J. Math. Pures Appl. 80, 10 (2001) 1045–1067

S0021-7824(01)01221-1/FLA

PERSISTENCE OF LOWER-DIMENSIONAL TORI UNDER THE FIRST MELNIKOV'S NON-RESONANCE CONDITION ☆

Junxiang XU [a], **Jiangong YOU** [b]

[a] *Department of Applied Mathematics, Southeast University, Nanjing 210018, PR China*
[b] *Department of Mathematics, Nanjing University, Nanjing 210093, PR China*

Manuscript received 09 April 2001

ABSTRACT. – In this paper we prove the persistence of lower-dimensional invariant tori of integrable equations after Hamiltonian perturbations under the first Melnikov's non-resonance condition. The proof is based on an improved KAM machinery which works for the angle variable dependent normal form. By an example, we also show the necessity of the Melnikov's first non-resonance condition for the persistence of lower dimensional tori. © 2001 Éditions scientifiques et médicales Elsevier SAS

Keywords: Invariant tori, KAM iteration, Hamiltonian systems, Small divisor conditions

1. Introduction

The dynamics of integrable Hamiltonian systems is simple in the sense that all the compact energy surface are foliated by invariant tori which carries quasiperiodic motions of the corresponding Hamiltonian equations. But integrable Hamiltonian systems are rather rare in the whole family of Hamiltonian systems. One of the landmarks in dynamical systems, especially in Hamiltonian dynamical systems, is the KAM (Kolmogorov–Arnold–Moser) theory, which discovered that, for all Hamiltonian systems in an open neighborhood of a nondegenerate integrable Hamiltonian systems, the quasi-periodic motions in invariant tori are typical (See Arnold [1] and the references therein). Later Melnikov [8] formulated a KAM type persistence result for lower-dimensional tori of nearly integrable Hamiltonian systems. His result is based on infinite many non-resonance conditions, each involving tangential frequencies and two of normal frequencies. In recent years, problems in construction of quasi-periodic solutions of Hamiltonian partial differential equations and the study of dynamics in the resonant zone of nearly integrable systems ask for a KAM theory under weaker restriction on the frequencies. Our aim in this paper is to prove the persistence of lower-dimensional tori under the so called first Melnikov's non-resonance condition, in which at most one of the normal frequencies is involved.

☆ The work was supported by the National Natural Science Foundation of China (19925107, 19701007) and the Special Funds for Major State Basic Research Projects (973 projects).

E-mail addresses: xujun@seu.edu.cn (J. Xu), jyou@nju.edu.cn (J. You).

Consider a family of Hamiltonian systems of the form:

$$\begin{aligned} &\dot{x} = \omega(\xi) + P_y, \qquad \dot{y} = -P_x, \\ &\dot{u}_j = \Omega_j(\xi)v_j + P_{v_j}, \qquad \dot{v}_j = -\Omega_j(\xi)u_j - P_{u_j}, \end{aligned} \tag{1.1}$$

with the Hamiltonian

$$H = \sum_{1\leqslant j\leqslant n} \omega_j(\xi)y_j + \frac{1}{2}\sum_{j=1}^{m} \Omega_j(\xi)\big(u_j^2 + v_j^2\big) + P,$$

defined in the symplectic space $(x, y, u, v) = (x_1, \ldots, x_n; y_1, \ldots, y_n; u_1, \ldots, u_m; v_1, \ldots, v_m) \in (T^n \times R^n \times R^m \times R^m, \sum_{j=1}^n \mathrm{d}x_j \wedge \mathrm{d}y_j + \sum_{j=1}^m \mathrm{d}u_j \wedge \mathrm{d}v_j)$, $1 \leqslant n, m < +\infty$; T^n is the usual n-torus; $P = P(x, y, u, v; \xi)$ is a small perturbation; $\omega(\xi) = (\omega_1(\xi), \omega_2(\xi), \ldots, \omega_n(\xi)) \subset R^n$ are called tangential frequencies depending on a parameter $\xi \in O$; O is a bounded open set of R^l. $\Omega_1(\xi), \Omega_2(\xi), \ldots, \Omega_m(\xi)$ are called normal frequencies usually depending on ξ.

For our convenience, we use the complex conjugated coordinates $z = (u + \mathrm{i}v)/\sqrt{2}$, $\bar{z} = (u - \mathrm{i}v)/\sqrt{2}$, $\mathrm{i} = \sqrt{-1}$. The Hamiltonians then read as

$$H = H(x, y, z, \bar{z}; \xi) = N + P = \big\langle \omega(\xi), y\big\rangle + \sum_{j=1}^{m} \Omega_j(\xi)z_j\bar{z}_j + P(x, y, z, \bar{z}, \xi). \tag{1.2}$$

$N = \langle \omega(\xi), y\rangle + \sum_{j=1}^m \Omega_j(\xi)z_j\bar{z}_j$ is usually called the normal form part while $P = P(x, y, z, \bar{z}; \xi)$ is called the perturbation. $x = (x_1, x_2, \ldots, x_n) \in T^n$ and $y = (y_1, y_2, \ldots, y_n) \in R^n$ are the conjugate angular variables and the action variables, $z = (z_1, z_2, \ldots, z_m) \in C^n$ and $\bar{z} = (\bar{z}_1, \bar{z}_2, \ldots, \bar{z}_m) \in C^n$ are a pair of conjugate complex variables, the parameter $\xi \in O \subset R^l$. The symplectic form in $(x, y, z, \bar{z})$-space is $\sum_{j=1}^n \mathrm{d}x_j \wedge \mathrm{d}y_j + \mathrm{i}\sum_{j=1}^m \mathrm{d}z_j \wedge \mathrm{d}\bar{z}_j$. The Hamiltonian systems are written as

$$\dot{x} = H_y, \qquad \dot{y} = -H_x, \qquad \dot{\bar{z}} = \mathrm{i}H_z, \qquad \dot{z} = -\mathrm{i}H_{\bar{z}}. \tag{1.3}$$

We assume that H is analytic [1] in $(x, y, z, \bar{z}, \xi)$. If $P = 0$, the Hamiltonian system at each $\xi \in O$ is integrable and has an invariant torus $T^n \times \{0\} \times \{0\} \times \{0\}$ with the frequency $\omega(\xi)$. When P is sufficiently small, the persistence of invariant tori has been extensively studied by many authors. In the sixties, Melnikov [8,9] announced that if $\omega(\xi)$ and Ω satisfy the non-resonance conditions, [2, 3]

$$\big\langle \omega(\xi), k\big\rangle \neq 0, \tag{1.4}$$

[1] Analyticity in ξ is not essential, actually, it is enough to assume the C^L Whitney-smooth in the parameter ξ. In Whitney smooth case, (1.4)–(1.7) are assume to be satisfied "almost every where" in Lebesgue measure sense.

[2] Melnikov's original formulation is:

$$\langle \omega(\xi), k\rangle + \langle \Omega(\xi), l\rangle > \frac{\alpha}{|k|^\tau},$$

for $k \in Z^n$, $l \in Z^m$, $|l| \leqslant 2$, $|k| + |l| \neq 0$. In analytic case, assumptions (1.4)–(1.7) imply that the above non-resonance conditions hold for majority of parameters of Lebesgue measure $(1 - O(\alpha))\mathrm{meas}\,O$. In Melnikov's papers, $\omega(\xi)$ is simply $\xi \in O \subset R^n$. For general case, we refer to Rüssmann [12], Sevryuk [13] and Xu, You and Qiu [17] for details.

[3] In (1.4)–(1.7), "$\neq 0$" means "not identical 0" as a function of ξ.

(1.5) $$\langle\omega(\xi),k\rangle+\Omega_j(\xi)\neq 0,$$

(1.6) $$\langle\omega(\xi),k\rangle-\Omega_i(\xi)-\Omega_j(\xi)\neq 0,\quad |k|+|i-j|\neq 0,$$

(1.7) $$\langle\omega(\xi),k\rangle+\Omega_i(\xi)-\Omega_j(\xi)\neq 0,\quad |k|+|i-j|\neq 0,$$

for all $\xi\in O$, $\forall k\in Z^n$, $\forall i,j=1,2,\dots,m$, then for the vast majority of the parameters ξ in Lebesgue measure sense, H at ξ possesses a linearly stable invariant torus. This result was proved in detail later independently by Eliasson [5], Kuksin [6]. Kuksin and Pöschel also generalized the result to infinite-dimensional Hamiltonian systems [10,11,7]. The condition (1.4) is a version of standard non-resonant condition of KAM theory, which is certainly necessary. The non-resonant condition (1.5) is generally referred as the first Melnikov's condition, while (1.6) and (1.7) are referred as the second Melnikov's non-resonance condition. In their proofs, besides (1.4) both the first Melnikov's non-resonance condition (1.5) and the second Melnikov's non-resonance condition (1.6) + (1.7) are needed.

If all normal frequencies Ω_i are independent of ξ, Bourgain proved the existence of quasiperiodic solutions for nearly integrable Hamiltonian systems (1.2) without the second Melnikov's condition (see [2,3]). His proof based on a method introduced by Craig and Wayne [4]. In [18] and [15] the authors improved the standard KAM machinery so as to prove a persistence result for the multiple normal frequency case, which is a special case that the second Melnikov's condition do not hold (more precisely, $|k|+|i-j|\neq 0$ is replaced by $|k|\neq 0$). The result applies to the constant normal frequency case considered by Bourgain [3]. Actually, for the constant normal frequency case, both the first and the second Melnikov's condition are satisfied for 'most' parameters ξ (in the sense of Lebesgue measure) for $k\neq 0$. In [16] we further show that if conditions (1.4), (1.5) and (1.6) hold, by a nonlinear symplectic map, the Hamiltonian can be reduced to the normally multiplicity case, then one can directly apply the result in [18] to get the persistence.

In this paper we will further prove the persistence of lower-dimensional tori without assuming the second Melnikov's non-resonance condition (1.6) and (1.7). The result is optimal since one can easily construct an example to show that the first Melnikov's non-resonance condition is necessary.

2. Main results

We first give some notations and assumptions. Denote a complex neighborhood of $T^n\times\{0\}\times\{0\}\times\{0\}$ by

$$D(s,r)=\left\{(x,y,z,\bar z)\mid |\mathrm{Im}\,x|\leqslant s,\ |y|\leqslant r^2,\ |z|_2\leqslant r,\ |\bar z|_2\leqslant r\right\},$$

where $|\mathrm{Im}\,x|=\max_{1\leqslant i\leqslant n}|\mathrm{Im}\,x_i|$, $|y|=\max_{1\leqslant i\leqslant n}|y_i|$, and $|\cdot|_2$ denotes the Euclidean norm.

If $P=P(x,y,z,\bar z;\xi)$ is analytic in $(x,y,z,\bar z,\xi)\in D(s,r)\times O$, then P can be expanded as

$$P=\sum_{k,l,q,\bar q}P_{k,l,q,\bar q}(\xi)\mathrm{e}^{\mathrm{i}\langle k,x\rangle}y^l z^q\bar z^{\bar q}.$$

Define [4]:

$$\|P\|_{s,r}^L=\sup_{|y|\leqslant r^2,|z|_2\leqslant r,|\bar z|_2\leqslant r}\left|\sum_{k,l,q,\bar q}\|P_{k,l,q,\bar q}(\xi)\|^L\mathrm{e}^{s|k|}y^l z^q\bar z^{\bar q}\right|.$$

[4] $\|\cdot\|^L$ denotes C^L norm.

Denote the Hamiltonian vector field of P by X_P, i.e., $X_P = (P_y, -P_x, -\mathrm{i}P_{\bar z}, \mathrm{i}P_z)$. Define $\|P_y\|_{s,r}^L = \max_{1\leqslant i\leqslant n}\|P_{y_i}\|_{s,r}^L$ and $\|P_z\|_{2;s,r}^L = (\sum_{j=1}^m (\|P_{z_j}\|_{s,r}^L)^2)^{1/2}$. $\|P_x\|_{s,r}^L$ and $\|P_{\bar z}\|_{2;s,r}^L$ are similarly defined. A weight norm of X_P is defined by:

$$\|X_P\|_{r;s,r}^L = \|P_y\|_{s,r}^L + \frac{1}{r^2}\|P_x\|_{s,r}^L + \frac{1}{r}\|P_{\bar z}\|_{2;s,r}^L + \frac{1}{r}\|P_z\|_{2;s,r}^L.$$

THEOREM 1. – *Suppose that Hamiltonian H in* (1.2) *is analytic in* $(x, y, z, \bar z, \xi) \in D(s,r) \times O$, $N = \langle\omega(\xi), y\rangle + \sum_{j=1}^m \Omega_j(\xi) z_j \bar z_j$ *satisfying the first Melnikov's non-resonance condition* (1.4) *and* (1.5). *Then for sufficiently small $\alpha > 0$, there exists $\varepsilon > 0$ depending on Ω, α, O, n, m such that if $\|X_P\|_{r;s,r}^L \leqslant \varepsilon$ with $L \geqslant m^2$, the following holds true: There exists a nonempty subset O_α of O such that for $\forall \xi \in O_\alpha$, there exists a real analytic symplectic map* [5]

$$\Phi(\cdot;\xi): D_{s/2,r/2} \to D_{s,r}$$

which transforms H into the following form

$$H \circ \Phi = N_* + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P_*,$$

where

$$N_* = \big\langle \omega_*(\xi), y\big\rangle + \sum_{j=1}^m \Omega_{*j}(\xi) z_j \bar z_j, \qquad \mathcal{A} = \sum_{i,j} a_{ij}(\xi) z_i \bar z_j,$$

$$\mathcal{B} = \sum_{i,j} b_{ij}(\xi) z_i z_j \mathrm{e}^{\mathrm{i}\langle \tilde k_{ij}, x\rangle}, \qquad \bar{\mathcal{B}} = \sum_{i,j} \bar b_{ij}(\xi) \bar z_i \bar z_j \mathrm{e}^{-\mathrm{i}\langle \tilde k_{ij}, x\rangle}$$

and

$$P_* = \sum_{2|l|+|q|+|\bar q|\geqslant 3} P_{*klq\bar q}\mathrm{e}^{\mathrm{i}\langle k,x\rangle} y^l z^q \bar z^{\bar q},$$

$\Omega_{*j} \in \{\Omega_1, \Omega_2, \ldots, \Omega_m\}$, $\forall 1 \leqslant j \leqslant m$. *Hence, for $\xi \in O_\alpha$, $\Phi(T^n, \xi)$ is an invariant torus of* (1.3) *with the frequency ω_* satisfying $\|\omega_*(\xi) - \omega(\xi)\|^L \leqslant 2\varepsilon$ and $|\langle\omega_*(\xi), k\rangle| > \frac{\alpha}{(|k|+1)^n}$. Moreover, we have* $\mathrm{mes}(O - O_\alpha) \leqslant c\alpha^{1/m^2}$, *where c is independent of α.*

Remark. – From the proof, we will see that $\mathcal{B}, \bar{\mathcal{B}}$ is zero when (1.5) and (1.6) are satisfied. $\mathcal{A}$ can be absorbed to $\sum_{j=1}^m \Omega_{*j}(\xi) z_j \bar z_j$ by a symplectic coordinator transformation, which makes the normal frequencies shift a little bit. The obtained torus in this case is linearly stable.

Remark. – Our Hamiltonian is somewhat more general than that considered by Bourgain [3] as we allow the dependence of normal frequencies on ξ. If the normal frequencies are independent of ξ, besides the persistence, we also obtain the linearly stability of the invariant torus by our approach. As for the proof, both Bourgain's and ours use a sequence of change of variables to reduce the perturbed Hamiltonian to $\langle\omega, y\rangle + \langle A(x)u, u\rangle + \mathrm{O}(y^2) + \mathrm{O}(yu) + \mathrm{O}(u^3)$ with $u = (z, \bar z)$, which has an invariant torus; The difference is, in Bourgain's approach, one kept all the x-dependent second-order terms unsolved and waved them to the normal form part, which surely makes the homological equations in KAM iteration steps more complicated. As a result, a multiscale analysis is needed to control the inverse of a linear operator. By his approach, the

[5] Φ is not analytic but Whitney smooth in ξ.

second Melnikov's non-resonance condition is avoided. In our approach, we observed that, after excluding a small set of parameters, most of the second Melnikov's non-resonance conditions are satisfied automatically although we assume only the first Melnikov's non-resonance condition. For example, if all the normal frequencies are constant (i.e., independent of ξ) as considered by Bourgain, the second Melnikov's conditions are satisfied for $k \neq 0$ outside a small set of parameters. This observation makes it possible to solve most of the x-dependent second-order terms of the form $a_{ij}(x)z_i z_j$, $a_{ij}(x)\bar{z}_i\bar{z}_j$, $a_{ij}(x)\bar{z}_i z_j$ in the homological equations. Only the terms corresponding to the resonances $\langle k, \omega(\xi)\rangle \pm 2\Omega_i(\xi) = 0$, which has finite many, have to be waved to the normal form part. As a consequence, the normal form part is much more simpler although it depends on x, and then the homological equations are easier to solve in the KAM iteration steps. In fact, the homological equations can be reduced to some elementary linear algebraic equtions with an uniform dimension bound at each KAM step. This approach avoids the multiscale analysis for bounding the corresponding Green's functions.

Consider a Hamiltinian with a special normal form part of the following:

$$H = N + P = \langle \omega(\xi), y\rangle + \sum_{j=1}^{m} \langle \Omega_j \tilde{I}_j z_j, \bar{z}_j\rangle + P, \tag{2.1}$$

where z_j and $\bar{z}_j$ are d_j vectors, and

$$\tilde{I}_j = I_j \quad \text{for } 1 \leqslant j \leqslant \bar{m}, \qquad \tilde{I}_j = \operatorname{diag}(I_{j1}, -I_{j2}) \quad \text{for } \bar{m} < j \leqslant m$$

are d_j-order matrices with I_j being the unit matrices of order d_j, where the unitary matrix I_{j1} or I_{j2} is allowed to be zero-order matrix and in this case $\tilde{I}_j = I_{j1}$ or $I_j = -I_{j2}$. Moreover, $\Omega_1, \Omega_2, \ldots, \Omega_m$ satisfy:

$$\langle \omega(\xi), k\rangle \neq 0, \tag{2.2}$$

$$\langle \omega(\xi), k\rangle + \Omega_j(\xi) \neq 0, \tag{2.3}$$

$$\langle \omega(\xi), k + k_j\rangle - \Omega_i(\xi) - \Omega_j(\xi) \neq 0, \quad |k| + |i - j| \neq 0, \tag{2.4}$$

$$\langle \omega(\xi), k\rangle + \Omega_i(\xi) - \Omega_j(\xi) \neq 0, \quad |k| + |i - j| \neq 0, \tag{2.5}$$

for all $\xi \in O$, $\forall k \in Z^n$, $\forall i, j = 1, 2, \ldots, m$ with some fixed $k_1, \ldots, k_{\bar{m}}$; $k_{\bar{m}+1} = 0, \ldots, k_m = 0$.

We will give a detail proof for the following:

THEOREM 2. – *Suppose that $H = N + P$, with the special form* (2.1) *satisfying* (2.2)–(2.5), *is real analytic in $(x, y, z, \bar{z}, \xi) \in D(s, r) \times O$. Then for sufficiently small $\alpha > 0$, there exists $\varepsilon = \varepsilon(\alpha) > 0$ such that if $\|X_P\|^L_{r;s,r} \leqslant \varepsilon$ with $L = \max_{1 \leqslant j \leqslant m} d_j^2$, the following holds true*: *There exists a nonempty subset O_α of O and for $\forall \xi \in O_\alpha$, there exists a real analytic symplectic map*

$$\Phi(\cdot; \xi) : D_{s/2, r/2} \to D_{s,r}, \quad \|\Phi - \mathrm{Id}\|^L \leqslant c\varepsilon$$

which transforms H into the following

$$H \circ \Phi = N_* + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P_*,$$

where

$$N_* = \langle \omega_*(\xi), y\rangle + \sum_{j=1}^{m} \langle \Omega_j(\xi) \tilde{I}_j z_j, \bar{z}_j\rangle, \qquad \mathcal{A} = \sum_{j=1}^{m} \langle A_j(\xi) z_j, \bar{z}_j\rangle,$$

$$\mathcal{B} = \sum_{j=1}^{m} B_j(\xi)\langle z_j, z_j\rangle e^{i\langle k_j, x\rangle}, \qquad \bar{\mathcal{B}} = \sum_{j=1}^{m} \langle \bar{B}_j(\xi)\bar{z}_j, \bar{z}_j\rangle e^{-i\langle k_j, x\rangle}$$

and

$$P_* = \sum_{2|l|+|q|+|\bar{q}|\geqslant 3} P_{*klq\bar{q}} e^{i\langle k,x\rangle} y^l z^q \bar{z}^{\bar{q}}.$$

Hence, for $\xi \in O_\alpha$, $\Phi(T^n, \xi)$ *is an invariant torus of* (2.1) *at* ξ *with the frequency* ω_* *satisfying* $\|\omega_* - \omega(\xi)\|^L \leqslant 2\varepsilon$ *and* $|\langle \omega_*(\xi), k\rangle| > \frac{\alpha}{(|k|+1)^n}$. *Moreover, we have* $\operatorname{mes}(O - O_\alpha) \leqslant c\alpha^{1/L}$, *where* c *is independent of* α.

Remark. – If $\bar{m} = 0$, i.e., all $k_j = 0$ and $\dim \tilde{I}_j = 1$ for all j, it is exactly the Melnikov Theorem proved by Kuksin and Eliasson. The case $\bar{m} = 0$ is the multiple normal frequency case considered in [18]. If $\bar{m} \neq 0$,

$$2\Omega_j = \langle k_j, \omega(\xi)\rangle, \quad 0 \neq k_j \in Z^n,\ j = 1, 2, \ldots, \bar{m},$$

are the trouble makers of our KAM iteration since some of x-dependent terms corresponding to this kind of resonance have to be moved to the normal form part.

Remark. – The above theorem implies that the obtained torus is linearly stable if $\bar{m} = 0$ and each $\tilde{I}_j$, $j = 1, \ldots, m$, is positive definite or negative definite.

We shall give a proof for the case $\omega(\xi) = \xi$, since there is no essential difficulties for general case if one incorporates the ideas and techniques of Rüssmann [12] and Xu, You and Qiu [17].

The paper is arranged as follows: we first prove that, by a nonlinear symplectic change of variables, the Hamiltonian (1.2) can be reduced to the special form (2.1), thus Theorem 1 is a consequence of Theorem 2. Then we give a full proof for Theorem 2.

3. Proof of Theorem 1

In this section, we prove that Theorem 2 implies Theorem 1. Firstly, we reorganize the normal frequencies according to the following equivalent relation:

DEFINITION 3.1. – *In the set* $\{\Omega_1, \ldots, \Omega_m\}$[6] *of normal frequencies, we introduce the following equivalence relation*: *Two normal frequencies* Ω_i *and* Ω_j *are said to be equivalent, denoted by* $\Omega_i \equiv \Omega_j$, *if there is a* $k_i^j \in Z^n$ *such that*

$$\Omega_j = \Omega_i + \langle \omega, k_i^j\rangle.$$

We denote by $[\Omega_i]$ *the equivalence class of* Ω_i *in* $\{\Omega_1, \ldots, \Omega_m\}$.

Now we relabel the normal frequencies $\{\Omega_1, \ldots, \Omega_m\}$ according to the following procedure:

1. Firstly, pick up a double resonant normal frequency in $\{\Omega_1, \ldots, \Omega_m\}$, i.e., $2\Omega_1^1 = \langle k_1, \omega\rangle$ for some $k_1 \in Z^n$, reindex it by Ω_1^1; After it we take *all* normal frequencies, which are equivalent to Ω_1^1, relabel them by $\{\Omega_1^2, \ldots, \Omega_1^{d_1}\}$; The situation that there is no double resonant frequency is simpler and has been considered in [18].

2. The remained $m - d_1$ normal frequencies $\{\Omega_1, \ldots, \Omega_m\} \setminus \{\Omega_1^1, \ldots, \Omega_1^{d_1}\}$ are neither equivalent to Ω_1^1 nor $-\Omega_1^1$ since $\Omega_1^1 \equiv -\Omega_1^1$. Repeat the above procedure for the remained

[6] The Ω_j are objects being distinct from one another by means of the index j even if they are equal as functions.

$m - d_1$ normal frequencies. All double-resonant frequencies in $\{\Omega_1, \ldots, \Omega_m\}$, relabeled by $\{\Omega_i^j, i = 1, \ldots, s_1, j = 1, \ldots, d_{s_1}\}$, will be picked out at the s_1 step.

3. Now all the remained $m - \sum_{i=1}^{s_1} d_i$ normal frequencies $\{\Omega_1, \ldots, \Omega_m\} \setminus \{\Omega_i^j, i = 1, \ldots, s_1;\ j = 1, \ldots, d_i\}$ are not double-resonant. Pick up one from the remained $m - \sum_{i=1}^{s_1} d_i$ normal frequencies, relabel it by Ω_{s+1}^1. After it we pick up all those normal frequencies which are equivalent to Ω_{s+1}^1,[7] and reindex them as $\{\Omega_{s+1}^j, j = 2, \ldots, \bar{d}_{s+1}\}$. If there are normal frequencies in the remained set $\{\Omega_1, \ldots, \Omega_m\} \setminus \{\Omega_i^j, i = 1, \ldots, s_1 + 1;\ j = 1, \ldots, d_i\}$ which are equivalent to $-\Omega_{s+1}^1$, we pick up them all, and reindex them by Ω_{s+1}^j, $j = \bar{d}_{s+1} + 1, \ldots, d_{s+1}$.

Remark. – Ω_{s+1}^1 may not be a normal frequency, but it does not matter. It is used *only* for the relabel of the normal frequencies $\{\Omega_1, \ldots, \Omega_m\}$.

4. The remained $m - \sum_{i=1}^{s+1} d_i$ normal frequencies do not belong to any of $[\Omega_i^1], [-\Omega_i^1]$, $i = 1, \ldots, s_1 + 1$. Repeat the step 3, until $\{\Omega_1, \ldots, \Omega_m\}$ are run out at some step, say s.

It follows that, by the above procedure, $\{\Omega_1, \ldots, \Omega_m\}$ can be relabeled as the following:

$$\{\Omega_1^1, \ldots, \Omega_1^{d_1}, \ldots, \Omega_s^1, \ldots, \Omega_s^{d_s}\}, \quad \sum_{i=1}^{s} d_i = m,$$

which is permutation of $\{\Omega_1, \ldots, \Omega_m\}$ such that

$$\begin{aligned} \Omega_i^j &= \Omega_i^1 + \langle \omega, k_i^j \rangle, \quad \forall 1 \leqslant j \leqslant d_i, \\ 2\Omega_i^1 &= \langle \omega, k_i \rangle, \quad \forall 1 \leqslant i \leqslant s_1, \end{aligned} \tag{3.1}$$

$$\begin{aligned} \Omega_i^j &= \Omega_i^1 + \langle \omega, k_i^j \rangle, \quad 1 \leqslant j \leqslant \bar{d}_i, \\ \Omega_i^j &= -\Omega_i^1 + \langle \omega, k_i^j \rangle, \quad \bar{d}_i + 1 \leqslant j \leqslant d_i, \\ 2\Omega_i^1 &\neq \langle \omega, k \rangle, \quad s_1 + 1 \leqslant i \leqslant s, \end{aligned} \tag{3.2}$$

for some $k_i^j \in Z^n$, and $\forall k \in Z^n$.

Now it is easy to check that $\{\Omega_i^1, i = 1, \ldots, s\}$ satisfy the non-resonance conditions (2.2)–(2.5), i.e.,

$$\begin{aligned} &\langle \omega(\xi), k \rangle + \Omega_i^1(\xi) \neq 0, \\ &\langle \omega(\xi), k + k_j \rangle - \Omega_i^1(\xi) - \Omega_j^1(\xi) \neq 0, \\ &\langle \omega(\xi), k \rangle + \Omega_i^1(\xi) - \Omega_j^1(\xi) \neq 0, \end{aligned}$$

with some fixed $k_1, \ldots, k_{s_1}$; $k_{s_1+1} = 0, \ldots, k_s = 0$, for all $k \in Z^n, i, j = 1, 2, \ldots, s$ with $|k| + |i - j| \neq 0$.

We will show that, by a nonlinear symplectic coordinates transformation, the normal form in Theorem 1 can be transformed into the special case (2.1) in Theorem 2. For this purpose, we need the following elementary observations:

[7] Here we don't care the order of Ω_{s+1}^j, $j = 2, \ldots, \bar{d}_{s+1}$ since it does not matter.

LEMMA 3.1. – *For any* $k_1, k_2, \ldots, k_m \in Z^n$*, non-singular* $m \times m$ *matrix* S *with* $S^T \bar{S} = I$*, the map* $\Psi : (\theta, I, z, \bar{z}) \to (\theta_+, I_+, z_+, \bar{z}_+)$ *defined by*:

$$
\begin{aligned}
\theta_+ &= \theta, \\
I_+ &= I \ + \sum_{j=1}^{m} z_j \bar{z}_j k_j, \\
z_+ &= SEz, \\
\bar{z}_+ &= \bar{S}\bar{E}\bar{z}
\end{aligned}
\tag{3.3}
$$

is symplectic, where diagonal matrix

$$E = E(k_1, k_2, \ldots, k_m) = \operatorname{diag}\big(\mathrm{e}^{\mathrm{i}\langle k_1, \theta\rangle}, \mathrm{e}^{\mathrm{i}\langle k_2, \theta\rangle}, \ldots, \mathrm{e}^{\mathrm{i}\langle k_m, \theta\rangle}\big),$$

i *stands for* $\sqrt{-1}$.

The proof of the above lemma is elementary, we refer to [16].

COROLLARY 1. – *Suppose that* $N = \langle \omega(\xi), I\rangle + \sum_{j=1}^{m} \Omega_j(\xi) z_j \bar{z}_j$ *be a real Hamiltonian. For any* $\tilde{\Omega}_j(\xi) \equiv \Omega_j(\xi)$*,* N *can be transformed to the integrable Hamiltonian of the form*

$$\big\langle \omega(\xi), I\big\rangle + \sum_{j=1}^{m} \tilde{\Omega}_j(\xi) z_j \bar{z}_j$$

by a symplectic change of variables.

For the case considered in this paper, we let

$$E_i = \operatorname{diag}\big(\mathrm{e}^{\mathrm{i}\langle k_i^1, x\rangle}, \ldots, \mathrm{e}^{\mathrm{i}\langle k_i^{d_i}, x\rangle}\big)$$

and $E = \operatorname{diag}(E_1, E_2, \ldots, E_s)$, and define a map $\Psi : (x, y, z, \bar{z}) \to (\theta, I, w, \bar{w})$ by

$$\theta = x, \qquad I = y + \sum_{i=1}^{m} \sum_{j=1}^{d_i} z_i^j \bar{z}_i^j k_i^j, \qquad w = Ez, \qquad \bar{w} = \bar{E}\bar{z},$$

where $\bar{E}$ is the complex conjugate matrix of E. Then

$$H \circ \Psi = N \circ \Psi + P \circ \Psi = N_0 + P_0,$$

where

$$N_0 = \sum_{j=1}^{n} \omega_j(\xi) I_j + \sum_{i=1}^{s_1} \Omega_i^1 \langle w_i, \bar{w}_i\rangle + \sum_{i=s_1+1}^{s} \big\langle \Omega_i^1 \tilde{I}_i w_i, \bar{w}_i\big\rangle,$$

where $w_i = (z_i^1 \mathrm{e}^{\mathrm{i}\langle k_i^1, x\rangle}, \ldots, z_i^{d_i} \mathrm{e}^{\mathrm{i}\langle k_i^{d_i}, x\rangle})$ are d_i vectors. If we still denote the variables of transformed Hamiltonian by $(x, y, z, \bar{z})$ instead of $(\theta, I, w, \bar{w})$, we get the special Hamiltonian (2.1). Thus Theorem 1 follows from Theorem 2.

Remark. – With the above symplectic map we easily see that the first Melnikov's condition (1.5) is necessary. Obviously, if there is a j_0 such that $\Omega_{j_0} = 0$, in $N = \langle\omega, I\rangle + \sum_{j=1}^{m} \Omega_j z_j \bar z_j$, then the invariant tori would disappear after the small perturbation $\varepsilon(z_{j_0} + \bar z_{j_0})$ for $\forall\varepsilon \neq 0$. This is a degenerate case. If (1.5) is not satisfied, i.e., there exist j_0 and $k_0 \in Z^n$ such that $\Omega_{j_0} = \langle\omega, k_0\rangle$, N can be transformed to the degenerate case (with a zero normal frequency) by a symplectic map. In fact, the symplectic map $\Psi : (\theta_+, I_+, z_+, \bar z_+) \to (\theta, I, z, \bar z)$ of the following form:

$$\begin{aligned} &\theta_+ = \theta, \qquad I_+ = I + k_0 z_{j_0} \bar z_{j_0}, \\ &z_{+j} = z_j, \qquad \bar z_{+j} = \bar z_j, \quad \forall j \neq j_0, \\ &z_{+j_0} = z_{j_0} \mathrm{e}^{\mathrm{i}\langle k_0, \theta\rangle}, \qquad \bar z_{+j_0} = \bar z_{j_0} \mathrm{e}^{-\mathrm{i}\langle k_0, \theta\rangle}, \end{aligned}$$

takes N to

$$N \circ \Psi = \langle\omega, I_+\rangle + \sum_{j\neq j_0} \Omega_j z_{+j} \bar z_{+j},$$

with a zero normal frequency. This shows that the first Melnikov's condition is necessary for the persistence of lower-dimensional invariant tori.

4. Proof of Theorem 2

Our proof follows in principle the usual KAM iteration. Since we have to remain some x-dependent second-order terms of z and $\bar z$, which can not be eliminated in KAM steps due to the second-order resonance between tangential frequencies and the normal frequencies, the homological equation is more complicated. The main idea of this paper is to decompose the homological equation into some linear algebraic systems then solve it, even if the normal form part is x-dependent.

4.1. Outline of KAM step

Below we give the ideas of one KAM iteration step. In the following, all the quantities represent the quantities in the νth KAM step. The quantities with subscript $+$ represent the quantity in the $(\nu + 1)$th KAM step.

Let

$$\mathcal{A} = \sum_{j=1}^{m} \langle A_j z_j, \bar z_j\rangle = \langle Az, \bar z\rangle, \quad A = \mathrm{diag}(A_1, A_2, \ldots, A_m), \tag{4.1}$$

where $z = (z_1, z_2, \ldots, z_m)$, A_j is a $d_j \times d_j$ matrix depending on ξ. Let

$$\mathcal{B} = \sum_{j=1}^{m} \langle B_j z_j, z_j\rangle \mathrm{e}^{\mathrm{i}\langle k_j, x\rangle} = \langle BGz, z\rangle, \tag{4.2}$$

where $B = \mathrm{diag}(B_1, B_2, \ldots, B_m)$, B_j is a $d_j \times d_j$ matrix depending on ξ;

$$G = \mathrm{diag}\big(\mathrm{e}^{\mathrm{i}\langle k_1, x\rangle} I_1, \ldots, \mathrm{e}^{\mathrm{i}\langle k_2, x\rangle} I_2, \ldots, \mathrm{e}^{\mathrm{i}\langle k_m, x\rangle} I_m\big), \tag{4.3}$$

with I_j being the $d_j \times d_j$ unit matrix and $k_j = 0$ for $\bar{m} < j \leqslant m$. Similarly,

$$\bar{\mathcal{B}} = \sum_{j=1}^{m} \langle \bar{B}_j \bar{z}_j, \bar{z}_j \rangle e^{-i\langle k_j, x\rangle} = \langle \bar{B}\bar{G}\bar{z}, \bar{z}\rangle.$$

At each KAM step, we will consider a Hamiltonian of the form:

$$H = N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}} + P,$$

where $N = \langle \omega(\xi), I\rangle + \sum_{j=1}^{m} \langle \Omega_j(\xi)\tilde{I}_j z_j, \bar{z}_j\rangle$, ω and Ω depend on the parameter ξ and P is a small perturbation.

Moreover, we assume that [8]

$$\|X_{\mathcal{A}}\|^L_{r;s,r}, \qquad \|X_{\mathcal{B}}\|^L_{r;s,r}, \qquad \|X_{\bar{\mathcal{B}}}\|^L_{r;s,r} \leqslant 2\varepsilon_0, \qquad \|X_P\|^L_{r;s,r} \leqslant \varepsilon, \tag{4.4}$$

with $\varepsilon \leqslant \varepsilon_0$.

Remark. – The appearance of $\mathcal{A}, \mathcal{B}, \bar{\mathcal{B}}$ is due to the second-order resonance between ω and Ω, which makes some x-dependent quadratic terms in the perturbation can not be removed.

Truncate P as $P = R + \tilde{P}$, where

$$R = \sum_{k\in Z^n, 2|l|+|q+\bar{q}|\leqslant 2} P_{k,l,q,\bar{q}} e^{\sqrt{-1}\langle k,x\rangle} y^l z^q \bar{z}^{\bar{q}}, \qquad \tilde{P} = P - R.$$

It follows that $\|X_R\|^L_{r;s,r} \leqslant \varepsilon$. We further write R as

$$\begin{aligned} R &= R^0 + R^1 + R^{01} + R^{10} + R^{11} + R^{02} + R^{20} \\ &= R^0(x) + \langle R^1(x), I\rangle + \langle R^{10}(x), \bar{z}\rangle + \langle R^{01}(x), z\rangle \\ &\quad + \langle R^{11}(x)\bar{z}, z\rangle + \langle R^{02}(x)z, z\rangle + \langle R^{20}(x)\bar{z}, \bar{z}\rangle \\ &= \sum_k R^0_k e^{i\langle k,x\rangle} + \sum_k \langle R^1_k, y\rangle e^{i\langle k,x\rangle} + \sum_{k,j} \langle R^{10}_{j,k}, \bar{z}_j\rangle e^{i\langle k,x\rangle} + \sum_{k,j} \langle R^{01}_{j,k}, z_j\rangle e^{i\langle k,x\rangle} \\ &\quad + \sum_{k,i,j} \langle R^{11}_{ij,k}\bar{z}_j, z_i\rangle e^{i\langle k,x\rangle} + \sum_{k,i,j} \langle R^{02}_{ij,k}\bar{z}_j, \bar{z}_i\rangle e^{i\langle k,x\rangle} + \sum_{k,i,j} \langle R^{20}_{ij,k} z_j, z_i\rangle e^{i\langle k,x\rangle}, \end{aligned} \tag{4.5}$$

where $R^{10}_{j,k}$ and $R^{01}_{j,k}$ are d_j-dimensional vectors, $R^{11}_{ij,k}$, $R^{20}_{ij,k}$ and $R^{02}_{ij,k}$ are $d_j \times d_j$ matrices. Since our Hamiltonian comes from a real function, we have R^{11} is Hermitian, R^{20}, R^{02} are symmetric with $\bar{R}^{02} = R^{20}$; $(\bar{R}^{11}_{ij,k})^{\mathrm{T}} = R^{11}_{ji,-k}$, $(\bar{R}^{20}_{ij,k})^{\mathrm{T}} = R^{02}_{ji,-k}$.

[8] From now on, $\|\cdot\|$ means C^L norm in Whitney's sense, since at each KAM step except the first one, the parameter set is no longer an open set.

Let

$$
\begin{aligned}
N_+ &= N + \sum \langle R_0^1, y\rangle, \qquad \mathcal{A}_+ = \sum_{j=1}^{m} \big\langle \big(A_j + R_{jj,0}^{11}\big)\bar z_j, z_j\big\rangle, \\
\mathcal{B}_+ &= \sum_{j=1}^{m} \big\langle \big(B_j + R_{jj,k_j}^{02}\big) z_j, z_j\big\rangle \mathrm{e}^{\mathrm{i}\langle k_j, x\rangle}, \qquad \bar{\mathcal{B}}_+ = \sum_{j=1}^{m} \big\langle \big(\bar B_j + R_{jj,-k_j}^{20}\big)\bar z_j, \bar z_j\big\rangle \mathrm{e}^{-\mathrm{i}\langle k_j, x\rangle}.
\end{aligned} \tag{4.6}
$$

$N_+ - N$, $\mathcal{A}_+ - \mathcal{A}$, $\mathcal{B} - \mathcal{B}$, $\bar{\mathcal{B}}_+ - \bar{\mathcal{B}}$ are terms in R which can not be solved at each KAM step and must be waved to the normal form part. Denote by:

$$
\begin{aligned}
\tilde R = {} & \sum_k R_k^0 \mathrm{e}^{\mathrm{i}\langle k,x\rangle} + \sum_{k\neq 0} \langle R_k^1, y\rangle \mathrm{e}^{\mathrm{i}\langle k,x\rangle} + \sum_{k,j} \langle R_{j,k}^{10}, \bar z_j\rangle \mathrm{e}^{\mathrm{i}\langle k,x\rangle} \\
& + \sum_{k,j} \langle R_{j,k}^{01}, z_j\rangle \mathrm{e}^{\mathrm{i}\langle k,x\rangle} + \sum_{|k|+|i-j|\neq 0} \langle R_{ij,k}^{11} \bar z_j, z_i\rangle \mathrm{e}^{\mathrm{i}\langle k,x\rangle} \\
& + \sum_{|k+k_j|+|i-j|\neq 0} \langle R_{ij,k}^{02} \bar z_j, \bar z_i\rangle \mathrm{e}^{\mathrm{i}\langle k,x\rangle} + \sum_{|k-k_j|+|i-j|\neq 0} \langle R_{ij,k}^{20} z_j, z_i\rangle \mathrm{e}^{\mathrm{i}\langle k,x\rangle}.
\end{aligned} \tag{4.7}
$$

Up to a constant, we have

$$
H = N_+ + \mathcal{A}_+ + \mathcal{B}_+ + \bar{\mathcal{B}}_+ + \tilde R + \tilde P.
$$

At each KAM step we will construct a symplectic map Φ such that $H_+ = H \circ \Phi = N_+ + \mathcal{A}_+ + \mathcal{B}_+ + \bar{\mathcal{B}}_+ + P_+$ with P_+ being much smaller.

4.2. Constructing the symplectic change of variables

As usual, we construct the desired symplectic map Φ by the time 1-map of the flow X_F^t of a Hamiltonian vector field X_F. It follows that

$$
\begin{aligned}
H \circ \Phi = {} & N_+ + \mathcal{A}_+ + \mathcal{B}_+ + \bar{\mathcal{B}}_+ + \{N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}}, F\} + \tilde R + \int_0^1 \{R, F\} \circ X_F^t \,\mathrm{d}t \\
& + \int_0^1 (1-t)\{\{N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}}, F\}, F\} \circ X_F^t \,\mathrm{d}t + \tilde P \circ \Phi \\
= {} & N_+ + \mathcal{A}_+ + \mathcal{B}_+ + \bar{\mathcal{B}}_+ + \{N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}}, F\} + \tilde R + P_+.
\end{aligned}
$$

We shall prove that

$$
\{N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}}, F\} + \tilde R = 0 \tag{4.8}
$$

is solvable and P_+ is much smaller. Taking F as the solution of the above equation, the time 1 map of the flow X_F^t is the desired map.

To solve the above (4.8), we let

$$
\begin{aligned}
F = {} & F^0(x) + \langle F^1(x), y\rangle + \langle F^{10}(x), \bar z\rangle + \langle F^{01}(x), z\rangle \\
& + \langle F^{20}(x)\bar z, \bar z\rangle + \langle F^{02}(x) z, z\rangle + \langle F^{11}(x)\bar z, z\rangle \\
= {} & \sum_k F^0_k e^{i\langle k,x\rangle} + \sum_{k\neq 0} \langle F^1_k, y\rangle e^{i\langle k,x\rangle} + \sum_{k,j} \langle F^{10}_{j,k}, \bar z_j\rangle e^{i\langle k,x\rangle} \\
& + \sum_{k,j} \langle F^{01}_{j,k}, z_j\rangle e^{i\langle k,x\rangle} + \sum_{|k|+|i-j|\neq 0} \langle F^{11}_{ij,k}\bar z_j, z_i\rangle e^{i\langle k,x\rangle} \\
& + \sum_{|k+k_j|+|i-j|\neq 0} \langle F^{02}_{ij,k}\bar z_j, \bar z_i\rangle e^{i\langle k,x\rangle} + \sum_{|k-k_j|+|i-j|\neq 0} \langle F^{20}_{ij,k} z_j, z_i\rangle e^{i\langle k,x\rangle},
\end{aligned}
\tag{4.9}
$$

where $F^{10}_{j,k}$ and $F^{01}_{j,k}$ are d_j-dimensional vectors, $F^{11}_{ij,k}$, $F^{20}_{ij,k}$ and $F^{02}_{ij,k}$ are $d_j \times d_j$ matrice, F^{11} is Hermitian, F^{20}, F^{02} are symmetric with $\bar F^{02} = F^{20}$; $(\bar F^{11}_{ij,k})^{\mathrm T} = F^{11}_{ji,-k}$, $(\bar F^{20}_{ij,k})^{\mathrm T} = F^{02}_{ji,-k}$.

It follows that

$$
\begin{aligned}
& -\{N + \mathcal A + \mathcal B + \bar{\mathcal B}, F\} \\
& = \partial_\omega F^0 + \langle \partial_\omega F^1, y\rangle && (4.10)\\
& \quad + \langle \partial_\omega F^{10}, \bar z\rangle + i\langle A_*^T F^{10} - 2\bar G\bar B F^{01}, \bar z\rangle && (4.11)\\
& \quad + \langle \partial_\omega F^{01}, z\rangle - i\langle A_* F^{01} - 2GBF^{10}, z\rangle && (4.12)\\
& \quad + \langle \partial\omega F^{11}\bar z, z\rangle + i\langle (F^{11}A_* - A_*F^{11} + 4GBF^{20} - 4F^{02}\bar B\bar G)\bar z, z\rangle && (4.13)\\
& \quad + \langle \partial_\omega F^{02} z, z\rangle - i\langle (A_*F^{02} + F^{02}A_*^{\mathrm T} - GB(F^{11})^{\mathrm T} - F^{11}BG - BGW) z, z\rangle && (4.14)\\
& \quad + \langle \partial_\omega F^{20} e\bar z, \bar z\rangle + i\langle (A_*^{\mathrm T}F^{20} + F^{20}A_* - \bar G\bar B F^{11} - (F^{11})^{\mathrm T}\bar B\bar G - \bar B\bar G\bar W)\bar z, \bar z\rangle, && (4.15)
\end{aligned}
$$

where $\partial_\omega F = \langle \omega, F_x\rangle$,

$$
W = \operatorname{diag}\big(\langle k_1, F^1\rangle I_1, \ldots, \langle k_j, F^1\rangle I_j, \ldots, \langle k_m, F^1\rangle I_m\big)
\tag{4.16}
$$

and $A_* = \operatorname{diag}(\Omega_1 \tilde I_1, \Omega_2 \tilde I_2, \ldots, \Omega_m \tilde I_m) + A$.

We proceed to find a F solving (4.8). By (4.8), (4.10), we have:

$$
\partial_\omega F^0 = \tilde R^0(x), \qquad \partial_\omega F^1 = \tilde R^1(x),
\tag{4.17}
$$

i.e.,

$$
F^j_k = \frac{1}{i\langle \omega, k\rangle}\tilde R^j_k, \quad j = 0, 1,\ k \neq 0.
$$

If the following small divisor conditions

$$
\big|\langle \omega, k\rangle\big| \geqslant \frac{\alpha}{|k|^\tau}, \quad k \neq 0,\ \tau > n - 1,
\tag{4.18}
$$

hold, we have

$$
\|F^j_k\|^L \leqslant \frac{\alpha^{L+1}}{|k|^{(L+1)\tau + L}}\|\tilde R^j_k\|^L, \quad k \neq 0,\ j = 0, 1.
$$

Thus we have:

$$\frac{1}{r^2}\|F^0\|^L_{s-\rho} \leqslant \frac{c}{r^2\alpha^{L+1}\rho^v}\|\tilde{R}^0\|^L_s \leqslant \frac{c\varepsilon}{\alpha^{L+1}\rho^v}, \qquad (4.19)$$
$$\|F^1_j\|^L_{s-\rho} \leqslant \frac{c}{\alpha^{L+1}\rho^v}\|\tilde{R}^1_j\|^L_s \leqslant \frac{c\varepsilon}{\alpha^{L+1}\rho^v}$$

with $v \geqslant (L+1)\tau + n + L$.

Now we consider 1-th terms of z and $\bar{z}$. By (4.8), (4.11), (4.12), we have:

$$\partial\omega F^{10} + \mathrm{i}\big(A^{\mathrm{T}}_* F^{10} - 2\bar{G}\bar{B}F^{01}\big) = \tilde{R}^{10}, \qquad (4.20)$$
$$\partial\omega F^{01} - \mathrm{i}\big(A_* F^{01} - 2GBF^{10}\big) = \tilde{R}^{01}.$$

Comparing the coefficients of $z_j, \bar{z}_j$, we get:

$$-\mathrm{i}\partial_\omega F^{10}_j + A^{\mathrm{T}}_{*j}F^{10}_j - 2\mathrm{e}^{-\mathrm{i}\langle k_j,x\rangle}\bar{B}_j F^{01}_j = -\mathrm{i}\tilde{R}^{10}_j,$$
$$-\mathrm{i}\partial\omega F^{01}_j - A_{*j}F^{01}_j + 2\mathrm{e}^{\mathrm{i}\langle k_j,x\rangle}B_j F^{10}_j = -\mathrm{i}\tilde{R}^{01}_j.$$

Going to the Fourier coefficients, we have:

$$\langle\omega,k\rangle F^{10}_{j,k} + A^{\mathrm{T}}_{*j}F^{10}_{j,k} - 2\bar{B}_j F^{01}_{j,k+k_j} = -\mathrm{i}\tilde{R}^{10}_{j,k},$$
$$\langle\omega,k\rangle F^{01}_{j,k} - A_{*j}F^{01}_{j,k} + 2B_j F^{10}_{j,k-k_j} = -\mathrm{i}\tilde{R}^{01}_{j,k}.$$

Replacing k by $k+k_j$ in the above second equation and noting that $A_{*j} = \Omega_j\tilde{I}_j + A_j$, we get:

$$\big[\langle\omega,k\rangle I_j + \Omega_j\tilde{I}_j + A^{\mathrm{T}}_j\big]F^{10}_{j,k} - 2\bar{B}_j F^{01}_{j,k+k_j} = -\mathrm{i}\tilde{R}^{10}_{j,k}, \qquad (4.21)$$
$$2B_j F^{10}_{j,k} + \big[\langle\omega,k+k_j\rangle I_j - \Omega_j\tilde{I}_j - A_j\big]F^{01}_{j,k+k_j} = -\mathrm{i}\tilde{R}^{01}_{j,k+k_j}.$$

(4.21) is a $2d_j$-dimensional linear algebraic system with the coefficient matrix $M_1(k) = \mathrm{Diag}((\langle\omega,k\rangle + \Omega_j)I, \langle\omega,k+k_j\rangle - \Omega_j)I) + P(\xi)$, where I is a unit matrix of order $\sum_{j=1}^m d_j$, and $P(\xi)$ is small with $\|P\|^L \leqslant c\varepsilon_0$. If[9]

$$\big\|\big(M_1(k)\big)^{-1}\big\| < \frac{|k|^\tau}{\alpha} \qquad (4.22)$$

for $|k| \neq 0$, together with

$$\|\tilde{R}^{10}_{j,k}\|^L, \|\tilde{R}^{01}_{j,k+k_j}\|^L \leqslant cr\varepsilon\mathrm{e}^{-|k|s},$$

we have:

$$\|F^{10}_{j,k}\|^L, \|F^{01}_{j,k+k_j}\|^L \leqslant c\left[\frac{|k|^\tau}{\alpha}\right]^{L+1}|k|^L\mathrm{e}^{-|k|s}r\varepsilon,$$

and thus

$$\frac{1}{r}\|F^{01}\|^L_{2;s-\rho,r} \leqslant \frac{c\varepsilon}{\alpha^{L+1}\rho^v}, \qquad \frac{1}{r}\|F^{10}\|^L_{2;s-\rho,r} \leqslant \frac{c\varepsilon}{\alpha^{L+1}\rho^v}, \qquad (4.23)$$

where and below $v = \tau(L+1) + L + n + 1$ and c is a constant only depending on τ, n, m, K.

[9] The norm $\|\cdot\|$ for matrix is defined to be the maximum of the absolute value of the eigenvalues.

In the following, we consider the quadratic terms of z and $\bar{z}$. (4.8) together with (4.13), (4.14), (4.15) yields:

$$
\begin{aligned}
&-\mathrm{i}\langle\omega, F^{11}_x\rangle + F^{11}A_* - A_*F^{11} + 4GBF^{20} - 4F^{02T}\bar{B}\bar{G} = -\mathrm{i}\tilde{R}^{11},\\
&-\mathrm{i}\langle\omega, F^{02}_x\rangle - A_*F^{02} - F^{02}A_*^T + GBF^{11T} + F^{11}BG - BGW = -\mathrm{i}\tilde{R}^{02},\\
&-\mathrm{i}\langle\omega, F^{20}_x\rangle + A_*^T F^{20} + F^{20}A_* - \bar{G}\bar{B}F^{11} - F^{11T}\bar{B}\bar{G} + \bar{B}\bar{G}\bar{W} = -\mathrm{i}\tilde{R}^{20}.
\end{aligned}
\tag{4.24}
$$

Denote by $\tilde{F}^{11} = F^{11} - \frac{1}{2}W$. By (4.16) W is a real diagonal block matrix, we have $\bar{W} = W$. Moreover, (4.17) implies $\langle\omega, W_x\rangle = \mathrm{diag}(\langle k_1, \tilde{R}^1\rangle I_1, \ldots, \langle k_j, \tilde{R}^1\rangle I_j, \ldots, \langle k_m, \tilde{R}^1\rangle I_m)$. Adding the transpose in both sides of the first equation to (4.24) we have the following systems:

$$
\begin{aligned}
&-\mathrm{i}\partial_\omega \tilde{F}^{11} + \tilde{F}^{11}A_* - A_*\tilde{F}^{11} + 4GBF^{20} - 4F^{02}\bar{B}\bar{G} = -i\tilde{R}^{11}_*,\\
&-\mathrm{i}\partial_\omega \big(\tilde{F}^{11}\big)^{\mathrm{T}} + A_*^{\mathrm{T}}\big(\tilde{F}^{11}\big)^{\mathrm{T}} - \big(\tilde{F}^{11}\big)^{\mathrm{T}}A_*^{\mathrm{T}} + 4F^{20}BG - 4\bar{G}\bar{B}F^{02} = -\mathrm{i}\big(\tilde{R}^{11}_*\big)^{\mathrm{T}},\\
&-\mathrm{i}\partial_\omega F^{02} - A_*F^{02} - F^{02}A_*^{\mathrm{T}} + GB\big(\tilde{F}^{11}\big)^T + \tilde{F}^{11}BG = -\mathrm{i}\tilde{R}^{02},\\
&-\mathrm{i}\partial_\omega F^{20} + A_*^{\mathrm{T}}F^{20} + F^{20}A_* - \bar{G}\bar{B}\tilde{F}^{11} - \big(\tilde{F}^{11}\big)^{\mathrm{T}}\bar{B}\bar{G} = -\mathrm{i}\tilde{R}^{20},
\end{aligned}
\tag{4.25}
$$

where for simplicity, we let $\tilde{R}^{11}_* = \tilde{R}^{11} + \hat{R}^{11}$ with $\hat{R}^{11} = \frac{1}{2}\partial_\omega W = \frac{1}{2}\langle\omega, W_x\rangle$.

Let

$$
Q = (Q_1, Q_2, Q_3, Q_4) = \big(\tilde{F}^{11}_{ij,k}, \big(\tilde{F}^{11}_{ji,k-k_i+k_j}\big)^{\mathrm{T}}, F^{02}_{ij,k+k_j}, F^{20}_{ij,k-k_i}\big),
$$

$$
C = (C_1, C_2, C_3, C_4) = \big(\tilde{R}^{11}_{*ij,k}, \big(\tilde{R}^{11}_{*ji,k-k_i+k_j}\big)^{\mathrm{T}}, R^{02}_{ij,k+k_j}, R^{20}_{ij,k-k_i}\big).
$$

Comparing the Fourier coefficients in (4.25), we have:

$$
\begin{aligned}
&\langle\omega, k\rangle Q_1 + Q_1A_{*j} - A_{*i}Q_1 + 4B_iQ_4 - 4Q_3\bar{B}_j = -\mathrm{i}C_1,\\
&\langle\omega, k - k_i + k_j\rangle Q_2 + A_{*i}^{\mathrm{T}}Q_2 - Q_2A_{*j}^{\mathrm{T}} + 4Q_4B_j - 4\bar{B}_iQ_3 = -\mathrm{i}C_2,\\
&\langle\omega, k + k_j\rangle Q_3 - A_{*i}Q_3 - Q_3A_{*j}^{\mathrm{T}} + B_iQ_2 + Q_1B_j = -\mathrm{i}C_3,\\
&\langle\omega, k - k_i\rangle Q_4 + A_{*i}^{\mathrm{T}}Q_4 + Q_4A_{*j} - \bar{B}_iQ_1 - Q_2\bar{B}_j = -\mathrm{i}C_4
\end{aligned}
\tag{4.26}
$$

for $|k| + |i - j| \neq 0$, which is a linear algebraic systems of order $4d_id_j$ for fixed k, i, j. If the above equation has a unique solution, the solution $\tilde{F}^{11}_{ij,k}, F^{02}_{ij,k+k_j}, \tilde{F}^{11\mathrm{T}}_{ji,k-k_i+k_j}, F^{20}_{ij,k-k_i}$ of (4.26) has the required symmetry since $(\bar{\tilde{F}}^{11}_{ij,k})^{\mathrm{T}}, (\bar{F}^{02}_{ij,k+k_j})^{\mathrm{T}}, (\bar{\tilde{F}}^{11}_{ji,k-k_i+k_j}), (\bar{F}^{20}_{ij,k-k_i})^{\mathrm{T}}$ satisfy the same linear systems as $\tilde{F}^{11}_{ji,-k}, F^{20}_{ji,-k-k_j}, (\tilde{F}^{11}_{ij,-k-k_j+k_i})^{\mathrm{T}}, F^{02}_{ji,-k+k_i}$.

The coefficient matrix of (4.26), denoted by $M_2(k)$, is a perturbed diagonal matrix with the diagonal elements $\langle\omega, k + k_j\rangle - \Omega_i(\xi) - \Omega_j(\xi), \langle\omega, k\rangle + \Omega_i(\xi) - \Omega_j(\xi),\ |k| + |i - j| \neq 0$. The size of the perturbation is ε.

By Cauchy estimates, we have $\|C\|^L \leqslant c\mathrm{e}^{-|k|s}\varepsilon$. If

$$
\big\|\big(M_2(k)\big)^{-1}\big\| < \frac{|k|^\tau}{\alpha}
\tag{4.27}
$$

for $|k| \neq 0$, the above linear equations are resolvable and

$$\|Q\|^L \leqslant c\frac{|k|^{\tau(L+1)+L}}{\alpha^{L+1}}\|C\|^L \leqslant c\frac{|k|^{\tau(L+1)+L}}{\alpha^{L+1}}e^{-|k|s}\varepsilon. \tag{4.28}$$

By $\tilde{F}^{11} = F^{11} - \frac{1}{2}W$ and (4.19), for F_k^{11}, F_k^{20} and F_k^{02}, we have the same estimates as Q in (4.28). Finally we have:

$$\begin{aligned} &\frac{1}{r}\|F^{20}\bar{z}\|^L_{2;s-\rho,r} \leqslant \frac{c\varepsilon}{\alpha^{L+1}\rho^v}, \qquad \frac{1}{r}\|F^{02}z\|^L_{2;s\rho,r} \leqslant \frac{c\varepsilon}{\alpha^{L+1}\rho^v}, \\ &\frac{1}{r}\|F^{11}\bar{z}\|^L_{2;s-\rho,r} \leqslant \frac{c\varepsilon}{\alpha^{L+1}\rho^v}, \qquad \frac{1}{r}\|F^{11\mathrm{T}}z\|^L_{2;s\rho,r} \leqslant \frac{c\varepsilon}{\alpha^{L+1}\rho^v}. \end{aligned} \tag{4.29}$$

Combining (4.19)–(4.29), we find a F such that

$$\{N + \mathcal{A} + \mathcal{B} + \bar{\mathcal{B}}, F\} + \tilde{R} = 0.$$

Moreover,

$$\|X_F\|^L_{r;s-\rho,r} \leqslant \frac{c\varepsilon}{\alpha^{L+1}\rho^v}, \tag{4.30}$$

if $\xi \in O_+$ satisfies (4.18), (4.22) and (4.27).

Thus the transformation $\Phi = X_F^1$, transforms H into $H_+ = H \circ \Phi = N_+ + \mathcal{A}_+ + \mathcal{B}_+ + \bar{\mathcal{B}}_+ + P_+$, where $N_+ = \langle \omega_+, y\rangle + \langle \Omega\bar{z}, z\rangle$, $\mathcal{A}_+ = \mathcal{A} + \hat{\mathcal{A}}$, $\mathcal{B}_+ = \mathcal{B} + \hat{\mathcal{B}}$ and

$$P_+ = \int_0^1 \{R_t, F\} \circ X_F^t \,\mathrm{d}t + \tilde{P} \circ \Phi,$$

with $R_t = R + (1-t)\tilde{R}$. Moreover, we have:

$$\|\omega_+ - \omega\|^L \leqslant \varepsilon, \qquad \|X_{\mathcal{A}_+} - X_{\mathcal{A}}\|_{r;D(s-\rho,r)} \leqslant \varepsilon,$$

$$\|X_{\mathcal{B}_+} - X_{\mathcal{B}}\|_{r;D(s-\rho,r)} \leqslant \varepsilon, \qquad \|X_{\bar{\mathcal{B}}_+} - X_{\bar{\mathcal{B}}}\|_{r;D(s-\rho,r)} \leqslant \varepsilon.$$

Remark. – By the Whitney's extension theorem in [14], a function defined on O_α can be extended to O such that all the estimates still hold on O, so we always regard all functions of ξ in the KAM step to be defined on O and ignore the domain in the estimates. But it makes sense only for $\xi \in O_\alpha$.

4.3. Estimates for the new perturbation

To complete the KAM step we have to estimate the new perturbation P_+. For $\eta \leqslant 1/8$,

$$\|X_{\tilde{P}}\|^L_{\eta r;D(s,4\eta r)} \leqslant 2\eta\|X_{\tilde{P}}\|^L_{r;D(s,r)} \leqslant 2\eta\varepsilon,$$

where and below $\|\cdot\|_{r;D(s,r)} = \|\cdot\|_{r;s,r}$. To consider another term in P_+ we first estimate the sympelctic map X_F^t. By Lemma A.3 we have:

LEMMA 4.1. – *If X_F satisfies* (4.30) *and* $\frac{2c\varepsilon}{\alpha^{L+1}\rho^{\upsilon+1}} \leqslant 1$, *then we have*:

$$\frac{1}{\rho}\left\|X_F^t - \mathrm{id}\right\|^L_{r;D(s-2\rho,\frac{r}{2})}, \qquad \left\|\mathcal{D}X_F^t - \mathrm{Id}\right\|^L_{r,r;D(s-3\rho,\frac{r}{4})} \leqslant cE$$

for $|t| \leqslant 1$ *with* $E = \frac{\varepsilon}{\alpha^{L+1}\rho^{\upsilon+1}}$. *Moreover,*

$$\rho\left\|\mathcal{D}^2 X_F^t - \mathrm{Id}\right\|^L_{r,r,r;D(s-4\rho,\frac{r}{8})} \leqslant cE,$$

where $\mathcal{D}$ is the differentiation operator with respect to $(x,y,z,\bar{z})$, $\|\cdot\|_{\bar{r},r}$ is the operator norm defined by $\|\mathcal{L}\|_{\bar{r},r} = \sup_{W\neq 0}\frac{\|\mathcal{L}W\|_{\bar{r}}}{\|W\|_r}$ and $\mathcal{D}^2$ is the 2-order differentiation operator. c is independent of the KAM step.

The Hamiltonian vectorfield of new perturbation is

$$X_{P_+} = \left(X_F^1\right)^*(X_{\tilde{P}}) + \int_0^1 \left(X_F^t\right)^*[X_{R_t}, X_F]\,\mathrm{d}t,$$

where $(X_F^t)^*$ is the cotangent mapping of X_F^t. By the constructing of X_F^t, it follows that for each $\xi \in O_+$,

$$X_F^t(\cdot,\xi) : D(s-5\rho,\eta r) \to D(s-4\rho,2\eta r) \quad \text{for } |t| \leqslant 1.$$

By Lemma A.4, if η is sufficiently small, then

$$\|X_{P_+}\|^L_{\eta r;D(s-5\rho,\eta r)} \leqslant 4\|X_{\tilde{P}}\|^L_{\eta r;D(s-4\rho,2\eta r)} + 4\int_0^1 \left\|[X_{R_t}, X_F]\right\|^L_{\eta r,D(s-4\rho,2\eta r)}\,\mathrm{d}t$$

By Cauchy's inequality and Lemma 4.1,

$$\begin{aligned}\left\|[X_{R_t}, X_F]\right\|^L_{\eta r;D(s-4\rho,2\eta r)} &\leqslant \frac{1}{\eta^2}\|\mathcal{D}X_{R_t}X_F + \mathcal{D}X_F X_{R_t}\|^L_{r;D(s-4\rho,2\eta r)} \\ &\leqslant \frac{c}{\eta^2\rho}\|X_{R_t}\|^L_{r;D(s,r)}\|X_F\|^L_{r;D(s-\rho,r)} \leqslant \frac{c\varepsilon^2}{\eta^2\alpha^{L+1}\rho^{\upsilon+1}}.\end{aligned}$$

So

$$\|X_{P_+}\|^L_{\eta r;D(s-5\rho,\eta r)} \leqslant 8\eta\varepsilon + \frac{c\varepsilon^2}{\eta^2\alpha^{L+1}\rho^{\upsilon+1}} \leqslant c\eta\varepsilon = \varepsilon_+ \tag{4.31}$$

with $\eta^3 = E = \frac{\varepsilon}{\alpha^{L+1}\rho^{\upsilon+1}}$ and $\varepsilon_+ = c\eta\varepsilon$.

The KAM step is now completed.

4.4. Iteration lemma and convergence

For a given s,ε,r in the introduction, we define some sequences inductively depending on s,ε,r.

$$\begin{aligned}
&\varepsilon_1=\varepsilon, \qquad r_1=r, \quad s_1=s, \qquad \alpha_1=\alpha,\\
&E_1=\frac{\varepsilon_1}{\alpha_1^{L+1}\rho_1^{v+1}}, \qquad \eta_1=E_1^{1/3}, \qquad \rho_1=\frac{s_1}{20},\\
&N_1=N, \qquad \mathcal{A}_1=0, \qquad \mathcal{B}_1=\bar{\mathcal{B}}_1=0;
\end{aligned}$$

$$\varepsilon_{\nu+1}=c\eta_\nu\varepsilon_\nu, \qquad r_{\nu+1}=\eta_\nu r_\nu, \qquad s_{\nu+1}=s_\nu-5\rho_\nu, \qquad \rho_{\nu+1}=\frac{1}{20}\rho_\nu;$$

$$\alpha_{\nu+1}=\frac{\alpha}{2}\big(1+2^{-\nu}\big), \qquad E_{\nu+1}=\varepsilon_{\nu+1}\alpha_{\nu+1}^{L+1}\rho_{\nu+1}^{v+1}, \qquad \eta_{\nu+1}=E_{\nu+1}^{1/3}.$$

Let

$$D_\nu=D(s_\nu,r_\nu).$$

The proceeding analysis may be summarized as the following iteration lemma:

LEMMA 4.2. – *If* $\varepsilon\leqslant\varepsilon(\alpha)=\alpha^{L+1}s^{v+1}/20^{v+1}2c^3$ *the following holds for all* $\nu\geqslant 1$: *Suppose* $H_\nu=H\circ\Phi^\nu=N_\nu+\mathcal{A}_\nu+\mathcal{B}_\nu+\bar{\mathcal{B}}_\nu+P_\nu$, *where*

$$N_\nu=\langle\omega_\nu,y\rangle+\langle\Omega z,\bar z\rangle, \mathcal{A}_\nu=\langle A_\nu z,\bar z\rangle,$$

defined on $D_\nu\times O_\nu$ *with*

$$\mathcal{B}_\nu=\langle B_\nu Gz,z\rangle, \qquad \bar{\mathcal{B}}_\nu=\langle\bar B_\nu\bar G\bar z,\bar z\rangle,$$

$$\|\omega_\nu-\omega_{\nu-1}\|^L\leqslant\varepsilon_\nu, \qquad \|X_{\mathcal{A}_\nu}-X_{\mathcal{A}_{\nu-1}}\|_{r_\nu;D_\nu}\leqslant\varepsilon_\nu, \tag{4.32}$$

$$\|X_{\mathcal{B}_\nu}-X_{\mathcal{B}_{\nu-1}}\|_{r_\nu;D_\nu}\leqslant\varepsilon_\nu, \qquad \|X_{\bar{\mathcal{B}}_\nu}-X_{\bar{\mathcal{B}}_{\nu-1}}\|_{r_\nu;D_\nu}\leqslant\varepsilon_\nu. \tag{4.33}$$

O_ν *is the set such that for* $\xi\in O_\nu$, *the small divisor conditions*

$$\big|\langle\omega_\nu,k\rangle\big|\geqslant\frac{\alpha_\nu}{|k|^\tau}, \quad \forall 0\neq k\in Z^n,\ \tau>n-1,$$

hold and the matrices $M_{1,\nu}, M_{2,\nu}$ *are invertible such that the estimates in* (4.22) *and* (4.27) *are satisfied in the* ν*th KAM iteration step.*

Finally, we assume that

$$\|X_{P_\nu}\|_{r_\nu,D_\nu,O_\nu}\leqslant\varepsilon_\nu.$$

Then, there is a subset $O_{\nu+1}\subset O_\nu$,

$$O_{\nu+1}=O_\nu\setminus\bigcup_{|k|\geqslant 2^\nu}\mathcal{R}_k^{\nu+1}(\alpha_\nu),$$

where

$$\mathcal{R}_k^{\nu+1}(\alpha_{\nu+1})=\left\{\xi\in O_\nu \mid \big|\langle k,\omega_{\nu+1}\rangle^{-1}\big|,\ \big\|M_{1,\nu+1}^{-1}\big\|,\ \big\|M_{2,\nu+1}^{-1}\big\|>\frac{|k|^\tau}{\alpha_\nu}\right\},$$

with $\omega_{\nu+1}=\omega_\nu+P_{0l0}^\nu$, $M_{1,\nu+1}=M_{1,\nu}+\mathrm{O}(\varepsilon_\nu)$ *and* $M_{2,\nu+1}=M_{2,\nu}+\mathrm{O}(\varepsilon_\nu)$, *and a symplectic change of variables*

$$\Phi_\nu: D_{\nu+1}\times O_{\nu+1}\to D_\nu, \tag{4.34}$$

such that $H_{\nu+1} = H_\nu \circ \Phi_\nu$, defined on $D_{\nu+1} \times O_{\nu+1}$, satisfies the same assumptions with $\nu + 1$ in place of ν.

By Lemma 4.1 for $\forall \xi \in O_{\nu+1}$ we have the map $\Phi_\nu : D_{\nu+1} \to D_\nu$ satisfying

$$\frac{1}{\rho_\nu}\|\Phi_\nu - \mathrm{id}\|^L_{r_\nu;D_{\nu+1}},\ \|\mathcal{D}\Phi_\nu - \mathrm{Id}\|^L_{r_\nu,r_\nu;D_{\nu+1}},\ \rho_\nu\big\|\mathcal{D}^2\Phi_\nu\big\|^L_{r_\nu,r_\nu,r_\nu;D_{\nu+1}} \leqslant cE_\nu. \tag{4.35}$$

Let $\Phi^\nu = \Phi_1 \circ \Phi_2 \circ \cdots \circ \Phi_\nu$, thus $H_\nu = H \circ \Phi^\nu = N_\nu + \mathcal{A}_\nu + \mathcal{B}_\nu + \bar{\mathcal{B}}_\nu + P_\nu$, where

$$N_\nu = \langle \omega_\nu, y\rangle + \langle \Omega z, \bar{z}\rangle, \qquad \mathcal{A}_\nu = \langle A_\nu z, \bar{z}\rangle,$$

$$\mathcal{B}_\nu = \langle B_\nu G z, z\rangle, \qquad \bar{\mathcal{B}}_\nu = \langle \bar{B}_\nu \bar{G}\bar{z}, \bar{z}\rangle.$$

It follows that $X_{H\circ\Phi^\nu} = D\Phi^\nu \cdot X_{H_\nu}$, i.e.,

$$\big\|\big(\Phi^\nu\big)^* X_{H_\nu} - X_{N_\nu+\mathcal{A}_\nu+\mathcal{B}_\nu+\bar{\mathcal{B}}_\nu}\big\| = \|X_{P_\nu}\|. \tag{4.36}$$

Let $O_\alpha = \bigcap_{\nu\geqslant 1} O_\nu$. By Lemma 4.2 we have:

$$\|X_{P_\nu}\|^L_{r_\nu;D_\nu} \leqslant \varepsilon_\nu, \qquad \|\omega_{\nu+1} - \omega_\nu\|^L \leqslant \varepsilon_\nu, \qquad \|X_{\mathcal{A}_{\nu+1}} - X_{\mathcal{A}_\nu}\|^L_{r_\nu;D_\nu} \leqslant \varepsilon_\nu, \tag{4.37}$$

$$\|X_{\mathcal{B}_{\nu+1}} - X_{\mathcal{B}_\nu}\|^L_{r_\nu;D_\nu} \leqslant \varepsilon_\nu, \qquad \|X_{\bar{\mathcal{B}}_{\nu+1}} - X_{\bar{\mathcal{B}}_\nu}\|^L_{r_\nu;D_\nu} \leqslant \varepsilon_\nu, \tag{4.38}$$

for $\xi \in O_\alpha$.

Since that $E_{\nu+1} = cE_\nu^{4/3}$, we have $c^3E_{\nu+1} \leqslant (c^3E_\nu)^{4/3} \leqslant \cdots \leqslant (c^3E_1)^{(4/3)^\nu}$. If $\varepsilon \leqslant \varepsilon(\alpha) = \alpha^{L+1}s^{\nu+1}/20^{\nu+1}2c^3$, we have $c^3E_1 \leqslant 1/2$ and thus $E_{\nu+1} \leqslant c^{-3}(\frac{1}{2})^{(4/3)^\nu}$. It follows that

$$\sum_{\nu\geqslant 2}\varepsilon_\nu \leqslant \sum_{\nu\geqslant 2}\alpha_\nu^{L+1}\rho_\nu^{\nu+1}E_\nu \leqslant 2\varepsilon_2 \leqslant 2c\eta_1\varepsilon.$$

Let $\varepsilon(\alpha)$ sufficiently small such that $2c\eta_1 \leqslant 1$. Thus we have $\sum_{\nu\geqslant 2}\varepsilon_\nu \leqslant \varepsilon$. By (4.37) it follows that

$$\big\|\omega_{\nu+1} - \omega(\xi)\big\|^L \leqslant (\varepsilon_1 + \varepsilon_2 + \cdots + \varepsilon_\nu) \leqslant 2\varepsilon. \tag{4.39}$$

Since $\|X_{\mathcal{A}_2}\|_{r_2;D_2} \leqslant \varepsilon$, we have $\|X_{\mathcal{A}_{\nu+1}} - X_{\mathcal{A}_2}\|_{r_{\nu+1};D_{\nu+1}} \leqslant \sum_{2\leqslant\nu'\leqslant\nu}\varepsilon_{\nu'} \leqslant \varepsilon$, so

$$\|X_{\mathcal{A}_{\nu+1}}\|_{r_{\nu+1};D_{\nu+1}} \leqslant 2\varepsilon. \tag{4.40}$$

Similarly, by (4.38) we have:

$$\|X_{\mathcal{B}_{\nu+1}}\|_{r_{\nu+1};D_{\nu+1}} \leqslant 2\varepsilon, \qquad \|X_{\bar{\mathcal{B}}_{\nu+1}}\|_{r_{\nu+1};D_{\nu+1}} \leqslant 2\varepsilon. \tag{4.41}$$

Now we prove $\{\Phi^\nu\}$ is convergent on $D_* \times O_\alpha = \bigcap_{\nu\geqslant 1} D_\nu \times O_\nu$ with $D_* = D(\frac{1}{2}s) \times \{0,0,0\}$.

By the constructing of Φ_ν, Φ_ν maps $D_{\nu+1}$ into $D(s_\nu - 4\rho_\nu, 2\eta_\nu r_\nu) \subset D(s_\nu - 2\rho_\nu, \frac{1}{2}r_\nu)$. Since the distance $\|\cdot\|_{r_\nu}$ from $D(s_\nu - 5\rho_\nu, 2\eta_\nu r_\nu)$ to the boundary of $D(s_\nu - 3\rho_\nu, \frac{1}{2}r_\nu)$ is more than ρ_ν, by the definitions of the norm $\|\cdot\|^L$, if E_1 is sufficiently small, we have

$\|\Phi_{\nu-1}\circ\Phi_\nu-\mathrm{id}\|^L_{r_{\nu-1};D_{\nu+1}}\leqslant\|\Phi_{\nu-1}-\mathrm{id}\|^L_{r_{\nu-1};D_\nu}$. Inductively it follows that for any $\nu\geqslant 1$ and $\nu'\geqslant 1$,

$$\|\Phi_\nu\circ\Phi_{\nu+1}\circ\cdots\circ\Phi_{\nu+\nu'}-\mathrm{id}\|^L_{r_\nu;D_{\nu+\nu'+1}}\leqslant\|\Phi_\nu-\mathrm{id}\|^L_{r_\nu;D_{\nu+1}}.$$

Since $\Phi^{\nu+1}=\Phi^\nu\circ\Phi_{\nu+1}$, we have:

$$\left\|\Phi^{\nu+1}-\Phi^\nu\right\|^L_{r_1;D_{\nu+2}}\leqslant\left\|\mathcal{D}\Phi^\nu\right\|^L_{r_1,r_\nu;D_{\nu+1}}\left\|\Phi_{\nu+1}-\mathrm{id}\right\|_{r_\nu;D_{\nu+2}}.$$

By the definition of the norm $\|\cdot\|_{r,s}$, we have $\|AB\|_{r,s}\leqslant\|A\|_{r,r}\|B\|_{s,s}$ for $r\geqslant s$. We have:

$$\begin{aligned}\left\|\mathcal{D}\Phi^\nu\right\|^L_{r_1,r_\nu;D_{\nu+1}}&\leqslant|\mathcal{D}\Phi_1\|_{r_1,r_1;D_2}\|^L\|\mathcal{D}\Phi_2\|^L_{r_2,r_2;D_3}\cdots\|\mathcal{D}\Phi\|^L_{r_\nu,r_\nu;D_{\nu+1}}\\&\leqslant\prod_{\nu'=1}^{\nu}(1+cE_{\nu'})<+\infty.\end{aligned}\tag{4.42}$$

So

$$\left\|\Phi^{\nu+1}-\Phi^\nu\right\|_{r_1;D_{\nu+2}}\leqslant c\|\Phi_{\nu+1}-\mathrm{id}\|_{r_\nu;D_{\nu+2}}\leqslant cE_\nu,$$

thus $\{\Phi^\nu\}$ is convergent on $D_*\times O_\alpha$, say, to Φ.

Now we can do in the same way as in [6,9,10] to prove the convergence of Φ^ν on $D(\frac12 s,\frac12 r)\times O_\alpha$, so we only give the ideas of proof and refer to [6,9,10] for details. We can use the estimates about $\mathcal{D}\Phi_\nu$ and $\mathcal{D}^2\Phi_\nu$ to prove $\{\mathcal{D}\Phi^\nu\}$ and $\{\mathcal{D}^2\Phi^\nu\}$ are convergent on $D_*\times O_\alpha$. By the constructing of Φ_ν, we know that Φ_ν are of quadratic order in y,z and $\bar z$, and so are their composition mappings Φ^ν. Thus, $\{\Phi^\nu\}$, $\{\mathcal{D}\Phi^\nu\}$, $\{\mathcal{D}^2\Phi^\nu\}$ are convergent on $D_*\times O_\alpha$ implies that $\{\Phi^\nu\}$ is actually convergent on $D(\frac12 s,\frac12 r)\times O_\alpha$.

Since

$$\|X_{P_\nu}\|^L_{r_\nu;D_\nu}\leqslant\varepsilon_\nu\quad\text{and}\quad\lim_{\nu\to\infty}\|X_{P_\nu}-X_{P_*}\|^L_{r_\nu;D_\nu}=0,$$

it follows that $X_{P_*}=0$ on $D_*\times O_\alpha$ and $\frac{\partial^{l+q+\bar q}P_*}{\partial y^l\partial z^q\partial\bar z^{\bar q}}|_{D_*}=0$ for $2|l|+|q+\bar q|\leqslant 2$. So $P_*=\sum_{k\in Z^n,2|l|+|q+\bar q|\geqslant 3}P_{*klq\bar q}y^lz^q\bar z^{\bar q}\mathrm{e}^{\mathrm{i}\langle k,\omega\rangle}$.

Let $\lim_{\nu\to+\infty}\Phi^\nu=\Phi$. Then

$$X_{H\circ\Phi}=X_{N_*+\mathcal{A}_*+\mathcal{B}_*+\bar{\mathcal{B}}_*+P_*}$$

on $D(\frac12 s,\frac12 r)\times O_\alpha$, where $N_*=\lim_{\nu\to\infty}N_\nu=\langle\omega_*,y\rangle+\langle\Omega z,\bar z\rangle$ and $\mathcal{A}_*=\lim_{\nu\to\infty}\mathcal{A}_\nu=\langle A_*z,\bar z\rangle$, $\mathcal{B}_*=\lim_{\nu\to\infty}\mathcal{B}_\nu=\langle B_*z,z\rangle$, $\bar{\mathcal{B}}_*=\lim_{\nu\to\infty}\bar{\mathcal{B}}_\nu=\langle\bar B_*\bar z,\bar z\rangle$. Since $\omega_*=\lim_{\nu\to\infty}\omega_\nu$ and $\mathcal{A}_*=\lim_{\nu\to\infty}\mathcal{A}_\nu$, from (4.37) it follows that $\|\omega_*-\omega(\xi)\|^L\leqslant 2\varepsilon$ and $\|X_{\mathcal{A}_*}\|^L_{r/2;D(s/2,r/2)}\leqslant 2\varepsilon$. Similarly, by (4.38) we have $\|X_{\mathcal{B}_*}\|^L_{r/2;D(s/2,r/2)}\leqslant 2\varepsilon$. $\|X_{\bar{\mathcal{B}}_*}\|^L_{r/2;D(s/2,r/2)}\leqslant 2\varepsilon$.

From the above iteration it is easy to see that the map Φ is close to the identity map with $\|\Phi-\mathrm{Id}\|^L\leqslant c\varepsilon$.

4.5. Measure estimate

The final thing left is to check the Lebegue measure of $O_\alpha=\bigcap_{\nu=1}^\infty O_\nu$. We will prove that for sufficiently small $\alpha>0$ it is not empty. According to the Iteration lemma,

$$O-O_\alpha=O-\bigcap_{\nu=1}^\infty O_\nu=\bigcup_{\nu=1}^\infty(O_\nu-O_{\nu-1})\subset\bigcup_{\nu\geqslant1,k>0}\mathcal{R}^\nu_k,\tag{4.43}$$

where

$$\mathcal{R}_k^\nu(\alpha_\nu) = \left\{\xi \in O_{\nu-1} \mid \left|\langle k, \omega_\nu\rangle^{-1}\right|,\ \left\|M_{1,\nu}^{-1}\right\|,\ \left\|M_{2,\nu}^{-1}\right\| > \frac{|k|^\tau}{\alpha_\nu}\right\},$$

with $\omega_\nu = \omega_{\nu-1} + P_{0l0}^{\nu-1}$ and $M_{1,\nu} = M_{1,\nu-1} + O(\varepsilon_{\nu-1})$ and $M_{2,\nu} = M_{2,\nu-1} + O(\varepsilon_{\nu-1})$.

It follows from Lemma A.1 that, for large $k > K_0$, $\mathcal{R}_k^\nu$ is of the measure $c(\alpha_\nu |k|^{L-\tau})^{1/L}$. For $k < K_0$, the small divisor conditions are always satisfied since $M_{1,\nu}, M_{2,\nu}$ are small perturbations of non-singular diagonal matrices.

By the Iteration lemma, $M_{1,\nu}(k), M_{2,\nu}(k)$ are $\varepsilon_{\nu-1}$-perturbation of $M_{1,\nu-1}(k), M_{2,\nu-1}(k)$, it follows that for $\xi \in O_{\nu-1}$,

$$\left\|\left(M_{1,\nu}(k)\right)^{-1}\right\|, \left\|\left(M_{2,\nu}(k)\right)^{-1}\right\| < \left(\frac{|k|^\tau}{\alpha_\nu}\right)^{1/L}$$

are automatically satisfied for $|k| < 2^\nu$ (provided that ε is small enough), which means $\mathcal{R}_k^\nu = \emptyset$ for $|k| < 2^\nu$. It follows that

$$\begin{aligned}\text{measure}(O - O_\alpha) &= \sum_{\nu=1}^{\infty} \text{measure}(O_\nu - O_{\nu-1}) \\ &\leqslant \sum_{\nu=1}^{\infty} \sum_{|k| \geqslant 2^\nu} \text{measure}\mathcal{R}_k^\nu \\ &\leqslant \left(\sum_{\nu=1}^{\infty} 2^{-\nu}\right) c\alpha^{1/L} = O\left(\alpha^{1/L}\right),\end{aligned} \tag{4.44}$$

provided that $\tau > (n+2)L$.

Appendix

LEMMA A.1. – *Suppose $\|\omega - \omega(\xi)\|^L \leqslant \varepsilon$ and $f_j(\xi)$ is a L-th continuously differentiable function and $\|f_j(\xi)\|^L \leqslant M$ $(1 \leqslant j \leqslant L)$. Let $P(\xi) = (p_{ij}(\xi))$ be a $L \times L$ matrix, L-th continuously differentiable with respect to $\xi \in O$ with*

$$\|P\|^L = \max_{|\beta| \leqslant L} \max_{1 \leqslant i,j \leqslant L} \sup_{\xi \in O} \left|\frac{\partial^\beta p_{ij}}{\partial \xi^\beta}\right| \leqslant \varepsilon.$$

Let $\mathcal{R}_k(\alpha)$ be the set such that

$$\left\|D^{-1}\right\| > \frac{c|k|^\tau}{\alpha}, \quad \forall \tau > nL,$$

where

$$D = \langle \omega, k\rangle I + \text{diag}\left(f_1(\xi), f_2(\xi), \ldots, f_L(\xi)\right) + P(\xi).$$

Then if $\varepsilon > 0$ is sufficiently small, there is a sufficiently large integer K depending on M, such that for $\forall \alpha > 0$ and $|k| > K$

$$\text{mes}\left(\mathcal{R}_k(\alpha)\right) \leqslant \left(\frac{c\alpha}{|k|^{\tau - L}}\right)^{1/L},$$

where c is a constant depending on ε, M.

Proof. – Since $\|D\| = \mathrm{O}(|k|)$, we know the norm of the inverse of a matrix is controlled by $|k|^L$ times of the lower bound of its determinant. In fact,

$$\mathcal{R}_k(\alpha) \subset \left\{\xi \colon \ |\det D| < c\frac{\alpha}{|k|^{\tau-L}}\right\}.$$

Note that

$$g(k,\xi) = \det(D) = \prod_{j=1}^{L}\big(\langle\omega,k\rangle + f_j(\xi)\big) + \sum_{l=0}^{L-1} a_l \prod_{j=1}^{L}\big(\langle\omega,k\rangle + f_j(\xi)\big)_j^l,$$

where $\|a_l\|^L \leqslant c\varepsilon$ and $l = (l_1, l_2, \ldots, l_L)$, $l_j = 0$ or 1 and $\sum_{j=1}^{L} l_j \leqslant L-1$. Note that there exists sufficiently large $K > 0$ such that if $|k| \geqslant K$ and ε is sufficiently small,

$$\begin{aligned}\left|\frac{\partial^L}{\partial\nu^L} g(k,\xi)\right| &\geqslant \left|\prod_{j=1}^{L}\left(\left\langle\frac{\partial}{\partial\nu}\omega, k\right\rangle + \frac{\partial}{\partial\nu} f_j(\xi)\right)\right| - c\varepsilon|k|^L \\ &\geqslant \big[(1-c\varepsilon)|k| - M\big]^L - c\varepsilon|k|^L,\end{aligned}$$

where $\partial^L/\partial\nu^L$ is the L-th direction derivative along the direction $\nu = k/|k|$ at ξ. Thus if ε is sufficiently small, it follows that $|\frac{\partial^L g(\xi)}{\partial\nu^L}| \geqslant \frac{1}{4}|k|^L$, which implies

$$\operatorname{mes}\big(\mathcal{R}_k(\alpha)\big) \leqslant c\left(\frac{\alpha}{|k|^{\tau-L}}\right)^{1/L} \operatorname{diam}(O)^{n-1},$$

where c is a constant only depending on M, ε, n, K and independent of α, k.

Remark. – If for $|k| \leqslant K$, $|\langle\omega,k\rangle + f_j(\xi)| \geqslant \alpha > 0$, then for sufficiently small $\varepsilon > 0$ depending on α, if $\|P\| \leqslant \varepsilon$, the matrix D is reversible and for $\forall\xi \in O$, we have $\|D^{-1}\| \leqslant c/\alpha_0$.

LEMMA A.2. – *Let D be a matrix depending on ξ and $\|D\|^L \leqslant M$. If A is reversible with $\|D^{-1}\| \leqslant N$, then $\|D^{-1}\|^L \leqslant cM^L N^{L+1}$, where c is a constant depending on L.*

Proof. – By differentiating the two sides of the equation $D^{-1}D = I$, we have $(D^{-1})' = D^{-1}D'D^{-1}$. So

$$\big\|\big(D^{-1}\big)'\big\| \leqslant \|D'\| \cdot \big\|D^{-1}\big\|^2 \leqslant MN^2.$$

Inductively it follows that $\|D^{-1}\|^L \leqslant cM^L N^{L+1}$.

LEMMA A.3. – *Let D be an open domain in a complex Banach space E with the norm $\|\cdot\|$. $X : (\Phi,\xi) \in D \times O \to E$ is a parameter dependent vector field on D, which is analytic in Φ on D and belongs to $C^L(O)$ in ξ. If $\sup_{\Phi\in D}\|X(\Phi,\cdot)\|^L \leqslant \rho$, then for each $\xi \in O$, its flow $\Phi^t(\cdot,\xi)$ exists on $D_{-\rho}$ for $|t| \leqslant 1$ and maps $D_{-2\rho}$ into $D_{-\rho}$, where $D_{-\rho} = \{\Phi \mid \Phi \in D, \operatorname{dist}(\Phi,\partial D) > \rho\}$ with ∂D indicating the boundary of D. Moreover, on $D_{-2\rho}$*

$$\big\|\Phi^t(\cdot,\xi) - \mathrm{id}\big\|^L \leqslant 2\sup_{\Phi\in D}\big\|X(\Phi,\cdot)\big\|^L \quad \textit{for } |t| \leqslant 1.$$

Proof. – The existence of the flow $\Phi^t(\cdot,\xi)$ ($|t| \leqslant 1$) is well known from general theory of differential equations. Below we estimate Φ^t. By integrating the equation of Φ^t, we have:

$$\Phi^t(\cdot,\xi) - \mathrm{id} = \int_0^t X\big(\Phi^s(\cdot,\xi),\xi\big)\,\mathrm{d}s, \quad |t| \leqslant 1.$$

Obviously, $\|\Phi^t(\cdot,\xi) - \mathrm{id}\| \leqslant \sup_{\Phi\in D}\|X(\Phi,\xi)\|$ on $D_{-2\rho}$. For the L-norms, we only prove the case $L = 1$ since for $L > 1$ the results can be obtained easily by induction and we omit the details. When $L = 1$ we have:

$$\big\|\Phi^t - \mathrm{id}\big\|^L \leqslant \int_0^t \left(\|X\|^L + \left\|\frac{\partial X}{\partial \Phi}\right\| \cdot \big\|\Phi^s\big\|^L\right)\mathrm{d}s.$$

By generalized Cauchy's inequality for analytic function on Banach space, it follows that $\|\frac{\partial X}{\partial\Phi}\| \leqslant \frac{1}{\rho}\sup_{\Phi\in D}\|X(\Phi,\cdot)\|$ on $D_{-\rho}$. Thus on $D_{-2\rho}$ we have:

$$\big\|\Phi^t - \mathrm{id}\big\|^L \leqslant \int_0^t \left(\sup_{\Phi\in D}\big\|X(\Phi,\cdot)\big\|^L + \frac{1}{\rho}\sup_{\Phi\in D}\big\|X(\Phi,\cdot)\big\|^*\big\|\Phi^s - \mathrm{id}\big\|^L\right)\mathrm{d}s \tag{A.1}$$

$$\leqslant \sup_{\Phi\in D}\big\|X(\Phi,\cdot)\big\|^L + \int_0^t \big\|\Phi^s - \mathrm{id}\big\|^L\,\mathrm{d}s. \tag{A.2}$$

By Gronwall's inequality we have for $|t| \leqslant 1$,

$$\big\|\Phi^t - \mathrm{id}\big\|^L \leqslant 2\sup_{\Phi\in D}\big\|X(\Phi,\cdot)\big\|^L.$$

LEMMA A.4. – *If a Hamiltonian vector field $W(\cdot,\xi)$ is analytic on $V = D(s-2\rho, 3\eta r)$ depending on the parameter ξ with $\|W\|^L_{r,V} < +\infty$, and $\Phi = X^t_F : U = D(s-4\rho,\eta r) \to \bar{V} = D(s-3\rho, 2\eta r)$, then $\Phi^*W = (\mathcal{D}\Phi)^{-1}W\circ\Phi$. Moreover, if*

$$\frac{1}{\rho}\|\Phi - \mathrm{id}\|^L_{\eta r,U},\ \|\mathcal{D}\Phi - \mathrm{Id}\|^L_{\eta r,\eta r,U} \leqslant cE \leqslant \frac{1}{2},$$

we have $\|\Phi^*W\|^L_{\eta r,U} \leqslant 4\|W\|^L_{\eta r,V}$.

Proof. – Since Φ is symplectic, it follows that $\Phi^*W = (\mathcal{D}\Phi)^{-1}W\circ\Phi$ and $\|\Phi^*W\|^L_{\eta r,U} \leqslant 2\|(\mathcal{D}\Phi)^{-1}\|^L_{\eta r,\eta r,U}\|W\circ\Phi\|^L_{\eta r,U}$. By Cauchy's inequality it follows that $\|W\circ\Phi\|^L_{\eta r,U} \leqslant 2\|W\|^L_{\eta r,V}$ and

$$\big\|(\mathcal{D}\Phi)^{-1}\big\|^L_{\eta r,\eta r,U} \leqslant 1 + \|\mathcal{D}\Phi - \mathrm{Id}\|^L_{\eta r,\eta r,U} + \big(\|\mathcal{D}\Phi - \mathrm{Id}\|^L_{\eta r,\eta r,U}\big)^2 + \cdots \leqslant 2.$$

Hence $\|\Phi^*W\|^L_{\eta r,U} \leqslant 4\|W\|^L_{\eta r,V}$.

Acknowledgement

The authors thank the referee for valuable suggestions on the presentation of Section 3.

REFERENCES

[1] V.I. ARNOLD (Ed.), *Dynamical Systems*, Encyclopaedia of Mathematical Sciences, Vol. 3, Sringer-Verlag, 1988.

[2] J. BOURGAIN, Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE, *Internat. Math. Res. Notices* (1994) 475–497.

[3] J. BOURGAIN, On Melnikov's persistency problem, *Math. Res. Lett.* 4 (1997) 445–458.

[4] W. CRAIG, C.E. WAYNE, Newton's method periodic solutions of nonlinear wave equations, *Comm. Pure Appl. Math.* 46 (1993) 1409–1498.

[5] L.H. ELIASSON, Perturbations of stable invariant tori for Hamiltonian systems, *Ann. Sci. Norm. Superior Pisa* 15 (1988) 115–147.

[6] S.B. KUKSIN, Perturbation theory for quasiperiodic solutions of infinite dimensional Hamiltonian systems, and its applications to the Korteweg–de Vries equations, *Mat. Sb.* 136 (178) (3) (1988); English transl. in *Math. USSR Sb.* 64 (1989) 397–417.

[7] S.B. KUKSIN, *Nearly Integrable Infinite Dimensional Hamiltonian Systems*, Lecture Notes in Math., Vol. 1556, Springer, Berlin, 1993.

[8] V.K. MELNIKOV, On some cases of conservation of conditionally periodic motions under a small change of the Hamiltonian function, *Soviet Math. Dokl.* 6 (1965) 1592–1596.

[9] V.K. MELNIKOV, A family of conditionally periodic solutions of a Hamiltonian systems, *Soviet Math. Dokl.* 9 (1968) 882–886.

[10] J. PÖSCHEL, On elliptic lower dimensional tori in Hamiltonian systems, *Math. Z.* 202 (4) (1989) 559–608.

[11] J. PÖSCHEL, A KAM-theorem for some nonlinear partial differential equations, *Ann. Scuola. Norm. Sup. Pisa Cl. Sci.* 23 (1996) 119–148.

[12] H. RÜSSMANN, On twist Hamiltonian, Talk on the Colloque International: Mécanique Céleste et Systèmes Hamiltoniens, Marseille, 1990.

[13] M.B. SEVRYUK, KAM-stable Hamiltonians, *J. Dynamics Control Systems* 1 (1995) 351–366.

[14] H. WHITNEY, Analytical extensions of differentiable functions defined in closed sets, *Trans. Amer. Math. Soc.* 36 (1934) 63–89.

[15] J. XU, Perturbations of elliptic lower dimensional tori for small perturbation of degenerate integrable Hamiltonian systems, *J. Math. Anal. Appl.* 208 (1997) 372–387.

[16] J. XU, J. YOU, A sympelctic map and its application to persistence of lower dimensional tori, Preprint, 1999.

[17] J. XU, J. YOU, Q. QIU, Invariant tori of nearly integrable Hamiltonian systems with degeneracy, *Math. Z.* 226 (1997) 375–386.

[18] J. YOU, Perturbations of lower dimensional tori for Hamiltonian systems, *J. Differential Equations* 152 (1999) 1–29.