PERSISTENCE OF THE HYPERBOLIC LOWER DIMENSIONAL
NON-TWIST INVARIANT TORUS IN A CLASS OF
HAMILTONIAN SYSTEMS

LEI WANG
Department of Mathematics and Physics, Hefei University
Hefei, MO 230601, China

QUAN YUAN
School of Machinery and Electronic Information
China University of Geosciences, Wuhan, MO 430074, China

JIA LI
School of Mathematical Physics, Xuzhou Institute of Technology
Xuzhou, MO 221000, China

(Communicated by Erik Van vleck)

ABSTRACT. We consider a class of nearly integrable Hamiltonian systems with
Hamiltonian being \(H(\theta, I, u, v) = h(I) + \frac{1}{2} \sum_{j=1}^{m} \Omega_j (u_j^2 - v_j^2) + f(\theta, I, u, v)\). By
introducing external parameter and KAM methods, we prove that, if the fre-
quency mapping has nonzero Brouwer topological degree at some Diophantine
frequency, the hyperbolic invariant torus with this frequency persists under
small perturbations.

1. Introduction and main results. Consider a real analytic Hamiltonian

\[ H(\theta, I, \zeta) = h(I) + \frac{1}{2} \langle A\zeta, \zeta \rangle + f(\theta, I, \zeta). \]

where \((\theta, I, \zeta) \in T^n \times D \times \mathbb{R}^{2m}\), with \(T^n\) being the usual \(n\)-torus obtained from \(\mathbb{R}^n\)
by identifying coordinates modulo \(2\pi\) and \(D\) a bounded open domain in \(\mathbb{R}^n\) which
is homeomorphic to an open unit ball in \(\mathbb{R}^n\); \(A\) is \(2m \times 2m\) non-singular, constant
matrices; the functions \(f(\theta, I, \zeta)\) is a small perturbation. Let \(\zeta = (u, v) \in \mathbb{R}^{2m}\).
The associated symplectic form is \(\sum_{i=1}^{n} d\theta_i \wedge dI_i + \sum_{j=1}^{m} du_j \wedge dv_j\).

If \(f = 0\), system (1) possesses a family of invariant tori \(T^n \times \{I_0\} \times \{0\} \times \{0\}\)
for all \(I_0 \in D\) and the flow on each torus is given by \(\theta(t) = \theta_0 + \omega(I_0) t\) with
\(\omega(I_0) = h_t(I_0)\) as its tangential frequency.

Some of the invariant tori can be destroyed by an arbitrarily small perturbation.\nWhether some invariant tori can persist under small perturbation is an important
problem in the perturbation theory of Hamiltonian systems and is studied by many
authors. If \(m = 0\), that is, when there is no normal frequency, Kolmogorov [8]
and Arnold [1] obtained that if \(\omega = h_I\) satisfies the Diophantine condition then the

2000 Mathematics Subject Classification. Primary: 37-XX, 70-XX; Secondary: 70H08.
Key words and phrases. Invariant tori, non-degeneracy conditions, KAM theory, Hamiltonian
systems.

* Corresponding author: woaidongda@163.com.
invariant torus with the frequency $\omega$ can persist under small perturbation. Later on, the result was extended by many authors to the case where the frequency $\omega$ depends on some parameter in a degenerate way, that is, the rank of the Jacobian matrix of $\omega$ with respect to the parameter is less than $n$, see [4, 5, 15].

When $m \neq 0$, if all the eigenvalues of $JA$ are not on imaginary axis, the tori are called hyperbolic. If all the eigenvalues of $JA$ are nonzero pure imaginary, the tori are called elliptic. In the hyperbolic case, the persistence of invariant tori for the perturbed systems has been extensively studied. Under Kolmogorov’s non-degeneracy condition, i.e.,

$$\det(\partial \omega/\partial I) = \det(h_{II}) \neq 0, \forall I \in D,$$

for a fixed Diophantine frequency $\omega_0$(see (3) below), Moser [11] proved that, if all the eigenvalues of $JA$ are distinct, the perturbed system still has a hyperbolic invariant tori with $\omega_0$ as its tangential frequency (in this case, we say that the torus persists under small perturbations). Later, Graff[7] and Zehnder[24] generalized Moser’s result by allowing multiple eigenvalues of $A$ under the same non-degeneracy condition.

In the sixties, the persistence of elliptic tori in the non-degenerate case was first observed by Melnikov [10] and later proved by Eliasson [6] and Pöschel [12]. The above results demand that the eigenvalues of $JA$ are different. However, Xu [19] proved the persistence of elliptic invariant tori for a class of nearly integrable Hamiltonian systems, where $JA$ has multiple eigenvalues. In [9] and [14], Kuksin and Pöschel extended the result to infinite dimensional Hamiltonian systems arising from some partial differential equations, such as nonlinear string wave equations.

In a recent paper of Xu and You [22], under a weaker condition than Kolmogorov’s non-degeneracy condition, they proved the persistence of the non-twist torus(see the following paragraph for the definition) in the classically nearly integrable Hamiltonian systems with one pair of conjugate variables. Motivated by [22], in this paper we consider the following reduced Hamiltonian

$$H(\theta, I, u, v) = h(I) + \frac{1}{2} \sum_{j=1}^{m} \Omega_j (u_j^2 - v_j^2) + f(\theta, I, u, v),$$

where $\text{Re}(\Omega_j) \neq 0$, $j = 1, 2, ..., m$. The corresponding Hamiltonian system reads as

$$\begin{align*}
\dot{\theta} &= H_I = h_I(I) + f_I(\theta, I, u, v) \\
\dot{I} &= -H_\theta = -f_\theta(\theta, I, u, v) \\
\dot{u} &= H_u = \Omega u + f_u(\theta, I, u, v) \\
\dot{v} &= H_v = \Omega v - f_v(\theta, I, u, v).
\end{align*}$$

(2)

Here $\Omega = \text{diag}(\Omega_1, \Omega_2, ..., \Omega_m)$.

Consider an unperturbed hyperbolic torus of system (2) ($f \equiv 0$) with tangential frequency $\omega_0 = \omega(I_0) = h_I(I_0)$. If $h_{II}(I_0) = 0$, we call this unperturbed hyperbolic torus non-twist. In this paper, we will prove that if $\omega_0$ is a Diophantine frequency and the topological degree $\text{deg}(\omega, D, \omega_0) \neq 0$, then the perturbed system (2) still has a hyperbolic invariant torus with $\omega_0$ as its tangential frequency, i.e., the torus persists under small perturbations.

Denote the complex neighbourhood of $T^n \times \{0\} \times \{0\} \times \{0\}$ by

$$D(s, r) = \{\theta : |\text{Im} \theta|_{\infty} < s\} \times \{I : |I|_1 < r^2\} \times \{u : |u|_2 < r\} \times \{v : |v|_2 < r\} \subset \mathbb{C}^n/2\pi \mathbb{Z}^n \times \mathbb{C}^n \times \mathbb{C}^m \times \mathbb{C}^m,$$
Theorem 1.1. Suppose that $D = 0$ where $\omega$ is a Diophantine frequency. Then there exists a sufficiently small positive constant $\epsilon > 0$ such that for any $\omega_0$ with $\omega_0 \neq \omega_0$, there is a perturbation $\epsilon$ of $\omega_0$ such that the system does not have a non-typical invariant torus. Moreover, the system (2) defined by $h(I) = h_1(I)$ and $\omega_0 = \omega(I_0)$, $I_0 \in D$ is a non-degenerate system.

Remark 1. We give an example as the application of Theorem 1.1. Consider the Hamiltonian system (2) defined by $h(I) = \langle \omega_0, I \rangle + I_1^2 + I_2^2 + \cdots + I_n^2$, where $\omega_0$ is a Diophantine frequency. Then $\omega(I) = h_1(I) = \omega_0 + (4I_1^2, 4I_2^2, \cdots, 4I_n^2)$, $\omega_1(I) = 12\text{diag}(I_1^2, I_2^2, \cdots, I_n^2)$. Obviously, the Kolmogorov’s non-degeneracy condition cannot be satisfied due to $\omega_1(0) = 0$, so the previous KAM theorem under the Kolmogorov’s non-degeneracy condition cannot be applied. However, we have $\text{det}(\omega_1(I)) = 12^n I_1^2 I_2^2 \cdots I_n^2 > 0, \forall I \in D \setminus \{0\}$, which implies $\text{deg}(\omega, D, \omega_0) \neq 0$. Thus, Theorem 1.1 can be applied.

As far as we know, the weakest non-degeneracy condition which can ensure the existence of a family of invariant tori is the Rüssmann’s non-degeneracy condition (see[16, 21]). This says that $\omega(p)$ does not fall into any hyperplane through the origin. However, one cannot obtain any information on the persistence of a fixed Diophantine frequency under this non-degeneracy condition. So, although the Rüssmann’s non-degeneracy condition can apply to the above example, it cannot tell us whether the frequency $\omega_0$ is broken under small perturbation.

Remark 2. It is obvious that Theorem 1.1 includes the result obtained by Moser[11]. In fact, if $A$ is hyperbolic with all eigenvalues being distinct, then the corresponding Hamiltonian system of (1) can be reduced to system (2). At the same time, according to the properties of the topological degree, it is not hard to see that $\text{det}(h_1(I)) \neq 0$ on $D$ implies $\text{deg}(\omega, D, \omega_0) \neq 0$. This implies the assertion. Moreover, if system (1) can be reduced to system (2) by allowing multiple eigenvalues of $JA$, then for the same reasons as above, our result partly includes the results obtained by Graff[7] and Zehnder[24].
Remark 3. In view of the topological degree theory, it is easy to see that if a Diophantine frequency $\omega^*$ is in a sufficiently small neighborhood of $\omega_0$, then the hyperbolic invariant tori with $\omega^*$ as its tangential frequency can also persist.

2. The parameterized form of Theorem 1.1 and its proof.

2.1. The parameterized form of Theorem 1.1. Our methods are to use the standard part of KAM iteration [13] to prove our result. At first, by linearizing the Hamiltonian system (2) at the invariant tori, we consider a parameterized Hamiltonian system instead. For any $\xi \in D$, let

$$I = \xi + y, \quad \theta = x, \quad \frac{u_i - v_i}{\sqrt{2}} = z_i, \quad \frac{u_i + v_i}{\sqrt{2}} = \bar{z}_i, \quad (i = 1, 2, \ldots, m).$$

Then,

$$H(\theta, I, u, v) = h(\xi + y) + (\Omega z, \bar{z}) + f(x, \xi + y, z, \bar{z})$$

$$= e + (\omega(\xi), y) + (\Omega z, \bar{z}) + P(\xi; x, y, z, \bar{z}),$$

where $e = h(\xi)$, $\omega(\xi) = h_1(I)$, $P(\xi; x, y, z, \bar{z}) = \int_{\xi} f(y, \xi) + f(x, \xi + y, z, \bar{z})$ and $\xi \in D$ is regarded as a parameter. Here $e$ is an energy constant and has no influence on the dynamical properties of the Hamiltonian system, so we usually omit it in the following discussion; $\omega : \xi \rightarrow \omega(\xi)$ is called the frequency mapping; and $P$ is a small perturbation term. Let

$$\Pi = \{\xi \in D \mid \text{dist}(\xi, \partial D) \geq \sigma\},$$

where $\sigma(> r)$ is a small constant. Let $\Pi_{\sigma}$ be the complex closed neighborhood of $\Pi$ in $\mathbb{C}^n$ with radius $\sigma$, that is

$$\Pi_{\sigma} = \{\xi \in \mathbb{C}^n \mid \text{dist}(\xi, \Pi) \leq \sigma\}.$$

Now the Hamiltonian $H(\xi; x, y, z, \bar{z})$ is real analytic in $(\xi; x, y, z, \bar{z})$ on $\Pi_{\sigma} \times D(s, r)$. The corresponding Hamiltonian system becomes

$$\begin{cases}
\dot{x} = H_y = \omega(\xi) + P_y(\xi; x, y, z, \bar{z}) \\
\dot{y} = -H_x = -P_x(\xi; x, y, z, \bar{z}) \\
\dot{z} = H_z = \Omega z + P_z(\xi; x, y, z, \bar{z}) \\
\dot{\bar{z}} = -H_{\bar{z}} = -\Omega \bar{z} - P_{\bar{z}}(\xi; x, y, z, \bar{z}).
\end{cases} \quad (4)$$

Thus, the persistence of hyperbolic invariant tori for system (2) is reduced to the persistence of hyperbolic invariant tori for the family of Hamiltonian systems (4) depending on the parameter $\xi \in \Pi$.

We expand $P(\xi; x, y, z, \bar{z})$ as Fourier series with respect to $x$ as follows:

$$P(\xi; x, y, z, \bar{z}) = \sum_{k \in \mathbb{Z}^n} P_k(\xi; y, z, \bar{z}) e^{i\sqrt{-1}(k,x)},$$

where $P_k(\xi; y, z, \bar{z}) = \sum_{l \in \mathbb{N}^n, p, q \in \mathbb{N}^m} P_{klpq}(\xi) y^l z^p \bar{z}^q$. Define

$$\|P\|_{\Pi_{\sigma} \times D(s, r)} = \sum_k \|P_k\|_{\sigma,r} e^{s|k|},$$

where $\|P_k\|_{\sigma,r} = \sup_{\xi \in \Pi_{\sigma}} \max_{|y| \leq r, |z| \leq r, |\bar{z}| \leq r} \sum_{l \in \mathbb{N}^n, p, q \in \mathbb{N}^m} |P_{klpq}(\xi)| y^l \bar{z}^p \bar{z}^q$. 

Theorem 2.1. Let $H(\xi; x, y, z, \bar{z}) = \langle \omega(\xi), y \rangle + \langle \Omega z, \bar{z} \rangle + P(\xi; x, y, z, \bar{z})$ be real analytic on $\Pi_\sigma \times D(s, r)$. Let $\omega_0 = \omega(\xi_0)$, with $\xi_0 \in \Pi$. Suppose that $\omega_0$ satisfies (3) and that $\deg(\omega, \Pi, \omega_0) \neq 0$. Then there exists a sufficiently small constant $\epsilon > 0$ such that if $\|P\|_{\Pi_\sigma \times D(s, r)} \leq \epsilon$, there exists $\xi_* \in \Pi$ such that the Hamiltonian system (4) has a hyperbolic invariant torus at $\xi = \xi_*$ with $\omega_0$ as its tangential frequency.

Remark 4. Theorem 1.1 can obviously follow from Theorem 2.1, and this technique was first introduced by Pöschel in [13].

2.2. The proof of Theorem 2.1 by introducing external parameter. The key idea is to introduce an artificial external parameter as used in [12, 21, 13]. Consider the following Hamiltonian system.

$$\begin{cases}
\dot{x} = H_y = \omega(\xi) + \lambda + P_y(\xi; x, y, z, \bar{z}) \\
\dot{y} = -H_x = -P_x(\xi; x, y, z, \bar{z}) \\
\dot{z} = H_z = \Omega\bar{z} + P_z(\xi; x, y, z, \bar{z}) \\
\dot{\bar{z}} = -H_{\bar{z}} = -\Omega z - P_{\bar{z}}(\xi; x, y, z, \bar{z}),
\end{cases}$$

where $H = H(\xi, \lambda; x, y, z, \bar{z}) = \langle \omega(\xi), \lambda \rangle + \langle \Omega z, \bar{z} \rangle + P(\xi; x, y, z, \bar{z})$. Obviously, system (4) is a special form of Hamiltonian system (5) with $\lambda = 0$. We will first give a KAM theorem for system (5) with parameters $(\xi, \lambda)$ and then prove Theorem 2.1.

Let

$$d = \max_{\xi, \eta \in \Pi_\sigma} |\omega(\xi) - \omega(\eta)|$$

and define

$$B(\omega, d) = \{ \lambda \in \mathbb{C}^n | \text{dist}(\lambda, \omega) < d \}.\]

Let $O = (\bigcup_{\xi \in \Pi} B(\omega(\xi), d)) \cap \mathbb{R}^n$. It follows that

$$\omega(\Pi) = \{ \omega(\xi) | \xi \in \Pi \} \subset O.$$

Let

$$O_\alpha = \{ \Delta \in O | |\langle k, \Delta \rangle| \geq \frac{\alpha}{|k|} \}, \forall k \in \mathbb{Z}^n \setminus \{0\} \}$$

and $O_{\alpha, \delta} = B(O_\alpha, \delta)$. Let $K > 0$ and $\delta = \frac{\alpha}{2|K|+1}$. Then, for all $\Delta \in O_{\alpha, \delta}$ it follows that

$$|\langle k, \Delta \rangle| \geq \frac{\alpha}{2|k|}, \quad 0 < |k| \leq K.$$

Let $M = \Pi_\sigma \times B(0, 2d + 1)$. The Hamiltonian $H(\xi, \lambda; x, y, z, \bar{z})$ is real analytic on $M \times D(s, r)$.

Theorem 2.2. There exists a small $\epsilon > 0$ such that if $\|P\|_{\Pi_\sigma \times D(s, r)} \leq \epsilon$, then we have a Cantor-like family of analytic curves in $M$,

$$\{ \Gamma_\Delta : \lambda = \lambda(\xi), \xi \in \Pi, \Delta \in O_\alpha \},$$

which are determined implicitly by the equation

$$\lambda + \omega(\xi) + h(\xi, \lambda) = \Delta,$$

where $h(\xi, \lambda)$ is a $C^\infty$-smooth function on $M$ with $|h(\xi, \lambda)| \leq \frac{\epsilon}{2}$ and $|h_\lambda(\xi, \lambda)| + |h_\xi(\xi, \lambda)| \leq \frac{1}{2}$, and a parameterized family of symplectic mappings

$$\Phi(\xi, \lambda; t) : D(s/2, r/2) \to D(s, r), \quad (\xi, \lambda) \in \Gamma = \bigcup_{\Delta \in O_\alpha} \Gamma_\Delta,$$
where $\Phi$ is $C^\infty$-smooth in $(\xi, \lambda)$ on $\Gamma$ in the sense of Whitney and analytic in $(x, y, z, \bar{z})$ on $D(s/2, r/2)$, such that for each $(\xi, \lambda) \in \Gamma_\Delta$,

$$H(\xi, \lambda; \Phi(\xi, \lambda; x, y, z, \bar{z})) = \langle \Delta, y \rangle + \langle \Omega z, \bar{z} \rangle + \hat{P}(\xi, \lambda; x, y, z, \bar{z}),$$

where $\hat{P}(\xi, \lambda; x, y, z, \bar{z})$ contains only the terms of $y^p\bar{z}^q$ with $2|l| + |p| + |q| > 2$ and the terms of $z^p\bar{z}^q$ with $|p| + |q| = 2$. Therefore, system (5) has hyperbolic invariant tori $\Phi(\xi, \lambda; T^n, 0, 0, 0)$ with $\Delta$ as their tangential frequencies.

Due to the complexity of the proof of Theorem 2.2, we first prove Theorem 2.1 by Theorem 2.2 and leave the proof of Theorem 2.2 until next section. In fact, let $\Delta = \omega_0$ and then we have an analytic curve $\Gamma_{\omega_0} : \xi \in \Pi \rightarrow \lambda(\xi)$, implicitly determined by the equation $\lambda + \omega(\xi) + h(\xi, \lambda) = \omega_0$. This yields

$$\lambda(\xi) = \omega_0 - \omega(\xi) + \hat{\lambda}(\xi), \forall \xi \in \Pi,$$

where $\hat{\lambda}(\xi) = h(\xi, \lambda(\xi))$. Furthermore, if $\epsilon$ is sufficiently small, we have $|\hat{\lambda}(\xi)| \leq 8\epsilon/r^2$ and $|\hat{\lambda}(\xi)| \leq 4\epsilon/r$. From the assumption it follows that

$$\deg(w_0 - \omega, \Pi, 0) \neq 0.$$

So if $\epsilon$ is sufficiently small we have

$$\deg(\lambda, \Pi, 0) = \deg(w_0 - \omega, \Pi, 0) \neq 0.$$

Then we have $\xi_* \in \Pi$ such that $\lambda(\xi_*) = 0$. Therefore, system (4) with $\xi = \xi_*$ has a hyperbolic invariant torus $\Phi(\xi_*, 0; T^n, 0, 0, 0)$ with $\omega_0$ as its tangential frequency. Thus, Theorem 2.1 holds. \hfill $\square$

3. The proof of Theorem 2.2 by KAM theories. Now it remains to prove Theorem 2.2. Our method is the KAM iteration and the idea is similar to [16, 21, 13, 3, 18, 2, 17, 20, 23, 25].

3.1. KAM-step. We note that the procedure of KAM-step is standard. We summarize the result for our KAM-step in the following lemma.

**Lemma 3.1** (Iteration Lemma). *Consider the Hamiltonian

$$H(\xi, \lambda; x, y, z, \bar{z}) = N(\xi, \lambda; y, z, \bar{z}) + P(\xi, \lambda; x, y, z, \bar{z}),$$

where $N(\xi, \lambda; y, z, \bar{z}) = (\Delta(\xi, \lambda), y) + \langle \Omega z, \bar{z} \rangle + Q(\xi, \lambda; x, z, \bar{z})$ with $\Delta(\xi, \lambda) = \omega(\xi) + \lambda + h(\xi, \lambda)$. Assume that the following conditions hold:

**C1.** $N$ and $P$ are real analytic on $M$ and $M \times D(s, r)$, respectively. For $0 < E < 1$, $0 < \rho < s/5$, $P$ satisfies

$$\|P\|_{M \times D(s, r)} \leq \epsilon = \alpha \rho^{\epsilon+n+1} E.$$

**C2.** The function $h$ satisfies

$$|h(\lambda, \xi, \lambda)| + |h(\xi, \lambda)| \leq \frac{1}{2}, \forall (\xi, \lambda) \in M$$

and for each $\Delta \in \mathcal{O}_\alpha$ the equation

$$\Delta(\xi, \lambda) = \omega(\xi) + \lambda + h(\xi, \lambda) = \Delta$$

defines implicitly an analytic mapping

$$\lambda : \xi \in \Pi_\sigma \rightarrow \lambda(\xi) \in B(0, 2d + 1)$$
such that $\Gamma_{\Delta} = \{(\xi, \lambda(\xi))|\xi \in \Pi_\sigma}\subset M$. Moreover, for $K > 0$ satisfying $e^{-K\rho} = E$, 
$\delta = \frac{\alpha}{2K^{r+1}}$, $L = 1 + \max_{\xi \in \Pi_{\sigma}}|\omega(\xi)|$ and $\delta = \frac{\alpha}{L}$, we have

$$U(\Gamma_{\Delta}, \delta) = \{(\xi, \lambda') \in \Pi_{\sigma} \times \mathbb{C}^n | |\lambda' - \lambda(\xi)| \leq \delta\} \subset M.$$

Then, there exist $M_+ \subset M$ and $D(s_+, r_+) \subset D(s, r)$ such that for every $(\xi, \lambda) \in M_+$ there exists a symplectic mapping

$$\Phi(\xi, \lambda; \cdot, \cdot, \cdot) : D(s_+, r_+) \to D(s, r)$$

with $\Phi$ real analytic on $M_+ \times D(s_+, r_+)$, such that

$$H_+(\xi, \lambda; x, y, z, \bar{z}) = H(\xi, \lambda; \Phi(\xi, \lambda; x, y, z, \bar{z})) = N_+(\xi, \lambda; y, z, \bar{z}) + P_+(\xi, \lambda; x, y, z, \bar{z})$$

where $N_+(\xi, \lambda; y, z, \bar{z}) = \langle \Delta_+(\xi, \lambda), y \rangle + \langle \Omega z, \bar{z} \rangle + Q_+(\xi, \lambda; x, y, z, \bar{z})$ with $\Delta_+(\xi, \lambda) = \omega(\xi) + \lambda + h(\xi, \lambda) + \hat{h}(\xi, \lambda)$, $Q_+ = Q + \hat{Q}$, $\hat{Q}$ is defined in the proof below, and the following conclusions hold:

(i) The new perturbation term $P_+$ satisfies

$$\|P_+\|_{M_+ \times D(s_+, r_+)} \leq \epsilon_+ = \alpha\rho_+^{r+n+1} E_+,$$

$$\|\hat{Q}\|_{M_+ \times D(s_+, r_+)} \leq \epsilon_+ = \alpha\rho_+^{r+n+1} E_+,$$

with

$$s_+ = s - 5\rho, \; \eta = \sqrt{E}, \; \rho_+ = \frac{1}{2}\rho, \; r_+ = \eta r, \; E_+ = cE^\frac{4}{r},$$

and

$$M_+ = \{(\xi, \lambda') \in \mathbb{C}^n \times \mathbb{C}^n | \xi \in \Pi_{\sigma - \frac{4}{r}\delta}, (\xi, \lambda) \in \Gamma, |\lambda' - \lambda| \leq \frac{1}{2}\delta\} \subset M,$$

where $\Gamma = \bigcup_{\Delta \in O_{\sigma}} \Gamma_{\Delta}$. Moreover, for the mapping $\Phi$ we have the following estimates:

$$\|W(\Phi - id)\|_{M_+ \times D(s_+, r_+)} \leq cE$$

and

$$\|W(D\Phi - Id)W^{-1}\|_{M_+ \times D(s_+, r_+)} \leq cE,$$

where $D$ is the differentiation operator with respect to $(x, y, z, \bar{z})$ and $W = \text{diag}(\frac{1}{\rho} I_n, \frac{1}{\rho} I_m, \frac{1}{\rho} I_n, \frac{1}{\rho} I_m)$ with $I_n$ and $I_m$ being the $n$-th unit matrix and $m$-th unit matrix, respectively.

(ii) $\hat{h}$ satisfies

$$|\hat{h}| \leq \frac{4\epsilon}{\tau^2} = 4\alpha\rho_+^{r+n+1} E, \; \forall (\xi, \lambda) \in M$$

and

$$|\hat{h}_\xi(\xi, \lambda)| + |\hat{h}_\z(\xi, \lambda)| \leq \frac{16\epsilon}{\tau^2\delta}, \; \forall (\xi, \lambda) \in M_+.$$

Thus, if

$$8\alpha\rho_+^{r+n+1} E \leq \frac{1}{4}\delta,$$

then the equation $\Delta_+(\xi, \lambda) = \omega(\xi) + \lambda + h_+(\xi, \lambda) = \Delta$ with $h_+(\xi, \lambda) = h(\xi, \lambda) + \hat{h}(\xi, \lambda)$ determines implicitly an analytic mapping

$$\lambda_+ : \xi \in \Pi_{\sigma_+} \to \lambda_+(\xi) \in B(0, 2d + 1)$$

with $\sigma_+ = \sigma - \frac{1}{2}\delta$, which satisfies

$$|\lambda_+(\xi) - \lambda(\xi)| \leq \frac{8\epsilon}{\tau^2} = 8\alpha\rho_+^{r+n+1} E \leq \frac{1}{4}\delta$$
and 

$$\Gamma^+_\Delta = \{ (\xi, \lambda_+(\xi)) | \xi \in \Pi_{s_+} \} \subset M_+.$$  \hspace{1cm} (9)

Let $\delta_+ = \frac{\alpha}{2K^+_1}$ with $K_+$ satisfying $e^{-\rho_+K_+} = E_+$. If 

$$\delta_+ < \frac{1}{4} \delta,$$  \hspace{1cm} (10)

then for all $\Delta \in O_\alpha$ we have $U(\Gamma^+_\Delta, \delta_+) \subset M_+$.

Remark 5. Lemma 3.1 is actually one step in our KAM iteration. If (7) and (10) hold and $h_+$ satisfies (6), then the assumptions C1 and C2 hold for the transformed Hamiltonian $H_+$ and so the KAM step can iterate.

Proof of the iteration lemma. First, we give an equivalent deformation to the Hamiltonian of system (5) as follows:

$$H(\xi, \lambda; x, y, z, \bar{z}) = N(\xi, \lambda; y, z, \bar{z}) + P(\xi, \lambda; x, y, z, \bar{z})$$

$$= (\Delta, y) + (\Omega, \bar{z}) + Q(\xi, \lambda; x, z, \bar{z}) + P'(\xi, \lambda; x, y, z, \bar{z}),$$

where $Q + P' = P$. Here $Q$ contains only the terms of $z^p \bar{z}^q$ with $|p| + |q| = 2$, i.e.,

$$Q(\xi, \lambda; x, z, \bar{z}) = \frac{1}{2} (Q_1(\xi, \lambda; x)z, \bar{z}) + (Q_2(\xi, \lambda; x)z, \bar{z}) + \frac{1}{2} (Q_3(\xi, \lambda; x)\bar{z}, \bar{z}).$$

For simplifying notations, we still denote $(\Delta, y) + (\Omega, \bar{z}) + Q(\xi, \lambda; x, z, \bar{z})$ and $P'$ by $N$ and $P$, respectively.

A. Truncation. Consider the Fourier series of $P$ on $M \times D(s, r)$

$$P = \sum_{k \in \mathbb{Z}^n, l \in \mathbb{N}^n, p \in \mathbb{N}^n} P_{klp}(\xi, \lambda) y^l z^p \bar{z}^q e^{\sqrt{-1}(k,x)}.$$ 

Let 

$$R^K = \sum_{k \in \mathbb{Z}^n, |k| \leq K} (P_{k000}(\xi, \lambda) + (P_{k100}(\xi, \lambda), y)$$

$$+ (P_{k010}(\xi, \lambda), z) + (P_{k001}(\xi, \lambda), \bar{z})) e^{\sqrt{-1}(k,x)}$$

$$= R_0(\xi, \lambda; x) + (R_1(\xi, \lambda; x), y) + (R_2(\xi, \lambda; x), z) + (\bar{R}_2(\xi, \lambda; x), \bar{z}).$$

Let $P - R^K = D_1 + D_2$, where 

$$D_1 = \sum_{|k| > K, |l| + |p| + |q| = 1} P_{klpq}(\xi, \lambda) y^l z^p \bar{z}^q e^{\sqrt{-1}(k,x)}, \forall k \in \mathbb{Z}^n,$$

$$D_2 = \sum_{2(|l| + |p| + |q|) > 2, |p| + |q| \neq 2} P_{klpq}(\xi, \lambda) y^l z^p \bar{z}^q e^{\sqrt{-1}(k,x)}, \forall k \in \mathbb{Z}^n.$$ 

By above definitions and Lemma 4.2 in Appendix, we have the following estimates

$$\|R^K\|_{M \times D(s, r)} \leq 2\|P\|_{M \times D(s, r)} < 2\epsilon$$

and 

$$\|D_1\|_{M \times D(s - \rho, r)} \leq cK^ne^{-K\rho}\epsilon, \|D_2\|_{M \times D(s, 2\eta r)} < c\eta^3 \epsilon.$$  \hspace{1cm} (11)

B. Homology equation. Now we try to find a symplectic coordinate transformation to eliminate as many terms in $R^K$ as possible. The transformation is generated by a Hamiltonian flow mapping at time 1, that is, $\Phi = X_F^1|_{t=1}$, where $F = F_0(\xi, \lambda; x) + \langle F_1(\xi, \lambda; x), y \rangle + \langle F_2(\xi, \lambda; x), z \rangle + \langle \bar{F}_2(\xi, \lambda; x), \bar{z} \rangle$ defined in smaller domain $D(s -$
According to Lemma 4.4 in Appendix, (14) exists unique solution $F$. Then (14) and (15) can be transformed into the following form

$$2 \rho, r)$$ is the generation function. The original method of the definition of $F$ stems from [1, 13]. It follows that [24, 12, 20]

$$H \circ \Phi = N \circ X_F^{t=1} + R^K \circ X_F^{t=1} + (P - R^K) \circ X_F^{t=1}$$

$$= N + \{N, F\} + R^K + \int_0^1 (1 - t) \{\{N, F\}, F\} \circ X_F^t \, dt$$

$$+ \int_0^1 \{R^K, F\} \circ X_F^t \, dt + (P - R^K) \circ X_F^{t=1}$$

$$= N + \{N, F\} + R^K + \int_0^1 \{R^K, F\} \circ X_F^t \, dt + (P - R^K) \circ \Phi,$$

where $\{\cdot, \cdot\}$ is the Poisson bracket and $R^K_n = (1 - t) \{N, F\} + R^K$. Then

$$\{N, F\} = \langle N_x, F_y \rangle - \langle N_y, F_x \rangle + \langle N_z, F_z \rangle$$

$$= \langle Q_x, F_1(x) \rangle - \langle \Delta, (F_0)_x \rangle + (F_1)_x y + (F_2)_x z + (F_2)_x \bar{z}$$

$$+ \langle \Omega z, F_2 \rangle + \langle Q_z, F_2 \rangle - \langle \Omega z, F_2 \rangle - \langle Q_z, F_2 \rangle$$

$$= \langle Q_x, F_1 \rangle - \partial_\Delta F_0 + F_1 y + F_2 z + F_2 \bar{z} - \langle \Omega z, F_2 \rangle$$

$$- \langle Q_2 z, F_2 \rangle + \langle Q_1 z, F_2 \rangle + \langle \Omega z, F_2 \rangle + \langle Q_2 \bar{z}, F_2 \rangle - \langle Q_3 \bar{z}, F_2 \rangle.$$  

In order to get rid of the first-order of $R^K$ on $y, z, \bar{z}$, we choose $F$ such that

$$\{N, F\} - \langle Q_x, F_1 \rangle + R - [R_0] - [R_1] y = 0,$$

where $[R_i]$ denotes the average of $R_i$ on $T^n (i = 0, 1)$. More precisely, we have

$$\partial_\Delta F_0 = R_0 - [R_0],$$

$$\partial_\Delta F_1 = R_1 - [R_1],$$

$$\partial_\Delta F_2 + \Omega F_2 + \langle Q_2, F_2 \rangle - \langle Q_1, F_2 \rangle = R_2,$$

$$\partial_\Delta \bar{F}_2 - \Omega \bar{F}_2 - \langle Q_2, F_2 \rangle + \langle Q_3, F_2 \rangle = \bar{R}_2.$$  

It follows that

$$\Delta : (\xi, \lambda) \in U(\Gamma, \delta) \to \Delta(\xi, \lambda) \in O_{\alpha, \delta}.$$  

According to Lemma 4.3 in Appendix, it is easy to see that (12) exists unique solution $F_0(\xi, \lambda; x)$ and (13) exists unique solution $F_1(\xi, \lambda; x)$. Moreover, we have

$$\|F_0\|_{U(\Gamma, \delta) \times D(s-2\rho, r)} \leq c \frac{\epsilon}{\alpha \rho^{r+n}}, \quad \|F_1\|_{U(\Gamma, \delta) \times D(s-2\rho, \bar{r})} \leq c \frac{\epsilon}{\alpha \rho^{r+n} r^2}. \quad (16)$$

Let

$$g(x) = \left( \begin{array}{c} R_2 \\ \bar{R}_2 \end{array} \right), \quad f(x) = \left( \begin{array}{c} F_2 \\ \bar{F}_2 \end{array} \right), \quad \Theta = \left( \begin{array}{cc} -\Omega & 0 \\ 0 & \Omega \end{array} \right), \quad \bar{\Theta} = \left( \begin{array}{cc} -Q_2 & Q_1 \\ -Q_3 & Q_2 \end{array} \right).$$

Then (14) and (15) can be transformed into the following form

$$g = \langle \Delta, \partial_x f \rangle - (\Theta + \bar{\Theta}) f.$$  

According to Lemma 4.4 in Appendix, (14) exists unique solution $F_2(\xi, \lambda; x)$ and (15) exists unique solution $\bar{F}_2(\xi, \lambda; x)$. Furthermore, by Cauchy estimates (Lemma 4.1 in Appendix), we have

$$\|F_2\|_{U(\Gamma, \delta) \times D(s-2\rho, \bar{r})} \leq c \frac{\epsilon}{\bar{r}}, \quad \|\bar{F}_2\|_{U(\Gamma, \delta) \times D(s-2\rho, \bar{r})} \leq c \frac{\epsilon}{\bar{r}}.$$  

(17)
Thus, by the symplectic mapping \( \Phi = X^t_{P} \), we get
\[
H \circ \Phi = (\Delta, y) + (\Omega z, \bar{z}) + Q + [R_0] + ([R_1], y)
\]
\[+ (Q_z, F_1) + \int_0^1 \{ R_t, F \} \circ X^t_{P} dt + (P - R) \circ \Phi = N_+ + P_+,
\]
where
\[N_+ = e_+(\xi, \lambda) + (\Delta_+, y) + (\Omega z, \bar{z}) + Q_+(\xi, \lambda; x, z, \bar{z}),
\]
\[P_+ = \int_0^1 \{ R_t, F \} \circ X^t_{P} dt + (P - R) \circ \Phi + (Q_z, F_1) - \bar{Q},
\]
where \( e_+(\xi, \lambda) = [R_0], \Delta_+ = \Delta + [R_1], \bar{Q} \) means the terms of \( z^p \bar{z}^q \) in \( \int_0^1 \{ R_t, F \} \circ X^t_{P} dt + (P - R) \circ \Phi \) with \(|p| + |q| = 2\), and \( Q_+ = Q + \bar{Q} \).

C. Estimates for the symplectic mapping. By \( \| R \|_{M \times D(s, r)} \leq 2\epsilon \) and Cauchy estimates, we get
\[
\| R_0 \|_{U(\Gamma, \delta) \times D(s - \rho, r)} \leq 2\epsilon,
\]
\[
\| R_1 \|_{U(\Gamma, \delta) \times D(s - \rho, \bar{z})} \leq \frac{4\epsilon}{r},
\]
(18)
\[
\| R_2 \|_{U(\Gamma, \delta) \times D(s - \rho, \bar{z})} \leq \frac{2\epsilon}{r}.
\]
(19)
Meanwhile, by (16) and (17), we have
\[
\| F \|_{U(\Gamma, \delta) \times D(s - 2\rho, \bar{z})} \leq \frac{\epsilon}{\alpha \rho^2 + n + \epsilon} + \frac{\epsilon}{\alpha \rho^2 + n + \epsilon} + \frac{\epsilon}{\alpha \rho^2 + n + \epsilon} + \frac{\epsilon}{\alpha \rho^2 + n + \epsilon}.
\]
Combining this with Cauchy estimates, we have
\[
\| F \|_{U(\Gamma, \delta) \times D(s - 3\rho, \bar{z})} \leq \frac{\epsilon}{\alpha \rho^2 + n + \epsilon},
\]
(20)
\[
\| F \|_{U(\Gamma, \delta) \times D(s - 2\rho, \bar{z})} \leq \frac{\epsilon}{\alpha \rho^2 + n + \epsilon},
\]
(21)
Thus, if \( 0 < cE \leq \eta \leq \frac{\epsilon}{\alpha} \), then for all \((\xi, \lambda) \in U(\Gamma, \delta)\), we have
\[
\Phi(\xi, \lambda; \cdot, \cdot, \cdot, \cdot) = X^t_{P} \big|_{t=1} : D(r_{\frac{r}{4}}, s - 4\rho) \rightarrow D(r_{\frac{r}{2}}, s - 3\rho),
\]
where
\[
\Phi(\xi, \lambda; x, y, z, \bar{z}) = (\varphi(\xi, \lambda; x), \psi(\xi, \lambda; x, y, z, \bar{z}), b(\xi, \lambda; x, z), \bar{b}(\xi, \lambda; x, \bar{z})),
\]
where \( \psi \) is affine linear transformation on \( y, z \) and \( \bar{z} \); \( b \) and \( \bar{b} \) are the translation transformation on \( z \) and \( \bar{z} \) respectively. Meanwhile, for all \((\xi, \lambda) \in U(\Gamma, \delta)\), it is easy to get
\[
|\varphi - id| \leq \frac{\epsilon}{\alpha \rho^2 + n + \epsilon} \leq cE \rho < \rho,
\]
\[
|\psi - id| \leq \frac{\epsilon}{\alpha \rho^2 + n + \epsilon} \leq cE \rho^2 < \eta^2 r^2,
\]
\[
|b - id| \leq \frac{\epsilon}{\alpha \rho^2 + n + \epsilon} \leq cp \rho < \rho,
\]
\[
|\bar{b} - id| \leq \frac{\epsilon}{\alpha \rho^2 + n + \epsilon} \leq cp \rho < \rho.
\]
Thus, for all \((\xi, \lambda) \in U(\Gamma, \delta)\), we have
\[
\Phi(\xi, \lambda; \cdot, \cdot, \cdot, \cdot) = X^t_{P} \big|_{t=1} : D(\eta \rho, s - 5\rho) \rightarrow D(2\eta \rho, s - 4\rho)
\]

and $D\Phi = \begin{pmatrix} \varphi_x & 0 & 0 & 0 \\ \psi_x & \psi_y & \psi_z & \psi_y \\ b_x & 0 & Id & 0 \\ b_x & 0 & 0 & Id \end{pmatrix}$. By Cauchy estimates, we have
\[
\|\varphi_x - Id\|_{U(\Gamma, \delta) \times D(s-5, \eta r)} < c \frac{\epsilon}{\alpha \rho^{r+n+1} r^2},
\]
\[
\|\psi_y - Id\|_{U(\Gamma, \delta) \times D(s-5, \eta r)} < c \frac{\epsilon}{\alpha \rho^{r+n+1} r^2},
\]
\[
\|\psi_z\|_{U(\Gamma, \delta) \times D(s-5, \eta r)} < c \frac{\epsilon}{\alpha \rho^{r+n+1} r},
\]
\[
\|\psi_y\|_{U(\Gamma, \delta) \times D(s-5, \eta r)} < c \frac{\epsilon}{\alpha \rho^{r+n+1} r},
\]

Thus, the estimates for $\Phi$ hold.

D. Estimates of the error terms. From $\hat{h}(\xi, \lambda) = [R_1]$, we have
\[
\|\hat{h}\|_M = \|R_1\|_M \leq \frac{\|P\|_{M \times D(s, r)}}{r^2} \leq \frac{\epsilon}{r^2},
\]
indicating that the estimate for $\hat{h}$ of in conclusion (ii) holds. Noticing that the set $O_\alpha$ is closed, it is easy to see that $M_+$ is also closed. Furthermore, we have $\text{dist}(M_+, \partial M) \geq \frac{\epsilon}{2 \delta}$. By Cauchy estimates, the estimates for $\hat{h}_\xi$ and $\hat{h}_\lambda$ hold. Moreover, by (7) and the implicit function theorem, if $|h_{+\lambda}(\xi, \lambda)| \leq \frac{1}{2}, \forall (\xi, \lambda) \in M,$

the equation
\[
\Delta_+ (\xi, \lambda) = \omega(\xi) + \lambda + h_+(\xi, \lambda) = \Delta \in O_\alpha
\]
determines implicitly an analytic mapping
\[
\lambda_+ : \xi \in \Pi_{\sigma_+} \to \lambda_+(\xi) \in B(0, 2d + 1).
\]
Then
\[
\omega(\xi) + \lambda_+(\xi) + h(\xi, \lambda_+(\xi)) + \hat{h}(\xi, \lambda_+(\xi)) = \Delta.
\]
Combining this with $\omega(\xi) + \lambda(\xi) + h(\xi, \lambda(\xi)) = \Delta$, we get
\[
|\lambda_+(\xi) - \lambda(\xi)| \leq \|h_\lambda\|_{\lambda_+(\xi) - \lambda(\xi)} + 4\frac{\epsilon}{r^2} \leq \frac{1}{2} |\lambda_+(\xi) - \lambda(\xi)| + \frac{4\epsilon}{r^2},
\]
which means
\[
|\lambda_+(\xi) - \lambda(\xi)| \leq \frac{8\epsilon}{r^2} \leq \frac{\delta}{4}.
\]
Meanwhile, we have
\[
\Gamma_+^\Delta = \{(\xi, \lambda_+(\xi))|\xi \in \Pi_{\sigma_+}\} \subset M_+.
\]
Thus, the statements (8) and (9) hold. By (10), it is easy to see that $U(\Gamma_+^\Delta, \delta_+) \subset M_+$. Thus, conclusion (ii) holds.
Now, it remains to prove conclusion (i). Since $F$ satisfies the following expressions

$$\{N, F\} + R^K = [R_0] + \langle [R_1], y\rangle + (Q_x, F_1)$$

and

$$\langle Q_x, F_1\rangle, F\rangle = \langle \partial(Q_x, F_1) \rangle \partial_x F_1 + \langle \partial(Q_x, F_1) \rangle \partial_z F_2 - \langle \partial(Q_x, F_1) \rangle \partial_z F_2,$$

it follows that $\hat{Q}$ stems from $\langle \partial(Q_x, F_1) \rangle F_1.$

For all $(\xi, \lambda) \in M,$ we have

$$\{N, F\}, F\rangle = \{[R_0] + \langle [R_1], y\rangle + (Q_x, F_1) - R^K, F\}$$

$$= \{[R_0], F\} + \langle [R_1], y\rangle + \{Q_x, F_1\}, F\rangle - \{R^K, F\}.$$

By (18), (19), (20), and (21), we have the following estimate

$$\|\{R^K, F\}\|_{U(\Gamma, \delta) \times D(s-3\rho, r)} \leq \|R^K\|_{U(\Gamma, \delta) \times D(s-3\rho, r)} \|F\|_{U(\Gamma, \delta) \times D(s-3\rho, r)}$$

$$+ \|R^K\|_{U(\Gamma, \delta) \times D(s-3\rho, r)} \|F\|_{U(\Gamma, \delta) \times D(s-3\rho, r)}$$

$$+ \|R^K\|_{U(\Gamma, \delta) \times D(s-3\rho, r)} \|F\|_{U(\Gamma, \delta) \times D(s-3\rho, r)}$$

$$+ \|R^K\|_{U(\Gamma, \delta) \times D(s-3\rho, r)} \|F\|_{U(\Gamma, \delta) \times D(s-3\rho, r)}$$

$$< 2c \frac{\epsilon}{\rho^\alpha r + n + 1} + c \frac{\epsilon}{\rho^\alpha r + n + 1} + c \frac{\epsilon}{\rho^\alpha r + n + 1} + c \frac{\epsilon}{\rho^\alpha r + n + 1}$$

$$< c \frac{\epsilon}{\rho^\alpha r + n + 1}.$$

For the similar reasons, we have

$$\|\{[R_0], F\}\|_{U(\Gamma, \delta) \times D(s-3\rho, r)} = 0 < c \frac{\epsilon^2}{\rho^\alpha r + n + 1}$$

$$\|\{[R_1], y\}, F\rangle\|_{U(\Gamma, \delta) \times D(s-3\rho, r)} \leq c \frac{\epsilon}{\rho^\alpha r + n + 1} \leq c \frac{\epsilon^2}{\rho^\alpha r + n + 1}$$

$$\|\partial(Q_x, F_1)\|_{U(\Gamma, \delta) \times D(s-3\rho, r)} \leq c \frac{\epsilon^2}{\rho^\alpha r + n + 1} \leq c \frac{\epsilon^2}{\rho^\alpha r + n + 1}$$

$$\|\partial(Q_x, F_1)\|_{U(\Gamma, \delta) \times D(s-3\rho, r)} < c \frac{\epsilon^2}{\rho^\alpha r + n + 1}$$

$$\|\{Q_x, F_1\}, F\rangle\|_{U(\Gamma, \delta) \times D(s-4\rho, r)} \leq c \frac{\epsilon^3}{\rho^\alpha r^2 + n + 2 + 2r^2} < c \frac{\epsilon^3}{\rho^\alpha r^2 + n + 2 + 2r^2}$$

Then we have

$$\|\hat{Q}\|_{U(\Gamma, \delta) \times D(s-5\rho, r)} \leq \|\partial(Q_x, F_1)\|_{U(\Gamma, \delta) \times D(s-4\rho, r)} < c \frac{\epsilon^2}{\rho^\alpha r + n + 1}$$

$$\|\{Q_x, F_1\}, F\rangle\|_{U(\Gamma, \delta) \times D(s-3\rho, r)} < c \frac{\epsilon^2}{\rho^\alpha r + n + 1}.$$
and
\[
\|\{N, F\} \|_{U(\Gamma, \delta) \times D(s-3\rho, \xi)} \\
\leq \|\{R_0, F\} \|_{U(\Gamma, \delta) \times D(s-3\rho, \xi)} + \|\{Q_x, F, k\} \|_{U(\Gamma, \delta) \times D(s-3\rho, \xi)} \\
+ \|\{R^K, F\} \|_{U(\Gamma, \delta) \times D(s-3\rho, \xi)} \leq c_2 \frac{c^2}{\alpha \rho^{r+n+1} r^2}.
\]

Thus, we obtain
\[
\| (1-t)\{N, F\} + R, F \|_{U(\Gamma, \delta) \times D(s-3\rho, \xi)} < c_2 \frac{c^2}{\alpha \rho^{r+n+1} r^2}.
\]

By (11), we have
\[
\| (P - R^K) \circ \Phi \|_{U(\Gamma, \delta) \times D(s-5\rho, \eta r)} \\
\leq \| P - R^K \|_{U(\Gamma, \delta) \times D(s-4\rho, 2\eta r)} \\
\leq \| D_1 \|_{U(\Gamma, \delta) \times D(s-4\rho, 2\eta r)} + \| D_2 \|_{U(\Gamma, \delta) \times D(s-4\rho, 2\eta r)} \\
< c(K^n e^{-K_\rho} + \eta^2) \epsilon.
\]

As usual, by the choice of the parameters, we have the following estimate
\[
\| P_+ \|_{M_+ \times D(s-5\rho, \eta r)} = \int_{0}^{1} \{R^K, F\} \circ X^t dt + (P - R^K) \circ \Phi \\
- \tilde{Q} + \{Q_x, F, k\} \|_{M_+ \times D(s-5\rho, \eta r)} \\
\leq c[\frac{c^2}{\alpha \rho^{r+n+1} r^2} + (K^n e^{-K_\rho} + \eta^2) \epsilon] \leq \alpha r_+ \rho_+^{r+n+1} E_+ = \epsilon_+.
\]

Then we also have
\[
\| \tilde{Q} \|_{M_+ \times D(s-5\rho, \eta r)} \leq \epsilon_+.
\]

Thus conclusion (i) is proved.

3.2. Iteration. Now we choose some suitable parameters so that the above iteration can go on infinitely.

At the initial step, let \( \rho_0 = \frac{s}{20}, r_0 = r, \epsilon_0 = \alpha r_0^2 \rho_0^{r+n+1} E_0, \) and \( \eta_0 = E_0^{\frac{3}{2}}. \) Let \( K_0 \) satisfy \( e^{-K_0 \rho_0} = E_0. \)

Assume the above parameters are all well defined for \( j. \) Then, we define \( \rho_{j+1} = \frac{2}{j} \rho_j, \)
\( \eta_j = E_j^{\frac{3}{2}}, r_{j+1} = \eta_j r_j, E_{j+1} = c E_j^{\frac{3}{2}}. \) Moreover, \( \epsilon_{j+1}, \eta_{j+1}, \) and \( K_{j+1} \) are defined similar to the initial step \( (\epsilon_{j+1} = \alpha r_{j+1}^2 \rho_{j+1}^{r+n+1} E_{j+1}, \eta_{j+1} = E_{j+1}^{\frac{3}{2}}, K_{j+1} = \frac{-\ln E_{j+1}}{\rho_{j+1}}). \)

Let \( M_0 = \Pi_0 \times B(0, 2d+1), D_0 = D(s_0, r_0) \) and \( H_0 = H. \) By the KAM-step, we have a sequence of closed sets \( \{M_j\} \) with \( M_{j+1} \subset M_j \) and a sequence of symplectic mappings \( \{\Phi_j\} \) such that for each \( (\xi, \lambda) \in M_{j+1}, \Phi_j(\xi, \lambda; \cdots, \cdots) : D_{j+1} \rightarrow D_j, \) where \( D_j = D(s_j, r_j). \) Furthermore, we have the following estimates
\[
\| W_j(\Phi_j - \text{id}) \|_{M_{j+1} \times D_{j+1}} \leq c E_j, \| W_j(D\Phi_j - \text{id}) W_j^{-1} \|_{M_{j+1} \times D_{j+1}} \leq c E_j.
\]
Let $\Phi^j = \Phi_0 \circ \Phi_1 \circ \cdots \circ \Phi_{j-1}$ with $\Phi^0 = id$ and

$$H_j = H \circ \Phi^j = N_j + P_j,$$

where $N_j = (\Delta_j(\xi, \lambda), y) + (\Omega z, \tilde{z}) + Q_j$ with $\Delta_j(\xi, \lambda) = \omega(\xi) + \lambda h_j(\xi, \lambda)$ and $Q_j = Q_{j-1} + \tilde{Q}_j$. Let $\delta_j = \frac{\alpha}{2K_j^2}$, $\delta = \frac{\delta}{2}$ with $L = 1 + \max_{\xi \in \Pi_{\sigma}} |\omega(\xi)|$, and $\sigma_j = \sigma_{j-1} - \frac{1}{2}\delta_{j-1}$ with $\sigma_0 = \sigma$. From the iteration lemma, we know that for $\Delta \in O_{\alpha}$ and $\xi \in \Pi_{\sigma_j}$ the equation

$$\Delta_j(\xi, \lambda) = \omega(\xi) + \lambda h_j(\xi, \lambda) = \Delta$$

on $M_j$ defines implicitly an analytic mapping $\lambda = \lambda_j(\xi)$, whose graph forms an analytic curve $\Gamma^{j}_\Delta$ in $M_j$.

Let $\Gamma_j = \bigcup_{\Delta \in O_{\alpha}} \Gamma^{j}_\Delta$. We have

$$M_{j+1} = \{ (\xi, \lambda') \in \mathbb{C}^n \times \mathbb{C}^n | \xi \in \Pi_{\sigma_{j+1}}, (\xi, \lambda) \in \Gamma_j, |\lambda' - \lambda| \leq \frac{1}{2}\delta_j \}$$

Evidently, it follows that $M_{j+1} \subset M_j$ and $\text{dist}(M_{j+1}, \partial M_j) \geq \frac{1}{2}\delta_j$.

Let $\hat{h}_j(\xi, \lambda) = \Delta_{j+1}(\xi, \lambda) - \Delta_j(\xi, \lambda)$. Then we have

$$|\hat{h}_j(\xi, \lambda)| \leq \frac{4\epsilon_j}{r_j^2}, \forall (\xi, \lambda) \in M_j$$

and

$$|\hat{h}_j(\xi, \lambda)| + |\hat{h}_j(\xi, \lambda)| \leq \frac{16\epsilon_j}{\delta_j r_j^2}, \forall (\xi, \lambda) \in M_{j+1}.$$

Moreover, we have

$$|\lambda_{j+1}(\xi) - \lambda_j(\xi)| \leq \frac{8\epsilon_j}{r_j^2}, \forall (\xi, \lambda) \in M_{j+1}.$$

Thus it follows that

$$||P_j||_{M_j \times D_j} \leq \epsilon_j = \alpha r_j^{n+1} r_j^2 E_j.$$

3.3. Convergence of the iteration. Now we prove convergence of the KAM iteration. In view if the structure of $\Phi_j$ and the relevant estimates, using the same way as in [22] and [13], we get that the sequence $\{ \Phi^j \}$ converges to $\Phi$ on $D\left(\frac{s}{2}, \frac{s}{2}\right)$ and

$$||W_0(\Phi - id)||_{M_j \times D\left(\frac{s}{2}, \frac{s}{2}\right)} \leq cE_0.$$ 

Now we consider the convergence of $\{h_j\}$. Let $B_j = \frac{8\epsilon_j}{\delta_j r_j^2}$. It follows that

$$\frac{B_{j+1}}{B_j} = \frac{(K_{j+1} r_{j+1})^{n+1}}{(K_{j} r_{j})^{n+1}} \left( \frac{K_{j}}{K_{j+1}} \right)^n \frac{E_{j+1}}{E_j} = \left( \frac{1}{2} \right)^n \frac{a_{j+1}}{a_j} e^{-a_{j+1}} e^{-a_j},$$

where $a_j = K_j r_j$. By the iteration $E_{j+1} = cE_j^4$, if $E_0$ is sufficiently small, $E_j$ are all sufficiently small and $a_j$ are all sufficiently large. Since the function $a^{n+1} e^{-a}$ is decreasing on $a$ in interval $(\tau + 1, \infty)$, there exists a sufficiently small $E_0 > 0$ such that

$$\frac{B_{j+1}}{B_j} \leq \left( \frac{1}{2} \right)^n \frac{a_{j+1}}{a_j} e^{-a_{j+1}} e^{-a_j} \leq \frac{1}{4}.$$
Similarly, by \(\xi, \lambda\), it is easy to see that for all \((\xi, \lambda)\)

\[
B_j \leq \left(\frac{1}{4}\right)^j B_0 \leq \frac{1}{4}, \quad \forall j \geq 0.
\]

Then the assumption (7) holds. Moreover,

\[
\frac{\delta_{j+1}}{\delta_j} = \left(\frac{K_j}{K_{j+1}}\right)^{r+1} = \left(\frac{1}{2}\right)^{r+1}\left(\frac{a_j}{a_{j+1}}\right)^{r+1} \leq \frac{1}{4}.
\]

This implies the condition (10).

Let \(\sigma_* = \sigma - \frac{1}{2} \sum_{j=0}^{\infty} \delta_j\). From \(\frac{\delta_{j+1}}{\delta_j} \leq \frac{1}{4}\), we obtain \(\sigma_* \geq \sigma - \frac{2}{3} \delta_0\). If \(E_0\) is sufficiently small such that \(\delta_0 \leq \varsigma\), then we have \(\sigma_* \geq \frac{1}{3} \sigma\).

By iteration we know \(h_j = \sum_{i=0}^{j-1} h_i\). Combining this with the estimates for \(h_i\), it is easy to see that for all \((\xi, \lambda) \in M_j\),

\[
|h_j(\xi, \lambda)| \leq \sum_{i=0}^{j-1} \frac{1}{2} \delta_i B_i \leq \frac{1}{2} \delta_0 \sum_{i=0}^{j-1} B_i \leq \frac{1}{2} \cdot \frac{3}{4} \delta_0 B_0 \leq \frac{8\epsilon}{r^2}.
\]

Similarly, by \(\rho_0 = \frac{-\ln E_0}{K_0}\), it follows that for all \((\xi, \lambda) \in M_j\),

\[
\left|\frac{\partial h_j(\xi, \lambda)}{\partial \xi}\right| + \left|\frac{\partial h_j(\xi, \lambda)}{\partial \lambda}\right| \leq \sum_{i=0}^{j-1} B_i \leq 2B_0 = \frac{8\epsilon \rho_0 r^{n+1} E_0}{\alpha_0 \frac{1}{2K_0^{r+1}}} = 16K_0^{-n} E_0 \left(\frac{1}{r^{n+1}}\right) E_0 \leq 16L(-\ln E_0)^{r+1} E_0
\]

Since \((-\ln E_0)^{r+1} E_0 \to 0(E_0 \to 0)\), so if \(E_0\) is sufficiently small we have

\[
|h_j(\xi, \lambda)| + |h_j(\xi, \lambda)| \leq \frac{1}{2}, \forall (\xi, \lambda) \in M_j,
\]

and so the assumption (6) holds for all \(j\).

Let \(h = \lim_{j \to \infty} h_j\). Then for \((\xi, \lambda) \in M_*\) we have

\[
|h(\xi, \lambda)| \leq \frac{8\epsilon}{r^2}
\]

and

\[
|h(\xi, \lambda)| + |h(\xi, \lambda)| \leq 16L(-\ln E_0)^{r+1} E_0 \leq \frac{1}{2}.
\]

Now, in the same way, we show that \(\{\lambda_j\}\) is also convergent on \(\Pi_{\sigma_*}\). In fact, for \(i > j\) it follows that

\[
|\lambda_i(\xi) - \lambda_j(\xi)| \leq |\lambda_i(\xi) - \lambda_i(\xi)| + |\lambda_i(\xi) - \lambda_i(\xi)| + \cdots + |\lambda_j(\xi) - \lambda_j(\xi)|
\]

\[
\leq \sum_{i=j}^{i-1} B_i \delta_i \leq 2B_j \delta_j \leq \frac{\delta_j}{2}.
\]

Since \(\lim_{j \to \infty} \delta_j = 0\), so \(\{\lambda_j\}\) is uniformly convergent on \(\Pi_{\sigma_*}\). Let \(\lambda_j(\xi) \to \lambda(\xi)\), \(\xi \in \Pi_{\sigma_*}\). Since \(\Gamma^\ast_{\Delta} = \{(\xi, \lambda(\xi))| \xi \in \Pi_{\sigma_*}\} \subset M_j\) and \(\lambda_j\) are all analytic on \(\Pi_{\sigma_*}\), so is the limit \(\lambda\). Let \(i \to \infty\). Then we get

\[
|\lambda(\xi) - \lambda_j(\xi)| \leq \frac{\delta_j}{2}.
\]

This implies that \(\Gamma^\ast_{\Delta} = \{(\xi, \lambda(\xi))| \xi \in \Pi_{\sigma_*}\} \subset M_j\) and so \(\Gamma^\ast = \bigcup_{\Delta \in O_0} \Gamma^\ast_{\Delta} \subset M_j\).

Hence, \(\Gamma^\ast \subset M_* = \bigcap_{j \geq 0} M_j\). Evidently, for all \((\xi, \lambda) \in \Gamma^\ast_{\Delta}\) we have

\[
\lambda + \omega(\xi) + h(\xi, \lambda) = \Delta.
\]
Finally, we consider the convergence of \( \{Q_j\} \) and \( \{P_j\} \). Since 
\[
P_j = P_{0j}(\xi, \lambda; x) + P_{1j}(\xi, \lambda; x)y + P_{2j}(\xi, \lambda; x)z + P_{2j}(\xi, \lambda; x)\bar{z} + \cdots,
\]
where the dots denote the terms of \( y^lz^p\bar{z}^q \) with \( 2|l| + |p| + |q| > 2 \). On \( M_j \times D_j \), by Cauchy estimates, we have
\[
\|P_{0j}\| < c\epsilon_j \to 0, \quad \|P_{1j}\| < \frac{c\epsilon_j}{r_j^2} \to 0, \quad \|P_{2j}\| < \frac{c\epsilon_j}{r_j} \to 0, \quad \|\bar{P}_{2j}\| < \frac{c\epsilon_j}{r_j^2} \to 0.
\]
Let \( P_j \to P_* \). Then \( P_* \) contains only the terms of \( y^lz^p\bar{z}^q \) with \( 2|l| + |p| + |q| > 2 \).

Since \( \|Q_{j+1} - Q_j\| < c\epsilon_j \), for \( i > j \) we obtain
\[
\|Q_i - Q_j\| \leq \|Q_i - Q_{i-1}\| + \|Q_{i-1} - Q_{i-2}\| + \cdots + \|Q_{j+1} - Q_j\|
\]
\[
\leq \sum_{k=1}^{i-j} \epsilon_{j+k} < c\epsilon_{j+1} \to 0.
\]
Then \( Q_i \to Q_* \) and so \( Q_* \) contains only the terms of \( z^p\bar{z}^q \) with \( |p| + |q| = 2 \).

Furthermore, we have
\[
N_j \to N_* = \langle \Delta, y \rangle + \langle \Omega z, \bar{z} \rangle + Q_* (\xi, \lambda; x, z, \bar{z}).
\]

Then
\[
H \circ \Phi = \langle \Delta, y \rangle + \langle \Omega z, \bar{z} \rangle + Q_* + P_*.
\]
We denote \( Q_* + P_* \) by \( \bar{P} \). Then Theorem 2.2 holds.

4. Appendix. In this section we state several lemmas. Some of the lemmas describe properties of the norm \( \| \cdot \|_{D(s, r)} \). The most of proofs are very similar to [12, 21] and even simpler, hence we omit some of them.

Denote by \( A_{s, r} \) the space of all real analytic function defined on \( D(s, r) \). For every \( f \) belonging to \( A_{s, r} \), we can write it as Fourier series \( f = \sum_{k \in \mathbb{Z}^n} f_k e^{\sqrt{-1}(k, x)} \),
\[
where \( f_k = \sum_{l \in \mathbb{N}^n, p, q \in \mathbb{N}^n} f_k e^{\sqrt{-1}l \cdot y^p\bar{z}^q} \). Let \( A^0_{s, r} = \{ f | f \in A_{s, r}, [f] = 0 \} \), where \([f]\) denotes the average of \( f \), i.e., \([f] = \frac{1}{(2\pi)^n} \int_{\mathbb{T}^n} f \, dx\).

Lemma 4.1 (Cauchy estimates). Let \( f \in A_{s, r} \). Then, for every \( \rho \) in \((0, s)\) and \( \sigma \) in \((0, r)\), we have
\[
\|f_x\|_{D(s-\rho, r)} \leq \frac{1}{\epsilon \rho} \|f\|_{D(s, r)}, \quad \|f_y\|_{D(s, r-\sigma)} \leq \frac{1}{(2s - \sigma)\sigma} \|f\|_{D(s, r)}.
\]
\[
\|f_x\|_{D(s, r-\sigma)} \leq \frac{1}{\sigma} \|f\|_{D(s, r)} , \quad \|f_{x}\|_{D(s, r-\sigma)} \leq \frac{1}{\sigma} \|f\|_{D(s, r)}.
\]

Lemma 4.2. Let \( f \in A_{s, r} \), \( T_K(f) = \sum_{|k| \leq K} f_k e^{\sqrt{-1}(k, x)} \). Then we have
\[
\|f - T_K f\|_{D(s-\rho, r)} \leq cK^n e^{-K\rho} \|f\|_{D(s, r)}, \quad 0 \leq \rho \leq s,
\]
where the constant \( c \) only depends on \( n \).

Lemma 4.3. Let \( \Xi_{\alpha, r} = \{ \omega | \langle k, \omega \rangle \geq \frac{\alpha}{|k|}, \forall 0 \neq k \in \mathbb{Z}^n \} \). Let \( \omega \in \Xi_{\alpha, r} \) and \( g(x) \in A^0_{s, r} \). Then the equation
\[
\partial \omega f = \sum_{j=1}^{n} \omega_j f_{x_j} = \langle \omega, \partial_x f \rangle = g(x)
\]
exists unique solution $f(x) \in \bigcup_{0 < \rho < s} A^0_{s-\rho, r}$ with
\[
\|f\|_{D(s-\rho, r)} \leq \frac{c}{\alpha \rho^{r+\tau}} \|g\|_{D(s, r)},
\]
where the constant $c$ depends only on $\tau$ and $n$.

For simplifying the description of the following lemma, we denote
\[
\mathcal{V}_{s,r} = \{f|f = (f_1, f_2, ..., f_{2m})^T, f_i \in A_{s,r}, 1 \leq i \leq 2m\}
\]
and
\[
\mathcal{M}_{s,r} = \{\Theta|\Theta = (\tilde{\lambda}_{ij})_{2m \times 2m}, \tilde{\lambda}_{ij} \in A_{s,r}, 1 \leq i, j \leq 2m\}.
\]
Moreover, the norm of $f \in \mathcal{V}_{s,r}$ and $\Theta \in \mathcal{M}_{s,r}$ are defined as
\[
\|f\|_{D(s, r)} = \sum_{i=0}^{2m} \|f_i\|_{D(s, r)} \quad \text{and} \quad \|\Theta\|_{D(s, r)} = 2m \times \max_{1 \leq i,j \leq 2m} \|\tilde{\lambda}_{ij}\|_{D(s, r)},
\]
respectively.

**Lemma 4.4.** Let $\lambda_1, \ldots, \lambda_{2m}$ be the eigenvalues of constant matrix $\Theta_0$. If $|\text{Re}\lambda_i| \geq \mu > 0$, $i = 1, 2, ..., 2m$ and $g \in \mathcal{V}_{s,r}$, then there exists a sufficiently small positive constant $\epsilon_0 = \epsilon_0(\Theta_0)$ such that for arbitrary $\tilde{\Theta} \in \mathcal{M}_{s,r}$, if $\|\tilde{\Theta}\|_{D(s, r)} \leq \epsilon_0$, the equation
\[
\langle \omega, \partial_x f \rangle - (\Theta_0 + \tilde{\Theta}) f = g,
\]
where $\langle \omega, \partial_x f \rangle = (\langle \omega, \partial_x f_1 \rangle, \ldots, \langle \omega, \partial_x f_{2m} \rangle)^T$, exists unique solution $f \in \mathcal{V}_{s,r}$ with $\|f\|_{D(s, r)} \leq c\|g\|_{D(s, r)}$. Here the constant $c$ depends only on $\mu$.

**Proof.** Let $L$ be a linear operator defined on $\mathcal{V}_{s,r}$ with $Lf = \langle \omega, \partial_x f \rangle - \Theta f$. By the Fourier expansions, it is easy to know that the eigenvalues of $L$ are $\sqrt{-1}(k, \omega) - \lambda_j, 1 \leq j \leq 2m$. From the assumption, it follow that $L$ exists bounded reversible operator $L^{-1} : \mathcal{V}_{s,r} \rightarrow \mathcal{V}_{s,r}$. Moreover, there exists a sufficiently small positive constant $\epsilon_0$ such that $\|L^{-1}\| \leq \frac{1}{2\epsilon_0}$. Let $\mathcal{T}f = L^{-1}(g + \tilde{\Theta} f)$. Then, for arbitrary $f_1, f_2 \in \mathcal{V}_{s,r}$, we have
\[
\|\mathcal{T}f_1 - \mathcal{T}f_2\|_{D(s, r)} = \|L^{-1}(\tilde{\Theta} f_1 - f_2)\|_{D(s, r)} \leq \|L^{-1}\| \|\tilde{\Theta}\|_{D(s, r)} \|f_1 - f_2\|_{D(s, r)} \leq \frac{1}{2\epsilon_0} \epsilon_0 \|f_1 - f_2\|_{D(s, r)} = \frac{1}{2} \|f_1 - f_2\|_{D(s, r)}.
\]
Therefore, the operator $\mathcal{T}$ has a unique fixed point $f \in \mathcal{V}_{s,r}$ by the contraction principle. Furthermore, we have
\[
\|f\|_{D(s, r)} = \|\mathcal{T}f\|_{D(s, r)} = \|L^{-1}(g + \tilde{\Theta} f)\|_{D(s, r)} \leq \|L^{-1}\| \|g + \tilde{\Theta} f\|_{D(s, r)} \leq \|L^{-1}\| \|g\|_{D(s, r)} + \|\tilde{\Theta}\|_{D(s, r)} \|f\|_{D(s, r)} \leq \frac{1}{2\epsilon_0} (\|g\|_{D(s, r)} + \epsilon_0 \|f\|_{D(s, r)}).
\]
Then
\[
\|f\|_{D(s, r)} \leq \frac{1}{\epsilon_0} \|g\|_{D(s, r)} \leq c\|g\|_{D(s, r)}.
\]

**Acknowledgments.** The authors thank Prof. Junxiang Xu (Department of Mathematics, Southeast University, China) for his patience guidance for this paper, and also thank the reviewers for their very careful reviews and valuable suggestions.

This work is supported by the National Natural Science Foundation of China (Tian Yuan Special Foundation, 11526177), the National Natural Science Foundation of China (11302080), the Natural Science Foundation for Colleges and Universities in Anhui Province (KJ2016A602) and Jiangsu Province (15KJD110005).
REFERENCES

[1] V.I. Arnold, Proof of a theorem by A. N. Kolmogorov on the persistence of quasi-periodic motions under small perturbations of the Hamiltonian, *Uspehi Mat. Nauk*, 18 (1963), 13–40.

[2] Q.Y. Bi and J.X. Xu, Persistence of lower dimensional hyperbolic invariant tori for nearly integrable symplectic mappings, *Qual. Theory Dyn. Syst.*, 13 (2014), 269–288.

[3] H. Broer, G. Huitema and M. Sevryuk, Quasi-periodic motions in families of dynamical systems, in *Lecture Notes in Mathematics*, Springer, Berlin, 1645, 1996.

[4] A.D. Bruno, Analytic form of differential equations, *Trans. Moscow Math. Soc.*, 25 (1971), 131–288.

[5] C. Cheng and Y. Sun, Existence of KAM tori in degenerate Hamiltonian systems, *J. Differential Equations*, 114 (1994), 288–335.

[6] L.H. Eliasson, Perturbations of stable invariant tori for Hamiltonian systems, *Ann. Scuola Norm. Sup. Pisa.*, 15 (1988), 115–147.

[7] S.M. Graff, On the conservation of hyperbolic invariant tori for Hamiltonian systems, *J. Differ. Eqs.*, 15 (1974), 1–69.

[8] A.N. Kolmogorov, On conservation of conditionally periodic motions for a small change in Hamilton’s function, *Dokl. Akad. Nauk. SSSR*, 98 (1954), 527–530.

[9] S.B. Kuksin, Nearly integrable infinite dimensional Hamiltonian systems, in *Lecture Notes in Mathematics*, Springer-Verlag, Berlin, 1556 (1993).

[10] 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.

[11] J. Moser, Convergent series expansions for quasi-periodic motions, *Math. Ann.*, 169 (1967), 136–176.

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

[13] J. Pöschel, A lecture on the classical KAM theorem, *School on Dynamical Systems*, May 1992.

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

[15] H. Rüssmann, On twist Hamiltonians. Talk on the Colloque International: Mécanique céleste et systèmes hamiltoniens, *Marseille*, 1990.

[16] H. Rüssmann, Invariant tori in non-degenerate nearly integrable Hamiltonian systems, *Regular and Chaotic Dynamics*, 6 (2001), 119–204.

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

[18] X. Wang, J. Xu and D. Zhang, Persistence of lower dimensional elliptic invariant tori for a class of nearly integrable reversible systems, *Discrete Contin. Dyn. Syst., Ser. B*, 14 (2010), 1237–1249.

[19] J. Xu, Persistence of elliptic lower dimensional invariant Tori for small perturbation of degenerate integrable Hamiltonian systems, *Journal of Mathematical Analysis and Applications*, 208 (1997), 372–387.

[20] J.X. Xu, J.G. You and Q.J. Qiu, Invariant tori for nearly integrable Hamiltonian systems with degeneracy, *Mathematische Zeitschrift*, 226 (1997), 375–387.

[21] J.X. Xu and J.G. You, Gevrey-smoothness of invariant tori for analytic nearly integrable Hamiltonian systems under Rüssmann’s non-degeneracy condition, *Journal of Differential Equations*, 235 (2007), 609–622.

[22] J.X. Xu and J.G. You, Persistence of the non-twist torus in nearly integrable Hamiltonian systems, *Pro Math Amer Soc.*, 138 (2010), 2385–2395.

[23] J.G. You, A KAM theorem for hyperbolic-type degenerate lower dimensional tori in Hamiltonian systems, *Commun. Math. Phys.*, 192 (1998), 145–168.

[24] E. Zehnder, Generalized implicit function theorems with applications to some small divisor problem. I and II, *Commun. Pure Appl. Math.*, 28 (1975), 91–140. 29 (1976), 49–111.

[25] D. Zhang and J. Xu, On invariant tori of vector field under weaker non-degeneracy condition *Nonlinear Differ. Equ. Appl.*, 22 (2015), 1381–1394.

Received August 2015; revised February 2016.

E-mail address: wanglei@hfuu.edu.cn
E-mail address: 769473876@qq.com
E-mail address: woaidongda@163.com