Lagrangian tori near resonances of near-integrable Hamiltonian systems

Medvedev A. G.,* Neishtadt A. I.,† Treschev D. V.‡

December 24, 2014

Abstract

We study families of Lagrangian tori that appear in a neighborhood of a resonance of a near-integrable Hamiltonian system. Such families disappear in the "integrable" limit $\varepsilon \to 0$. Dynamics on these tori is oscillatory in the direction of the resonance phases and rotating with respect to the other (non-resonant) phases.

We also show that, if multiplicity of the resonance equals one, generically these tori occupy a set of a large relative measure in the resonant domains in the sense that the relative measure of the remaining "chaotic" set is of order $\sqrt{\varepsilon}$. Therefore for small $\varepsilon > 0$ a random initial condition in a $\sqrt{\varepsilon}$-neighborhood of a single resonance occurs inside this set (and therefore generates a quasi-periodic motion) with a probability much larger than in the "chaotic" set.

We present results of numerical simulations and discuss the form of projection of such tori to the action space.

At the end of Section 4 we discuss the relationship of our results and a conjecture that tori (in a near-integrable Hamiltonian systems) occupy all the phase space except a set of measure $\sim \varepsilon$.

1 Projection of trajectories to the action space

Consider a symplectic map

$$(y, x) \mapsto (y_+, x_+), \qquad y \in \mathbb{R}^2, \quad x \in \mathbb{T}^2 \tag{1.1}$$

$$y_+ = y - \varepsilon\, \partial V / \partial x, \quad x_+ = x + y_+, \qquad V = V(x). \tag{1.2}$$

For $\varepsilon = 0$ the dynamics is integrable and (y, x) are action-angle variables on the phase space $\mathbb{R}^2 \times \mathbb{T}^2$. We choose for definiteness

$$V = a_1 \cos(x_1 + \varphi_1) + a_2 \cos(x_2 + \varphi_2) + a_3 \cos(x_1 - x_2 + \varphi_3)$$

*Lomonosov Moscow State University
†Space Research Institute of RAS, Department of Mathematical Sciences, Loughborough University
‡Steklov Mathematical Institute of RAS

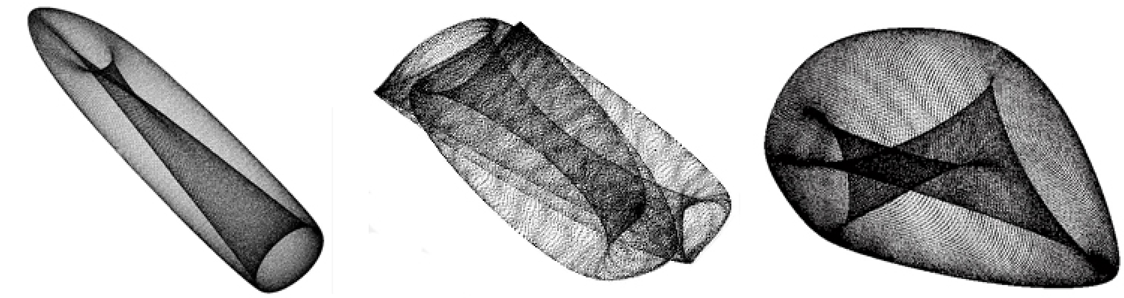

Figure 1: KAM-tori

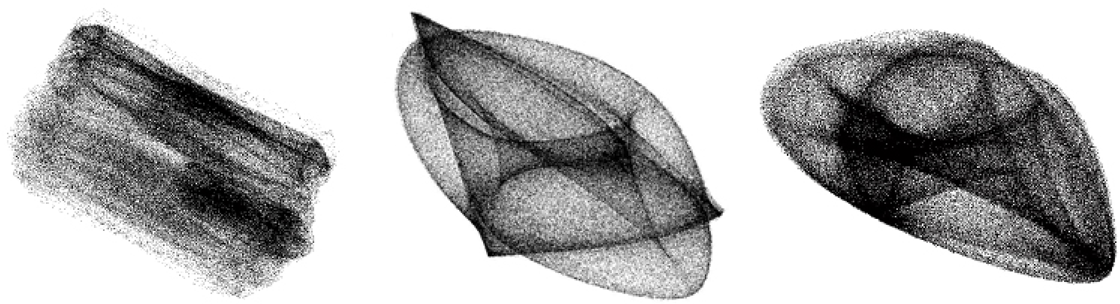

Figure 2: Chaotic trajectories: near KAM-tori (center,right) and near a low order resonance (left)

with constant $a_1, a_2, a_3, \varphi_1, \varphi_2, \varphi_3$. We study trajectories numerically having in mind the idea to observe some interesting images and then find theoretic explanations. To visualize objects (closures of trajectories) lying in the 4-dimensional phase space, we project them to the action plane $\mathbb{R}^2 = \{y\}$. We fix $a_j \sim 1$, φ_j, $\varepsilon \sim 1/10$, take a random initial condition and look at the result. The symplectic map (1.1)-(1.2) belongs to the family of maps introduced by Claude Froeschlé in [6], there is also the numerical simulation.

Typical images are presented on Fig. 1. They can be easily identified with ordinary KAM-tori. We see typical singularities (folds and pleats) of Lagrangian projections (these singularities are presented for example, in [3]).

It is also easy to observe chaotic trajectories, Fig. 2. The trajectories presented in Fig. 2, right and center, are situated near KAM-tori, but chaotic effects create a small "defocusing" in comparison with tori from Fig. 1.

Trying more initial conditions, it is possible to obtain more interesting objects, Fig. 3, which look like closed ribbons. Further attempts lead to more exotic images, Fig. 4, looking as finite sequences of spots.

In this paper we discuss a mechanism which generates such structures. In particular we show that these objects form sets of positive measure in the phase space. Hence probability to observe them is also positive.

On the end of this section let us give some examples. Fix $\varphi_1 = \varphi_2 = 1$, $\varphi_3 = 2$, $a_1 = 1$, $a_2 = 1$, $b = 0,8$, $\varepsilon = 0,15$ and $y_1 = 1$, $y_2 = 3$. If $x_1 = 0,74$, $x_2 = 1,685$ we obtain the right picture on Fig. 1, if $x_1 = 0,74$, $x_2 = 1,684$ we obtain the right picture on Fig. 4, if $x_1 = 1,3595$, $x_2 = 1,7$ we obtain the right picture on Fig. 2, finally if $x_1 = 0,74$, $x_2 = 1,707$ we obtain the right picture on Fig. 3.

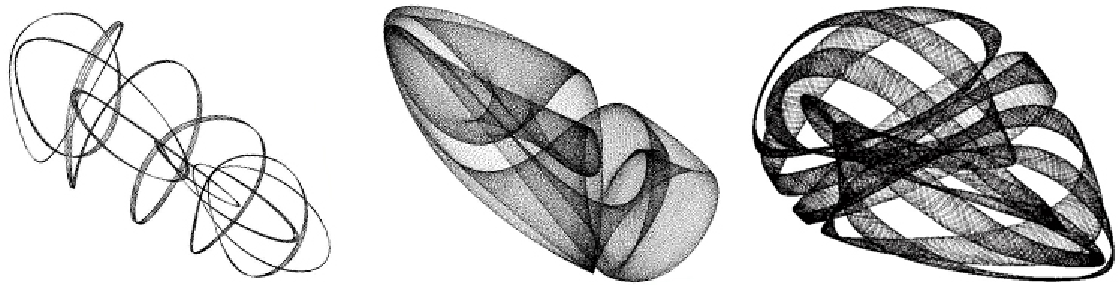

Figure 3: Closed "ribbons"

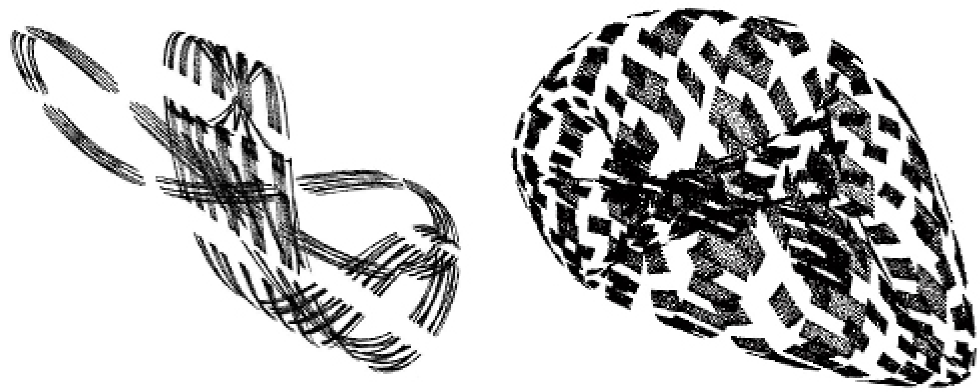

Figure 4: Another type of quasi-periodic trajectories, “sequences of spots”

2 Lower-dimensional tori near resonances

As usual it is more convenient to write formulas for flows although numerics are faster, simpler and more precise for maps. Consider a real-analytic near-integrable Hamiltonian system

$$\dot{X} = \partial H/\partial Y, \quad \dot{Y} = -\partial H/\partial X, \qquad Y \in \mathcal{D} \subset \mathbb{R}^N, \quad X \in \mathbb{T}^N. \tag{2.1}$$

$$H = H_0(Y) + \varepsilon H_1(Y, X) + O(\varepsilon^2), \qquad \varepsilon \geq 0. \tag{2.2}$$

Below we use the following notation for such a system:

$$(P, \omega_P, H), \qquad P = \mathcal{D} \times \mathbb{T}^N, \quad \omega_P = dY \wedge dX,$$

where the symplectic manifold (P, ω_P) is the phase space.

The map (1.1)–(1.2) can be regarded as the Poincaré map for some system (2.1)–(2.2) with $N = 3$ on an energy level $M_h = \{H = h = \text{const}\}$ (see for example [13]). Hence the dimension of the phase space $6 = 2N$ drops by 1 because of the reduction to M_h and by another 1 because of the passage to a (hyper) surface $\Sigma \subset M_h$ transversal to the flow. Below $N \geq 3$ is arbitrary.

The vector $\nu(Y) = \partial H_0/\partial Y$ is called an unperturbed frequency. For a fixed $Y = Y^0$ we have a fixed frequency $\nu^0 = \nu(Y^0)$. Any equation

$$\langle \nu^0, K \rangle = 0, \qquad K \in \mathbb{Z}^N \setminus \{0\} \tag{2.3}$$

is called a resonance. The word "resonance" is also attributed to the integer vector K, satisfying (2.3).

Given a constant $\nu^0 \in \mathbb{R}^N$ all the corresponding resonances K (together with $0 \in \mathbb{Z}^N$) form a resonance $\mathbb{Z}$-module

$$g(\nu^0) = \{K \in \mathbb{Z}^N : K \text{ is a resonance}\} \cup \{0\}.$$

If $g(\nu^0) = 0$, the frequency vector ν^0 is said to be nonresonant. We define

$$l = l(\nu^0) = \operatorname{rank}(g(\nu^0)), \quad m = m(\nu^0) = N - l,$$

where $\operatorname{rank}(g(\nu^0))$ is the number of generators in $g(\nu^0)$. Informally speaking, l is the number of independent resonances for the frequency vector ν^0. Invariant torus is called resonant (nonresonant) if the dynamics on the torus is quasi-periodic with a resonant (nonresonant) frequency vector. Any torus

$$\mathbb{T}^N_{Y^0} = \{(Y, X) : Y = Y^0\}$$

is invariant with respect to the unperturbed flow

$$(Y, X) \mapsto (Y, X + \nu(Y)t), \qquad t \in \mathbb{R}.$$

Then $\mathbb{T}^N_{Y^0}$ is foliated by the m-tori

$$\mathbb{T}^m_{Y^0,X^0} = \text{closure}\big(\{(Y, X) : Y = Y^0,\ X = X^0 + \nu^0 t,\ t \in \mathbb{R}\}\big).$$

Computing frequency vector corresponding to the unperturbed quasi-periodic motion on $\mathbb{T}^m_{Y^0,X^0}$, it is easy to show that the tori $\mathbb{T}^m_{Y^0,X^0}$ are non-resonant.

Note that in the non-resonant case ($l = 0$) we have $\mathbb{T}^m_{Y^0,X^0} = \mathbb{T}^m_{Y^0}$ for any $X^0 \in \mathbb{T}^N$. If $l > 0$ then $m \neq N$ and the foliation is non-trivial.

If $l > 0$ then a generic perturbation destroys $\mathbb{T}^m_{Y^0}$, [11]. However generically some tori $\mathbb{T}^m_{Y^0,X^0}$ survive a perturbation even if ν^0 is resonant. To present the corresponding result, we fix a $\mathbb{Z}$-module $g \subset \mathbb{Z}^n$. Consider the resonance set

$$\Sigma_g = \{Y \in \mathcal{D} : g(\nu(Y)) = g\}.$$

Under natural non-degeneracy conditions[1] Σ_g is a real-analytic submanifold, $\dim \Sigma_g = m$.

It is convenient to study system (2.1)–(2.2) in a $\sqrt{\varepsilon}$-neighborhood of the torus $\mathbb{T}^N_{Y^0}$, $Y^0 \in \Sigma_g$ by using the scaling

$$Y = Y^0 + \sqrt{\varepsilon}\widetilde{Y}, \quad X = \widetilde{X}, \qquad H(Y, X, \varepsilon) = H_0(Y^0) + \sqrt{\varepsilon}\widetilde{H}(\widetilde{Y}, \widetilde{X}, \sqrt{\varepsilon}).$$

Then the system (P, ω_P, H) turns to the system $(\widetilde{P}, \omega_{\widetilde{P}}, \widetilde{H})$,

$$\widetilde{H} = \langle \nu^0, \widetilde{Y}\rangle + \tfrac{1}{2}\sqrt{\varepsilon}\langle \Pi\widetilde{Y}, \widetilde{Y}\rangle + \sqrt{\varepsilon}\, H_1(Y^0, \widetilde{X}) + O(\varepsilon),$$
$$\widetilde{P} = \mathbb{R}^N \times \mathbb{T}^N, \quad \omega_{\widetilde{P}} = d\widetilde{Y} \wedge d\widetilde{X}, \quad \Pi = H''_{0\,YY}(Y^0).$$

[1] the functions $\langle K^{(j)}, \partial/\partial Y\rangle H_0(Y)$ are independent in $\mathcal{D}$ where $K^{(1)}, \dots, K^{(l)}$ are generators of g

The tori $\mathbb{T}^m_{Y^0,X^0}$ which survive the perturbation are generated by fixed points of some Hamiltonian system which is obtained from the initial one by using averaging, neglecting some higher order perturbative terms, and by reduction of the order. Now we turn to description of these steps.

For any function $f = \sum_{K\in\mathbb{Z}^N} f^K e^{i\langle K,X\rangle}$ consider the averaging

$$\langle f\rangle_g = \sum_{K\in g} f^K e^{i\langle K,X\rangle}. \tag{2.4}$$

Hence, $\langle\cdot\rangle_g$ is a projector, removing all nonresonant Fourier harmonics.

Now it is natural to perform the following standard coordinate change:

$$(\widetilde{Y},\widetilde{X}) \mapsto (\widehat{Y},\widehat{X}), \qquad \widetilde{Y} = \widehat{Y} + \sqrt{\varepsilon}\,\widetilde{S}_{\widetilde{X}}, \quad \widetilde{X} = \widehat{X},$$

where $\widetilde{S} = \widetilde{S}(\widetilde{X})$ is a solution of the (co)homologic equation

$$\langle \nu^0, \widetilde{S}_{\widetilde{X}}\rangle + H_1(Y^0,\widetilde{X}) = v(Y^0,\widetilde{X}), \qquad v(Y,X) = \langle H_1(Y,X)\rangle_g. \tag{2.5}$$

Under standard Diophantine conditions a real-analytic solution of equation (2.5) exists and unique up to a g-invariant additive term $\langle\widetilde{S}\rangle_g$. For example, $\widetilde{S}$ can be chosen so that $\langle\widetilde{S}\rangle_g = 0$.

In the new coordinates we have the system $(\widehat{P},\omega_{\widehat{P}},\widehat{H})$,

$$\widehat{P} = \mathbb{R}^N\times\mathbb{T}^N, \quad \omega_{\widehat{P}} = d\widehat{Y}\wedge d\widehat{X}, \quad \widehat{H} = H^g + O(\varepsilon),$$
$$H^g = \langle\nu^0,\widehat{Y}\rangle + \frac{\sqrt{\varepsilon}}{2}\langle\Pi\widehat{Y},\widehat{Y}\rangle + \sqrt{\varepsilon}\,v(Y^0,\widehat{X}).$$

Consider the approximate system $(\widehat{P},\omega_{\widehat{P}},H^g)$, which is usually called the partially averaged system. Any critical point X^0 of the "potential" $v(Y^0,X)$ generates an invariant torus $\mathbb{T}^m_{Y^0,X^0}$. To study linearization of the averaged system on $\mathbb{T}^m_{Y^0,X^0}$ it is convenient to consider the reduced system. First we recall general invariant construction (see [2]) and then give a more explicit coordinate form.

The system $(\widehat{P},\omega_{\widehat{P}},H^g)$ admits the symmetry group $G\cong\mathbb{T}^m$. The Lie algebra associated with G is naturally identified with

$$g_\perp = \{Q\in\mathbb{R}^N : \langle Q,K\rangle = 0 \text{ for any } K\in g\}.$$

Note that $\nu^0\in g_\perp$ and $\operatorname{rank} g_\perp = m$.

Action of G on $\widehat{P}$ is Poissonian and the corresponding momentum map $\mathcal{M}_G:\widehat{P}\to g_\perp^*$ is as follows:

$$\mathcal{M}_G(\widehat{Y},\widehat{X})\,Q = \langle Q,\widehat{Y}\rangle \quad \text{for any } Q\in g_\perp.$$

Reduction with respect to G means

(a) fixing values of the first integrals $\langle Q,\widehat{Y}\rangle = \langle Q,\gamma\rangle$ for any $Q\in g_\perp$, where $\gamma = \text{const}\in\mathbb{R}^N$,

(b) passage to the quotient phase space

$$\mathcal{P} = \widehat{P}_\gamma/G, \qquad \widehat{P}_\gamma = \big\{(\widehat{Y},\widehat{X})\in\widehat{P} : \langle Q,\widehat{Y}\rangle = \langle Q,\gamma\rangle \text{ for all } Q\in g_\perp\big\}.$$

The reduced phase space has a canonical symplectic structure ω_γ [2] while the Hamiltonian $\mathcal{H}^g : \mathcal{P} \to \mathbb{R}$ is determined by the commutative diagram

$$\begin{array}{ccc} \widehat{\mathcal{P}}_\gamma & \xrightarrow{\text{pr}} & \mathcal{P} \\ {\scriptstyle H^g|_{\widehat{\mathcal{P}}_\gamma}} \searrow & & \swarrow {\scriptstyle \mathcal{H}^g} \\ & \mathbb{R} & \end{array}$$

where pr is the natural projection. The torus $\mathbb{T}^m_{Y^0,X^0}$ turns to a fixed point p in the reduced system $(\mathcal{P}, \omega_\gamma, \mathcal{H}^g)$. Flow of the system $(P, \omega_P, \hat{H}^g)$ near $\mathbb{T}^m_{Y^0,X^0}$ is essentially determined by linear approximation of $(\mathcal{P}, \omega_\gamma, \mathcal{H}^g)$ at p. To obtain this approximation, we turn to the coordinate form of the above order reduction.

Let Θ be an $N \times l$ matrix formed by the integer vectors $K^{(1)}, \dots, K^{(l)}$, generators of g. This matrix is not unique: for any $M \in SL(l, \mathbb{Z})$ (an integer $l \times l$ matrix with unit determinant) one may take ΘM instead of Θ. We assume that the quadratic form determined by Π is non-degenerate on $g_\mathbb{R}$, where $g_\mathbb{R} \subset \mathbb{R}^N$ is the natural extension of $g(\nu^0)$ to a linear subspace. Equivalently $\Theta^T \Pi \Theta$ is a non-degenerate $l \times l$-matrix.

Then we can take as local coordinates on $\mathcal{P}$ the variables $\eta \in \mathbb{R}^l$, $\xi \in \mathbb{T}^l$ such that

$$\widehat{Y} = \Theta(\eta - \eta^0) + \gamma, \quad \xi = \Theta^T X, \qquad \eta^0 = (\Theta^T \Pi \Theta)^{-1} \Theta^T \Pi \gamma.$$

Here the constant η^0 is chosen for convenience to remove from $\mathcal{H}^g$ a term linear in η.

The form ω_γ, the fixed point p, and the function $\mathcal{H}^g$ are as follows:

$$\omega_\gamma = d\eta \wedge d\xi, \quad \eta(p) = 0, \ \xi(p) = \Theta^T X^0, \quad \mathcal{H}^g = \frac{\sqrt{\varepsilon}}{2} \langle \Theta^T \Pi \Theta \eta, \eta \rangle + \sqrt{\varepsilon}\, v_{\mathcal{P},Y^0}(\xi) + h_\gamma.$$

Here $v_{\mathcal{P},Y^0} : \mathbb{T}^l \to \mathbb{R}$ is the unique function, satisfying the identity[2]

$$v_{\mathcal{P},Y^0}(\Theta^T X) = v(Y^0, X).$$

The constant $h_\gamma = \frac{1}{2} \langle \Pi(\gamma - \Theta\eta^0), (\gamma - \Theta\eta^0) \rangle$ can be ignored.

Theorem 1 ([8]). *Suppose that* $\det \Pi \neq 0$ *in* D. *Then for any sufficiently small* $\varepsilon > 0$ *there exists a set* $\Lambda_\varepsilon \subset \Sigma_g$ *such that for each* $Y^0 \in \Lambda_\varepsilon$ *and for each nondegenerate critical point* ξ^0 *of* $v_{\mathcal{P},Y^0}$ *the perturbed system admits a real-analytic invariant m-torus* $\mathbb{T}^m_{Y^0,\xi^0}(\varepsilon)$. *This torus is close to* $\mathbb{T}^m_{Y^0,X^0}$, *where* X^0 *is any point satisfying the equation* $\xi^0 = \Theta^T X^0$. *Moreover,* $\mathbb{T}^m_{Y^0,\xi^0}(\varepsilon)$ *carries a quasi-periodic motion with the same frequency vector.*

The perturbed invariant m-tori constitute a finite number of m-parameter Whitney smooth families. The relative Lebesgue measure of Λ_ε *on the surface*

$$\{Y \in \Sigma_g : v_{\mathcal{P},Y^0} \text{ has nondegenerate critical points}\}$$

tends to 1 as $\varepsilon \to 0$.

---

[2] Explicit formula for $v_\mathcal{P}$ is as follows. Let the Fourier expansion for v be

$$v(Y, X) = \sum_{K \in g} v^K(Y) e^{i\langle K, X \rangle} = \sum_{j \in \mathbb{Z}^l} v^{\Theta j}(Y) e^{i\langle \Theta j, X \rangle}.$$

Then $v_{\mathcal{P},Y}(\xi) = \sum_{j \in \mathbb{Z}^l} v^{\Theta j}(Y) e^{i\langle j, \xi \rangle}$.

Hamiltonian of the linear approximation for $(\mathcal{P}, \omega_\gamma, \mathcal{H}^g)$ at the fixed point p is

$$\mathcal{H}^g_{lin} = \frac{\sqrt{\varepsilon}}{2}\langle \Theta^T \Pi \Theta \eta, \eta\rangle + \frac{\sqrt{\varepsilon}}{2}\langle V\zeta, \zeta\rangle, \qquad V = \frac{\partial^2 v_{\mathcal{P},Y^0}}{\partial \xi^2}(0), \quad \zeta = \xi - \xi^0.$$

Therefore eigenvalues $\pm\lambda_1, \dots, \pm\lambda_l$ of the fixed point $\eta = \zeta = 0$ in the system $(\mathcal{P}, \omega_\gamma, \mathcal{H}^g_{lin})$ satisfy the equation

$$\det(\varepsilon \Theta^T \Pi \Theta V + \lambda^2) = 0.$$

If all λ_j are purely imaginary, p and the corresponding torus $\mathbb{T}^m_{Y^0,\xi^0}(\varepsilon)$ are said to be normally elliptic. The "opposite" case is normally hyperbolic, where no λ_j is purely imaginary. In general situation the torus $\mathbb{T}^m_{Y^0,\xi^0}(\varepsilon)$ has an associated centre manifold.

For $m = 1$ Theorem 1 was proven by Poincaré [11]. In this case no small divisors appear and the proof is based on the ordinary implicit function theorem. The equation $m = N$ corresponds to the (ordinary) KAM-theorem for Lagrangian tori.

Hyperbolic case with arbitrary m is presented in [18]. In [4] the case of arbitrary m and arbitrary normal behavior of the perturbed tori is treated. In [8] it is shown that an additional condition (the so called g-nondegeneracy of H_0), introduced in [4], can be skipped. A statement analogous to Theorem 1 should be true in infinite dimension but as far as we know this has not been proven yet.

Since for non-trivial g all the tori $\mathbb{T}^m_{Y^0,\xi^0}(\varepsilon)$ are lower-dimensional ($m < N$), their total measure in the phase space $D \times \mathbb{T}^N$ vanishes. In other words they are practically invisible in numerical experiments. This does not mean that they are inessential for dynamics. For example, hyperbolic lower-dimensional tori and their asymptotic manifolds are known as elementary links which form transition chains, forming a basis for the Arnold diffusion.

3 Visible objects

In this paper we are interested in visible objects. More precisely, in invariant tori of dimension N. Geometry of their projections to the action space depends on the order of a resonance at which these tori appear.

No resonance. For example, such objects are ordinary (N-dimensional) KAM-tori. If $\varepsilon > 0$ is small, KAM tori form a large Cantorian set: the measure of the complement $\mathcal{C}(\varepsilon)$ to this set in $D \times \mathbb{T}^N$ does not exceed a quantity of order $\sqrt{\varepsilon}$, [2, 7, 10, 12, 17], in the case $N = 2$ the measure of $\mathcal{C}(\varepsilon)$ is exponentially small in ε [10]. The measure estimates of $\mathcal{C}(\varepsilon)$ for degenerate systems are contained in [1, 9, 19, 14, 15].

In the first $\sqrt{\varepsilon}$ approximation the projection of a KAM-torus to the action space $\mathcal{D}$ has the form $\widehat{Y} = 0$ under the condition $l = \operatorname{rank} g = 0$. In the original coordinates we have:

$$\big\{ Y \in \mathcal{D} : Y = Y^0 + \sqrt{\varepsilon}\widetilde{S}_{\widetilde{X}}(\widetilde{X}), \quad \widetilde{X} \in \mathbb{T}^N \big\},$$

where $\widetilde{S}$ satisfies (2.5).

The discrete system (1.1)–(1.2) can be obtained from a Hamiltonian system (2.1)–(2.2) with $N = 3$ degrees of freedom on an energy level $H = \text{const}$ by passing to the Poincaré map on the section $\{X_3 = 0\}$. Hence, the first in $\sqrt{\varepsilon}$ approximation the objects presented in Fig. 1 are sets of the form

$$\big\{(Y_1, Y_2) : Y_j = Y_j^0 + \sqrt{\varepsilon}\widetilde{S}_{\widetilde{X}_j}(\widetilde{X}_1, \widetilde{X}_2, 0), \quad j = 1, 2, \quad (\widetilde{X}_1, \widetilde{X}_2) \in \mathbb{T}^2\big\}. \tag{3.1}$$

For a random initial condition the probability to occur on one of such torus is greater than $1 - C_0\sqrt{\varepsilon}$ for some $C_0 > 0$.

The "resonance" set $\mathcal{C}(\varepsilon)$ contains other families of quasi-periodic motions. Here it is reasonable to distinguish the case of a single resonance ($\operatorname{rank} g = 1$) and the case of a multiple resonance ($\operatorname{rank} g > 1$).

Single resonance. In this case dimension of the commutative symmetry group G is $N - 1$. Therefore the system (P, ω_P, H^g) is completely integrable. Informally speaking, it is a product of $N - 1$ "rotators" and a "pendulum" (such representation in a small neighborhood of a single resonance is discussed in [3]), where the pendulum is determined by the reduced system with one degree of freedom $(\mathcal{P}, \omega_\gamma, \mathcal{H}^g)$. The variables η and ξ are 1-dimensional while vector $\Theta = (\Theta_1, \dots, \Theta_N)^T$ is integer.

Solutions on which the pendulum motions are rotations are ordinary KAM-tori while solutions on which the pendulum oscillates, lie in $\mathcal{C}(\varepsilon)$. When we say that the pendulum oscillates, we mean that in the system $(\mathcal{P}, \omega_\gamma, \mathcal{H}^g)$ the angular variable ξ changes periodically in an interval $J \subset \mathbb{T}^1$, $J \neq \mathbb{T}^1$. In the first approximation in $\sqrt{\varepsilon}$ the corresponding trajectory of (1.1)–(1.2) fills the set

$$\big\{(Y_1, Y_2) : Y_j = Y_j^0 + \sqrt{\varepsilon}\widetilde{S}_{\widetilde{X}_j}(\widetilde{X}_1, \widetilde{X}_2, 0), \quad j = 1, 2, \quad \Theta_1\widetilde{X}_1 + \Theta_2\widetilde{X}_2 \in J, \quad (\widetilde{X}_1, \widetilde{X}_2) \in \mathbb{T}^2\big\}.$$

This is a "ribbon-like" subset of (3.1). This explains the structure of sets in Fig. 3.

Since the system (P, ω_P, H^g) is integrable, the set of phase points in a $\sqrt{\varepsilon}$-neighborhood of the torus $\mathbb{T}^N_{Y^0}$ lying outside invariant N-tori of the original system (P, ω_P, H), has a small relative measure. Precise statement, Theorem 2, is given in Section 4. Hence if an initial condition is taken randomly the probability to obtain a quasi-periodic orbit like Fig. 3 is not less than $C_1\sqrt{\varepsilon}$, $\quad C_1 > 0$, because the width of a resonance domain corresponding to a single resonance is $\sim \sqrt{\varepsilon}$. Pictures analogous to Fig. 3 can be found in [16]. Similar tori considered in context of Arnold diffusion in [5].

Multiple resonance. If $\operatorname{rank} g > 1$, the systems $(\widehat{P}, \omega_{\widehat{P}}, H^g)$ and $(\mathcal{P}, \omega_\gamma, \mathcal{H}^g)$ are generically non-integrable. Therefore existence of invariant l-tori in the latter one is not straightforward provided the energy $\mathcal{H}^g$ is not very big or not very small. A standard source for such tori is a neighborhood of a totally elliptic fixed point. However, a totally elliptic fixed point may not exist if the "kinetic energy" $\frac{1}{2}\sqrt{\varepsilon}\langle\Theta^T\Pi\Theta\eta, \eta\rangle$ is indefinite: a simple example is

$$\mathcal{H}^g = \sqrt{\varepsilon}\,(\eta_1\eta_2 + \cos\xi_1 + \cos\xi_2).$$

System (1.1)–(1.2) corresponds to a positive definite kinetic energy and trajectories presented in Fig. 4 present nonlinear versions of small oscillations near a totally elliptic periodic orbit in the corresponding system with 3 degrees of freedom. A random initial condition lies on one of such tori with probability of order ε, because the measure of a resonant domain corresponding to a double resonance is of order ε. Unlike the case of a single resonance only a small portion of this domain is filled with tori, in general.

4 Invariant N-tori at a single resonance

Putting $N = n+1$, consider the Hamiltonian system (2.1)–(2.2) in a neighborhood of the resonance $\Sigma_g \times \mathbb{T}^{n+1}$, where g is generated by the vector $K^0 \in \mathbb{Z}^{n+1} \setminus \{0\}$ with relatively prime components. Hence, we plan to study invariant tori, located in the vicinity of a single resonance

$$\Sigma = \Sigma_g = \{Y \in \mathcal{D} : \langle K^0, \nu(Y)\rangle = 0\}. \tag{4.1}$$

We assume that the unperturbed system is non-degenerate and K^0 is not a light-like vector:

$$\det H''_{0YY} \neq 0, \qquad Y \in \mathcal{D}, \tag{4.2}$$

$$\mathcal{A}(Y) \equiv \big\langle K^0, H''_{0YY}\, K^0 \big\rangle \neq 0. \tag{4.3}$$

Note that (4.3) means that for any Y satisfying (4.1) the function $\lambda \mapsto H_0(Y - \lambda K^0)$ has a non-degenerate critical point $\lambda = 0$.

Now our aim is to give a defenition of the oscillatory part of the resonance domain and to introduce convenient notation for the main KAM theorem.

By (4.3) the resonant set $\Sigma \subset \mathcal{D}$ is a smooth hypersurface transversal to the constant vector field K^0. The equation

$$\frac{d}{d\lambda} H_0(Y - \lambda K^0) \equiv \big\langle K^0, \nu(Y - \lambda K^0)\big\rangle = 0 \tag{4.4}$$

has a real-analytic solution $\lambda = \lambda(Y)$ in $\mathcal{D}$ near Σ.

We have a smooth map $\chi : U(\Sigma) \to \Sigma$, where $U(\Sigma)$ is a neighborhood of the resonance Σ and $\chi(Y) = Y - \lambda(Y) K^0$.

Let $\langle\cdot\rangle^{K^0} \equiv \langle\cdot\rangle_g$ be the operator of resonant averaging

$$f = \sum_{K \in \mathbb{Z}^{n+1}} f^K\, e^{i\langle K, X\rangle} \;\mapsto\; \langle f\rangle^{K^0} = \sum_{j=-\infty}^{\infty} f^{jK^0}\, e^{ijq}, \qquad q = \langle K^0, X\rangle.$$

Consider the Hamiltonian system (P, ω_P, H_{K^0}),

$$H_{K^0}(Y, X) = H_0(Y) + \varepsilon\, \mathbf{u}(Y, q), \quad \mathbf{u}(Y, q) = \big\langle H_1\big(\chi(Y), X\big)\big\rangle^{K^0}. \tag{4.5}$$

The function $\mathbf{u}$ is 2π-periodic in the resonant phase q and

$$\mathbf{u}(Y + \lambda K^0, q) = \mathbf{u}(Y, q) \quad \text{for any } \lambda \text{ in a neighborhood of } 0 \in \mathbb{R}.$$

In fact, below the quantity $\lambda(Y) K^0 = Y - \chi(Y)$ (some sort of distance to the resonant surface) will be of order $\sqrt{\varepsilon}$.

Since $H_0(Y) = H_0(\chi(Y)) + \frac{1}{2}\mathcal{A}(\chi(Y))\, \lambda^2(Y) + O(\lambda^3(Y))$, Hamiltonian (4.5) can be presented in the form

$$H_{K^0}(Y, X) = H_0(\chi(Y)) + \frac{1}{2}\mathcal{A}(\chi(Y))\, \lambda^2(Y) + \varepsilon \mathbf{u}(Y, q) + O(\lambda^3(Y)). \tag{4.6}$$

For any vector $J \in \mathbb{R}^{n+1}$ such that $\langle K^0, J\rangle = 0$ the function $\langle Y, J\rangle$ is a first integral. Therefore the system (P, ω_P, H_{K^0}) is completely integrable. It is responsible for the dynamics of the original system near the resonance (4.1). Below we only deal with the oscillatory part $\mathcal{D}_{os}$ of the resonance domain, where $\mathcal{D}_{os} \subset \mathcal{D} \times \mathbb{T}^{n+1}$ is defined as follows. Let $q_{\min}(Y)$ and $q_{\max}(Y)$ be points of global minimum and maximum of $\mathbf{u}(Y, q)$ for fixed $Y \in \Sigma$:

$$\mathbf{u}(Y, q_{\min}(Y)) = \min_{q\in\mathbb{T}} \mathbf{u}(Y, q), \quad \mathbf{u}(Y, q_{\max}(Y)) = \max_{q\in\mathbb{T}} \mathbf{u}(Y, q). \tag{4.7}$$

Then we define

$$\begin{aligned}\mathcal{D}_{os} \ = \ \big\{&(Y, X) \in \mathcal{D} \times \mathbb{T}^{n+1} \ : \\ &\varepsilon\mathbf{u}(Y, q_{\min}(Y)) < \frac{1}{2}\mathcal{A}(\chi(Y))\,\lambda^2(Y) + \varepsilon\mathbf{u}(Y, q) < \varepsilon\mathbf{u}(Y, q_{\max}(Y))\big\}.\end{aligned} \tag{4.8}$$

If $\mathbf{u}(Y, q)$ is not a constant as a function of q, the domain $\mathcal{D}_{os}$ belongs to an $O(\sqrt{\varepsilon})$-neighborhood of the resonance $\Sigma \times \mathbb{T}^{n+1}$. On almost any orbit of this flow located in $\mathcal{D}_{os}$ the resonant phase q oscillates between two quantities q_1 and q_2, depending on initial conditions and such that

$$|q_2 - q_1| < 2\pi, \quad \dot q|_{q=q_1} = \dot q|_{q=q_2} = 0.$$

These orbits lie on $(n + 1)$-dimensional Lagrangian tori. Below we prove that for small values of ε all these tori except a set of a small measure survive the perturbation.

To formulate the result, for any Y in a neighborhood of Σ consider the Hamiltonian system

$$\dot q = \sqrt{\varepsilon}\,\hat{\mathbf{h}}'_p(\chi(Y), p, q), \quad \dot p = -\sqrt{\varepsilon}\,\hat{\mathbf{h}}'_q(\chi(Y), p, q). \tag{4.9}$$

with one degree of freedom, the Hamiltonian

$$\hat{\mathbf{h}}(\chi(Y), p, q) = \frac{1}{2}\mathcal{A}\big(\chi(Y)\big)p^2 + \mathbf{u}(Y, q)$$

and the symplectic structure $\frac{1}{\sqrt{\varepsilon}}dp \wedge dq$. The point Y is regarded as a parameter. Recall that by (4.3) $\mathcal{A} \neq 0$. This system coincides with $(\mathcal{P}, \omega_\gamma, \mathcal{H}^g)$ from Section 3 written in slightly other terms.

Proposition 4.1 *The point $\chi(Y)$ is a constant of motion in the system (P, ω_P, H_{K^0}). The variables $\lambda(Y) = \sqrt{\varepsilon}p$ and q in the averaged system satisfy equations (4.9) up to $O(\varepsilon + \lambda^2)$.*

Indeed, the first statement of Proposition 4.1 follows from the relation $\dot Y \parallel K_0$.

Applying the operator $\langle K^0, \partial/\partial Y\rangle$ to (4.4) we get: $\langle K^0, \lambda_Y(Y)\rangle = 1 + O(\lambda)$. Then by (4.6) in system (P, ω_P, H_{K^0})

$$\begin{aligned}\dot\lambda \ &= \ \lambda'_Y\,\dot Y = -\varepsilon\mathbf{u}'_q(Y, q) + O(\varepsilon\lambda) = -\varepsilon\hat{\mathbf{h}}'_q + O(\varepsilon\lambda), \\ \dot q \ &= \ \langle K^0, H'_{K^0\,Y}\rangle = \mathcal{A}(\chi(Y))\,\lambda + O(\lambda^2) + O(\varepsilon) = \sqrt{\varepsilon}\,\hat{\mathbf{h}}'_p + O(\varepsilon + \lambda^2).\end{aligned}$$

■

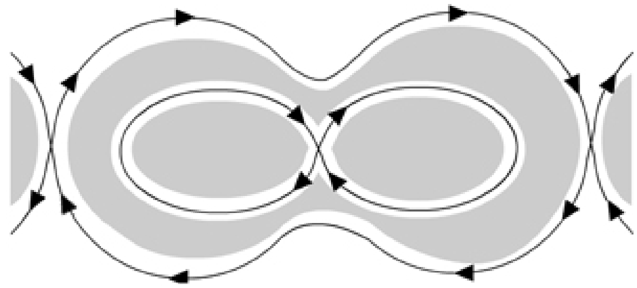

Figure 5: Phase portrait of system (4.9). Section of the oscillatory domain $\mathcal{D}_0$ by $\{\chi(Y) = \text{const}\}$ is marked grey.

The projection $\pi : U(\Sigma \times \mathbb{T}^{n+1}) \to \Sigma \times \mathbb{R} \times \mathbb{T}$,

$$\pi(Y, X) = (\chi(Y), p, q), \qquad p = \varepsilon^{-1/2}\lambda(Y), \quad q = \langle K^0, X\rangle$$

maps the domain $\mathcal{D}_{os}$ to $\hat{\mathcal{D}}_{os} = \pi(\mathcal{D}_{os})$, see Fig. 5.

Let the closed curve $\gamma(Y, p, q)$ be the connected component of the set

$$\big\{(\tilde{p}, \tilde{q}) : \hat{\mathbf{h}}(\chi(Y), \tilde{p}, \tilde{q}) = \hat{\mathbf{h}}(\chi(Y), p, q)\big\}$$

containing the point (p, q). We define the action variable I and the Hamiltonian function $\mathbf{h}$:

$$I = I(Y, p, q) = \frac{1}{2\pi} \int_{\gamma(Y,p,q)} \tilde{p}\, d\tilde{q}, \qquad \mathbf{h}(\chi(Y), I) = \hat{\mathbf{h}}(\chi(Y), p, q). \tag{4.10}$$

Note that if the energy levels $\mathbf{h}(\chi(Y), p, q) = \text{const}$ are not connected (i.e., consist of several curves γ), the function $\mathbf{h}$ is not single-valued.

If $\gamma = \gamma(Y, p, q)$ is a closed smooth curve, the torus

$$\mathbb{T}^{n+1}_{Y^0,\gamma}(\varepsilon) = \big\{(Y, X) : \chi(Y) = Y^0, (\varepsilon^{-1/2}\lambda(Y), q) \in \gamma\big\} \tag{4.11}$$

is invariant for the system (P, ω_P, H_{K^0}) up to terms of order $O(\varepsilon + \lambda^2)$. For small $\varepsilon > 0$ majority of tori (4.11) survive the perturbation and exist in the original system. Any surviving torus has to satisfy several additional conditions.

(1) The frequency vector $\nu_{Y^0,\gamma}$ associated with $\mathbb{T}^{n+1}_{Y^0,\gamma}$ is Diophantine, i.e. for some $\tau > 0$ and $C_\tau > 0$

$$|\langle K, \nu_{Y^0,\gamma}\rangle| > \frac{C_\tau}{|K|^{n+1+\tau}}, \quad K \in \mathbb{Z}^{n+1} \setminus \{0\}. \tag{4.12}$$

This is a standard assumption which holds on almost all tori.

(2) The system (P, ω_P, H_{K^0}) is nondegenerate on $\mathbb{T}^{n+1}_{Y^0,\gamma}$.

This condition essentially means that

$$|\mathbf{h}'_I| < c, \quad |\mathbf{h}''_{II}| < c, \quad |\mathbf{h}'_I|^{-1} < c, \quad |\mathbf{h}'_{II}|^{-1} < c. \tag{4.13}$$

Therefore we have to replace $\mathcal{D}_{os}$ by a smaller domain $\mathcal{D}_0$ by throwing out a small neighborhood of asymptotic manifolds (where the tori degenerate) and a small neighborhood of tori on which the twist conditions (4.13) is violated. Let μ be the measure in the phase space $\mathcal{D} \times \mathbb{T}^{n+1}$ generated by the symplectic structure ω_P. Then the measures $\mu(\mathcal{D})$ and $\mu(\mathcal{D}_0)$ are both of order $O(\sqrt{\varepsilon})$.

Theorem 2 *Suppose that the system (P, ω_P, H) is real-analytic. If $\varepsilon_0 > 0$ is sufficiently small, then for all positive $\varepsilon \leq \varepsilon_0$ any Diophantine torus $\mathbb{T}^{n+1}_{Y^0,\gamma}(\varepsilon) \subset \mathcal{D}_0$ survives the perturbation. For some constant $C > 0$ independent of ε, measure of union of such tori in $\mathcal{D}_0$ is not less than $\mu(\mathcal{D}_0) - C\varepsilon$.*

The measure of invariant tori in a near-integrable Hamiltonian system. It is well known that measure of the complement to the KAM tori does not exceed a quantity of order $\sqrt{\varepsilon}$ ([2, 7, 10, 12, 17]). In order to prove this one has to construct some KAM procedure and at each step of the procedure remove from the phase space a small resonant strip (the measure of this strip is $\sim \sqrt{\varepsilon}$). The total measure of all strips is of order $\sqrt{\varepsilon}$.

In Theorem 2 we consider a resonant strip of width $\sim \sqrt{\varepsilon}$. We remove from this strip the set where the system degenerates and prove that the remaining part of the strip has a lot of tori (the relative measure of "chaotic" set is of order $\sqrt{\varepsilon}$). It would be interesting to modify the proof, considering weaker non-degeneracy conditions (4.13) to improve estimates of the measure of the complement to the tori of a near-integrable Hamiltonian systems. Here it is natural to remind a conjecture (see [2]) that tori occupy all the phase space except a set of measure $\sim \varepsilon$.

5 Preliminaries

Beginning from this place till the end of the paper we prove Theorem 2.

- All vectors by default are regarded as columns. For any $u \in \mathbb{R}^m$ and any $m \times m$-matrix A we use the notation

$$|u| = \max_{1 \leq j \leq m} |u_j|, \quad |A| = \max_{0 \neq u \in \mathbb{R}^m} \frac{|Au|}{|u|}.$$

  The brackets $\langle\,,\rangle$ denote the standard Euclidean scalar product: $\langle u, v\rangle = \sum_{j=1}^m u_j v_j$.

- μ denotes the standard Lebesgue measure on $\mathbb{R}^m$.

- Prime denotes a partial derivative e.g., $f'_{y_k} = \partial f / \partial y_k$. If $y \in \mathbb{R}^m$, and $f : \mathbb{R}^m \to \mathbb{R}$ then f'_y is regarded as a vector and f''_{yy} as a matrix.

- For any function $f : \mathbb{T}^m \to \mathbb{R}$ we define its average

$$\langle f \rangle = \frac{1}{(2\pi)^m} \int_{\mathbb{T}^m} f(x)\, dx. \tag{5.1}$$

  The same notation is used if f depends on other variables. In this case to avoid misunderstanding we use for (5.1) the notation $\langle\ \rangle_x$.

- Below $c_1, c_2, \dots$ denote positive constants. If c_j depends on another constant, say, α, we write $c_j(\alpha)$. Dependence on the dimension n is not indicated.

To present the system (P, ω_P, H^g) in a form convenient for application of KAM procedure, we have to perform several preliminary coordinate changes.

(a). Consider a matrix $M \in GL(n+1, \mathbb{Z})$ such that K^0 is its last column. In the new coordinates

$$\hat{Y} = M^{-1} Y, \quad \hat{X} = M^T X$$

resonance (4.1) takes the form

$$\widehat{\nu}_{n+1}(\hat{Y}) \equiv \widehat{H}'_{0\,\hat{Y}_{n+1}}(\hat{Y}) = 0, \qquad \widehat{H}_0(\hat{Y}) = H_0(Y). \tag{5.2}$$

To have more convenient coordinates in a neighborhood of this resonance, we solve the first equation (5.2) with respect to $\hat{Y}_{n+1}$. This can be done locally because by (4.3) $\widehat{H}''_{0\,\hat{Y}_{n+1}\hat{Y}_{n+1}} \neq 0$. We denote the result

$$\hat{Y}_{n+1} = G(\hat{Y}_1, \dots, \hat{Y}_n), \qquad \hat{\nu}_{n+1}\big(\hat{Y}_1, \dots, \hat{Y}_n, G(\hat{Y}_1, \dots, \hat{Y}_n)\big) \equiv 0.$$

(b). Consider the change of the variables

$$y = (y_1, \dots, y_n) = (\hat{Y}_1, \dots, \hat{Y}_n), \quad p = \varepsilon^{-1/2}(\hat{Y}_{n+1} - G(y)),$$
$$x = (x_1, \dots, x_n) = \big(\hat{X}_1 + G'_{y_1} q, \dots, \hat{X}_n + G'_{y_n} q\big), \quad q = \hat{X}_{n+1}.$$

We are interested in motions which are oscillatory in the coordinate q. Therefore it is not necessary to assume periodicity of this change with respect to q. Below q lies in an interval while the variables x are still angular: $x \in \mathbb{T}^n$.

The resonance Σ in the new coordinates locally takes the form $\{p = 0\}$ while y are local coordinates on Σ.

Then symplectic structure and the Hamiltonian take the form

$$\omega = dy \wedge dx + \sqrt{\varepsilon}\, dp \wedge dq,$$
$$\Lambda(y) + \varepsilon\big(\tfrac{1}{2} A(y) p^2 + u(y, q)\big) + \varepsilon U_1(y, x, q) + \varepsilon^{3/2} U_2(y, p, x, q, \sqrt{\varepsilon}), \tag{5.3}$$

where

$$\Lambda(y) = \widehat{H}_0(y, G(y)), \quad A(y) = \widehat{H}''_{0\,\hat{Y}_{n+1}\hat{Y}_{n+1}}(y, G(y)),$$

u, U_1, U_2 are real-analytic, and average of U_1 with respect to x vanishes: $\langle U_1 \rangle_x = 0$.

By (4.3),(4.5)

$$A(y) = \mathcal{A}\big(\chi(Y)\big), \quad u(y, q) = \mathbf{u}(\chi(Y), q). \tag{5.4}$$

Neglecting the terms $\varepsilon U_1 + \varepsilon^{3/2} U_2$, we obtain an integrable system which can be regarded as a skew-product of an n-dimensional rotator in variables y, x and a (generalized) pendulum in variables p, q.

We put (see (4.7))

$$\hat{q}_{\min}(y) = q_{\min}(\chi(Y)), \quad \hat{q}_{\max}(y) = q_{\max}(\chi(Y)).$$

Then the domain $\mathcal{D}_{os}$ (see (4.8)) takes the form

$$\mathcal{D}_{os} = \big\{(y,p,x,q) : u(y,\hat{q}_{\min}) < \frac{1}{2}A(y)p^2 + u(y,q) < u(y,\hat{q}_{\max})\big\}. \qquad (5.5)$$

(c). Let $W(y,I,q)$ be a generating function which introduces action-angle variables I,φ in domain (5.5) for the system with one degree of freedom, the symplectic structure $dp\wedge dq$ and the Hamiltonian $\frac{1}{2}A(y)p^2+u(y,q)$, variables y are parameters y:

$$p = W'_q, \quad \varphi = W'_I, \quad h(y,I) = \frac{1}{2}A(y)p^2 + u(y,q).$$

Then the canonical change with the generating function $\hat{y}x + \sqrt{\varepsilon}W(\hat{y},I,q)$

$$y=\hat{y}, \quad \hat{x} = x + \sqrt{\varepsilon}W'_{\hat{y}}, \quad p = W'_q, \quad \varphi = W'_I$$

transforms the symplectic structure to $d\hat{y}\wedge d\hat{x} + \sqrt{\varepsilon}dI\wedge d\varphi$ and Hamiltonian (5.3) to

$$\Lambda(\hat{y}) + \varepsilon h(\hat{y},I) + \varepsilon\hat{V}(\hat{y},I,\hat{x},\varphi,\sqrt{\varepsilon}), \quad \langle\hat{V}\rangle_x = O(\sqrt{\varepsilon}), \qquad (5.6)$$

where the functions $h,\Lambda,\hat{V}$ are real-analytic. The function h satisfies the equation

$$h(y,I) = \mathbf{h}(\chi(Y),I), \qquad Y\in U(\Sigma),$$

where $\mathbf{h}$ is defined in (4.10). Below we skip hats for brevity.

6 Initial KAM Hamiltonian

For any set $D\subset\mathbb{R}^{n+1}$ let $U_a(D)\subset\mathbb{C}^{n+1}$ be the following neighborhood:

$$U_a(D) = \{(y+\eta, I+\zeta) : (y,I)\in D, |\eta|<\sqrt{\varepsilon}a, |\zeta|<a\}, \qquad \eta\in\mathbb{C}^n,\ \zeta\in\mathbb{C}.$$

For any function $f : D\to\mathbb{R}$ which admits a real-analytic extension to $U_a(D)$ we put

$$|f|_a = \sup_{z\in U_a(D)} |f(z)|.$$

This norm is anisotropic in y and I directions.

Let $U_b(\mathbb{T}^{n+1})$ be a complex neighborhood of $\mathbb{T}^{n+1}$

$$U_b(\mathbb{T}^{n+1}) = \{(x+\xi,\varphi+\kappa) : (x,\varphi)\in\mathbb{T}^{n+1},\ |\xi|<b, |\kappa|<b\}, \qquad \xi\in\mathbb{C}^n,\ \kappa\in\mathbb{C}.$$

For any function $f:\mathbb{T}^{n+1}\to\mathbb{R}$ which admits a real-analytic extension to $U_b(\mathbb{T}^{n+1})$ we put

$$|f|_b = \sup_{z\in U_b(\mathbb{T}^{n+1})} |f(z)|.$$

For functions, real analytic on $D\times\mathbb{T}^{n+1}$ we define $|f|_{a,b}$ as the corresponding double supremums over

$$U_{a,b}(D\times\mathbb{T}^{n+1}) = U_a(D)\times U_b(\mathbb{T}^{n+1}).$$

Consider the Hamiltonian system with the symplectic structure

$$dy \wedge dx + \sqrt{\varepsilon}\, dI \wedge d\varphi \tag{6.1}$$

and the real-analytic Hamiltonian (see (5.6))

$$H_0 = \Lambda(y) + \varepsilon h_0(y, I) + \varepsilon v_0(y, I, \varphi, \sqrt{\varepsilon}) + \varepsilon u_0(y, I, x, \varphi, \sqrt{\varepsilon}), \quad v_0(y, I, \varphi, 0) = 0,\ \langle u_0 \rangle_x = 0,$$

where $v_0 = \langle V \rangle_x$, $u_0 = V - \langle V \rangle_x$ and the points (y, I, x, φ) lie in a complex neighborhood

$$U_{a_0,b_0}(D_0 \times \mathbb{T}^{n+1}), \qquad D_0 \subset \mathbb{R}^{n+1}$$

for some $a_0, b_0 > 0$.

The above analyticity assumptions mean that there exist $\varepsilon_0, \bar{s}_0, s_0 > 0$ such that for any $0 \le \varepsilon < \varepsilon_0$

$$|\Lambda|_{a_0} \le \bar{s}_0, \quad |h_0|_{a_0} \le c_h, \quad |u_0|_{a_0,b_0} \le s_0, \quad |v_0|_{a_0,b_0} \le \sqrt{\varepsilon}\mathbf{c}. \tag{6.2}$$

Assumptions of Theorem 2 imply the following non-degeneracy conditions:

$$\underline{c}_\Lambda \le |\det \Lambda''_{yy}| \le \bar{c}_\Lambda, \quad |\Lambda''_{yy}|_{a_0} \le c_\Lambda, \quad |\Lambda''^{-1}_{yy}|_{a_0} \le c_\Lambda, \quad |h'_{0\,I}|_{a_0} \le c'_h, \quad |h'_{0\,I}|^{-1}_{a_0} \le c'_h, \tag{6.3}$$

$$|h''_{0\,II}|_{a_0} \le c'_h, \quad |h''_{0\,II}|^{-1}_{a_0} \le c'_h, \quad |\sqrt{\varepsilon} h''_{0\,Iy}|_{a_0} \le c''_h, \quad |\varepsilon h''_{0\,yy}|_{a_0} \le c''_h. \tag{6.4}$$

Let ε be sufficiently small, then we can assume that c''_h is small because $c''_h \sim \sqrt{\varepsilon}$. Below we assume that $c''_h \le c''_{h\,0}(\underline{c}_\Lambda, c_\Lambda, c'_h)$.

7 The Hamiltonian H_m

Below all functions depend smoothly on ε. For brevity we do not write ε in their arguments.

As usual KAM procedure includes a converging sequence $\mathcal{F}_0, \mathcal{F}_1, \ldots$ of coordinate changes and a converging sequence of Hamiltonians $H_0, H_1, \ldots$

$$\mathcal{F}_m : U_{a_{m+1},b_{m+1}}(D_{m+1} \times \mathbb{T}^{n+1}) \to U_{a_m,b_m}(D_m \times \mathbb{T}^{n+1}), \quad H_{m+1} = H_m \circ \mathcal{F}_m.$$

Consider an increasing sequence $\{N_j \in \mathbb{Z}\}$ (N_j is the maximal order of a resonance essential on the j-th step), a decreasing sequence $\{\lambda_j > 0\}$ ($\sqrt{\varepsilon}\,\lambda_j$ determines the width of resonance strips on the j-th step, $N_{-1} = 0$) and the function $\mathbf{j} : \mathbb{N} \to \mathbb{N}$ defined by the inequality

$$N_{\mathbf{j}(r)-1} < r \le N_{\mathbf{j}(r)} \quad \text{for all } r > 0. \tag{7.1}$$

Then $\mathbf{j}(r)$ is the number of the first step on which the resonance of order r is essential.

Consider two positive decreasing sequences a_m, b_m,

$$a_m = a_{m+1} + 6\sigma_m, \quad b_m = b_{m+1} + 3\delta_m \tag{7.2}$$

Suppose that on the m-th step we have the Hamiltonian

$$H_m = \Lambda(y) + \varepsilon h_m(y, I) + \varepsilon v_m(y, I, \varphi) + \varepsilon u_m(y, I, x, \varphi), \qquad \langle u_m \rangle_x = 0. \tag{7.3}$$

The function H_m is defined in a complex neighborhood $U_{a_m,b_m}(D_m \times \mathbb{T}^{n+1})$

$$D_{m+1} = D_m \setminus \bigcup_{|K|\le N_m,\, k\ne 0} U_{a_m}(Q_{K,m}), \qquad K = \begin{pmatrix} k \\ k_0 \end{pmatrix} \in \mathbb{Z}^{n+1}, \quad k \in \mathbb{Z}^n, \tag{7.4}$$

where the resonant strips $Q_{K,m}$ are defined with the help of the sequences N_j and λ_j:

$$Q_{K,m} = \big\{(y,I) \in U_{a_m}(D_m) : \big|\langle \nu_m(y,I), K\rangle\big| \le \lambda_{\mathbf{j}(|K|)}(1+2^{-m-1})\sqrt{\varepsilon}\big\}, \tag{7.5}$$

$$\nu_m(y,I) = \begin{pmatrix} \Lambda'_y(y) + \varepsilon h'_{m\,y}(y,I) \\ \sqrt{\varepsilon}\, h'_{m\,I}(y,I) \end{pmatrix}. \tag{7.6}$$

Remark 7.1 *Further we show that for any $K \in \mathbb{Z}^{n+1}$, $|K| \le N_{m-1}$*

$$Q_{K,m} = \oslash.$$

Equation (7.4) which defines the domains D_m can be represented as

$$D_{m+1} = D_m \setminus \cup_{N_{m-1}<|K|\le N_m} U_{a_m}(Q_{K,m}).$$

Proposition 7.1 *For any $m = 1, 2, \dots$*

$$\mu(D_0 \setminus D_m) \le c_\mu \sqrt{\varepsilon},$$

where $c_\mu > 0$ is independent of ε.

Inductive assumptions. For $(y,I) \in U_{a_m}(D_m)$ the following estimates hold:

$$|v_m|_{a_m,b_m} \le s_m, \quad |u_m|_{a_m,b_m} \le s_m, \tag{7.7}$$

$$|h_m|_{a_m} \le (2-2^{-m})c_h, \quad |h'_{m\,I}|^{-1}_{a_m} \le (2-2^{-m})c'_h, \quad |h''_{m\,II}|^{-1}_{a_m} \le (2-2^{-m})c'_h, \tag{7.8}$$

$$|h''_{m\,II}|_{a_m} \le (2-2^{-m})c'_h, \quad |\sqrt{\varepsilon} h''_{m\,Iy}|_{a_m} \le (2-2^{-m})c''_h, \quad |\varepsilon h''_{m\,yy}|_{a_m} \le (2-2^{-m})c''_h. \tag{7.9}$$

8 The KAM-step

For any natural N and a periodic function

$$f : \mathbb{T}^{n+1} \to \mathbb{R}, \qquad f(x,\varphi) = \sum_{K=(k,k_0)\in\mathbb{Z}^{n+1}} f^K e^{i\langle k,x\rangle + ik_0\varphi}$$

we define the cut off

$$\Pi_N f(x,\varphi) = \sum_{|K|\le N,\, k\ne 0} f^K e^{i\langle k,x\rangle + ik_0\varphi}. \tag{8.1}$$

Then by Lemma 12.1 for any real-analytic f such that $|f|_b < \infty$ and for any $\delta \in (0,b)$

$$\big|f - \langle f\rangle_x - \Pi_N f\big|_{b-\delta} \le \frac{C}{\delta}\Big(N + \frac{1}{\delta}\Big)^n e^{-N\delta}\, |f|_b. \tag{8.2}$$

By using Hamiltonian (7.3), we introduce the canonical[3] change of variables $(y, I, x, \varphi) \mapsto (\hat{y}, \hat{I}, \hat{x}, \hat{\varphi})$, determined by the generating function $\hat{y}x + \sqrt{\varepsilon}\big(\hat{I}\varphi + S(\hat{y}, \hat{I}, x, \varphi)\big)$:

$$y = \hat{y} + \sqrt{\varepsilon} S'_x, \quad I = \hat{I} + S'_\varphi, \quad \hat{x} = x + \sqrt{\varepsilon} S'_{\hat{y}}, \quad \hat{\varphi} = \varphi + S'_{\hat{I}},$$

where the arguments $(\hat{y}, \hat{I})$ are supposed to lie in $U_{a_m - \sigma_m}(D_{m+1})$.

Remark 8.1 *$|\cdot|_{a_m-\sigma, b_m-\delta}$ (resp. $|\cdot|_{a_m-\sigma}$) denotes the norm in $U_{a_m-\sigma, b_m-\delta}(D_{m+1} \times \mathbb{T}^{n+1})$ (resp. $U_{a_m-\sigma}(D_{m+1})$).*

By definition the function S is a solution of the homologic equation

$$\Big\langle \nu_m(\hat{y}, \hat{I}), \begin{pmatrix} S'_x \\ S'_\varphi \end{pmatrix}(\hat{y}, \hat{I}, x, \varphi) \Big\rangle = -\sqrt{\varepsilon}\, \Pi_{N_m} V_m(\hat{y}, \hat{I}, x, \varphi), \qquad V_m = v_m + u_m. \tag{8.3}$$

Proposition 8.1 *For any $m = 0, 1, \dots$ there exists a solution of (8.3) where*

$$|S|_{a_m, b_m - \delta_m} \le L_m s_m, \qquad L_m = \sum_{j=0}^{m} 2 \frac{N_j^{n+1}}{\lambda_j} e^{-(n+1)\delta_m N_{j-1}}. \tag{8.4}$$

The Hamiltonian (7.3) takes the form

$$\widetilde{H}_m = \Lambda(\hat{y}) + \varepsilon h_m(\hat{y}, \hat{I}) + \varepsilon v_m(\hat{y}, \hat{I}, \hat{\varphi}) + \varepsilon \tilde{v}_m(\hat{y}, \hat{I}, \hat{\varphi}) + \varepsilon \tilde{u}_m(\hat{y}, \hat{I}, \hat{x}, \hat{\varphi}), \qquad \langle \tilde{u}_m \rangle_x = 0. \tag{8.5}$$

Remark 8.2 *Below we show that that*

$$|S'_x|_{a_m - \sigma_m, b_m - 2\delta_m} \le \sigma_m, \quad |S'_\varphi|_{a_m - \sigma_m, b_m - 2\delta_m} \le \sigma_m, \tag{8.6}$$

$$|\sqrt{\varepsilon} S'_y|_{a_m - \sigma_m, b_m - 2\delta_m} \le \delta_m, \quad |S'_I|_{a_m - \sigma_m, b_m - 2\delta_m} \le \delta_m. \tag{8.7}$$

Estimates (8.6)–(8.7) imply that the coordinate change is well-defined for $(\hat{y}, \hat{I}, \hat{x}, \hat{\varphi}) \in U_{a_m - \sigma_m, b_m - 2\delta_m}(D_{m+1} \times \mathbb{T}^{n+1})$.

Proposition 8.2 *For $m \ge 1$ estimates (8.6)–(8.7) imply the inequalities*

$$|\tilde{v}_m + \tilde{u}_m|_{a_m - \sigma_m, b_m - 2\delta_m} \le \tilde{s}_m, \tag{8.8}$$

$$\tilde{s}_m = (c_\Lambda + c'_h)\Big(\frac{L_m s_m}{\delta_m}\Big)^2 + (n+2)\frac{L_m s_m^2}{\sigma_m \delta_m} + \frac{C}{\delta_m}\Big(N_m + \frac{1}{\delta_m}\Big)^n e^{-N_m \delta_m} s_m. \tag{8.9}$$

9 An additional step

Consider the symplectic transformation $(\hat{I}, \hat{\varphi}) \mapsto (\bar{I}, \bar{\varphi})$ with generating function $\bar{I}\hat{\varphi} + \sqrt{\varepsilon}\widetilde{S}(\bar{y}, \bar{I}, \hat{\varphi})$ which introduces action-angle variables in the system with one degree of freedom and Hamiltonian

$$h_m(\hat{y}, \hat{I}) + v_m(\hat{y}, \hat{I}, \hat{\varphi}) = h_{m+1}(\bar{y}, \bar{I}). \tag{9.1}$$

[3] i.e., preserving symplectic structure (6.1)

The variables $\hat{y} = \bar{y}$ are regarded as parameters. We extend this map to a canonical transformation of the whole phase space:

$$\hat{y} = \bar{y}, \quad \hat{I} = \bar{I} + \widetilde{S}'_{\hat{\varphi}}, \quad \bar{x} = \hat{x} + \sqrt{\varepsilon}\widetilde{S}'_{\bar{y}}, \quad \bar{\varphi} = \hat{\varphi} + \widetilde{S}'_{\bar{I}}. \tag{9.2}$$

Then Hamiltonian (8.5) takes the form

$$H_{m+1} = \Lambda(\bar{y}) + \varepsilon h_{m+1}(\bar{y}, \bar{I}) + \varepsilon v_{m+1}(\bar{y}, \bar{I}, \bar{\varphi}) + \varepsilon u_{m+1}(\bar{y}, \bar{I}, \bar{x}, \bar{\varphi}), \qquad \langle u_{m+1}\rangle_{\bar{x}} = 0.$$

Proposition 9.1 *Suppose that*

$$|v_0|_{a_0} \le \sqrt{\varepsilon}\mathbf{c} \le \frac{\sigma_0\delta_0}{2c'_h}, \quad |v_m|_{a_m-\sigma_m, b_m-2\delta_m} \le s_m, \quad s_m \le \frac{\sigma_m\delta_m}{4c'_h}, \qquad m \ge 1. \tag{9.3}$$

Then for any $m \ge 0$

$$|\tilde{S}'_{\hat{\varphi}}|_{a_m-4\sigma_m, b_m-2\delta_m} \le 8c'_h s'_m, \quad |\tilde{S}|_{a_m-4\sigma_m, b_m-2\delta_m} \le 16\pi c'_h s'_m, \tag{9.4}$$

$$|h_{m+1} - h_m|_{a_m-4\sigma_m, b_m-2\delta_m} \le \frac{8c_h c'_h s'_m}{\sigma_m}, \quad |u_{m+1} + v_{m+1}|_{a_{m+1}, b_{m+1}} \le s'_m, \tag{9.5}$$

where $s'_0 = \sqrt{\varepsilon}\mathbf{c}$ *and* $s'_m = s_m$ *for all* $m \ge 1$.

Remark 9.1 *Below we show that*

$$|\widetilde{S}'_{\bar{I}}|_{a_{m+1}, b_{m+1}} \le \delta_m, \quad |\sqrt{\varepsilon}\widetilde{S}'_{\bar{y}}|_{a_{m+1}, b_{m+1}} \le \delta_m, \tag{9.6}$$

$$|\widetilde{S}'_{\hat{\varphi}}|_{a_{m+1}, b_{m+1}} \le \sigma_m. \tag{9.7}$$

Estimates (9.6),(9.7) imply that the coordinate change (9.2) is well-defined for $(\bar{y}, \bar{I}, \bar{x}, \bar{\varphi}) \in U_{a_{m+1}, b_{m+1}}(D_{m+1} \times \mathbb{T}^{n+1})$.

10 The sequences $\sigma_m, \delta_m, s_m, L_m, N_m, \lambda_m$

We define σ_m and δ_m by (7.2) and put

$$\sigma_m = \frac{a'_0}{6}\, 2^{-(2n+5)(m+1)}, \quad \delta_m = \frac{b_0}{3}\, 2^{-(m+1)}, \quad a'_0 = \frac{a_0}{2^{2n+5} - 1} \tag{10.1}$$

$$s_m \;=\; s_0\, e^{-c_s m - 2^m}, \quad N_m = c_N\, 2^{2m}, \quad \lambda_m \;=\; c_\lambda\, 2^{-(2n+2+\tau)m}, \tag{10.2}$$

where $\tau \in (0, 1)$. The constants a_0, b_0, s_0 are a priori fixed. We can choose only c_N, c_s, c_λ and ε. First we fix c_s, then we define $c_N = c_N(c_s)$ and $c_\lambda = c_\lambda(c_N, c_s)$. Below we explain how to do this.

Proposition 10.1 *Suppose that the sequences* $\delta_m, N_m,$ *and* λ_m *are defined by (10.1) and (10.2). Then the sequence* L_m*, defined by (8.4) satisfies the estimate*

$$L_m \le \frac{c_N^{n+1}}{c_\lambda} 2^{(4n+5)(m+2)}. \tag{10.3}$$

To show that our choice of the sequences $a_m, b_m, \sigma_m, \delta_m, s_m, L_m, N_m, \lambda_m$ makes the procedure converging, we have to check that assumptions (7.7)–(7.9), (8.6), (8.7) and (9.6), (9.7) hold. The remaining part of this section contains this check.

10.1 Several estimates

By using (10.1)–(10.3) we obtain:

$$
\begin{aligned}
\left(\frac{L_m s_m}{\delta_m}\right)^2 &\leq 9\frac{4^{8n+11}c_N^{2(n+1)}2^{(8n+12)m}s_0}{b_0^2c_\lambda^2}s_0e^{-2c_sm-2^{m+1}} \leq c_{Ls}2^{(8n+12)m}e^{-c_s(m-1)}s_{m+1},\\
\frac{L_m s_m^2}{\delta_m\sigma_m} &\leq 18\frac{2^{10n+16}c_N^{n+1}2^{(6n+11)m}s_0}{b_0a_0'c_\lambda}s_0e^{-2c_sm-2^{m+1}} \leq c_{Ls}2^{(8n+12)m}e^{-c_s(m-1)}s_{m+1},
\end{aligned}
$$

where $c_{Ls} = \max\left(9\frac{4^{8n+11}c_N^{2(n+1)}s_0}{b_0^2c_\lambda^2}, 18\frac{2^{10n+16}c_N^{n+1}s_0}{b_0a_0'c_\lambda}\right)$. For the third term of (8.9) we have

$$
\begin{aligned}
\frac{C}{\delta_m}\left(N_m+\frac{1}{\delta_m}\right)^n e^{-N_m\delta_m}\,s_m &\leq \frac{6C}{b_0}\left(c_N+\frac{6}{b_0}\right)^n 2^{(2n+1)m}e^{-b_0c_N2^m/6}s_0e^{-c_sm-2^m}\\
&\leq \frac{6C}{b_0}\left(c_N+\frac{6}{b_0}\right)^n e^{(2n+1)m+c_s-(b_0c_N/6-2)2^m-2^m}s_{m+1}.
\end{aligned}
$$

10.2 Inequalities (7.7)

Rewrite (8.9) in terms of s_{m+1}

$$
\begin{aligned}
\tilde{s}_m \leq\ & (c_\Lambda+c_h')c_{Ls}2^{(8n+12)m}e^{-c_s(m-1)}s_{m+1}+(n+2)c_{Ls}2^{(8n+12)m}e^{-c_s(m-1)}s_{m+1}\\
&+\frac{6C}{b_0}\left(c_N+\frac{6}{b_0}\right)^n e^{(2n+1)m+c_s-(b_0c_N/6-2)2^m-2^m}s_{m+1}.
\end{aligned}
$$

If $c_s > c_{s_0} = 16n+24$ and $c_\lambda > c_{\lambda\,0}(a_0, b_0, s_0, c_s, c_N, c_\Lambda)$, then for all $m \geq 0$ we have

$$
c_{Ls}2^{(8n+12)m}e^{-c_s(m-1)} \leq \frac{1}{3}\max((c_\Lambda+c_h')^{-1}, (n+2)^{-1}).
$$

For sufficently large $c_{N\,0}$ for all $c_N > c_{N\,0}(b_0, c_s)$ we obtain

$$
\frac{6C}{b_0}\left(c_N+\frac{6}{b_0}\right)^n e^{(2n+1)m+c_s-(b_0c_N/6-2)2^m-2^m} \leq \frac{1}{3}.
$$

This implies $\tilde{s}_m \leq s_{m+1}$. From Proposition 9.1 and the last inequality follow estimates (7.7).

10.3 Inequalities (7.8),(7.9)

By (9.5) $h_1 - h_0 = O(\sqrt{\varepsilon})$. Hence it is sufficient to check (7.8) and (7.9) only for $m \geq 1$. For large $c_s > c_{s\,1} = c_{s\,1}(c_h', a_0, s_0)$:

$$
\begin{aligned}
|h_{m+1}-h_m|_{a_{m+1}+2\sigma_m} &\leq \frac{8c_hc_h's_m}{\sigma_m} \leq 48c_h\frac{2^{2n+5}c_h's_0}{a_0'}e^{-(c_s-2n-5)m-2^m} \leq c_h2^{-m-1},\\
|h''_{m+1\,II}-h''_{m\,II}|_{a_{m+1}} &\leq \frac{8c_hc_h's_m}{\sigma_m^3} \leq 1728\frac{2^{6n+15}c_hc_h's_0}{a_0'^2}e^{-(c_s-6n-15)m-2^m}\\
&\leq c_h'2^{-m-1}.
\end{aligned}
$$

Note, that for $c_s > c_{s\,2} = c_{s\,2}(c_h, c'_h, a_0, s_0)$ we have

$$\begin{aligned}
|h'_{m+1\,I}|_{a_{m+1}+\sigma_m} &\geq |h'_{m\,I}|_{a_{m+1}+\sigma_m} - |h'_{m+1\,I} - h'_{m\,I}|_{a_{m+1}+\sigma_m} \\
&\geq \frac{1}{(2-2^{-m})c'_h} - \frac{8c_h c'_h s_m}{\sigma_m^2} \geq \frac{1}{(2-2^{-m})c'_h} - 288\frac{2^{4n+10}c_h c'_h s_0}{a_0'^2}e^{-(c_s-4n-10)m-2^m} \\
&\geq \frac{1}{(2-2^{-m})c'_h} - \frac{2^m}{(2^{m+1}-1)(2^{m+2}-1)c'_h} = \frac{1}{(2-2^{-m-1})c'_h}.
\end{aligned}$$

Consider the first inequality (7.9). For $c_s > c_{s\,3} = c_{s\,3}(c_h, c'_h, a_0, s_0)$

$$\begin{aligned}
|h''_{m+1\,II}|_{a_{m+1}} &\geq |h''_{m\,II}|_{a_{m+1}+\sigma_m} - |h''_{m+1\,II} - h''_{m\,II}|_{a_{m+1}} \\
&\geq \frac{1}{(2-2^{-m})c'_h} - \frac{8c_h c'_h s_m}{\sigma_m^3} \geq \frac{1}{(2-2^{-m})c'_h} - 1728\frac{2^{6n+15}c_h c'_h s_0}{a_0'^2}e^{-(c_s-6n-15)m-2^m} \\
&\geq \frac{1}{(2-2^{-m})c'_h} - \frac{2^m}{(2^{m+1}-1)(2^{m+2}-1)c'_h} = \frac{1}{(2-2^{-m-1})c'_h}.
\end{aligned}$$

For the last two inequalities (7.9) let $c_s > c_{s\,4} = c_{s\,4}(c_h, c'_h, c''_h, a_0)$. Then

$$\begin{aligned}
|\sqrt{\varepsilon}h''_{m+1\,Iy} - \sqrt{\varepsilon}h''_{m\,Iy}|_{a_{m+1}} &\leq \sqrt{\varepsilon}\frac{8nc_h c'_h s_m}{\sqrt{\varepsilon}\sigma_m^3} \leq 1728\frac{2^{6n+15}nc_h c'_h s_0}{a_0'^2}e^{-(c_s-6n-15)m-2^m} \leq c''_h 2^{-m-1}, \\
|\varepsilon h''_{m+1\,yy} - \varepsilon h''_{m\,yy}|_{a_{m+1}} &\leq \varepsilon\frac{8nc_h c'_h s_m}{\varepsilon\sigma_m^3} \leq c''_h 2^{-m-1}.
\end{aligned}$$

For sufficently large $c_s > \max(c_{s\,1}, c_{s\,2}, c_{s\,3}, c_{s\,4})$ the exponents $e^{-(c_s-2n-5)m-2^m}$, $e^{-(c_s-4n-10)m-2^m}$, $e^{-(c_s-6n-15)m-2^m}$ are small and all inequalities (7.8),(7.9) hold.

10.4 Inequalities (8.6),(8.7)

Note, that for $c_\lambda > c_{\lambda 1}(c_N, c_s, b_0, a_0, s_0)$:

$$\begin{aligned}
\frac{L_m s_m}{\delta_m} &\leq 3\frac{2^{8n+11}c_N^{n+1}2^{(4n+6)m}s_0 e^{-c_s m-2^m}}{c_\lambda b_0} \leq \frac{1}{6}a'_0 2^{-(2n+5)(m+1)} = \sigma_m, \\
\frac{L_m s_m}{\sigma_m} &\leq 6\frac{2^{10n+15}c_N^{n+1}2^{(6n+10)m}s_0 e^{-c_s m-2^m}}{c_\lambda a'_0} \leq \frac{1}{3}b_0 2^{-(m+1)} = \delta_m.
\end{aligned}$$

This implies (8.6) and (8.7).

10.5 Inequalities (9.3), (9.6),(9.7)

The first inequality in (9.3) holds for small ε. Note, that for $c_s > c_{s\,5}(s_0, c'_h, a_0, b_0)$ we obtain

$$s_m = s_0 e^{-c_s m-2^m} \leq \frac{\sigma_m\delta_m}{4c'_h} = \frac{a'_0 b_0}{72c'_h}2^{-(2n+6)(m+1)}, \quad m \geq 1.$$

By Proposition 9.1:

$$|\widetilde{S}|_{a_m-4\sigma_m, b_m-2\delta_m} \leq 16\pi c'_h s'_m, \quad |\widetilde{S}'_{\hat{\varphi}}|_{a_m-4\sigma_m, b_m-2\delta_m} \leq 8c'_h s'_m.$$

For $m=0$ inequalities (9.3), (9.6), (9.7) hold if ε is sufficently small. Consider $m\ge 1$. For $c_s > c_{s\,6}(s_0, c'_h, a_0, b_0)$

$$
\begin{aligned}
|\widetilde{S}'_{\hat I}|_{a_m-5\sigma_m, b_m-2\delta_m} &\le 16\pi c'_h \frac{s_m}{\sigma_m} \le 96\pi c'_h \frac{s_0 2^{2n+5}}{a'_0} e^{-(c_s-2n-5)m-2^m} \le \frac{b_0}{3} 2^{-(m+1)} = \delta_m,\\
|\sqrt{\varepsilon}\widetilde{S}'_{\hat y}|_{a_m-5\sigma_m, b_m-2\delta_m} &\le \sqrt{\varepsilon}16\pi c'_h \frac{s_m}{\sqrt{\varepsilon}\sigma_m} \le \delta_m,\\
|\widetilde{S}'_{\hat\varphi}|_{a_m-4\sigma_m, b_m-2\delta_m} &\le 8c'_h s_m \le 8c'_h s_0 e^{-c_s m-2^m} \le \frac{a'_0}{6} 2^{-(2n+5)(m+1)} = \sigma_m.
\end{aligned}
$$

We choose the constants in the following way. Fix $c_s > \max(c_{s\,1}, \ldots c_{s\,6})$, $c_N > c_N(c_s)$ and $c_\lambda = c_\lambda(c_N, c_s) > \max(c_{\lambda\,0}, c_{\lambda\,1})$, we obtain for $m \ge 0$ inequalities (7.7), (8.6), (8.7) and for $m\ge 1$ we obtain (7.8), (7.9), (9.3),(9.6),(9.7). Finally, for sufficently small ε inequalities (7.8), (7.9), (9.3),(9.6),(9.7) hold for $m=0$.

11 Proofs

11.1 Diophantine conditions (4.12)

Using (10.2) we obtain

$$
|\langle \nu, K\rangle| > \sqrt{\varepsilon}\lambda_{j(K)} = c_\lambda 2^{-(2n+2+\tau)j(K)} > \frac{c_\lambda c_N^{n+1+\tau}}{N_{j(K)}^{n+1+\tau}} = \frac{C_\tau}{|K|^{n+1+\tau}}, \quad C_\tau = c_\lambda c_N^{n+1+\tau}.
$$

11.2 Proof of Proposition 7.1

Proposition 11.1 *For any $K\in\mathbb{Z}^{n+1}$, $|K|\le N_m$*

$$
Q_{K,m+1} = \oslash.
$$

Consider the scaled frequency map

$$
\tilde\nu_m : (y, I) \mapsto \Big(\Lambda'_y(y) + \varepsilon h'_{m\,y}(y, I), h'_{m\,I}(y, I)\Big).
$$

In comparison with (7.6) we remove the multiplier $\sqrt{\varepsilon}$ at $h'_{m\,I}$. It's Jacobi matrix equals

$$
J_m(y, I) = \begin{pmatrix} \Lambda''_{yy} + \varepsilon h''_{m\,yy} & h''_{m\,Iy} \\ \varepsilon h''_{m\,yI} & h''_{m\,II} \end{pmatrix}. \tag{11.1}
$$

Proposition 11.2 *For some positive constants $\underline{C}_J$ and $\overline{C}_J$*

$$
|\det J_m(y, I)|_{a_m} \le \overline{C}_J, \quad |\det J_m(y, I)|^{-1}_{a_m} \le \underline{C}_J. \tag{11.2}
$$

Estimates (11.2) imply the following inequality for measure of the domain $\tilde\nu_m(D_m)$:

$$
\mu\Big(\tilde\nu_m(D_m)\Big) < \overline{C}_J\,\mu(D_m).
$$

Consider the vector $(\omega_y, \sqrt{\varepsilon}\omega_I)$, $\quad \omega_y \in \mathbb{R}^n$, $\quad \omega_I \in \mathbb{R}$ and set

$$Q^{\omega}_{K,m} = \Big\{(\omega_y, \omega_I) \in \tilde{\nu}_m(D_m) : \big|\langle \omega_y, k\rangle + \sqrt{\varepsilon}\omega_I k_0\big| \le \lambda_{\mathbf{j}(|K|)}(1 + 2^{-m-1})\sqrt{\varepsilon}\Big\}. \quad (11.3)$$

The set $Q^{\omega}_{K,m} \subset \mathbb{R}^{n+1}$ is a strip between two planes

$$\frac{\langle \omega_y, k\rangle}{\sqrt{\langle k,k\rangle + \varepsilon k_0^2}} + \frac{\sqrt{\varepsilon}\omega_I k_0}{\sqrt{\langle k,k\rangle + \varepsilon k_0^2}} = \pm\frac{\sqrt{\varepsilon}\lambda_{\mathbf{j}(|K|)}(1 + 2^{-m-1})}{\sqrt{\langle k,k\rangle + \varepsilon k_0^2}}.$$

Using (8.1) we have, that $\langle k,k\rangle \ge 1$ and the the distance between the planes is not more than $4\lambda_{\mathbf{j}(|K|)}\sqrt{\varepsilon}$. The measure estimates are

$$\begin{aligned}
\mu\Big(Q^{\omega}_{K,m}\Big) &\le 4\lambda_{\mathbf{j}(|K|)}\sqrt{\varepsilon}\,\overline{C}_J\, C_D,\\
\mu\Big(Q_{K,m} \cap D_m\Big) &\le \mu\Big(\tilde{\nu}_m^{-1}(Q^{\omega}_{K,m})\Big) \le 4\lambda_{\mathbf{j}(|K|)}\sqrt{\varepsilon}\,\underline{C}_J\,\overline{C}_J\, C_D,
\end{aligned}$$

where C_D depends on diameter and dimension of D_0.

Consider estimates for the measure of $D_m \cap U_{a_m}(Q_{K,m})$. Let $|y'| \le \sqrt{\varepsilon}\sigma_m$, $|I'| \le \sigma_m$. Then

$$\begin{aligned}
|\langle \nu_m(y+y', I+I') - \nu_m(y,I), K\rangle| \le\; & n|\Lambda''_{yy}||K|\sqrt{\varepsilon}\sigma_m + n|\varepsilon h''_{yy}||K|\sqrt{\varepsilon}\sigma_m\\
& + n|\sqrt{\varepsilon}h''_{yI}||K|\sqrt{\varepsilon}\sigma_m + n|\sqrt{\varepsilon}h''_{yI}||K|\sqrt{\varepsilon}\sigma_m\\
& + n|\sqrt{\varepsilon}h''_{yI}||K|\sqrt{\varepsilon}\sigma_m \le \sqrt{\varepsilon}C_\psi N_m\sigma_m,
\end{aligned}$$

where $C_\psi = C_\psi(c_\Lambda, c'_h, c''_h, n)$.

Consider the extension of $Q^{\omega}_{K,m}$

$$Q^{\omega+}_{K,m} = \Big\{(\omega_y, \omega_I) \in \tilde{\nu}_m(D_m) : \big|\langle \omega_y, k\rangle + \sqrt{\varepsilon}\omega_I k_0\big| \le (2\lambda_{\mathbf{j}(|K|)} + C_\psi N_m\sigma_m)\sqrt{\varepsilon}\Big\}. (11.4)$$

Note that $\Big(D_m \cap U_{a_m}(Q_{K,m})\Big) \subset \tilde{\nu}_m^{-1}(Q^{\omega+}_{K,m})$. Finally

$$\mu\Big(D_m \cap U_{a_m}(Q_{K,m}\Big) \le \mu(\tilde{\nu}_m^{-1}(Q^{\omega+}_{K,m})) \le 4(\lambda_{\mathbf{j}(|K|)} + C_\psi N_m\sigma_m)\sqrt{\varepsilon}\,\underline{C}_J\,\overline{C}_J C_D. \quad (11.5)$$

We have $(D_m \setminus D_{m+1}) \subset \cup_{N_{m-1}<|K|\le N_m}\Big(D_m \cap U_{a_m}(Q_{K,m})\Big)$. Using (11.5) we obtain

$$\begin{aligned}
\mu\Big(D_0 \setminus D_{m+1}\Big) &\le \sum_{i=0}^{m}\mu\Big(\cup_{N_{i-1}<|K|\le N_i}\Big(D_i \cap U_{a_i}(Q_{K,i})\Big)\Big)\\
&\le 4\sqrt{\varepsilon}\,\underline{C}_J\,\overline{C}_J C_D\sum_{i=0}^{m} N_i^{n+1}(\lambda_i + C_\psi N_i\sigma_i).
\end{aligned}$$

The proposition holds for $c_\mu = 4\underline{C}_J\,\overline{C}_J C_D\sum_{i=0}^{m} N_i^{n+1}(\lambda_i + C_\psi N_i\sigma_i)$. To finish the proof we need to check

$$\sum_{i=0}^{+\infty} N_i^{n+1}(\lambda_i + C_\psi N_i\sigma_i) < +\infty. \quad (11.6)$$

Using (10.1) and (10.2) we obtain

$$\sum_{i=0}^{+\infty} N_i^{n+1}(\lambda_i + C_\psi N_i\sigma_i) = \sum_{i=0}^{+\infty}\Big(c_N^{n+1}c_\lambda 2^{-\tau i} + \frac{1}{6}a_0 C_\psi c_N^{n+2}2^{-i}\Big) < +\infty.$$

∎

11.3 Proof of Proposition 8.1

Solution of equation (8.3) has the form

$$S = \sum_{|K| \le N_m,\, k \ne 0} S^K \, e^{i\langle k,x\rangle + ik_0\varphi}, \quad S^K = \frac{i\sqrt{\varepsilon}\, V_m^K(\hat{y}, \hat{I})}{\langle \nu_m(\hat{y}, \hat{I}), K\rangle}, \tag{11.7}$$

where $\nu_m(\hat{y}, \hat{I})$ is determined by (7.6).

The first inequality (7.7) means that

$$|V_m^K e^{i\langle k,x\rangle + ik_0\varphi}|_{a_m, b_m - \delta} \le 2s_m e^{-(|k_1| + |k_2| + \ldots + |k_0|)\delta}, \qquad 0 \le \delta \le b_m.$$

Then

$$|S|_{a_m, b_m - \delta_m} \le \sum_{j=0}^{m} \sum_{N_{j-1} < |K| \le N_j,\, k \ne 0} \frac{2s_m}{\lambda_j} \, e^{-\delta_m (n+1) N_{j-1}} \le L_m s_m.$$

■

11.4 Proof of Proposition 8.2

In this section for brevity we write V, h instead of V_m, h_m and $a, b, \sigma, \delta, N, L, s, \tilde{s}$ instead of a_m, b_m, σ_m, δ_m, N_m, L_m, s_m, $\tilde{s}_m$.

The function $\widetilde{V}_m = \tilde{v}_m + \tilde{u}_m$ can be presented in the form

$$\widetilde{V}_m(\hat{y}, \hat{I}, \hat{x}, \hat{\varphi}) = R_1 + R_2 + R_3 + R_4 + R_5, \tag{11.8}$$

$$\begin{aligned}
R_1 &= \frac{1}{\varepsilon}\Big(\Lambda(y) - \Lambda(\hat{y}) - \langle \Lambda'_y(\hat{y}), \sqrt{\varepsilon} S'_x\rangle\Big), \\
R_2 &= h(y, I) - h(\hat{y}, \hat{I}) - \big\langle h'_y(\hat{y}, \hat{I}), \sqrt{\varepsilon} S'_x\big\rangle - h'_I(\hat{y}, \hat{I})\, S'_\varphi, \\
R_3 &= V(y, I, x, \varphi) - V(\hat{y}, \hat{I}, x, \varphi), \\
R_4 &= V(\hat{y}, \hat{I}, x, \varphi) - \langle V\rangle_x(\hat{y}, \hat{I}, \varphi) - \Pi_N V(\hat{y}, \hat{I}, x, \varphi), \\
R_5 &= \langle V\rangle_x(\hat{y}, \hat{I}, \varphi) - \langle V\rangle_x(\hat{y}, \hat{I}, \hat{\varphi}).
\end{aligned}$$

By Proposition (8.1) the first term in (11.8) satisfies the estimate

$$|R_1|_{a-\sigma, b-2\delta} \le \frac{1}{2}|\Lambda''_{yy}|_a \, |S'_x|^2_{a-\sigma, b-2\delta} \le \frac{c_\Lambda}{2}\left(\frac{Ls}{\delta}\right)^2.$$

To estimate the second one we use (7.8),(7.9):

$$\begin{aligned}
|R_2|_{a-\sigma, b-2\delta} &\le \frac{1}{2}\Big(|\varepsilon h''_{yy}|_a \, |S'_x|^2_{a-\sigma, b-2\delta} + 2n|\sqrt{\varepsilon} h''_{yI}|_a \, |S'_x S'_\varphi|_{a-\sigma, b-2\delta} + |h''_{II}|_a \, |S'_\varphi|^2_{a-\sigma, b-2\delta}\Big) \\
&\le c''_h n \left(\frac{Ls}{\delta}\right)^2 + 2nc''_h \left(\frac{Ls}{\delta}\right)^2 + c'_h \left(\frac{Ls}{\delta}\right)^2 = (c'_h + 3nc''_h)\left(\frac{Ls}{\delta}\right)^2
\end{aligned}$$

The third term is estimated by (7.7):

$$\begin{aligned}
|R_3|_{a-\sigma, b-2\delta} &\le \sqrt{\varepsilon} n |V'_y|_{a-\sigma, b-\delta} |S'_x|_{a-\sigma, b-2\delta} + |V'_I|_{a-\sigma, b-\delta} |S'_\varphi|_{a-\sigma, b-2\delta} \\
&\le (n+1)\, \frac{Ls^2}{\sigma\delta}.
\end{aligned}$$

By (8.2)

$$|R_4|_{a-\sigma,b-2\delta} \le \frac{C}{\delta}\Big(N + \frac{1}{\delta}\Big)^n e^{-N\delta}\, s.$$

Finally

$$|R_5|_{a-\sigma,b-2\delta} \le |V'_\varphi|_{a-\sigma,b-\delta}|S'_I|_{a-\sigma,b-\delta} \le \frac{Ls^2}{\sigma\delta}.$$

Note, that $3nc''_h \le \frac{c_\Lambda}{2}$. Therefore

$$|\tilde u_m + \tilde v_m|_{a-\sigma,b-2\delta} \le \tilde s,$$

where

$$\tilde s = (c_\Lambda + c'_h)\Big(\frac{Ls}{\delta}\Big)^2 + (n+2)\frac{Ls^2}{\sigma\delta} + \frac{C}{\delta}\Big(N + \frac{1}{\delta}\Big)^n e^{-N\delta}\, s.$$

■

11.5 Proof of Proposition 9.1

By (6.2), (6.3), (7.7), (7.8), and (8.8), for any $m \ge 1$ we have:

$$|h_0|_{a_0} \le c_h, \quad |h'_{0I}|^{-1}_{a_0} \le c'_h, \quad |v_0|_{a_0,b_0} \le \sqrt{\varepsilon}\mathbf{c}, \quad |h_m|_{a_m} \le 2c_h, \quad |h'_{mI}|^{-1}_{a_m} \le 2c'_h, \tag{11.9}$$

$$|v_m|_{a_m-\sigma_m,b_m-2\delta_m} \le s_m, \quad |\tilde v_m + \tilde u_m|_{a_m-\sigma_m,b_m-2\delta_m} \le \tilde s_m. \tag{11.10}$$

Applying Lemma 12.2 to equation (9.1), we obtain:

$$|\widetilde S'_\varphi|_{a_m-4\sigma_m,b_m-2\delta_m} \le 8c'_h|v_m|_{a_m-\sigma_m,b_m-2\delta_m}, \quad |\widetilde S|_{a_m-4\sigma_m,b_m-2\delta_m} \le 16\pi c'_h|v_m|_{a_m-\sigma_m,b_m-2\delta_m}, \tag{11.11}$$

$$\widetilde S'_x = 0, \quad |h_m - h_{m+1}|_{a_m-4\sigma_m,b_m-2\delta_m} \le 8c_h c'_h \frac{|v_m|_{a_m-\sigma_m,b_m-2\delta_m}}{\sigma_m}. \tag{11.12}$$

Then estimates (9.5) follow from (11.9), (11.10), and (11.12).

We have:

$$v_{m+1}(\bar y, \bar I, \bar\varphi) + u_{m+1}(\bar y, \bar I, \bar x, \bar\varphi) = \tilde v_m(\bar y, \bar I + \widetilde S'_{\hat\varphi}, \bar\varphi - \widetilde S'_{\hat I}) + \tilde u_m\big(\bar y, \bar I + \widetilde S'_{\hat\varphi}, \bar x - \sqrt{\varepsilon}\widetilde S'_{\bar y}, \bar\varphi - \widetilde S'_{\hat I}\big).$$

Since by (11.11)

$$\begin{aligned}
|\widetilde S'_{\hat\varphi}|_{a_{m+1},b_{m+1}+2\delta_m} &\le 8c'_h s_m \le \sigma_m,\\
|\sqrt{\varepsilon}\widetilde S'_{\bar y}|_{a_{m+1},b_{m+1}+\delta_m} &\le \sqrt{\varepsilon}\frac{16\pi c'_h s_m}{\sigma_m} \le \delta_m,\\
|\widetilde S'_{\bar I}|_{a_{m+1},b_{m+1}+\delta_m} &\le \frac{16\pi c'_h s_m}{\sigma_m} \le \delta_m,
\end{aligned}$$

we have:

$$|v_{m+1} + u_{m+1}|_{a_{m+1},b_{m+1}} \le \tilde s_m \le s_{m+1}.$$

■

11.6 Proof of Proposition 10.1

By equations (8.4) and (10.2) we have

$$L_m \leq \sum_{j=0}^{m} \frac{2N_j^{n+1}}{\lambda_j} < \frac{2c_N^{n+1}}{c_\lambda} \sum_{j=0}^{m} 2^{(4n+5)j} < \frac{c_N^{n+1}}{c_\lambda} 2^{(4n+5)(m+2)}.$$

■

11.7 Proof of Proposition 11.1

A point $(y, I) \in Q_{K,m+1}$ if

$$\left|\left\langle \nu_{m+1}(y, I), K\right\rangle\right| \leq \lambda_{\mathbf{j}(|K|)}(1 + 2^{-m-2})\sqrt{\varepsilon}.$$

It is sufficient to show that $(y, I) \in Q_{K,m}$, i.e.

$$\left|\left\langle \nu_m(y, I), K\right\rangle\right| \leq \lambda_{\mathbf{j}(|K|)}(1 + 2^{-m-1})\sqrt{\varepsilon}. \tag{11.13}$$

We have the inequality

$$\left|\langle \nu_m(y, I), K\rangle - \langle \nu_{m+1}(y, I), K\rangle\right| \leq (n + 1)W\,|K|,$$
$$W = |\nu_m(y, I) - \nu_{m+1}(y, I)|.$$

By (9.5)

$$W \leq \varepsilon|h'_{m+1\,y} - h'_{m\,y}|_{a_{m+1},b_{m+1}} + \sqrt{\varepsilon}|h'_{m+1\,I} - h'_{m\,I}|_{a_{m+1},b_{m+1}} \leq \sqrt{\varepsilon}\frac{16c_h c'_h s'_m}{\sigma_m^2}.$$

Therefore for any $|K| \leq N_m$ we have the estimate

$$(n + 1)W|K| \leq (n + 1)\frac{16c_h c'_h s'_m}{\sigma_m^2} N_m,$$

where $s'_0 = \sqrt{\varepsilon}\mathbf{c}$ and $s'_m = s_m$ for all $m \geq 1$.

It remains to check the estimate

$$(n + 1)\frac{16c_h c'_h s'_m}{\sigma_m^2} N_m \leq \lambda_m 2^{-m-2}.$$

Let $c_\lambda > c_\lambda(c_s, c_N)$. For $m = 0$ we have

$$(n + 1)\frac{16c_h c'_h s'_0}{\sigma_0^2} N_0 \leq \lambda_0 2^{-2}.$$

and for $m \geq 1$

$$(n + 1)\frac{36c_h c'_h s_0 c_N}{a_0'^2} 2^{4n+14} 2^{(4n+12)m} e^{-c_s m - 2^m} \leq c_\lambda 2^{-(2n+2+\tau)m}.$$

■

11.8 Proof of Proposition 11.2

Suppose that the arguments of functions Λ, h lie in $U_{a_m}(D_m)$. Let us expand the Jacobian J_m with respect to the last column

$$\begin{aligned} \det J_m(y,I) &= \det \begin{pmatrix} \Lambda''_{yy} + \varepsilon h''_{m\,yy} & \sqrt{\varepsilon} h''_{m\,Iy} \\ \sqrt{\varepsilon} h''_{m\,yI} & h''_{m\,II} \end{pmatrix} = \\ &= \det\Big(\Lambda''_{yy} + \varepsilon h''_{m\,yy}\Big) h''_{m\,II} + \sum_{i=1}^{n} (-1)^{n+1+i} \sqrt{\varepsilon} h''_{m\,Iyi} M_{i,n+1}. \end{aligned}$$

Here $M_{i,n+1}$ is the $(i, n+1)$ minor matrix of $J_m(y,I)$. Using (6.3), (6.4), (7.8) and (7.9) we obtain

$$|\det J_m(y,I)| \le n!\Big(|\Lambda''_{yy}| + |\varepsilon h''_{m\,yy}|\Big)^n |h''_{m\,II}| + nn! |\sqrt{\varepsilon} h''_{m\,Iy}|^2 \Big(|\Lambda''_{yy}| + |\varepsilon h''_{m\,yy}|\Big)^{n-1} \le \overline{C}_J,$$

where $\underline{C}_J$ is some constant depending on c_λ, c'_h, c''_h and n.

Return to estimate for the Jacobian. For sufficiently small c''_h

$$c''_h \le c''_{h\,1} = \min\left(\frac{\underline{c}_\Lambda}{2^{n+1} nn! c_\Lambda^{n-1}}, \frac{c_\Lambda}{2}\right), \quad |\varepsilon h''_{m\,yy}| \le 2c''_h$$

we have

$$\left|\det\Big(\Lambda''_{yy} + \varepsilon h''_{m\,yy}\Big)\right| > |\det \Lambda''_{yy}| - nn! |\varepsilon h''_{m\,yy}| \Big(|\Lambda''_{yy}| + |\varepsilon h''_{m\,yy}|\Big)^{n-1} \ge \frac{1}{2}\underline{c}_\Lambda,$$

Note, that $|h''_{m\,II}| \ge \frac{1}{2c'_h}$ and for

$$c''_h \le c''_{h\,2} = \min\left(\frac{\underline{c}_\Lambda}{2^{n+4} nn! c_\Lambda^{n-1} c'_h}, c''_{h\,1}\right)$$

we have

$$\left|\sum_{i=1}^{n} (-1)^{n+1+i} \sqrt{\varepsilon} h''_{m\,Iyi} M_{i,n+1}\right| \le \frac{\underline{c}_\Lambda}{8c'_h}.$$

Finally

$$|\det J(y,I)|_{a_m} \ge \frac{\underline{c}_\Lambda}{8c'_h} = \underline{C}_J^{-1}.$$

■

12 Further technical statements

12.1 Lemma on a cut off

Lemma 12.1 *For any real-analytic function f on $U_b(\mathbb{T}^{m+1})$ and any $\delta \in (0,b)$*

$$\big|f - \langle f \rangle_x - \Pi_N f\big|_{b-\delta} \le \frac{C}{\delta}\Big(N + \frac{1}{\delta}\Big)^n e^{-N\delta} |f|_b.$$

where the constant C depends only on n.

Proof. The Fourier coefficients (8.1) satisfy the inequalities

$$|f^K| \le e^{-b|K|} |f|_b.$$

Then the equation

$$f - \langle f \rangle_x - \Pi_N f = \sum_{|K|>N,\, k\neq 0} f^K e^{i\langle k,x\rangle + k_0\varphi}$$

implies

$$|f - \langle f \rangle_x - \Pi_N f|_{b-\delta} \le |f|_b \sum_{|K|>N,\, k\neq 0} e^{-\delta|K|}.$$

The sum in the right-hand side does not exceed

$$c_1 \int_{x\in\mathbb{R}^{n+1},\, |x|>N} e^{-\delta|x|}\, dx \le \frac{c_2}{\delta^{n+1}} \int_{\delta N}^{\infty} s^n e^{-s}\, ds \le \frac{c_3}{\delta^{n+1}} (1+\delta N)^n e^{-\delta N},$$

where c_1, c_2, c_3 depend only on n.

12.2 Lemma on the action-angle variables

Lemma 12.2 *Let h and v be real-analytic functions, defined in complex neighborhoods of $[-\alpha,\alpha]$ and $[-\alpha,\alpha]\times\mathbb{T}$ respectively. Let the canonical change $(I, \varphi \bmod 2\pi) \mapsto (\bar I, \bar\varphi \bmod 2\pi)$, determined by the generating function $\bar I\varphi + S(\bar I, \varphi)$, $\langle S\rangle_\varphi = 0$, be such that*

$$h(I) + v(I,\varphi) = h_*(\bar I), \tag{12.1}$$

$$|h|_a \le c, \quad |h'|_a^{-1} \le c', \quad |v|_{a,b} \le \tfrac{\sigma}{2c'}, \quad 0 < \sigma < a/2 \tag{12.2}$$

Then

$$|S'_\varphi|_{a-3\sigma,b} \le 4c'|v|_{a,b}, \quad |S|_{a-3\sigma,b} \le 8\pi c'|v|_{a,b}, \quad |h - h_*|_{a-3\sigma} \le 2cc' \frac{|v|_{a,b}}{\sigma}. \tag{12.3}$$

Proof. Equation that determines $\bar I$ is well-known:

$$\bar I(r) = \frac{1}{2\pi} \int_0^{2\pi} I(r,\varphi)\, d\varphi. \tag{12.4}$$

Here $I(r,\varphi)$ is the solution of the equation

$$h(I) + v(I,\varphi) = h(r).$$

We use r as a constant which fixes the energy $h(r)$.

By Lemma 12.3 the function $I = I(r,\varphi)$ is as follows:

$$I = r + f(r,\varphi), \quad |f|_{a-2\sigma,b} \le 2c'|v|_{a,b}. \tag{12.5}$$

Moreover,

$$I \in U_{a-2\sigma-\delta}([-\alpha,\alpha]) \quad \text{implies} \quad r \in U_{a-\delta}([-\alpha,\alpha]) \quad \text{for any } \delta \in [2\sigma, a-\sigma]. \tag{12.6}$$

Equations (12.4) and (12.5) imply

$$I(r,\varphi)-\bar I(r)=f(r,\varphi)-\langle f\rangle_\varphi(r), \tag{12.7}$$
$$r=r(\bar I),\quad I-\bar I=S'_\varphi(\bar I,\varphi). \tag{12.8}$$

Combining (12.7) and (12.8), we obtain:

$$S(\bar I,\varphi)=\int_0^\varphi \big(I(r,\varphi)-\bar I\big)\,d\varphi=\int_0^\varphi \big(f(r,\varphi)-\langle f\rangle_\varphi(r)\big)\,d\varphi.$$

Therefore

$$|S'_\varphi|_{a-3\sigma,b}\le 4c'|v|_{a,b},\quad |S|_{a-3\sigma,b}\le 8\pi c'|v|_{a,b}.$$

By using the equation $h(r)=h_*(\hat I)$, we have:

$$|h(I)-h_*(I)|_{a-3\sigma}\ \le\ |h(r+f(r,\varphi))-h(r)|_{a-2\sigma}\le |h'|_{a-\sigma}|f|_{a-2\sigma,b}\le \frac{c}{\sigma}2c'|v|_{a,b}.$$

■

12.3 A version of the implicit function theorem

Lemma 12.3 *Let the real-analytic functions h,v, defined in a complex neighborhood of the interval $\mathcal{I}\subset\mathbb{R}$, satisfy the estimates*

$$|h'|_a^{-1}\le c',\quad |v|_a\le\frac{\sigma}{2c'},\qquad 0<\sigma<\frac{a}{2}. \tag{12.9}$$

Then the equation

$$h(I)+v(I)=h(r),\qquad I\in U_{a-\sigma}(\mathcal{I}) \tag{12.10}$$

implies

$$I=r+f(r),\qquad |f|_{a-2\sigma}\le 2c'|v|_a\le\sigma, \tag{12.11}$$

where $f(r)$ is the real-analytic function, $r\in U_{a-2\sigma}(\mathcal{I})$.

Proof. Applying the map h^{-1} to (12.10), we get:

$$I+u(I)=r,\qquad u(I)=h^{-1}\big(h(I)+v(I)\big)-I.$$

If $|v|_a\le\sigma/c'$, the function u is defined in a complex neighborhood of $\mathcal{I}$ and admits the estimate

$$|u|_{a-\sigma}\le|h'|_a^{-1}|v|_a\le c'|v|_a.$$

The function $I=I(r)$, defined by (12.11), is a fixed point of the operator

$$I(r)\mapsto\Phi(I(r),r)=r-u(I(r)).$$

This operator is contracting with respect to the norm $|\cdot|_{a-2\sigma}$ because by (12.9)

$$|\Phi'_I|_{a-2\sigma}=|u'_I|_{a-2\sigma}\le\frac{c'|v|_a}{\sigma}\le\frac12.$$

Therefore $I-r=f(r)$, where $|f(r)|_{a-2\sigma}\le 2|u|_{a-\sigma}\le 2c'|v|_a$. ■

Acknowledgements

The work was supported by grants RFBR 13-01-00251, 05-01-01119, 13-01-12462, State contract no. 8223 "Dynamic instability and catastrophe" (RF) and RF Program for the State Support of Leading Scientific Schools NSh-2964.2014.1.

The authors are grateful to M. Sevryuk and G. Piftankin for useful discussions and comments.

References

[1] V.I. Arnold, "Small denominators and problems of stability of motion in classical and celestial mechanics", Russian Math. Surveys, 18:6 (1963), 85–191

[2] V. I. Arnold, V. V. Kozlov, A. I. Neishtadt, *Mathematical aspects of classical and celestial mechanics*, Dynamical systems. III, Encyclopaedia Math. Sci., 3, 3rd edition, Springer-Verlag, Berlin, 2006

[3] V. I. Arnold, *Mathematical methods of classical mechanics*, Grad. Texts in Math., 60, Springer-Verlag, New York, NY, 1978

[4] F. Cong, T. Kupper, Y. Li and J. You, "KAM-type theorem on resonant surfaces for nearly integrable Hamiltonian systems", J. Nonlinear Sci. 10 (2000), no. 1, 49–68

[5] A. Delshams, R. de la Llave, T. M. Seara "A geometric mechanism for diffusion in Hamiltonian systems overcoming the large gap problem: heuristics and rigorous verification on a model", Mem. Amer. Math. Soc. 179(844) 2006

[6] Cl. Froeschlé, "Numerical Study of a Four-Dimensional Mapping", Astron. Astrophys. 16, 172–189 (1972)

[7] V. F. Lazutkin, "To Moser's theorem on invariant curves", Vopr. Dinamich. Teor. Rasprostr. Seism. Voln. 14 (1974), 109–120 (Russian).

[8] Y. Li and Y. Yi, "A quasi-periodic Poincare's theorem", Math. Ann. 326 (2003), no. 4, 649–690

[9] M. Mazzocco, "KAM theorem for generic analytic perturbations of the Euler system", Z. angew. Math. Phys. 48 (1997), 193–219

[10] A. I. Neishtadt, "Estimates in the Kolmogorov theorem on conservation of conditionally periodic motions", J. Appl. Math. Mech. 45 (1982), no. 6, 766–772

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

[12] J. Pöschel, "Integrability of Hamiltonian systems on Cantor sets", *Comm. Pure Appl. Math.*, v. 35 (1982), no. 1, 653–695

[13] A. V. Pronin, D.V. Treschev, "On the inclusion of analytic maps into analytic flows. Regular and Chaotic Dynamics", v. 2 (1997), no. 2, 14–24

[14] P. Randall, "A KAM theorem for some degenerate Hamiltonian systems", 1998, MP_ARC: 98–19

[15] M. B. Sevryuk, "KAM-stable Hamiltonians", J. Dynam. Control Systems, V. 1, N. 3, 1995, 351–366

[16] C. Simó "Dynamical properties of the figure eight solution of the three-body problem", in Proceedings of the Celestial Mechanics Conference dedicated to D. Saari for his 60th birthday, Evanston, 1999, ed. A. Chenciner et al, 209–228, *Contemporary Mathematics 292*, AMS, 2000

[17] N. V. Svanidze, "Small perturbations of an integrable dynamical system with an integral invariant", Proc. Steklov Inst. Math. 147 (1981), 127–151

[18] D. Treschev, "A mechanism for the destruction of resonance tori in Hamiltonian systems", Math. USSR-Sb. 1991, v. 68 (1991), no. 1, 181–203

[19] U. Vaidya, I. Mezić, "Existence of invariant tori in three dimensional maps with degeneracy", Physica D 241 (2012), no. 13, 1136–1145