Unbounded growth of energy in nonautonomous Hamiltonian systems

Abstract

The result of J. Mather on the existence of trajectories with unbounded energy for time periodic Hamiltonian systems on a torus is generalized to a class of multidimensional Hamiltonian systems with Hamiltonian polynomial in momenta. It is assumed that the leading homogeneous term of the Hamiltonian is autonomous and the corresponding Hamiltonian system has a hyperbolic invariant torus possessing a transversal homoclinic trajectory. Under certain Melnikov type condition, the existence of trajectories with unbounded energy is proved. Instead of the variational methods of Mather, a geometrical approach based on KAM theory and the Poincaré–Melnikov method is used. This makes it possible to study a more general class of Hamiltonian systems, but requires additional smoothness assumptions on the Hamiltonian.

Mathematics Subject Classification 58F05

1 Introduction

The paper deals with a simplified version of the phenomenon which is usually called the Arnold diffusion. This phenomenon was first discovered by Arnold in his famous example [1]. The term Arnold diffusion was introduced by Chirikov [7]. A brief description is as follows. Consider an analytic time periodic Hamiltonian system

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

with Hamiltonian

$$H(q,p,t,\varepsilon) = H_0(p) + \varepsilon H_1(q,p,t,\varepsilon), \qquad q \in \mathbf{T}^m, \quad p \in \mathbf{R}^m, \tag{1.2}$$

where $\mathbf{T}^m = \mathbf{R}^m/(2\pi\mathbf{Z})^m$ is the m-dimensional torus. For $\varepsilon = 0$, the system is integrable and the action p is constant along the trajectories. The problem is to evaluate the action for small $\varepsilon > 0$. The investigation of Hamiltonian systems (1.1)–(1.2) was called by Poincaré [27] the basic problem of dynamics.

Suppose that the unperturbed Hamiltonian H_0 is not too degenerate: for example, H_0 is a convex function. Then KAM theory implies that for $m = 1$ the action p will be almost constant: $|p(t) - p(0)| \leq C\sqrt{\varepsilon}$ for all $t \in \mathbf{R}$. Indeed, invariant two-dimensional KAM-tori separate the three-dimensional extended phase space.

For $m \geq 2$ the argument based on KAM theory is no longer true. Arnold conjectured [1] that for $m \geq 2$ the action can change considerably: for a generic analytic perturbation H_1 there exists $C > 0$ such that for arbitrary small $\varepsilon > 0$ there exists a solution with $|p(T) - p(0)| \geq C$ for some T. The drift of the action variable is called diffusion because it is very slow for small ε and in numerical experiments is similar to a random walk. As proved by Nekhoroshev [25], for analytic H such that H_0 satisfies a rather weak steepness condition, the diffusion (if it exists) is exponentially slow: a noticeable change of the action p requires an exponentially large time $T \geq ce^{b/\varepsilon^a}$ for some constants $a, b, c > 0$. For quasi-convex Hamiltonians the results of Lochak [20] and Pöschel [28] imply that $a \geq \frac{1}{2(m+1)}$.

Arnold [1] established diffusion for an example of a system with the Hamiltonian depending on two parameters ε, μ, where $\mu > 0$ is exponentially small in ε. This example formally does not belong to the class of Hamiltonian systems (1.1)–(1.2). Also the perturbation was taken to be very degenerate: the system possesses a smooth family of invariant tori.

In the C^∞ case the existence of the Arnold diffusion was proved by Douady [12]. The average diffusion speed in smooth systems in general is greater than exponentially small, but less than any power of the small parameter. A generic smooth case is much simpler to analyze than the analytic one.

As far as we know, for the basic problem of dynamics in the form of Poincaré, there are still no examples of more or less general form published where the Arnold diffusion is established in analytic systems. The main problem is to estimate the intersection angle of stable and unstable manifolds for hyperbolic invariant tori. In some cases the angle can be approximated by the Poincaré–Melnikov integral, which turns out to be exponentially small [9, 15] (see also [6, 30]). In the case of exponentially small separatrix splitting the Poincaré–Melnikov integral in general does not give the main term of the asymptotic formula for the splitting rate [33]. One can rely on the prediction of the splitting rate given by the Poincaré–Melnikov integral only provided that the perturbation is chosen in some specific way [10, 30]. Another problem in the proof of the Arnold diffusion are the gaps in the Cantor family of hyperbolic invariant tori. So, the proof of the Arnold diffusion for a generic analytic Hamiltonian system is still absent.

Bessi [4] used variational methods to study the Arnold diffusion in the original Arnold example and its modifications. The trajectory with large increment of the action variable was constructed as a point of local minimum for the action functional. The construction of a box in the function space, where the existence of a minimum point was established, again relied on the Poincaré–Melnikov integral. Hence applying this method to the basic problem of dynamics encounters similar difficulties.

J. Mather suggested a general variational method for constructing connecting orbits for invariant sets of positive definite time periodic Hamiltonian systems [23]. This approach seems very promising. The method is based on a generalization of the Peierls barrier function in the Aubry–Mather theory to the multidimensional case. For a Hamiltonian system which is close to an integrable one, the barrier function is an analog of the Poincaré–Melnikov integral.

Mather conjectured that by using these methods one can prove the following statement. Let M be a compact manifold with Riemannian metric $\|\dot q\|^2 = \langle A(q)\dot q, \dot q\rangle$, $\dot q \in T_q M$. Consider a classical Hamiltonian system with smooth time-periodic Hamiltonian

$$H(q,p,t) = \frac{1}{2}\|p\|^2 + V(q,t), \qquad q \in M, \quad p \in T_q^* M, \tag{1.3}$$

on $T^*M \times \mathbf{T}$, where $\mathbf{T} = \mathbf{R}/2\pi\mathbf{Z}$ and $\|p\|^2 = \langle p, A^{-1}(q)p\rangle$. The conjecture is that for a generic potential V, there exists a trajectory with $\|p(t)\| \to \infty$ as $t \to \infty$.

In an unpublished manuscript [23] Mather proved this conjecture for M a two-dimensional torus with a generic Riemannian metric. Next we formulate his result.

There exists a minimizing closed geodesic σ in an arbitrary simple homotopy class Γ of closed curves in M. If the minimizing geodesic in Γ is unique, it follows from the results of Morse [24] that σ possesses a homoclinic trajectory γ. Suppose that σ is nondegenerate and the homoclinic trajectory γ is transversal. There exist $a, b \in \mathbf{R}$ such that $\operatorname{dist}(\gamma(t), \sigma(t+a)) \to 0$ as $t \to -\infty$, and $\operatorname{dist}(\gamma(t), \sigma(t+b)) \to 0$ as $t \to +\infty$. The numbers a, b are defined mod the period l of the geodesic σ. Without loss of generality suppose that the average of V over σ vanishes: $\int_0^l V(\sigma(\tau), t)\, d\tau = 0$ for all t. Then the Poincaré–Melnikov function

$$I(t) = \lim_{T\to\infty} \left(\int_{-T}^{T} V(\gamma(\tau), t)\, d\tau - \int_{-T+a}^{T+b} V(\sigma(\tau), t)\, d\tau \right), \qquad t \in \mathbf{T}, \tag{1.4}$$

is correctly defined, i.e., it does not depend on the choice of a, b. Mather proved that if I is nonconstant, then there exists a trajectory such that $\|p(t)\| \to \infty$ as $t \to \infty$. Note that for a generic metric and a generic potential all these conditions are satisfied.

Mather's result may be regarded as a simplified version of the Arnold diffusion. Indeed, for large $\|p\|$ of order ε^{-1}, perform the canonical transformation

$$p = \hat p/\varepsilon, \quad H = \widehat H/\varepsilon^2, \quad t = \varepsilon \hat t.$$

Then the new Hamiltonian

$$\widehat{H} = \frac{1}{2}\|\hat{p}\|^2 + \varepsilon^2 V(q, \varepsilon\hat{t})$$

looks similar to (1.2). The diffusion in Mather's system is not exponentially slow, because the unperturbed system (geodesic flow) has a hyperbolic invariant torus σ (in (1.2), the hyperbolic tori appear only for $\varepsilon \neq 0$), and the time dependence is slow.

In the present paper we obtain a multidimensional version of Mather's result. We study Hamiltonians of a more general form compared with (1.3) and replace the periodic orbit by a hyperbolic invariant torus. Instead of the variational methods, we use a more traditional approach based on some results from KAM theory and the classical Poincaré–Melnikov method. Because of the method we use, stronger regularity assumptions are necessary. Below we assume the Hamiltonian to be real-analytic. Certainly, our results hold also for sufficiently smooth H, but the corresponding KAM technique will be more complicated.

Similar multidimensional results can be obtained also by variational methods, following the approach of Mather. This approach does not need analyticity. However, it requires convexity assumptions on the Hamiltonian, and the invariant torus must be a minimizing set in Mather's sense [22]. Such a torus always possesses a homoclinic trajectory [5]. However, there are no variational results on the existence of minimizing tori. Up to now, all existence results are based on KAM theory. We do not discuss variational methods in this paper.

After this paper was submitted, there appeared the preprint [11], where Mather's result is proved by geometrical methods similar to the ones used in the present paper.

2 Main result

Let a real-analytic m-dimensional manifold M be the configuration space of the Hamiltonian system with Hamiltonian $H = H(q, p, t)$, where $q \in M$ and $p \in T_q^* M$ is the corresponding momentum. The function $H \in C^\omega(T^*M \times \mathbf{T})$ is assumed to be 2π-periodic in time t. The symplectic structure on the phase space $\mathcal{M} = T^*M$ is standard: $\omega = dp \wedge dq$. In local symplectic coordinates, the Hamiltonian system takes the usual form (1.1).

Suppose that H satisfies the following conditions (H1)–(H4).

(H1) H is polynomial in the momenta, p, with analytic coefficients and its leading homogeneous form, H_0, does not depend on time.

Then H has the form

$$H(q, p, t) = H_0(q, p) + \widetilde{H}_0(q, p) + H_1(q, p, t) + \widetilde{H}_1(q, p, t), \tag{2.1}$$

where the functions H_0 and H_1 are homogeneous in p of order $k_0 > 1$, $k_1 \geq 0$ respectively, $k_0 > k_1$, and

$$\deg_p \widetilde{H}_0(q,p) < k_0, \quad \deg_p \widetilde{H}_1(q,p,t) < k_1.$$

The basic example one can keep in mind is as follows:

$$H = T(q,p) + V(q,t), \tag{2.2}$$

where T is a quadratic form in p. In applications T is usually positive definite. Then we have the natural system (1.3) studied by Mather. The first term in the Hamiltonian (2.2) is usually associated with the kinetic energy and the second one with the potential energy of the system.

(H2) Assume that the autonomous system with Hamiltonian H_0 has a hyperbolic nondegenerate Diophantine $(m-1)$-dimensional invariant torus $\mathcal{N} \subset \mathcal{M}$. This means, in particular, that $\mathcal{N}$ is the image of a smooth embedding $\Phi : \mathbf{T}^{m-1} \to \mathcal{M}$ and the phase flow $g_t|_{\mathcal{N}}$ is quasiperiodic with Diophantine frequency vector $\nu \in \mathbf{R}^{m-1}$.

We postpone the definition of a hyperbolic invariant torus to the next section. Here we just recall that the hyperbolic invariant torus $\mathcal{N}$ has m-dimensional stable Γ^+ and unstable Γ^- invariant asymptotic manifolds, and the flows $g_t|_{\Gamma^+}$ and $g_{-t}|_{\Gamma^-}$ contract $\Gamma^\pm$ to $\mathcal{N}$ with a nonzero Lyapunov exponent. Now we continue to formulate the conditions of the main theorem.

The stable and unstable asymptotic manifolds $\Gamma^\pm$ of the hyperbolic torus $\mathcal{N}$ in general intersect along several homoclinic trajectories.

(H3) Assume that there exists a transversal homoclinic trajectory $\gamma : \mathbf{R} \to \mathcal{M}$, i.e., Γ^+ and Γ^- intersect transversely at every point, $\gamma(t)$, on the corresponding energy level, i.e., the intersection $T_{\gamma(t)}\Gamma^+ \cap T_{\gamma(t)}\Gamma^-$ is one-dimensional and is generated by $\dot\gamma(t)$.

The last assumption is

(H4) The Poincaré–Melnikov function

$$\begin{aligned} I(t) &= \lim_{T\to+\infty} \Big[\int_{-T}^{T} (H_1(\gamma(\tau),t) - \overline{H}_1(t))\, d\tau - \chi(\gamma(T),t) + \chi(\gamma(-T),t) \Big] \\ &= \int_{-\infty}^{\infty} (H_1(\gamma(\tau),t) - \overline{H}_1(t) - \dot\chi(\gamma(\tau),t))\, d\tau \end{aligned} \tag{2.3}$$

is nonconstant.

Here $\overline{H}_1(t)$ is the average of the function $H_1(\,\cdot\,,t)$ along the torus $\mathcal{N}$:

$$\overline{H}_1(t) = \frac{1}{(2\pi)^{m-1}} \int_{\mathbf{T}^{m-1}} H_1(\Phi(x),t)\, dx, \tag{2.4}$$

and χ is an arbitrary smooth function on $\mathcal{M} \times \mathbf{T}$ such that for any $w \in \mathcal{N}$,

$$\dot{\chi}(w,t) = H_1(w,t) - \overline{H}_1(t).$$

The time derivative $\dot{\chi} = \{H_0, \chi\}$ is taken with respect to the Hamiltonian system with Hamiltonian H_0, with t regarded as a parameter. Equivalently, the function $f = \chi \circ \Phi$ is a solution of the equation

$$\partial_\nu f(x,t) = H_1(\Phi(x),t) - \overline{H}_1(t), \qquad \partial_\nu = \sum_j \nu_j \frac{\partial}{\partial x_j}.$$

Due to the Diophantine condition, this equation has a unique analytic solution with zero average. Obviously, the integral (2.3) converges and is independent of how the function χ is defined outside $\mathcal{N} \times \mathbf{T}$. It is easy to check that for $m = 2$ and H of the form (1.3), the Poincaré functions (1.4) and (2.3) coincide.

Theorem 2.1 *Suppose that assumptions (H1)–(H4) are satisfied. Then there exists a solution $(q(t), p(t)) \in T^*M$ of the Hamiltonian system such that $q(t)$ lies in a compact domain in M while $\|p(t)\| \to \infty$ as $t \to \infty$.*

Here $\|\cdot\|$ is the norm defined by any Riemannian metric on M. In the case of the "physical" Hamiltonian (2.2) we can say that the energy H of the solution $(q(t), p(t))$ tends to infinity.

Remark 2.1 *Theorem 2.1 holds for $\mathcal{N}$ a hyperbolic invariant torus of arbitrary dimension if it is assumed to be reducible (see the definition in the next section). Note that hyperbolic tori obtained by KAM-theory are always reducible [16, 37]. Probably a similar result holds for irreducible tori, but, as far as we know, there is no KAM-type result in the literature which can be applied in this case.*

Remark 2.2 *It is possible to estimate how fast $\|p(t)\|$ tends to infinity along the solution in Theorem 2.1. A rough estimate shows that there exists a solution with polynomial average speed: for any $\varepsilon > 0$, $\|p(t)\| > \varepsilon^{-1}$ for $t > T(\varepsilon)$, where $T(\varepsilon) \le c\varepsilon^{-a}$. The constant $a > 0$ is sufficiently large. By using the methods of [21], one can get an upper estimate for a. However, the estimates we are able to get are far from being optimal. To get more realistic estimates, probably one should use variational methods [4]. Note that the energy theorem $\dot{H} = \partial H/\partial t$ implies that $T(\varepsilon) \ge c\varepsilon^{k_1 - k_0}$. Hence $a \ge k_0 - k_1$. As in Mather's result, diffusion speed is polynomial in ε, because the system with Hamiltonian H_0 has a hyperbolic behavior near $\mathcal{N}$.*

Remark 2.3 *If M is an analytic two-dimensional compact Riemannian manifold of genus greater than one, then there always exists a hyperbolic periodic geodesic possessing a transverse homoclinic trajectory [18]. Hence no conditions on the metric are needed: for a generic analytic V the Hamiltonian system with Hamiltonian (2.2) has a trajectory with unbounded energy.*

An idea of the proof of Theorem 2.1 is as follows. First, we establish the existence of an infinite set of hyperbolic m-dimensional invariant tori $\{\mathcal{N}_\varepsilon\} \subset \mathcal{M} \times \mathbf{T}$ for all $\varepsilon > 0$ belonging to a Cantor set $\mathcal{E} \subset \mathbf{R}$ with a density point at 0. The frequency vector associated with the torus $\{\mathcal{N}_\varepsilon\}$ is $(\varepsilon^{1-k_0}\nu, 1) \in \mathbf{R}^m$, where the frequency 1 corresponds to the variable $t \in \mathbf{T}$ and ε is assumed to be small enough. When $\varepsilon \to 0$, $\|p|_{\mathcal{N}_\varepsilon}\| \to \infty$ as ε^{-1}, while the projection of the torus $\mathcal{N}_\varepsilon$ to the configuration space M remains close to the projection of $\mathcal{N}$.

The tori $\mathcal{N}_\varepsilon$ are small perturbations of the torus $\mathcal{N}$ in the following sense. Let $h_\varepsilon : \mathcal{M} \times \mathbf{T} \to \mathcal{M} \times \mathbf{T}$ be the homothety:

$$h_\varepsilon(q, p, t) = (q, \varepsilon p, t). \tag{2.5}$$

Then for small $\varepsilon > 0$ the torus $N_\varepsilon = h_\varepsilon(\mathcal{N}_\varepsilon)$ is close to $\mathcal{N} \times \mathbf{T}$. Let $\Gamma_\varepsilon^\pm \subset \mathcal{M} \times \mathbf{T}$ be the local stable and unstable manifolds of the torus $\mathcal{N}_\varepsilon$. Then the manifolds $W_\varepsilon^\pm = h_\varepsilon(\Gamma_\varepsilon^\pm)$ are close to $\Gamma^\pm \times \mathbf{T}$. Moreover, the torus N_ε and the manifolds $W_\varepsilon^\pm$ are Whitney smooth functions of $\varepsilon \in \mathcal{E}$.

Second, we show that each torus $\mathcal{N}_\varepsilon$ has a topologically transversal homoclinic trajectory γ_ε. If I has a nondegenerate critical point, then γ_ε is transversal, and the minimal "intersection angle" of Γ_ε^+ and Γ_ε^- can be estimated in powers of ε. However, the intersection can be nontransversal if the Poincaré function I is degenerate. Then we establish the existence of a heteroclinic connection between pairs of tori $\mathcal{N}_\varepsilon$ and $\mathcal{N}_{\varepsilon'}$ for small and sufficiently close to one another values of ε and ε', i.e., the existence of a topologically transversal heteroclinic orbit in $\Gamma_\varepsilon^- \cap \Gamma_{\varepsilon'}^+$, asymptotic to $\mathcal{N}_\varepsilon$ as $t \to -\infty$, and to $\mathcal{N}_{\varepsilon'}$ as $t \to \infty$.

This gives a decreasing sequence $\varepsilon_n \in \mathcal{E}$, $\varepsilon_n \to 0$, such that $\Gamma_{\varepsilon_n}^-$ has a topologically transversal intersection with $\Gamma_{\varepsilon_{n+1}}^+$. As shown by Arnold [1], the existence of such a chain of heteroclinic connections implies the existence of the trajectory $(q(t), p(t))$ we need.

We prove Theorem 2.1 in sections 4–5. In section 3 the definition and basic properties of hyperbolic invariant tori are discussed. In section 4 the existence of a Whitney family $\mathcal{N}_\varepsilon$ of invariant tori is proved. In section 5 the existence of intersection points for Γ_ε and $\Gamma_{\varepsilon'}$ with small and close $\varepsilon, \varepsilon'$ is established. The Appendices contain proofs of several technical results used in the paper.

3 Definition and basic properties of a hyperbolic torus

In this section we put aside our main problem and discuss the general concept of hyperbolicity for an invariant torus of a Hamiltonian system. The notion of a hyperbolic torus is widely used in the literature, but, as far as we know, there is no universally accepted definition. Arnold [1] named the hyperbolic tori as whiskered tori.

Let a $2m$-dimensional symplectic manifold $(\mathcal{M}, \omega)$ be the phase space of an autonomous Hamiltonian system with Hamiltonian $\mathcal{H}$. Denote by $v_\mathcal{H}$ the corresponding Hamiltonian vector field on $\mathcal{M}$ and by $g_t : \mathcal{M} \to \mathcal{M}$ the phase flow of the system. Let $\mathcal{N} \subset \mathcal{M}$ be an invariant n-dimensional torus with frequency vector $\nu \in \mathbf{R}^n$. Thus $\mathcal{N}$ is the image of a smooth embedding $\Phi : \mathbf{T}^n \to \mathcal{M}$ and the equality

$$\partial_\nu \Phi = D\Phi \cdot \nu = v_{\mathcal{H}} \circ \Phi$$

is satisfied. Equivalently, $\Phi(x + \nu t) \equiv g_t(\Phi(x))$. Consider the linearized phase flow

$$Dg_t : T_{\mathcal{N}}\mathcal{M} \to T_{\mathcal{N}}\mathcal{M}, \qquad v \in T_w\mathcal{M} \mapsto Dg_t(w)v \in T_{g_t(w)}\mathcal{M}, \quad w \in \mathcal{N}.$$

It is defined by a linear differential equation which is called the variational equation for the invariant torus $\mathcal{N}$.

Definition. The torus $\mathcal{N}$ is called hyperbolic if $n < m$ and there exist two smooth $(m-n)$-dimensional subbundles $E^\pm$ of the bundle $T_{\mathcal{N}}\mathcal{M}$ such that

1. $E^\pm$ are invariant for the linearized phase flow:

$$Dg_t(w)E^\pm_w = E^\pm_{g_t(w)}, \qquad w \in \mathcal{N}, \quad t \in \mathbf{R}.$$

2. The linearized flow is contracting on E^+ and expanding on E^-, i.e., for some positive constants C and λ, we have

$$\begin{aligned} \|Dg_t(w)v\| &\le Ce^{-\lambda t}\|v\|, \qquad w \in \mathcal{N}, \quad v \in E^+_w, \quad t \ge 0, \\ \|Dg_{-t}(w)v\| &\le Ce^{-\lambda t}\|v\|, \qquad w \in \mathcal{N}, \quad v \in E^-_w, \quad t \ge 0. \end{aligned}$$

Here $\|\cdot\|$ is any Riemannian metric on $\mathcal{M}$. Since $\mathcal{N}$ is compact, the definition does not depend on the choice of the metric. Passing to a finite covering of the torus $\mathcal{N}$ (and its tubular neighborhood), without loss of generality we can assume that the bundles $E^\pm$ are trivial.

Almost the same definition of a hyperbolic torus takes place in the nonautonomous case provided the Hamiltonian depends on time periodically with period τ. The difference is that instead of $\mathcal{M}$ one should take the extended phase space $\mathcal{M} \times (\mathbf{R}/\tau\mathbf{Z})$. In particular, the phase flow g_t acts on $\mathcal{M} \times (\mathbf{R}/\tau\mathbf{Z})$. In this case any hyperbolic torus of the system has $2\pi/\tau$ as the frequency corresponding to the rotation of the time variable. Hence the frequency vector equals $(\nu, 2\pi/\tau)$.

Definition. If the vector ν is nonresonant, the corresponding hyperbolic torus is called nonresonant. If ν is Diophantine, i.e., there exist positive constants α and β such that

$$|\langle \nu, k\rangle| \ge \alpha |k|^{-\beta} \qquad \text{for any nonzero} \quad k \in \mathbf{Z}^n, \tag{3.1}$$

the torus is called Diophantine.

Here $\langle\, , \rangle$ is the standard Euclidean scalar product. In the nonautonomous case the nonresonance and Diophantine properties for a hyperbolic torus are formulated in terms of the frequency vector $(\nu, 2\pi/\tau)$.

> **Definition.** A hyperbolic torus $\mathcal{N}$ is called nondegenerate if all bounded solutions of the variational equation are tangent to $\mathcal{N}$. Thus $\|Dg_t(w)v\| \le C$, $w \in \mathcal{N}$, for all t implies that $v \in T_w\mathcal{N}$.

For $n = 1$ the definition of a nondegenerate hyperbolic torus coincides with the definition of a hyperbolic periodic orbit in a Hamiltonian system.

As a model example consider the autonomous system with Hamiltonian

$$\mathcal{H}(x, y, z_-, z_+) = \langle \nu, y \rangle + \langle Ay, y \rangle/2 + \langle \Omega(x) z_-, z_+ \rangle + O_3(y, z), \tag{3.2}$$

on $\mathcal{M} = \mathbf{T}^n\{x\} \times \mathbf{R}^n\{y\} \times \mathbf{R}^{m-n}\{z_-\} \times \mathbf{R}^{m-n}\{z_+\}$, where the variables $y \in \mathbf{R}^n$, $z_\pm \in \mathbf{R}^{m-n}$ are defined in the vicinity of zero;

$$\omega = dy \wedge dx + dz_+ \wedge dz_-;$$

the symmetric $(n \times n)$-matrix A is constant; the matrix $\Omega(x) + \Omega^T(x)$ is positive definite [1] for any $x \in \mathbf{T}^n$ and the function $O_3(y, z_\pm)$, which can also depend on x, has at least third order in y and $z_\pm$.

Proposition 3.1 *The invariant torus $\mathcal{N} = \{(x, y, z) : y = 0, z_\pm = 0\}$ is hyperbolic.*

Indeed, the Hamiltonian equations have the form

$$\begin{cases} \dot{x} &= \nu + Ay & +O_2(y, z), \\ \dot{y} &= & O_2(y, z), \\ \dot{z}_- &= \Omega(x) z_- & +O_2(y, z), \\ \dot{z}_+ &= -\Omega^T(x) z_+ & +O_2(y, z). \end{cases}$$

Hence, we can put $\Phi(x) = (x, 0, 0, 0)$. The variational equations on

$$\mathbf{T}_{\mathcal{N}}\mathcal{M} = \mathbf{T}^n\{x\} \times (\mathbf{R}^n\{\xi\} \oplus \mathbf{R}^n\{\eta\} \oplus \mathbf{R}^{m-n}\{\zeta_+\} \oplus \mathbf{R}^{m-n}\{\zeta_-\})$$

take the form

$$\begin{cases} \dot{x} &= \nu, \\ \dot{\xi} &= A\eta, \\ \dot{\eta} &= 0, \\ \dot{\zeta}_- &= \Omega(x)\zeta_-, \\ \dot{\zeta}_+ &= -\Omega^T(x)\zeta_+. \end{cases} \tag{3.3}$$

Let us show that the subbundles $E^{\pm} = \mathbf{T}^n \times \mathbf{R}^{m-n}\{\zeta_{\pm}\}$ satisfy the definition of a hyperbolic torus. The bundles $E^{\pm}$ are obviously invariant. The last equation (3.3) implies that

$$\frac{d}{dt}\langle\zeta_+,\zeta_+\rangle = -\langle(\Omega(x)+\Omega^T(x))\zeta_+,\zeta_+\rangle \le -\lambda\langle\zeta_+,\zeta_+\rangle, \tag{3.4}$$

where

$$\lambda = \min_{x\in\ ^n}\{\text{minimal eigenvalue of the operator } \Omega(x)+\Omega^T(x)\}.$$

Since $\Omega+\Omega^T$ is positive definite, $\lambda > 0$. According to the estimate (3.4), the subbundle E^+ is contracting. Analogously

$$\frac{d}{dt}\langle\zeta_-,\zeta_-\rangle = \langle(\Omega(x)+\Omega^T(x))\zeta_-,\zeta_-\rangle \ge \lambda\langle\zeta_-,\zeta_-\rangle.$$

Hence, E^- is expanding.

> **Definition.** The coordinates (x, y, z) in the vicinity of a hyperbolic torus $\mathcal{N}$ are called canonical for $\mathcal{N}$. If canonical coordinates for a hyperbolic torus exist, the torus is called reducible.

Canonical coordinates for a hyperbolic torus of a nonautonomous system are defined analogously. The only difference is that the matrix Ω and the terms $O_3(y, z)$ in the Hamiltonian (3.2) can depend periodically on time. In this case the torus is $(n+1)$-dimensional and the frequency vector is $(\nu, 2\pi/\tau) \in \mathbf{R}^{n+1}$, where τ is the period of the Hamiltonian with respect to t.

Proposition 3.2 *A reducible hyperbolic torus is nondegenerate if and only if* $\det A \neq 0$.

Indeed, in the autonomous case any solution lying on $\mathcal{N}$ has the form $(x, y, z) = (x^0 + \nu t, 0, 0)$, where x^0 is the initial condition. Any bounded solution of the variational equations (3.3) has the form

$$(x,\xi,\eta,\zeta) = (x^0+\nu t, \xi^0 + A\eta^0 t, \eta^0, 0), \quad A\eta^0 = 0.$$

All these solutions are tangent to $\mathcal{N}$ if and only if the matrix A is nondegenerate. In the nonautonomous case the argument is analogous.

Lemma 3.1 *If the symplectic structure is exact, any $(m-1)$-dimensional Diophantine hyperbolic torus of an autonomous Hamiltonian system is reducible. Moreover, there exist canonical coordinates such that Ω is constant.*

The proof of the lemma is given in Appendix A.

Below we consider the case of $(m-1)$-dimensional hyperbolic tori. However, the majority of the assertions we present, hold in the general case as well.

Reducible Diophantine tori survive small perturbations of the system. More precisely, assume that the Hamiltonian

$$\mathcal{H}_0 = \mathcal{H}_0(x, y, z_-, z_+, \varepsilon) = \langle \nu, y\rangle + \langle Ay, y\rangle/2 + \lambda z_- z_+ + O_3(y, z)$$

is analytic in the complex domain

$$D_{\rho,\varepsilon_0} = \{(x, y, z_\pm, \varepsilon) : |\mathrm{Im}\, x| \le \rho, |y| \le \rho, |z_\pm| \le \rho, |\varepsilon| \le \varepsilon_0\}$$

and $\max_{D_{\rho,\varepsilon_0}} |\mathcal{H}_0| = \chi_0$. Assume also that $\lambda_0 = \lambda|_{\varepsilon=0}$ is a positive constant, the matrix $A_0 = A|_{\varepsilon=0}$ is constant and for some matrix norm $\|\cdot\|$,

$$\|A_0\| \le a_1, \quad \|A_0^{-1}\| \le a_2.$$

Let the vector ν be independent of ε and satisfy the Diophantine condition (3.1).

Theorem 3.1 *There exists a constant $\chi > 0$, depending only on ρ, χ_0, λ_0, α, β, a_1, a_2, such that for any real-analytic perturbation $\varepsilon\mathcal{H}_1(x, y, z_-, z_+, \varepsilon)$, satisfying the inequality*

$$\max_{D_{\rho,\varepsilon_0}} |\varepsilon\mathcal{H}_1| \le \chi$$

the system with Hamiltonian $\mathcal{H}_0 + \varepsilon\mathcal{H}_1$ has a hyperbolic reducible torus $\mathcal{N}(\varepsilon)$ with the same frequency vector ν. The torus $\mathcal{N}(\varepsilon)$ is analytic in ε. Moreover, the change $(x, y, z_\pm) \mapsto (\hat{x}, \hat{y}, \hat{z}_\pm)$ introducing canonical coordinates for $\mathcal{N}(\varepsilon)$ is analytic in ε.

Corollary 3.1 *In the new coordinates*

$$\mathcal{H}_0 + \varepsilon\mathcal{H}_1 = \mathrm{const} + \langle \nu, \hat{y}\rangle + \langle \hat{A}(\varepsilon)\hat{y}, \hat{y}\rangle/2 + \hat{\lambda}(\varepsilon)\hat{z}_-\hat{z}_+ + O_3(\hat{y}, \hat{z}),$$

where the functions $\hat{A}$, $\hat{\lambda}$ are real-analytic and $\hat{A}(0) = A_0$, $\hat{\lambda}(0) = \lambda_0$.

The proof of Theorem 3.1 is presented in [16, 37].

4 Hyperbolic tori in the domain of large momenta

Now we turn to the proof of Theorem 2.1. Let $\mathcal{N}$ be a hyperbolic nondegenerate Diophantine $(m-1)$-dimensional torus of the system with Hamiltonian H_0 and $\nu \in \mathbf{R}^{m-1}$ its frequency vector. In this section we show that the nonautonomous system with Hamiltonian (2.1) has a smooth Cantor family of m-dimensional hyperbolic tori $\mathcal{N}_\varepsilon \subset \mathcal{M} \times \mathbf{T}$ with the frequency vectors $(\varepsilon^{1-k_0}\nu, 1)$ and the corresponding asymptotic manifolds $\Gamma_\varepsilon^\pm$ also form a smooth Cantor family. In fact we perform a scaling h_ε in $\mathcal{M} = T^*M$ given by (2.5) and construct the torus $N_\varepsilon = h_\varepsilon(\mathcal{N}_\varepsilon) \subset \mathcal{M} \times \mathbf{T}$ and the images $W_\varepsilon^\pm = h_\varepsilon(\Gamma_\varepsilon^\pm) \subset \mathcal{M} \times \mathbf{T}$ of the stable and unstable asymptotic manifolds $\Gamma_\varepsilon^\pm$.

First, we extend the scaling h_ε of T^*M to a canonical transformation. Let $\hat{H} = \varepsilon^{k_0} H \circ h_{\varepsilon^{-1}}$. Then

$$\omega - dH \wedge dt = \varepsilon^{-1} h_\varepsilon^*(\omega - \hat{H} \wedge d\tau),$$

where we denoted $t = \mu\tau$ with $\mu = \varepsilon^\kappa$, $\kappa = k_0 - 1$. Thus the homothety h_ε maps the trajectories of the system with Hamiltonian H to the trajectories of the system with Hamiltonian $\hat{H}$, parameterized by the new time τ.

The new Hamiltonian $\hat{H}$ takes the form

$$\hat{H}(w,t,\varepsilon) = H_0(w) + \varepsilon H_0'(w,\varepsilon) + \varepsilon^s H_1(w,t) + \varepsilon^{s+1} H_1'(w,t,\varepsilon), \tag{4.1}$$

where we denoted $w = (q,p) \in \mathcal{M}$ and $s = k_0 - k_1$. The functions $H_0'(q,p,\varepsilon) = \varepsilon^{k_0-1}\tilde{H}_0(q,\varepsilon^{-1}p)$ and $H_1'(q,p,t,\varepsilon) = \varepsilon^{k_1-1}\tilde{H}_1(q,\varepsilon^{-1}p,t)$ are polynomial in ε.

We obtain an analytic Hamiltonian system on $\mathcal{M}$ with time-periodic Hamiltonian slowly depending on time τ. It will be regarded as an autonomous differential equation on $\mathcal{M} \times \mathbf{T}$ of the form

$$\dot{w} = v_{\hat{H}}(w,t,\varepsilon), \quad \dot{t} = \mu, \tag{4.2}$$

where dot denotes the derivative with respect to τ.

Any hyperbolic torus $N \subset \mathcal{M} \times \mathbf{T}$ of the system (4.2) corresponds to the hyperbolic torus $\mathcal{N} = h_{\varepsilon^{-1}}(N)$ of the initial system. Furthermore, if the frequencies of N equal (ν, μ), then the frequencies of $\mathcal{N}$ are $(\mu^{-1}\nu, 1)$.

Consider the set

$$\mathcal{E} = \{\varepsilon \in (0,\varepsilon_0) : |\langle j,\nu\rangle + l\varepsilon^\kappa| \geq \varepsilon^r(|j|+|l|)^{-\beta_*},\ j \in \mathbf{Z}^{m-1},\ l \in \mathbf{Z},\ |j|+|l| \neq 0\},$$

where $\varepsilon_0 > 0$ is sufficiently small, and the constants r, β_* satisfy the inequalities

$$r \geq \kappa, \quad \beta_* \geq \beta, \quad \beta_* > m-1. \tag{4.3}$$

Below we construct invariant tori for $\varepsilon \in \mathcal{E}$. The set $\mathcal{E}$ occupies a large relative measure on the interval $(0,\varepsilon_0)$.

Proposition 4.1 *Suppose that $\varepsilon_0 > 0$ is small enough.*[2] *For any β_* satisfying (4.3), there exists a positive constant, c, depending only on β_* and m, such that for any $\varepsilon \in (0,\varepsilon_0)$,*

$$\operatorname{meas}([\varepsilon/2,\varepsilon] \setminus \mathcal{E}) < c\varepsilon^{1-\kappa+r},$$

where meas *is the Lebesgue measure on* $\mathbf{R}$.

$\varepsilon \quad < \alpha$

This number-theoretic result is proved in Appendix B.

Corollary 4.1 *Suppose that* $\varepsilon \in (0, \varepsilon_0)$ *and* $c\varepsilon^{1-\kappa+r} \le \varepsilon/2$. *Then there exists* $\varepsilon' \in \mathcal{E}$ *such that* $0 < \Delta\varepsilon = \varepsilon - \varepsilon' < c\,\varepsilon^{1-\kappa+r}$.

Indeed, since $\operatorname{meas}([\varepsilon/2, \varepsilon] \setminus \mathcal{E}) < c\varepsilon^{1-\kappa+r}$, the set $\mathcal{E} \cap (\varepsilon - c\varepsilon^{1-\kappa+r}, \varepsilon)$ is not empty.

Below we assume that $1 - \kappa + r > 2k_0 - k_1$.

To formulate the main result of the present section we briefly recall the definition of Whitney smoothness [36, 32, 29]. Let $\Omega \subset \mathbf{R}$ be a closed set.

> **Definition.** $C^\alpha(\Omega)$ is the class of all functions $u^{(0)} : \Omega \to \mathbf{R}$ together with the "derivatives" $u^{(k)} : \Omega \to \mathbf{R}$, $0 < k < \alpha$, such that for some $N \in \mathbf{R}$
>
> $$|u^{(k)}(x)| \le N,$$
> $$|u^{(k)}(x) - P_k(x, y)| \le N|x - y|^{\alpha - k}$$
>
> for all $x, y \in \Omega$ and $0 \le k < \alpha$. Here
>
> $$P_k(x, y) = \sum_{k+l<\alpha} \frac{1}{l!} u^{(k+l)}(y)(x - y)^l$$
>
> is the analogue of the k-th Taylor polynomial.
>
> The set $C^\infty(\Omega)$ of Whitney smooth functions is defined as $\cap_{\alpha>0} C^\alpha(\Omega)$.

According to Whitney's theorem [36], any Whitney smooth function on Ω can be extended to a smooth function on $\mathbf{R}$.

Put

$$\mathcal{E}^\kappa = \{\mu \in (0, \varepsilon_0^\kappa) : \mu^{1/\kappa} \in \mathcal{E}\}. \tag{4.4}$$

Below it will be convenient to regard the parameters ε and μ as independent. To have Diophantine conditions for the frequency vector (ν, μ), we assume that

$$(\mu, \varepsilon) \in \mathcal{P} = \{(\mu, \varepsilon) : \mu \in \mathcal{E}^\kappa,\ 0 < \varepsilon \le \mu^{1/\kappa}\}. \tag{4.5}$$

Theorem 4.1 *Let* $\varepsilon_0 > 0$ *be sufficiently small. Then for any* $(\mu, \varepsilon) \in \mathcal{P}$, *the system (4.2) has an invariant torus* $N_{\mu,\varepsilon} \subset \mathcal{M} \times \mathbf{T}$ *with the frequency vector* (ν, μ). *The tori* $N_{\mu,\varepsilon}$ *and the corresponding asymptotic manifolds* $W^\pm_{\mu,\varepsilon}$ *are analytic submanifolds of the phase space. They are analytic in* ε *and* C^∞ *Whitney smooth in* μ.

The last assertion means that there exists a map

$$F : \mathbf{T}^{m-1}\{x\} \times \mathbf{T}\{t\} \times (-\delta, \delta)\{z_-\} \times (-\delta, \delta)\{z_+\} \times \mathcal{P}\{\mu, \varepsilon\} \to \mathcal{M} \times \mathbf{T},$$

analytic in the phase variables, t and ε, and C^∞ Whitney smooth in μ such that the manifolds

$$\begin{array}{rcl} N_{\mu,\varepsilon} & = & F(\mathbf{T}^m \times \{0\} \times \{0\} \times \{\mu,\varepsilon\}), \\ W^+_{\mu,\varepsilon(\mathrm{loc})} & = & F(\mathbf{T}^m \times \{0\} \times (-\delta,\delta) \times \{\mu,\varepsilon\}), \\ W^-_{\mu,\varepsilon(\mathrm{loc})} & = & F(\mathbf{T}^m \times (-\delta,\delta) \times \{0\} \times \{\mu,\varepsilon\}). \end{array} \tag{4.6}$$

are invariant, and the restriction of the system to $W^\pm_{\mu,\varepsilon(\mathrm{loc})}$ has the form

$$\left\{ \begin{array}{rcl} \dot{x} & = & \nu, \\ \dot{t} & = & \mu, \\ \dot{z}_\pm & = & \mp z_\pm \lambda(\mu,\varepsilon), \end{array} \right.$$

where $\lambda(0,0) = \lambda_0 > 0$. Thus the phase flow on $W^\pm_{\mu,\varepsilon(\mathrm{loc})}$ is given by

$$g_\tau \circ F(x,t,z_-,0,\mu,\varepsilon) = F(x+\nu\tau, t+\mu\tau, e^{\lambda(\mu,\varepsilon)\tau} z_-, 0, \mu, \varepsilon),$$
$$g_\tau \circ F(x,t,0,z_+,\mu,\varepsilon) = F(x+\nu\tau, t+\mu\tau, 0, e^{-\lambda(\mu,\varepsilon)\tau} z_+, \mu, \varepsilon).$$

To prove Theorem 4.1, it is sufficient to prove

Proposition 4.2 *There exist normal symplectic coordinates $x, y, z_\pm$ in a neighborhood of the torus $N_{\mu,\varepsilon}$, analytic in ε and Whitney smooth in μ, such that the new Hamiltonian has the Eliasson normal form [14]*

$$H^*(x,y,z_-,z_+,t,\mu,\varepsilon) = \langle \nu, y\rangle + \lambda(\mu,\varepsilon) z_- z_+ + O_2(y, z_- z_+). \tag{4.7}$$

Hence the Hamiltonian system takes the form

$$\left\{ \begin{array}{rcl} \dot{x} & = & \nu + O_1(y) + O_2(z_- z_+), \\ \dot{t} & = & \mu, \\ \dot{y} & = & Ay + O_2(y, z_- z_+), \\ \dot{z}_\pm & = & \mp z_\pm \lambda(\mu,\varepsilon) + z_\pm O_1(y, z_- z_+). \end{array} \right.$$

The proof of Proposition 4.2 is sketched in Appendix C.

5 Splitting of the invariant manifolds

We continue to study the system (4.2) with Hamiltonian $\hat{H}$ as in (4.1). By assumptions (H2)–(H3), the unperturbed system with Hamiltonian H_0 has a hyperbolic reducible Diophantine invariant torus $\mathcal{N}$ possessing a transversal homoclinic orbit γ. In this section the dimension $\dim \mathcal{N} = n < m$ can be arbitrary.

Since in general the stable and unstable manifolds are only immersed submanifolds in $\mathcal{M}$, their definition should be more precise. Let $\Gamma^\pm_{\mathrm{loc}}$ be the local stable and unstable manifolds of the torus $\mathcal{N}$. They exist by Theorem 4.1. Take $T > 0$ so large that the homoclinic trajectory γ is contained in the union of the manifolds $\Gamma^- = g_T(\Gamma^-_{\rm loc})$ and $\Gamma^+ = g_{-T}(\Gamma^+_{\rm loc})$, where $g_t : \mathcal{M} \to \mathcal{M}$ is the phase flow of the unperturbed system. Then $\Gamma^\pm$ are positively and negatively invariant, respectively. We can assume that $\gamma(0) \in \Gamma^+ \cap \Gamma^-$ an d

$$\Gamma^+ \cap \Gamma^- \supset \Delta = \{\gamma(t) \in \mathcal{M} : |t| < \delta\}.$$

Let $\mathcal{E}^\kappa$ be the set (4.4). First, we take $\mu \in \mathcal{E}^\kappa$ independent of ε and consider the perturbed Hamiltonian system as a differential equation depending on two parameters μ and $\varepsilon \in [0, \mu^{1/\kappa}]$.

Let $N_{\mu,\varepsilon}$ and $W^\pm_{\mu,\varepsilon({\rm loc})}$ be the manifolds (4.6) constructed in Theorem 4.1. The manifolds $W^\pm_{\mu,\varepsilon}$ are defined as $W^\pm_{\mu,\varepsilon} = \phi^{\mp T}(W^\pm_{\mu,\varepsilon({\rm loc})})$, where $\phi^\tau = \phi^\tau_{\mu,\varepsilon} : \mathcal{M} \times \mathbf{T} \to \mathcal{M} \times \mathbf{T}$ is the phase flow of system (4.2).

Next we formulate a version of the classical Poincaré–Melnikov result. Let χ_μ be an arbitrary smooth function on $\mathcal{M} \times \mathbf{T}$ such that

$$\dot\chi_\mu|_{\mathcal{N}\times} \quad = H_1 - \overline{H}_1(t), \tag{5.1}$$

where $\dot\chi_\mu = \{H_0, \chi_\mu\} + \mu D_t \chi_\mu$ and the average $\overline{H}_1(t)$ is defined in (2.4). Equivalently, for the function $f(x,t) = \chi_\mu(\Phi(x), t)$,

$$\partial f = H_1(\Phi(x), t) - \overline{H}_1(t), \qquad \partial = \partial_\nu + \mu D_t. \tag{5.2}$$

This equation has an unique analytic solution with zero average for $\mu \in \mathcal{E}^\kappa$ due to the Diophantine condition (3.1). In fact, substracting from H a function of the time, we can always assume that $\overline{H}_1(t) \equiv 0$.

Theorem 5.1 *Let t_0 be a minimum or a maximum point of the Poincaré function*

$$I_\mu(t) = \int_{-\infty}^{\infty} (H_1(\gamma(\tau), t + \mu\tau) - \overline{H}_1(t + \mu\tau) - \dot\chi_\mu(\gamma(\tau), t + \mu\tau))\, d\tau. \tag{5.3}$$

If t_0 is nondegenerate, then for sufficiently small $\varepsilon > 0$ the invariant manifolds $W^\pm_{\mu,\varepsilon}$ intersect transversely along a curve $\gamma_{\mu,\varepsilon} \subset M \times \mathbf{T}$ which is close to the curve $\gamma_\mu = \{(\gamma(\tau), t_0 + \mu\tau) : |\tau| < \delta\}$. If t_0 is degenerate, but I_μ is nonconstant, then the intersection of $W^\pm_{\mu,\varepsilon}$ along $\gamma_{\mu,\varepsilon} \subset M \times \mathbf{T}$ is either transversal or topologically transversal.

An intersection point w of the manifolds $W^\pm_{\mu,\varepsilon}$ is called topologically transversal if there exists a homeomorphism of a neighborhood of w such that the images of $W^\pm_{\mu,\varepsilon}$ intersect transversely.

Note that $I_\mu(t)$ is analytic in t and Whitney smooth in μ, and $I_\mu = I + O(\mu)$, where I is the Poincaré function (2.3).

Theorem 5.1 is essentially classical [27, 19, 34]. However, probably it was not formulated explicitly in the situation we study. So, for completeness we include the proof in Appendix D. In fact we will need a more quantitative version of Theorem 5.1 (Proposition D.1).

Put $\mu = \varepsilon^\kappa$. Note that we performed the transformation h_ε, which was different for different ε. Now we perform the inverse transformation and study the intersection of the invariant manifolds $\Gamma_\varepsilon^\pm = h_{\varepsilon^{-1}}(W^\pm_{\mu,\varepsilon})$ corresponding to different values of $\varepsilon, \varepsilon' \in \mathcal{E}$.

Theorem 5.2 *Let ε_0 be sufficiently small.*

- *There exists $c > 0$ such that the manifolds Γ_ε^- and $\Gamma_{\varepsilon'}^+$ intersect for $\varepsilon, \varepsilon' \in \mathcal{E}$ provided that $\Delta\varepsilon = |\varepsilon - \varepsilon'| < c\varepsilon^{2k_0 - k_1}$.*
- *If t_0 is a nondegenerate critical point for I, the manifolds Γ_ε^- and $\Gamma_{\varepsilon'}^+$ intersect transversely along a curve close to $h_{\varepsilon^{-1}}\gamma_\mu$.*
- *If t_0 is a degenerate minimum or maximum point for I, then Γ_ε^- and $\Gamma_{\varepsilon'}^+$ intersect transversely or topologically transversely along a curve close to $h_{\varepsilon^{-1}}\gamma_\mu$.*

The proof is given in Appendix D.

Let $r > 3k_0 - k_1 - 2$. By Corollary 4.1, we can construct a transition chain of heteroclinic orbits. Let $\varepsilon_0, \varepsilon_1, \dots \in \mathcal{E}$, $\varepsilon_n \to 0$, be a decreasing sequence such that $\varepsilon_n - \varepsilon_{n+1} < c\varepsilon_n^{2k_0 - k_1}$. Then for each n the manifold $\Gamma^-_{\varepsilon_{n-1}}$ intersects $\Gamma^+_{\varepsilon_n}$ topologically transversely at some point w_n. It can be assumed that $w_n \in \Gamma^+_{\varepsilon_n(\mathrm{loc})}$. This implies Theorem 2.1 exactly as in Arnold's work [1]. Let $\phi^\tau : \mathcal{M} \times \mathbf{T} \to \mathcal{M} \times \mathbf{T}$ be the phase flow of the system.

Lemma 5.1 *There exists a point $w \in \mathcal{M}$ and a sequence $\tau_n \to \infty$ such that* $\mathrm{dist}\,(\phi^{\tau_n}(w), w_n) \to 0$.

Proof. Take a closed ball U_0 around any point $z_0 \in \Gamma^+_{\varepsilon_0}$. Then the normal form (4.7), the nonresonance of the frequency vector (ν, μ), and the analogue of the λ-lemma [26] for hyperbolic tori [8, 13, 21], imply that for some $t_1 > 0$ the image $\phi^{t_1}(U_0)$ is a neighborhood of a point $z_1 \in \Gamma^+_{\varepsilon_1}$.

Indeed, take a short segment $\sigma \subset \Gamma^+_{\varepsilon_1}$ of a curve passing through the point w_1. Then σ crosses $\Gamma^-_{\varepsilon_0}$. By Proposition 4.2 and the λ-lemma proved in [8, 13], $\overline{\cup_{\tau > 0}\phi^{-\tau}(\sigma)} \supset \Gamma^+_{\varepsilon_0}$. Hence $\phi^{-t_1}(\sigma)$ crosses U_0 for some $t_1 > 0$, which gives the point z_1 we need.

Hence $\phi^{t_1}(U_0)$ contains a closed ball U_1 around the point $z_1 \in \Gamma^+_{\varepsilon_1}$. By induction, we get a sequence $t_n > 0$, and a sequence of closed balls U_n with centers $z_n \in \Gamma^+_{\varepsilon_n}$ such that $\phi^{t_n}(U_{n-1})$ contains the ball U_n.

The sequence of compact sets $V_n = \phi^{-\tau_n}(U_n) \subset U_0$, $\tau_n = t_1 + \dots + t_n$, is nested: $V_1 \supset V_2 \supset \dots \supset V_n \supset \dots$, and hence $V_\infty = \cap_n V_n \neq \emptyset$. We can take for

w any point in V_∞. The condition dist $(\phi^{\tau_n}(w), w_n) \to 0$ can be satisfied by the choice of the curves $\sigma = \sigma_n$. ∎

The solution $(q(t), p(t))$ with the initial point w_0 satisfies the assertion of Theorem 2.1. Such a construction gives a very exaggerated estimate for the time τ_n. Better estimates can be obtained by using the methods of [4, 8, 21].

A Proof of Lemma 3.1

The proof is based on several auxiliary propositions.

Proposition A.1 *Let $\mathcal{N}$ be a nonresonant hyperbolic torus. Then:*

- *The torus $\mathcal{N}$ is isotropic i.e., the restriction of the symplectic structure to the bundle $T\mathcal{N}$ vanishes.*
- *For any $w \in \mathcal{N}$ the subspaces $T_w\mathcal{N} \oplus E_w^\pm$ are Lagrangian: for any vectors $v, v' \in T_w\mathcal{N} \oplus E_w^\pm$, $w \in \mathcal{N}$, we have $\omega(v, v') = 0$.*
- *The form ω defines a nondegenerate bilinear form on $E_w^+ \times E_w^-$: if $v_+ \in E_w^+$ and $\omega(v_+, v_-) = 0$ for all $v_- \in E_w^-$, then $v_+ = 0$.*

Proof. The first statement follows from the Herman theorem [17]. To prove the second one, take any two vectors $v, v' \in T_w\mathcal{N} \oplus E_w^+$. Then

$$\omega(v, v') = \omega(Dg_t v, Dg_t v') \to 0, \qquad t \to \infty.$$

Indeed, for any vector $v \in T_w\mathcal{N} \oplus E_w^+$,

$$\|Dg_t v - Dg_t u\| \to 0, \quad \|Dg_t u\| \leq C,$$

where $u \in T_w\mathcal{N}$ is the component of v tangent to the torus. The last statement follows from the first two and the nondegeneracy of ω. ∎

Recall that passing to a finite covering of a tubular neighborhood of the torus $\mathcal{N}$, we can assume that the stable and unstable bundles $E^\pm$ are trivial.

Proposition A.2 *In a neighborhood of the torus $\mathcal{N}$ on $\mathcal{M}$ there exist coordinates $(x, y, z_-, z_+) \in \mathbf{T}^n \times \mathbf{R}^n \times \mathbf{R}^{m-n} \times \mathbf{R}^{m-n*}$ such that:*

- $\Phi(x) = (x, 0, 0, 0)$ *for any* $x \in \mathbf{T}^n$,
- $\omega = dy \wedge dx + dz_+ \wedge dz_-$,
- $E^+ = \mathbf{T}^n \times \mathbf{R}^{m-n*}$, $E^- = \mathbf{T}^n \times \mathbf{R}^{m-n}$.

Proof. By definition, $\mathcal{N} = \Phi(\mathbf{T}^n)$. Let x be the standard coordinate in $\mathbf{T}^n$. Since the bundle E^- is trivial, for any $w = \Phi(x)$ we have an isomorphism $E_w^- = \mathbf{R}^{m-n}\{z_-\}$. By Proposition A.1, $V = E_w^- \oplus E_w^+$ is a symplectic space. Hence $E_w^+ = E_w^{-*}$ and $V = \mathbf{R}^{m-n}\{z_-\} \oplus \mathbf{R}^{m-n^*}\{z_+\}$ with the standard symplectic structure $dz_+ \wedge dz_-$. Let $V^\perp = \{v \in T_w\mathcal{M} \mid \omega(v,u) = 0 \ \forall u \in V\}$ be the symplectic complement of V. Then V is a symplectic space of dimension $2n$, and it contains the Lagrangian subspace $T_w\mathcal{N}$. Hence there exists a symplectic isomorphism $V^\perp = T_w\mathcal{N} \oplus T_w^*\mathcal{N} = \mathbf{R}^n\{x\} \oplus \mathbf{R}^{n*}\{y\}$. Thus, for any $w = \Phi(x) \in \mathcal{N}$, we have a symplectic isomorphism

$$T_w\mathcal{M} = V^\perp \oplus V = \mathbf{R}^n\{x\} \oplus \mathbf{R}^{n*}\{y\} \oplus \mathbf{R}^{m-n}\{z_-\} \oplus \mathbf{R}^{m-n^*}\{z_+\},$$

and the bilinear form $\omega|_{T_w\mathcal{M}}$ takes the form $dy \wedge dx + dz_+ \wedge dz_-$.

Now using, for example, the exponential map $T_w\mathcal{M} \to \mathcal{M}$, we obtain a map $f : \mathbf{T}^n\{x\} \times \mathbf{R}^{n*}\{y\} \times \mathbf{R}^{m-n}\{z_-\} \times \mathbf{R}^{m-n^*}\{z_+\} \to \mathcal{M}$ defining the coordinates $x, y, z_\pm$ in a tubular neighborhood of the torus $\mathcal{N}$ such that

$$f^*\omega = dy \wedge dx + dz_+ \wedge dz_- + O(y,z).$$

Put

$$\omega^0 = dy \wedge dx + dz_+ \wedge dz_-.$$

The symplectic structures $f^*\omega$ and ω^0 coincide on the torus $\mathbf{T}^n$. Hence, they are homotopic in a neighborhood of $\mathbf{T}^n$ in the class of symplectic structures. For example, the homotopy can be chosen as $af^*\omega + (1-a)\omega^0$, $0 \le a \le 1$. According to the generalized Darboux theorem [3], there exists a diffeomorphism f_0 in a neighborhood of the torus, such that

- f_0 and Df_0 equal the identity on $\mathbf{T}^n$;
- $f_0^*(f^*\omega) = \omega^0$.

The map $f \circ f_0$ defines the coordinates we need. Proposition A.2 is proved.

Now suppose that $n = m - 1$.

Proposition A.3 *Suppose that the vector ν is Diophantine, the function $\psi : \mathbf{T}^{m-1} \to \mathbf{R}$ is real-analytic and λ is a real constant. In the case $\lambda = 0$ suppose also that the average of ψ over the torus $\mathbf{T}^{m-1}$ vanishes. Then there exists a unique function $\varphi : \mathbf{T}^{m-1} \to \mathbf{R}$ (in the case $\lambda = 0$ the average of φ is assumed to vanish) satisfying the equation*

$$\partial_\nu \varphi(x) + \lambda\varphi(x) = \psi(x), \qquad \partial_\nu = \langle \nu, \partial/\partial x \rangle. \tag{A.1}$$

Moreover, φ is real-analytic.

The proof can be obtained by using the Fourier expansion of the equation (A.1). Note that we need the Diophantine condition only for $\lambda = 0$.

Proposition A.4 *The coordinates (x, y, z_-, z_+) from Proposition A.2 can be transformed to canonical ones for the torus $\mathcal{N}$.*

Proof. Proposition A.2 implies that the restriction of the system to the torus $\mathcal{N}$ takes the form

$$\dot{x} = \nu = \mathcal{H}_u|_N, \quad \dot{y} = 0 = -\mathcal{H}_x|_{\mathcal{N}}, \quad \dot{z}_\pm = 0 = \mp\mathcal{H}_{z_\mp}|_{\mathcal{N}}.$$

Hence,

$$\begin{aligned}\mathcal{H} &= \langle \nu, y\rangle + O_2(y, z_\pm)\\ &= \langle \nu, y\rangle + \langle A(x)y, y\rangle/2 + \langle b_+(x), y\rangle z_+ + \langle b_-(x), y\rangle z_-\\ &\qquad + c_-(x)z_-^2/2 + c(x)z_-z_+ + c_+(x)z_+^2/2 + O_2(y, z_\pm),\end{aligned}$$

where the $(m-1)\times(m-1)$-matrix A, $(m-1)$-dimensional vectors $b_\pm$, and $c_\pm, c$ are real-analytic functions of $x \in \mathbf{T}^{m-1}$.

The corresponding differential equations are as follows

$$\left\{\begin{array}{lcll}\dot{x} &=& \nu + Ay + b_+z_+ + b_-z_- & +O_2(y,z),\\ \dot{y} &=& & O_2(y,z),\\ \dot{z}_- &=& cz_- + c_+z_+ + \langle b_+, y\rangle & +O_2(y,z),\\ \dot{z}_+ &=& -c_-z_- - cz_+ - \langle b_-, y\rangle & +O_2(y,z).\end{array}\right.$$

By Proposition A.2, the manifolds $\{y = 0, z_- = 0\}$ and $\{y = 0, z_+ = 0\}$ are invariant in the linear approximation. This means that $c_- \equiv c_+ \equiv 0$. According to the definition of a hyperbolic torus, the average $\overline{c}$ over the torus $\mathbf{T}^{m-1}$ of the function c is positive.

Let $a = e^\varphi$, where $\varphi = \varphi(x)$ is a solution of the equation

$$\partial_\nu\varphi(x) + c(x) = \overline{c}.$$

We assume that the average of φ vanishes. According to Proposition A.3, φ is real-analytic. Consider the canonical change $(x, y, z_-, z_+) \mapsto (x^*, y^*, z_-^*, z_+^*)$ determined by the generating function

$$S(x, z_-, y^*, z_+^*) = \langle x, y^*\rangle + a(x)z_-z_+^*.$$

Direct calculations show that in the new variables

$$\begin{aligned}\mathcal{H} &= \langle \nu, y^*\rangle + \langle A^*(x^*)y^*, y^*\rangle/2 + \langle b_+^*(x^*), y^*\rangle z_+^* + \langle b_-^*(x^*), y^*\rangle z_-^*\\ &\qquad + \overline{c}z_-^*z_+^* + O_2(y^*, z_\pm^*),\end{aligned}$$

where A^* and $b_\pm^*$ are some new real-analytic functions.

To cancel the terms containing $b_\pm^*$, we perform the change $(x^*, y^*, z_-^*, z_+^*) \mapsto (\hat{x}, \hat{y}, \hat{z}_-, \hat{z}_+)$ with the generating function

$$S^*(x^*, z_-^*, \hat{y}, \hat{z}_+) = \langle x^*, \hat{y}\rangle + z_-^*\hat{z}_+ + \langle g_-(x^*), \hat{y}\rangle z_-^* + \langle g_+(x^*), \hat{y}\rangle\hat{z}_+,$$

where the vector-functions $g_\pm$ are solutions of the equations

$$\partial_\nu g_\pm(x) \mp \overline{c} g_\pm(x) + b^*_\pm(x) = 0.$$

In the new variables

$$\mathcal{H} = \langle \nu, \hat{y} \rangle + \langle \hat{A}(\hat{x})\hat{y}, \hat{y} \rangle/2 + \overline{c}\hat{z}_-\hat{z}_+ + O_2(\hat{y}, \hat{z}_\pm)$$

with some matrix-valued function $\hat{A}$.

The change $(\hat{x}, \hat{y}, \hat{z}_-, \hat{z}_+) \mapsto (x, y, z_-, z_+)$ with the generating function

$$\hat{S}(\hat{x}, \hat{z}, y, z_+) = \langle \hat{x}, y \rangle + \hat{z}_- z_+ + \langle B(\hat{x})y, y \rangle,$$

where the matrix B is a solution of the equation

$$\partial_\nu B(x) + \hat{A}(x) = \overline{A}, \qquad \overline{A} = \frac{1}{(2\pi)^{m-1}} \int_{\ {}^{m-1}} \hat{A}(x)\, dx,$$

defines coordinates canonical for $\mathcal{N}$. Indeed,

$$\mathcal{H} = \langle \nu, y \rangle + \langle \overline{A}y, y \rangle/2 + \overline{c} z_- z_+ + O_2(y, z_\pm).$$

Proposition A.4 and Lemma 3.1 are proved.

B Proof of Proposition 4.1

The set $[\varepsilon/2, \varepsilon] \setminus \mathcal{E}$ is the infinite union $\cup_{|k|+|l|\neq 0} I_{k,l}$,

$$I_{k,l} = \{\sigma \in [\varepsilon/2, \varepsilon] : |\langle k, \nu \rangle + l\sigma^\kappa| < \sigma^r(|k| + |l|)^{-\beta_*}\}.$$

Note that $I_{k,l} \subset J_{k,l}$,

$$J_{k,l} = \{\sigma \in [\varepsilon/2, \varepsilon] : |\langle k, \nu \rangle + l\sigma^\kappa| < \varepsilon^r(|k| + |l|)^{-\beta_*}\}.$$

The sets $J_{k,0}$ are empty since ν is Diophantine, $\beta_* \geq \beta$ and it is possible to assume that $\varepsilon^r < \alpha$.

Consider the case $l \neq 0$. The set $J_{k,l}$ is the image of the interval (may be empty)

$$\Big[-\frac{\langle k, \nu \rangle}{l} - \frac{\varepsilon^r}{|l|(|k| + |l|)^{\beta_*}}, -\frac{\langle k, \nu \rangle}{l} + \frac{\varepsilon^r}{|l|(|k| + |l|)^{\beta_*}} \Big] \cap \Big[\Big(\frac{\varepsilon}{2}\Big)^\kappa, \varepsilon^\kappa \Big]$$

with respect to the map $x \mapsto x^{1/\kappa}$. The Lipschitz constant of this map on $[(\varepsilon/2)^\kappa, \varepsilon^\kappa]$ is $\kappa^{-1}(\varepsilon/2)^{1-\kappa}$. Hence, the measure of $J_{k,l}$ does not exceed

$$\frac{1}{\kappa}\Big(\frac{\varepsilon}{2}\Big)^{1-\kappa} \frac{2\varepsilon^r}{|l|(|k| + |l|)^{\beta_*}}.$$

Now Proposition 4.1 follows from the inequality

$$\sum_{l \neq 0, k \in \ {}^{m-1}} \frac{1}{|l|(|k| + |l|)^{\beta_*}} < \infty \quad \text{for} \quad \beta_* > m - 1.$$

C Proof of Proposition 4.2

The proof is a combination of standard methods of perturbation theory and KAM-theory [2]. For $(\varepsilon, \mu) \in \mathcal{P}$, the transformation F from the new to the old variables has the form

$$F = f_{\text{hyp}} \circ f_{(K)} \circ f_{\text{norm}}. \tag{C.1}$$

The maps f_{hyp} and $f_{(K)}$ are analytic in all the arguments (including μ). The map f_{norm} is also analytic in the phase variables and ε, but it is defined only for $(\mu, \varepsilon) \in \mathcal{P}$. However, it is C^∞ Whitney smooth in μ.

The construction of the map (C.1) is as follows. The map f_{hyp} introduces canonical coordinates $(x_{(0)}, y_{(0)}, z_{(0)\pm})$ for the hyperbolic torus $\mathcal{N}$ of the system with Hamiltonian H_0. Such a map exists according to Lemma 3.1. The Hamiltonian (4.1) takes the form

$$\begin{aligned} H(x_{(0)}, y_{(0)}, z_{(0)}, t, \mu, \varepsilon) = \langle \nu, y_{(0)} \rangle + \langle A_0 y_{(0)}, y_{(0)} \rangle / 2 + \lambda_0 z_{(0)-} z_{(0)+} \\ + O_3(y_{(0)}, z_{(0)\pm}) + O(\varepsilon) \end{aligned}$$

with some positive constant λ_0 and nondegenerate matrix A_0.

Proposition C.1 *For any $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$ and any natural k there exists a 2π-periodic in $t = \mu\tau$ canonical change*

$$f_{(k)} : (x_{(k)}, y_{(k)}, z_{(k)\pm}, t) \mapsto (x_{(0)}, y_{(0)}, z_{(0)\pm}, t)$$

such that in the new coordinates the Hamiltonian is as follows:

$$\begin{aligned} H_k(x_{(k)}, y_{(k)}, z_{(k)}, t, \mu, \varepsilon) \quad = \quad & \langle \nu, y_{(k)} \rangle + \langle A_k y_{(k)}, y_{(k)} \rangle / 2 + \lambda_k z_{(k)-} z_{(k)+} \\ & + O_3(y_{(k)}, z_{(k)\pm}) + O(\varepsilon \mu^k), \end{aligned} \tag{C.2}$$

where the functions $A_k = A_k(\mu, \varepsilon)$, $\lambda_k = \lambda_k(\mu, \varepsilon)$ are real-analytic at zero and $A_k(0,0) = A_0$, $\lambda_k(0,0) = \lambda_0$. The map $f_{(k)}$ is analytic in all the arguments, ε and μ.

Proof. We construct the sequence of changes

$$(x_{(0)}, y_{(0)}, z_{(0)\pm}, t) \mapsto (x_{(1)}, y_{(1)}, z_{(1)\pm}, t) \mapsto \ldots \mapsto (x_{(k)}, y_{(k)}, z_{(k)\pm}, t).$$

The map $(x_{(j)}, y_{(j)}, z_{(j)\pm}) \mapsto (x_{(j+1)}, y_{(j+1)}, z_{(j+1)\pm})$ can be taken as follows. Let us fix the parameter t in the Hamiltonian $H_j(x_{(j)}, y_{(j)}, z_{(j)\pm}, t, \mu, \varepsilon)$. We obtain an autonomous system, depending on the parameter t, which is an $O(\varepsilon\mu^j)$-perturbation of the system with Hamiltonian

$$\langle \nu, y_{(j)} \rangle + \langle A_j y_{(j)}, y_{(j)} \rangle / 2 + \lambda_j z_{(j)-} z_{(j)+} + O_3(y_{(j)}, z_{(j)\pm}).$$

According to Theorem 3.1, there exists a change, real-analytic and $O(\varepsilon\mu^j)$-close to the identity, introducing in the system with Hamiltonian H_j canonical coordinates $x_{(j+1)}, y_{(j+1)}, z_{(j+1)\pm}$ for the hyperbolic torus with frequencies ν. The corresponding generating function S_j depends analytically and 2π-periodically on t.

Now consider the nonautonomous change with the generating function S_j, in which we put $t = \mu\tau$. Since $\partial S_j/\partial\tau = O(\varepsilon\mu^{j+1})$, the transformed Hamiltonian will be changed by $O(\varepsilon\mu^{j+1})$. Hence the new variables can be regarded as $(x_{(j+1)}, y_{(j+1)}, z_{(j+1)\pm})$. Proposition C.1 is proved.

For what follows it is sufficient to take $k = 6r$.

Lemma C.1 (KAM-step) *There exists a real-analytic in ε, in the phase variables (including t), and C^∞ Whitney smooth in μ canonical map*

$$f_{\rm norm} : (x, y, z_\pm, t) \mapsto (x_{(6r)}, y_{(6r)}, z_{(6r)\pm}, t), \qquad (\mu, \varepsilon) \in \mathcal{P},$$

that transforms the Hamiltonian (C.2)$|_{k=6r}$ to the normal form (4.7).

Proposition 4.2 follows from Lemma C.1 and Proposition C.1.

Sketch of the proof of Lemma C.1. The map $f_{\rm norm}$ is obtained by means of a cumbersome but standard procedure [6, 14]. First, we transform the Hamiltonian (C.2)$|_{k=6r}$ to the Graff normal form [6, 16, 14], and then pass to the Eliasson normal form (4.7), see [14].

Smoothness in μ is more delicate. However, the corresponding argument is also standard. A similar result can be found for example, in [31]. The main difference of the assertion we need from the one presented in [31] is that in our case the frequency vector (ν, μ) of the hyperbolic torus has a small component μ. In particular, the corresponding Diophantine conditions should contain the small parameter μ. However, this circumstance is compensated by the smallness condition for the perturbation (it is of order $\varepsilon\mu^{6r}$).

Here we only present a rapidly convergent Newton-type procedure which can serve as a basis for the proof of Lemma C.1, without proving the convergence. First, we transform the Hamiltonian (C.2)$|_{k=6r}$ to the Graff normal form

$$H'(x, y, z, t, \mu, \varepsilon) = \langle \nu, y\rangle + \lambda z_- z_+ + \frac{1}{2}\langle Ay, y\rangle + O_3(y, z). \qquad \text{(C.3)}$$

To this end we perform an infinite sequence of nonautonomous canonical coordinate changes ϕ_j and define the transformation of variables as $\phi_0 \circ \phi_1 \circ \phi_2 \circ \cdots$. Let us determine the $(j+1)$-st change $\phi_j : (X, Y, Z, t) \mapsto (x, y, z, t)$. For brevity we do not write ε and μ among the arguments of the functions.

Let the Hamiltonian at the j-th step be

$$H = \langle \nu, y\rangle + \lambda z_- z_+ + \frac{1}{2}\langle Ay, y\rangle + h + \varepsilon\mu^{6r} g, \qquad h(x, y, z, t) = O_3(y, z).$$

Our goal is to cancel g. The change $\phi_j : (X, Y, Z, t) \mapsto (x, y, z, t)$ is determined by the generating function $S = S(x, Y, z_-, Z_+, t)$ as follows:

$$x = X - S_Y,\ y = Y + c + S_x,\ z_- = Z_- - S_{Z_+},\ z_+ = Z_+ + S_{z_-}. \qquad \text{(C.4)}$$

where c is a constant vector. Let

$$S = a + \langle b, Y\rangle + b_- z_- + b_+ Z_+ + \tfrac{1}{2}\langle BY, Y\rangle$$
$$+\langle c_-, Y\rangle z_- + \langle c_+, Y\rangle Z_+ + \tfrac{1}{2}\left\langle C\begin{pmatrix} z_- \\ Z_+ \end{pmatrix}, \begin{pmatrix} z_- \\ Z_+ \end{pmatrix}\right\rangle,$$

where the scalar functions $a(x,t)$, $b_\pm(x,t)$, the vector functions $b(x,t)$, $c_\pm(x,t)$, and the symmetric matrix functions $B(x,t)$, $C(x,t)$ on $\mathbf{T}^{m+1}$ satisfy the equations

$$\begin{array}{rcl}
\partial a + \varepsilon\mu^{6r} g & = & \alpha, \\
\partial b + Ac + Aa_x + \varepsilon\mu^{6r} g_y & = & 0, \\
\partial b_\pm \mp \lambda b_\pm + \varepsilon\mu^{6r} g_{z_\pm} & = & 0, \\
\partial B + Ab_x + (Ab_x)^T + \varepsilon\mu^{6r} g_{yy} + h_{yy} & = & \beta, \\
\partial c_\pm \mp \lambda c_\pm + Ab_{\pm x} + \varepsilon\mu^{6r} g_{yz_\pm} + h_{yz_\pm} & = & 0, \\
\partial C + \Lambda C + C\Lambda + \varepsilon\mu^{6r} g_{zz} + h_{zz} & = & \operatorname{diag}(\gamma, -\gamma).
\end{array} \tag{C.5}$$

The differential operator ∂ is as in (5.2), $\Lambda = \operatorname{diag}(\lambda, -\lambda)$, the quantities α, β, γ depend only on ε and μ. The function g and its derivatives are calculated in (C.5) at the point $(x, y = 0, z = 0, t)$, and the derivatives of h in (C.5) are taken at $(x, y = c + a_x, z_- = -b_+, z_+ = b_-, t)$. The homological equations (C.5) are linear, and can be solved for $\mu \in \mathcal{E}^\kappa$ by using Fourier series.

The corresponding generating function S is "of order" $O(\varepsilon\mu^{3r} g)$. The new Hamiltonian equals $\widehat{H} = \mu S_t + H$:

$$\widehat{H}(X, Y, Z, t) = \langle \nu, Y\rangle + \hat{\lambda} Z_- Z_+ + \frac{1}{2}\langle \hat{A} Y, Y\rangle + \hat{h} + \varepsilon\mu^{6r}\hat{g}, \qquad \hat{h} = O_3(Y, Z).$$

The function $\hat{g}$ turns out to be at least quadratic with respect to g and S. In fact $\hat{g}$ is "of order" $\hat{g} = O(\varepsilon g^2)$. Hence $\hat{g} \ll g$. This implies rapid convergence of the sequence ϕ_j.

Now let us describe the passage from the Graff normal form (C.3) to the following one:

$$H''(x, y, z, t) = \langle \nu, y\rangle + \lambda z_- z_+ + O(y, z_- z_+) O(y, z). \tag{C.6}$$

We again construct an infinite sequence of nonautonomous canonical coordinate changes f_k and define the transformation as $f_0 \circ f_1 \circ f_2 \circ \cdots$. Let us determine the $(k+1)$-st change $f_k : (X, Y, Z, t) \mapsto (x, y, z, t)$.

Let the Hamiltonian at the k-th step be

$$H = \langle \nu, y\rangle + \lambda z_- z_+ + \varphi_-(x, z_-, t) + \varphi_+(x, z_+, t)$$
$$+\langle y, \psi(x, y, z_-, z_+, t)\rangle + z_- z_+ \theta(x, y, z_-, z_+, t),$$

where $\varphi_\pm = O_3(z_\pm)$, $\psi = O(y, z)$, $\theta = O(y, z)$. Note that at the first step the functions $\varphi_\pm$, ψ, θ can be assumed to be small, because otherwise we perform the change

$$(x, y, z, t) \mapsto (x, c^2 y, cz, t), \quad H' \mapsto c^2 H'$$

with sufficiently small constant $c > 0$.

The change $f_k : (X, Y, Z, t) \mapsto (x, y, z, t)$ is determined by the generating function

$$S = S(x, Y, z_-, Z_+, t) = a_-(x, z_-, t) + a_+(x, Z_+, t)$$

as in (C.4) with $c = 0$, where the following equalities hold:

$$\begin{aligned} \partial a_- + \langle \psi_-, a_{-x} \rangle + z_-(\lambda + \theta_-) a_{-z_-} + \varphi_- = 0, \\ \partial a_+ + \langle \psi_+, a_{+x} \rangle - z_+(\lambda + \theta_+) a_{+z_+} + \varphi_+ = 0. \end{aligned} \tag{C.7}$$

Here we denoted

$$\begin{aligned} \psi_- = \psi(x, 0, z_-, 0, t), \quad \psi_+ = \psi(x, 0, 0, z_+, t), \\ \theta_- = \theta(x, 0, z_-, 0, t), \quad \theta_+ = \theta(x, 0, 0, z_+, t), \end{aligned}$$

and substituted z_+ instead of Z_+ in a_+. The solutions $a_\pm$ of the homological equations (C.7) have the form

$$\begin{aligned} a_-(x, z_-, t) = -\int_{-\infty}^0 \varphi_- \circ g_-^s(x, z_-, t)\, ds, \\ a_+(x, z_+, t) = \int_0^\infty \varphi_+ \circ g_+^s(x, z_+, t)\, ds, \end{aligned} \tag{C.8}$$

where $g_\pm^s$ are the phase flows of the systems

$$\begin{cases} \frac{dx}{ds} = \nu + \psi_-, \\ \frac{dz_-}{ds} = z_-(\lambda + \theta_-), \\ \frac{dt}{ds} = \mu, \end{cases} \quad \text{and} \quad \begin{cases} \frac{dx}{ds} = \nu + \psi_+, \\ \frac{dz_+}{ds} = -z_+(\lambda + \theta_+), \\ \frac{dt}{ds} = \mu. \end{cases}$$

The new Hamiltonian equals $\widehat{H} = \mu S_t + H$:

$$\begin{aligned} \widehat{H} = \langle \nu, Y \rangle + \lambda Z_- Z_+ + \hat{\varphi}_-(X, Z_-, t) + \hat{\varphi}_+(X, Z_+, t) \\ + \langle Y, \hat{\psi}(X, Y, Z_-, Z_+, t) \rangle + Z_- Z_+ \hat{\theta}(X, Y, Z_-, Z_+, t), \end{aligned}$$

where again $\hat{\varphi}_\pm = O_3(Z_\pm)$, $\hat{\psi} = O(Y, Z)$, $\hat{\theta} = O(Y, Z)$. Moreover, the functions $\hat{\varphi}_\pm$ are quadratic in $\varphi_\pm$ and $a_\pm$. Due to (C.8), $a_\pm$ and $\varphi_\pm$ are of the same order. Hence, $\hat{\varphi}_\pm$ are quadratic in $\varphi_\pm$, and the sequence f_k rapidly converges.

Analogously we pass from the normal form (C.6) to the Eliasson normal form (4.7). The corresponding generating function is

$$S = \langle Y, b_-(x, z_-, t) + b_+(x, Z_+, t) \rangle + z_- Z_+ (c_-(x, z_-, t) + c_+(x, Z_+, t)).$$

D The Poincaré–Melnikov formula

In this Appendix we prove Theorems 5.1–5.2. Since the manifolds $\Gamma^\pm$ are Lagrangian, by the Weinstein theorem [35], there exist neighborhoods $V_\pm$ of $\Gamma^\pm$ in

$$T^*\Gamma^\pm = \{(q, p) : q \in \Gamma^\pm, p \in T_q^*\Gamma^\pm\}$$

with the standard symplectic structure $dp \wedge dq$ and analytic symplectic embeddings $\phi_\pm : V_\pm \to U_\pm \subset M$ onto open sets such that $\phi_\pm|_{\Gamma^\pm} = \mathrm{id}_{\Gamma^\pm}$. Here, $\Gamma^\pm$ are identified with the zero sections of $T^*\Gamma^\pm$:

$$\Gamma^\pm = \{(q,p) \in T^*\Gamma^\pm : p = 0\}.$$

From now on we identify the neighborhoods $U_\pm$ and $V_\pm$ of $\Gamma^\pm$. The choice of the diffeomorphisms $\phi_\pm$ is not canonical, and the construction below will depend on them. Further on it will be convenient to assume that $U_+ \cap U_-$ consists of two connected components U and V, where U is a tubular neighborhood of $\mathcal{N}$, and V a contractible neighborhood of the point $\gamma(0)$.

Since we have fixed $\mu \in \mathcal{E}^\kappa$, we will drop it from the notation and write $N_\varepsilon = N_{\mu,\varepsilon}$ and $W_\varepsilon^\pm = W_{\mu,\varepsilon}^\pm$.

Recall the definition of the action for the invariant torus N_ε. Since $\mathcal{M} = T^*M$, the symplectic structure ω is exact: $\omega = d\lambda$. Consider the Poincaré–Cartan 1-form

$$\Lambda = \lambda - \mu^{-1}\hat{H}\, dt$$

on $\mathcal{M} \times \mathbf{T}$. Since the torus $N_\varepsilon \subset \mathcal{M} \times \mathbf{T}$ is isotropic [17], the closed 1-form $\Lambda|_{N_\varepsilon}$ defines a cohomology class $[\Lambda|_{N_\varepsilon}] \in H^1(N_\varepsilon, \mathbf{R})$. The torus N_ε is the image of a map $\Phi_\varepsilon : \mathbf{T}^{n+1} \to \mathcal{M} \times \mathbf{T}$, $\Phi_\varepsilon(x,t) = (\Psi_{\varepsilon,t}(x), t)$. The action J_ε of the torus N_ε is the pull back of the class $[\Lambda|_{N_\varepsilon}]$:

$$J_\varepsilon = [\Phi_\varepsilon^* \Lambda] \in H^1(\mathbf{T}^{n+1}, \mathbf{R}) = \mathbf{R}^{n+1}.$$

Then $J_\varepsilon = (b_\varepsilon, \mu^{-1}c_\varepsilon) \in \mathbf{R}^{n+1}$ with $b_\varepsilon \in \mathbf{R}^n$ the cohomology class of the form $\Psi_{\varepsilon,t}^*\lambda$. Replacing $\hat{H}$ by $\hat{H} - c_\varepsilon$, we can always put $c_\varepsilon = 0$. Let us show that we can assume that b_ε does not depend on ε.

Lemma D.1 *There exists a nonexact canonical transformation (symplectic embedding)*

$$f_\varepsilon : U_+ \cup U_- \to \mathcal{M}, \qquad f_0 = \mathrm{id}, \quad 0 \le \varepsilon \le \varepsilon_0,$$

such that the action on the torus $(f_\varepsilon \times \mathrm{id}\ )^{-1} N_\varepsilon$ *is independent of* ε.

Note that in the new variables, the action is defined by means of the same form λ (not by means of the form $f_\varepsilon^*\lambda$).

Proof. We construct a map f_ε such that $[\Psi_{\varepsilon,t}^* f_\varepsilon^* \lambda] = b_0 \in \mathbf{R}^n$. There exists a 1-form β_ε on $U_+ \cup U_-$ such that $d\beta_\varepsilon = \omega$ and $[\Psi_{\varepsilon,t}^* \beta_\varepsilon] = b_\varepsilon$ for all $t \in \mathbf{T}$. Indeed, such forms $\beta_\varepsilon^\pm$ can be defined on $U_\pm$, since $U_\pm$ is contractible on $\mathcal{N}$. We can assume that they coincide on U. Since the neighborhood V of the point $\gamma(0)$ is contractible, $\beta_\varepsilon^+ - \beta_\varepsilon^- = d\phi_\varepsilon$ on V. Now extend the function ϕ_ε to $U_+ \cup U_-$ so that $\phi_\varepsilon|_U = 0$ and put $\beta_\varepsilon = \beta_\varepsilon^-$ on U_- and $\beta_\varepsilon = \beta_\varepsilon^+ - d\phi_\varepsilon$ on U_+.

We look for the map f_ε such that the 1-form $f_\varepsilon^*\beta_\varepsilon - \beta_0$ is exact. Define a locally Hamiltonian vector field ξ_ε on $U_+ \cup U_-$ by the formula $i_{\xi_\varepsilon}\omega = -D_\varepsilon \beta_\varepsilon$. Then the corresponding phase flow f_ε satisfies the condition, because

$D_\varepsilon(f_\varepsilon^*\beta_\varepsilon) = d(f_\varepsilon^*\beta_\varepsilon(\xi_\varepsilon))$. Maybe here the neighborhoods $U_\pm$ and $\varepsilon_0 > 0$ should be decreased. ■

Since we are only interested in a neighborhood $U_+ \cup U_-$ of $\Gamma^+ \cup \Gamma^-$, from now on we assume that the action of the torus N_ε is independent of ε. In the new variables, the Hamiltonian $\hat{H}$ is replaced by $\hat{H} \circ f_\varepsilon - c_\varepsilon$. Since f_ε is time-independent and $f_0 = \mathrm{id}$, the function H_0 changes by a constant, and H_1 does not change. Hence the conditions of Theorem 2.1 do not change in the new variables.

For sufficiently small ε, the manifolds $W_\varepsilon^\pm$ of the torus N_ε can be represented as graphs:

$$W_\varepsilon^\pm = \{(q,p,t) \in T^*\Gamma^\pm \times \mathbf{T} : p = h^\pm(q,t,\varepsilon)\},$$

where $h^\pm$ are smooth covector fields on $\Gamma^\pm$. For $\varepsilon = 0$, we have $W_0^\pm = \Gamma^\pm \times \mathbf{T}$, and so $h^\pm(q,t,0) \equiv 0$. Since the Hamiltonian H in (4.1) is autonomous up to terms of order ε^s, the manifolds $W_\varepsilon^\pm$ are t-independent up to terms of order $O(\varepsilon^s)$. Hence

$$h^\pm(q,\tau,\varepsilon) = h_0^\pm(q,\varepsilon) + \varepsilon^s h_1^\pm(q,t) + \varepsilon^{s+1} h_2^\pm(q,t,\varepsilon). \tag{D.1}$$

The manifolds $W_\varepsilon^\pm$ are Lagrangian. Thus the Poincaré–Cartan 2-form

$$\omega - \mu^{-1}\, d\hat{H} \wedge dt$$

vanishes on $W_\varepsilon^\pm$. Since $\omega = dp \wedge dq$ on $U_\pm \subset T^*\Gamma^\pm$, the 1-form

$$\sigma_\varepsilon^\pm = \langle h^\pm(q,t,\varepsilon), dq\rangle - \mu^{-1}\hat{H}(q, h^\pm(q,t,\varepsilon), t, \varepsilon)\, dt$$

on $\Gamma^\pm \times \mathbf{T}$ is closed.

Lemma D.2 *The 1-forms $\sigma_\varepsilon^\pm$ on $\Gamma^\pm \times \mathbf{T}$ are exact.*

Proof. Since $\Gamma^\pm$ are retractible on $\mathcal{N}$, the cohomology classes $[\sigma_\varepsilon^\pm] \in H^1(\Gamma^\pm \times \mathbf{T}) = \mathbf{R}^{n+1}$ coincide, up to a constant, with the action J_ε of the torus N_ε. Indeed, the 1-form $\langle p, dq\rangle$ on $U_\pm$ differs from λ by a closed 1-form independent of ε. By Lemma D.1, the classes $[\sigma_\varepsilon^\pm] \in \mathbf{R}^{n+1}$ are independent of ε. Since $\sigma_0^\pm = 0$, the forms $\sigma_\varepsilon^\pm$ are exact. ■

Thus there exist generating functions $S^\pm$ on $\Gamma^\pm \times \mathbf{T}$ such that

$$\sigma_\varepsilon^\pm = dS^\pm(q,t,\varepsilon).$$

Hence

$$W_\varepsilon^\pm = \{(q,p,t) \in T^*\Gamma^\pm \times \mathbf{T} : p = \nabla_q S^\pm(q,t,\varepsilon)\} \tag{D.2}$$

and $S^\pm$ satisfy the Hamilton–Jacobi equation

$$\mu D_t S^\pm + H^\pm(q, \nabla_q S^\pm(q,t,\varepsilon), t, \varepsilon) = 0. \tag{D.3}$$

We write

$$H^{\pm} = \hat{H} \circ \phi_{\pm} = H_0^{\pm}(w) + \varepsilon {H'}_0^{\pm}(w,\varepsilon) + \varepsilon^s H_1^{\pm}(w,t) + \varepsilon^{s+1} {H'}_1^{\pm}(w,t,\varepsilon)$$

instead of $\hat{H}$, because we performed different canonical transformations $\phi_{\pm}$ in the domains $U_{\pm}$. By (D.1),

$$S^{\pm} = S_0^{\pm}(q,\varepsilon) + \varepsilon^s S_1^{\pm}(q,t) + \varepsilon^{s+1} S_2^{\pm}(q,t,\varepsilon).$$

The functions $S_0^{\pm}$ are not uniquely defined because some terms of order $O(\varepsilon^{s+1})$ can be moved to $S_2^{\pm}$. To make the expansion unique, let $S_0^{\pm}$ be the solutions of the Hamilton–Jacobi equation

$$(H_0^{\pm} + \varepsilon {H'}_0^{\pm})(q, \nabla_q S_0^{\pm}(q,\varepsilon), \varepsilon) = 0.$$

The solutions exist and are analytic in ε. Indeed, equations $p = \nabla_q S_0^{\pm}(q,\varepsilon)$ represent the stable and unstable manifolds of the hyperbolic invariant torus with frequency ν of the autonomous Hamiltonian system with Hamiltonian $H_0 + \varepsilon H_0'$.

Putting together the expansions for H and $S^{\pm}$ and collecting terms of order ε^s yields

$$\mu D_t S_1^{\pm}(q,t) + \langle D_p H_0^{\pm}(q,0), \nabla_q S_1^{\pm}(q,t)\rangle + H_1^{\pm}(q,0,t) = 0.$$

Denote by $g_\tau^{\pm}$ the phase flow of the unperturbed Hamiltonian system restricted to $\Gamma^{\pm}$. Then the equation can be rewritten as

$$\dot{S}_1^{\pm} = \frac{d}{d\tau}\bigg|_{\tau=0} S_1^{\pm}(g_\tau^{\pm}(q), t+\mu\tau) = -H_1^{\pm}(q,0,t).$$

Without loss of generality assume that the average $\overline{H}_1(t)$ equals zero. Introducing the function χ_μ on $\mathcal{M} \times \mathbf{T}$ satisfying condition (5.1) on $\mathcal{N} \times \mathbf{T}$ and arbitrary elsewhere, we rewrite this equation as

$$\frac{d}{d\tau}\bigg|_{\tau=0} (S_1^{\pm} + \chi_\mu)(g_\tau^{\pm}(q), t+\mu\tau) = -H_1^{\pm}(q,0,t) + \dot{\chi}_\mu.$$

Since the frequency vector (ν,μ) is nonresonant, $S_1^{\pm} + \chi_\mu$ is constant on $\mathcal{N} \times \mathbf{T}$, and we can assume that this constant is zero. Then

$$S_1^{\pm}(q,t) + \chi_\mu(q,t) = \pm \int_0^{\pm\infty} (H_1^{\pm}(g_\tau^{\pm}(q), t+\mu\tau) - \dot{\chi}_\mu(g_\tau^{\pm}(q), t+\mu\tau))\, d\tau.$$

By (5.1), the integral is exponentially convergent.

If the point $q = \gamma(s)$, $|s| < \delta$, belongs to the intersection curve Δ of the invariant manifolds $\Gamma^{\pm}$, both functions $S_1^{\pm}$ are defined at q and

$$S_1^{+}(q,t) - S_1^{-}(q,t) = \int_{-\infty}^{\infty} (H_1(\gamma(s+\tau), t+\mu\tau) - \dot{\chi}_\mu(\gamma(s+\tau), t+\mu\tau))\, d\tau.$$

Such a formula was obtained by Poincaré [27] in a similar situation. Thus

$$S_1^+(\gamma(0),t) - S_1^-(\gamma(0),t) = I_\mu(t), \tag{D.4}$$

where $I_\mu(t)$ is the Poincaré function (5.3).

Now let us study the intersection of the invariant manifolds W_ε^+ and W_ε^- near the point $(\gamma(0),t_0)$. Then we can use both variables $(p_-,q_-) \in T^*\Gamma^-$ and $(p_+,q_+) \in T^*\Gamma^+$. However, if the subspaces $T_w\Gamma^+$ and $T_w^*\Gamma^-$ are not transversal, these coordinates are not convenient. Thus we introduce new symplectic variables x,y in a neighborhood of $\gamma(0)$.

Lemma D.3 *There exist canonical coordinates $x_1,\dots,x_m$, $y_1,\dots,y_m$ in a neighborhood V of $\gamma(0)$ such that:*

- $\gamma(0) = (0,0)$*;*
- *the manifold $\{x=0\}$ is transversal to both manifolds $\Gamma^\pm$ at $\gamma(0)$;*
- *the curve Δ is given by the equations $\{y_1 = \cdots = y_m = 0, x_1 = \cdots = x_{m-1} = 0\}$.*

We omit the proof.

Define the crossection $\Sigma \subset V$ of the curve $\Delta = \gamma(-\delta,\delta)$ by the equation $\Sigma = \{x_m = 0\}$. We will use on $\Sigma \times \mathbf{T} \subset \mathcal{M} \times \mathbf{T}$ the coordinates $x = (x_1,\dots,x_{m-1})$, $y = (y_1,\dots,y_{m-1})$, t, $\hat{H}$. Note that they are independent variables in a neighborhood of $\gamma(0)\times\mathbf{T}$, because v_{H_0} is tangent to Δ.

Proposition D.1 *In the variables $(x,y,t,\hat{H})$, the manifolds $Z_{\mu,\varepsilon}^\pm = W_{\mu,\varepsilon}^\pm \cap (\Sigma \times \mathbf{T})$ are given by the equations*

$$Z_{\mu,\varepsilon}^\pm = \{(x,y,t,\hat{H}) : y = \nabla_x F_{\mu,\varepsilon}^\pm(x,t), \hat{H} = \mu D_t F_{\mu,\varepsilon}^\pm(x,t)\},$$

where

- *the functions $F_{\mu,\varepsilon}^\pm$ are analytic in all the variables except μ, Whitney smooth in μ and have the form*

$$F_{\mu,\varepsilon}^\pm(x,t) = F_0^\pm(x,\varepsilon) + \varepsilon^s F_1^\pm(x,t,\mu) + O(\varepsilon^{s+1}); \tag{D.5}$$

- *the function $G_{\mu,\varepsilon} = F_{\mu,\varepsilon}^+ - F_{\mu,\varepsilon}^-$ has the form*

$$G_{\mu,\varepsilon}(x,\tau) = G_0(x,\varepsilon) + \varepsilon^s G_1(x,t,\mu) + O(\varepsilon^{s+1}), \tag{D.6}$$

  where $G_0(x,0)$ has a nondegenerate critical point at $x=0$ and

$$G_1(0,t,\mu) = I_\mu(t). \tag{D.7}$$

Thus the Hessian matrix of $G_{\mu,\varepsilon}$ at the point $(0, t_0)$ has the form

$$\begin{pmatrix} A + O(\varepsilon) & O(\varepsilon^s) \\ O(\varepsilon^s) & \varepsilon^s I''_\mu(t_0) + O(\varepsilon^{s+1}) \end{pmatrix}, \qquad \det A \neq 0.$$

Proposition D.1 implies Theorem 5.1. Indeed, intersection points of the manifolds $Z^\pm_{\mu,\varepsilon}$ in a neighborhood of the point $(\gamma(0), t_0) \in \Sigma \times \mathbf{T}$ correspond to critical points of the function $G_{\mu,\varepsilon} = F^+_{\mu,\varepsilon} - F^-_{\mu,\varepsilon}$. If t_0 is a nondegenerate minimum point for I_μ, the existence of a critical point for small $\varepsilon > 0$ follows from the implicit function theorem. The degenerate case is also easy, because for $\varepsilon = 0$ the Hessian degenerates in just one direction, and so the problem is reduced to finding a critical point of a function of one variable which is a perturbation of I_μ. The existence of a topologically transversal intersection point follows from the fact that this function has a local minimum or maximum near t_0. Theorem 5.1 is proved. ∎

Proof of Proposition D.1. Again we drop μ from the notation. Suppose that in the variables x, y the manifolds $W^\pm_\varepsilon$ are locally given by the equations

$$W^\pm_\varepsilon = \{(x, y, t) : y = \nabla_x F^\pm_\varepsilon(x, t)\}, \tag{D.8}$$

where

$$\mu D_t F_\varepsilon(x, t) + \hat{H}(x, \nabla_x F^\pm_\varepsilon(x, t), t, \varepsilon) = 0. \tag{D.9}$$

Then the manifolds $\Gamma^\pm$ are locally given by:

$$\Gamma^\pm = \{(x, y) : y = \nabla F^\pm_0(x)\},$$

where $\nabla F^\pm_0|_\Delta = 0$, because $\Delta \subset \{y = 0\}$. Let us show that

$$D_t F^\pm_\varepsilon(0, t) = D_t S^\pm_\varepsilon(\gamma(0), t) + O(\varepsilon^{s+1}). \tag{D.10}$$

We will prove (D.10) for F^+_ε. Since V is contractible,

$$\langle p_+, dq_+\rangle - \langle y, dx\rangle = d\psi(x, y),$$

where ψ is a function on V. Since $p_+ = 0$ and $y = 0$ on Δ, we have $d\psi|_\Delta = 0$. By (D.8), (D.9), (D.2), and (D.3),

$$(\langle y, dx\rangle - \mu \hat{H}\, dt)|_{W^+_\varepsilon} = dF^+_\varepsilon(x, t),$$
$$(\langle p_+, dq_+\rangle - \mu \hat{H}\, dt)|_{W^+_\varepsilon} = dS^+_\varepsilon(q_+, t).$$

Here the variables x and q_+ corresponding to the same point in W^+_ε are related: $q_+ = f_\varepsilon(x, t)$. Hence, up to a constant,

$$F^+_\varepsilon(x, t) = S^+_\varepsilon(f_\varepsilon(x, t), t) + \psi(x, \nabla_x F^+_\varepsilon). \tag{D.11}$$

Since W_ε^+ is independent of t up to terms of order $O(\varepsilon^s)$,

$$D_t F_\varepsilon^+ = O(\varepsilon^s), \quad D_t f_\varepsilon = O(\varepsilon^s).$$

Since $D_q S_\varepsilon^+ = O(\varepsilon)$, $f_0(0,t) = \gamma(0)$ and $d\psi|_\Delta = 0$, differentiating equation (D.11), we get (D.10).

Let $G_{\mu,\varepsilon} = F_\varepsilon^+ - F_\varepsilon^-$. Equations (D.4) and (D.10) imply that

$$D_t G_{\mu,\varepsilon}(0,t) = \varepsilon^s I'_\mu(t) + O(\varepsilon^{s+1}).$$

The function $G_{\mu,\varepsilon}$ has the form (D.6), and $G_0(x,0) = F_0^+(x) - F_0^-(x)$. Since the manifolds $\Gamma^\pm \cap \Sigma \cap \{H_0 = 0\}$ intersect transversely in $\Sigma \cap \{H_0 = 0\}$ at the point $\gamma(0)$, Proposition D.1 is proved. ■

Proof of Theorem 5.2. Let $0 < \varepsilon' < \varepsilon$, $\mu = \varepsilon^\kappa$, $\mu' = \varepsilon'^\kappa$, and $\Delta\varepsilon = \varepsilon - \varepsilon' < c\varepsilon^{2k_0 - k_1}$. Recall that we performed the homothety h_ε and a symplectic transformation f_ε that are different for different ε, and also added to $\hat{H}$ a constant c_ε. Let $\psi_\varepsilon = f_\varepsilon \circ h_\varepsilon$. Then $W_\varepsilon^\pm = \psi_\varepsilon \Gamma_\varepsilon^\pm$. Hence the intersection of Γ_ε^- and $\Gamma_{\varepsilon'}^+$ corresponds to the intersection of W_ε^- and $\psi_\varepsilon \psi_{\varepsilon'}^{-1} W_{\varepsilon'}^+$. Since $\psi_\varepsilon \psi_{\varepsilon'}^{-1} = \mathrm{id} + O(\Delta\varepsilon/\varepsilon)$, the variables $x, y, t, \hat{H}$ and $x', y', t, \hat{H}'$ corresponding to ε and ε', respectively, are related as follows:

$$x' = x + O(\Delta\varepsilon/\varepsilon), \quad y' = y + O(\Delta\varepsilon/\varepsilon), \quad \hat{H}' = \hat{H} + O(\Delta\varepsilon/\varepsilon).$$

Hence the intersection points of the manifolds Γ_ε^- and $\Gamma_{\varepsilon'}^+$ correspond to the solutions of the equations

$$\begin{gathered} y = \nabla_x F_{\mu,\varepsilon}^-(x,t), \quad y' = y + O(\Delta\varepsilon/\varepsilon) = \nabla_x F_{\mu',\varepsilon'}^+(x + O(\Delta\varepsilon/\varepsilon), t), \\ \mu D_t F_{\mu,\varepsilon}^-(x,t) = \mu' D_t F_{\mu',\varepsilon'}^+(x',t) + O(\Delta\varepsilon/\varepsilon). \end{gathered} \tag{D.12}$$

Hence

$$\nabla_x F_{\mu,\varepsilon}^-(x,t) = \nabla_x F_{\mu',\varepsilon'}^+(x + O(\Delta\varepsilon/\varepsilon), t) + O(\Delta\varepsilon/\varepsilon).$$

By (D.5) and (D.6), this equation takes the form

$$\nabla_x G_0(x,0) = O(\varepsilon, \Delta\varepsilon/\varepsilon).$$

Since 0 is a nondegenerate critical point for G_0, there exists a solution

$$x = \xi(t, \varepsilon, \Delta\varepsilon/\varepsilon) = O(\varepsilon, \Delta\varepsilon/\varepsilon).$$

By (D.6) and (D.12),

$$D_t F_1^+(\xi + O(\Delta\varepsilon/\varepsilon), t, \mu') = D_t F_1^-(\xi, t, \mu) + O(\varepsilon, \Delta\varepsilon/(\mu\varepsilon^{s+1})). \tag{D.13}$$

Since by (D.7),

$$D_t F_1^+(0, t, \mu') - D_t F_1^-(0, t, \mu) = I'(t) + O(\mu, \mu'),$$

and for some natural d,

$$I(t) = a(t-t_0)^{2d} + O_{2d+1}(t-t_0), \qquad a \neq 0,$$

equation (D.13) has a solution $t = t(\varepsilon, \Delta\varepsilon)$ close to t_0 provided that ε and $\Delta\varepsilon/(\mu\varepsilon^{s+1}) = \Delta\varepsilon/\varepsilon^{2k_0-k_1}$ are sufficiently small. The solution is nondegenerate if t_0 is a nondegenerate critical point. In the degenerate case, there can be several solutions, but at least one of them has odd multiplicity, and hence corresponds to a topologically transversal intersection. Theorem 5.2 is proved.

Acknowledgements

This work was supported by RFBR Grant 96-15-96855 and INTAS Grant 93-339 ext. The second author was supported also by RFBR grant 96-15-96855. This paper was completed during the permanence of the first author at the University of Barcelona in February 1998. We would like to thank A. Delshams, E. Fontich, T. Seara and C. Simó for stimulating discussions.

References

[1] Arnold V. I., Instability of dynamical systems with several degrees of freedom, *Sov. Mat. Dokl.* **5**, 581–585 (1964)

[2] Arnold V. I., Kozlov V. V., Neishtadt A. I., *Mathematical aspects of classical and celestial mechanics,* Enciklopedia of Math. Sciences, Vol. 3, Springer-Verlag (1989)

[3] Arnold V. I. and Givental A. B., *Symplectic geometry,* Encyclopedia of Math. Sciences, Vol. 4, Springer-Verlag (1989)

[4] Bessi U., An approach to Arnold's diffusion through the calculus of variations. *Nonlin. Anal. TMA* **20**, 1303–1318 (1996)

[5] Bolotin S. V., Homoclinic orbits to invariant tori of Hamiltonian systems, *Amer. Math. Soc. Transl., Ser. 2,* **168**, 21–90 (1995)

[6] Chierchia L. and Gallavotti G., Drift and diffusion in phase space, *Ann. de l'Institut H. Poincaré (Physique Theorique)* **60**, 1–144 (1994)

[7] Chirikov B. V., A universal instability of many-dimensional oscillation systems. *Phys. Rep.* **52**, No. 5, 263–379 (1979)

[8] Cresson J., A λ-lemma for partially hyperbolic tori and the obstruction property, *Lett. Math. Phys.* **42**, 363–377 (1997)

[9] Delshams A., Gelfreich V., Jorba A., Seara T. M, Exponentially small splitting of separatrices under fast quasiperiodic forcing, *Comm. Math. Phys.*, **189**, 35–71 (1997)

[10] Delshams A. and Seara T. M., An asymptotic expression for the splitting of separatrices of the rapidly forced pendulum. *Comm. Math. Phys.* **150**, 433–463 (1992)

[11] Delshams A., de la Llave R., Seara T. M., A geometric approach to the existence of orbits with unbounded energy in generic periodic perturbations by potential of generic geodesic flows on $\mathbf{T}^2$, *Preprint* (1998)

[12] Douady R., Stabilité ou instabilité des points fixes elliptiques, *Ann. Sci. Ec. Norm. Super.* IV Ser. **21**, 1–46 (1988)

[13] Fontich E. and Martín P., Arnold diffusion in perturbations of analytic integrable Hamiltonian systems, *Preprint* (1998)

[14] Eliasson L. H., Biasymptotic solutions of perturbed integrable Hamiltonian systems, *Bol. Soc. Bras. Mat.* **25**, 57–76 (1994)

[15] Gallavotti G., Twistless KAM tori, quasi flat homoclinic intersections, and other cancellations in the perturbation series of certain completely integrable Hamiltonian systems. A review. *Rev. Math. Phys.* **6**, 342–411 (1994)

[16] Graff S., On the conservation of hyperbolic invariant tori for Hamiltonian systems, *J. Diff. Eq.* **15**, 1–69 (1974)

[17] Herman M., Inégalités "a priori" pour des tores lagrangiens invariants par des difféomorphismes symplectiques, *Inst.-Hautes-Études-Sci.-Publ.-Math.* **70**, 47–101 (1990)

[18] A. Katok, Entropy and closed geodesics, *Ergod. Th. and Dynam. Syst.* **2**, 339–367 (1982)

[19] Kozlov V. V., Integrability and non-integrability in Hamiltonian mechanics, *Russ. Math. Surveys* **38**, 1–76 (1983).

[20] Lochak P., Canonical perturbation theory via simultaneous approximation, *Russ. Math. Surv.* **47**, No. 6, 57–133 (1992)

[21] Marco J. P., Transition le long chaines de tores invariants pour les systèmes hamiltoniens analytiques, *Ann. Inst. H. Poincaré Phys. Theor.* **64**, 205–252 (1996)

[22] Mather J., Variational construction of connecting orbits, *Ann. Inst. Fourier* **43**, 1349–1386 (1993)

[23] Mather J., Talk given at the conference in dynamical systems, Oberwolfach (1997)

[24] Morse M., A fundamental class of geodesics in any closed surface of genus greater than one, *Trans. Amer. Math. Soc.*, **26**, 25–61 (1924)

[25] Nekhoroshev N. N., An exponential estimate of the time of stability in Hamiltonian systems close to integrable, *Russ. Math. Surveys* **32**, No. 6, 1–65 (1977)

[26] Palis J. and de Melo W., *Geometric Theory of Dynamical Systems,* Springer-Verlag (1982)

[27] Poincaré H., *Les métodes nouvelles de la mécanique céleste.* V. 1–3. Paris: Gauthier–Villars, 1892, 1893, 1899.

[28] Pöschel J., Nekhoroshev estimates for quasi-convex Hamiltonian systems. *Math. Z.* **213**, 187–216 (1993)

[29] Pöschel J., Integrability of Hamiltonian systems on Cantor sets. *Comm. Pure Appl. Math.* **35**, 653–695 (1982)

[30] Rudnev M. and Wiggins S., Existence of exponentially small separatrix splitting and homoclinic and heteroclinic connections between whiskered tori in weakly hyperbolic near-integrable Hamiltonian systems, *Physica* D **114**, 3–80 (1998)

[31] Rudnev M. and Wiggins S., KAM theory near multiplicity one resonant surfaces in perturbations of a-priori stable Hamiltonian systems. *Journal of Nonlin. Science*, **7**, 177–209 (1997)

[32] Stein E., *Singular Integrals and Differentiability Properties of Functions,* Princeton (1970)

[33] Treschev D. V., Separatrix splitting for a pendulum with rapidly oscillating suspension point. *Russ. J. Math. Phys.* **5**, No. 1, 63–98 (1997)

[34] Treschev D. V., Hyperbolic tori and asymptotic surfaces in Hamiltonian systems, *Russ. J. Math. Phys.* **2**, No. 1, 93–110 (1994)

[35] Weinstein A., Lagrangian submanifolds and Hamiltonian systems, *Ann. Math.* **98**, 377–410 (1973)

[36] Whitney H., Analytic extensions of differentiable functions defined in closed sets, *Trans. A.M.S.* **36**, 63–89 (1934)

[37] Zehnder E., Generalized implicit function theorem with applications to some small divisor problems II, *Comm. Pure Appl. Math.* **29**, 49–111 (1976)