $[t_1, K/\varepsilon]$ such that for $t_1 \leq t \leq t_{**}$ the following estimates hold:

$$(p(t), q(t), z(t)) \in D - c_1^{-1}, \ h(t) \geq c_6\varepsilon, \ \frac{1}{2}H(\varepsilon t) \leq h(t) \leq 2H(\varepsilon t). \quad (3.10)$$

Lemma 3.5 allows to define moments of time $t_2, ..., t_{n^*}$ of consecutive arrivals of the point $(p(t), q(t))$ at the ray $C\eta$, where n^* is the maximal number n such that $t_n < t_{**}$. Denote $h_n = h(t_n)$ and, analogously, z_n, j_n, H_n, Z_n, J_n. From (3.6) we have that for $t_n \leq t \leq t_{n+1}$ the following estimates are satisfied:

$$|h(t) - h_n| = O(\varepsilon), \ \frac{1}{2}h_n \leq h(t) \leq 2h_n,$$
$$|z(t) - z_n| + |j(t) - j_n| = O(\varepsilon \ln h_n). \qquad (3.11)$$

Lemma 3.2 allows to estimate the right hand sides of the averaged system:

$$\dot{H} = O(\varepsilon/\ln H), \ \dot{Z} = O(\varepsilon), \ \dot{J} = O(\varepsilon).$$

Considering motion in the averaged system for $t_n \leq t \leq t_{n+1}$ and making use of estimates (3.6), (3.10), we get for $t_n \leq t \leq t_{n+1}$

$$\begin{aligned} |H(\varepsilon t) \ - \ H_n| &= O(\varepsilon), \\ |Z(\varepsilon t) \ - \ Z_n| + |J(\varepsilon t) - J_n| &= O(\varepsilon \ln h_n). \end{aligned} \qquad (3.12)$$

If $n = n^*$, then (3.11), (3.12) hold for $t_{n^*} \leq t \leq t_{**}$.

From formulas

$$\begin{aligned} j_{n+1} - j_n &= \varepsilon \int_{t_n}^{t_{n+1}} \left(\frac{\partial I}{\partial h} \left(\frac{\partial E}{\partial q} f_1 + \frac{\partial E}{\partial p} f_2 + \frac{\partial E}{\partial z} f_3 \right) + \frac{\partial I}{\partial z} f_3 \right) dt, \\ z_{n+1} - z_n &= \varepsilon \int_{t_n}^{t_{n+1}} f_3 \, dt, \end{aligned}$$

19

making use of (3.3), (3.4), (3.7), (3.11), we get

$$j_{n+1} = j_n + \varepsilon T(h_n, z_n)F(h_n, z_n) + \varepsilon^2 O(h_n^{-1}),$$
$$z_{n+1} = z_n + \varepsilon T(h_n, z_n)\Phi(h_n, z_n) + \varepsilon^2 O(h_n^{-1}),$$

where $\varepsilon F(h, z)$ and $\varepsilon \Phi(h, z)$ are right hand sides of the averaged equations for I and z respectively:

$$2\pi F(h, z) = \oint_{E=h} \left(\frac{\partial E}{\partial q} f_1^0 + \frac{\partial E}{\partial p} f_2^0 + \frac{\partial E}{\partial z} f_3^0 \right) dt - \frac{1}{T} \oint_{E=h} \frac{\partial E}{\partial z} dt \cdot \oint_{E=h} f_3^0 \, dt,$$

$$\Phi(h, z) = \frac{1}{T} \oint_{E=h} f_3^0 \, dt.$$

In the averaged system

$$\dot{J} = \varepsilon F(H, Z), \quad \dot{Z} = \varepsilon \Phi(H, Z).$$

Therefore, making use of (3.3), (3.4), (3.6), (3.12), we get

$$
\begin{aligned}
J_{n+1} - J_n &= \varepsilon \int_{t_n}^{t_{n+1}} F(H, Z)dt = \varepsilon \int_{t_n}^{t_{n+1}} F(H_n, Z_n)dt + \varepsilon^2 O\left(\frac{h_n^{-1}}{\ln h_n} \right) \\
&= \varepsilon T(h_n, z_n)F(H_n, Z_n) + \varepsilon^2 O(h_n^{-1}), \\
Z_{n+1} - Z_n &= \varepsilon \int_{t_n}^{t_{n+1}} \Phi(H, Z)dt = \varepsilon \int_{t_n}^{t_{n+1}} \Phi(H_n, Z_n)dt + \varepsilon^2 O\left(\frac{h_n^{-1}}{\ln h_n} \right) \\
&= \varepsilon T(h_n, z_n)\Phi(H_n, Z_n) + \varepsilon^2 O(h_n^{-1}).
\end{aligned}
$$

So

$$J_{n+1} = J_n + \varepsilon T(h_n, z_n)F(H_n, Z_n) + \varepsilon^2 O(h_n^{-1}),$$
$$Z_{n+1} = Z_n + \varepsilon T(h_n, z_n)\Phi(H_n, Z_n) + \varepsilon^2 O(h_n^{-1}).$$

Denote

$$u_n = |j_n - J_n| + |z_n - Z_n|.$$

From the previous estimates by means of (3.3), (3.5), we get

$$u_{n+1} \leq u_n + |h_n - H_n|\varepsilon O(h_n^{-1} \ln^{-1} h_n) \qquad (3.13)$$
$$+|z_n - Z_n|\varepsilon O(\ln h_n) + \varepsilon^2 O(h_n^{-1}).$$

We have

$$2\pi \partial I/\partial h = T > c_4^{-1}|\ln h|, \ \partial I/\partial z = O(1).$$

Therefore, solving relation $I = I(h, z)$ for h we get h as a function of I, z for which

$$\frac{\partial h}{\partial I} = \frac{2\pi}{T} = O(\ln^{-1} h), \ \frac{\partial h}{\partial z} = -\frac{\partial I/\partial z}{\partial I/\partial h} = O(\ln^{-1} h).$$

Therefore, by Lagrange's formula,

$$h_n - H_n = \left(\frac{\partial h}{\partial I}\right)_* (j_n - J_n) + \left(\frac{\partial h}{\partial z}\right)_* (z_n - Z_n) =$$
$$= (|j_n - J_n| + |z_n - Z_n|)\, O(\ln^{-1} h_n).$$

Symbol " $*$ " here means that derivatives are calculated at some point in the straight line interval with endpoints (j_n, z_n) and (J_n, Z_n). Using this estimate in (3.13), we get

$$u_{n+1} \leq \left[1 + \varepsilon O(h_n^{-1}\ln^{-2} h_n)\right] u_n + \varepsilon^2 O(h_n^{-1}). \tag{3.14}$$

Consecutive use of this relation gives

$$u_{n+1} \leq \left[\prod_{s=1}^{n}(1 + \varepsilon|O(h_s^{-1}\ln^{-2} h_s)|\right] \left(u_1 + \varepsilon^2 \sum_{s=1}^{n} |O(h_s^{-1})|\right).$$

In accordance with (3.8), $h_s - h_{s+1} > \frac{1}{2}c_3^{-1}\varepsilon$. Therefore

$$\sum_{s=1}^{n} \varepsilon^2|O(h_s^{-1})| \leq \varepsilon^2 c_8 \sum_{s=1}^{n} h_s^{-1} \leq 2c_8 c_3 \varepsilon \sum_{s=1}^{n} h_s^{-1}(h_s - h_{s+1})$$

$$< \varepsilon c_9 \int_{h_{n+1}}^{h_1} h^{-1}dh = \varepsilon c_9 \ln(h_1/h_{n+1}) = \varepsilon O(\ln h_{n+1}).$$

In an analogous way, from the convergence of $\int_0^{1/2} h^{-1}\ln^{-2} h\, dh$, we get

$$\prod_{s=1}^{n} \left(1 + \varepsilon|O(h_s^{-1}\ln^{-2} h_s)|\right) = \exp\left[\sum_{s=1}^{n} \ln(1 + \varepsilon|O(h_s^{-1}\ln^{-2} h_s)|\right]$$

$$\leq \exp(\varepsilon \sum_{s=1}^{n} O(h_s^{-1}\ln^{-2} h_s)) = O(1).$$

21

Using these estimates in (3.14) and taking into account that $u_1 = O(\varepsilon)$, we get for $n = 1, ..., n^* - 1$:

$$u_{n+1} = |j_{n+1} - J_{n+1}| + |z_{n+1} - Z_{n+1}| = O(\varepsilon \ln h_{n+1}).$$

From here we get with the help of (3.10), (3.11), (3.12) for $t_1 \leq t \leq t_{**}$:

$$\begin{aligned}
|j(t) - J(\varepsilon t)| &= O(\varepsilon \ln H(\varepsilon t)), && (3.15) \\
|z(t) - Z(\varepsilon t)| &= O(\varepsilon \ln H(\varepsilon t)), \quad |h(t) - H(\varepsilon t)| < c_{10}\varepsilon.
\end{aligned}$$

Let us choose $k_1 = 2c_{10} + c_6$. While $H(\varepsilon t) \geq k_1 \varepsilon$, the condition (3.10) can not be violated. Hence, the estimates in Proposition 2.1 for $h(t)$ are proved.

III. The estimate for $z(t)$ in (3.15) is less accurate, than in formulation of Proposition 2.1. Now we will improve this estimate. Denote

$$\begin{aligned}
(\Delta z)_n &= z_n - Z_n, \quad (\Delta h)_n = h_n - H_n, \\
a_n &= -\oint_{E=h_n} \left(\frac{\partial E}{\partial q} f_1^0 + \frac{\partial E}{\partial p} f_2^0 + \frac{\partial E}{\partial z} f_3^0 \right) dt, \quad b_n = \oint_{E=h_n} (f_3^0 - f_{3C}^0) dt.
\end{aligned}$$

Here $f_{3C}^0 = f_{3C}(z)$ is the value of the function f_3^0 at the saddle point C. The integrals are calculated for $z = z_n$. We have an identity

$$\begin{aligned}
\dot{z} - \dot{Z} &= \varepsilon[(f_{3C}^0(z) - f_{3C}^0(Z)) + (f_3^0(p, q, z) - f_{3C}^0(z)) && (3.16) \\
&\quad - \frac{1}{T(H, Z)} \oint_{E=H} (f_3^0(p, q, Z) - f_{3C}^0(Z)) dt + (f_3(p, q, z, \varepsilon) - (f_3^0(p, q, z)))].
\end{aligned}$$

Let us integrate both sides of this identity with respect to t from t_n to t_{n+1} and estimate the right hand side making use of already established estimates (3.15) and Lemmas 3.2, 3.3, 3.5. In the left hand side we have $(\Delta z)_{n+1} - (\Delta z)_n$. Terms in the right hand side are

$$\begin{aligned}
\int_{t_n}^{t_{n+1}} (f_{3C}^0(z) - f_{3C}^0(Z)) dt &= O(\varepsilon \ln^2 h_n), \\
\int_{t_n}^{t_{n+1}} (f_3^0(p, q, z) - f_{3C}^0(z)) dt &= b_n + O(\varepsilon h_n^{-1/2}), \\
\int_{t_n}^{t_{n+1}} \frac{dt}{T(H, Z)} \left(\oint_{E=H} (f_3^0(p, q, Z) - f_{3C}^0(Z)) dt \right) &= b_n \int_{t_n}^{t_{n+1}} \frac{dt}{T(H, Z)} + O(\varepsilon \ln^2 h_n), \\
\int_{t_n}^{t_{n+1}} (f_3^0(p, q, z, \varepsilon) - f_3^0(p, q, z)) dt &= O(\varepsilon \ln h_n).
\end{aligned}$$

Finally, we have

$$(\Delta z)_{n+1} - (\Delta z)_n = \varepsilon b_n \left(1 - \int_{t_n}^{t_{n+1}} \frac{dt}{T(H, Z)} \right) + \varepsilon^2 O(h_n^{-1/2}). \qquad (3.17)$$

In an analogous way we get

$$(\Delta h)_{n+1} - (\Delta h)_n = -\varepsilon a_n \left(1 - \int_{t_n}^{t_{n+1}} \frac{dt}{T(H, Z)} \right) + \varepsilon^2 O(h_n^{-1/2}). \qquad (3.18)$$

In accordance with (3.2), $a_n > c_3^{-1}$. Let us denote $\mu_n = a_n^{-1} b_n$. It follows from (3.17), (3.18) that

$$(\Delta z)_{n+1} - (\Delta z)_n = -[(\Delta h)_{n+1} - (\Delta h_n)]\mu_n + \varepsilon^2 O(h_n^{-1/2}).$$

Consecutive use of this relation gives

$$(\Delta z)_{n+1} = (\Delta z)_1 - \sum_{s=1}^{n} [(\Delta h)_{s+1} - (\Delta h)_s)]\mu_s + \varepsilon^2 \sum_{s=1}^{n} O(h_s^{-1/2}).$$

Moreover,

$$\sum_{s=1}^{n} [(\Delta h)_{s+1} - (\Delta h)_s]\mu_s = (\Delta h)_{n+1}\mu_n - (\Delta h)_1\mu_1 - \sum_{s=2}^{n} (\Delta h)_s(\mu_s - \mu_{s-1}).$$

The definition of μ_s, Lemma 3.2 and estimates (3.11) imply that $\mu_s - \mu_{s-1} = O(\varepsilon \ln h_s)$. As $(\Delta h)_{n+1} = O(\varepsilon)$, $(\Delta z)_1 = O(\varepsilon)$, $\mu_n = O(1)$, so

$$z_{n+1} - Z_{n+1} \equiv (\Delta z)_{n+1} = O(\varepsilon) + \varepsilon^2 \sum_{s=1}^{n} O(h_s^{-1/2}) = O(\varepsilon).$$

Integrating both sides of (3.16) with respect to time from t_n to some $t \in (t_n, t_{n+1})$ and estimating the right hand side, we get for $t_n \leq t \leq t_{n+1}$

$$z(t) - Z(\varepsilon t) = z_n - Z_n + O(\varepsilon)$$

(in the case when $n = n^*$ this estimate is valid for $t_{n^*} \leq t < t_{**}$). Thus, for $0 \leq t \leq t_{**}$ we have $z(t) - Z(\varepsilon t) = O(\varepsilon)$. Hence, this estimate is valid while $H(\varepsilon t) \geq k_1 \varepsilon$ (as the last inequality certainly holds for $0 \leq t \leq t_{**}$). The Proposition 2.1 is proved.

23

Remark. In the proof of the last estimate representation (3.16) was used. We can not use this representation in problems where separatrices connect different saddle points, and at these points the function $f_3^0(p, q, z)$ has different values. The accuracy of the description of behaviour of z in these cases is worse than $O(\varepsilon)$ in general. This accuracy is given by the obtained above estimate (3.15): $z(t) - Z(\varepsilon t) = O(\varepsilon \ln H(\varepsilon t))$.

IV. The proof of Proposition 2.2 is completely analogous to the given above proof of the Proposition 2.1. But final estimates are different. The reason is as follows. As above, in nn. I, II, we prove that

$$|j(t) - J(\varepsilon t)| + |z(t) - Z(\varepsilon t)| = O(\varepsilon) + O\left(\varepsilon \ln \frac{H(\varepsilon t)}{H(\varepsilon t')}\right)$$

under assumptions of Proposition 2.2 (we will omit subscript ν and superscript " ' " in the notation (H_ν', Z_ν')). Here t' is the initial moment of time for the motion in Proposition 2.2. Because $H(\varepsilon t') \sim \varepsilon$, we have

$$|j(t) - J(\varepsilon t)| + |z(t) - Z(\varepsilon t)| = O(\varepsilon \ln \varepsilon)$$

(for Proposition 2.1 there was $H(\varepsilon t') \sim 1$, and there was used the estimate $O(\varepsilon \ln H(\varepsilon t))$ in the right hand side of this equality). As in n. II, to estimate $|h(t) - H(\varepsilon t)|$ from here, we use Lagrange's formula

$$h(t) - H(\varepsilon t) = \left(\frac{\partial h}{\partial I}\right)_* (j(t) - J(\varepsilon t)) + \left(\frac{\partial h}{\partial z}\right)_* (z(t) - Z(\varepsilon t)).$$

The symbol " * " here indicates that derivatives are calculated at some point in the straight line interval with endpoints $(j(t), z(t))$ and $(J(\varepsilon t), Z(\varepsilon t))$. As by assumption $\frac{1}{2}|H(\varepsilon t)| \leq |h(t)| \leq 2|H(\varepsilon t)|$, we have

$$\left(\frac{\partial h}{\partial I}\right)_* = \frac{2\pi}{T_*} = O\left(\frac{1}{1 + |\ln |H(\varepsilon t)||}\right),$$

$$\left(\frac{\partial h}{\partial z}\right)_* = -\left(\frac{\partial I/\partial z}{\partial I/\partial h}\right)_* = O\left(\frac{1}{1 + |\ln |H(\varepsilon t)||}\right).$$

Therefore

$$h(t) - H(\varepsilon t) = O\left(\frac{\varepsilon \ln \varepsilon}{1 + |\ln |H(\varepsilon t)||}\right),$$

and this is the assertion of Proposition 2.2 concerning an accuracy of description of h.

Now, the estimate of z can be improved as in n. III, and we get

$$z(t) - Z(\varepsilon t) = O\left(\frac{\varepsilon \ln \varepsilon}{1 + |\ln |H(\varepsilon t)||}\right).$$

This is the assertion of Proposition 2.2 concerning the accuracy of description of z.

3.4 Proofs of Lemmas on unperturbed motion

3.4.1 Preliminary estimates

Let η, ξ be such variables that the quadratic part of the Hamiltonian near the saddle C is $\frac{1}{2}\omega_0(z)(\eta^2 - \xi^2)$, $\omega_0 > 0$, and the transformation $(p, q) \mapsto (\eta, \xi)$ is a canonical transformation containing z as a parameter. Let us denote $\mathcal{H}(\eta, \xi, z)$ the Hamiltonian E expressed via η, ξ, z.

Lemma 3.7 *Let $|\xi| < d_1^{-1}$, $|\eta| < d_1^{-1}$.*
If $\mathcal{H}(\eta, \xi, z) > 0$, then

$$
\begin{aligned}
\eta &= \pm\sqrt{2\omega_0^{-1}\mathcal{H} + \xi^2} + O(\mathcal{H} + \xi^2), \\
\partial\mathcal{H}/\partial\eta &= \omega_0\eta + O(\eta^2), \\
d_2^{-1}\sqrt{\mathcal{H} + \xi^2} &< |\partial\mathcal{H}/\partial\eta| < d_2\sqrt{\mathcal{H} + \xi^2}.
\end{aligned}
\tag{3.19}
$$

If $\mathcal{H}(\eta, \xi, z) < 0$, then

$$
\begin{aligned}
\xi &= \pm\sqrt{2\omega_0^{-1}|\mathcal{H}| + \eta^2} + O(|\mathcal{H}| + \eta^2), \\
\partial\mathcal{H}/\partial\xi &= -\omega_0\xi + O(\xi^2), \\
d_2^{-1}\sqrt{|\mathcal{H}| + \eta^2} &< |\partial\mathcal{H}/\partial\xi| < d_2\sqrt{|\mathcal{H}| + \eta^2}.
\end{aligned}
$$

Proof.
Consider the case $\mathcal{H} > 0$. The case $\mathcal{H} < 0$ is analogous. We have

$$
\begin{aligned}
\mathcal{H} &= \frac{1}{2}\omega_0(\eta^2 - \xi^2) + O(\eta^3) + O(\eta^2\xi) + O(\eta\xi^2) + O(\xi^3), \\
\eta^2(1 + O(\eta) + O(\xi)) &= 2\omega_0^{-1}\mathcal{H} + \xi^2 + \xi^2(O(\eta) + O(\xi)), \\
\eta^2 &= \frac{2\omega_0^{-1}\mathcal{H} + \xi^2}{1 + O(\eta) + O(\xi)} + \xi^2(O(\eta) + O(\xi)).
\end{aligned}
$$

25

From this equality we derive that

$$\nu_1^{-1}|\xi| < \nu_1^{-1}\sqrt{\mathcal{H} + \xi^2} < |\eta| < \nu_1\sqrt{\mathcal{H} + \xi^2},$$
$$\eta^2 = 2\omega_0^{-1}\mathcal{H} + \xi^2 + O((\sqrt{\mathcal{H} + \xi^2})^3),$$
$$\eta = \pm\sqrt{2\omega_0^{-1}\mathcal{H} + \xi^2} + O(\mathcal{H} + \xi^2).$$
$$\partial\mathcal{H}/\eta = \omega_0\eta + O(|\eta|^2 + |\xi|^2) = \omega_0\eta + O(\eta^2),$$
$$d_2^{-1}\sqrt{\mathcal{H} + \xi^2} < \left|\frac{\partial\mathcal{H}}{\partial\eta}\right| < d_2\sqrt{\mathcal{H} + \xi^2}.$$

Lemma 3.7 is proved.

Corollary 3.5 *a) For $|\xi| < d_3^{-1} < d_1^{-1}$, $0 \le h \le d_4^{-1}$ the equation $\mathcal{H}(\eta, \xi, z) = h$ defines a unique $\eta = \tilde{\eta}(h, \xi, z)$ such that $0 \le \tilde{\eta}(h, \xi, z) < \frac{1}{2}d_2^{-1}$. For $h + \xi^2 > 0$ the function $\tilde{\eta}$ is smooth and*

$$\frac{\partial\tilde{\eta}}{\partial h} = \frac{1}{\partial\mathcal{H}/\partial\eta}, \quad \frac{\partial\tilde{\eta}}{\partial z} = -\frac{\partial\mathcal{H}/\partial z}{\partial\mathcal{H}/\partial\eta}. \tag{3.20}$$

The same equation $\mathcal{H}(\eta, \xi, z) = h$ defines also a unique $\eta = \tilde{\tilde{\eta}}(h, \xi, z)$ such that $-\frac{1}{2}d_2^{-1} < \tilde{\tilde{\eta}} \le 0$ with analogous properties.

b) For $|\eta| < d_3^{-1} < d_1^{-1}$, $-d_4^{-1} \le h \le 0$ the equation $\mathcal{H}(\eta, \xi, z) = h$ defines a unique $\xi = \tilde{\xi}(h, \eta, z)$ such that $0 \le \tilde{\xi}(h, \eta, z) < 1/2d_2$. For $|h| + \eta^2 > 0$ the function $\tilde{\xi}$ is smooth and

$$\frac{\partial\tilde{\xi}}{\partial h} = \frac{1}{\partial\mathcal{H}/\partial\xi}, \quad \frac{\partial\tilde{\xi}}{\partial z} = -\frac{\partial\mathcal{H}/\partial z}{\partial\mathcal{H}/\partial\xi}.$$

The same equation $\mathcal{H}(\eta, \xi, z) = h$ defines also a unique $\xi = \tilde{\tilde{\xi}}(h, \xi, z)$ such that $-\frac{1}{2}d_2^{-1} < \tilde{\tilde{\xi}} \le 0$ with analogous properties.

3.4.2 Proof of Lemma 3.1

To save notation we consider the case when φ does not depend on ε: $\varphi = \varphi(p, q, z)$. Let $0 < h < \frac{1}{2}d_4^{-1}$. Denote

$$R(h) = \oint_{E=h} \varphi dt, \quad R(0) = \oint_{l_3} \varphi dt.$$

26

1^0. Split $R(h)$ into integrals $R'(h)$ and $R''(h)$ over the segments of the curve $E = h$, situated inside and outside of the rectangle $|\xi| < d_3^{-1}$, $|\eta| < d_1^{-1}$ respectively. Split $R'(h)$ into integrals $R'_i(h), i = 1, 2, 3, 4$, where $R'_i(h)$ is calculated over the segment of the curve $E = h$, situated into the ith quadrant of the coordinate system $C\xi\eta$. Split $R(0)$ in the same way.

2^0. The integral $R''(h)$ is calculated over the arcs separated from the singularity (i.e. from the point C). It is easy to check that

$$R''(h) = R''(0) + O(h).$$

3^0. Estimate $R'_1(h)$. Let us use ξ as an independent variable in this integral. According to Corollary 3.5 we get

$$R'_1(h) = \int_0^{d_3^{-1}} \left(\frac{\varphi}{\dot{\xi}}\right)_h d\xi = \int_0^{d_3^{-1}} \left(\frac{\varphi}{\partial\mathcal{H}/\partial\eta}\right)_h d\xi,$$

where the subscript " h " indicates that we should plug $\eta = \tilde{\eta}(h, \xi, z)$ into the integrand. The function φ has the form

$$\varphi = \alpha\xi + \beta\eta + \varphi_2(\eta, \xi, z), \ \ \varphi_2 = O(\xi^2 + \eta^2), \ \ \alpha = \alpha(z), \ \ \beta = \beta(z).$$

Represent

$$R'_1(h) \ - \ R'_1(0) = \alpha \int_0^{d_3^{-1}} \left[\frac{\xi}{(\partial\mathcal{H}/\partial\eta)_h} - \frac{\xi}{(\partial\mathcal{H}/\partial\eta)_0}\right] d\xi \tag{3.21}$$

$$+ \ \beta \int_{\sqrt{h}}^{d_3^{-1}} \left[\left(\frac{\eta}{\partial\mathcal{H}/\partial\eta}\right)_h - \left(\frac{\eta}{\partial\mathcal{H}/\partial\eta}\right)_0\right] d\xi \ + \beta \int_0^{\sqrt{h}} \left[\left(\frac{\eta}{\partial\mathcal{H}/\partial\eta}\right)_h - \left(\frac{\eta}{\partial\mathcal{H}/\partial\eta)}\right)_0\right] d\xi$$

$$+ \int_{\sqrt{h}}^{d_3^{-1}} \left[\left(\frac{\varphi_2}{\partial\mathcal{H}/\partial\eta}\right)_h - \left(\frac{\varphi_2}{\partial\mathcal{H}/\partial\eta}\right)_0\right] d\xi + \int_0^{\sqrt{h}} \left[\left(\frac{\varphi_2}{\partial\mathcal{H}/\partial\eta}\right)_h - \left(\frac{\varphi_2}{\partial\mathcal{H}/\partial\eta}\right)_0\right] d\xi.$$

Estimate the last integral in (3.21). According to Lemma 3.7, the integrand in this integral is $O(\sqrt{h})$. Therefore, this integral is $O(h)$.

Estimate the next to last integral in (3.21). Using Lagrange's formula and Corollary 3.5, we get, that this integral is equal to

$$\int_{\sqrt{h}}^{d_3^{-1}} h \left(\frac{\partial}{\partial h}\left(\frac{\varphi_2}{\partial\mathcal{H}/\partial\eta}\right)\right)_{h_*} d\xi = h \int_{\sqrt{h}}^{d_3^{-1}} \left(\frac{\partial}{\partial\eta}\left(\frac{\varphi_2}{\partial\mathcal{H}/\partial\eta}\right)\frac{1}{\partial\mathcal{H}/\partial\eta}\right)_{h_*} d\xi,$$

27

where $0 < h_* < h$. By Lemma 3.7, the integrand is $O((h_* + \xi^2)^{-1/2})$. Therefore, this integral is

$$O(h) \int_{\sqrt{h}}^{d_3^{-1}} \frac{d\xi}{\sqrt{h_* + \xi^2}} d\xi = O(h) \int_{\sqrt{h}}^{d_3^{-1}} \frac{d\xi}{\xi} = O(h \ln h).$$

Consider the third integral in (3.21). It is equal to

$$\frac{1}{\omega_0} \int_0^{\sqrt{h}} \left[\left(\frac{\eta}{\eta + \chi_2} \right)_h - \left(\frac{\eta}{\eta + \chi_2} \right)_0 \right] d\xi = \frac{1}{\omega_0} \int_0^{\sqrt{h}} \frac{(\eta)_h (\chi_2)_0 - (\eta)_0 (\chi_2)_h}{(\eta + \chi_2)_h (\eta + \chi_2)_0} d\xi.$$

Here χ_2 is a function of η, ξ, z; expansion of χ_2 with respect to η, ξ starts with second order terms. By Lemma 3.7, the integrand is $O(\sqrt{h})$. The integral is, therefore, $O(h)$.

Estimate the second integral in (3.21). Using Lagrange's formula and Corollary 3.5, we get that this integral is equal to

$$\frac{h}{\omega_0^2} \int_{\sqrt{h}}^{d_3^{-1}} \left(\frac{\partial}{\partial \eta} \left(\frac{\eta}{\eta + \chi_2} \right) \cdot \frac{1}{\eta + \chi_2} \right)_{h_*} d\xi = \frac{h}{\omega_0^2} \int_{\sqrt{h}}^{d_3^{-1}} \left(\frac{\chi_2 - \eta\, \partial\chi_2/\partial\eta}{(\eta + \chi_2)^3} \right)_{h_*} d\xi.$$

The integrand here is $O(1/\xi)$. Therefore the second integral in (3.21) is $O(h \ln h)$.

Estimate the first integral in (3.21). It is equal to

$$\frac{1}{\omega_0} \int_0^{d_3^{-1}} \left[\frac{\xi}{\sqrt{2\omega_0^{-1}\mathcal{H} + \xi^2}} - \frac{\xi}{|\xi|} \right] d\xi$$

$$+ \frac{1}{\omega_0} \int_0^{d_3^{-1}} \xi \left[\left(\frac{\omega_0 \sqrt{2\omega_0^{-1}\mathcal{H} + \xi^2} - \partial\mathcal{H}/\partial\eta}{\sqrt{2\omega_0^{-1}\mathcal{H} + \xi^2}\, \partial\mathcal{H}/\partial\eta} \right)_h - \left(\frac{\omega_0 |\xi| - \partial\mathcal{H}/\partial\eta}{|\xi| \partial\mathcal{H}/\partial\eta} \right)_0 \right] d\xi.$$

We can estimate the second integral in this expression by splitting it into two integrals, from 0 to $\sqrt{h}$ and from $\sqrt{h}$ to d_3^{-1}. This gives that this integral is $O(h \ln h)$. So we get

$$R_1'(h) - R_1'(0) = \frac{\alpha}{\omega_0} \int_0^{d_3^{-1}} \left(\frac{\xi}{\sqrt{2\omega_0^{-1}h + \xi^2}} - \frac{\xi}{|\xi|} \right) d\xi + O(h \ln h).$$

28

Analogous estimates are valid for $R'_2, ..., R'_4$. In particular,

$$R'_4(h) - R'_4(0) = \frac{\alpha}{\omega_0} \int\limits_{-d_3^{-1}}^{0} \left(\frac{\xi}{\sqrt{2\omega_0^{-1}h + \xi^2}} - \frac{\xi}{|\xi|} \right) d\xi + O(h \ln h).$$

Therefore

$$R'_1(h) + R'_4(h) - R'_1(0) - R'_4(0) = O(h \ln h)$$

(we use here that the integral of an odd function over a symmetric with respect to 0 interval is equal to 0). The same estimate holds for $R'_2 + R'_3$. Taking into account the estimate for R'', we finally get

$$R(h) = R(0) + O(h \ln h).$$

Lemma 3.1 is proved.

3.4.3   Proof of Lemma 3.2

To save notation we consider the case when the functions φ, ψ do not depend on ε: $\varphi = \varphi(p, q, z), \psi = \psi(p, q, z)$. The proof uses the same scheme, as the proof of the Lemma 3.1 above. The part of the curve $E = h$, situated outside of the neighbourhood of the point C, gives a contribution $O(1)$ to any integral in (3.3). In a neighbourhood of the saddle point C we split the curve $E = h$ into four segments situated in four quadrants of the coordinate system $C\xi\eta$. The contribution to the integral of each of these segments is estimates in the same way. For certainty, let us consider the segment situated in the first quadrant. The corresponding integral we denote $R^{(i)}$, where i is the number of the integral in the Lemma 3.2, $i \neq 5$. To estimate integrals $R^{(i)}$ we introduce ξ as an independent variable in the integral and use estimates

in Lemma 3.7 and Corollary 3.5. We have

$$
R^{(1)} = \int_0^{d_3^{-1}} \left(\frac{|\varphi|}{\partial \mathcal{H}/\partial \eta} \right)_h d\xi = \int_0^{d_3^{-1}} \left(\frac{O(\eta)}{|\eta|} \right)_h d\xi = \int_0^{d_3^{-1}} O(1) d\xi = O(1).
$$

$$
R^{(2)} = \int_0^{d_3^{-1}} \left(\frac{|\psi|}{\partial \mathcal{H}/\partial \eta} \right)_h d\xi = \int_0^{d_3^{-1}} \frac{O(1)}{\sqrt{h + \xi^2}} d\xi = O(\ln h).
$$

$$
R^{(3)} = \frac{\partial}{\partial h} \int_0^{d_3^{-1}} \left(\frac{\psi}{\partial \mathcal{H}/\partial \eta} \right)_h d\xi = \int_0^{d_3^{-1}} \left(\frac{\partial}{\partial \eta} \left(\frac{\psi}{\partial \mathcal{H}/\partial \eta} \right) \frac{1}{\partial \mathcal{H}/\partial \eta} \right)_h d\xi
$$

$$
= \int_0^{d_3^{-1}} \frac{O(1) d\xi}{(h + \xi^2)^{3/2}} = O\left(\frac{1}{h} \right).
$$

$$
R^{(4)} = \frac{\partial}{\partial z} \int_0^{d_3^{-1}} \left(\frac{\psi}{\partial \mathcal{H}/\partial \eta} \right)_h d\xi
$$

$$
= \int_0^{d_3^{-1}} \left(\frac{\partial}{\partial z} \left(\frac{\psi}{\partial \mathcal{H}/\partial \eta} \right) - \frac{\partial}{\partial \eta} \left(\frac{\psi}{\partial \mathcal{H}/\partial \eta} \right) \frac{\partial \mathcal{H}/\partial z}{\partial \mathcal{H}/\partial \eta} \right)_h d\xi =
$$

$$
= \int_0^{d_3^{-1}} \frac{O(1)}{\sqrt{h + \xi^2}} d\xi = O(\ln h).
$$

$$
R^{(6)} = \frac{\partial}{\partial z} \int_0^{d_3^{-1}} \left(\frac{\varphi}{\partial \mathcal{H}/\partial \eta} \right)_h d\xi
$$

$$
= \int_0^{d_3^{-1}} \left(\frac{\partial}{\partial z} \frac{\varphi}{\partial \mathcal{H}/\partial \eta} - \frac{\partial}{\partial \eta} \left(\frac{\varphi}{\partial \mathcal{H}/\partial \eta} \right) \frac{\partial \mathcal{H}/\partial z}{\partial \mathcal{H}/\partial \eta} \right)_h d\xi = \int_0^{d_3^{-1}} O(1) d\xi = O(1).
$$

In order to estimate the integral (5) in Lemma 3.2 it is useful to combine the integrals over the segments of the curve $E = h$ situated in a neighbourhood of the point C in the 1st and 4th quadrants. Denote the corresponding

integral $R^{(5)}$. Then

$$\mathcal{H} = \frac{1}{2}\omega_0(\eta^2 - \xi^2) + \chi_3(\eta, \xi, z), \quad \chi_3 = O(|\xi|^3 + |\eta|^3),$$
$$\varphi = \alpha\xi + \beta\eta + \varphi_2(\eta, \xi, z), \quad \varphi_2 = O(\xi^2 + \eta^2),$$

$$
\begin{aligned}
R^{(5)} &= \frac{\partial}{\partial h} \int_{-d_3^{-1}}^{d_3^{-1}} \left(\frac{\varphi}{\partial\mathcal{H}/\partial\eta}\right)_h d\xi = \int_{-d_3^{-1}}^{d_3^{-1}} \left(\frac{\partial}{\partial\eta}\left(\frac{\varphi}{\partial\mathcal{H}/\partial\eta}\right)\frac{1}{\partial\mathcal{H}/\partial\eta}\right)_h d\xi \\
&= \int_{-d_3^{-1}}^{d_3^{-1}} \left(\frac{(\beta + \partial\varphi_2/\partial\eta)(\omega_0\eta + \partial\chi_3/\partial\eta) - (\alpha\xi + \beta\eta + \varphi_2)(\omega_0 + \partial^2\chi_3/\partial\eta^2)}{(\partial\mathcal{H}/\partial\eta)^3}\right)_h d\xi \\
&= -\int_{-d_3^{-1}}^{d_3^{-1}} \frac{\omega_0\alpha\xi d\xi}{(\partial\mathcal{H}/\partial\eta)_h^3} + \int_{-d_3^{-1}}^{d_3^{-1}} \frac{O(1)}{\sqrt{h+\xi^2}}d\xi = -\omega_0\alpha \int_{-d_3^{-1}}^{d_3^{-1}} \frac{\xi d\xi}{(\partial\mathcal{H}/\partial\eta)_h^3} + O(\ln h) \\
&= -\omega_0\alpha \int_{-d_3^{-1}}^{d_3^{-1}} \frac{\xi d\xi}{(\omega_0\sqrt{2\omega_0^{-1}h + \xi^2})^3} \\
&\quad - \omega_0\alpha \int_{-d_3^{-1}}^{d_3^{-1}} \xi\left(\frac{1}{(\omega_0\eta + \partial\chi_3/\partial\eta)_h^3} - \frac{1}{(\omega_0\sqrt{2\omega_0^{-1}h + \xi^2})^3}\right)d\xi + O(\ln h) = \\
&= \int_{-d_3^{-1}}^{d_3^{-1}} \xi\frac{O(\eta^4)d\xi}{(h+\xi^2)^3} + O(\ln h) = \int_{-d_3^{-1}}^{d_3^{-1}} \frac{O(1)d\xi}{\sqrt{h+\xi^2}} + O(\ln h) = O(\ln h).
\end{aligned}
$$

Combining estimates for different segments, we get estimates of Lemma 3.2.

3.4.4   Proof of Lemma 3.3

Let $0 < h < \frac{1}{2}d_4^{-1}$. We have

$$T = T(h) = \oint_{E=h} dt.$$

31

Let us split $T(h)$ into two integrals, $T'(h)$ and $T''(h)$, calculated over the segments of the curve $E = h$ situated respectively inside the domain $|\xi| < d_3^{-1}, |\eta| < d_3^{-1}$ and outside this domain. Let us, in addition, split $T'(h)$ into integrals $T'_{1,4}(h)$ and $T'_{2,3}(h)$, calculated over the segments situated in the 4th and 1st and, respectively, in the 2nd and 3rd quadrants of the system $C\xi\eta$. Integral $T''(h)$ is calculated over the segments of the curve separated from the point C. Therefore, $T''(h) = T''(0) + O(h)$. To estimate $T'_{4,1}(h)$, let us introduce ξ as an independent variable:

$$
\begin{aligned}
T'_{4,1}(h) &= \int_{-d_3^{-1}}^{d_3^{-1}} \frac{d\xi}{(\partial \mathcal{H}/\partial \eta)_h} \\
&= \int_{-d_3^{-1}}^{d_3^{-1}} \frac{d\xi}{\omega_0 \sqrt{2\omega_0^{-1}h + \xi^2}} + \int_{-d_3^{-1}}^{d_3^{-1}} \frac{\omega_0\sqrt{2\omega_0^{-1}h + \xi^2} - (\partial \mathcal{H}/\partial \eta)_h}{\omega_0\sqrt{2\omega_0^{-1}h + \xi^2}\,(\partial \mathcal{H}/\partial \eta)_h} d\xi.
\end{aligned}
$$

The integrand in the last integral remains bounded as $h \to 0$. Acting as in the proof of Lemma 3.1 we get that this integral can be calculated at $h = 0$ (i.e. over the separatrix) with an accuracy $O(h \ln h)$. Another integral, forming a part of $T'_{4,1}(h)$, has an explicit form:

$$
\begin{aligned}
\int_{-d_3^{-1}}^{d_3^{-1}} \frac{d\xi}{\omega_0 \sqrt{2\omega_0^{-1}h + \xi^2}} &= \frac{1}{\omega_0} \ln(\xi + \sqrt{2\omega_0^{-1}h + \xi^2}\,)\Big|_{-d_3^{-1}}^{d_3^{-1}} \tag{3.22}\\
&= -\frac{1}{\omega_0} \ln h + \frac{1}{\omega_0} \ln \frac{\omega_0}{2} + \frac{2}{\omega_0} \ln(d_3^{-1} + \sqrt{d_3^{-2} + 2\omega_0^{-1}h}\,) \\
&= -\frac{1}{\omega_0} \ln h + \frac{1}{\omega_0} \ln \frac{\omega_0}{2} + \frac{2}{\omega_0} \ln(d_3^{-1}) + O(h)\,.
\end{aligned}
$$

Analogous estimate holds for $T'_{2,3}(h)$. Combining these estimates we get the assertion of Lemma 3.3.

The calculation of asymptotic expansion of the function $T_i(h)$, the period of the motion along the trajectory $E = h < 0$ situated in the region G_i, $i = 1, 2$, is treated by the same method, but as independent variable near the saddle the variable η is used. In this calculation the sum of integrals over separatrices in asymptotic expansions of T_1 and T_2 coincides with the sum of integrals over separatrices for asymptotic expansion of the period of motion

in the the region G_3. The integral, which is calculated in this way, is reduced to the same form (3.22); the only difference in that h is replaced with $|h|$, and ξ is replaced with η. Hence the assertion of the Remark to Lemma 3.3 is valid.

3.4.5 Proof of Lemma 3.4

Let us denote ψ_c the value of the function ψ at the point C. Then

$$\frac{\partial}{\partial h}\frac{1}{T}\oint_{E=h}\psi dt = \frac{\partial}{\partial h}\frac{1}{T}\oint_{E=h}(\psi-\psi_c)dt\,. \tag{3.23}$$

Calculating the derivative and making use of the estimates in Lemma 3.2, its corollary, and corollary of Lemma 3.3, we get the result of the Lemma 3.4. The transformation (3.23) can not be used in the problems where the boundary of the domain contains several saddle points. But the result of the lemma is valid in these cases too. Let us describe briefly the corresponding proof. As in Lemma 3.3 we can prove expansions

$$T = \alpha\ln h + \chi,\ \ \chi = \chi(h,z) = O(1),\ \ \partial\chi/\partial h = O(\ln h)\,,$$

$$\oint_{E=h}\psi dt = \beta\ln h + \mu,\ \ \mu = \mu(h,z) = O(1),\ \ \partial\mu/\partial h = O(\ln h)\,.$$

Now

$$\frac{\partial}{\partial h}\frac{1}{T}\oint_{E=h}\psi dt \ = \ \frac{1}{T^2}\left[\left(\frac{\beta}{h}+\frac{\partial\mu}{\partial h}\right)(\alpha\ln h + \chi) - (\beta\ln h + \mu)\left(\frac{\alpha}{h}+\frac{\partial\chi}{\partial h}\right)\right] =$$

$$= \ \frac{1}{T^2}O(h^{-1}) = O(h^{-1}\ln^{-2}h)\,.$$

3.5 Proofs of Lemmas on perturbed motion

In this subsection Lemma 3.5 is proved. The proof of Lemma 3.6 is completely analogous, and it is omitted.

In accordance with Lemma 3.2 (1)

$$\oint_{E=h}|\chi|dt < \nu,$$

where

$$\chi = \chi(p, q, z, \varepsilon) = \frac{\partial E}{\partial q} f_1 + \frac{\partial E}{\partial p} f_2 + \frac{\partial E}{\partial z} f_3 \, .$$

Let us introduce $c_6 = 3\nu_1$, $c_7 = 2\max(c_2, c_4)$. Denote $\eta(t), \xi(t)$ the values of η, ξ at the point $(p(t), q(t), z(t))$. Let $c_6\varepsilon < h(t') < c_7^{-1}$. Denote t_{1*} the supremum of moments of time t_1 such that for $t' \le t \le t_1$ the solution $(p(t), q(t), z(t))$ is defined and meets the conditions

$$|\xi(t)| < d_3^{-1}, \ |\eta(t)| < d_1^{-1}, \ z(t) \in B - \frac{3}{2}c_1^{-1}, \tag{3.24}$$

$$\frac{1}{2}h(t') < h(t) < 2h(t').$$

Denote $\xi_{1*} = \xi(t_{1*})$. For $t' \le t \le t_{1*}$ we have

$$\dot{\xi} = \partial\mathcal{H}/\partial\eta + O(\varepsilon) > d_2^{-1}\sqrt{\frac{1}{2}h(t') + \xi^2} + O(\varepsilon) > \frac{1}{2}d_2^{-1}\sqrt{\frac{1}{2}h(t') + \xi^2} \, ,$$

$$t - t' = \int_0^{\xi(t)} \frac{d\xi}{\dot{\xi}} < 2d_2 \int_0^{d_3^{-1}} \frac{d\eta}{\sqrt{\frac{1}{2}h(t') + \xi^2}} = O(\ln h(t')) \, , \tag{3.25}$$

$$|z(t) - z(t')| = O(\varepsilon \ln h(t')) \, ,$$

$$|h(t) - h(t')| \le \varepsilon \int_{t'}^{t} |\chi| dt < 2d_2\varepsilon \int_0^{\xi(t)} \frac{|\chi| d\xi}{\sqrt{\frac{1}{2}h(t') + \xi^2}} = \int_0^{d_3^{-1}} O(1) d\varepsilon = O(\varepsilon) \, .$$

The obtained inequalities allow to get more accurate estimate for $h(t_{1*})$:

$$h(t_{1*}) - h(t') = \varepsilon \int_{t'}^{t_{1*}} \chi dt = \varepsilon \int_0^{\xi_{1*}} \frac{\chi}{\partial\mathcal{H}/\partial\eta + O(\varepsilon)} d\xi =$$

$$= \varepsilon \int_0^{\xi_{1*}} \left(\frac{\chi}{\partial\mathcal{H}/\partial\eta}\right)_{\substack{E=h(t') \\ z=z(t') \\ \varepsilon=0}} d\xi + \varepsilon \int_0^{\xi_{1*}} \left[\left(\frac{\chi}{\partial\mathcal{H}/\partial\eta + O(\varepsilon)}\right)_{\substack{E=h(t) \\ z=z(t)}} - \left(\frac{\chi}{\partial\mathcal{H}/\partial\eta}\right)_{\substack{E=h(t') \\ z=z(t') \\ \varepsilon=0}}\right] d\xi \, .$$

By means of (3.25) we get that the integrand in the second integral is

$$\frac{O(\varepsilon)}{h(t') + \xi^2} + \frac{O(\varepsilon \ln h(t'))}{\sqrt{h(t') + \xi^2}} \, .$$

34

So

$$h(t_{1*}) - h(t') = \varepsilon \int_0^{\xi_{1*}} \left(\frac{\chi}{\partial \mathcal{H}/\partial \eta} \right)_{\substack{E=h(t') \\ z=z(t') \\ \varepsilon=0}} d\xi + O(\varepsilon^2/\sqrt{h(t')})\,.$$

In this expression the integral is calculated over the segment of the unperturbed trajectory in the unperturbed motion. Therefore

$$|h(t_{1*}) - h(t')| < \varepsilon \nu_1 \le \frac{1}{2}\varepsilon c_6, \quad \frac{1}{2}h(t') < h(t_{1*}) < 2h(t')\,.$$

Therefore at $t = t_{1*}$ the conditions on z, h, η in (3.24) are satisfied as strict inequalities. By definition of t_{1*} there should be $\xi(t_{1*}) = d_3$.

In the further motion the phase point evidently makes the curve that is close to the unperturbed trajectory $E = h(t'), z = z(t')$, and arrives at the segment $\xi = d_3^{-1}$, $-d_1^{-1} < \eta < 0$. Along this curve $h = h(t') + O(\varepsilon)$, $z = z(t') + O(\varepsilon \ln h(t'))$. Along any part of this curve the change of E with an accuracy $O(\varepsilon^2 \ln \varepsilon)$ is equal to the integral of the function $\varepsilon \chi(p, q, z, 0)$ over the segment of the unperturbed trajectory and, therefore, this change does not exceed $\frac{3}{2}\varepsilon \nu_1$. Therefore in this motion $\frac{1}{2}h(t') < h(t) < 2h(t')$.

Further motion is considered in an analogous manner. The phase point arrives first at the segment $\xi = -d_3^{-1}$, $-d_1^{-1} < \eta < 0$, then it makes the curve close to the unperturbed trajectory, arrives at the segment $\xi = -d_3^{-1}$, $0 < \eta < d_1^{-1}$ and, finally, at a moment of time $t'' = t' + O(\ln h(t'))$ arrives at the ray $C\eta$ having $0 < \eta < d_1^{-1}$. For $t' \le t \le t''$ estimates (3.6) are satisfied and the estimate

$$|j(t) - j(t')| \le \left(\frac{\partial I}{\partial h} \right)_{\substack{h=h_* \\ z=z_*}} |h(t) - h(t')| + \left(\frac{\partial I}{\partial h} \right)_{\substack{h=h_* \\ z=z_*}} |z(t) - z(t')| = O(\varepsilon \ln h(t'))$$

is valid (here $h_* \in (\frac{1}{2}h(t'), 2h(t'))$, $z_* = z(t') + O(\varepsilon \ln h(t'))$).

For the function $\varphi(p, q, z, \varepsilon)$, which vanishes identically at the point C, the integral along the motion is estimated in the same manner as above for the function χ. This gives the second estimate (3.7).

To estimate the integral along the motion of the function ψ in (3.7), we split it into integrals over the defined above segments of the trajectory, situated either far from or near the saddle. Integrals over the segments, situated far from the saddle, coincide with accuracy $O(\varepsilon \ln h(t'))$ with integrals over the segments of the unperturbed trajectory $E = h(t')$. For the integral over

the segment of the perturbed trajectory with $0 < \xi < d_3^{-1}$, $\eta > 0$, situated near the saddle, we have

$$\int\limits_0^{t_{1*}} \psi\, dt = \int\limits_0^{d_3^{-1}} \left(\frac{\psi}{\partial\mathcal{H}/\partial\eta + O(\varepsilon)} \right)_{\substack{E=h(t)\\z=z(t)}} d\xi = \int\limits_0^{d_3^{-1}} \left(\frac{\psi}{\partial\mathcal{H}/\partial\eta} \right)_{\substack{E=h(t')\\z=z(t')\\\varepsilon=0}} d\xi +$$

$$+ \int\limits_0^{d_3^{-1}} \left[\left(\frac{\psi}{\partial\mathcal{H}/\partial\eta + O(\varepsilon)} \right)_{\substack{E=h(t)\\z=z(t)}} - \left(\frac{\psi}{\partial\mathcal{H}/\partial\eta} \right)_{\substack{E=h(t')\\z=z(t')\\\varepsilon=0}} \right] d\xi\, .$$

Taking into account already proved estimates (3.6) and Lemma 3.7, we estimate the integrand in the last integral as

$$\frac{O(\varepsilon)}{(h(t') + \xi^2)^{3/2}} + \frac{O(\varepsilon \ln h(t'))}{h(t') + \xi^2}\, .$$

The integral of this function is $O(\varepsilon/h(t'))$. Therefore

$$\int\limits_0^{t_{1*}} \psi\, dt = \int\limits_0^{d_3^{-1}} \left(\frac{\psi}{\partial\mathcal{H}/\partial\eta} \right)_{\substack{E=h(t')\\z=z(t')\\\varepsilon=0}} + O(\varepsilon/h(t'))\, .$$

The integral in the right hand side is just the integral of ψ along the unperturbed trajectory $E = h(t'), z = z(t')$ when $\varepsilon = 0$. The integrals of ψ along other segments of the trajectory near the saddle are estimated in analogous manner. Combining these estimates we get the first estimate (3.7).

3.6   Proofs of Lemmas on averaged system

In this subsection the proofs of Lemmas 2.2 and 2.3 on the motion in averaged system are given. These proofs are based on Lemmas on unperturbed motion in Subsection 3.1.

3.6.1   Proof of Lemma 2.2

a) Existence of the solution with initial condition $h = 0, z = z_*$ at $\tau = \tau_*$ follows from the standard existence theorem for ODEs (see, e.g., [17], p. 21) as the right hand side of the averaged system is continuous.

36

To prove a uniqueness let us suppose that there are two solutions, $(H^{(1)}(\tau), Z^{(1)}(\tau))$ and $(H^{(2)}(\tau), Z^{(2)}(\tau))$, say in the region G_3 (i.e. $\tau < \tau_*$), crossing the separatrix at $\tau = \tau_*, z = z_*$. Denote $J^{(i)}(\tau) = I(H^{(i)}(\tau), Z^{(i)}(\tau))$, $i = 1, 2$. Denote $u(\tau) = |(J^{(2)}(\tau) - J^{(1)}(\tau)| + |(Z^{(2)}(\tau) - Z^{(1)}(\tau)|$. Suppose that $u(\tau_{**}) \neq 0$ for some $\tau_{**} < \tau_*$. Then $u(\tau) \neq 0$ for $\tau_{**} < \tau < \tau_*$ (in the opposite case we have a contradiction with the standard uniqueness theorem). We may assume that $H^{(i)}(\tau_{**}) < 1/2$, $i = 1, 2$. From the formulas for the right hand side of the averaged equations for z, J, (2.3), (2.7) and from Lemmas 3.1 - 3.4 we get that for $\tau_{**} < \tau < \tau_*$

$$\frac{du(\tau)}{d\tau} = O\left(\frac{1}{H_*(\tau)\ln^3 H_*(\tau)}\right) u(\tau),$$

where $H_*(\tau) = \min\{H^{(1)}(\tau), H^{(2)}(\tau)\}$. Therefore for $\tau_1 \in (\tau_{**}, \tau_*)$ we have

$$u(\tau_{**}) = u(\tau_1) \exp\left(\int_{\tau_1}^{\tau_{**}} O\left(\frac{1}{H_*(\tau)\ln^3 H_*(\tau)}\right) d\tau\right).$$

But according to the averaged equation for h (2.3), estimate (3.2), and Lemma 3.2 we have

$$\frac{dH^{(i)}}{d\tau} < \nu_1^{-1}/\ln H^{(i)}.$$

Therefore

$$\int_{\tau_1}^{\tau_{**}} O\left(\frac{1}{H_*(\tau)\ln^3 H_*(\tau)}\right) d\tau = O\left(\int_0^{1/2} \frac{dh}{h\ln^2 h}\right) = O(1)$$

and

$$U(\tau_{**}) = U(\tau_1)O(1). \tag{3.26}$$

But $U(\tau_1) \to 0$ as $\tau_1 \to \tau_*$. Therefore $U(\tau_{**}) = 0$ in contradiction with our hypothesis. The uniqueness is proved.

b) Assertion b) of the Lemma 2.2 is an evident corollary of formula (2.7) and condition (2.2), as it was discussed at the end of Subsection 2.3.

3.6.2 Proof of Lemma 2.3

We will omit index "ν" at solutions (H_ν, Z_ν) and (H'_ν, Z'_ν). There exists a moment of the slow time $\bar{\tau}_0 > \tau_0, \bar{\tau}_0 = \tau_0 + O(\delta\ln\delta)$ such that

$$-\nu_1\delta < H(\bar{\tau}_0) < -\delta, \quad -\nu_1\delta < H'(\bar{\tau}_0) < -\delta.$$

Making use of the formula

$$\frac{dZ}{d\tau} = f_{3C}(Z) + \frac{1}{T} \oint_{E=h} (f_3^0 - f_{3C})dt \tag{3.27}$$

we get that

$$Z(\bar{\tau}_0) - Z'(\bar{\tau}_0) = O(\delta).$$

For $\tau_0 \leq \tau \leq \bar{\tau}_0$ we have

$$|H(\tau)| + |H'(\tau)| + |Z(\tau) - Z'(\tau)| = O(\delta). \tag{3.28}$$

Therefore, the estimates of Lemma 2.3 are valid for $\tau_0 \leq \tau \leq \bar{\tau}_0$. Denote

$$\begin{aligned} J(\tau) &= I(H(\tau), Z(\tau)), \quad J'(\tau) = I(H'(\tau), Z'(\tau)), \\ U(\tau) &= |J(\tau) - J'(\tau)| + |Z(\tau) - Z'(\tau)|. \end{aligned}$$

Then $U(\bar{\tau}_0) = O(\delta \ln \delta)$.

Exactly as in (3.26), we get that $U(\tau) = O(U(\bar{\tau}_0))$ for $\bar{\tau}_0 \leq \tau \leq K$. So

$$|J(\tau) - J'(\tau)| = O(\delta \ln \delta), \ |Z(\tau) - Z'(\tau)| = O(\delta \ln \delta). \tag{3.29}$$

From here we get that

$$|H(\tau) - H'(\tau)| < \nu_2 \frac{\delta|\ln \delta|}{1 + |\ln |H_*(\tau)||}, \tag{3.30}$$

where $H_*(\tau)$ lies between $H(\tau)$ and $H'(\tau)$.

Let us consider two cases: $H(\tau) \geq -2\nu_2\delta|\ln \delta|$ and $H(\tau) < -2\nu_2\delta|\ln \delta|$.

If $H(\tau) \geq -2\nu_2\delta|\ln \delta|$, then, from (3.30), $H'(\tau) > -\nu_2\delta|\ln \delta|$, $H_*(\tau) > -\nu_2\delta|\ln \delta|$ and

$$|H(\tau) - H'(\tau)| = O(\delta). \tag{3.31}$$

If $H(\tau) < -2\nu_2\delta|\ln \delta|$, then, from (3.30), $\frac{1}{2}|H(\tau)| < |H'(\tau)| < 2|H(\tau)|$, $\frac{1}{2}|H(\tau)| < |H_*(\tau)| < 2|H(\tau)|$ and, therefore,

$$|H(\tau) - H'(\tau)| = O\left(\frac{\delta \ln \delta}{1 + |\ln |H(\tau)||}\right). \tag{3.32}$$

In view of (3.28), (3.31), (3.32) we get for $\tau_0 \leq \tau \leq K$

$$|H(\tau) - H'(\tau)| = O\left(\delta + \frac{\delta|\ln \delta|}{1 + |\ln |H(\tau)||}\right),$$

i.e. the estimate of Lemma 2.3 for $|H(\tau) - H'(\tau)|$.

Now we should improve the estimate (3.29) for Z. From equality (3.27), considered for $Z(\tau)$ and $Z'(\tau)$, Lemma 3.4, and (3.29), (3.30), making use of Lagrange formula, we get for $\bar\tau_0 < \tau < \bar\tau_0 + \nu_3^{-1}$

$$\frac{d}{dt}(Z - Z') = O(Z - Z') + O\left(\frac{\delta \ln \delta}{H_*(\tau) \ln^3 |H_*(\tau)|}\right),$$

where $H_*(\tau)$ lies between $H(\tau)$ and $H'(\tau)$.

From here we get for $\bar\tau_0 \leq \tau \leq \bar\tau_0 + \nu_3^{-1}$

$$Z(\tau) - Z'(\tau) = O(\delta) + O\left(\int_{\bar\tau_0}^{\tau} \frac{\delta \ln \delta \, d\tau}{H_*(\tau) \ln^3 |H_*(\tau)|}\right) = O(\delta) + O\left(\int_{\delta}^{|H_*(\tau)|} \frac{\delta \ln \delta \, dh}{h \ln^2 h}\right),$$

or

$$|Z(\tau) - Z'(\tau)| = O\left(\delta + \frac{\delta |\ln \delta|}{|\ln |H_*(\tau)||}\right). \tag{3.33}$$

We know that for $\bar\tau_0 + \nu_3^{-1} \leq \tau \leq K$ the following estimate is valid:

$$|Z(\tau) - Z'(\tau)| = O(\delta \ln \delta). \tag{3.34}$$

In view of (3.28), (3.33), (3.34) we have for $\tau_0 \leq \tau \leq K$

$$|Z(\tau) - Z'(\tau)| = O\left(\delta + \frac{\delta |\ln \delta|}{1 + |\ln |H(\tau)||}\right).$$

Lemma 2.3 is proved.

4 Passage through a narrow vicinity of separatrices

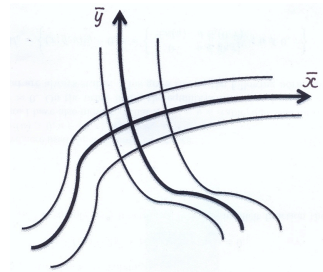In this section Proposition 2.3 of Subsection 2.5 is proved.

39

Figure 5: New coordinates.

4.1 Preliminary transformation

Below in all estimates $z \in B - c_1^{-1}$, where the constant c_1 is chosen in such a way that $3c_1^{-1}$-neighbourhood of the set $\{z\colon z = Z_\nu(\tau),\ \nu = 1, 2;\ \tau \in [0, K]\}$ belongs to B, and $3c_1^{-1}$-neighbourhood of the set $\{(p, q, z)\colon z = Z_\nu(\tau),\ E(p, q, z) = H_\nu(\tau),\ \nu = 1, 2;\ \tau \in [0, K]\}$ belongs to D. In some neighbourhood of the point C we can introduce new coordinates $\bar{x}, \bar{y}$ instead of p, q in such a way that the equations $\bar{x} = 0$ and $\bar{y} = 0$ define separatrices as it is shown in Fig. 5, and $\partial(p, q)/\partial(\bar{y}, \bar{x})|_C = 1$ (we do not need the transformation $(p, q) \mapsto (\bar{y}, \bar{x})$ be symplectic). In the new coordinates the function E has the form

$$E = -\omega_0 \bar{y} \bar{x}(1 + O(|\bar{y}| + |\bar{x}|)), \quad \omega_0 = \omega_0(z) > c_2^{-1}.$$

Lemma 4.1 *For $z \in B - c_1^{-1}$, $|\bar{y}| < c_3^{-1}, |\bar{x}| < c_3^{-1}$ there exists a smooth transformation of variables $\bar{F}\colon y, x, z \to \bar{y}, \bar{x}, z$ such that $\bar{y} = y + O(\varepsilon)$, $\bar{x} = x + O(\varepsilon)$, and in the variables y, x the perturbed motion is described by equations*

$$\dot{x} = -\omega_0 x(1 + O(\varepsilon + |x| + |y|)), \quad \dot{y} = \omega_0 y(1 + O(\varepsilon + |x| + |y|)).$$

This Lemma is a direct corollary of a theorem by N. Fenichel (see [12], Theorem 11.1). Coordinates of such type as y, x here are often called Fenichel's coordinates.

Lemma 4.1 allows to define in the domain $|\bar{y}| < c_3^{-1}, |\bar{x}| < c_3^{-1}$ the transformation of variables $F\colon y, x \to p, q$ depending on z as a parameter. The
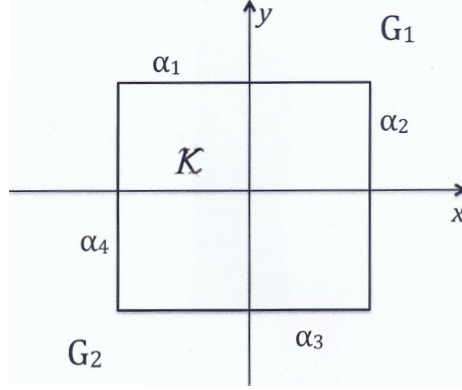
40

Figure 6: The square $\mathcal{K}$.

inverse transformation $F^{-1}\colon p, q \to y, x$ is defined in the c_4^{-1}-neighbourhood of the point C in the plane q, p.

4.2 Lemmas on perturbed motion

Consider in the plane x, y a square $\mathcal{K} = \{x, y\colon |x| < 2c_6\sqrt{\varepsilon}, \ |y| < 2c_6\sqrt{\varepsilon}, \}$. The constant c_6 is chosen below in Lemma 4.2 and Subsection 4.3. Let $\alpha_1, \ldots \alpha_4$ be sides of this square, the numeration of the sides corresponds to Fig. 6. Denote $\tilde{\mathcal{K}} = \tilde{\mathcal{K}}(z) = F[\mathcal{K}]$, $\tilde{\alpha}_i = \tilde{\alpha}_i(z) = F[\alpha_i]$, $i = 1, \ldots, 4$. For a point $(p(t), q(t))$, that belongs to c_4^{-1}-neighbourhood of the point C, we denote $(y(t), x(t)) = F^{-1}(p(t), q(t))$. The following two Lemmas show that points that started to move at sides of the square $\tilde{\mathcal{K}}$ far from its corners and not too close to the x-axis, will return to the sides of $\tilde{\mathcal{K}}$. In Fig. 7 the trajectories I and II correspond to Lemmas 4.2 and 4.3, respectively.

Lemma 4.2 *If* $z(t') \in B - 2c_1^{-1}$, $(p(t'), q(t')) \in \tilde{\alpha}_i(z(t')), i = 1, 3$, $|x(t')| < c_6\sqrt{\varepsilon}$, *then there exists a moment of time* t'' *such that*

$$(p(t''), q(t'')) \in \tilde{\alpha}_{i+1}(z(t'')), \ t'' - t' = O(\ln \varepsilon), \ \ h(t') - h(t'') > c_7^{-1}\varepsilon.$$

Remark. If we increase the value of the constant c_6, it would result in increasing of values of the constants $c_j, j > 7$. The constant c_7 can be chosen independently of c_6.
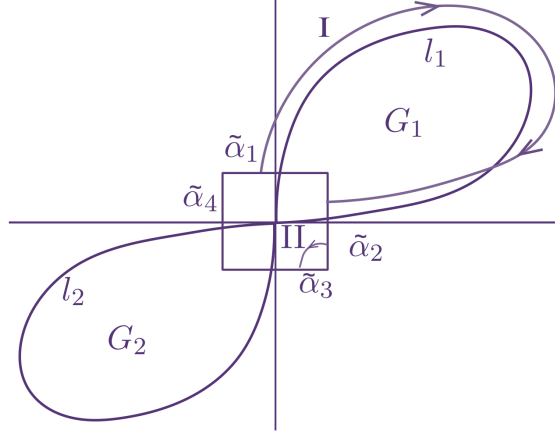
Figure 7: Trajectories I and II.

Lemma 4.3 *Let $z(t') \in B - 2c_1^{-1}$, $(p(t'), q(t')) \in \tilde{\alpha}_2(z(t')) \cup \tilde{\alpha}_4(z(t'))$, $|y(t')| > \varepsilon^{r_1}$, where r_1 is any given in advance positive number. If $(p(t'), q(t')) \in \tilde{\alpha}_2(z(t'))$, $y(t') > \varepsilon^{r_1}$, then there exists a moment of time t'' such that*

$$(p(t''), q(t'')) \in \tilde{\alpha}_1(z(t'')).$$

In analogous way,

if $(p(t'), q(t')) \in \tilde{\alpha}_2(z(t'))$, $y(t') < -\varepsilon^{r_1}$, then $(p(t''), q(t'')) \in \tilde{\alpha}_3(z(t''))$,
if $(p(t'), q(t')) \in \tilde{\alpha}_4(z(t'))$, $y(t') > \varepsilon^{r_1}$, then $(p(t''), q(t'')) \in \tilde{\alpha}_1(z(t''))$,
if $(p(t'), q(t')) \in \tilde{\alpha}_4(z(t'))$, $y(t') < -\varepsilon^{r_1}$, then $(p(t''), q(t'')) \in \tilde{\alpha}_3(z(t''))$.

In all cases $t'' - t' = O(\ln \varepsilon)$, $h(t'') - h(t') = O(\varepsilon^{3/2})$.

The following assertion shows that a phase point that stared to move from a diagonal of the square $\tilde{\mathcal{K}}$ not too close to the origin of the coordinate system will arrive to a side of $\tilde{\mathcal{K}}$.

Lemma 4.4 *If $z(t') \in B - 2c_1^{-1}$, $(p(t'), q(t')) \in \tilde{\mathcal{K}}(z(t'))$, $y(t') = -x(t') > \varepsilon$, then there exists a moment of time t'' such that*

$$(p(t''), q(t'')) \in \tilde{\alpha}_1(z(t'')), \quad t'' - t' = O(\ln \varepsilon), \ h(t'') - h(t') = O(\varepsilon^{3/2}).$$

42

Solutions of the averaged system with initial conditions from W^δ at $\tau = 0$ cross the separatrix at τ, z such that $|\tau - \hat{\tau}_*| < c_8\delta$, $|z - \hat{z}_*| < c_9\delta$. Consider the surface Λ in the space x, y, z:

$$\Lambda = \{x, y, z \colon |x| = 2c_6\sqrt{\varepsilon}, \ |z - \hat{z}_*| < 2c_9\delta, \ |y| < \varepsilon^{r_1}\}.$$

Let $\tilde{\Lambda}$ be the image of this surface under the map $x, y, z \mapsto p, q, z$. Denote Ξ the set in the space p, q, z swept by $\tilde{\Lambda}$ during the shift along the trajectories of the perturbed system (2.1) for time $t \in (-\hat{\tau}_* - 2c_8\delta, -\hat{\tau}_* + 2c_8\delta)$ before the first arrival to the boundary of the region $D - c_1^{-1}$.

Lemma 4.5
$$\text{mes } \Xi = O(\varepsilon^{r_1 - 1/2}\text{mes } W^\delta).$$

These Lemmas are proved in Subsections 4.4.2, 4.4.3

4.3   Proof of Proposition 2.3

Define $r_1 = r + 1/2$, where r is the integer number introduced before the formulation of Theorem 1. Introduce $w = W^\delta \cap \Xi$, $W_3^\delta = W^\delta \setminus w$, where Ξ is defined at the end of Subsection 4.2. According to Lemma 4.5 mes $w = O(\varepsilon^r\text{mes } W^\delta)$. Consider motion of a phase point $(p(t), q(t), z(t))$ that starts in W_3^δ at the moment of time $t = 0$. In accordance with Proposition 2.1 the moment of time t_- is defined such that $h(t_-) = 2k_1\varepsilon$. Then there exists a moment of time t'_- preceding t_- such that the point $(p(t'_-), q(t'_-))$ lies in the c_4^{-1}-neighbourhood of the saddle C for $z = z(t'_-)$, and $y(t'_-) = -x(t'_-) > 0$, $c_5\varepsilon < h(t'_-) < 2c_5\varepsilon$. The proof of the last assertion is completely analogous to the proofs of Lemmas 3.5, 3.6 and is omitted here.

Now let us choose the constant c_6 that defines the size of the square $\mathcal{K}$ in Subsection 4.2 (here we use the Remark to Lemma 4.2, which allows to increase the value of c_6). Choose c_6 such that in the segments α_1 and α_3 the inequality $-2k_2\varepsilon < E < 4c_5\varepsilon$ implies the inequality $|x| < c_6\sqrt{\varepsilon}$ for $z \in B - c_1^{-1}$. In the segments α_1 and α_3 according to Lemma 4.1 we have $|E| = 2\omega_0 c_6\sqrt{\varepsilon}\,|x| + O(\varepsilon^{3/2})$, $\omega_0 > c_2^{-1}$. Choose $c_6^2 > 4c_2\max(2c_5, k_2)$. Then for $|x| > c_6\sqrt{\varepsilon}$ we have $|E| > \max(4c_5\varepsilon, 2k_2)$. So, this choice of c_6 meets our condition. It is easy to check that under the same choice of c_6 the inequality $|E| > 2c_5\varepsilon$ is satisfied on the line $x = -y$ outside of $\mathcal{K}$.

Let us return to the motion of the point $(p(t), q(t), z(t))$. It is evident that $y(t') > \varepsilon$ (otherwise $h(t') = O(\varepsilon^2) < c_5\varepsilon$). In accordance with Lemma

43

4.4 there exists a moment of time $t''_- = t'_- + O(\ln \varepsilon)$ such that $(p(t''_-), q(t''_-)) \in \tilde{\alpha}_1(z(t''_-))$ and $h(t''_-) = h(t'_-) + O(\varepsilon^{3/2})$. In particular $\frac{1}{2}c_5\varepsilon < h(t''_-) < 3c_5\varepsilon$ and, therefore, $|x(t''_-)| < c_6\sqrt{\varepsilon}$ due to the choice of c_6. Now Lemmas 4.2, 4.3 allow, while their hypotheses are satisfied, to definite inductively the moments of time t_s of consecutive arrivals of our phase point to the sides $\tilde{\alpha}_1$ or $\tilde{\alpha}_3$ of the square; $t_1 = t''_-$. In accordance with Lemmas 4.2, 4.3, $h(t_s) - h(t_{s+1}) > \frac{1}{2}c_7^{-1}\varepsilon$. As $|E| < c_{10}\sqrt{\varepsilon}$ in the square $\mathcal{K}$, so $s < 4c_7c_{10}$. As $t_{s+1} - t_s = O(\ln \varepsilon)$, so $z(t_s) - z(t_1) = O(\ln \varepsilon)$ and, therefore, $z(t_s) \in B - \frac{5}{2}c_1^{-1}$. If $|x(t_s)| < c_6\sqrt{\varepsilon}$, then the moment of time t_{s+1} do exists (here we use the following: because of the choice of initial conditions and the definition of the set w, the hypothesis $|y| > \varepsilon^{r_1}$ of Lemma 4.3 can not be violated). But the sequence $\{t_s\}$ should be finite: $s < 4c_7c_{10}$. Therefore, there exists a number s_* such that $|x(t_{s_*})| \geq c_6\sqrt{\varepsilon}$. Then because of the choice of c_6 it should be $h(t_{s_*}) < -k_2\varepsilon$. By continuity there exists a moment of time $t_+ \in (t_-, t_{s_*})$ such that $h(t_+) = -k_2\varepsilon$, as it was stated in Proposition 2.3.

Estimate $h(t), H_\nu(\varepsilon t), z(t), Z_\nu(\varepsilon t)$ for $t_- \leq t \leq t_+$. For $t \in (t_-, t_+)$ we have $h(t) = O(\varepsilon)$ by definitions of t_-, t_+. As $H_\nu(\varepsilon t_-) = O(\varepsilon)$, $t_+ - t_- = O(\ln \varepsilon)$ and, in correspondence with estimates of Lemmas 3.2, 3.3, $\dot{H}_\nu = O(\varepsilon/\ln|H_\nu|)$ for $H_\nu \neq 0$, so $H_\nu(\varepsilon t) = O(\varepsilon)$ for $t \in (t_-, t_+)$. As $\dot{z} = O(\varepsilon), \dot{Z}_\nu = O(\varepsilon)$, so for $t \in (t_-, t_+)$ we have

$$z(t) - Z_\nu(\varepsilon t) = O(\varepsilon \ln \varepsilon).$$

Improve the last estimate. From identity (3.16) we get

$$\dot{z} - \dot{Z}_\nu = \varepsilon O(|z - Z_\nu|) + \varepsilon O(1/\ln|H_\nu|) + \varepsilon f_3^0(p, q, z) - f_{3C}^0(z) + O(\varepsilon^2),$$

where f_{3C}^0 is the value of the function $f_3^0(p, q, z)$ at the point C. Calculate integrals of the left and right sides of this relation with respect to time from t_s to some $t \in (t_s, t_{s+1})$. Making use of already established estimates for H_ν, $z - Z_\nu$, we get

$$z(t) - Z_\nu(\varepsilon t) = z(t_s) - Z_\nu(\varepsilon t_s) + O(\varepsilon) + \varepsilon \int_{t_s}^t (f_3^0(p, q, z) - f_{3C}^0(z))dt.$$

The integrand in the last integral vanishes at the point C. Analogously to Lemma 3.5, this integral is $O(1)$. Therefore, for $t \in (t_s, t_{s+1})$ we have

$$z(t) - Z_\nu(\varepsilon t) = z(t_s) - Z_\nu(\varepsilon t_s) + O(\varepsilon).$$

Analogous estimate holds for $t_1 \leq t \leq t_-$. As $s < 4c_7c_9$ and, in accordance with Proposition 2.1, $z(t_-) - Z_\nu(\varepsilon t_-) = O(\varepsilon)$, so $z(t) - Z_\nu(\varepsilon t) = O(\varepsilon)$ for $t_- \leq t \leq t_+$. This was the assertion of Proposition 2.3.

Remark. In the proof of the last estimate the representation (3.16) was used. We can not use this representation in problems, where separatrices connect different saddle points, and at these points the function $f_3^0(p, q, z)$ has different values. The accuracy of description of z in these cases is, in general, given by proved above estimate $z(t) - Z_\nu(\varepsilon t) = O(\varepsilon \ln \varepsilon)$ (see also Remark at the end of Subsection 3.3, part III).

4.4 Proofs of Lemmas on motion in a narrow vicinity of separatrices

4.4.1 Proof of Lemma 4.2

Let, for certainty, $(p(t'), q(t')) \in \tilde\alpha_1(z(t'))$. Then for any choice of c_6 there exists a moment of time $t'' > t'$, $t'' = t' + O(\ln \varepsilon)$ such that the point $(p(t''), q(t''))$ lies in the c_4^{-1}-neighbourhood of the saddle C, and $|x(t'')| = 2c_6\sqrt{\varepsilon}$, $h(t'') = h(t') - \varepsilon\Theta_2(z(t')) + O(\varepsilon^{3/2})$. The proof of this assertion is analogous to the proof of Lemma 3.5, and we omit it. Last estimate shows, in particular, that $c_7^{-1}\varepsilon \leq h(t') - h(t'') \leq d_1\varepsilon$, where c_7 and d_1 do not depend on choice of c_6. As $h(t') = O(\varepsilon)$, so $h(t'') = O(\varepsilon)$ and, therefore, $y(t'') = O(\sqrt{\varepsilon})$. Now

$$h(t') = -2\omega_0 c_6 \sqrt{\varepsilon} x(t') + O(\varepsilon^{3/2}),$$
$$h(t'') = -2\omega_0 c_6 \sqrt{\varepsilon} y(t'') + O(\varepsilon^{3/2}).$$

From here

$$y(t'') - x(t') = \frac{h(t') - h(t'')}{2\omega_0 c_6 \sqrt{\varepsilon}} + O(\varepsilon).$$

Choose $c_6^2 > d_1 c_2$. Then

$$|y(t'') - x(t')| < \frac{d_1 c_2}{2} \frac{\sqrt{\varepsilon}}{c_6} + O(\varepsilon) < \frac{1}{2} c_6 \sqrt{\varepsilon} + O(\varepsilon) < c_6 \sqrt{\varepsilon}.$$

Therefore, if $|x(t')| < c_6\sqrt{\varepsilon}$, then $|y(t'')| < 2c_6\sqrt{\varepsilon}$. This means that $(p(t''), q(t'')) \in \tilde\alpha_2(z(t''))$. This was the assertion of the Lemma 4.2.

4.4.2 Proofs of Lemmas 4.3, 4.4

Let, for certainty, $(p(t'), q(t')) \in \tilde{\alpha}_2(z(t'))$, $y(t') > \varepsilon^{r_1}$. Denote t'' the supremum of the moments of time $\bar{t}$ such that for $t' \le t \le \bar{t}$ the following conditions hold:

$$z(t) \in B - c_1^{-1}, \quad -2c_6\sqrt{\varepsilon} \le x(t) \le 2c_6\sqrt{\varepsilon} + \varepsilon, \quad \frac{1}{2}\varepsilon^{r_1} \le y(t) \le 2c_6\sqrt{\varepsilon}. \quad (4.1)$$

According to Lemma 4.1, for $t' \le t \le t''$ we have $\dot{y} > \frac{1}{2}c_2^{-1}y$. Therefore $t'' - t' = O(\ln \varepsilon)$ and $z(t'') \in B - \frac{3}{2}c_2^{-1}$. Then, again in correspondence with Lemma 4.1, on the segment $x = 2c_6\sqrt{\varepsilon}$, $|y| < 2c_6\sqrt{\varepsilon}$ we have $\dot{x} < -c_6c_2^{-1}\sqrt{\varepsilon}$. On the segment $x = -2c_6\sqrt{\varepsilon}$, $|y| < 2c_6\sqrt{\varepsilon}$ we have $\dot{x} > c_6c_2^{-1}\sqrt{\varepsilon}$. Therefore, for $t' \le t \le t''$ conditions for $x(t)$ in (4.1) hold as strict inequalities, and $|x(t'')| < 2c_6\sqrt{\varepsilon}$. In correspondence with the estimate $\dot{y} > \frac{1}{2}c_2^{-1}y$ we have $y(t'') > \varepsilon^{r_1}$. As at the moment $t = t''$ the phase point should arrive to the boundary of the domain defined in (4.1), there exists the only possibility: $y(t'') = 2c_6\sqrt{\varepsilon}$, i.e. $(p(t''), q(t'')) \in \tilde{\alpha}_1(z(t''))$.

In accordance with Lemma 4.1 for $t' \le t \le t''$ we have

$$h(t) = -\omega_0(z(t))y(t)x(t) + O(\varepsilon^{3/2}), \quad \frac{d}{dt}(\omega_0(z(t))y(t)x(t)) = O(\varepsilon^2 + |x|^3 + |y|^3).$$

Therefore, $h(t'') - h(t') = O(\varepsilon^{3/2})$, and this was the assertion of Lemma 4.3.

The proof of Lemma 4.4 follows the same lines.

4.4.3 Proof of Lemma 4.5

Denote $\tilde{\Lambda} \circ t$ the time-t schift of the set $\tilde{\Lambda}$ along trajectories of the perturbed system (2.1). Denote $\tilde{\Lambda} \circ [a_1, a_2] = \cup_{t \in [a_1, a_2]}\tilde{\Lambda} \circ t$. Assume for simplicity of explanation that $\tilde{\Lambda} \circ [-K/\varepsilon, 0] \in D$ (it is easy to avoid this restriction by considering only the part of the set $\tilde{\Lambda} \circ [-K/\varepsilon, 0]$ which does not leave D). For the set Ξ we have

$$\Xi = \left(\tilde{\Lambda} \circ \left[\frac{-\hat{\tau}_* - 2c_8\delta}{\varepsilon}, \frac{-\hat{\tau}_* + 2c_8\delta}{\varepsilon}\right]\right) \cap (D - c_1^{-1}).$$

Consider $\Xi_0 = \tilde{\Lambda} \circ [-d_1^{-1}, 0]$. For small enough d_1^{-1} the points of this set lie in $B - \frac{3}{2}c_1^{-1}$ with respect to z and in $\frac{1}{2}c_4^{-1}$-neighbourhood of the saddle point C with respect to (p, q). Therefore the set Ξ_0 can be considered in the

variables x, y, z of Lemma 4.1. As in these variables the phase flux through $\tilde{\Lambda}$ is $O(\varepsilon^{r_1+1/2}\delta^{l-2})$, so the phase volume of the set Ξ_0 is $O(\varepsilon^{r_1+1/2}\delta^{l-2})$, if the volume element is $dxdydz$. As the Jacobian of the transformation $x, y, z \mapsto p, q, z$ is $O(1)$, so the same estimate of the phase volume is valid if the volume element is $dpdqdz$. Consider now sets

$$
\begin{aligned}
\Xi_j &= \Xi_0 \circ (-jd_1^{-1}), \\
j &= \left[\frac{(\hat{\tau}_* - 2c_8\delta)d_1}{\varepsilon}\right] - 1, ..., \left[\frac{(\hat{\tau}_* + 2c_8\delta)d_1}{\varepsilon}\right] + 1.
\end{aligned}
$$

We have mes $\Xi_j = O(\text{mes } \Xi_0)$, as the divergence of the right hand side of system (2.1) is $O(\varepsilon)$, and so the distortion of the phase volume in this system during the time $O(1/\varepsilon)$ does not exceed $O(1)$ times. As $\Xi \subset \cup_j \Xi_j$, we get mes $\Xi = O(\varepsilon^{r_1-1/2}\delta^{l-1}) = O(\varepsilon^{r_1-1/2}\text{mes } W^\delta)$. Lemma 4.5 is proved.

5 Calculation of measures captured into different regions at the separatrix

In this section Propositions 2.4 and 2.5 of Subsection 2.5 are proved.

5.1 Preliminary constructions

Denote $I = I(\nu, z, h)$ the value of the action variable for the trajectory $E = h$ of the unperturbed system in the region G_ν, $\nu = 1, 2, 3$. Since in each of the regions G_ν the dependence of I on h is monotonous, we can rewrite in any G_ν the averaged system (2.3) as a system of differential equations for I, z. Making use of gluing of solutions of averaged system at the separatrix (see Subsection 2.3), we can consider the averaged system with respect to the variable $\gamma = (\nu, z, I)$. Separatrix crossing leads to a jump of values ν and I. Denote $\bar{g}_1^\tau(\gamma)$ the shift of the point γ for slow time $\tau = \varepsilon t$ along the trajectory of such averaged system, $\bar{g}_1^0(\gamma) = \gamma$, and index "1" indicates that the solutions with $\nu = 3$ and $\nu = 1$ are glued at the separatrix. Denote $W = W^\delta$, $W_\nu = W_\nu^\delta$, $\alpha = (p, q, z)$. Let $\Pi(\cdot)$ be the standard projection from the α-space to the γ-space; $\Pi(\alpha) = \gamma$. Denote

$$
\Gamma = \Pi[W], \ \Gamma_1^K = \bar{g}_1^K[\Gamma], W_1^K = \Pi^{-1}[\Gamma_1^K].
$$

We will denote points from $\Gamma, \Gamma_1^K, W, W_1^K$ as $\gamma^0, \gamma^K, \alpha^0 = (p^0, q^0, z^0), \alpha^K = (p^K, q^K, z^K)$ respectively.

5.2   Proof of the Proposition 2.4

Denote $g^t(\alpha)$ the time-t shift of a point $\alpha = (p, q, z)$ along the trajectory of system (2.1). Consider $g^t(\alpha^K)$ for $\alpha^K \in W_1^K$. Estimates in Propositions 2.1, 2.2, 2.3 were proved for initial conditions from W. Analogous estimates are valid for initial conditions from W_1^K, if we consider the motion in reverse direction in time. From this we get the following assertion: there exists a set $w_1 \subset W_1^K$, mes $w_1 = O(\varepsilon^r \delta^{l-1})$, such that for initial conditions $\alpha^K \in W_1^K \setminus w_1$ the behaviour of h, z along $g^t(\alpha^K)$, $-K\varepsilon \leq t \leq 0$, is described with an accuracy $O(\varepsilon \ln \varepsilon)$ by the solution of averaged system glued of solutions in G_1 and G_3 (passage from G_1 to G_2 is impossible for $\alpha^K \in W_1^K \setminus w_1$). It follows from these estimates that there exists a set $W_1'^K \subseteq W_1^K$ such that $g^{-K\varepsilon^{-1}}[W_1'^K] \subseteq W$, mes$(W_1^K \setminus W_1'^K) = O(\varepsilon \ln \varepsilon \, \delta^{l-2})$. Denote $W_1' = g^{-K\varepsilon^{-1}}[W_1'^K]$. The definition of W_2 implies that

$$W_1' \subseteq W \setminus W_2 = W_1 \cup w$$

A standard calculation of change of a phase volume along a motion gives us

$$\text{mes}\, W_1' = \int\limits_{W_1'^K} \exp\left(-\varepsilon \int\limits_0^{K\varepsilon^{-1}} \left(\frac{\partial f_1}{\partial q} + \frac{\partial f_2}{\partial p} + \frac{\partial f_3}{\partial z}\right) dt\right) dp^K dq^K dz^K. \quad (5.1)$$

Here the outer integral is calculated with respect to initial (or, better to say, "final") conditions from the set $W_1'^K$. The inner integral is calculated with respect to time along a solution of system (2.1) with given "final" condition. In the integral with respect to time it is reasonable to replace exact solution with the averaged one $\bar{g}_1^{-\tau}(\gamma^K)$, and to estimate the accuracy of this approximation. The result is described by the following Lemma, which is proved in Subsection 5.4.1.

Lemma 5.1 *For $\alpha^K \in W_1'^K$ the following estimate holds:*

$$\varepsilon \int_0^{K\varepsilon^{-1}} \left(\frac{\partial f_1}{\partial q} + \frac{\partial f_2}{\partial p} + \frac{\partial f_3}{\partial z} \right) dt = \int_0^K \mathcal{F}(\bar{\gamma}) d\tau + O(\varepsilon \ln \varepsilon),$$

$$\text{where } \bar{\gamma} = \bar{g}_1^{-\tau}(\gamma^K), \ \gamma^K = \Pi(\alpha^K),$$

$$\mathcal{F}(\gamma) = \frac{1}{T} \oint_\gamma \left(\frac{\partial f_1^0}{\partial q} + \frac{\partial f_2^0}{\partial p} + \frac{\partial f_3^0}{\partial z} \right) dt.$$

The integral with index $\gamma = (\nu, Z, J)$ is calculated along the level line of the Hamiltonian E in the region $G_\nu(z)$, and the "action" for this level line is equal to J. The parameter along the level line is time t of the unperturbed motion, T is the period of this motion.

Making use of Lemma 5.1 we get

$$\text{mes } W_1' = \int_{W_1'^K} \exp\left(- \int_0^K \mathcal{F}(\bar{\gamma}) d\tau \right) dp^K dq^K dz^K + O(\varepsilon \ln \varepsilon \ \delta^{l-2}).$$

Replace domain of integration $W_1'^K$ with W_1^K in the last expression. This gives an additional error $O(\varepsilon \ln \varepsilon \ \delta^{l-2})$. We get

$$\text{mes } W_1' = \int_{W_1^K} \exp\left(- \int_0^K F(\bar{\gamma}) d\tau \right) dp^K dq^K dz^K + O(\varepsilon \ln \varepsilon \ \delta^{l-2}).$$

In the last expression we can use action-angle variables J^K, φ^K instead of p^K, q^K as independent variables. As integrand does not depend on φ^K, we get

$$\text{mes } W_1' = 2\pi \int_{\Gamma_1^K} \exp\left(- \int_0^K \mathcal{F}(\bar{\gamma}) d\tau \right) dJ^K dz^K + O(\varepsilon \ln \varepsilon \ \delta^{l-2}).$$

Let us make in the outer integral the transformation of variables $(J^K, z^K) \mapsto (J^0, z^0)$ by means of the formula $\bar{g}_1^K(\gamma^0) = \gamma^K$, where $\gamma^0 = (3, z^0, J^0)$, $\gamma^K = (1, z^K, J^K)$. We get

$$\text{mes } (W_1') = 2\pi \int_\Gamma \Omega_{0,K}(\gamma^0) dJ^0 dz^0 + O(\varepsilon \ln \varepsilon \ \delta^{l-2}),$$

where

$$\Omega_{0,K}(\gamma^0) = \exp\left(-\int\limits_0^K \mathcal{F}(\bar\gamma)d\tau\right)\frac{\partial(J^K, z^K)}{\partial(J^0, z^0)}, \quad \bar\gamma = \bar{g}_1^\tau(\gamma^0).$$

Lemma 5.2
$$\Omega_{0,K}(\gamma^0) = P_1(z_*) = \Theta_1(z_*)/\Theta_3(z_*),$$

where $z_* = z_*(\gamma^0)$ is the value of z at the moment of the separatrix crossing for the solution of averaged system $\bar{g}_1^\tau(\gamma^0)$.

This Lemma is proved in Subsection 5.4.2.

Because of Lemma 5.2

$$\text{mes } W_1' = 2\pi \int\limits_\Gamma P_1(z_*)dJ^0 dz^0 + O(\varepsilon \ln \varepsilon\; \delta^{l-2}). \tag{5.2}$$

As $W_1' \subseteq W_1 \cup w$ and mes $w = O(\varepsilon^r\; \delta^{l-1})$, so

$$\text{mes } W_1' < \text{mes } W_1 + O(\varepsilon^r\; \delta^{l-1}).$$

In completely analogous way, but for the index $\nu = 2$, we get that there exists a set $W_2' \subseteq W_2 \cup w$ such that

$$\text{mes } W_2' = 2\pi \int\limits_\Gamma P_2(z_*)\, dJ^0 dZ^0 + O(\varepsilon \ln \varepsilon\; \delta^{l-2}), \tag{5.3}$$

$$\text{mes } W_2' < \text{mes } W_2 + O(\varepsilon^r\; \delta^{l-1}).$$

Consider relations

$$\begin{aligned}
\text{mes } W_1' &< \text{mes } W_1 + O(\varepsilon^r\; \delta^{l-1}), \tag{5.4}\\
\text{mes } W_2' &< \text{mes } W_2 + O(\varepsilon^r\; \delta^{l-1}),\\
\text{mes } W_1 + \text{mes } W_2 &= \text{mes } W + O(\varepsilon^r\; \delta^{l-1}),\\
\text{mes } W_1' + \text{mes } W_2' &= \text{mes } W + O(\varepsilon \ln \varepsilon\; \delta^{l-2}).
\end{aligned}$$

To get the last equality, it is enough to add (5.2) and (5.3), and to take into account that $P_1 + P_2 = 1$. From (5.4) we get

$$\begin{aligned}
\text{mes } W_1 &= \text{mes } W - \text{mes } W_2 + O(\varepsilon^r\delta^{l-1}) = \text{mes } W_1' + \text{mes } W_2' - \text{mes } W_2 + O(\varepsilon \ln \varepsilon\delta^{l-2})\\
&< \text{mes } W_1' + O(\varepsilon \ln \varepsilon\; \delta^{l-2}),
\end{aligned}$$

and, on the other hand,

$$\text{mes } W_1 > \text{mes } W_1' + O(\varepsilon^r \ \delta^{l-1}).$$

Therefore we get

$$\text{mes } W_1 = \text{mes } W_1' + O(\varepsilon \ln \varepsilon \ \delta^{l-2}) = 2\pi \int_\Gamma P_1(z_*) \, dJ^0 dz^0 + O(\varepsilon \ln \varepsilon \ \delta^{l-2}) =$$

$$= \int_W P_1(z_*) dp^0 dq^0 dz^0 + O(\varepsilon \ln \varepsilon \ \delta^{l-2}),$$

and analogous expression for W_2. Proposition 2.4 is proved.

5.3 Proof of Proposition 2.5

Let $I, \varphi \bmod 2\pi$ be the action-angle variables of the system with Hamiltonian E in a neighbourhood of the set W^δ. Consider in system (2.1) variable φ as a new time. For I, z we get a nonautonomous system of $l - 1$ equations. Consider for this system extended phase space with space variables I, z, φ and a new time $\vartheta: d\vartheta/d\varphi = 1$. Now for φ, I, z we have

$$\varphi' = 1, \ I' = \varepsilon f_4, \ z' = \varepsilon f_5.$$

Here "prime" denotes derivative with respect to ϑ, $f_i = f_i(\varphi, I, z, \varepsilon)$, $j = 4, 5$, are smooth functions. Denote u^ϑ the operator of the shift along the trajectories of this system during the time ϑ. Consider the sequence of the sets
$$u^{2s\delta}(U^\delta), \ s = 0, 1, ..., N - 1, N = [\pi/\delta].$$

Only adjoining sets in this sequence can intersect each other. The measure of any such intersection is $O(\varepsilon\delta^{l-1})$. Making use of the fact that the shift along the trajectories of this system during the time $O(1)$ distorts measure only with a coefficient $1 + O(\varepsilon)$, we get

$$\text{mes } \bigcup_{s=0}^{N-1} u^{2s\delta}(U^\delta) = N \text{ mes } U^\delta + O(\varepsilon\delta^{l-2}). \tag{5.5}$$

Analogous reasoning for sets U_ν^δ, $\nu = 1, 2$, gives the estimate

$$\text{mes } \bigcup_{s=0}^{N-1} u^{2s\delta}(U_\nu^\delta) = N \text{ mes } U_\nu^\delta + O(\varepsilon\delta^{l-2}). \tag{5.6}$$

Then

$$\text{mes } (W^\delta \triangle \bigcup_{s=0}^{N-1} u^{2s\delta}(U^\delta)) = O(\varepsilon\delta^{l-2} + \delta^l), \tag{5.7}$$

$$\text{mes } (W_\nu^\delta \triangle \bigcup_{s=0}^{N-1} u^{2s\delta}(U_\nu^\delta)) = O(\varepsilon\delta^{l-2} + \delta^l).$$

Here it is taken into account that $u^\vartheta(U_\nu^\delta) \cap W^\delta \subseteq W_\nu^\delta \cup w$; $\triangle$ is the symbol of symmetric difference of sets. From (5.5) - (5.7) we get

$$\text{mes } U^\delta = \frac{1}{N} \text{ mes } W^\delta + O(\varepsilon\delta^{l-1} + \delta^{l+1}),$$

$$\text{mes } U_\nu^\delta = \frac{1}{N} \text{ mes } W_\nu^\delta + O(\varepsilon\delta^{l-1} + \delta^{l+1}),$$

$$\frac{\text{mes } U_\nu^\delta}{\text{mes } U^\delta} = \frac{\text{mes } W_\nu^\delta}{\text{mes } W^\delta} + O(\delta + \frac{\varepsilon}{\delta}).$$

From here and from the result of Proposition 2.4 we get

$$\frac{\text{mes } U_\nu^\delta}{\text{mes } U^\delta} = \frac{\Theta_\nu(\hat{z}_*)}{\Theta_3(\hat{z}_*)} + O\left(\delta + \frac{\varepsilon|\ln\varepsilon|}{\delta}\right).$$

This was the assertion of Proposition 2.5.

5.4 Proofs of Lemmas on measure estimates

5.4.1 Proof of Lemma 5.1

Denote

$$\chi = \chi(p, q, z, \varepsilon) = \frac{\partial f_1}{\partial q} + \frac{\partial f_2}{\partial p} + \frac{\partial f_3}{\partial z}, \quad \chi^0 = \chi(p, q, z, 0).$$

We have

$$\varepsilon \int_0^{K\varepsilon^{-1}} \chi dt = \varepsilon \int_0^{K\varepsilon^{-1}} \chi^0 dt + O(\varepsilon) \tag{5.8}$$

$$= \varepsilon \int_0^{t'_-} \chi^0 dt + \varepsilon \int_{t'_-}^{t'_+} \chi^0 dt + \varepsilon \int_{t'_+}^{K\varepsilon^{-1}} \chi^0 dt + O(\varepsilon),$$

where t'_- and t'_+ are the moments of the time analogous to the moments t_-, t_+ introduced in Subsection 2.5, but for initial conditions from W_1^K. As $t'_- - t'_+ = O(\ln \varepsilon)$, so the second term in the right hand side of the last equality is $O(\varepsilon \ln \varepsilon)$.

For $0 \leq t \leq t'_-$ we will consider motion round by round, as it was done in Section 3. Suppose for simplicity of the exposition that during any round the phase point crosses the ray $C\eta$ just one time, and that the motion takes place in the region $|E| < 1/2$. Denote $t_1 < t_2 < ... < t_N$ the successive moments of the crossing of $C\eta$, $t_i \in (0, t'_-)$. According to Lemma 3.5 and its Corollary, we have

$$\frac{\varepsilon}{t_{i+1} - t_i} \int_{t_i}^{t_{i+1}} \chi^0 dt = \frac{\varepsilon}{T(h(t_i), z(t_i))} \oint_{E=h(t_i)} \chi^0 dt + \varepsilon^2 O\left(h^{-1}(t_i) \ln^{-1} h(t_i)\right).$$

$$\tag{5.9}$$

According to Propositions 2.1 - 2.3

$$|h(t_i) - H_1(t_i)| + |z(t_i) - Z_1(t_i)| = O\left(\frac{\varepsilon \ln \varepsilon}{\ln H_1(t_i)}\right), \tag{5.10}$$

$$\frac{1}{2} H_1(t_i) < h(t_i) < 2H_1(t_i).$$

From (5.9), (5.10), making use of Lemma 3.4 and estimate (4) of Lemma 3.2 we get

$$\frac{\varepsilon}{t_{i+1} - t_i} \int_{t_i}^{t_{i+1}} \chi^0 dt = \frac{\varepsilon}{T(H_1(t_i), Z_1(t_i))} \oint_{\substack{E=H_1(t_i) \\ z=Z_1(t_i)}} \chi^0 dt + O\left(\frac{\varepsilon^2 \ln \varepsilon}{H_1(t_i) \ln^3 H_1(t_i)}\right).$$

Let us multiplay the left and right hand sides of this equality by $t_{i+1} - t_i = O(\ln H(t_i))$ and sum up the obtained estimates. Taking into account that $t_1 = O(1)$, $t_N = t'_- + O(\ln \varepsilon)$, we get

$$\varepsilon \int_0^{t'_-} \chi^0 dt = \sum_{i=1}^{N} \frac{1}{T(H_1(t_i), Z_1(t_i))} \left(\oint_{\substack{E=H_1(t_i) \\ z=Z_1(t_i)}} \chi^0 dt \right) \varepsilon(t_{i+1} - t_i) + O(\varepsilon \ln \varepsilon).$$

The sum in the right hand side can be represented as an integral with an accuracy $O(\varepsilon \ln \varepsilon)$. Therefore, we have

$$\varepsilon \int_0^{t'_-} \chi^0 dt = \int_0^{\tau_*} \mathcal{F}(\bar{\gamma}) d\tau + O(\varepsilon \ln \varepsilon),$$

where τ_* is the moment of the separatrix crossing in the averaged system. In the analogous way

$$\varepsilon \int_{t'_+}^{K\varepsilon^{-1}} \chi^0 dt = \int_{\tau_*}^{K} \mathcal{F}(\bar{\gamma}) d\tau + O(\varepsilon \ln \varepsilon).$$

Results of Lemma 5.1 follow from these estimates and (5.8).

5.4.2   Proof of Lemma 5.2

Let τ_3 and τ_1 be any numbers such that $0 \le \tau_3 < \tau_* < \tau_1 \le K$. Here $\tau_* = \tau_*(\gamma^0)$ is the moment of the separatrix crossing for $\bar{g}_1^{\tau}(\gamma^0)$. Denote

$$\bar{g}_1^{\tau_3}(\gamma^0) = \gamma^{(3)} = (3, z^{(3)}, J^{(3)}), \quad \bar{g}_1^{\tau_1}(\gamma^0) = \gamma^{(1)} = (1, z^{(1)}, J^{(1)}).$$

Each of values $\gamma^0, \gamma^{(3)}, \gamma^{(1)}, \gamma^K$ defines all others. These values can be defined also through τ_*, z_*. Denote

$$\Omega_{\tau_3, \tau_1} = \exp \left(-\int_{\tau_3}^{\tau_1} \mathcal{F}(\bar{\gamma}) d\tau \right) \frac{\partial(J^{(1)}, z^{(1)})}{\partial(J^{(3)}, z^{(3)})}. \tag{5.11}$$

Similarly define Ω_{0, τ_3} and $\Omega_{\tau_1, K}$. Then

$$\Omega_{0, K} = \Omega_{0, \tau_3} \Omega_{\tau_3, \tau_1} \Omega_{\tau_1, K}.$$

Lemma 5.3 $\Omega_{0,\tau_3} = 1$, $\Omega_{\tau_1, K} = 1$.

This Lemma is proved in Subsection 5.4.3.

Corollary 5.1 $\Omega_{0,K} = \Omega_{\tau_3,\tau_1}$.

Now let τ_3 and τ_1 tend to τ_*. The first multiplier in (5.11) tends to 1. Therefore

$$\Omega_{0,K} = \lim_{\substack{\tau_3 \to \tau_*-0 \\ \tau_1 \to \tau_*+0}} \frac{\partial(J^{(1)}, z^{(1)})}{\partial(J^{(3)}, z^{(3)})} = \lim_{\tau_1 \to \tau_*+0} \frac{\partial(J^{(1)}, z^{(1)})}{\partial(\tau_*, z_*)} \bigg/ \lim_{\tau_1 \to \tau_*-0} \frac{\partial(J^{(3)}, z^{(3)})}{\partial(\tau_*, z_*)} .$$

Let us calculate these limits. Denote F and Φ the right hand sides of the averaged equations for I and z respectively. Then

$$J^{(1)} = \frac{1}{2\pi} S_1(z_*) + \int_{\tau_*}^{\tau_1} F d\tau , \quad z^{(1)} = z_* + \int_{\tau_*}^{\tau_1} \Phi d\tau .$$

Integrals here are calculated along the solution of the averaged system $\bar{g}_1^\tau(\gamma^0)$, and $S_1 = S_1(z)$ is the area of the region $G_1(z)$. Making use of the formulas for the right hand sides of averaged system (2.3), (2.7), we get

$$\lim_{\tau_1 \to \tau_*+0} \frac{\partial J^{(1)}}{\partial \tau_*} = -\lim_{\tau_1 \to \tau_*+0} F = \frac{1}{2\pi} \left(\Theta_1(z_*) + \left(\oint_{l_1} \frac{\partial E}{\partial z} dt \right) f_{3c}^0 \right) ,$$

$$\lim_{\tau_1 \to \tau_*+0} \frac{\partial z^{(1)}}{\partial \tau_*} = -\lim_{\tau_1 \to \tau_*+0} \Phi = -f_{3c}^0,$$

$$\lim_{\tau_1 \to \tau_*+0} \frac{\partial J^{(1)}}{\partial z_*} = \frac{1}{2\pi} \frac{\partial S_1}{\partial z_*} = -\frac{1}{2\pi} \oint_{l_1} \frac{\partial E}{\partial z} dt,$$

$$\lim_{\tau_1 \to \tau_*+0} \frac{\partial z^{(1)}}{\partial z_*} = \mathbf{1}_{l-2}.$$

Here $\mathbf{1}_{l-2}$ is the unit $(l-2) \times (l-2)$ matrix; f_{3c}^0 is the value of the function f_3^0 at the saddle point C for $z = z_*$. Making use of these relations, we get

$$\lim_{\tau_1 \to \tau_*+0} \frac{\partial(J^{(1)}, z^{(1)})}{\partial(\tau_*, z_*)} = \begin{vmatrix} \frac{1}{2\pi} \left(\Theta_1(z_*) + \left(\oint_{l_1} \frac{\partial E}{\partial z} dt \right) f_{3c}^0 \right) & -\frac{1}{2\pi} \oint_{l_1} \frac{\partial E}{\partial z} dt \\ -f_{3c}^0 & \mathbf{1}_{l-2} \end{vmatrix} = \frac{1}{2\pi} \Theta_1(z_*) .$$

55

Similarly,

$$\lim_{\tau_3 \to \tau_* - 0} \frac{\partial(J^{(3)}, z^{(3)})}{\partial(\tau_*, z_*)} = \frac{1}{2\pi} \Theta_3(z_*).$$

Finally we have

$$\Omega_{0,K} = \Theta_1(z_*)/\Theta_3(z_*).$$

This was the assertion of Lemma 5.2.

5.4.3 Proof of Lemma 5.3

To avoid long calculations with derivatives we will use known results about the averaging method. Denote Γ_0 a neighbourhood of the point $(3, z^0, J^0)$ such that $\bar{g}_1^\tau(\gamma)$ does not cross the separatrix for $0 \le \tau \le \tau_1$ and $\gamma \in \Gamma_0$. Let $W_0 = \Pi^{-1}(\Gamma_0)$, where Π is the standard projection from p, q, z-space to $\gamma = (\nu, z, J)$-space. The reasoning of the Subsection 5.1 shows that there exists a set $W_0' \in W_0$ such that

$$\begin{aligned}
\text{mes } W_0' &= \text{mes } W_0 + O(\varepsilon), \\
\text{mes } W_0' &= 2\pi \int_{\Gamma_0} \Omega_{0,\tau_3} dJ^0 dz^0 + O(\varepsilon).
\end{aligned}$$

From here

$$\text{mes } W_0 - 2\pi \int_{\Gamma_0} \Omega_{0,\tau_3} dJ^0 dz^0 = O(\varepsilon).$$

As the left hand side does not depend on ε, so

$$\text{mes } W_0 = 2\pi \int_{\Gamma_0} \Omega_{0,\tau_3} dJ^0 dz^0 = \int_{W_0} \Omega_{0,\tau_3} dp^0 dq^0 dz^0.$$

Therefore $\Omega_{0,\tau_3} = 1$. Similarly, $\Omega_{\tau_1,K} = 1$. Lemma 5.3 is proved.

5.5 A rule for calculation of probabilities

A heuristic reasoning of [14, 21], which leads to the formulas of Section 2 for probabilities of capture into different regions, is exposed in this subsection. This reasoning can be used as, in some sense, a rule, as it allows to calculate probabilities in general case of systems of form (2.1), for other than in Fig. 2 types of phase portraits. This reasoning is justified by Proposition 5.1 of this subsection.

5.5.1   A scheme of calculation of probabilities

The following reasoning does not pretend to be rigorous. A corresponding rigorous assertion is formulated at the end of this subsection. Let $C\xi\eta = C\xi\eta(z)$ be the system of principal axes for the saddle point C oriented as in Fig. 2. Let a phase point $(p(t), q(t))$ start moving at a moment of time $t = t'$ with $z = z'$, $E = h'$ from the ray $C\eta$. Denote $\Theta_\nu = \Theta_\nu(z')$. The point $(p(t), q(t))$ first makes a curve l'_2, which is close to $l_2(z)$. At the end of this curve

$$E = h'' = h' + \int_{l'_2} \frac{dE}{dt} dt \approx h' + \int_{l_2} \frac{dE}{dt} dt = h' - \varepsilon\Theta_2\,.$$

If $0 < h' < \varepsilon\Theta_2$, then $h'' < 0$, i.e. the phase point is captured into the region G_2. If $h' > \varepsilon\Theta_2$, then in further motion the phase point makes a curve l'_1, which is close to $l_1(z')$. At the end of this curve $E = h''' \approx h' - \varepsilon(\Theta_1 + \Theta_2)$. If $\varepsilon\Theta_2 < h' < \varepsilon(\Theta_1 + \Theta_2)$, then $h''' < 0$, i.e. the phase point is captured into the region G_1. If $h' > \varepsilon(\Theta_1 + \Theta_2)$, then $h''' > 0$, i.e. the phase point comes back to $C\eta$. Introduce intervals $\varkappa_1 = (\varepsilon\Theta_2, \ \varepsilon(\Theta_1 + \Theta_2))$, $\varkappa_2 = (0, \varepsilon\Theta_2,)$, $\varkappa_3 = (\varepsilon(\Theta_1 + \Theta_2), \infty)$, $\varkappa_\nu = \varkappa_\nu(z')$. In accordance with the previous explanation, points with $h' \in \varkappa_3$ will come back to $C\eta$ and after several rounds will arrive to $\varkappa_1 \cup \varkappa_2$. Points with $h' \in \varkappa_\nu$, $\nu = 1, 2$, will be captured into G_ν. The measure of the subset of U^δ, which will be captured into G_ν, is proportional to the length of the interval $\varkappa_\nu(\hat{z}_*)$ (because the majority of points from U^δ have $z \approx \hat{z}_*$ when cross $C\eta$ for the last time, and the phase flux through $\varkappa_\nu$ is equal in the principal approximation to the length of $\varkappa_\nu$). Thus

$$Q_\nu(\hat{M}_0) = \frac{\text{length } \varkappa_\nu}{\text{length } \varkappa_1 + \text{length } \varkappa_2} = \frac{\Theta_\nu(\hat{z}_*)}{\Theta_1(\hat{z}_*) + \Theta_2(\hat{z}_*)}\,.$$

The following assertion corresponds to the previous reasoning.

Proposition 5.1 *Let at a moment of time t' a point $(p(t'), q(t'))$ lie on the axis $C\eta(z')$ in k_3^{-1}-neighbourhood of the point C, and $z' \in B - k_4^{-1}$, $h' = E(p(t'), q(t'), z(t'))$. Introduce intervals (Fig. 8)*

$$\begin{aligned}
\varkappa'_1 &= (\varepsilon\Theta_2 + k_5\varepsilon^{3/2}, \ \varepsilon(\Theta_1 + \Theta_2) - k_5\varepsilon^{3/2}), \\
\varkappa'_2 &= (k_5\varepsilon^{3/2}, \varepsilon\Theta_2 - k_5\varepsilon^{3/2}), \\
\varkappa'_3 &= (\varepsilon(\Theta_1 + \Theta_2) + k_5\varepsilon^{3/2}, k_6^{-1})\,.
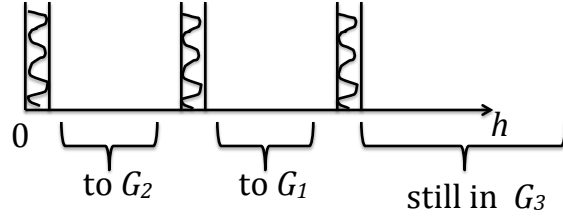\end{aligned}$$

Figure 8: For calculation of probabilities.

Then the following holds.

1^0*. If* $h' \in \varkappa'_\nu$*,* $\nu = 1, 2$ *then the phase point does not cross* $C\eta$ *again for* $t > t'$*. There exists* $t_\nu = t' + O(\ln \varepsilon)$ *such that* $(p(t_\nu), q(t_\nu)) \in G_\nu(z(t_\nu))$*,* $h(t_\nu) = -k_2 \varepsilon$*.*

2^0*. If* $h' \in \varkappa'_3$*, then there exists* $t_3 > t'$ *such that* $(p(t_3), q(t_3)) \in C\eta(z(t_3))$*,* $h' - h(t_3) > k_7^{-1}\varepsilon$*.*

Remarks.

1. Constant k_2 was introduced in Proposition 2.1.

2. Making use of Proposition 5.1 it is possible to prove formula for the probability (2.13). It is possible to prove also the result analogous to Proposition 2.4 but with more rough estimate: $O(\sqrt{\varepsilon})$ instead of $O(\varepsilon \ln \varepsilon)$.

3. The proof of the formula for the probability in previous sections uses essentially that separatrices divide phase space into three regions. But sometimes, because of an additional symmetry of the problem, a system has several saddle points connected by separatrices, and in these cases phase space can be divided into four or more regions (see, for example, [22]). Formulas for probabilities for these cases can be obtained by means of a reasoning, analogous to that used at the beginning of this section. These formulas can be justified by means of an assertion analogous to Proposition 5.1.

5.5.2 Proof of Proposition 5.1

Let us restrict ourselves by proving of assertion 1^0 for $\nu = 2$. The proofs of other assertions are completely analogous. Denote

$$\chi = \chi(p, q, z, \varepsilon) = \frac{\partial E}{\partial q} f_1 + \frac{\partial E}{\partial p} f_2 + \frac{\partial E}{\partial z} f_3.$$

58

Choose any k_4 such that $B - k_4^{-1}$ is not empty. For $z \in B - \frac{1}{2}k_4^{-1}$ we have

$$\oint_{l_2} |\chi| dt < c_1 \,.$$

We may assume that the quadratic part of the Hamiltonian E near the saddle point C in the variables η, ξ has the form $\frac{1}{2}\omega_0(z)(\eta^2 - \xi^2)$, $\omega_0 > 0$. Denote $x = (\xi - \eta)/\sqrt{2}$, $y = (\xi + \eta)/\sqrt{2}$ (cf. Fig. 7). Let $\Psi = \Psi(y, x, z)$ be the Hamiltonian E, expressed through y, x, z:

$$\Psi = -\omega_0 yx + O(|y|^3 + |x|^3) \,.$$

Lemma 5.4 *Let $z \in B - \frac{1}{2}k_4^{-1}$, $|x| < d_1^{-1}, |y| < d_1^{-1}$. If $|y| \geq |x|$, then*

$$\frac{\partial \Psi}{\partial x} = -\omega_0 y + O(y^2) \,.$$

If $|y| < |x|$, then

$$\frac{\partial \Psi}{\partial y} = -\omega_0 x + O(x^2) \,.$$

The proof is evident.

Corollary 5.2 *For $0 < |y| < d_2^{-1} < d_1^{-1}$, $|h| < d_3^{-1}$, $|x| \leq y$ the equation $\Psi(y, x, z) = h$ defines a unique $x = \tilde{x}(y, h, z)$ such that $|x| < d_1^{-1}$. Function $\tilde{x}$ is smooth and*

$$\frac{\partial \tilde{x}}{\partial h} = \frac{1}{\partial \Psi / \partial x}, \quad \frac{\partial \tilde{x}}{\partial z} = -\frac{\partial \Psi / \partial z}{\partial \Psi / \partial x} \,.$$

For $0 < |x| < d_2^{-1}$, $|h| < d_3^{-1}$, $|y| \leq |x|$ the equation $\Psi = h$ defines in the analogous manner $y = \tilde{y}(x, h, z)$.

Denote $y(t)$, $x(t)$ the values of y, x at the point $p(t), q(t), z(t)$. We have $y(t') = -x(t') > 0$. Let $\varepsilon^{3/2} < h(t') < \varepsilon \Theta_2(z')$. Denote t_{1*} the supremum of moments of time $t_1 > t'$ such that for $t' \leq t \leq t_1$ the solution $p(t), q(t), z(t)$ is defined and meets the conditions

$$
\begin{aligned}
|x(t)| &< d_1^{-1}, \ 0 < y(t) < d_2^{-1}, \ z(t) \in B - \frac{3}{4}k_4^{-1}, && (5.12) \\
|y(t)| &> |x(t)|, \ y(t) > \varepsilon^{7/8}
\end{aligned}
$$

(the set of such t_1 is not empty, as $\dot{y}(t') > -\dot{x}(t')$). Denote $y_{1*} = y(t_{1*})$.
For $t' \le t \le t_{1*}$ we have

$$\dot{y} \;=\; \omega_0 y + O(y^2) + O(\varepsilon) > c_2^{-1} y, \qquad\qquad (5.13)$$

$$t_{1*} - t' \;=\; \int\limits_{y(t')}^{y_{1*}} \frac{dy}{\dot{y}} < c_2 \int\limits_{y(t')}^{y_{1*}} \frac{dy}{y} < c_2 \ln \frac{y_{1*}}{y(t')} = O(\ln \varepsilon),$$

$$z(t) - z(t') \;=\; O(\varepsilon \ln \varepsilon),$$

$$|h(t) - h(t')| \;\le\; \varepsilon \int\limits_{t'}^{t} |\chi| dt < c_2 \varepsilon \int\limits_{y(t')}^{y(t)} \frac{|\chi| dy}{y} = \varepsilon \int\limits_{y(t')}^{y(t)} O(1) dy = O(\varepsilon).$$

The obtained inequalities show that at the moment of time t_{1*} all conditions in (5.12) but the inequality $y(t) < d_2^{-1}$ are satisfied with some margins. Therefore $y(t_{1*}) = d_2^{-1}$.

The obtained inequalities allow to give more accurate estimate of $h(t_{1*})$:

$$h(t_{1*}) \;-\; h(t') = \varepsilon \int\limits_{t'}^{t_{1*}} \chi dt = \varepsilon \int\limits_{y(t')}^{d_2^{-1}} \frac{\chi dy}{-\partial\Psi/\partial x + O(\varepsilon)}$$

$$=\; \varepsilon \int\limits_{y(t')}^{d_2^{-1}} \left(-\frac{\chi}{\partial\Psi/\partial x} \right)_{\substack{E=0 \\ z=z' \\ \varepsilon=0}} dy - \varepsilon \int\limits_{y(t')}^{d_2^{-1}} \left[\left(\frac{\chi}{\partial\Psi/\partial x + O(\varepsilon)} \right)_{\substack{E=h(t) \\ z=z(t)}} - \left(\frac{\chi}{\partial\Psi/\partial x} \right)_{\substack{E=0 \\ z=z' \\ \varepsilon=0}} \right] dy.$$

By means of (5.12), Lemma 5.4 and its Corollary, the integrand in the second integral for $y > \sqrt{\varepsilon}$ is estimated as

$$O\left(\frac{\varepsilon}{y^2} \right) + O\left(\frac{\varepsilon \ln \varepsilon}{y} \right).$$

For $\varepsilon^{7/8} < y < \sqrt{\varepsilon}$ the integrands in both integrals are $O(1)$. Therefore

$$h(t_{1*}) - h(t') = -\varepsilon \int\limits_{0}^{d_2^{-1}} \left(\frac{\chi}{\partial\Psi/\partial x} \right)_{\substack{E=0 \\ z=z(t') \\ \varepsilon=0}} dy + O(\varepsilon^{3/2}).$$

Here the integral is calculated over a segment of the unperturbed separatrix.

In the further motion the phase point makes a curve situated in $O(\varepsilon \ln \varepsilon)$-neighbourhood of the unperturbed separatrix for $z = z'$ and arrives at the segment $x = d_2^{-1}, |y| < d_1^{-1}$ at a moment of time $t_{2*} = t_{1*} + O(1)$ having $E = O(\varepsilon)$. The change of E along this curve with an accuracy $O(\varepsilon^2 \ln \varepsilon)$ is equal to the integral of function $\varepsilon \chi(p, q, z', 0)$ along the corresponding part of the unperturbed separatrix.

Then, through the time $O(\ln \varepsilon)$, at some moment of time t_{3*}, the phase point arrives either at the ray $x = y + c_3 \varepsilon > 0$ or at the ray $x = -y + c_3 \varepsilon > 0$. The motion for $t \in (t_{2*}, t_{3*})$ is considered in completely analogous way to that for $t' \le t \le t_{1*}$. The change of E with an accuracy $O(\varepsilon^{3/2})$ is equal to the integral of $\varepsilon \chi(p, q, z', 0)$ along the segment of the unperturbed separatrix with $0 \le x \le d_2^{-1}$, $|y| < d_1^{-1}$. Therefore $h(t_{3*}) = h(t') - \varepsilon \Theta_2(z') + O(\varepsilon^{3/2})$.

If

$$k_5 \varepsilon^{3/2} < h(t') < \varepsilon \Theta_2(z') - k_5 \varepsilon^{3/2} \,,$$

then

$$-\varepsilon \Theta_2(z') + \frac{1}{2} k_5 \varepsilon^{3/2} < h(t_{3*}) < -\frac{1}{2} k_5 \varepsilon^{3/2}$$

Therefore at the moment of time t_{3*} the phase point lies in the region $G_2(z(t_{3*}))$. Condition $h(t_{3*}) < -\frac{1}{2} k_5 \varepsilon^{3/2}$ and Lemma 5.4 allow to estimate $\dot\eta$ from below by a value of order $\varepsilon^{3/4}$. Making use of this estimate we can show that at a moment of time $t'' = t_{3*} + O(\varepsilon^{1/4})$ the phase point arrives at the ray $C\xi$ having $h(t'') = O(\varepsilon)$, $h(t'') < -\frac{1}{4} k_5 \varepsilon^{3/2}$.

Further motion is considered in an analogous manner. While $-k_2 \varepsilon < h(t) < 2c_1 \varepsilon$ the phase point moves round by round making curves near the unperturbed separatrix l_2. One round takes time $O(\ln \varepsilon)$, the value of E during one round decays by $\varepsilon \Theta_2(z') + O(\varepsilon^{3/2}) > c_3^{-1} \varepsilon$. Therefore, there exists a moment of time $t_\nu = t' + O(\ln \varepsilon)$ such that $h(t_\nu) = -k_2 \varepsilon$, $(p(t_\nu), q(t_\nu)) \in G_\nu(z(t_\nu))$. This is the assertion of Proposition 5.1.

A   Appendix. Perturbations of polyintegrable systems and separatrix crosings

The goal of this Appendix is to give a general description of the problem of separatrix crossing in single-frequency systems and to demonstrate, that under rather general assumptions the study of this problem can be reduced to study of separatrix crossing in system (2.1). The exposition here follows mainly [4], Subsection 6.1.10.

A natural framework for studying one-frequency averaging is the framework of perturbations of polyintegrable superintegrable (also called Nambu) systems[2]. In this problem the equations of motion have the form

$$\dot{x} = v(x, \varepsilon), \quad x \in D \subseteq \mathbb{R}^l, \quad 0 < \varepsilon \ll 1, \quad v(x, \varepsilon) = v_0(x) + v_1(x, \varepsilon). \qquad (A.1)$$

Here D is a bounded domain in $\mathbb{R}^l$. We assume that the unperturbed ($\varepsilon = 0$) system is polyintegrable, i.e. it has $l - 1$ smooth first integrals $H_1, \ldots, H_{l-1}$ which are independent almost everywhere in D. We assume that the domain D contains, together with each point, also the entire connected component of the common level set (a level line) of the integrals passing through this point. Then a level line on which the first integrals are independent is a smooth closed curve. In any domain filled by such level lines system (A.1) can be reduced to the standard form of a system with one rotating phase.

To introduce a framework for separatrix crossing we assume that:
a) the rank of the Jacobi matrix of the map $\mathcal{H} : D \to \mathbb{R}^{l-1}$ given by $\mathcal{H}(x) = (H_1(x), \ldots, H_{l-1}(x))$ is equal to $l - 1$ everywhere but on a smooth $l - 2$ dimensional surface, where it equals to $l - 2$;
b) at each point, where the rank equals $l-2$, the restriction of one of integrals onto the joint level of other integrals has a non-degenerate critical point;
c) at equilibrium positions of the unperturbed system (A.1) two eigenvalues are non-zero real numbers (the other eigenvalues are equal to 0 because of the existence of the integrals).

Then points, where the rank of the map $\mathcal{H}$ equals $l - 2$, coincide with equilibria of system (A.1) for $\varepsilon = 0$, the sum of non-zero eigenvalues equals 0 for such an equilibrium. We call separatrices the common level lines that pass through these points as well as a union of such level lines. Under the action of the perturbation phase points can cross separatrices.

---

[2]See [3] for description of properties of polyintegrable systems.

Assume that functions $H_1, \ldots, H_{l-2}$ are independent on separatrices. The values $z_1, \ldots, z_{l-2}$ of these functions from some ball in $\mathbb{R}^{l-2}$ can be taken as new variables. Joint levels of these functions form $l-2$-parametric family of 2-dimensional surfaces $\mathcal{S}_z$, $z = (z_1, \ldots, z_{l-2})$. Unperturbed dynamics on each of these surfaces is described by a Hamiltonian system with one degree of freedom for which the restriction E of the function H_{l-1} onto this surface is Hamilton's function, but the symplectic structure may be non-canonical. The phase portrait of each of these systems contains a saddle point and passing through it separatrices. In a neighbourhood of separatrices the phase portrait has the same form as in Fig. 2 and can be considered as a portrait in $\mathbb{R}^2$. (Notice that this does not depend on topology of $\mathcal{S}_z$. For example, the phase portrait of the pendulum, Fig. 1, should be considered on a cylinder, but a neighbourhood of separatrices can be put in $\mathbb{R}^2$ as a neighbourhood of separatrices of the form shown in Fig. 2.)

Let $\tilde{p}, q$ be Cartesian coordinates in $\mathbb{R}^2$. In these coordinates, in a neighbourhood of separatrices, the symplectic structure has a form $\mu(\tilde{p}, q, z)d\tilde{p} \wedge dq$, $\mu(\tilde{p}, q, z) \neq 0$. Define in a neighbourhood of separatrices a function $p = p(\tilde{p}, q, z)$ such that $\partial p/\partial \tilde{p} = \mu(\tilde{p}, q, z)$. In the variables p, q the symplectic structure takes the canonical form $dp \wedge dq$, and equation (A.1) takes the form (2.1). Thus, the results in Subsection 2.4 for system (2.1) describe also separatrix crossing for (A.1).

One can also consider separatrix crossings directly for perturbations of a polyintegrable system. The phase space of the averaged system is the set of common level lines of the integrals of the unperturbed system, which has the natural structure of a manifold with singularities [6] (singularities correspond to a separatrix). The averaged system approximately describes the evolution of the slow variables - values of the integrals of the unperturbed system. The probabilities of falling into different domains after a separatrix crossing are expressed in terms of ratios of the quantities

$$\tilde{\Theta}_i(z) = -\oint_{l_i(z)} \left(\beta_1(z)\frac{\partial H_1}{\partial x} + \ldots + \beta_{l-1}(z)\frac{\partial H_{l-1}}{\partial x} \right) v_1(x, 0)dt \,,$$

where z parametrises the surface of singular points ("saddles") of the unperturbed system, β_j are coefficients such that the expression inside the parentheses in the integrand vanishes at singular points, and $l_i = l_i(z)$ is a separatrix.

References

[1] *Arnold V. I.* Small denominators and problems of stability of motion in classical and celestial mechanics. Russ. Math. Surv., **18**, 6, 85-191 (1963)

[2] *Arnold V. I.* Mathematical Methods of Classical Mechanics: Graduate Texts in Mathematics 60. Springer-Verlag, New York (1978), x+462 pp.

[3] *Arnold V. I.* Poly-integrable flows. St. Petersburg Math. J., **4**, 6, 1103-1110 (1993)

[4] *Arnold V. I., Kozlov V. V., Neishtadt A. I.* Mathematical Aspects of Classical and Celestial Mechanics. Dynamical systems, III, Encyclopaedia Math. Sci., vol. 3, Springer-Verlag, Berlin (2006), xiv+518 pp.

[5] *Bogolyubov N. N., Mitropolskij Yu. A.* Asymptotic Methods in the Theory of Non-Linear Oscillations. Hindustan Publ. Corp., Delhi; Gordon and Breach Sci. Publ., New York (1961), x+537 pp.

[6] *Bolsinov A. V., Fomenko A. T.* Introduction to the Topology of Integrable Hamiltonian Systems. Nauka, Moscow (1997), 352 pp. (Russian)

[7] *Bourland F. J., Haberman R.* Separatrix crossing: time-invariant potentials with dissipation. SIAM J. Appl. Math., **50**, 6, 1716-1744 (1990)

[8] *Brin M., Freidlin M.* On stochastic behavior of perturbed Hamiltonian systems. Ergod. Th. & Dynam. Sys., **20**, 55-76 (2000)

[9] *Cary J. R., Escande D. F., Tennyson, J. L.* Adiabatic invariant change due to separatrix crossing. Phys. Rev. A, **34**, 5, 4256-4275 (1986)

[10] *Cary J. R., Skodje R. T.* Phase change between separatrix crossings. Phys. D, **36**, 3, 287-316 (1989)

[11] *Diminnie D., Haberman R.* Slow passage through homoclinic orbits for the unfolding of a saddle-center bifurcation and the change in the adiabatic invariant. Phys. D, **162**, 1/2, 34-52 (2002)

[12] *Fenichel N.* Geometric singular perturbation theory for ordinary differential equations. J. of Differ. Equat., **31**, 53–98 (1979)

[13] *Freidlin M. I.* Random and deterministic perturbations of nonlinear oscillators. Doc. Math., J. DMV, Extra Vol. ICM Berlin 1998, vol. III, 223–235 (1998)

[14] *Goldreich P., Peale S.* Spin-orbit coupling in the Solar System. Astron. J., **71**, 6, 425–438 (1966)

[15] *Gurevich A. V., Tsedilina E. E.* Long Distance Propagation of Hf Radio Waves (Physics and Chemistry in Space). Springer-Verlag, Berlin (1985), 344 pp.

[16] *Haberman R.* Slow passage through a transcritical bifurcation for Hamiltonian systems and the change in action due to a nonhyperbolic homoclinic orbit. Chaos, **10**, 3, 641-648 (2000)

[17] *Hartman P.* Ordinary Differential Equations. Classics in Applied Mathematics, 38. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (2002) xx+612 pp.

[18] *Kiselev O. M., Glebov S. G.* An asymptotic solution slowly crossing the separatrix near a saddle-centre bifurcation point. Nonlinearity, **16**, 1, 327-362 (2003)

[19] *Landau L. D., Lifshits E. M.* Course of Theoretical Physics. Vol. 1: Mechanics. Addison-Wesley, Reading MA (1960), 165 pp.

[20] *Lebovitz N.R., Pesci A.* Dynamic bifurcation in Hamiltonian systems with one degree of freedom. SIAM J. Appl. Math., **55**, 4, 1117-1133 (1995)

[21] *Lifshitz I. M., Slutskin A. A., Nabutovskii V.M.* The scattering of charged quasi-particles from singularities in p-space. Soviet Phys. Dokl., **6**, 238-240 (1961)

[22] *Neishtadt A.I.* On the evolution of rotation of a rigid body under the action of a sum of a constant and a dissipative perturbing moments. Izv. Akad. Nauk SSSR, Mekh. Tverd. Tela, **6**, 30-36 (1980) (Russian)

[23] *Neishtadt A. I.* Change of an adiabatic invariant at a separatrix. Sov. J. Plasma Phys., **12**, 568-573 (1986)

[24] *Neishtadt A. I.* On the change in the adiabatic invariant on crossing a separatrix in systems with two degrees of freedom. J. Appl. Math. Mech., **51**, 5, 586-592 (1987)

[25] *Neishtadt A. I.* Problems of Perturbation Theory for Non-Linear Resonant Systems. Doktor Diss., Moscow Univ., Moscow (1989), 342 pp. (Russian)

[26] *Neishtadt A. I.* Probability phenomena due to separatrix crossings. Chaos, **1**, 1, 42-48 (1991)

[27] *Timofeev A.V.* On the constancy of an adiabatic invariant when the nature of the motion changes. Sov. Phys., JETP, **48**, 656-659 (1978)

[28] *Wolansky G.* Limit theorem for a dynamical system in the presence of resonances and homoclinic orbits. J. of Differ. Equat., 1990, **83**, 2, 300–335.