Lagrangian tori near resonances of near-integrable Hamiltonian systems

A G Medvedev\(^1,5\), A I Neishtadt\(^2,3\) and D V Treschev\(^4\)

\(^1\) Lomonosov Moscow State University, Leninskie Gory 1, Moscow 119991, Russia
\(^2\) Department of Mathematical Sciences, Space Research Institute of RAS, 84/32 Profsoyuznaya Str, Moscow 117997, Russia
\(^3\) Loughborough University, Loughborough, Leicestershire LE11 3TU, UK
\(^4\) Steklov Mathematical Institute of RAS, Gubkina str. 8, Moscow 119991, Russia

E-mail: medvedev_ag@inbox.ru, A.Neishtadt@lboro.ac.uk and treschev@mi.ras.ru

Received 1 September 2014, revised 24 December 2014
Accepted for publication 18 March 2015
Published 19 May 2015

Recommended by R de la Llave

Abstract

We study families of Lagrangian tori that appear in a neighbourhood of a resonance of a near-integrable Hamiltonian system. Such families disappear in the ‘integrable’ limit \( \varepsilon \to 0 \). Dynamics on these tori are 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 the order \( \sqrt{\varepsilon} \). Therefore, for small \( \varepsilon > 0 \) a random initial condition in a \( \sqrt{\varepsilon} \)-neighbourhood 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 \( \sim \varepsilon \).

Keywords: Lagrangian tori, KAM theory, resonant domains, Hamiltonian systems, chaotic dynamics
Mathematics Subject Classification: 70H08, 37J40

\(^5\) Author to whom any correspondence should be addressed.

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
1. Projection of trajectories to the action space

Consider a symplectic map
\[(y, x) \mapsto (y_+, x_+), \quad y \in \mathbb{R}^2, \quad x \in \mathbb{T}^2 \quad (1.1)\]
\[y_+ = y - \epsilon \partial V / \partial x, \quad x_+ = x + y_+, \quad V = V(x). \quad (1.2)\]

For \(\epsilon = 0\) 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
\[V = a_1 \cos(x_1 + \varphi_1) + a_2 \cos(x_2 + \varphi_2) + a_3 \cos(x_1 - x_2 + \varphi_3)\]
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 \(a_j \sim 1, \varphi_j, \epsilon \sim 1/10\), 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 in figure 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 (figure 2). The trajectories presented in figure 2, right and centre, are situated near KAM tori, but chaotic effects create a small ‘defocusing’ in comparison with tori from figure 1.

Trying more initial conditions, it is possible to obtain more interesting objects, figure 3, which look like closed ribbons. Further attempts lead to more exotic images, figure 4, looking as finite sequences of spots.
In this paper, we discuss a mechanism that generates such structures. In particular, we show that these objects form sets of positive measures in the phase space. Hence, the probability to observe them is also positive.

At the end of this section, let us give some examples. Fix $\phi_1 = \phi_2 = 1$, $\phi_3 = 2$, $a_1 = 1$, $a_2 = 1$, $b = 0$, $\varepsilon = 0$, $\gamma_1 = 1$, $\gamma_2 = 3$. If $x_1 = 0$, $x_2 = 1$, $685$ we obtain the right picture in figure 1, if $x_1 = 0$, $x_2 = 1$, $684$ we obtain the right picture in figure 4, if $x_1 = 1$, $3595$, $x_2 = 1$, $7$ we obtain the right picture in figure 2, finally if $x_1 = 0$, $x_2 = 1$, $707$ we obtain the right picture in figure 3.

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

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

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

Below we use the following notation for such a system:

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

where the symplectic manifold $(P, \omega_P)$ 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
\[
\langle \nu^0, K \rangle = 0, \quad K \in \mathbb{Z}^N \setminus \{0\}
\]
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 \( \nu^0 \) is said to be non-resonant. 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 \( \nu^0 \). Invariant torus is called resonant (non-resonant) if the dynamics on the torus is quasi-periodic with a resonant (non-resonant) frequency vector. Any torus
\[
T^{\nu^0}_Y = \{(Y, X) : Y = Y^0\}
\]
is invariant with respect to the unperturbed flow
\[
(Y, X) \mapsto (Y, X + \nu(Y)t), \quad t \in \mathbb{R}.
\]
Then \( T^{\nu^0}_Y \) is foliated by the \( m \)-tori
\[
T^{\nu^0}_y = \text{ closure} \left \{ \{(Y, X) : Y = Y^0, \ X = X^0 + \nu^0 t, \ t \in \mathbb{R} \} \right \}.
\]
By computing the frequency vector corresponding to the unperturbed quasi-periodic motion on \( T^{\nu^0}_y \), it is easy to show that the tori \( T^{\nu^0}_y \) are non-resonant.

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

If \( l > 0 \) then a generic perturbation destroys \( T^{\nu^0}_Y \) [11]. However, generically some tori \( T^{\nu^0}_y \) survive a perturbation even if \( \nu^0 \) is resonant. To present the corresponding result, we fix a \( \mathbb{Z} \)-module \( g \subset \mathbb{Z}^n \). Consider the resonance set
\[
\Sigma_g = \{ Y \in D : g(\nu(Y)) = g \}.
\]
Under natural non-degeneracy conditions\(^6\) \( \Sigma_g \) is a real-analytic submanifold, \( \dim \Sigma_g = m \).

It is convenient to study system (2.1)–(2.2) in a \( \sqrt{\epsilon} \)-neighbourhood of the torus \( T^{\nu^0}_Y \), \( Y^0 \in \Sigma_g \) by using the scaling
\[
Y = Y^0 + \sqrt{\epsilon} \tilde{Y}, \quad X = \tilde{X}, \quad H(Y, X, \epsilon) = H_0(Y^0) + \sqrt{\epsilon} \tilde{H}(\tilde{Y}, \tilde{X}, \sqrt{\epsilon}).
\]
Then the system \((P, \omega_P, H)\) turns to the system \((\tilde{P}, \omega_{\tilde{P}}, \tilde{H})\),
\[
\tilde{H} = \langle \nu^0, \tilde{Y} \rangle + \frac{1}{2} \sqrt{\epsilon} (\Pi \tilde{Y}, \tilde{Y}) + \frac{1}{2} \tilde{H}_1(\nu^0, \tilde{X}) + O(\epsilon),
\]
\[
\tilde{P} = \mathbb{R}^N \times \mathbb{R}^N, \quad \omega_{\tilde{P}} = d\tilde{Y} \wedge d\tilde{X}, \quad \Pi = H^0_{YY}(Y^0).
\]
The tori \( T^{\nu^0}_y \) 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

\(^6\) The functions \( \langle K^{(i)}, \partial / \partial Y \rangle H_0(Y) \) are independent in \( D \) where \( K^{(1)}, \ldots, K^{(i)} \) are generators of \( g \).
higher order perturbative terms, and by reduction of the order. Now we turn to description of these steps.

For any function \( f = \sum_{K \in \mathbb{Z}^n} f^K e^{i(K \cdot X)} \) consider the averaging

\[
\langle f \rangle_x = \sum_{K \in \mathbb{Z}^n} f^K e^{i(K \cdot X)}.
\]  

(2.4)

Hence, \( \langle \cdot \rangle_x \) is a projector, removing all non-resonant Fourier harmonics.

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

\[
(\tilde{Y}, \tilde{X}) \mapsto (\hat{Y}, \hat{X}), \quad \hat{Y} = \tilde{Y} + \sqrt{\varepsilon} \tilde{S}_\varepsilon, \quad \hat{X} = \tilde{X},
\]

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

\[
\langle v^0, \tilde{S}_\varepsilon \rangle + H_1(Y^0, \tilde{X}) = v(Y^0, \tilde{X}), \quad v(Y, X) = \langle H_1(Y, X) \rangle_x.
\]  

(2.5)

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

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

\[
\hat{P} = \mathbb{R}^N \times T^N, \quad \omega_{\hat{P}} = d\hat{Y} \wedge d\hat{X}, \quad \hat{H} = H^\varepsilon + O(\varepsilon),
\]

\[
H^\varepsilon = \langle v^0, \hat{Y} \rangle + \frac{\varepsilon}{2} \langle \Pi \hat{Y}, \hat{Y} \rangle + \sqrt{\varepsilon} v(Y^0, \hat{X}).
\]

Consider the approximate system \((\tilde{P}, \omega_{\tilde{P}}, H^\varepsilon)\), which is usually called the partially averaged system. Any critical point \( X^0 \) of the ‘potential’ \( v(Y^0, X) \) generates an invariant torus \( T^m_{Y^0, X^0} \). To study linearization of the averaged system on \( T^m_{Y^0, X^0} \) it is convenient to consider the reduced system. First we recall the general invariant construction (see [2]) and then give a more explicit coordinate form.

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

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

Note that \( v^0 \in g_\perp \) and \( \text{rank } g_\perp = m \).

Action of \( G \) on \( \hat{P} \) is Poissonian and the corresponding momentum map \( M_G : \hat{P} \to g_\perp^* \) is as follows:

\[
M_G(\hat{Y}, \hat{X}) Q = \langle Q, \hat{Y} \rangle \quad \text{for any } Q \in g_\perp.
\]

Reduction with respect to \( G \) means

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

(b) passage to the quotient phase space

\[
\mathcal{P} = \hat{P}/G, \quad \hat{P}/G = \{ (\tilde{Y}, \tilde{X}) \in \hat{P} : \langle Q, \tilde{Y} \rangle = \langle Q, \gamma \rangle \text{ for all } Q \in g_\perp \}.
\]

The reduced phase space has a canonical symplectic structure \( \omega_{\gamma} \) [2] while the Hamiltonian \( \mathcal{H}^\varepsilon : \mathcal{P} \to \mathbb{R} \) is determined by the commutative diagram

\[
\begin{array}{ccc}
\hat{P}/G & \xrightarrow{\text{pr}} & \mathcal{P} \\
\downarrow \mathcal{H}^\varepsilon & & \uparrow \mathbb{R} \\
\mathbb{R} & \swarrow \mathcal{H}^\varepsilon \end{array}
\]

where \( \text{pr} \) is the natural projection. The torus \( T^m_{Y^0, X^0} \) turns to a fixed point \( p \) in the reduced system \((\mathcal{P}, \omega_{\gamma}, \mathcal{H}^\varepsilon)\). The flow of the system \((\mathcal{P}, \omega_{\gamma}, \mathcal{H}^\varepsilon)\) near \( T^m_{Y^0, X^0} \) is essentially determined by linear approximation of \((\mathcal{P}, \omega_{\gamma}, \mathcal{H}^\varepsilon)\) at \( p \). To obtain this approximation, we turn to the coordinate form of the above order reduction.
Let $\Theta$ be an $N \times l$ matrix formed by the integer vectors $K^{(1)}, \ldots, K^{(l)}$, 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 $\Theta M$ instead of $\Theta$. We assume that the quadratic form determined by $\Theta$ 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 $P$ the variables $\eta \in \mathbb{R}^l$, $\xi \in \mathbb{T}^l$ such that

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

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

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

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

Here $v_{\eta^0,y^0} : \mathbb{T}^l \to \mathbb{R}$ is the unique function, satisfying the identity\(^7\)

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

The constant $h_\eta = \frac{1}{2} (\Pi (\gamma - \Theta \eta^0), (\gamma - \Theta \eta^0))$ 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 non-degenerate critical point $\xi^0$ of $v_{\eta^0,y^0}$ the perturbed system admits a real-analytic invariant $m$-torus $\mathbb{T}^m_{Y^0,X^0}(\varepsilon)$. 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, $\mathbb{T}^m_{Y^0,X^0}(\varepsilon)$ 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 $\Lambda_\varepsilon$ on the surface

$$\{Y \in \Sigma_g : v_{\eta^0,y^0} \text{ has nondegenerate critical points}\}
$$
tends to 1 as $\varepsilon \to 0$.

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

$$\mathcal{H}^\varepsilon_{lin} = \frac{\sqrt{\varepsilon}}{2} (\Theta^T \Pi \Theta \eta, \eta) + \frac{\sqrt{\varepsilon}}{2} (V \xi, \xi), \quad V = \frac{\partial v_{\eta^0,y^0}}{\partial \xi^2}(0), \quad \xi = \xi^0 - \eta.$$

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

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

If all $\lambda_j$ are purely imaginary, $p$ and the corresponding torus $\mathbb{T}^m_{Y^0,X^0}(\varepsilon)$ are said to be normally elliptic. The ‘opposite’ case is normally hyperbolic, where no $\lambda_j$ is purely imaginary. In general situation, the torus $\mathbb{T}^m_{Y^0,X^0}(\varepsilon)$ 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.

A hyperbolic case with arbitrary $m$ is presented in [18]. In [4] the case of arbitrary $m$ and arbitrary normal behaviour of the perturbed tori is treated. In [8], it is shown that an additional condition (the so-called $g$-non-degeneracy of $H_0$), introduced in [4], can be skipped.

\(^7\) Explicit formula for $v_\eta$ is as follows. Let the Fourier expansion for $v$ be

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

Then $v_{\eta^0,y^0}(\xi) = \sum_{j \in \mathbb{Z}^l} v^{(j)}(Y)e^{i(j, \xi)}$. 

2110
A statement analogous to theorem 1 should be true in infinite dimension but as far as we know this has not yet been proven.

Since for non-trivial $g$ all the tori $T^m_{Y_i^\varepsilon}(\varepsilon)$ are lower dimensional ($m < N$), their total measure in the phase space $D \times T^N$ 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 $N$ with these tori appears.

In this paper we are interested in visible objects. More precisely, in invariant tori of dimension $3$. Visible objects this has not yet been proven.

A statement analogous to theorem 1 should be true in infinite dimension but as far as we know this has not yet been proven.

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 $N$ with these tori appears.

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 $N$ with these tori appears.

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 $N$ with these tori appears.

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 $N$ with these tori appears.

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 $N$ with these tori appears.

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 $N$ with these tori appears.

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 $N$ with these tori appears.

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 $N$ with these tori appears.

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 $N$ with these tori appears.

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 $N$ with these tori appears.

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 $N$ with these tori appears.

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.
Since the system \((P, \omega_P, H^g)\) is integrable, the set of phase points in a \(\sqrt{\varepsilon}\)-neighbourhood of the torus \(\mathbb{T}^N_{\nu_0}\) lying outside invariant \(N\)-tori of the original system \((P, \omega_P, H)\) has a small relative measure. A 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 figure 3 is not less than \(\sqrt{\varepsilon} C_1\), \(C_1 > 0\), because the width of a resonance domain corresponding to a single resonance is \(\sim \sqrt{\varepsilon}\). Pictures analogous to figure 3 can be found in [16]. Similar tori are considered in the context of Arnold diffusion in [5].

**Multiple resonance.** If \(\text{rank } g > 1\), the systems \((\hat{P}, \omega_{\hat{P}}, H^g)\) and \((P, \omega_P, H^g)\) are generically non-integrable. Therefore, the existence of invariant \(l\)-tori in the latter one is not straightforward provided the energy \(H^g\) is not very big or not very small. A standard source for such tori is a neighbourhood of a totally elliptic fixed point. However, a totally elliptic fixed point may not exist if the ‘kinetic energy’ \(\sqrt{\varepsilon} \langle \Theta^{1/2} \Pi \Theta_{\eta}, \eta \rangle\) is indefinite: a simple example is

\[
H^g = \sqrt{\varepsilon} (\eta_1^2 + \cos \xi_1 + \cos \xi_2).
\]

System (1.1)–(1.2) corresponds to a positive definite kinetic energy and trajectories presented in figure 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 \(\varepsilon\), because the measure of a resonant domain corresponding to a double resonance is of order \(\varepsilon\). 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 neighbourhood of the resonance \(\Sigma_1 \times \mathbb{T}^{n+1}\), 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

\[
\Sigma = \Sigma_\varepsilon = \{Y \in \mathcal{D} : \langle K^0, v(Y) \rangle = 0\}.
\]  

(4.1)

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

\[
\det H^g_{0Y} \neq 0, \quad Y \in \mathcal{D},
\]  

(4.2)

\[
A(Y) \equiv \langle K^0, H''_{0Y} K^0 \rangle \neq 0.
\]  

(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 definition 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

\[
\frac{d}{d\lambda} H_0(Y - \lambda K^0) \equiv \langle K^0, v(Y - \lambda K^0) \rangle = 0
\]  

(4.4)

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

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

Let \(\langle \cdot \rangle^K\equiv \langle \cdot \rangle^g\) be the operator of resonant averaging

\[
f = \sum_{K \in \mathbb{Z}^{n+1}} f^K e^{i(K,X)} \mapsto \langle f^K \rangle^0 = \sum_{j=-\infty}^{\infty} f^j e^{ijq}, \quad q = \langle K^0, X \rangle.
\]
Consider the Hamiltonian system \((P, \omega_P, H_{K^0})\),
\[ H_{K^0}(Y, X) = H_0(Y) + \varepsilon u(Y, q). \]  
(4.5)
The function \(u\) is \(2\pi\)-periodic in the resonant phase \(q\) and
\[ u(Y + \lambda K^0, q) = u(Y, q) \quad \text{for any } \lambda \text{ in a neighbourhood of } 0 \in \mathbb{R}. \]
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 \(H_0(Y) = H_0(\chi(Y)) + \frac{1}{2}A(\chi(Y))\lambda^2(Y) + O(\lambda^3(Y))\), Hamiltonian (4.5) can be presented in the form
\[ H_{K^0}(Y, X) = H_0(\chi(Y)) + \frac{1}{2}A(\chi(Y))\lambda^2(Y) + \varepsilon u(Y, q) + O(\lambda^3(Y)). \]  
(4.6)

For any vector \(J \in \mathbb{R}^{n+1}\) such that \((K^0, J) = 0\) the function \((Y, J)\) is a first integral. Therefore, the system \((P, \omega_P, H_{K^0})\) 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 \(u(Y, q)\) for fixed \(Y \in \Sigma\):
\[ u(Y, q_{\min}(Y)) = \min_{q \in \mathbb{T}} u(Y, q), \quad u(Y, q_{\max}(Y)) = \max_{q \in \mathbb{T}} u(Y, q). \]  
(4.7)
Then we define
\[ \mathcal{D}_{os} = \{(Y, X) \in \mathcal{D} \times \mathbb{T}^{n+1} : \varepsilon u(Y, q_{\min}(Y)) < \frac{1}{2}A(\chi(Y))\lambda^2(Y) + \varepsilon u(Y, q) < \varepsilon u(Y, q_{\max}(Y))\}. \]  
(4.8)
If \(u(Y, q)\) is not a constant as a function of \(q\), the domain \(\mathcal{D}_{os}\) belongs to an \(O(\sqrt{\varepsilon})\)-neighbourhood 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 \(\varepsilon\) all these tori except a set of a small measure survive the perturbation.

To formulate the result, for any \(Y\) in a neighbourhood of \(\Sigma\) consider the Hamiltonian system
\[ \dot{q} = \sqrt{\varepsilon} \hat{H}_p(\chi(Y), p, q), \quad \dot{p} = -\sqrt{\varepsilon} \hat{H}_q(\chi(Y), p, q) \]  
(4.9)
with one degree of freedom, the Hamiltonian
\[ \hat{H}(\chi(Y), p, q) = \frac{1}{2}A(\chi(Y))p^2 + u(Y, q) \]  
and the symplectic structure \(\sqrt{\varepsilon} dp \wedge dq\). The point \(Y\) is regarded as a parameter. Recall that by (4.3) \(A \neq 0\). This system coincides with \((P, \omega_Y, \hat{H}_Y^\varepsilon)\) from section 3 written in slightly other terms.

**Proposition 4.1.** The point \(\chi(Y)\) is a constant of motion in the system \((P, \omega_P, H_{K^0})\). The variables \(\lambda(Y) = \sqrt{\varepsilon} p\) and \(q\) in the averaged system satisfy equation (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 \((K^0, \partial/Y)\) to (4.4) we get: \((K^0, \lambda_Y(Y)) = 1 + O(\lambda)\). Then by (4.6) in system \((P, \omega_P, H_{K^0})\)
\[ \dot{\lambda} = \lambda_Y \dot{Y} = -\varepsilon u_q(Y, q) + O(\varepsilon \lambda) = -\varepsilon \hat{H}_q + O(\varepsilon \lambda), \]
\[ \dot{q} = (K^0, H_{K^0}^\varepsilon(Y)) = A(\chi(Y))\lambda + O(\lambda^2) + O(\varepsilon) = \sqrt{\varepsilon} \hat{H}_p + O(\varepsilon + \lambda^2). \]
The projection $\pi : U(\Sigma \times \mathbb{T}^{n+1}) \to \Sigma \times \mathbb{R} \times \mathbb{T}$,

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

depicts the domain $D_{os}$ to $\hat{D}_{os} = \pi(D_{os})$, see figure 5.

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

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

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

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

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

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

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

is invariant for the system $(P, \omega_P, H_{K^0})$ 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 $v_{Y^0, \gamma}$ associated with $\mathbb{T}^{n+1}_{\gamma(Y, p, q)}$ is Diophantine, i.e. for some $\tau > 0$ and $C_\tau > 0$

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

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

2. The system $(P, \omega_P, H_{K^0})$ is non-degenerate on $\mathbb{T}^{n+1}_{\gamma(Y, p, q)}$.

This condition essentially means that

$$|h'_\gamma| < c, \quad |h''_{\gamma}| < c, \quad |h'_\gamma|^{-1} < c, \quad |h''_{\gamma}|^{-1} < c. \quad (4.13)$$

Therefore, we have to replace $D_{os}$ by a smaller domain $D_0$ by throwing out a small neighbourhood of asymptotic manifolds (where the tori degenerate) and a small neighbourhood of tori on which the twist condition (4.13) is violated. Let $\mu$ be the measure in the phase space $D \times \mathbb{T}^{n+1}$ generated by the symplectic structure $\omega_P$. Then the measures $\mu(D)$ and $\mu(D_0)$ are both of order $O(\sqrt{\varepsilon})$. 

2114
Theorem 2. Suppose that the system \((P, \omega_P, H)\) is real-analytic. If \(\varepsilon_0 > 0\) is sufficiently small, then for all positive \(\varepsilon \leq \varepsilon_0\) any Diophantine torus \(\tilde{T}^{m+1}_n(\varepsilon) \subset D_0\) survives the perturbation. For some constant \(C > 0\) independent of \(\varepsilon\), the measure of union of such tori in \(D_0\) is not less than \(\mu(D_0) - C\varepsilon\).

The measure of invariant tori in a near-integrable Hamiltonian system. It is well known that the 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 the ‘chaotic’ set is of order \(\sqrt{\varepsilon}\)). It would be interesting to modify the proof, considering weaker non-degeneracy conditions \((4.13)\) to improve the estimates of the measure of the complement to the tori of near-integrable Hamiltonian systems. Here, it is natural to recall 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: \(\langle u, v \rangle = \sum_{j=1}^m u_j v_j\).
- \(\mu\) 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
  \[
  \langle f \rangle = \frac{1}{(2\pi)^m} \int_{\mathbb{T}^m} f(x) \, dx.
  \]
  The same notation is used if \(f\) depends on other variables. In this case to avoid misunderstanding we use for \((5.1)\) the notation \(\langle f \rangle_x\).
- Below \(c_1, c_2, \ldots\) denote positive constants. If \(c_j\) depends on another constant, say, \(\alpha\), we write \(c_j(\alpha)\). Dependence on the dimension \(n\) is not indicated.

To present the system \((P, \omega_P, H^\varepsilon)\) 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
\[
\tilde{\gamma}_{m+1}(\hat{Y}) = \tilde{\gamma}_{m+1}^0(\hat{Y}) = 0, \quad \hat{H}_0(\hat{Y}) = H_0(Y).
\]

\[2115\]
To have more convenient coordinates in a neighbourhood of this resonance, we solve the first equation (5.2) with respect to \( \dot{y}_{n+1} \). This can be done locally because by (4.3) \( \hat{H}_{0,\hat{y}_{n+1},\hat{y}_{n+1}} \neq 0 \). We denote the result
\[
\hat{y}_{n+1} = G(\hat{y}_1, \ldots, \hat{y}_n), \quad \dot{\hat{y}}_{n+1}(\hat{y}_1, \ldots, \hat{y}_n, G(\hat{y}_1, \ldots, \hat{y}_n)) \equiv 0.
\]

(b). Consider the change of the variables
\[
y = (y_1, \ldots, y_n) = (\hat{y}_1, \ldots, \hat{y}_n), \quad p = \epsilon^{-1/2}(\hat{y}_{n+1} - G(y)),
\]
\[
x = (x_1, \ldots, x_n) = (\hat{x}_1 + G'_{y_1}q, \ldots, \hat{x}_n + G'_{y_n}q), \quad q = \hat{x}_{n+1}.
\]

We are interested in motions which are oscillatory in the coordinate \( q \). Therefore, it is not necessary to assume the 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 \( \Sigma \) in the new coordinates locally takes the form \( \{p = 0\} \) while \( y \) are local coordinates on \( \Sigma \).

Then the symplectic structure and the Hamiltonian take the form
\[
\omega = dy \wedge dx + \sqrt{\epsilon} dp \wedge dq,
\]
\[
\Lambda(y) + \epsilon \left( \frac{1}{2} A(y)p^2 + u(y, q) \right) + \epsilon U_1(y, x, q) + \epsilon^{3/2} U_2(y, p, x, q, \sqrt{\epsilon}),
\]
(5.3)

where
\[
\Lambda(y) = \hat{H}_0(y, G(y)), \quad A(y) = \hat{H}_{0,\hat{y}_{n+1}}(y, G(y)),
\]
\[u, U_1, U_2 \text{ are real-analytic, and average of } U_1 \text{ with respect to } x \text{ vanishes: } \langle U_1 \rangle_x = 0.\]

By (4.3),(4.5)
\[
A(y) = A(\chi(Y)), \quad u(y, q) = u(\chi(Y), q).
\]

Neglecting the terms \( \epsilon U_1 + \epsilon^{3/2} U_2 \), 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 \( D_{\text{cs}} \) (see (4.8)) takes the form
\[
D_{\text{cs}} = \{y, p, x, q: u(y, \hat{q}_{\min}) < \frac{1}{2} A(y)p^2 + u(y, q) < u(y, \hat{q}_{\max})\}.
\]
(5.5)

(c). Let \( W(y, I, q) \) be a generating function which introduces action-angle variables \( I, \varphi \) 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 \( \hat{y}x + \sqrt{\epsilon} W(\hat{y}, I, q) \)
\[
y = \hat{y}, \quad \hat{x} = x + \sqrt{\epsilon} W'_I, \quad p = W'_{q'}, \quad \varphi = W'_I
\]
transforms the symplectic structure to \( d\hat{y} \wedge d\hat{x} + \sqrt{\epsilon} d\varphi \wedge dI \) and Hamiltonian (5.3) to
\[
\Lambda(\hat{y}) + \epsilon h(\hat{y}, I) + \epsilon \hat{V}(\hat{y}, I, \hat{x}, \varphi, \sqrt{\epsilon}), \quad \langle \hat{V} \rangle_x = O(\sqrt{\epsilon}),
\]
(6.1)

where the functions \( h, \Lambda, \hat{V} \) are real-analytic. The function \( h \) satisfies the equation
\[
h(y, I) = h(\chi(Y), I), \quad Y \in U(\Sigma),
\]
where \( h \) 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 neighbourhood:

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

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 neighbourhood of $\mathbb{T}^{n+1}$:

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

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 suprema 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

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

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

$$H_0 = \Lambda(y) + \varepsilon h_0(y, I) + \varepsilon \nu_0(y, I, \varphi, \sqrt{\varepsilon}) + \varepsilon u_0(y, I, x, \varphi, \sqrt{\varepsilon}),$$

$$\nu_0(y, I, \varphi, 0) = 0, \quad \langle u_0 \rangle_x = 0,$$

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

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

for some $a_0, b_0 > 0$.

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

$$|\Lambda|_{a_0} \leq \tilde{s}_0, \quad |h_0|_{a_0} \leq c_h, \quad |u_0|_{a_0,b_0} \leq s_0, \quad |\nu_0|_{a_0,b_0} \leq \sqrt{\varepsilon}c.$$

(6.2)

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

$$c_A \leq |\det \Lambda'_{yy}| \leq \tilde{c}_A, \quad |\Lambda'_{yy}|_{a_0} \leq c_A, \quad |\Lambda'^{-1}_{yy}|_{a_0} \leq c_A, \quad |h_{0I}|_{a_0} \leq c_h, \quad |h_{0I}|^{-1}_{a_0} \leq c_h'$$

$$h_{0II}|_{a_0} \leq c_h', \quad |h_{0II}|^{-1}_{a_0} \leq c_h', \quad |\sqrt{\varepsilon}h_{0II}|_{a_0} \leq c_h', \quad |\varepsilon h_{0II}|_{a_0} \leq c_h''.$$ 

(6.3)

Let $\varepsilon$ be sufficiently small, then we can assume that $c_h''$ is small because $c_h'' \sim \sqrt{\varepsilon}$. Below we assume that $c_h'' \leq c_h''(c_A, c_A, c_h')$. 

(6.4)
7. The Hamiltonian $H_m$

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

As usual the KAM procedure includes a converging sequence $F_0, F_1, \ldots$ of coordinate changes and a converging sequence of Hamiltonians $H_0, H_1, \ldots$

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

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 $j : \mathbb{N} \rightarrow \mathbb{N}$ defined by the inequality

$$N_{j(r)-1} < r \leq N_{j(r)} \quad \text{for all} \ r > 0.$$  

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

Consider two positive decreasing sequences $a_m, b_m$,

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

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

$$H_m = \lambda(y) + \varepsilon h_m(y, I) + \varepsilon v_m(y, I, \varphi) + \varepsilon u_m(y, I, x, \varphi), \quad \langle u_m \rangle_x = 0.$$  

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

$$D_{m+1} = D_m \setminus \bigcup_{|K| \leq N_{m+1}, k \neq 0} U_{a_k}(Q_{K,m}), \quad K = \binom{k}{0} \in \mathbb{Z}^{n+1}, \quad k \in \mathbb{Z}^n,$$  

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

$$Q_{K,m} = \{(y, I) \in U_{a_m}(D_m) : \left|\nu_m(y, I, K)\right| \leq \lambda_j(|K|) (1 + 2^{-m-1}) \sqrt{\varepsilon}\},$$  

$$v_m(y, I) = \left(\Lambda'_y(y) + \varepsilon h_{my}'(y, I), \frac{\sqrt{\varepsilon} h_{my}'(y, I)}{h_{my}'(y, I)}\right).$$  

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

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

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

$$D_{m+1} = D_m \setminus \bigcup_{N_{m-1} < |K| \leq N_m} U_{a_n}(Q_{K,m}).$$  

Proposition 7.1. For any $m = 1, 2, \ldots$

$$\mu(D_0 \setminus D_m) \leq c\mu \sqrt{\varepsilon},$$  

where $c\mu > 0$ is independent of $\varepsilon$.

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

$$|\nu_m|_{a_m, b_m} \leq s_m, \quad |u_m|_{a_m, b_m} \leq s_m, \quad |h_m|_{a_n} \leq (2 - 2^{-m}) c_h, \quad |h_{m,l}^-|_{a_n} \leq (2 - 2^{-m}) c_h', \quad |h_m''|_{a_n} \leq (2 - 2^{-m}) c_h'', \quad |\varepsilon h_{my}'|_{a_n} \leq (2 - 2^{-m}) c_h'''.$$  

$$|\sqrt{\varepsilon} h_{my}|_{a_n} \leq (2 - 2^{-m}) c_h, \quad |\varepsilon h_{my}'|_{a_n} \leq (2 - 2^{-m}) c_h', \quad |\varepsilon h_{my}''|_{a_n} \leq (2 - 2^{-m}) c_h''.$$

2118
8. The KAM step

For any natural $N$ and a periodic function

$$f : \mathbb{Z}_{N+1} \to \mathbb{R}, \quad f(x, \varphi) = \sum_{K=(k,k')} f^K e^{i(k,x)+ik\varphi}$$

we define the cut off

$$\Pi_N f(x, \varphi) = \sum_{|K| \leq N, k \not= 0} f^K e^{i(k,x)+ik\varphi}.$$  \hfill (8.1)

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

$$|f - (f)_x - \Pi_N f|_{b-\delta} \leq \frac{C}{\delta} \left( N + \frac{1}{\delta} \right)^n e^{-N\delta} |f|_b.$$  \hfill (8.2)

By using Hamiltonian (7.3), we introduce the canonical$^8$ change of variables $(y, I, x, \varphi) \mapsto (\hat{y}, \hat{I}, \hat{x}, \hat{\varphi})$, determined by the generating function $\hat{y}x + \sqrt{e} (\hat{I}\varphi + S(\hat{y}, \hat{I}, x, \varphi))$:

$$y = \hat{y} + \sqrt{e} S'_y, \quad I = \hat{I} + S'_y, \quad \hat{x} = x + \sqrt{e} S'_y, \quad \hat{\varphi} = \varphi + S'_y,$$

where the arguments $(\hat{y}, \hat{I})$ are supposed to lie in $U_{a_0 - \sigma,(D_{m+1})}$.

To definition the function $S$ is a solution of the homologic equation

$$\left\{v_m(\hat{y}, \hat{I}), S'_y(\hat{y}, \hat{I}, x, \varphi)\right\} = -\sqrt{e} \Pi_{N_m} V_m(\hat{y}, \hat{I}, x, \varphi), \quad V_m = v_m + u_m.$$  \hfill (8.3)

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

$$|S|_{a_0 - \sigma,(\Delta_{m+1})} \leq L_m s_m, \quad L_m = \sum_{j=0}^{n} \frac{N^{j+1}}{\lambda_j} e^{-\lambda_j a_0} s_{n-j}.$$  \hfill (8.4)

The Hamiltonian (7.3) takes the form

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

**Remark 8.2.** Below we show that

$$|S'_y|_{a_0 - \sigma,(\Delta_{m+1})} \leq \sigma_m, \quad |S'_y|_{a_0 - \sigma,(\Delta_{m+1})} \leq \sigma_m.$$  \hfill (8.6)

$$|\sqrt{e} S'_y|_{a_0 - \sigma,(\Delta_{m+1})} \leq \delta_m, \quad |S'_y|_{a_0 - \sigma,(\Delta_{m+1})} \leq \delta_m.$$  \hfill (8.7)

Estimates (8.6)–(8.7) imply that the coordinate change is well defined for $(\hat{y}, \hat{I}, \hat{x}, \hat{\varphi}) \in U_{a_0 - \sigma,(\Delta_{m+1})} (D_{m+1} \times T^{n+1})$.

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

$$|\tilde{v}_m + \tilde{u}_m|_{a_0 - \sigma,(\Delta_{m+1})} \leq \tilde{s}_m.$$  \hfill (8.8)

$$\tilde{s}_m = (c_\Lambda + c_\delta) \left( \frac{L_m s_m}{\delta_m} \right)^2 + (n+2) \frac{L_m s_m^2}{\sigma_m \delta_m} + C \delta_m \left( N_m + \frac{1}{\delta_m} \right)^n e^{-N_m \lambda_m} s_{m'}.$$  \hfill (8.9)

$^8$ That is, preserving the symplectic structure (6.1).
9. An additional step

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

\[
h_m(\tilde{y}, \tilde{I}) + \varepsilon v_m(\tilde{y}, \tilde{I}, \tilde{\varphi}) = h_{m+1}(\hat{y}, \hat{I}).
\]  

(9.1)

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

\[
\hat{y} = \tilde{y}, \quad \hat{I} = \tilde{I} + \sqrt{\varepsilon} \tilde{S}, \quad \hat{\varphi} = \tilde{\varphi} + \sqrt{\varepsilon} \tilde{S}_y.
\]  

(9.2)

Then Hamiltonian (8.5) takes the form

\[
H_{m+1} = \Lambda(\hat{y}) + \varepsilon h_{m+1}(\hat{y}, \hat{I}) + \varepsilon v_{m+1}(\hat{y}, \hat{I}, \hat{\varphi}) + \varepsilon u_{m+1}(\hat{y}, \hat{I}, \hat{\varphi}), \quad (u_{m+1})_{\hat{y}} = 0.
\]

Proposition 10.1. Suppose that

\[
|v_0|_{\text{loc}} \leq \sqrt{\varepsilon} \leq \frac{\sigma_0 h_0}{2c^2}, \quad |v_m|_{\text{loc}} \leq \sigma_m, \quad s_m \leq \frac{\sigma_m \delta_m}{4c^2}, \quad m \geq 1.
\]  

(9.3)

Then for any \( m \geq 0 \)

\[
|\tilde{S}_y|_{\text{loc}} \leq 8s_m s'_m, \quad |\tilde{S}|_{\text{loc}} \leq 16\pi(c^2 + s'_m),
\]  

(9.4)

\[
|h_{m+1} - h_m|_{\text{loc}} \leq \frac{8c^2 s'_m}{\sigma_m}, \quad |u_{m+1} + v_{m+1}|_{\text{loc}} \leq s'_m,
\]  

(9.5)

where \( s'_m = \sqrt{\varepsilon} s_m \) and \( s'_m = s_m \) for all \( m \geq 1 \).

Remark 9.1. Below we show that

\[
|\tilde{S}_y|_{\text{loc}} \leq \delta_m, \quad \sqrt{\varepsilon} |\tilde{S}|_{\text{loc}} \leq \delta_m,
\]  

(9.6)

\[
|\tilde{S}_y|_{\text{loc}} \leq \sigma_m.
\]  

(9.7)

Estimates (9.6), (9.7) imply that the coordinate change (9.2) is well defined for \((\tilde{y}, \tilde{I}, \tilde{\varphi}) \in U_{\lambda_m, \beta_m} \times U_{\bar{\lambda}, \bar{\beta}}(D_{m+1} \times T^{m+1})\).

10. The sequences \( \sigma_m, \delta_m, s_m, \lambda_m, N_m, \lambda_m \)

We define \( \sigma_m \) and \( \delta_m \) by (7.2) and put

\[
\sigma_m = \frac{a'_0}{6} 2^{-\frac{2}{2} + \frac{1}{2} + \frac{1}{2} + \frac{1}{2}}, \quad \delta_m = \frac{b_0}{3} 3^{-\frac{2}{2} + \frac{1}{2} + \frac{1}{2} + \frac{1}{2}}, \quad a'_0 = \frac{a_0}{2^{m+2}},
\]

(10.1)

\[
s_m = s_0 \epsilon^{-c_m - \frac{1}{2}}, \quad N_m = c_N 2^m, \quad \lambda_m = c_\lambda 2^{-m},
\]

(10.2)

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

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

\[
\lambda_m \leq \frac{c_N^{m+1}}{c_\lambda} 2^{\left(\frac{2}{2} + \frac{1}{2} + \frac{1}{2} + \frac{1}{2}\right)}.
\]  

(10.3)

To show that our choice of the sequences \( a_m, b_m, \sigma_m, \delta_m, s_m, \lambda_m, N_m, \lambda_m \) 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:
\[
\left( \frac{Lm \delta_m}{\delta_m} \right)^2 < \frac{\text{48} \text{e}^{2(\nu+1)\text{e}^{8(\nu+1)\text{e}^{12\nu}}} (\text{e}^{-2\nu} \text{e}^{-2m})}{b_0^2 c_L^2} \leq c_L^2 2^{(8\nu+12)m} e^{-c_L(m-1)s_{m+1}},
\]
\[
\frac{Lm \delta_m^2}{\delta_m^2} \leq 18 \frac{2^{(10n+16)\text{e}^{n+1}+1(6n+11)m}}{b_0 a_0 c_L^2} s_0 e^{-2c_L(m-1)s_{m+1}} \leq c_L^2 2^{(8\nu+12)m} e^{-c_L(m-1)s_{m+1}},
\]
where \( c_L = \max \left( 2^{(8\nu+12)m} e^{-c_L(m-1)s_{m+1}} \right) \). For the third term of (8.9) we have
\[
\frac{C}{\delta_m} \left( N_m + \frac{1}{\delta_m} \right)^n \leq \frac{6C}{b_0} \left( c_N + \frac{b_0}{b_0} \right)^n e^{-c_N(m-1)} \leq \frac{6C}{b_0} \left( c_N + \frac{b_0}{b_0} \right)^n e^{-c_N(m-1)}.
\]

10.2. Inequalities (7.7)

Rewrite (8.9) in terms of \( s_{m+1} \)
\[
\tilde{s}_m \leq (c_{\lambda} + c_i) c_L s_{m+1} + (n + 2) c_L s_{m+1}
\]
\[
+ \frac{6C}{b_0} \left( c_N + \frac{b_0}{b_0} \right)^n e^{2(\nu+1)m} e^{-c_L(m-1)s_{m+1}}.
\]
If \( c_s > c_0 = 16n + 24 \) and \( c_{\lambda} > c_{\lambda}/a_0(b_0, s_0, c_s) \), then for all \( m \geq 0 \) we have
\[
\frac{c_L}{b_0} \left( c_N + \frac{b_0}{b_0} \right)^n e^{2(\nu+1)m} e^{-c_L(m-1)s_{m+1}} \leq \frac{1}{3} \text{max}(c_{\lambda} + c_i)^{-1},
\]
\[ (n + 2)^{-1} \]
For sufficiently large \( c_{\lambda} \) for all \( c_{\lambda} > c_{\lambda}(b_0, s_0) \) we obtain
\[
\frac{6C}{b_0} \left( c_N + \frac{b_0}{b_0} \right)^n e^{2(\nu+1)m} e^{-c_L(m-1)s_{m+1}} \leq \frac{1}{3}.
\]
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) \( h_1 - h_0 = O(\sqrt{e}) \). Hence it is sufficient to check (7.8) and (7.9) only for \( m \geq 1 \). For large \( c_s > c_{s1} = c_s(c_L, a_0, s_0) \):
\[
|h_{m+1} - h_m|_{a_0, a_1, a_2} \leq \frac{8c_s c_L^2 s_m}{\sigma_m} \leq \frac{8c_s c_L^2 s_m}{a_0} \leq 48c_s \frac{2^{2n+5} c_L s_0}{a_0} e^{-(c_s-2n-5)m-2^n} \leq c_L 2^{-m-1},
\]
\[
|h''_{m+1} - h''_m|_{a_0, a_1, a_2} \leq \frac{8c_s c_L^2 s_m}{\sigma_m} \leq 1728 \frac{2^{6n+15} c_L^2 s_0}{a_0^2} e^{-(c_s-6n-15)m-2^n} \leq c_L^2 2^{-m-1}.
\]

Note, that for \( c_s > c_s = c_{s2}(c_L, a_0, s_0) \) we have
\[
|h'_{m+1} - h'_{m+1}|_{a_0, a_1, a_2} \leq \frac{1}{(2-2^{-m})c_{s2}^2} \left( h'_{m+1} - h'_{m+1} \right)_{a_0, a_1, a_2}
\]
\[
\geq \frac{1}{(2-2^{-m})c_{s2}^2} \frac{8c_s c_L^2 s_m}{\sigma_m^2} \geq \frac{1}{(2-2^{-m})c_{s2}^2}
\]
\[
- 288 \frac{2^{4n+10} c_L^2 s_0}{a_0^2} e^{-(c_s-4n-10)m-2^n} \leq \frac{1}{(2-2^{-m})c_{s2}^2} \frac{2^m}{(2m+1) - (2m+1)c_{s2}^2} = \frac{1}{(2-2^{-m})c_{s2}^2}.
\]
Consider the first inequality (7.9). For \( c_\varepsilon > c_\varepsilon 3 = c_\varepsilon 3 (c_\varepsilon h, c_\varepsilon h', a_0, s_0) \)
\[
|h_{m+1}''| |a_{m+1}| - |h_{m+1}'' - h_{m+1}''| |a_{m+1}|
\geq \frac{1}{(2 - 2^{-m})c_h^{'}} - \frac{8\sigma^3 m}{\sigma^4 m} \geq \frac{1}{(2 - 2^{-m})c_h^{'}} - \frac{1}{2^m} - \frac{1728}{a_0^2} 2^{6\epsilon + 15} c_h^{'2} c_h^{'2} e^{-(c_\varepsilon - 6\varepsilon - 15)m - 2^n}
\]

For the first two inequalities (7.9) let \( c_\varepsilon > c_\varepsilon 4 = c_\varepsilon 4 (c_\varepsilon h, c_\varepsilon h', c_\varepsilon h', a_0) \). Then
\[
|\sqrt{E} h_{m+1}'' | - |\sqrt{E} h_{m+1}'' | |a_{m+1}| \leq \sqrt{E} \sigma^3 m \leq \sqrt{E} \sigma^3 m \leq c_\varepsilon h^{'2} 2^{-m - 1}.
\]

For sufficiently large \( c_\varepsilon > \max (c_\varepsilon 1, c_\varepsilon 2, c_\varepsilon 3, c_\varepsilon 4) \) the exponents \( e^{-(c_\varepsilon - 2\varepsilon - 5)m - 2^n} \) are small and all inequalities (7.8), (7.9) hold.

10.4. Inequalities (8.6), (8.7)

Note, that for \( c_\varepsilon > c_\varepsilon 1 (c_N, c_N, b_0, a_0, s_0) \):
\[
|\sqrt{E} h_{m+1}'' | |a_{m+1}| \leq \sqrt{E} \sigma^3 m \leq \sqrt{E} \sigma^3 m \leq c_\varepsilon h^{'2} 2^{-m - 1}.
\]

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 \( \varepsilon \). Note, that for \( c_\varepsilon > c_\varepsilon 5 (s_0, c_\varepsilon h, a_0, b_0) \) we obtain
\[
\sigma_m \delta_m \leq \frac{s_m \sigma^4 m}{4c_\varepsilon} \leq \frac{1}{6} \frac{a_0^2}{2^{-(2\varepsilon + 5)(m + 1)}} = \sigma_m.
\]

By proposition 9.1:
\[
|\sqrt{E} h_{m+1}'' | |a_{m+1}| \leq \sqrt{E} \sigma^3 m \leq \sqrt{E} \sigma^3 m \leq c_\varepsilon h^{'2} 2^{-m - 1}.
\]

For \( m = 0 \) inequalities (9.3), (9.6), (9.7) hold if \( \varepsilon \) is sufficiently small. Consider \( m \geq 1 \). For \( c_\varepsilon > c_\varepsilon 6 (s_0, c_\varepsilon h, a_0, b_0) \)
\[
|\sqrt{E} h_{m+1}'' | |a_{m+1}| \leq \sqrt{E} \sigma^3 m \leq \sqrt{E} \sigma^3 m \leq \delta_m.
\]

We choose the constants in the following way. Fix \( c_\varepsilon > \max (c_\varepsilon 1, \ldots, c_\varepsilon 6) \), \( c_N > c_N (c_\varepsilon) \) and \( c_\varepsilon \varepsilon = c_N (c_\varepsilon) \varepsilon > \max (c_\varepsilon 0, c_\varepsilon 1) \), we obtain for \( m \geq 0 \) inequalities (7.7), (8.6), (8.7) and for \( m \geq 1 \) we obtain (7.8), (7.9), (9.3), (9.6), (9.7). Finally, for sufficiently small \( \varepsilon \) 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

\[
|\langle \nu, K \rangle| > \sqrt{\epsilon} \lambda_j(K) = c_j 2^{-2(2m+2\tau+1)j(K)} = \frac{c_j \epsilon^{m+1+\tau}}{N_j(K)} = \frac{C_\tau}{|K|^{m+1+\tau}}, \quad C_\tau = c_j \epsilon^{m+1+\tau}.
\]

11.2. Proof of proposition 7.1

Proposition 11.1. For any \( K \in \mathbb{Z}^{n+1}, |K| \leq N_m \)

\( Q_{K,m+1} = \emptyset \).

Consider the scaled frequency map \( \tilde{\nu}_m : (y, I) \mapsto (\Lambda'_y(y) + \epsilon h''_{my}(y, I), h''_{my}(y, I)) \).

In comparison with (7.6) we remove the multiplier \( \sqrt{\epsilon} \) at \( h''_{my} \). Its Jacobi matrix equals

\[
J_m(y, I) = \begin{pmatrix}
\Lambda''_{yy} + \epsilon h''_{my} & h''_{my} \\
\epsilon h''_{my} & h''_{mI}
\end{pmatrix}.
\]

(11.1)

Proposition 11.2. For some positive constants \( C_J \) and \( C_J^\prime \)

\[
|\det J_m(y, I)| \leq |C_J|, \quad |\det J_m(y, I)|^{-1} \geq |C_J^\prime|.
\]

Estimates (11.2) imply the following inequality for the measure of the domain \( \tilde{\nu}_m(D_m) \):

\[
\mu(\tilde{\nu}_m(D_m)) < C_J \mu(D_m).
\]

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

\[
Q^\omega_{K,m} = \left\{ (\omega_y, \omega_I) \in \tilde{\nu}_m(D_m) : |\langle \omega_y, K \rangle + \sqrt{\epsilon} \omega_I k_0| \leq \lambda_j(|K|) \right\}.
\]

(11.3)

The set \( Q^\omega_{K,m} \subset \mathbb{R}^{n+1} \) is a strip between two planes

\[
\frac{(\omega_y, k)}{\sqrt{(k, k) + \epsilon k_0^2}} \pm \frac{\sqrt{\epsilon} \omega_I k_0}{\sqrt{(k, k) + \epsilon k_0^2}} = \frac{\lambda_j(|K|) (1 + 2^{-m-1})}{\sqrt{\epsilon}}.
\]

Using (8.1) we have that \( (k, k) \geq 1 \) and the distance between the planes is not more than \( 4 \lambda_j(|K|) \sqrt{\epsilon} \). The measure estimates are

\[
\mu(\tilde{\nu}_m(D_m)) \leq 4 \lambda_j(|K|) \sqrt{\epsilon} C_J D,
\]

\[
\mu(\tilde{\nu}_m(D_m) \cap Q^\omega_{K,m}) \leq \mu(\tilde{\nu}_m^{-1}(Q^\omega_{K,m})) \leq 4 \lambda_j(|K|) \sqrt{\epsilon} C_J \bar{C}_J D,
\]

where \( C_D \) depends on the diameter and dimension of \( D_0 \).

Consider estimates for the measure of \( D_m \cap U_{\omega_0}(Q_{K,m}) \). Let \( |y'| \leq \sqrt{\epsilon} \sigma_m, |I'| \leq \sigma_m \).

Then

\[
|\langle \nu_m(y + y', I + I'), K \rangle| \leq n \left| \Lambda''_{yy} |K| \sqrt{\epsilon} \sigma_m + n \epsilon h''_{yy} |K| \sqrt{\epsilon} \sigma_m + n \epsilon h''_{yi} |K| \sqrt{\epsilon} \sigma_m + n \epsilon h''_{yi} |K| \sqrt{\epsilon} \sigma_m + n \epsilon h''_{yi} |K| \sqrt{\epsilon} \sigma_m \right.
\]

\[
\leq \sqrt{\epsilon} C_{\omega} N_m \sigma_m,
\]

where \( C_{\omega} = C_{\omega}(c_\lambda, c'_\lambda, c''_\lambda, n) \).
Consider the extension of $Q_{K,m}^\ast$

\[ Q_{K,m}^\ast = \{ (\omega_1, \omega_2) \in \mathcal{V}_m(D_m) : \left| \{ \omega_1, k \} + \sqrt{\varepsilon} \omega_2 k_0 \right| \leq (2\lambda_{\beta[K]} + C_\psi N_m \sigma_m) \sqrt{\varepsilon} \}. \]  

(11.4)

Note that $(D_m \cap U_{a_n}(Q_{K,m})) \subseteq \bar{v}_{m}^{-1}(Q_{K,m}^\ast)$. Finally,

\[ \mu(D_m \cap U_{a_n}(Q_{K,m})) \leq \mu(\bar{v}_{m}^{-1}(Q_{K,m}^\ast)) \leq 4(\lambda_{\beta[K]} + C_\psi N_m \sigma_m) \sqrt{\varepsilon} C_j \bar{C}_j C_D. \]  

(11.5)

We have $(D_m \setminus D_{m+1}) \subseteq \bigcup_{N_{m-1} < |\lambda| < N_m} (D_m \cap U_{a_n}(Q_{K,m})).$ Using (11.5) we obtain

\[ \mu(D_0 \setminus D_{m+1}) \leq \sum_{i=0}^{m} \mu \left( \bigcup_{N_{i-1} < |\lambda| < N_i} (D_i \cap U_{a_n}(Q_{K,i})) \right) \leq 4 \sqrt{\varepsilon} C_j \bar{C}_j C_D \sum_{i=0}^{m} N_{i+1}^{r+1}(\lambda, + C_\psi N_i \sigma_i). \]

The proposition holds for $c_\mu = 4C_j \bar{C}_j C_D \sum_{i=0}^{m} N_{i+1}^{r+1}(\lambda, + C_\psi N_i \sigma_i)$. To finish the proof we need to check

\[ \sum_{i=0}^{+\infty} N_{i+1}^{r+1}(\lambda, + C_\psi N_i \sigma_i) < +\infty. \]  

(11.6)

Using (10.1) and (10.2) we obtain

\[ \sum_{i=0}^{+\infty} N_{i+1}^{r+1}(\lambda, + C_\psi N_i \sigma_i) = \sum_{i=0}^{+\infty} \left( C_N \binom{N}{c_\lambda k} N^2 + 1 \right) a_0 C_\psi N_N \sigma_i < +\infty. \]  

11.3. Proof of proposition 8.1

The solution of equation (8.3) has the form

\[ S = \sum_{|K| \leq N_m, k \neq 0} \hat{S}^K e^{iK(x_i) + ik_0 \psi}, \quad \hat{S}^K = \frac{i\sqrt{\varepsilon} V_m^K(\hat{y}, \hat{I})}{\nu_m(\hat{y}, \hat{I}, K)}. \]  

(11.7)

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

The first inequality (7.7) means that

\[ |V_m^K e^{iK(x_i) + ik_0 \psi}|_{|a_m|, b_m, \delta} \leq 2a_m e^{-|x_i| + |k_0| + |\sigma|}, \quad 0 \leq \delta \leq b_m. \]

Then

\[ |S|_{|a_m|, b_m, \delta} \leq \sum_{j=0}^{m} \sum_{|K| \leq N_m, k \neq 0} \frac{2a_m}{|\lambda_j|} e^{-|\lambda_j| N_{j-1}} \leq L_m s_m. \]  

11.4. Proof of proposition 8.2

In this section for brevity we write $V, h$ instead of $V_m, h_m$ and $a, b, \sigma, \delta, N, L, s$ instead of $a_m, b_m, \sigma_m, \delta_m, N_m, L_m, s_m, \delta_m$.

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

\[ \tilde{V}_m(\hat{y}, \hat{I}, \hat{\psi}) = R_1 + R_2 + R_3 + R_4 + R_5, \]  

(11.8)

\[ R_1 = \frac{1}{E} \left( \Lambda(y) - \Lambda(\hat{y}) - \langle \Lambda'_\psi(y), \hat{\psi} S'_\psi \rangle \right), \]

\[ R_2 = h(y, I) - h(\hat{y}, \hat{I}) - \langle h'_\psi(y), \hat{I} \rangle, \]

\[ R_3 = V(y, I, x, \psi) - V(\hat{y}, \hat{I}, x, \psi), \]

\[ R_4 = V(\hat{y}, \hat{I}, x, \psi) - \langle V'_\psi(y), \hat{I}, \psi \rangle - \Pi_N V(\hat{y}, \hat{I}, x, \psi), \]

\[ R_5 = \langle V'_\psi(\hat{y}, \hat{I}, \psi) - \langle V'_\psi(\hat{y}, \hat{I}, \psi) \rangle. \]

2124
By proposition (8.1) the first term in (11.8) satisfies the estimate
\[ |R_1|_{a_1, a_2, b, 2b - 2a} \leq \frac{1}{2} \left| \Lambda_{\delta}'' \right| \left| S_{\delta}' \right|^2 \leq \frac{c_\Lambda}{2} \left( \frac{L_s}{\delta} \right)^2. \]
To estimate the second one we use (7.7), (7.9):
\[ |R_2|_{a_1, a_2, b, 2b - 2a} \leq \frac{1}{2} \left( |\varepsilon h''' y| \Lambda_{\delta}'' \left| S_{\delta}' \right|^2 \right) \leq \frac{c_\varepsilon}{2} \left( \frac{L_s}{\delta} \right)^2 + 2n \left( \frac{L_s}{\delta} \right)^2 + c_h \left( \frac{L_s}{\delta} \right)^2. \]
The third term is estimated by (7.7):
\[ |R_3|_{a_1, a_2, b, 2b - 2a} \leq \sqrt{e} \left( \frac{C}{\delta} \right)^n e^{-N^2 \delta}. \]
By (8.2)
\[ |R_4|_{a_1, a_2, b, 2b - 2a} \leq \frac{C}{\delta} \left( \frac{N + 1}{\delta} \right)^n e^{-N^2 \delta}. \]
Finally
\[ |R_5|_{a_1, a_2, b, 2b - 2a} \leq \frac{L_s^2}{\sigma \delta}. \]
Note that \( 3 n c_h^2 \leq \frac{c_h}{\sigma} \). Therefore
\[ \tilde{u}_m + \tilde{v}_m \leq \tilde{s}, \]
where
\[ \tilde{s} = (c_h + c_h^2) \left( \frac{L_s}{\delta} \right)^2 + (n + 1) \left( \frac{L_s^2}{\sigma \delta} \right) + C \left( \frac{N + 1}{\delta} \right)^n e^{-N^2 \delta}. \]

11.5. Proof of proposition 9.1

By (6.2), (6.3), (7.7), (7.8), and (8.8), for any \( m \geq 1 \) we have:
\[
\begin{align*}
|v_0|_{a_m} &\leq c_h, \\
|v_0|_{a_m, b_n} &\leq \sqrt{\varepsilon c}, \\
|v_m|_{a_m} &\leq 2c_h, \\
|v_m|_{a_m} &\leq 2c_h, \\
|v_m|_{a_{m-\sigma_0, b_{m-2a}}} &\leq \tilde{s}_m.
\end{align*}
\]
Applying lemma 12.2 to equation (9.1), we obtain
\[
\begin{align*}
|\tilde{S}_\psi|_{a_1 - \tilde{S}_\psi} &\leq 8 c_h \left| v_m \right|_{a_m - \sigma_0, b_{m-2a}} \leq 16 \varepsilon c_h |v_m|_{a_m - \sigma_0, b_{m-2a}}, \\
|\tilde{S}_\psi|_{a_1 - \tilde{S}_\psi} &\leq \sqrt{\varepsilon} \left| v_m \right|_{a_m - \sigma_0, b_{m-2a}} \leq \frac{16 \varepsilon c_h}{\sigma_m} |v_m|_{a_m - \sigma_0, b_{m-2a}}.
\end{align*}
\]
Then estimates (9.5) follow from (11.9), (11.10) and (11.12). We have
\[
|v_m|_{a_1 - \tilde{S}_\psi} + u_{m+1} (\tilde{y}, \tilde{\hat{I}}, \tilde{\hat{x}}, \tilde{\hat{\varphi}}) = \tilde{u}_m (\tilde{y}, \tilde{\hat{I}} + \tilde{\hat{S}}_\psi, \tilde{\hat{\varphi}} - \tilde{S}_\psi) + \tilde{u}_m (\tilde{y}, \tilde{\hat{I}} + \tilde{\hat{S}}_\psi, \tilde{\hat{x}} - \sqrt{\varepsilon} \tilde{S}_\psi, \tilde{\hat{\varphi}} - \tilde{S}_\psi). \]
Since by (11.11)
\[
\begin{align*}
|\tilde{S}_\psi|_{a_1, b_{m+1} + \tilde{S}_\psi} &\leq 8 c_h s_m \leq \sigma_m, \\
|\sqrt{\varepsilon} \tilde{S}_\psi|_{a_1, b_{m+1} + \tilde{S}_\psi} &\leq \frac{16 \varepsilon c_h}{\sigma_m} s_m \leq \delta_m, \\
|\tilde{S}_\psi|_{a_1, b_{m+1} + \tilde{S}_\psi} &\leq \frac{16 \varepsilon c_h}{\sigma_m} s_m \leq \delta_m,
\end{align*}
\]
we have
\[ |v_{m+1} + u_{m+1}|_{a_s, b_{m+1}} \leq s_m \leq s_{m+1}. \]
11.6. Proof of proposition 10.1
By equations (8.4) and (10.2) we have
\[ L_m \leq \sum_{j=0}^{m} \frac{2N^{n+1}}{\lambda_j} \leq \frac{2N^{n+1}}{c_\lambda} \sum_{j=0}^{m} 2(4n+5j) < \frac{c_{n+1}}{c_\lambda} 2^{(4n+5)(m+2)}. \]

11.7. Proof of proposition 11.1
A point \( (y, I) \in Q_{K,m+1} \) if
\[ \left| \langle \nu_{m+1}(y, I), K \rangle \right| \leq \lambda_j (|K|)(1 + 2^{-m-2}) \sqrt{\varepsilon}. \]
It is sufficient to show that \( (y, I) \in Q_{K,m} \), i.e.
\[ \left| \langle \nu_{m}(y, I), K \rangle \right| \leq \lambda_j (|K|)(1 + 2^{-m-1}) \sqrt{\varepsilon}. \] (11.13)
We have the inequality
\[ \left| \langle \nu_{m}(y, I), K \rangle - \langle \nu_{m+1}(y, I), K \rangle \right| \leq (n+1) |K| \varepsilon. \]
By (9.5)
\[ W \leq \varepsilon |h_{m+1}^{y} - h_{m}^{y}|_{i \neq m+1} + \sqrt{\varepsilon} |h_{m+1}^{I} - h_{m}^{I}|_{i \neq m+1} \leq \sqrt{\varepsilon} \frac{16c_\lambda c_{\nu} s_{m'}}{\sigma_{m}^2}. \]
Therefore, for any \(|K| \leq N_m\) we have the estimate
\[ (n+1) |W|_{K} \leq (n+1) \frac{16c_\lambda c_{\nu} s_{m'}}{\sigma_{m}^2} N_m, \]
where \( s_{m'} = \sqrt{\varepsilon} \) and \( s_{m'} = s_m \) for all \( m \geq 1 \).
It remains to check the estimate
\[ (n+1) \frac{16c_\lambda c_{\nu} s_{m'}}{\sigma_{m}^2} N_m \leq \lambda_j 2^{-m-2}. \]
Let \( c_\lambda > c_j (c_\nu, c_N) \). For \( m = 0 \) we have
\[ (n+1) \frac{16c_\lambda c_{\nu} s_{0}'}{\sigma_{0}^2} N_0 \leq \lambda_0 2^{-2}, \]
and for \( m \geq 1 \)
\[ (n+1) \frac{36c_\lambda c_{\nu} s_{m'} c_N}{\sigma_{m}^2} 2^{4n+14} |c| (4n+12)m e^{-c_{m}-2^m} \leq c_{\lambda} 2^{-(2n+2+\tau)m}. \]

11.8. Proof of proposition 11.2
Suppose that the arguments of functions \( \Lambda, h \) lie in \( U_{am}(D_m) \). Let us expand the Jacobian \( J_m \) with respect to the last column
\[ \det J_m(y, I) = \det \left( \Lambda''_{yy} + \varepsilon h''_{myy} \sqrt{\varepsilon} h''_{mly} \right) \]
\[ \det (A''_{yy} + \varepsilon h''_{myy}) h''_{ml} + \sum_{i=1}^{n} (-1)^{n+1+i} \sqrt{\varepsilon} h''_{mlyi} M_{i,n+1}. \]
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
\[ \left| \det J_m(y, I) \right| \leq n! \left( \left| \Lambda''_{yy} \right| + \left| \varepsilon h''_{myy} \right| \right) \left| h''_{ml} \right| + n n! \sqrt{\varepsilon} \left( \left| \Lambda''_{yy} \right| + \left| \varepsilon h''_{myy} \right| \right)^{n-1} \leq C_{\nu} J, \]
where \( C_{\nu} J \) is some constant depending on \( c_\lambda, c_{\nu}', c''_{\nu}, c''_h \) and \( n \).
Nonlinearity 28 (2015) 2105

Return to the estimate for the Jacobian. For sufficiently small \( c''_h \)
\[
c''_h \leq c''_1 = \min \left( \frac{c_A}{2n!}, \frac{c_A}{2} \right), \quad \left| \varepsilon h''_{m,yy} \right| \leq 2c''_h
\]
we have
\[
\left| \det \left( \Lambda''_{yy} + \varepsilon h''_{m,yy} \right) \right| > \left| \det \Lambda''_{yy} \right| - n n! |\varepsilon h''_{m,yy}| \left( |\Lambda''_{yy}| + |\varepsilon h''_{m,yy}| \right)^{n-1} \geq \frac{1}{2} \Lambda_1.
\]
Note that \( |h''_{m,II}| \geq \frac{1}{2} c'h \) and for
\[
c''_h \leq c''_2 = \min \left( \frac{c_A}{2n+4n!}, c''_1 \right)
\]
we have
\[
\left| \sum_{i=1}^{n} (-1)^{n+1+i} \sqrt{\varepsilon h''_{m,II}} M_{i,n+1} \right| \leq \frac{c_A}{8c'_h}.
\]
Finally
\[
|\det J(y, I)|_{\infty} \geq \frac{c_A}{8c'_h} = C^{-1}.
\]

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 \( \delta \in (0, b) \)
\[
|f - \langle f \rangle_x - \Pi_N f|_{b-\delta} \leq \frac{C}{\delta} \left( N + \frac{1}{\delta} \right)^n e^{-N\delta} |f|_b.
\]
where the constant \( C \) depends only on \( n \).

**Proof.** The Fourier coefficients (8.1) satisfy the inequalities
\[
|f^K| \leq e^{-b|K|} |f|_b.
\]
Then the equation
\[
f - \langle f \rangle_x - \Pi_N f = \sum_{|K|>N, k \neq 0} f^K e^{i\langle k, x \rangle + k_{\text{op}}}
\]
implies
\[
|f - \langle f \rangle_x - \Pi_N f|_{b-\delta} \leq |f|_b \sum_{|K|>N, k \neq 0} e^{-\delta|K|}.
\]
The sum on the right-hand side does not exceed
\[
c_1 \int_{\varepsilon \in \mathbb{R}^{m+1}, |\varepsilon|>N} e^{-\delta|\varepsilon|} \, d\varepsilon \leq \frac{c_2}{\delta^{m+1}} \int_{\delta N}^{\infty} s^n e^{-s} \, ds \leq \frac{c_3}{\delta^{m+1}} (1 + \delta N)^n e^{-\delta N},
\]
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 neighbourhoods of \([-\alpha, \alpha]\) and \([-\alpha, \alpha] \times \mathbb{T}\), respectively. Let the canonical change \((I, \varphi \mod 2\pi) \mapsto (\bar{I}, \bar{\varphi} \mod 2\pi)\), determined by the generating function \(\bar{I} \varphi + S(\bar{I}, \varphi)\), \(\langle S\rangle = 0\), be such that

\[
\begin{align*}
  h(I) + v(I, \varphi) &= h(\bar{I}), \quad (12.1) \\
  |h|_{\alpha} &\leq c, \quad |h'|_{\alpha}^{-1} \leq c', \quad |v|_{\alpha, \beta} \leq \frac{\sigma}{2c^2}, \quad 0 < \sigma < a/2. \quad (12.2)
\end{align*}
\]

Then

\[
\begin{align*}
  |S'|_{\alpha-3\sigma, \beta} &\leq 4c' |v|_{\alpha, \beta}, \quad |S|_{\alpha-3\sigma, \beta} \leq 8\pi c' |v|_{\alpha, \beta}, \quad |h - h_s|_{\alpha-3\sigma} \leq 2cc' |v|_{\alpha, \beta}. \quad (12.3)
\end{align*}
\]

**Proof.** The equation that determines \(\bar{I}\) is well known:

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

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:

\[
I = r + f(r, \varphi), \quad |f|_{\alpha-2\sigma, \beta} \leq 2c' |v|_{\alpha, \beta}. \quad (12.5)
\]

Moreover,

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

Equations (12.4) and (12.5) imply

\[
\begin{align*}
  I(r, \varphi) - \bar{I}(r) &= f(r, \varphi) - \langle f\rangle(\varphi)(r), \quad (12.7) \\
  r &= r(\bar{I}), \quad I - \bar{I} = S'_\varphi(\bar{I}, \varphi). \quad (12.8)
\end{align*}
\]

Combining (12.7) and (12.8), we obtain

\[
S(\bar{I}, \varphi) = \int_0^{\varphi} (I(r, \varphi) - \bar{I}) \, d\varphi = \int_0^{\varphi} (f(r, \varphi) - \langle f\rangle(\varphi)(r)) \, d\varphi.
\]

Therefore

\[
|S'_\varphi|_{\alpha-3\sigma, \beta} \leq 4c' |v|_{\alpha, \beta}, \quad |S|_{\alpha-3\sigma, \beta} \leq 8\pi c' |v|_{\alpha, \beta}.
\]

By using the equation \(h(r) = h_s(\bar{I})\), we have

\[
|h(I) - h_s(I)|_{\alpha-3\sigma} \leq |h(r + f(r, \varphi)) - h(r)|_{\alpha-2\sigma} \leq |h'|_{\alpha-\sigma} |f|_{\alpha-2\sigma, \beta} \leq \frac{c}{\sigma} 2c' |v|_{\alpha, \beta}. \quad \square
\]
**12.3. A version of the implicit function theorem**

**Lemma 12.3.** Let the real-analytic functions \( h, v \), defined in a complex neighbourhood of the interval \( I \subset \mathbb{R} \), satisfy the estimates

\[
|h'|_{a}^{-1} \leq c', \quad |v|_{a} \leq \frac{\sigma}{2c'}, \quad 0 < \sigma < \frac{a}{2}.
\]

Then the equation

\[
h(I) + v(I) = h(r), \quad I \in U_{a-\sigma}(I)
\]

implies

\[
I = r + f(r), \quad |f|_{a-2\sigma} \leq 2c'|v|_{a} \leq \sigma,
\]

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

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

\[
I + u(I) = r, \quad u(I) = h^{-1}(h(I) + v(I)) - I.
\]

If \( |v|_{a} \leq \sigma/c' \), the function \( u \) is defined in a complex neighbourhood of \( I \) and admits the estimate

\[
|u|_{a-\sigma} \leq |h'|_{a}^{-1}|v|_{a} \leq c'|v|_{a}.
\]

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)

\[
|\Phi'(I)|_{a-2\sigma} = |u'|_{a-2\sigma} \leq \frac{c'|v|_{a}}{\sigma} \leq \frac{1}{2}.
\]

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

---

**Acknowledgments**

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] Arnold V I 1963 Small denominators and problems of stability of motion in classical and celestial mechanics *Russ. Math. Surv.* **18** 85–191

[2] Arnold V I, Kozlov V V and Neishtadt A I 2006 Mathematical aspects of classical and celestial mechanics *Dynamical Systems* vol III, 3rd edn (*Encyclopaedia of Mathematical Sciences* vol 3) (Berlin: Springer)

[3] Arnold V I 1978 *Mathematical Methods of Classical Mechanics* (Graduate Texts in Mathematics vol 60) (New York: Springer)

[4] Cong F, Kupper T, Li Y and You J 2000 KAM-type theorem on resonant surfaces for nearly integrable Hamiltonian systems *J. Nonlinear Sci.* **10** 49–68

[5] Delshams A, de la Llave R and Seara T M 2006 A geometric mechanism for diffusion in Hamiltonian systems overcoming the large gap problem: heuristics and rigorous verification on a model *Mem. Am. Math. Soc.* **179** 141

[6] Froeschlé CI 1972 Numerical study of a four-dimensional mapping *Astron. Astrophys.* **16** 172–89

---

2129
[7] Lazutkin V F 1974 To Moser’s theorem on invariant curves Vopr. Dinamich. Teor. Rasprostr. Seism. Voln. 14 109–120 (in Russian)

[8] Li Y and Yi Y 2003 A quasi-periodic Poincare’s theorem Math. Ann. 326 649–90

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

[10] Neishtadt A I 1982 Estimates in the Kolmogorov theorem on conservation of conditionally periodic motions J. Appl. Math. Mech. 45 766–72

[11] Poincaré H 1892 Les méthodes nouvelles de la mécanique céleste vol 1 (Paris: Gauthier-Villars)

Poincaré H 1893 Les méthodes nouvelles de la mécanique céleste vol 2 (Paris: Gauthier-Villars)

Poincaré H 1899 Les méthodes nouvelles de la mécanique céleste vol 3 (Paris: Gauthier-Villars)

[12] Pöschel J 1982 Integrability of Hamiltonian systems on Cantor sets Commun. Pure Appl. Math. 35 653–95

[13] Pronin A V and Treschev D V 1997 On the inclusion of analytic maps into analytic flows Regular Chaot. Dyn. 2 14–24

[14] Randall P 1998 A KAM theorem for some degenerate Hamiltonian systems MPARC: 98–19

[15] Sevryuk M B 1995 KAM-stable Hamiltonians J. Dyn. Control Syst. 1 351–66

[16] Simó C 2000 Dynamical properties of the figure eight solution of the three-body problem Proc. Celestial Mechanics Conf. dedicated to D Saari for his 60th birthday (Evanston, IL, 1999) ed A Chenciner et al (Contemporary Mathematics vol 292) (Providence, RI: American Mathematical Society) pp 209–28

[17] Svanidze N V 1981 Small perturbations of an integrable dynamical system with an integral invariant Proc. Steklov Inst. Math. 147 127–51

[18] Treschev D 1991 A mechanism for the destruction of resonance tori in Hamiltonian systems Math. USSR-Sb. 68 181–203

[19] Vaidya U and Mezić I 2012 Existence of invariant tori in three dimensional maps with degeneracy Physica D 241 1136–45