ON THE EXISTENCE OF INVARIANT TORI IN NON-CONSERVATIVE DYNAMICAL SYSTEMS WITH DEGENERACY AND FINITE DIFFERENTIABILITY

XUEMEI LI

Key Laboratory of High Performance Computing and Stochastic Information Processing
Department of Mathematics, Hunan Normal University
Changsha, Hunan 410081, China

ZAIJIU SHANG

HLM, Academy of Mathematics and Systems Science, Chinese Academy of Sciences
School of Mathematical Sciences, University of Chinese Academy of Sciences
Beijing 100190, China

(Communicated by Chong-Qing Cheng)

Abstract. In this paper, we establish a KAM-theorem about the existence of invariant tori in non-conservative dynamical systems with finitely differentiable vector fields and multiple degeneracies under the assumption that the integrable part is finitely differentiable with respect to parameters, instead of the usual assumption of analyticity. We prove these results by constructing approximation and inverse approximation lemmas in which all functions are finitely differentiable in parameters.

1. Introduction. Consider the non-conservative dynamical systems of the form

\[
\begin{align*}
\dot{I}_1 &= \varepsilon^{q_1} A_1(\xi, \varepsilon) I_1 + \varepsilon^{q_2} g_1(I, \varphi; \xi, \varepsilon) \\
\dot{I}_2 &= \varepsilon^{q_3} A_2(\xi, \varepsilon) I_2 + \varepsilon^{q_4} g_2(I, \varphi; \xi, \varepsilon) \\
\dot{\varphi}_1 &= \varepsilon^{q_5} \omega_1(\xi, \varepsilon) + \varepsilon^{q_6} g_3(I, \varphi; \xi, \varepsilon) \\
\dot{\varphi}_2 &= \omega_2(\xi, \varepsilon) + \varepsilon^{q_7} g_4(I, \varphi; \xi, \varepsilon),
\end{align*}
\]

where \( I = \text{col}(I_1, I_2) \in \Omega \subset \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} = \mathbb{R}^{n_1}, \varphi = \text{col}(\varphi_1, \varphi_2) \in T^{n_21} \times T^{n_22} = T^{n_2}, q_j \geq 0 (j = 1, \ldots, 7), \xi \in \Pi \subset \mathbb{R}^{n_3} \) is the parameter, \( \varepsilon \) is a small perturbation parameter, all terms in (1.1) are finitely differentiable in both the coordinate variables and the parameters.

The integrable system of (1.1)

\[
\begin{align*}
\dot{I}_1 &= \varepsilon^{q_1} A_1 I_1, \\
\dot{I}_2 &= \varepsilon^{q_3} A_2 I_2, \\
\dot{\varphi}_1 &= \varepsilon^{q_5} \omega_1, \\
\dot{\varphi}_2 &= \omega_2
\end{align*}
\]

has an invariant \( n_2 \)-dimensional torus \( I = 0 \) with frequencies \( \omega = (\varepsilon^{q_5} \omega_1, \omega_2) \). The aim of the present paper is to examine the persistence of this family of rotational tori \( I = 0 \) under small perturbations (i.e., \( g_j \neq 0, j = 1, \ldots, 4 \)). System (1.1) has the following features: the perturbation terms \( g_j, j = 1, \ldots, 4 \) are only finitely differentiable, there exist small frequencies, small twist and higher-order degeneracy.
and the number of parameter variables is possibly less than the dimension of tori. To tackle these difficulties we establish a new KAM theorem for (1.1), which is an extension of Pöschel [27] for finitely differentiable Hamiltonian systems and Han, Li and Yi [18] for multiple-degenerate and analytic Hamiltonian systems to non-conservative dynamical systems, respectively.

In the context of finitely differentiable perturbations, the study on the persistence of quasi-periodic invariant tori has originated from the work of Moser [24] on area-preserving mappings of an annulus, which was extended to dissipative vector fields in [6] based on smoothing operator technique. Another important method, which can relax the requirement for regularity of perturbations, is to approximate a differentiable function by real analytic ones [25, 30, 40, 27, 35, 13, 1, 38]. Rüssmann [30] proved an optimal estimate result on approximating a differentiable function by analytic ones. Following this approach Zehnder [40] established a generalized implicit function theorem and applied it to the existence of parameterized invariant tori of nearly integrable Hamiltonian systems in finitely differentiable case, Pöschel [27] showed that on a Cantor set, invariant tori of the perturbed Hamiltonian system form a differentiable family in the sense of Whitney. The results and ideas of Moser and Pöschel are extended to the case of symplectic mappings by Shang [35] and to the case of lower dimensional elliptic tori by Chierchia and Qian [13], respectively. Wagener [38] extended the modifying terms theorem of Moser [26] (i.e., introducing additional parameters) to finitely differentiable and Gevrey regular vector fields.

The results mentioned above, except for [13], were restricted to the case where the integrable part is analytic in coordinate variables as well as in parameters. The integrable part in [13] is assumed to be Lipschitz with respect to parameters and the frequency map to be a Lipschitz homeomorphism. Of course, if the unperturbed (integrable) part and the perturbation are both of class $C^l$ with $l > 2n$ ($n$ is the number of degrees of freedom), it is reduced to the case where the integrable part is analytic and the perturbation is of class $C^l$ by regarding the initial values of action variables as parameters for Hamiltonian systems. In this paper, we shall extend the result and method of Pöschel [27] to the non-conservative system (1.1) with degeneracies, removing the restriction that the integrable part is analytic in parameters.

The perturbation was assumed to be $C^{333}$ originally in the work of Moser [24] on area-preserving mappings of an annulus, and then was weakened to $C^5$ by Rüssmann [30] and to $C^l (l > 3)$ (meaning that the perturbation is of class $C^3$ and the derivatives of order 3 are Hölder continuous) by Rüssmann [32] and Herman [19], where a counterexample for $l < 3$ was given. For improvements on weakening the regularity of perturbations in the Hamiltonian case we refer to [1] and references therein.

The above mentioned results were proved under the so-called non-degeneracy conditions. In the context of degenerate KAM theory, i.e., if Kolmogorov’s non-degeneracy or Arnold’s iso-energetic non-degeneracy condition is violated, Arnol’d [3] established a properly degenerate KAM theorem (refined by [14, 12]) to deal with quasi-periodic motions in the planetary many body problem. In this case the integrable part does not depend on the full set of action variables, and the non-degeneracy conditions are imposed additionally on the averaged perturbation. The ideas of Arnol’d [3] were extended to the resonant torus case in [10, 23] and the normal zero-frequency case in [16, 17, 15] for Hamiltonian systems and in [5, 22, 20] for non-conservative systems. Another method is to search for weaker non-degeneracy conditions concerning frequency maps, which have been studied in a
series of papers, for example, by Bruno [8], Cheng and Sun [9], Rüssmann [33, 34], Han, Li and Yi [18] for finite dimensional Hamiltonian systems, and Bambusi, Berti and Magistrelli [4] for infinite dimensional case. The weaker non-degeneracy condition in [9] is that the image of the frequency map in an open set includes a curved $C^{n+2}$ one-dimensional submanifold. Rüssmann [33, 34] pointed out that the weaker non-degeneracy condition means that the image of the frequency map does not lie in an $(n-1)$-dimensional linear subspace of $\mathbb{R}^n$ (this condition is also necessary in the analytic case). An interesting and real analytic Hamiltonian of the form

$$H(x,y,\varepsilon) = h_0(y^{n_0}) + \varepsilon^{m_1}h_1(y^{n_1}) + \cdots + \varepsilon^{m_a}h_a(y^{n_a}) + \varepsilon^{m_{a+1}}P(x,y,\varepsilon)$$

with the degeneracy involving several time scales was considered in [18]. The degeneracy in (1.1) is somewhat similar to the one in [18].

We will construct approximation and inverse approximation lemmas in which approximating and approximated functions are finitely differentiable in parameters and then prove the existence of invariant tori in (1.1) for the most of parameter values.

The KAM theorems in [7, 38] can be applied to quasi-periodic bifurcations (i.e., bifurcations of quasi-periodic invariant tori). Our results in this paper can be applied to investigating the persistence of quasi-periodic invariant tori in bifurcation theory of equilibrium points. To study the bifurcations of equilibria of a system of differential equations (ODEs, PDEs and functional differential equations), one usually reduces such a system to a lower-dimensional one on the center manifold by the Center Manifold Theorem. However the reduced system need not be analytic even if the original system is analytic. When the equilibrium is partially elliptic and the reduced subsystem has a normal form of Birkhoff type on the center manifold, then the truncated normal form may possess quasi-periodic invariant tori. In this case, a question arises naturally: do the original system (equivalently, the reduced system on the center manifold) have quasi-periodic invariant tori with the same dimension? This problem can be discussed by KAM theory and a careful study leads us to consider the existence of quasi-periodic tori of a system similar to (1.1). We will give the derivation and special forms of (1.1) corresponding to various bifurcations in a forthcoming paper.

2. Statement of results. Let $\Omega_1$ and $\Omega_2$ be convex open neighbourhoods of the origin in $\mathbb{R}^{n_1}$ and $\mathbb{R}^{n_2}$, respectively, $\Omega = \Omega_1 \times \Omega_2$, the parameter set $\Pi$ be a convex bounded open set of positive Lebesgue measure in $\mathbb{R}^{n_3}$. Let $|x|$ denote the maximum norm and $|x|_p$ the $p$-norm. In the following, $l$ and $\alpha$ represent the differentiability orders of functions in the space variables $(I,\varphi)$ and the parameter variables $\xi$, respectively.

**Definition 2.1.** Let $\alpha$ be a positive integer and $l > 0$, $C^{l,\alpha}(\Omega \times \mathbb{T}^{n_2}, \Pi)$ be the class of all functions $f$ on $\Omega \times \mathbb{T}^{n_2} \times \Pi$ whose partial derivatives $\partial_{\xi}^\beta f$ with respect to the parameter variable $\xi \in \Pi$ (which means the Whitney derivative if $\Pi$ is a closed set) for all $\beta, 0 \leq |\beta| \leq \alpha$ are of class $C^l$ in the space variable $x = (I,\varphi) \in \Omega \times \mathbb{T}^{n_2}$, that is, there is some positive constant $M$ such that the partial derivatives $D^k\left(\partial_{\xi}^\beta f\right)$ of $\partial_{\xi}^\beta f$ with respect to the space variable $x = (I,\varphi) \in \Omega \times \mathbb{T}^{n_2}$ satisfy

$$\left|D^k\left(\partial_{\xi}^\beta f(x,\xi)\right)\right| \leq M$$

(2.1)
and
\[
|D^k \left( \frac{\partial^\beta f(x, \xi)}{\partial \xi} \right) - D^k \left( \frac{\partial^\beta f(y, \xi)}{\partial \xi} \right)| \leq M|x - y|^{l - |l|}, \quad |k|_1 = |l| \tag{2.2}
\]
for all \( x, y \in \Omega \times T^{n_2}, \xi \in \Pi \) and all \( \beta, k \) with \( 0 \leq |\beta| \leq \alpha \), \( 0 \leq |k|_1 \leq |l| \), where \( |l| \) is the integer part of \( l \): \( l - |l| \in (0, 1) \), for nonnegative integer vectors \( k, \beta \),
\[
D^k = D^{k_1}_1 \circ D^{k_2}_2 \circ \cdots \circ D^{k_{n_2}}_{n_2}, \quad D_j^{k_i} = \frac{\partial^{k_i}}{\partial \xi_j^{k_i}}, \quad \partial^\beta f = \partial^{\beta_{1,1}} \cdots \partial^{\beta_{n_2}}.
\]
In addition, define a norm
\[
\|f\|_{l, \alpha; \Omega \times T^{n_2}, \Pi} = \inf M,
\]
it is the smallest \( M \) for which the inequalities (2.1) and (2.2) hold. \( C^{l, \alpha}(\Omega \times T^{n_2}, \Pi) \) is a Banach space with respect to the norm \( \| \cdot \|_{l, \alpha; \Omega \times T^{n_2}, \Pi} \), which is a generalization of the H"older space to a parameter-depending case. The norms \( \| \cdot \|_{l, \alpha; T^{n_2}, \Pi} \) and \( \| \cdot \|_{l, l, \alpha, T^{n_2}, \Pi} \) are defined in a similar way, which means that the associated function only depends on \( \varphi \in T^{n_2}, \xi \in \Pi \) and \( \xi \in \Pi \), respectively.

When \( l \) is a positive integer, we also introduce a generalization of the Zygmund space \( \tilde{C}^{l, \alpha}(\Omega \times T^{n_2}, \Pi) \) of all functions satisfying
\[
\left| D^k \left( \frac{\partial^\beta f(x, \xi)}{\partial \xi} \right) \right| \leq M, \quad 0 \leq |k|_1 \leq l - 1 \tag{2.3}
\]
and for \( |k|_1 = l - 1 \)
\[
\left| D^k \left( \frac{\partial^\beta f(x, \xi)}{\partial \xi} \right) + D^k \left( \frac{\partial^\beta f(y, \xi)}{\partial \xi} \right) - 2D^k \left( \frac{\partial^\beta f}{\partial \xi} \left( \frac{1}{2} (x + y), \xi \right) \right) \right| \leq M|x - y|, \tag{2.4}
\]
instead of (2.1) and (2.2), respectively, and the norm \( \|f\|_{\tilde{C}^{l, \alpha}(\Omega \times T^{n_2}, \Pi)} \) is the smallest \( M \) for which the inequalities (2.3) and (2.4) hold. For a non-integer \( l > 0 \), \( \tilde{C}^{l, \alpha}(\Omega \times T^{n_2}, \Pi) = C^{l, \alpha}(\Omega \times T^{n_2}, \Pi) \).

Sometimes we shall drop parameters from functions whenever there is no confusion.

a) Assume

\((H1)\) these non-negative constants \( q_1, \ldots, q_7 \) satisfy
\[
q_1 > q_3 \geq q_5, \quad q_7 \geq q_2 + q_5, \quad 0 < q_2 \leq \min\{q_4, q_5\};
\]

\((H2)\) \( \omega_i, A_i \in C^\alpha(\Pi) \) with some positive integer \( \alpha \) and \( A_i \) is a diagonalizable matrix, \( A_i(\xi, \varepsilon) = B_i(\xi, \varepsilon)A_i(\xi, \varepsilon)B_i^{-1}(\xi, \varepsilon) \) for some diagonal matrix \( A_i, i = 1, 2 \). Denote \( \omega(\xi, \varepsilon) = \text{col}(\varepsilon^{\alpha_1} \omega_1, \omega_2), A_1(\xi, \varepsilon) = \text{diag}(\lambda_1, \cdots, \lambda_{n_1}), \ A_2(\xi, \varepsilon) = \text{diag}(\lambda_{n_1+1}, \cdots, \lambda_{n_1}) \). Furthermore, assume that there are positive constants \( c_0, c_1 \) and \( \varepsilon^* \) such that for all \( \varepsilon \in (0, \varepsilon^*) \)
\[
\inf_{\xi \in \Pi} |\lambda_j(\xi, \varepsilon)| \geq c_0 \quad \text{and} \quad \inf_{\xi \in \Pi} |\lambda_j(\xi, \varepsilon) - \lambda_i(\xi, \varepsilon)| \geq c_0
\]
for \( i \neq j, 1 \leq i, j \leq n_1 \) or \( n_1 + 1 \leq i, j \leq n_1 \),
\[
\|B_i\|_{\alpha, \Pi}, \|B_i^{-1}\|_{\alpha, \Pi}, \|A_i\|_{\alpha, \Pi} \leq c_1, \quad \|\omega_i\|_{0, \Pi} = \sup_{\xi \in \Pi} |\omega_i| \leq c_1, \quad i = 1, 2,
\]
\[
\|\partial^\beta \omega\|_{\Pi} = \sup_{\xi \in \Pi} |\partial^\beta \omega(\xi, \varepsilon)| \leq c_1 \varepsilon^{\alpha_7}, \quad 1 \leq |\beta| \leq \alpha;
\]

\((H3)\) \( g_j \in C^{l, \alpha}(\Omega \times T^{n_2}, \Pi)(j = 1, \cdots, 4) \) with \( l > 2(\alpha + 1)(l + 2) + \alpha, \ i > \alpha n_2 - 1 \).

**Remark 2.1.** The requirement that \( A_i \) does not have multiple eigenvalues is not necessary, only for the sake of simplification. The difficulty caused by multiple eigenvalues may be overcome by the technique of Rüssmann [34].
Theorem 2.2. Suppose that the system (1.1) satisfies Assumptions (H1)-(H3). Then for any given $0 < \gamma \ll 1$, there is a sufficiently small $0 < \varepsilon_0 = O(\gamma^{-n_2})$ such that for $0 < \varepsilon \leq \varepsilon_0$, there exists a Cantor set $\Pi_\varepsilon \subset \Pi$ and for each $\xi \in \Pi_\varepsilon$, the system (1.1) admits a quasi-periodic invariant torus of the form $I_1 = \Phi_1(\varphi; \xi), I_2 = \Phi_2(\varphi; \xi), \varphi = \text{col}(\varphi_1, \varphi_2) \in \mathbb{T}^{n_2^1} \times \mathbb{T}^{n_2^2}$ with frequencies $\omega^{0}(\xi) = (\varepsilon^{n_2^1}_0 \omega^{0}_1(\xi), \omega^{0}_2(\xi))$, which is of class $C^0$ in $\xi \in \Pi_\varepsilon$ in the sense of Whitney and of class $C^1(l_1 = l - (\alpha + \mu + 2)(\nu + 1) - 2)$ in $\varphi \in \mathbb{T}^{n_2^2}$ together with derivatives up to order $\mu$ with respect to $\xi$ for $0 < \mu \leq \alpha$ (positive integer $\mu$), the frequency map $\omega^{0}(\xi)$ is of class $C^0$ in $\xi \in \Pi_\varepsilon$ in the sense of Whitney and satisfies

$$
\|(\Phi_1, \Phi_2)\|_{C^{0.5} - 1; \mathbb{T}^{n_2^1} \times \mathbb{T}^{n_2^2}, \Pi_\varepsilon} \leq C\varepsilon^{\varepsilon_0} \gamma^{-1} - (\mu + 1),
$$

and

$$
\|\omega^*_1 - \omega_1\|_{C^0_{\Pi_\varepsilon}(\Pi_\varepsilon)} \leq C\varepsilon^{\varepsilon_0}, \quad \|\omega^*_2 - \omega_2\|_{C^0_{\Pi_\varepsilon}(\Pi_\varepsilon)} \leq C\varepsilon^{\varepsilon_0}. \tag{2.6}
$$

Moreover, there exist closed subsets $\Pi_\nu$ of $\Pi$, frequency vectors $\omega^{\nu}(\xi) = \text{col}(\varepsilon^{\nu}_0 \omega^{\nu}_1(\xi), \varepsilon^{\nu}_0 \omega^{\nu}_2(\xi))$ and diagonal matrices $\Lambda^{\nu}(\xi) = \text{diag}(\varepsilon^{\nu}_0 \Lambda^{\nu}_1(\xi), \varepsilon^{\nu}_0 \Lambda^{\nu}_2(\xi))$ for $\nu = 1, 2, \ldots$, satisfying

$$
\text{with frequencies}
$$

$$
\|\omega^{\nu}_1 - \omega_1\|_{\alpha, \Pi_\nu} \leq C\varepsilon^{\varepsilon_0}, \|\omega^{\nu}_2 - \omega_2\|_{\alpha, \Pi_\nu} \leq C\varepsilon^{\varepsilon_0}, \|\Lambda^{\nu}_1 - \Lambda_1\|_{\alpha, \Pi_\nu} \leq C\varepsilon^{\varepsilon_0}, \|\Lambda^{\nu}_2 - \Lambda_2\|_{\alpha, \Pi_\nu} \leq C\varepsilon^{\varepsilon_0}
$$

and

$$
\Pi_\nu = \Pi_{\nu - 1} \setminus \bigcup_{k,m} \mathcal{R}^{\nu}_{km}(\gamma)
$$

such that $\Pi_\nu = \bigcap_{\nu = 1}^\infty \Pi_\nu$, where

$$
\mathcal{R}^{\nu}_{km}(\gamma) = \{\xi \in \Pi_{\nu - 1} : \sqrt{-1}(k, \omega^{\nu - 1}) + (m, \Lambda^{\nu - 1}) < \gamma^{1/2} \}
$$

for $m \in \mathbb{Z}^{n_2}, k_{\nu - 1} < |k|_2 \leq K_\nu, m = \{m \in \mathbb{Z}^{n_1} : |m|_1 \leq 2, \sum_{j=1}^{n_1} m_j = 0 \text{ or } 1\}, \omega^0 = \omega, \Lambda^0 = \Lambda \equiv \text{diag}(\varepsilon^{n_2^1} \Lambda_1, \varepsilon^{n_2^2} \Lambda_2), K_0 = 0$ and $K_\nu = [K_0^\nu] + 1, K_\nu^r = r^{-1}3^{n_2^1 - 1}(l + (n_2^1 + 1)(\nu + 1) - 1)\ln 3 + (n_2 + 1)\ln r + \ln C, \Pi_0$ is a closed subset of $\Pi$ whose distance to the boundary of $\Pi$ is at least equal to $\gamma$. Here $C$ and $c$ are constants independent of $\nu, \varepsilon$ and $\gamma$, $r$ is the radius of the neighbourhood $\Omega$, $\| \cdot \|_{C^0(\Pi_\varepsilon)}$ is the Whitney norm (see Appendix A.1), $C = 24(n_2^1)!n_2^{n_2^2}e^{-n_2^2}, [K_\nu^r]$ represents the integer part of $K_\nu^r$.

Here, we drop $\varepsilon$ from functions, the continuous differentiability of functions $\omega^{\nu}_i$ and $\Lambda^0_i (i = 1, 2)$ on the closed set $\Pi_\nu$ means that they are continuously differentiable in some neighbourhood of $\Pi_\nu$. Here and in the sequel, we also regard the $\Lambda$ as a column vector of its diagonal elements when $\Lambda$ is a diagonal matrix.

b) The Cantor set $\Pi_\nu$ is not empty and indeed the measure measure $\Pi \setminus \Pi_\nu \to 0$ as $\gamma \to 0$ as long as we impose proper non-degeneracy conditions on frequencies. Since in applications the non-degeneracy conditions on frequencies are different, Theorem 2.2 does not involve the measure estimate of $\Pi_\nu$, so that it can be used more widely. In the following theorems, we give some conditions to ensure that the Cantor set $\Pi_\nu$ is not empty.

By the assumption (H2), we can write $\omega_2$ as

$$
\omega_2(\xi, \varepsilon) = \omega_2 + \varepsilon^{n_2} \omega_2(\xi) + o(\varepsilon^{n_2}),
$$

where $\omega_2$ is independent of $\xi, o(\varepsilon^{n_2})$ represents infinitely small quantity of $\varepsilon^{n_2}$ up to $\alpha$-th derivatives. Denote $\Lambda_2(\xi) = \Lambda_2(\xi, \varepsilon)|_{\varepsilon = 0}$ and $\widetilde{\omega}(\xi) = \text{col}(\omega_1(\xi, 0), \omega_2(\xi))$ in the case $n_2^2 \neq 0$, $\widetilde{\omega}(\xi) = \omega_1(\xi, 0)$ in the case $n_2^2 = 0$. 
Theorem 2.3. Suppose that the system (1.1) satisfies the assumptions in Theorem 2.2, moreover assume that $n_3 + \cdots + n_3^2 \geq n_2$ and

(i) for all $\xi \in \Pi$

\[
\text{rank} \left( \frac{\partial |^{\beta}|_1}{\partial \xi^\beta} : 1 \leq |\beta|_1 \leq \alpha \right) = n_2 \quad \text{in Case } n_{22} = 0, \quad (2.8)
\]

\[
\text{rank} \left( \frac{\partial |^{\beta}|_1}{\partial \xi^\beta} : 1 \leq |\beta|_1 \leq \alpha \right) = n_2 \quad \text{in Case } n_{22} \neq 0, \quad (2.9)
\]

(ii) for all integer vectors $0 \neq k \in \mathbb{Z}^{n_2}, m = (m_1, \cdots, m_{n_{12}}) \in \mathbb{Z}^{n_{12}}$ with $1 \leq |m|_1 \leq 2$ and $m_1 + \cdots + m_{n_{12}} = 0$ or $-1$

\[
\text{meas} \left\{ \xi \in \Pi : \sqrt{-1}(k, \bar{\omega}_0 + \varepsilon q_3 \bar{\omega}(\xi)) + \varepsilon q_3 \langle m, \Im \bar{\Lambda}_2(\xi) \rangle = 0 \right\} = 0, \quad (2.10)
\]

where $\bar{\omega}_0 = \text{col}(0, \omega_2)$.

Then the Cantor set $\Pi_\gamma$ defined in Theorem 2.2 is of positive Lebesgue measure and $\text{meas}(\Pi_\gamma \cap \Pi_\gamma^0) \to 0$ as $\gamma \to 0$.

The conditions (2.8), (2.9), Lemma 6.5 and Remark A.2 imply that there is a constant $c_2 > 0$ such that

\[
\max_{0 \leq \mu \leq \alpha} \| D^\mu \langle b, \bar{\omega}(\xi) \rangle \| \geq c_2 \quad \text{in Case } n_{22} = 0, \quad (2.11)
\]

and

\[
\max_{1 \leq \mu \leq \alpha} \| D^\mu \langle b, \bar{\omega}(\xi) \rangle \| \geq c_2 \quad \text{in Case } n_{22} \neq 0, \quad (2.12)
\]

for all $\xi \in \Pi, b \in S_{n_{21}} \equiv \{ b \in \mathbb{R}^{n_2} : |b|_2 = 1 \}$.

Let

\[
K^* = \frac{32c_1}{c_2} n_3^2, \quad f_{k,m}(\xi) = \langle k, \bar{\omega}_0 + \varepsilon q_3 \bar{\omega}(\xi) \rangle + \varepsilon q_3 \langle m, \Im \bar{\Lambda}_2(\xi) \rangle,
\]

where $\Im \bar{\Lambda}_2$ is the imaginary part of $\bar{\Lambda}_2$.

Theorem 2.4. If in Theorem 2.3, the condition (ii) is replaced by

(ii)' there is a constant $c_3 > 0$ such that

\[
\max_{1 \leq \mu \leq \alpha} \| D^\mu f_{k,m}(\xi) \| \geq c_3 \varepsilon q_3
\]

for all $\xi \in \Pi, 0 < |k|_2 < K^*, 1 \leq |m|_1 \leq 2, m_1 + \cdots + m_{n_{12}} = 0$ or $-1$, then

\[
\text{meas} \Pi_\gamma = \text{meas} \Pi_\gamma - O(\gamma^{\frac{1}{3}})
\]

for sufficiently small $\gamma$.

Remark 2.2. If $q_3 > q_5 \geq 0$, or $q_3 = q_5$ and there a constant $c_2' > 0$ such that

\[
\inf_{\xi \in \Pi} |\langle m, \text{Re} \bar{\Lambda}_2(\xi) \rangle| \geq c_2' \quad \text{for } 1 \leq |m|_1 \leq 2, m_1 + \cdots + m_{n_{12}} = 0 \text{ or } -1,
\]

then the conditions (ii) and (ii)' in Theorems 2.3 and 2.4, respectively, may be removed, see the proof of Theorems 2.3 and 2.4, and Remark 5.1 in Section 5.

c) we consider a specific form of (1.1) for the case $n_{11} = n_{22} = 0, n_2 = n_3, q_3 = q_5 = 0$ and $q_4 = q_5 = 1$, which means that the first and fourth equations in (1.1) are absent and the number of parameter variables equals the dimension of tori, the equation (1.1) reads

\[
\begin{align*}
\dot{I} &= A(\xi) I + \varepsilon g_1(I, \varphi; \xi, \varepsilon) \\
\dot{\varphi} &= \omega(\xi) + \varepsilon g_2(I, \varphi; \xi, \varepsilon). 
\end{align*}
\]

(2.13)
Denote $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_{n_2})$ and $\omega = \text{col}(\omega_1, \ldots, \omega_{n_2})$, where $\lambda_1, \ldots, \lambda_{n_2}$ are the eigenvalues of $A$, $A(\xi) = B(\xi)\Lambda(\xi)B(\xi)^{-1}$. Assume $\omega, A \in C^1(\Pi)$, the map $\xi \to \omega(\xi)$ is a diffeomorphism between $\Pi$ and its image, and there exist positive constants $c_0, c_1$ and $c_4$ such that $||B||_{1;\Pi}, ||B^{-1}||_{1;\Pi}, ||A||_{1;\Pi}, ||\omega||_{1;\Pi} \leq c_1$.

Corollary 1. Suppose that the system (2.13) satisfies Assumptions (H2) and (H3). Then for any given $\varepsilon > 0$, there is a sufficiently small $\varepsilon^* > 0$ such that for $0 < \varepsilon \leq \varepsilon^*$, there exists a Cantor set $\Pi_\gamma \subset \Pi$ with positive Lebesgue measure (and the measure satisfies the estimate $\text{meas} \Pi_\gamma = \text{meas} \Pi - c\gamma$ if (2.16) replaces (2.15)) and for each $\xi \in \Pi_\gamma$, the system (2.13) possesses a quasi-periodic invariant torus $I = \Phi(\varphi; \xi), \varphi \in T^{n_2}$ consisting of quasi-periodic motions, which is of $C^{l_1}(l_1 = l - 4(n + 1) - 3)$ in $\varphi \in T^{n_2}$ and Lipschitz in $\xi$ $\in \Pi_\gamma$, where $c$ is a constant independent of $\gamma$ and $\varepsilon$.

Usually the normal form (integrable part) of (1.1) related bifurcation problems of actual models is only finitely differentiable, not analytic in the parameter $\xi$, and the frequency map is possibly degenerate so that we need the higher-order derivatives of the frequency map to estimate the Lebesgue measure of $\Pi_\gamma$ and obtain $\Pi_\gamma$ is the most part of $\Pi$. Hence, we need to establish an approximation lemma and the corresponding inverse approximation lemma in which a finitely differentiable function is approximated by a sequence of functions being analytic in space variables, but finitely differentiable in parameter variables. These comprise Section 3. The proofs of Theorems 2.2-2.4 are given in Sections 4 and 5, respectively.

3. Approximation lemmas. Zehnder [40] established the approximation and inverse approximation Lemmas on a finitely differentiable real function approximated by a sequence of real analytic functions, which was generalized to the anisotropic case by Pöschel [27], and was sharpened to covering the finitely differentiable and Gevrey regular cases by Wagener [38], respectively. Here, we give generalized versions of Zehnder’s approximation and inverse approximation lemmas finitely-smoothly depending on parameters, and obtain estimates of higher-order regularity.
In particular, for \( g \in \mathcal{C}^\infty (\mathbb{R}, \Pi) \) be the class of all functions of \((z, \xi)\) on \( \mathcal{U} \times \Pi \) which are analytic in \( z \in \mathcal{U} \) and \( \alpha \)-times continuously differentiable in \( \xi \in \Pi \). For \( g \in \mathcal{A}^\alpha (\mathcal{U}, \Pi) \), define

\[
|g|_{\mathcal{U}, \alpha; \Pi} = \sup_{|\beta|_1 \leq \alpha} \sup_{(z, \xi) \in \mathcal{U} \times \Pi} \left| \partial^{\beta}_{\xi} g(z, \xi) \right|.
\]

In particular, for \( \mathcal{U} = \{ z \in \mathbb{C}^m : \text{Im} z \equiv \sup_{1 \leq j \leq m} \text{Im} z_j < r \} \), we denote \(|g|_{\mathcal{U}, \alpha; \Pi} \) by \(|g|_{r, \alpha; \Pi} \).

Take an even function \( u_0 \in C^\infty_0 (\mathbb{R}) \), vanishing outside the interval \([-1, 1]\) and identically equal to 1 in a neighbourhood of 0 (see [38] for the construction of such a function). For \( x \in \mathbb{R}^m \), let \( u(x) = u_0(|x|^2) \) and \( \tilde{u} \) be the inverse Fourier transform of \( u \)

\[
\tilde{u}(z) = (2\pi)^{-m} \int_{\mathbb{R}^m} u(x) e^{-i(z, x)} dx.
\]

Let \( f \) be a real-valued function of class \( \mathcal{C}^{l, \alpha} (\mathbb{R}^m, \Pi) \) (see Definition 2.1), \( f_r (0 < r \leq 1) \) be defined by the convolution

\[
f_r (x, \xi) \equiv (S_r (f(\cdot, \xi)))(x) = r^{-m} \int_{\mathbb{R}^m} \tilde{u}(r^{-1}(x-y)) f(y, \xi) dy \quad (3.1)
\]

for \( x \in \mathbb{C}^m \). We list some properties of the analytic smoothing operator \( S_r \) in Section A.3 of the appendix, which will be used in the proof of the next lemma.

**Lemma 3.1.** Let \( f(x, \xi) \) be a real-valued function of class \( \mathcal{C}^{l, \alpha} (\mathbb{R}^m, \Pi) \) for some real number \( l > 0 \) and \( \alpha \in \mathbb{N} \), where \( \Pi \subset \mathbb{R}^n \) is an open set. Then for every \( r \in (0, 1] \), the function \( f_r (x, \xi) \) is \( \alpha \)-times continuously differentiable in \( \xi \in \Pi \), entire real analytic in \( x \in \mathbb{C}^m \) together with derivatives up to order \( \alpha \) with respect to \( \xi \), and satisfies

(i) \( \|f_r - f\|_{\mathcal{C}^{l, \alpha} (\mathbb{R}^m, \Pi)} \leq C_1 (l, p) r^{l-p} \|f\|_{\mathcal{C}^{l, \alpha} (\mathbb{R}^m, \Pi)} \) for all \( 0 \leq p < l \),

(ii) \( \|f_r - f_r\|_{\mathcal{C}^{l, \alpha} (\mathbb{R}^m, \Pi)} \leq C_2 (l, p) r^{l-p} \|f\|_{\mathcal{C}^{l, \alpha} (\mathbb{R}^m, \Pi)} \) for all \( 0 \leq p \leq l \) and \( 0 < r < r' \leq 1 \),

(iii) \( \|f_r\|_{\mathcal{C}^{l} (\mathbb{R}^m, \Pi)} \leq C_3 (l) \|f\|_{\mathcal{C}^{l, \alpha} (\mathbb{R}^m, \Pi)} \) where \( C_j (j = 1, 2, 3) \) are constants depending on \( l, p \) and the dimension \( m \). Moreover, \( f_r \) is \( \omega \)-periodic in some variable if in which \( f \) is \( \omega \)-periodic.

**Proof.** From (3.1) it is clear that \( f_r (x, \xi) \) is analytic in \( x \in \mathbb{C}^m \), and \( \alpha \)-times continuously differentiable in \( \xi \in \Pi \), taking real values on real variables \( x \). As differentiation may commute with integration in (3.1) for functions with bounded derivatives, we obtain \( \partial^{\beta}_{\xi} f_r = S_r (\partial^{\beta}_{\xi} f) \) for \( |\beta|_1 \leq \alpha \). Of course, we also have \( S_r (D^k f) = D^k (S_r (f)) \) for \( |k|_1 \leq l, k \in \mathbb{Z}^m \). Hence we only need to prove the estimates (i)-(iii) in the case without parameter-dependence. In the following, we will use \( C \) to denote some constant depending \( l, p \) and \( m \).

(i) The case where \( p \) is a integer, is proved by Chierchia [11], see Lemma 6.6 (f) in Appendix. Hence we only give the proof for the case \( p = q + \mu \leq l, \mu \in (0, 1), q \in \mathbb{Z}_+ \).

Denote \( g(x) = D^\beta f, |\beta|_1 = q \). Then by (a) and (b) in Lemma 6.6, we have for \( x, y \in \mathbb{R}^m \),

\[
\sup_{x \not= y} |x-y|^{-\mu} |(g - S_r g)(x) - (g - S_r g)(y)|
\]

\[
= \sup_{x \not= y} |x-y|^{-\mu} \left| \int_{\mathbb{R}^m} \tilde{u}(z) [g(x) - g(x-rz) - g(y) + g(y-rz)] dz \right| \equiv (*)
\]
Case I. $q = [l]$, the integer part of $l$. For $|x - y| \geq r$, by $g \in C^{l-q}$ and Lemma 6.6 (d), we have

\[
(*) \leq \sup_{x \neq y} |x - y|^{-\mu} \int_{\mathbb{R}^m} |\tilde{u}(z)||g(x) - g(x - rz)| + |g(y) - g(y - rz)| dz \\
\leq 2r^{l-p}||f||_{l;\mathbb{R}^m} \int_{\mathbb{R}^m} |\tilde{u}(z)||z|^{l-q}dz \leq Cr^{l-p}||f||_{l;\mathbb{R}^m}.
\]

For $|x - y| < r$, we also have

\[
(*) \leq \sup_{x \neq y} |x - y|^{-\mu} \left( ||f||_{l;\mathbb{R}^m} |x - y|^{l-q} + \int_{\mathbb{R}^m} |\tilde{u}(z)||g(x - rz) - g(y - rz)| dz \right) \\
\leq \left( 1 + \int_{\mathbb{R}^m} |\tilde{u}(z)|dz \right) |x - y|^{l-p}||f||_{l;\mathbb{R}^m} \leq Cr^{l-p}||f||_{l;\mathbb{R}^m}.
\]

Hence, $||g - S_r g||_{l;\mathbb{R}^m} \leq Cr^{l-p}||f||_{l;\mathbb{R}^m}$, which, combining with Lemma 6.6 (f) for the case of integers, implies (i) for the case $q = [l]$.

Case II. $q < [l]$. For $|x - y| \geq r$, using the Taylor’s formula of $h(rz) = g(x - rz) - g(y - rz)$ at $z = 0$ and Lemma 6.6 (c), we obtain

\[
(*) = \sup_{x \neq y} |x - y|^{-\mu} \left( \int_{\mathbb{R}^m} \tilde{u}(z)(-rz \cdot \nabla) |\tilde{u}(z)|(q(x - \theta rz) - g(y - \theta rz))dz \right) \\
\leq \sup_{x \neq y} |x - y|^{-\mu} |(l - q)|^{q-1} \sum_{|k| = |l| - q} \frac{1}{k!} \left| \int_{\mathbb{R}^m} \tilde{u}(z)z^k(D^k g(x - \theta rz) - D^k g(y - \theta rz))dz \right|.
\]

(3.2)

where $rz \cdot \nabla = \sum_{j=1}^m rz_j D_j$ and $\theta \in (0, 1)$. Thus, equivalently, we need to estimate the following expression

\[
(**) \equiv \sup_{x \neq y} |x - y|^{-\mu} \left( \int_{\mathbb{R}^m} \tilde{u}(z)z^k(D^k f(x - \theta rz) - D^k f(y - \theta rz))dz \right), |k + \beta|_1 = [l].
\]

By Lemma 6.6 (c), we get

\[
(**) = \sup_{x \neq y} |x - y|^{-\mu} \left( \int_{\mathbb{R}^m} \tilde{u}(z)z^k(D^k f(x - \theta rz) - D^k f(y - \theta rz))dz \right) \\
+ \int_{\mathbb{R}^m} \tilde{u}(z)z^k(D^k f(y) - D^k f(y - \theta rz))dz \leq Cr^{l-[l]-\mu}||f||_{l;\mathbb{R}^m}.
\]

(3.3)

For $|x - y| < r$, if $[l] - q \geq 2$, then similarly we have

\[
(*) \leq \sup_{x \neq y} |x - y|^{-\mu} |(l - q)|^{q-1} \sum_{|k| = [l] - q - 1} \frac{1}{k!} \left| \int_{\mathbb{R}^m} \tilde{u}(z)z^k(D^k g(x - \theta rz) - D^k g(y - \theta rz))dz \right|.
\]

(3.4)

The mean value theorem and Lemma 6.6 (c) deduce

\[
\sup_{x \neq y} |x - y|^{-\mu} \left( \int_{\mathbb{R}^m} \tilde{u}(z)z^k(D^k g(x - \theta rz) - D^k g(y - \theta rz))dz \right) \\
\leq \sup_{x \neq y} |x - y|^{-\mu} \sum_{|k'|_1 = 1} \left( \int_{\mathbb{R}^m} \tilde{u}(z)z^{k'}(x - y)^{k'}D^{k+k'} g(y - \theta rz + \theta_{kk'}(x - y))dz \right)
\]

where $\theta_{kk'}(x - y) = \frac{x - y}{|x - y|}$ for $|x - y| > 0$. If $|x - y| = 0$, then we have $\lim_{x \rightarrow y} \theta_{kk'}(x - y) = 1$.
where \(\theta_{kk'} \in (0, 1)\). If \([l] - q = 1\), then by the mean value theorem,

\[
(*) \quad \leq \sup_{x \neq y} |x - y|^{-\mu} \sum_{|k'| = 1} |x - y| |f|_{l,R^m} \int_{R^m} |\tilde{u}(z)||r z + (\theta_{1k'} - \theta_{2k'})(x - y)|^{\ell - [l]}dz
\]

\[
\leq C_{r^{l+1-[l]-\mu}} |f|_{l,R^m},
\]

(3.6)

where \(\theta_{1k'}, \theta_{2k'} \in (0, 1)\). Hence, (3.2)-(3.6) and Lemma 6.6 (f) imply (i) for the case \(q < [l]\).

Obviously, Lemma 6.6 (f) implies (ii), and the definition of \(f_r\) and Lemma 6.6 (b) and (d) imply (iii). \(\square\)

From Lemma 3.1, it follows the approximation lemma.

**Lemma 3.2. (Approximation Lemma)** Let \(f(x, \xi)\) be a real-valued function of class \(C^{l,\alpha}(\mathbb{R}^m, \Pi)\) for some real number \(l > 0\) and \(\alpha \in \mathbb{N}\), where \(\Pi\) is an open set, and let \(\{r_j\}_{j=0}^{\infty}\) be a monotonically decreasing sequence of positive numbers with \(r_0 \leq 1\) and tend to zero. Then there exists a sequence of functions \(\{f_j(z, \xi)\}_{j=0}^{\infty}\) tending to \(f(x, \xi)\) as \(j \to \infty\) for almost every \(x, \xi \in \Pi\), with derivatives up to order \(\alpha\) with respect to \(\xi\), starting with \(f_0 \equiv 0\), such that

\[
\lim_{j \to \infty} ||f_j - f||_{p,\alpha;\mathbb{R}^m, \Pi} = 0 \quad \text{for all } 0 \leq p < l
\]

and

\[
||f_j - f_{j-1}||_{r_j,\alpha;\Pi} \leq C_0 r_j^{\frac{l}{2}} ||f||_{l,\alpha;\mathbb{R}^m, \Pi} \quad \text{for } j \geq 1,
\]

where the constant \(C_0\) depends on \(l\) and the dimension \(m\). Moreover, the \(f_j\) is \(\omega\)-periodic in each variable in which \(f\) is \(\omega\)-periodic.

b) Now, we want to apply the approximation lemma to the proof of Theorem 2.2 and obtain sequences of real analytic functions approximating \(g_i(i = 1, \cdots, 4)\) in the equation (1.1).

Without loss of generality, we take

\[
\Omega = \{I = \text{col}(I_1, I_2) \in \mathbb{R}^{n_1} : |I| < 3\hat{r}\}
\]

for some constant \(0 < \hat{r} \leq 1\).

Let

\[
\Omega^* = \{I \in \mathbb{R}^{n_1} : |I| \leq 2\hat{r}\}, \quad r_j = \hat{r}3^{-j}, \quad j = 0, 1, 2, \cdots
\]

Define complex neighbourhoods \(U_j\) of \(\Omega^* \times \mathbb{T}^{n_2}\) for \(j = 0, 1, 2, \cdots\) by

\[
U_j = \{I, \varphi \in \mathbb{C}^{n_1} \times \mathbb{C}^{n_2} : \text{dist}(I, \Omega^*) < 3r_j, |\text{Im}\varphi| < 3r_j\} \equiv \Omega^* \times \mathbb{T}^{n_2} + (3r_j, 3r_j).
\]

We first expand the definition domain \(\Omega \times \mathbb{T}^{n_2} \times \Pi\) of \(g_i(i = 1, \cdots, 4)\) to \(\mathbb{R}^{n_1} \times \mathbb{T}^{n_2} \times \Pi\) in the following manner: we multiply \(g_i\) by a \(C^\infty\)-function on \(\mathbb{R}^{n_1}\) which identical 1
on $\Omega^*$ and vanishes outside $\Omega$. The obtained function belongs to $C^{l,\alpha}(\mathbb{R}^{n_1} \times \mathbb{T}^{n_2}, \Pi)$ and is equal to $g_i$ on $\Omega^* \times \mathbb{T}^{n_2} \times \Pi$, its norm is bounded by $c_i\|g_i\|_{C^{l,\alpha} \times \mathbb{T}^{n_2}, \Pi}$, where $c_i$ is a constant depending $l, n_1$ and the chosen $C^\infty$-function. Then by the approximation lemma (Lemma 3.2) we have the following corollary.

**Corollary 2.** If the system (1.1) satisfies Assumption (H3), then there exist sequences \(\{g_i^j(1, \varphi, \xi)\}_{j=0}^\infty\) of real analytic functions, being of class $C^{\alpha}$ in $\xi \in \Pi$, and entire, real analytic in $(I, \varphi) \in \mathcal{U}_0$, periodic in the variables $\varphi$ with periodic $2\pi$ together with derivatives up to order $\alpha$ with respect to $\xi$, starting with $g_i^0 \equiv 0$, such that

$$
\lim_{j \to \infty} \|g_i^j - g_i\|_{C^{l,\alpha} \times \mathbb{T}^{n_2}, \Pi} = 0 \quad \text{for all } 0 \leq p < l,
$$

$$
|g_i^j - g_i^{j-1}|_{C^{l-1,\alpha} \times \Pi} \leq C_0 r_j^p \quad \text{for } j \geq 1, i = 1, \ldots, 4,
$$

where $C_0$ is a constant depending only on $l, n_1, n_2, \tilde{r}$ and $c_i$.

c) Let $\Omega \subset \mathbb{R}^m$ be an open convex set, and $\Pi_0 \subset \mathbb{R}^n$ be a closed set,

$$
\mathcal{W}_j = \Omega + r_j, \quad \Pi_j = \bigcup_{\xi \in \Pi_0} \{ \xi \in \mathbb{R}^n : |\xi - \xi| < s_j \}, \quad j = 0, 1, 2, \ldots,
$$

where $r_j = r_0 \theta^j, 0 < \theta < 1$ and \(\{s_j\}_{j=0}^\infty\) is a monotonically decreasing sequence of positive numbers with $s_0 \leq 1$ and tend to zero.

**Lemma 3.3.** (Inverse Approximation Lemma) Let $\{f_j(x, \xi)\}_{j=0}^\infty$ be a sequence of functions such that $f_0 \equiv 0$, $f_j(x, \xi)$ is of class $C^{\alpha}$ in $\xi \in \Pi_j$, real analytic in $x \in \mathcal{W}_j$ together with derivatives up to order $\alpha$ with respect to $\xi$, and

$$
|f_j - f_{j-1}|_{C^{l,\alpha} \times \Pi_j} \leq M r_j^p \quad (3.7)
$$

for every $j \geq 1$ and some constant $M$. If there exists a constant $c'_0 > 0$ such that $r_j^p \leq c'_0 s_j^{\delta}$, then there is a unique function $f(x, \xi)$ being of class $C^{\alpha}$ in $\xi \in \Pi_0$ in the sense of Whitney (see Appendix A.1), and of class $C^{l}$ in $x \in \Omega$ together with derivatives up to order $\alpha - 1$ with respect to $\xi$ such that

$$
\|f\|_{C^{l,\alpha-1} \times \Pi} \leq C'_0 M \quad \text{and} \quad \lim_{j \to \infty} \|f - f_j\|_{C^{l,\alpha-1} \times \Pi} = 0 \quad \text{for all } 0 \leq p < l.
$$

Moreover, let $l = q + \mu, q \in \mathbb{Z}_+, \mu > 0$ and if $r_j^p \leq c'_1 s_j^{\delta}$ for some constant $c'_1$ and $0 < \delta \leq 1$, then we may require the $(\alpha - 1)$-order derivatives $\partial^\beta f(x, \xi)$ with $|\beta|_1 = \alpha - 1$ to be uniformly $\delta$-Hölder continuous in $\xi \in \Pi_0$ in the space $C^\mu(\Omega)$, that is,

$$
\|\partial^\beta f(\cdot, \xi) - \partial^\beta f(\cdot, \zeta)\|_{C^\mu(\Omega)} \leq C'_1 M|\xi - \zeta|^\delta \quad \text{for } \xi, \zeta \in \Pi_0, |\beta|_1 = \alpha - 1, \quad (3.8)
$$

where the constant $C'_0$ and $C'_1$ depend on $l, m, n, \theta, c'_0$ and $c'_1$, $C^\mu(\Omega)$ is the Zygmund space.

**Proof.** By a similar proof to that of Lemma 2.2 (ii) in [40] (also see the proof of Lemma 4.3 in [21], Theorem A.3 in [38]), we can obtain that there exist functions $f^{(\beta)} \in \hat{C}^l(\Omega), |\beta|_1 \leq \alpha$ such that

$$
\sup_{\xi \in \Pi} \|f^{(\beta)}(\cdot, \xi)\|_{C^\mu(\Pi)} \leq C'_0 M \quad \text{and} \quad \lim_{j \to \infty} \|\partial^\beta f_j(\cdot, \xi) - f^{(\beta)}(\cdot, \xi)\|_{C^\mu(\Omega)} = 0
$$
uniformly on $\Pi_0$ for all $0 \leq p < l$ and $|\beta| \leq \alpha$. Set $f(x, \xi) = f^{(\beta)}(x, \xi)$ with $\beta = 0$. To prove the rest of the lemma we only need to verify (3.8) and the compatibility conditions in the definition of Whitney derivatives (see Appendix A.1)

$$f^{(\beta)}(x, \xi) = \sum_{|\beta+k|_1 \leq \alpha-1} \frac{1}{k!} f^{(\beta+k)}(x, \xi) (\xi - \zeta)^k + R^\beta(x, \xi, \zeta) \tag{3.9}$$

with

$$\sup_{x \in \Omega} |R^\beta(x, \xi, \zeta)| \leq CM|\xi - \zeta|^{\alpha-|\beta|_1} \tag{3.10}$$

for all $\xi, \zeta \in \Pi_0$, $|\beta| \leq \alpha - 1$ and some finite constant $C$.

Set

$$h_j(x, \xi) = f_j(x, \xi) - f_{j-1}(x, \xi)$$

and

$$R^\beta_j(x, \xi, \zeta) = \partial^\beta_{\xi} h_j(x, \xi) - \sum_{|\beta+k|_1 \leq \alpha-1} \frac{1}{k!} \partial^\beta_{\xi+k} h_j(x, \zeta) (\xi - \zeta)^k$$

for $j \geq 1$, $|\beta|_1 \leq \alpha - 1$. Then

$$f^{(\beta)}(x, \xi) = \sum_{j=0}^{\infty} \partial^\beta_{\xi} h_j(x, \xi), \quad R^\beta(x, \xi, \zeta) = \sum_{j=0}^{\infty} R^\beta_j(x, \xi, \zeta), \quad |\beta|_1 \leq \alpha - 1. \tag{3.11}$$

If $s_{j_0+1} \leq |\xi - \zeta| < s_{j_0}$ for some positive integer $j_0$, then the line segment $L$ connecting $\xi$ to $\zeta$ is contained in $\Pi_j$ with $1 \leq j \leq j_0$, and the Taylor expansion implies

$$\sup_{x \in \Omega} |R^\beta_j(x, \xi, \zeta)| \leq C_1(\beta) M r^l_j |\xi - \zeta|^{\alpha-|\beta|_1}, \quad 1 \leq j \leq j_0.$$ 

And

$$\sup_{x \in \Omega} |R^\beta_j(x, \xi, \zeta)| \leq C_2(\beta) M r^l_j s_{j_0+1}^{-\alpha-|\beta|_1} |\xi - \zeta|^{\alpha-|\beta|_1}, \quad j \geq j_0 + 1.$$ 

Hence,

$$\sup_{x \in \Omega} |R^\beta(x, \xi, \zeta)| \leq M |\xi - \zeta|^{\alpha-|\beta|_1} \left( C_1 \sum_{j=1}^{j_0} r^l_j + C_2 \sum_{j=j_0+1}^{\infty} \left( \frac{r^l_j}{r^l_{j_0+1}} \right)^t \left( \frac{s_{j_0+1}^{-\alpha-|\beta|_1}}{s_1^{-\alpha-|\beta|_1}} \right) \right)$$

$$\leq CM |\xi - \zeta|^{\alpha-|\beta|_1}$$

If $|\xi - \zeta| \geq s_1$, then we also have

$$\sup_{x \in \Omega} |R^\beta(x, \xi, \zeta)| \leq C_2(\beta) M |\xi - \zeta|^{\alpha-|\beta|_1} \left( \frac{r^l_1}{s_1^{-\alpha-|\beta|_1}} \right) \sum_{j=1}^{\infty} \left( \frac{r_j}{r_1} \right)^t$$

$$\leq CM |\xi - \zeta|^{\alpha-|\beta|_1}.$$ 

Thus, we prove the compatibility conditions (3.9) and (3.10), and obtain $\partial^\beta_{\xi} f(x, \xi) = f^{(\beta)}(x, \xi)$ for $|\beta|_1 \leq \alpha - 1$.

Now, we prove (3.8). Let

$$u_j(x, \xi) = \partial^\beta_{\xi} h_j(x, \xi) \quad \text{and} \quad u(x, \xi) = \partial^\beta_{\xi} f(x, \xi), \quad |\beta|_1 = \alpha - 1.$$ 

Then the (3.11) implies

$$u(x, \xi) = \sum_{j=1}^{\infty} u_j(x, \xi) \quad \text{for} \quad (x, \xi) \in \Omega \times \Pi_0. \tag{3.12}$$
By the Cauchy inequality and (3.7), we have
\[ |D^k u_j|_{\Omega,1,\Pi_j} \leq C(k)M r_j^{-|k|}, \quad \text{for } |k|_1 \leq q, \tag{3.13} \]
where \( C(k) \) is a constant depending only on \( k \). By a similar proof to one for the compatibility and replacing (3.7) with (3.13), (3.12) implies
\[ \sup_{x \in \Omega} |D^k u(x, \xi) - D^k u(x, \zeta)| \leq C'_1 |\xi - \zeta|^\delta \quad \text{for } \xi, \zeta \in \Pi_0, |k|_1 \leq q. \]

The proof of the lemma is complete. \( \square \)

4. **Proof of Theorem 2.2.** We first introduce some notation so that the system (1.1) is written in a compact form. Denote
\[ A^0 = \diag(\varepsilon q_1 A_1, \varepsilon q_2 A_2), \quad B = \diag(B_1, B_2), \]
\[ \Lambda^0 = \diag(\varepsilon q_1 A_1, \varepsilon q_2 A_2), \quad \omega^0 = \col(\varepsilon q_1 \omega_1, \omega_2), \]
\[ P_1 = \diag(\varepsilon q_1 E_{n_1}, \varepsilon q_1 + q_2 - q_2 E_{n_1}), \]
\[ P_2 = \diag(\varepsilon q_1 + q_2 - q_2 E_{n_2}, \varepsilon q_1 - q_2 E_{n_2}), \quad P = \diag(P_1, P_2), \]
where \( E_n \) represents the \( n \times n \) identity matrix. Then the system (1.1) reads
\[ \left( \begin{array}{c} \dot{I} \\ \dot{\varphi} \end{array} \right) = \left( \begin{array}{c} A^0(\xi, \varepsilon) I \\ \omega^0(\xi, \varepsilon) \end{array} \right) + P G(I, \varphi, \xi, \varepsilon) \tag{4.1} \]
with \( G = \varepsilon q_0 \col(g_1, g_2, g_3, g_4) \).

a) **Outline of the Proof** We are going to prove Theorem 2.2 by employing the KAM iteration process. By Corollary 2 (see Section 3), we obtain a sequence of real analytic functions \( G^0 = 0, G^j = \varepsilon q_0 \col(g^0_1, g^0_2, g^0_3, g^0_4)(j = 1, 2, \ldots) \) approximating \( G \) and
\[ \lim_{j \to \infty} ||G^j - G||_{p,\alpha;\Omega^* \times T^{n_2};\Pi} = 0 \quad \text{for all } 0 \leq p < l, \tag{4.2} \]
\[ |G^j - G^{j-1}|_{|u|_{j-1,\alpha;\Pi}} \leq C_0 r_{j-1} ||G||_{l,\alpha;\Omega^* \times T^{n_2};\Pi} \quad \text{for } j \geq 1. \tag{4.3} \]
The definitions of \( G^*, \mathcal{U}_j \) etc are seen above Corollary 2. Denote \( G^1_j = \varepsilon q_0 \col(g^1_1, g^1_2) \) and \( G^2_j = \varepsilon q_0 \col(g^2_1, g^2_2) \). We truncate \( G^1 \) to its lower-degree terms
\[ \mathcal{L}(G^1) \equiv \begin{pmatrix} G^1_1(0, \varphi) + \partial_1 G^1_1(0, \varphi) I \\ G^1_2(0, \varphi) \end{pmatrix} \equiv \begin{pmatrix} u^0_1(\varphi) + u^0_1(\varphi) I + H^0_1 \\ u^0_2(\varphi) + H^0_2 \end{pmatrix} \]
and write (4.1) as
\[ \left( \begin{array}{c} \dot{I} \\ \dot{\varphi} \end{array} \right) = \left( \begin{array}{c} A^0 I \\ \omega^0 \end{array} \right) + P \left( \begin{array}{c} u^0_1(\varphi) + u^0_1(\varphi) I + H^0_1 \\ u^0_2(\varphi) + H^0_2 \end{array} \right) + P(G - G^1), \tag{4.4} \]
with \( \partial_1 f(I, \varphi) \) represents the partial derivative (Jacobian matrix) of \( f \) with respect to the variable \( I \). Here, we drop parameters from functions and will do this also in the sequel whenever there is no confusion.

Moreover, the Cauchy inequality (see Lemma A.3 in [28]) implies
\[ |u^0_1|_{r_0,\alpha;\Pi} \leq C_0 M \varepsilon q_2 r_0^{-l}, \quad |u^0_1|_{r_0,\alpha;\Pi} \leq C_0 M \varepsilon q_2 r_0^{-l-1}, \tag{4.5} \]
\[ |w^0|_{r_0,\alpha;\Pi} \leq C_0 M \varepsilon q_2 r_0^{-l-1}, \quad H^0_1 = O_{\mathcal{U}_l,\alpha;\Pi}(I^2), \quad H^0_2 = O_{\mathcal{U}_l,\alpha;\Pi}(I) \tag{4.6} \]
and
\[ |H^0_1|_{\mathcal{U}_l,\alpha;\Pi} \leq 2C_0 M \varepsilon q_2 r_0^{-l-2}, \quad |H^0_2|_{\mathcal{U}_l,\alpha;\Pi} \leq C_0 M \varepsilon q_2 r_0^{-l-1}, \tag{4.7} \]
where \( M \varepsilon q_2 = ||G||_{l,\alpha;\Omega^* \times T^{n_2};\Pi} \).

We want to look for a transformation \( T_1 \) to eliminate the lower-degree terms of \( PG^1 \) such that in new coordinates the lower-degree terms of analytic part in (4.4)
are much smaller than the old ones. Assume that at the \(\nu\)-th step of the process, we have already found a coordinate transformation \(T_\nu(\nu \geq 0)\) with \(T_0 = \text{Id}, \) the identity map\) such that the system \((4.1)\) is transformed into

\[
\begin{pmatrix}
\dot{\varphi} \\
\dot{\varphi} \\
\end{pmatrix} =
\begin{pmatrix}
A^\nu I \\
\omega^\nu \\
\end{pmatrix} + P \begin{pmatrix}
\tilde{u}_0^\nu(\varphi) + \tilde{u}_1^\nu(\varphi) I + \tilde{H}_1^\nu \\
\tilde{w}^\nu(\varphi) + H_2^\nu \\
\end{pmatrix} + P \mathcal{D}_\nu(G \circ T_\nu - G^\nu \circ T_\nu),
\]

where \(\tilde{H}_1^\nu = O(I^2), \tilde{H}_2^\nu = O(I), \mathcal{D}_\nu = P^{-1}(DT_\nu)^{-1}P,\) the circle \(\circ\) indicates composition of functions and \(DT\) the Jacobian matrix of \(T\) with respect to coordinate variables. Then we replace \(G^\nu\) with \(G^\nu + \nu^\nu\) which is closer to \(G,\) and the above equation is rewritten as

\[
\begin{pmatrix}
\dot{\varphi} \\
\dot{\varphi} \\
\end{pmatrix} =
\begin{pmatrix}
A^\nu I \\
\omega^\nu \\
\end{pmatrix} + P \begin{pmatrix}
u_0^\nu(\varphi) + \nu_1^\nu(\varphi) I + H_1^\nu \\
\tilde{w}^\nu(\varphi) + H_2^\nu \\
\end{pmatrix} + P \mathcal{D}_\nu(G \circ T_\nu - G^\nu + 1 \circ T_\nu),
\]

(4.8)

where

\[
\begin{pmatrix}
u_0^\nu(\varphi) + \nu_1^\nu(\varphi) I \\
\tilde{w}^\nu(\varphi) \\
\end{pmatrix} =
\begin{pmatrix}
\tilde{u}_0^\nu(\varphi) + \tilde{u}_1^\nu(\varphi) I \\
\tilde{w}^\nu(\varphi) \\
\end{pmatrix} + \mathcal{L}(\mathcal{D}_\nu(G^\nu + 1 \circ T_\nu - G^\nu \circ T_\nu)),
\]

\(H_3^\nu = O(I^2), \quad H_2^\nu = O(I)\).

We want to construct a coordinate change \(T^{\nu+1}\) to eliminate the lower-degree terms in \((4.8)\) such that the lower-degree terms of the next step are much smaller. Repetition of this process leads to a sequence of transformation \(T_\nu = T_{\nu-1} \circ T^\nu\) with \(T_0 = \text{Id}, \nu = 1, 2, \cdots,\) the limit transformation of which, if converges, reduces \((4.1)\) into a system without the lower-degree terms. Thus, we can obtain the quasi-periodic solution of \((4.1)\). The proof of convergence is due to the following iteration lemma which describes quantitatively the KAM iteration process.

b) Iteration Lemma Before stating the iteration lemma we first introduce the iterative sequences and notations used at each iteration step. Set

\[
\varepsilon_0 = \varepsilon^q_2, \quad ||G||_{I,\alpha;\Omega \times \mathbb{T}^{n_2}, \Pi} = M \varepsilon_0,
\]

\(\Omega = \{I \in \mathbb{R}^{n_1} : |I| < \tilde{r}\}, \quad \Omega^* = \{I \in \mathbb{R}^{n_1} : |I| \leq 2\tilde{r}\}, \quad \Omega_0 = \{I \in \mathbb{R}^{n_1} : |I| < \tilde{r}\}\)

with some constant \(0 < \tilde{r} \leq 1.\) For \(\nu \geq 1,\) let

(i) \(r_0 = \tilde{r}, \quad r_\nu = \tilde{r}^3^{-\nu},\)

\(\mathcal{U}_\nu = \Omega^* \times \mathbb{T}^{n_2} + (3r_\nu, 3r_\nu), \quad \mathcal{V}_\nu = \Omega_0 \times \mathbb{T}^{n_2} + (r_\nu, r_\nu), \quad \nu \geq 0,\)

\(\mathcal{V}_\nu^* = \Omega_0 \times \mathbb{T}^{n_2} + (2r_\nu, 2r_\nu) \subset \mathcal{V}_{\nu-1} \subset \mathcal{U}_\nu;\)

(ii) \(K_0 = 0, \quad K_\nu = [K'_\nu] + 1, \quad K'_\nu = 3^{\nu}r_0^{-1}(|\ln \tilde{C} + (n_2 + 1)\ln r_0| + (l + (n_2 + 1)\nu - \alpha) + 3); \quad \tilde{C} = 24(n_2)\mu_3^{n_2}e^{-\mu_2}, [K'_\nu] \) is the integer part of \(K'_\nu;\)

(iii) \(s_0 = \gamma, \quad s_\nu = \gamma(16c_3n_3\sqrt{n_2}K'_1)^{-1}, \quad \Pi^* = \{\xi \in \mathbb{R}^{n_3} : \text{dist}(\xi, \Pi_\nu) < s_\nu\};\)

(iv) \(\chi_\nu = r_\nu^{l-2(\alpha+1)(l+1)-\alpha-3}, \quad \Pi_\nu = \sum_{j=1}^\nu \chi_j, \) the assumption \(l > 2(\alpha + 1)(l + 2) + \alpha\) implies \(X_\nu = \sum_{j=1}^\nu \chi_j < \frac{1}{2};\)

(v) \(\delta_{\mu, \nu} = \gamma^{\nu-1}r_\nu^{l-2(\alpha+\mu+2)(l+1)-\alpha-3}C_0M\varepsilon_0, \quad 0 \leq \mu \leq \alpha;\)

(vi) \(f(I, \varphi, \xi) = O_{\mathcal{U}, \alpha; \Pi}(I^k)\) denotes a map which is real analytic in coordinate variables \((I, \varphi) \in \mathcal{U},\) continuously differentiable up to order \(\alpha\) in parameter \(\xi \in \Pi,\) and vanishes with \(I\)-derivatives up to order \(k - 1 \geq 0,\) and \(f\) and its \(\xi\)-derivatives up to order \(\alpha\) are bounded on \(\mathcal{U} \times \Pi.\)
Lemma 4.1. (Iteration Lemma) Assume that for the equation (4.8) with \( \nu \geq 0 \),
(v.1) (Frequency condition) let \( A^\nu = \text{diag}(\varepsilon^0 A_1^\nu, \varepsilon^0 A_2^\nu) \), \( A^\nu = \text{diag}(\varepsilon^0 A_1^\nu, \varepsilon^0 A_2^\nu) \), \( A^\nu = \text{diag}(\varepsilon^0 A_1^\nu, \varepsilon^0 A_2^\nu) \), \( A^\nu = \text{diag}(\varepsilon^0 A_1^\nu, \varepsilon^0 A_2^\nu) \), \( A^\nu = \text{diag}(\varepsilon^0 A_1^\nu, \varepsilon^0 A_2^\nu) \), \( A^\nu = \text{diag}(\varepsilon^0 A_1^\nu, \varepsilon^0 A_2^\nu) \), \( A^\nu = \text{diag}(\varepsilon^0 A_1^\nu, \varepsilon^0 A_2^\nu) \), \( A^\nu = \text{diag}(\varepsilon^0 A_1^\nu, \varepsilon^0 A_2^\nu) \), \( A^\nu = \text{diag}(\varepsilon^0 A_1^\nu, \varepsilon^0 A_2^\nu) \), \( A^\nu = \text{diag}(\varepsilon^0 A_1^\nu, \varepsilon^0 A_2^\nu) \), \( A^\nu = \text{diag}(\varepsilon^0 A_1^\nu, \varepsilon^0 A_2^\nu) \), \( A^\nu = \text{diag}(\varepsilon^0 A_1^\nu, \varepsilon^0 A_2^\nu) \), \( A^\nu = \text{diag}(\varepsilon^0 A_1^\nu, \varepsilon^0 A_2^\nu) \), \( A^\nu = \text{diag}(\varepsilon^0 A_1^\nu, \varepsilon^0 A_2^\nu) \), \( A^\nu = \text{diag}(\varepsilon^0 A_1^\nu, \varepsilon^0 A_2^\nu) \), \( A^\nu = \text{diag}(\varepsilon^0 A_1^\nu, \varepsilon^0 A_2^\nu) \), \( \Lambda^\nu = B_i \Lambda_i B_i^{-1} (i = 1, 2) \) and \( \omega^\nu = \text{col}(\varepsilon^0 \omega^\nu, \omega^\nu) \) satisfy, for \( \varepsilon \in (0, \varepsilon^* \] ,
\[
\inf_{\xi \in \Pi_{\nu}} |\lambda_j| \geq c_0 (1 - X_{\nu}) > \frac{c_0}{2}, \quad \inf_{\xi \in \Pi_{\nu}} |\lambda_j - \lambda| \geq c_0 (1 - X_{\nu}) > \frac{c_0}{2}
\] for \( i \neq j, 1 \leq i, j \leq n_{11} \), or \( n_{11} + 1 \leq i, j \leq n_1 \), and
\[
\|A^\nu\|_{\alpha; \Pi_{\nu}} \leq c_0 (1 + X_{\nu}) < 2c_1, \quad \|\partial^\beta \omega^\nu\|_{\Pi_{\nu}} \leq c_0 (1 + X_{\nu}) \varepsilon^{q_\beta} < 2c_1 \varepsilon^{q_\beta}, \quad 1 \leq |\beta| \leq \alpha,
\]
\[
|A^\nu_{\nu} - A^\nu_{\nu-1}|_{\alpha; \Pi_{\nu}} \leq c_0 C_0 M \varepsilon_0 \varepsilon^{b_1} r_{\nu}^{-\alpha+1}(i+2), \quad |\omega^\nu_{\nu} - \omega^\nu_{\nu-1}|_{\alpha; \Pi_{\nu}} \leq C_0 M \varepsilon_0 \varepsilon^{b_1+2} r_{\nu}^{-\alpha+1}(i+2), \nu \geq 1
\] for \( i = 1, 2 \), where \( \varepsilon_0 \) is a positive constant, \( c_0 \) and \( c_1 \) are given in Assumption (H2), \( b_1 = 0, b_2 = q_\nu - q_\nu, b_3 = q_\nu - q_\nu, b_4 = q_\nu - q_\nu \);
(v.2) (Small condition) the terms \( u^\nu_{\nu}, u^\nu_{\nu-1} \) and \( u^\nu \) satisfy the following estimates
\[
|u^\nu_{\nu} - u^\nu_{\nu-1}|_{\alpha; \Pi_{\nu}} \leq 4C^0 M \varepsilon_0 r_{\nu}^{-\alpha}, \quad |u^\nu_{\nu} - u^\nu_{\nu-1}|_{\alpha; \Pi_{\nu}} \leq 4C^0 M \varepsilon_0 r_{\nu}^{-\alpha}
\]
\[
|u^\nu - u^\nu_{\nu-1}|_{\alpha; \Pi_{\nu}} \leq 4C^0 M \varepsilon_0 r_{\nu}^{-\alpha}
\]
\[
H^\nu_{\nu}(I, \varphi, \xi) \quad \text{and} \quad H^\nu_{\nu}(I, \varphi, \xi) \quad \text{fulfill}
\]
\[
H^\nu_{\nu} = O_{\nu; \alpha; \Pi_{\nu}} (I^2), \quad H^\nu_{\nu} = O_{\nu; \alpha; \Pi_{\nu}} (I), \quad |H^\nu_{\nu} - H^\nu_{\nu-1}|_{\nu; \alpha; \Pi_{\nu}} \leq C_1 \varepsilon_0 \varepsilon_0 M \varepsilon_0
\] for \( \nu \geq 1, i = 1, 2 \);
(v.3) (Transformation) the transformation \( T^\nu : \nu \times \Pi^\nu_{\nu} \to \nu \) is real analytic in coordinate variables \( (I, \varphi) \in \nu \), and continuously differentiable up to order \( \alpha \) in the parameter \( \xi \in \Pi^\nu_{\nu} \), satisfies
\[
\|T^\nu - T^\nu_{\nu-1}|_{\nu; \alpha; \Pi^\nu_{\nu}} \leq (1 + X_{\nu}) C_1 C_0 M \varepsilon_0 M \varepsilon_0 r_{\nu}^{-\alpha}(i+2), \quad \gamma_{\nu} \leq C_{1, \nu} \varepsilon^{2q_\nu}, \nu \geq 0
\] with \( T_0 = \text{Id} \) and \( 0 \leq \mu < \alpha \), where \( C_1 \) is a constant independent of \( \nu \).

Then there exists a closed set \( \Pi_{\nu+1} \subset \Pi_{\nu} \)
\[
\Pi_{\nu+1} = \{ \xi \in \Pi_{\nu} : \sqrt{-1} (\xi, \omega^\nu) + (m, \Lambda^\nu) |\gamma^{\nu}| \leq M \varepsilon_0 |k|_{\nu} \}
\]
for all \( m \in \mathbb{M}, K_{\nu} < |k|_{\nu} \leq K_{\nu+1} \) (see Theorem 2.2 and (H3) for definitions of \( \mathbb{M} \) and \( \nu \), respectively) and a coordinate transformation
\[
T^\nu_{\nu+1} : \nu \times \Pi^\nu_{\nu+1} \to \nu \subset \nu \subset \nu_{\nu+1}
\]
in the form
\[
I = \rho + \nu^\nu_{\nu} (\phi, \xi) + \nu^\nu_{\nu} (\phi, \xi) \rho, \quad \varphi = \phi + \Phi^\nu (\phi, \xi)
\]
where \( \rho \) and \( \phi \) are new coordinate variables, and all terms in the transformation are real analytic in \( \phi \) and continuously differentiable in \( \xi \) up to order \( \alpha \), satisfy the estimates
\[
|\Phi^\nu|_{2r_{\nu+1}, \alpha; \Pi^\nu_{\nu+1}} \leq C_1 C_0 M \varepsilon_0 \varepsilon^{2q_\nu} r_{\nu+1}^{-\alpha-1},
\]
\[
|\nu^\nu_{\nu} - \nu^\nu_{\nu-1}|_{2r_{\nu+1}, \alpha; \Pi^\nu_{\nu+1}} \leq C_1 C_0 M \varepsilon_0 \varepsilon^{b_1+2} r_{\nu+1}^{-\alpha+1},
\]
\[
|\nu^\nu_{\nu} - \nu^\nu_{\nu-1}|_{2r_{\nu+1}, \alpha; \Pi^\nu_{\nu+1}} \leq C_1 C_0 M \varepsilon_0 \varepsilon^{b_1+2} r_{\nu+1}^{-\alpha+1},
\]
and
\[
|P^{-1} (DT^\nu_{\nu+1})^{-1} \nu^\nu_{\nu+1} \leq 1 + \chi_{\nu+1},
\]
\[
\left| \frac{\partial}{\partial \xi} \left( P^{-1}(DT^{\nu+1})^{-1}P \right) \right|_{\nu=0; \Pi} < \lambda \nu+1 \quad (4.17)
\]
for \(1 \leq |\beta| \leq \alpha\), such that the equation (4.8) is transformed into
\[
\left( \begin{array}{c}
\hat{\rho} \\
\phi
\end{array} \right) = \left( \begin{array}{c}
A^\nu(\phi) + D^\nu(\phi) + H^\nu \phi \\
\omega^\nu(\phi) + H^\nu \phi
\end{array} \right) + P \mathcal{D}^\nu(G \circ T^\nu - G^\nu \circ T^\nu+1)
\]
and the conditions (v.1)-(v.3) are satisfied by replacing \(\nu\) by \(\nu+1\) and \((I, \varphi)\) by \((\rho, \phi)\), respectively, where \(T^\nu = T_\nu \circ T^\nu+1, \mathcal{D}^\nu \equiv P^{-1}(DT^\nu+1)^{-1}P\).

c) **Proof of Theorem 2.2** Theorem 2.2 is easy to be proven by the Iteration Lemma and Inverse Approximation Lemma.

First the system (1.1) has been written in the form (4.4) just as (4.8) satisfying the conditions (v.1)-(v.3) with \(\nu = 0\) in the Iteration Lemma by Assumptions (H2) and (H3), (4.5) and (4.6). We use the Iteration Lemma inductively to obtain a sequence of transformations \(T^\nu\) mapping \(V_0 \times \Pi^\nu\) into \(V_0\) and satisfying the estimate (4.11). Noting that \(V_\nu\) and \(\Pi^\nu\) are exactly regarded as those neighbourhoods of the open convex set \(\Omega_0 \times T^{\nu+1} \subset \mathbb{R}^{n+1}\) and closed subset \(\Pi^\nu \subset \Pi\), respectively, and \(s^\nu/s^\nu_{\nu} \to 0\) as \(\nu \to \infty\) \((l_1 = l - (\alpha + \mu + 2)(\nu+1) - \alpha - 2\) and the positive integer \(\mu \leq \alpha\) by the definition of \(s^\nu\), the Inverse Approximation Lemma and Condition (v.3) imply that for every \(\xi \in \Pi^\nu\), the limit map \(T = \lim_{\nu \to \infty} T^\nu\) exists in \(C^\mu,\mu\nu-1(\Omega_0 \times T^{\nu+1}, \Pi)\) for \(0 < p < l_1\) and \(T : \Omega_0 \times T^{\nu+1} \times \Pi^\nu \to \Omega^\nu \times T^{\nu+1}\) for sufficiently small \(\varepsilon\), and is of the form
\[
T: \quad I = \rho + V_0(\phi, \xi) + V_1(\phi, \xi)\rho, \quad \varphi = \phi + \Phi(\phi, \xi)
\]
by (4.13), which is of class \(C^\mu\) in \(\xi \in \Pi^\nu\) in the sense of Whitney and of class \(\tilde{C}\tilde{\nu}\) in \(\varphi \in T^{\nu+1}\) together with derivatives up to order \(\mu\) with respect to \(\xi\) for \(0 < \mu \leq \alpha\).

Moreover, by (4.2), we obtain
\[
\lim_{\nu \to \infty} ||G \circ T^\nu - G^\nu \circ T^\nu||_{p,\mu-1; \Omega_0 \times T^{\nu+1}, \Pi^\nu} = 0 \quad \text{for all } 0 < p < l_1, 0 < \mu \leq \alpha
\]
(4.18)
and by Condition (v.3) and (17),
\[
|\mathcal{D}^\nu - \mathcal{D}^\nu|_{V_\nu, \Pi^\nu} \leq C_2 C_0 M_0 \varepsilon_0 \gamma^{- \mu - 1} r^\nu(\nu+1) - \alpha - 3
\]
(4.19)
where \(C_2\) is a constant independent of \(\nu, \gamma, \varepsilon_0\). It follows from (4.10), (4.18) and (4.19) that System (4.1) is transformed by \(T\) into the system
\[
\left\{ \begin{array}{l}
\hat{\rho} = A^\nu(\xi)\rho + P_1 O(\rho^2) \\
\hat{\phi} = \omega^\nu(\xi) + P_2 O(\rho)
\end{array} \right.
\]
(4.20)
for \((\rho, \phi) \in \Omega_0 \times T^{\nu+1}, \xi \in \Pi^\nu\), where \(A^\nu(\xi) = \text{diag}(\varepsilon^{\nu_1} A^\nu_1(\xi), \varepsilon^{\nu_2} A^\nu_2(\xi), \omega^\nu(\xi) = \text{col}(\varepsilon^{\nu_1} \omega^\nu_1(\xi), \varepsilon^{\nu_2} \omega^\nu_2(\xi)))\), \(A_\nu^\nu = \lim_{\nu \to \infty} A^\nu_\nu\) and \(\omega^\nu_i = \lim_{\nu \to \infty} \omega^\nu_i(i = 1, 2)\) exist by Condition (v.1) and are of class \(C^\alpha\) in \(\xi \in \Pi^\nu\) in the sense of Whitney by the Inverse Approximation Lemma since \(r^\nu(\nu+1)(\nu+2)(\nu+1) - \alpha - 3 \to 0\) as \(\nu \to \infty\). Thus, we obtain the quasi-periodic invariant torus of (4.1)
\[
I = V_0(\phi, \xi), \quad \varphi = \phi + \Phi(\phi, \xi), \quad \phi = \omega^\nu(\xi) t + \phi_0
\]
satisfying the estimates (2.5), (2.6) and (2.7) by Conditions (v.1) and (v.3). The rest of Theorem 2.2 can be derived immediately from the Iteration Lemma.

d) **Proof of Iteration Lemma** To simplify the notation, we denote quantities referring to \(\nu + 1\) with \(+\) such as \(u^{\nu+1}\) by \(u^+, r^\nu+1\) by \(r_+, \) and those referring to \(\nu\)
without the $\nu$ such as $u^\nu$ by $u$, $r_\nu$ by $r$. Substituting the transformation $T^+$ into (4.8), the transformation $T^+$ will be obtained by solving the homological equations

$$\begin{align*}
\partial \phi v_0 \cdot \omega - Av_0 &= P_1 \Gamma_{K^+} u_0(\phi), \\
\partial \phi v_1 \cdot \omega + v_1 A - Av_1 &= P_1 (\Gamma_{K^+} u_1(\phi) - B \text{diag}(B^{-1} \uhat_1(0) B) B^{-1}), \\
\partial \phi \Phi \cdot \omega &= P_2 (\Gamma_{K^+} w(\phi) - \what(0)),
\end{align*}$$

(4.21)-(4.23)

where $B = \text{diag}(B_1, B_2), \text{diag}(B^{-1} \uhat_1(0) B)$ denotes a diagonal matrix whose elements are the diagonal elements of $B^{-1} \uhat_1(0) B$, $\uhat_1(0)$ and $\what(0)$ denote the mean values (that is, the zero-order coefficients of the Fourier series expansions) of $u_1$ and $w$ over $\mathbb{T}^{n_2}$, respectively, $\Gamma_{K^+}$ is the truncation operator of the Fourier series expansions defined in Lemma 6.7 and the notation $\partial \phi f \cdot \omega$ means $\partial \phi f \cdot \omega = \sum_{j=1}^{n_2} t_j \frac{\partial f}{\partial \psi_j}$ for $\omega = \text{col}(t_1, \cdots, t_{n_2})$. Here, the homological equations are approximated by truncating the Fourier series expansions of $u_0, u_1$ and $w$ so that the solutions are defined on an open set of parameters. This idea is due to Arnol'd[2] and Pöschel[27].

d1) Solutions of (4.21)-(4.23) and estimates. Set

$$\Pi_+ = \{ \xi \in \Pi_\nu : |\sqrt{-1} \langle k, \omega \rangle + \langle m, \Lambda \rangle | \geq \gamma \varepsilon^{\nu} |k|_2^{-\epsilon}, m \in \mathbb{m}, K < |k|_2 \leq K^+ \}$$

and

$$\Pi_{\nu}^{s+} = \{ \xi \in \mathbb{R}^{n_3} : \text{dist}(\xi, \Pi_+) < s_+ \} \subset \Pi_\nu^{\nu}.$$

Lemma 4.2. For every $\xi \in \Pi_{\nu}^{s+}$, we have

$$|\sqrt{-1} \langle k, \omega(\xi) \rangle + \langle m, \Lambda(\xi) \rangle | \geq \frac{1}{4} \gamma \varepsilon^{\nu} |k|_2^{-\epsilon}, \quad 0 < |k|_2 \leq K^+, m \in \mathbb{m}. \quad (4.24)$$

Proof. We first prove

$$|\sqrt{-1} \langle k, \omega(\xi) \rangle + \langle m, \Lambda(\xi) \rangle | \geq \frac{1}{2} \gamma \varepsilon^{\nu} |k|_2^{-\epsilon}, \quad 0 < |k|_2 \leq K^+, m \in \mathbb{m} \quad (4.25)$$

for every $\xi \in \Pi_+$. Noting the fact that $\Pi_{x_j}^{l+1} r_j^{l-(\alpha+1)(\nu+2)-2} \to 0$ as $j \to \infty$, the (4.9) implies that for $0 < |k|_2 \leq K_j$, $1 \leq j \leq \nu$,

$$|\sqrt{-1} \langle k, \omega^j(\xi) - \omega^j(-1) \rangle + \langle m, \Lambda^j(\xi) - \Lambda^j(-1)(\xi) \rangle | \leq \varepsilon^{\nu} (\sqrt{2} \gamma + 2 \varepsilon_0) C_0 M \varepsilon_0 r_j^{l-(\alpha+1)(\nu+2)-2} < \gamma \varepsilon^{\nu} K_j^{-\epsilon}$$

for sufficiently small $\varepsilon_0$. As for $K_{j-1} < |k|_2 \leq K_j$, $|\sqrt{-1} \langle k, \omega^j(\xi) + \langle m, \Lambda^j(\xi) \rangle | \geq \gamma \varepsilon^{\nu} |k|_2^{-\epsilon}$, hence,

$$|\sqrt{-1} \langle k, \omega^\nu(\xi) \rangle + \langle m, \Lambda^\nu(\xi) \rangle | \geq \frac{1}{2} \gamma \varepsilon^{\nu} |k|_2^{-\epsilon}$$

which implies (4.25).
For every $\xi \in \Pi^+_{\nu} \subset \Pi^+_{\alpha}$, there is $\xi_0 \in \Pi_+$ such that $|\xi - \xi_0| < s_+$. The condition (v.1) and (4.25) imply
\[ |\sqrt{-\mathcal{L}}(k, \omega(\xi)) + (m, \Lambda(\xi))| \geq |\sqrt{-\mathcal{L}}(k, \omega(\xi_0)) + (m, \Lambda(\xi_0))| - |\sqrt{-\mathcal{L}}(k, \omega(\xi) - \omega(\xi_0)) + (m, \Lambda(\xi) - \Lambda(\xi_0))| \geq \frac{1}{2} \gamma \varepsilon^{\frac{q}{5}} |k|^{-\varepsilon^{-1}} - 2c_1n_3 \varepsilon^{\frac{q}{5}} (\sqrt{n_2} |k| + 2)s_+ \geq \frac{1}{4} \gamma \varepsilon^{\frac{q}{5}} |k|^{-\varepsilon^{-1}}. \]

The procedure of solving (4.21)-(4.23) is standard in KAM theory. Expanding the functions into the Fourier series in $\phi$, and substituting in (4.21)-(4.23) and comparing coefficients of the term $e^{\sqrt{-\mathcal{L}}(k, \phi)}$, one obtain the solutions
\[ v_0(\phi) = P_1 \sum_{|k| \leq K_+} B(\sqrt{-\mathcal{L}}(k, \omega) - \Lambda)^{-1} \bar{u}_0(k)e^{\sqrt{-\mathcal{L}}(k, \phi)}, \]  
\[ v_1(\phi) = P_1 \sum_{|k| \leq K_+} BV_1(k)B^{-1}e^{\sqrt{-\mathcal{L}}(k, \phi)}, \]  
\[ \Phi(\phi) = P_2 \sum_{0 < |k| \leq K_+} (\sqrt{-\mathcal{L}}(k, \omega))^{-1} \tilde{u}(k)e^{\sqrt{-\mathcal{L}}(k, \phi)}, \]
where
\[ (V_1(k))_{ij} = \begin{cases} (\sqrt{-\mathcal{L}}(k, \omega) + \varepsilon^{\varepsilon} \lambda_j - \varepsilon^{\varepsilon} \lambda_i)^{-1} (\tilde{U}_1(k))_{ij}, & |k| + |i - j| \neq 0, \\ 0, & |k| + |i - j| = 0, \end{cases} \]

$U_1(\phi) = B^{-1}u_1(\phi)B$, $a_i = q_i$ if $1 \leq i \leq n_{11}$, $= q_3$ if $n_{11} + 1 \leq i \leq n_1$, $(V_1(k))_{ij}$ and $(\tilde{U}_1(k))_{ij}$ represent elements of the matrices $V_1(k)$ and $\tilde{U}_1(k)$, respectively, and $u(k)$ is the $k$-order coefficients of the Fourier series expansions of $u$. Hence, Lemma 4.2, (4.26)-(4.28) and Conditions (v.1)-(v.2) imply that $v_0, v_1$ and $\Phi$ are real analytic in $\phi \in \mathcal{W} \equiv \mathcal{W}^{n_2} + 2r_2$, continuously differentiable up to order $\alpha$ in $\xi \in \Pi^+_{\nu}$. Meanwhile using Lemma 6.9, one easily gets the estimates (4.14)-(4.16) (we denote $|\cdot|_{r, \mu; \Pi^+_{\nu}}$ by $|\cdot|_{r, \mu; \nu}$, $|\cdot|_{r, \mu; \nu}$ by $|\cdot|_{r, \mu; \nu}$ for simplification) and
\[ |v_0|_{2r_+ + \mu; s} \leq C_1 \varepsilon^{\mu} \delta^{+\mu} r_+^{\alpha_1+3}, \quad |P_1^{-1} \partial_\phi v_0|_{2r_+ + \mu; s} \leq C_1 \varepsilon^{-\mu} \delta^{+\mu} r_+^{\alpha_1+3+1}, \]  
\[ |v_1|_{2r_+ + \mu; s} \leq C_1 \varepsilon^{\mu} \delta^{+\mu} r_+ + |P_1^{-1} v_1 P_1|_{2r_+ + \mu; s} \leq C_1 \varepsilon^{\mu} \delta^{+\mu} r_+ + |P_1^{-1} \partial_\phi v_1|_{2r_+ + \mu; s} \leq C_1 \varepsilon^{-\mu} \delta^{+\mu} r_+ \]
and
\[ |\Phi|_{2r_+ + \mu; s} \leq C_1 \delta^{+\mu} r_+^2, \quad |P_2^{-1} \partial_\phi \Phi|_{2r_+ + \mu; s} \leq C_1 \varepsilon^{-\mu} \delta^{+\mu} r_+ \]
for $0 \leq \mu \leq \alpha$ and an appropriate choice of the constant $C_1$ independent of $\nu$.

It is easy to see that when the $\varepsilon_0$ is sufficiently small, the transformation $T^+$ maps $\mathcal{V}_+$ into $\mathcal{V}_+^*$ in $\mathcal{V}$ and $\mathcal{V}_+^*$ into $\mathcal{V}$, respectively, and
\[ |T^+ - \text{Id}|_{\mathcal{V}_+^{*} + 0; s} \leq C_1 r_+ \delta^{+0}, \quad |\partial_\phi T^+|_{\mathcal{V}_+^{*} + 0; s} \leq C_1 r_+ \delta^{+\mu}, \quad 1 \leq |\beta|_1 = \mu \leq \alpha. \]

**d2** Proof of (4.17). Corresponding to the transformation $T^+$, we have its Jacobian matrix
\[ DT^+ = \begin{pmatrix} E_{n_1} + v_1 & \partial_\phi v_1 + \partial_\phi v_1 \rho \\ 0 & E_{n_2} + \partial_\phi \Phi \end{pmatrix} \] (4.33)
and the inverse
\[
(DT^+)^{-1} = \begin{pmatrix}
(E_n + v_1)^{-1} & -(E_n + v_1)^{-1}\partial_v v_0 + \partial_v v_1\rho (E_n + \partial_\Phi)^{-1} \\
0 & (E_n + \partial_\Phi)^{-1}
\end{pmatrix}
\]
(4.34).

Thus, (4.29)-(4.31) imply
\[
|P^{-1}(DT^+ - E)|_{V_+;a;\alpha} \leq 1 + \chi_+r_+^{\alpha+1},
\]
for 1 \leq \mu \leq \alpha. Noting that the derivatives of DT^+ with respect to the parameter \( \xi \) is sufficiently small and that for a matrix \( M(\xi) \) with a small norm, differentiating the left- and right-hand sides of \((E + M(\xi))^{-1}(E + M(\xi)) = E\) and using the Leibniz formula, we find
\[
\partial_\xi^\beta (E + M(\xi))^{-1} = -\sum_{k<\beta} \begin{pmatrix} \beta \\ k \end{pmatrix} \partial_\xi^k (E + M(\xi))^{-1} \cdot \partial_\xi^{\beta-k} (E + M(\xi)) \cdot (E + M(\xi))^{-1},
\]
where \( k, \beta \in \mathbb{Z}_+ \), \( \begin{pmatrix} \beta \\ k \end{pmatrix} = \left( \begin{array}{c} \beta_1 \\ k_1 \\ \vdots \\ \beta_n \\ k_n \end{array} \right) \), \( E \) is the identity matrix, the estimates (4.30) and (4.31) imply
\[
|P_1^{-1}(E_n + v_1)^{-1}P_1|_{2r_+;a;\alpha} \leq 1 + \chi_+r_+^{\alpha+1},
\]
(4.36)
\[
|P_2^{-1}(E_n + \partial_\Phi)^{-1}P_2|_{2r_+;a;\alpha} \leq 1 + \chi_+r_+^{\alpha+1},
\]
(4.37)
\[
|\partial_\xi^\beta (P_1^{-1}(E_n + v_1)^{-1}P_1)|_{2r_+;a;\alpha} \leq 2(1 + r_+\chi_+)^2 |P_1^{-1}v_1P_1|_{2r_+;\beta;\alpha},
\]
(4.38)
\[
|\partial_\xi^\beta (P_2^{-1}(E_n + \partial_\Phi)^{-1}P_2)|_{2r_+;a;\alpha} \leq 2(1 + r_+\chi_+)^2 |P_2^{-1}\partial_\Phi P_2|_{2r_+;\beta;\alpha},
\]
(4.39)
for 1 \leq |\beta|_1 \leq \alpha and sufficiently small \( \varepsilon_0 \). Hence, (4.17) follows from (4.34), (4.29), (4.30) and (4.36)-(4.39). Moreover, we have
\[
\begin{pmatrix}
|P_1^{-1}(E_n + v_1)^{-1}P_1F|_{V_+;a;\alpha} \\
|P_2^{-1}(E_n + \partial_\Phi)^{-1}P_2F|_{V_+;a;\alpha} \\
|P^{-1}(DT^+)^{-1}PF|_{V_+;a;\alpha}
\end{pmatrix} < (1 + r_+\chi_+)|F|_{V_+;a;\alpha},
\]
(4.40)
for suitable \( F \) which is real analytic in \( V_+ \) and continuously differentiable up to order \( \alpha \) in \( \xi \in \Pi_+^\varepsilon \).

Finally, we proceed to verify (v.3) with \( \nu + 1 \) replacing \( \nu \). As the transformation \( T^+ : V_+ \times \Pi_+^\varepsilon \to V_+ \times \Pi_+^\varepsilon \) (or \( V_+^\ast \times \Pi_+^\varepsilon - V \)) is real analytic in coordinate variables and continuously differentiable up to order \( \alpha \) in parameter \( \xi \), so is \( T_\nu = T \circ T^+ \).

We first prove (4.11) and (4.12) inductively. For \( \nu = 1 \), by (4.29)-(4.33), it implies
\[
|T_1 - \text{Id}|_{V_+;\alpha} \leq C_1 r_1 \delta_\lambda \mu < r_1 \chi_1, \quad |DT_1 - E|_{V_+;\alpha} \leq C_1 \delta_\lambda \mu < \chi_1
\]
and
\[
|P^{-1}(DT_1 - E)|_{V_+;\alpha} \leq C_1 \delta_\lambda \mu < \chi_1
\]
for 0 \leq \mu \leq \alpha. Assume that at the \( \nu \)-th step we have
\[
|T - T_{\nu-1}|_{V_+;\alpha} \leq (1 + X)C_1 r_\nu \delta_\nu \mu < r_\chi, \quad |DT - DT_{\nu-1}|_{V_+;\alpha} \leq (1 + X)C_1 \delta_\nu \mu < \chi
\]
and
\[
|P^{-1}(DT - DT_{\nu-1})|_{V_+;\alpha} \leq 2(1 + X)C_1 \delta_\nu \mu < \chi
\]
for 0 \leq \mu \leq \alpha, here we have omitted the subscript \( \nu \) from the quantities referring to \( \nu \). Then in view of the induction assumptions we obtain
\[
|DT|_{V_0;\alpha} \leq 1 + X, \quad |\partial_\xi^\beta DT|_{V_0;\alpha} \leq X \quad \text{for } 1 \leq |\beta|_1 \leq \alpha.
\]
(4.41)
Combining (4.32), (4.41) and Lemma 6.8 (i) we get
\[ |T_+ - T|_{\nu_+;\mu;\delta_{+\mu}} = |T \circ T^+ - T|_{\nu_+;\mu;\delta_{+\mu}} \leq (1 + X_+)C_1r_+\delta_{+\mu} < r_+\chi_+ \quad \text{for } 0 \leq \mu \leq \alpha, \]
which, together with the Cauchy inequality, implies
\[ |DT_+ - DT|_{\nu_+;\mu;\delta_{+\mu}} \leq (1 + X_+)C_1\delta_{+\mu} < \chi_+ \quad \text{for } 0 \leq \mu \leq \alpha. \]
Similarly, we have
\[ |P^{-1}DTP|_{\nu_+;\mu;\delta_{+\mu}} \leq 1 + X, \quad |\partial_\xi^3(P^{-1}DTP)|_{\nu_+;\mu;\delta_{+\mu}} \leq X \quad \text{for } 1 \leq |\beta| \leq \alpha \] (4.44) and
\[ |(P^{-1}DTP) \circ T^+ - P^{-1}DTP|_{\nu_+;\mu;\delta_{+\mu}} \leq (1 + \chi_+)(1 + X)C_1\delta_{+\mu} \quad \text{for } 0 \leq \mu \leq \alpha \] (4.45) by (4.32), (4.44) and Lemma 6.8 (i). Based on the observation
\[ P^{-1}(DT_+ - DT)P = ((P^{-1}DTP) \circ T^+ - P^{-1}DTP)(P^{-1}DT^+P) + P^{-1}DTP(P^{-1}(DT^+ - E)P), \]
from (4.45), (4.35), (4.44) and the Leibniz formula, it follows
\[ |P^{-1}(DT_+ - DT)P|_{\nu_+;\mu;\delta_{+\mu}} \leq 2(1 + X_+)C_1\delta_{+\mu} < \chi_+ \]
for $0 \leq \mu \leq \alpha$ and sufficiently small $\varepsilon_0$. Thus, we have proved (4.11) and (4.12) with $\nu + 1$.

Now, we show that $T_+$ maps $\mathcal{V}_+$ into $\mathcal{U}_+$. Noting the expression of $T_+$ in angle variable direction is independent of $\rho$, we set $T_+(\rho, \phi) = \text{col}(v_+(\rho, \phi), \Phi_+(\phi))$. In view of the induction hypotheses, (4.42) implies
\[ |T_+ - \text{Id}|_{\nu_+;\mu;\delta_{+\mu}} \leq \sum_{j=1}^{\nu+1} |T_j - T_{j-1}|_{\nu_+;\mu;\delta_{+\mu}} \leq \sum_{j=1}^{\nu+1} r_j < r_0, \]
and (4.43) implies
\[ |DT_+|_{\nu_+;\mu;\delta_{+\mu}} < 1 + X_+ < 2. \] (4.46)
Hence, the first component of $T_+$ is mapped into $\Omega^* + r_+$. For $\phi$ with $|\text{Im}\phi| < r_+$, there exists a $\phi_0 \in \mathbb{T}^{n_2}$ such that $|\phi - \phi_0| < r_+$. Therefore, the $\Phi_+$ being real analytic and (4.46) imply
\[ |\text{Im}\Phi_+(\phi)| = |\text{Im}(\Phi_+(\phi) - \Phi_+(\phi_0))| \leq |DT_+|_{\nu_+;\mu;\delta_{+\mu}}|\phi - \phi_0| < 2r_+. \]
Thus, for $\xi \in \Pi_+^*$, $T_+$ maps $\mathcal{V}_+ \times \mathcal{V}_+$ into $\Omega^* \times \mathcal{V}_+ \times (2r_+ + 2r_+) \subset \mathcal{U}_+$, as claimed. Furthermore, let $\mathcal{G}(\rho, \phi) = \mathcal{D}_+(G^\nu \circ T_+ - G^{\nu+1} \circ T_+)$, then (4.3), (4.11), (4.12), Lemma 6.8 (ii) and the Cauchy inequality imply
\[ |\mathcal{G}|_{\nu_+;\alpha;\delta_{+\mu}} \leq 2r_+^{l-\alpha}C_0\varepsilon_0 \] (4.47) and
\[ |\partial_\phi \mathcal{G}|_{\nu_+;\alpha;\delta_{+\mu}} \leq 2r_+^{l-\alpha-1}C_0\varepsilon_0 \] (4.48) for sufficiently small $\varepsilon_0$.

d4) Estimates of remainder terms. Denote \[ W(I, \varphi) = \text{col}(u_0(\varphi) + u_1(\varphi)I, w(\varphi)), \quad A^+ = A + P_1\tilde{A}, \]
\[ \Lambda^+ = \Lambda + P_1\tilde{\Lambda}, \quad \omega^+ = \omega + P_2\tilde{\omega}, \]
where \[ \tilde{A} = B(\text{diag}(B^{-1}\hat{u}_1(0)B))B^{-1}, \quad \tilde{\Lambda} = \text{diag}(B^{-1}\hat{u}_1(0)B), \quad \tilde{\omega} = \hat{w}(0). \]
Then the assumption (H2) implies that there is a constant $\tilde{c}_0 \geq 1$ such that
\[
\| \text{diag}(B^{-1}u_0(0)B) \|_{\alpha, \gamma} \leq \tilde{c}_0 \| u_1 \|_{r, \alpha, \gamma, x}, \quad \| \tilde{A} \|_{\alpha, \gamma} \leq \tilde{c}_0 \| u_1 \|_{r, \alpha, \gamma, x}.
\]
(4.49)
By Assumption (H1) and Condition (v.2), it is easy to see $A^+, \Lambda^+$ and $\omega^+$ satisfy (v.1) with $v$ replaced by $\nu + 1$. We have found the transformation $T^+$ which transforms the equation (4.8), by using (4.21)-(4.23), into the following one in the new variables
\[
\begin{pmatrix}
\dot{\rho} \\
\dot{\phi}
\end{pmatrix} = \begin{pmatrix}
A^+ \rho \\
\omega^+
\end{pmatrix} + P\mathcal{D}^+ \left\{ P^{-1}(E - DT^+)P \left( \dot{A}\rho \right) + (\text{Id} - \Gamma_{K^+})W(\rho, \phi) + W \circ T^+(\rho, \phi) - W(\rho, \phi) + \left( \begin{array}{c}
H_1 \\
H_2
\end{array} \right) \circ T^+(\rho, \phi) \right\}
+ P\mathcal{G}(\rho, \phi) + P\mathcal{D}_+(G - G^{\nu+2}) \circ T^+(\rho, \phi),
\]
where $\mathcal{D}^+ = (P^{-1}DT^+P)^{-1}, \mathcal{D}_+ = (P^{-1}DT_+P)^{-1}$. We use the notation $Lf(\rho, \phi)$ to denote the linear part of a function $f$ in $\rho$, that is
\[
Lf(\rho, \phi) = f(0, \phi) + \partial_\rho f(0, \phi) \rho
\]
and denote $\mathcal{G}(\rho, \phi) = \text{col}(\mathcal{G}_1(\rho, \phi), \mathcal{G}_2(\rho, \phi))$, rewrite the above equation in the form of (4.8)
\[
\begin{pmatrix}
\dot{\rho} \\
\dot{\phi}
\end{pmatrix} = \begin{pmatrix}
A^+ \rho \\
\omega^+
\end{pmatrix} + P \left( u_0^+(\phi) + u_1^+(\phi) \rho + H_1^+ \right) + P\mathcal{D}_+(G - G^{\nu+2}) \circ T^+(\rho, \phi),
\]
(4.50)
where
\[
w^+ = P_2^{-1}(E_{n_2} + \partial_\phi \Phi)^{-1}P_2 \left[ -P_2^{-1}\partial_\phi \Phi P_2 \omega + (\text{Id} - \Gamma_{K^+})w + w(\phi + \Phi) - w(\phi) + H_2(v_0, \phi + \Phi) + \mathcal{G}_2(0, \phi) \right],
\]
(4.51)
\[
u_0^+ = P_1^{-1}(E_{n_1} + v_1)^{-1}P_1 \left[ -P_1^{-1}\partial_\phi \nu_0 P_2(\omega^+ + w^+(\phi)) + (\text{Id} - \Gamma_{K^+})u_0 + u_0(\phi + \Phi) - u_0(\phi) + u_1(\phi + \Phi) v_0 + H_1(v_0, \phi + \Phi) + \mathcal{G}_1(0, \phi) \right],
\]
(4.52)
\[
u_1^+ = P_1^{-1}(E_{n_1} + v_1)^{-1}P_1 \left[ -(\text{Id} - \Gamma_{K^+})u_1 + \partial_\phi \mathcal{G}_1(0, \phi) - P_1^{-1}\partial_\phi \nu_0 (E_{n_2} + \partial_\phi \Phi)^{-1}P_2 Q \right],
\]
(4.53)
\[
Q = \partial_\phi H_2(v_0, \phi + \Phi)(E_{n_1} + v_1) + \partial_\phi \mathcal{G}_2(0, \phi),
\]
\[
\begin{pmatrix}
H_1^+ \\
H_2^+
\end{pmatrix} = \mathcal{D}^+(\text{Id} - L) \begin{pmatrix}
H_1 \circ T^+ + \mathcal{G}_1 \\
H_2 \circ T^+ + \mathcal{G}_2
\end{pmatrix}
+ \begin{pmatrix}
-P_1^{-1}(E_{n_1} + v_1)^{-1}(\partial_\phi \nu_0 (E_{n_2} + \partial_\phi \Phi)^{-1}P_2 Q \rho \\
-P_2^{-1}(E_{n_2} + \partial_\phi \Phi)^{-1}P_2 Q \rho
\end{pmatrix},
\]
or in another form,
\[
H_2^+ = P_2^{-1}(E_{n_2} + \partial_\phi \Phi)^{-1}P_2 \left[ H_2 \circ T^+(\rho, \phi) + \mathcal{G}_2(\rho, \phi) - H_2 \circ T^+(0, \phi) - \mathcal{G}_2(0, \phi) \right],
\]
(4.54)
\[
H_1^+ = P_1^{-1}(E_{n_1} + v_1)^{-1}P_1 \left[ (\text{Id} - L)(H_1 \circ T^+(\rho, \phi) + \mathcal{G}_1(\rho, \phi)) + \partial_\phi \mathcal{G}_1(0, \phi) \right]
+ P_1^{-1}\partial_\phi \nu_0 (E_{n_2} + \partial_\phi \Phi)^{-1}P_2 Q \rho - P_1^{-1}(\partial_\phi \nu_0 + \partial_\phi \nu_1 \phi)P_2 H_2^+.
\]
(4.55)
By (4.7) and (4.10), we have
\[
|H_i|_{r, \alpha; s} \leq 1 \sup \{C_0 M \varepsilon_0 + |H_0^i|_{d_4, \alpha; \Pi} < C_0 C_0 M \varepsilon_0, \quad i = 1, 2,
\]
(4.56)
where $\tilde{C}_0 = 2 + \sum_{j \geq 1} \chi_j < \infty$.

Now, we proceed to prove (v.2) for $\nu + 1$ and first estimate the three terms $u_0^+, u_1^+$ and $w^+$. We will use $C_\alpha$ to denote a constant only depending on $\alpha$. By using the Taylor expansions of $H_1$ and $H_2$, Lemma 6.8 (ii), (4.29)-(4.31), (4.10), (4.56), (4.47) and the Cauchy inequality we find

$$|H_1(v_0, \phi + \Phi)|_{r_+ \alpha; s^+} \leq C_\alpha \tilde{C}_0 C_0 M \varepsilon_0 r_+^{-(\alpha + 1)(\nu + 2)},$$

$$|H_2(v_0, \phi + \Phi)|_{r_+ \alpha; s^+} \leq C_\alpha \tilde{C}_0 C_0 M \varepsilon_0 r_+^{-(\alpha + 1)},$$

and

$$|Q|_{r_+ \alpha; s^+} \leq C_\alpha \tilde{C}_0 C_0 M \varepsilon_0 r_+^{-1}$$

for sufficiently small $\varepsilon_0$. From the Cauchy inequality, Condition (v.2), Lemma 6.8 (i) and (4.31), it follows

$$|u_0(\phi + \Phi) - u_0(\phi)|_{r_+ \alpha; s^+} \leq C_\alpha r_+^{-1} |u_0|_{r_+ \alpha; s} |\Phi|_{r_+ \alpha; s^+} \leq C_\alpha M \varepsilon_0 r_+^{l-\alpha},$$

$$|u_1(\phi + \Phi) - u_1(\phi)|_{r_+ \alpha; s^+} \leq C_\alpha r_+^{-1} |u_1|_{r_+ \alpha; s} |\Phi|_{r_+ \alpha; s^+} \leq C_\alpha M \varepsilon_0 r_+^{l-\alpha},$$

$$|w(\phi + \Phi) - w(\phi)|_{r_+ \alpha; s^+} \leq C_\alpha r_+^{-1} |w|_{r_+ \alpha; s} |\Phi|_{r_+ \alpha; s^+} \leq C_\alpha M \varepsilon_0 r_+^{l-\alpha}.$$  

Lemma 6.7, Condition (v.2) and the definition of $K_+$ imply

$$|(\mathrm{Id} - \Gamma_{K_+}) u_0|_{r_+ \alpha; s^+} \leq r_+ C_\alpha M \varepsilon_0 r_+^{l-\alpha},$$

$$|(\mathrm{Id} - \Gamma_{K_+}) u_1|_{r_+ \alpha; s^+} \leq C_\alpha M \varepsilon_0 r_+^{l-\alpha},$$

and (4.30), (4.31), Condition (v.2) and (4.49),

$$|P_1^{-1} v_1 P_1 \tilde{A}|_{r_+ \alpha; s^+} \leq C_\alpha |P_1^{-1} v_1 P_1|_{r_+ \alpha; s^+} |\tilde{A}|_{r_+ \alpha; s^+} \leq C_\alpha M \varepsilon_0 r_+^{l-\alpha},$$

$$|P_2^{-1} \partial_\phi \Phi P_2 \tilde{w}|_{r_+ \alpha; s^+} \leq C_\alpha |P_2^{-1} \partial_\phi \Phi P_2|_{r_+ \alpha; s^+} |w|_{r_+ \alpha; s^+} \leq C_\alpha M \varepsilon_0 r_+^{l-\alpha},$$

Combining the estimates (4.58), (4.63), (4.65), (4.67), (4.47) and (4.40) for $w^+$, we have

$$|w^+|_{r_+ \alpha; s^+} \leq C_\alpha M \varepsilon_0 r_+^{l-\alpha}$$

by (4.51) and choosing small $\varepsilon_0$. By the estimates (4.29), (4.30), (4.59), (4.40), (4.68) and Condition (v.2), we also have

$$|P_1^{-1} \partial_\phi v_0 P_2 (\tilde{w} + w^+)|_{r_+ \alpha; s^+} \leq C_\alpha |P_1^{-1} \partial_\phi v_0 P_2|_{r_+ \alpha; s^+} |\tilde{w} + w^+|_{r_+ \alpha; s^+} \leq C_\alpha M \varepsilon_0 r_+^{l-\alpha},$$

and

$$|P_1^{-1} \partial_\phi v_0 (E_{n_1} + \partial_\phi \Phi)^{-1} P_2 Q|_{r_+ \alpha; s^+} \leq C_\alpha |P_1^{-1} \partial_\phi v_0 P_2|_{r_+ \alpha; s^+} |Q|_{r_+ \alpha; s^+} \leq C_\alpha M \varepsilon_0 r_+^{l-\alpha}.$$
From Lemma 6.8 (ii), Condition (v.2), (4.29) and (4.30), it follows

\[ |u_1(\phi + \Phi)v_0|_{r_+^-,\alpha,s+} \leq C_\alpha |u_1(\phi + \Phi)|_{r_+^-,\alpha,s+} |v_0|_{r_+^-,\alpha,s+} \ll C_0 M \varepsilon_0 r_+^{-\alpha}, \quad (4.72) \]

\[ |u_1(\phi + \Phi)v_1|_{r_+^-,\alpha,s+} \leq C_\alpha |u_1(\phi + \Phi)|_{r_+^-,\alpha,s+} |v_1|_{r_+^-,\alpha,s+} \ll C_0 M \varepsilon_0 r_+^{-\alpha}. \quad (4.73) \]

On account of the estimates (4.57), (4.60), (4.61), (4.62), (4.64), (4.69)-(4.73) for \( u_0^+ \) and \( u_1^+ \), (4.40) and (4.47)-(4.48), and the expressions (4.52) and (4.53), we can choose \( \varepsilon_0 \) so small that our estimates yield

\[ |u_0^+|_{r_+^-,\alpha,s+} \leq 4C_0 M \varepsilon_0 r_+^{-\alpha}, \quad |u_1^+|_{r_+^-,\alpha,s+} \leq C_0 M \varepsilon_0 r_+^{-\alpha}. \]

To turn to the estimates of \( H_1^+ \) and \( H_2^+ \), by (4.54) and (4.55) we have

\[ H_1^+ - H - 1 = P_1^{-1}(E_{n_1} + v_1) - 1 P_1 \left[ H_1 \circ T^+(\rho, \phi) - H_1(\rho, \phi) - \partial_1 H_1(v_0, \phi + \Phi)(E_{n_1} + v_1) \rho \right. \]

\[ - H_1(v_0, \phi + \Phi) + \mathcal{G}_1(0, \phi) - \partial_\phi \mathcal{G}_1(0, \phi) \rho - P_1^{-1} v_1 P_1 H_1 \]

\[ + P_1^{-1} \partial_\phi v_0(E_{n_2} + \partial_\phi \Phi)^{-1} P_2 Q \rho - P_1^{-1} (\partial_\rho v_0 + \partial_\rho v_1 \rho) P_2 H_2^+. \]

After a short calculation, we find

\[ |H_1 \circ T^+ - H_1|_{\nu_+,\alpha;s+} \leq C_\alpha r_+^{-\alpha} |H_1|_{\nu_+,\alpha;s} |T^+ - \text{Id}|_{\nu_+,\alpha;s+} \ll \chi C_0 M \varepsilon_0 \]

and

\[ |H_2 \circ T^+ - H_2|_{\nu_+,\alpha;s+} \ll \chi C_0 M \varepsilon_0 \]

by Lemma 6.8 (i), the Cauchy inequality, (4.32) and (4.56). It implies

\[ |\mathcal{G}_1(\rho, \phi) - \mathcal{G}_1(0, \phi) - \partial_\rho \mathcal{G}_1(0, \phi) \rho|_{\nu_+,\alpha;s+} \ll \chi C_0 M \varepsilon_0 \]

and

\[ |\mathcal{G}_2(\rho, \phi) - \mathcal{G}_2(0, \phi)|_{\nu_+,\alpha;s+} \ll \chi C_0 M \varepsilon_0 \]

by (4.47),

\[ |P_1^{-1} v_1 P_1 H_1|_{\nu_+,\alpha;s+} \leq C_\alpha |P_1^{-1} v_1 P_1|_{r_+^-,\alpha,s+} |H_1|_{\nu_+,\alpha;s+} \ll \chi C_0 M \varepsilon_0 \]

and

\[ |P_2^{-1} \partial_\phi \Phi P_2 H_2|_{\nu_+,\alpha;s+} \ll \chi C_0 M \varepsilon_0 \]

by (4.30), (4.31) and (4.56), and

\[ |P_2^{-1} \partial_\phi v_0(E_{n_2} + \partial_\phi \Phi)^{-1} P_2 Q \rho|_{\nu_+,\alpha;s+} \leq C_\alpha |P_2^{-1} \partial_\phi v_0 P_2|_{r_+^-,\alpha,s+} |Q|_{r_+^-,\alpha,s+} \ll \chi C_0 M \varepsilon_0 \]

by (4.40), (4.29) and (4.59). The above estimates, (4.40), (4.58) and the expression (4.75) yield

\[ |H_2^+ - H_2|_{\nu_+,\alpha;s+} \leq \chi C_0 M \varepsilon_0, \quad |H_2^+|_{\nu_+,\alpha;s+} \leq (X_+ + r_0^{-1}) C_0 M \varepsilon_0 < \tilde{C}_0 C_0 M \varepsilon_0, \]

which, together with (4.29) and (4.30), implies

\[ |P_1^{-1} (\partial_\phi v_0 + \partial_\rho v_1 \rho) P_2 H_2^+|_{\nu_+,\alpha;s+} \ll \chi C_0 M \varepsilon_0. \]

Thus, the above estimates, (4.40), (4.57)-(4.60) and the expression (4.74) also yield

\[ |H_1^+ - H_1|_{\nu_+,\alpha;s+} \leq \chi C_0 M \varepsilon_0, \quad |H_1^+|_{\nu_+,\alpha;s+} \leq (X_+ + 2r_0^{-2}) C_0 M \varepsilon_0 < \tilde{C}_0 C_0 M \varepsilon_0. \]
Obviously,
\[ H_1^+ = \varepsilon_0 O_{V, \alpha; s} (\rho^2), \quad H_2^+ = \varepsilon_0 O_{V, \alpha; s} (\rho). \]
This completes the proof of the Iteration Lemma.

5. **Proof of Theorems 2.3 and 2.4.** By (2.7), (H1), (H2) and the Whitney extension theorem (see Lemma 6.3), still using \( \omega^\nu \) and \( \Lambda^\nu \) to denote their extensions, we have
\[ \| \Lambda^\nu \|_{\alpha; \Pi} \leq 2c_1 \varepsilon^{\rho_5} \leq 2c_1 \varepsilon^{\rho_5}, \]
and also
\[ \max_{0 \leq \mu \leq \alpha} \| D^\mu (b, \omega^\nu (\xi, \varepsilon)) \| \geq \frac{c_2}{2} \varepsilon^{\rho_5} \quad \text{in Case } n_{22} = 0, \]
\[ \max_{1 \leq \mu \leq \alpha} \| D^\mu (b, \omega^\nu (\xi, \varepsilon)) \| \geq \frac{c_2}{2} \varepsilon^{\rho_5} \quad \text{in Case } n_{22} \neq 0 \]
by (2.11) and (2.12), for sufficiently small \( \varepsilon \), all \( \xi \in \Pi, b \in S_{n_{21}, \nu}, \nu = 0, 1, 2, \ldots \).

Set
\[ f^\nu_{km} (\xi) = \langle k, \omega^\nu (\xi, \varepsilon) \rangle + \langle m, \Im \Lambda^\nu (\xi, \varepsilon) \rangle \]
for \( 0 \neq k \in \mathbb{Z}^{n_2}, K_\nu < |k|_2 \leq K_{\nu+1}; m \in \mathfrak{m} \) (defined in Theorem 2.2), \( \nu = 0, 1, 2, \ldots \).
Here \( \Im \Lambda^\nu \) is the imaginary part of \( \Lambda^\nu \). Then
\[ \mathcal{R}^\nu_{km} (\gamma) \subset \{ \xi \in \Pi : |f^\nu_{km} (\xi) | < \gamma \varepsilon^{\rho_5} |k|_2^{-s} \}, \quad \nu = 1, 2, \ldots, \]
where \( \Pi_0 \) is the closed subset of \( \Pi \) defined in Theorem 2.2.

**Lemma 5.1.** If \( |k|_2 \geq \frac{16}{c_2} c_1 |m|_1 n_3^{-1/2} \), \( 0 \neq k \in \mathbb{Z}^{n_2} \), then
\[ \text{meas} \mathcal{R}^\nu_{km} (\gamma) \leq c_5 (\text{diam} \Pi_0)^{n_3-1} (\gamma |k|_2^{-s-1})^2 \]
for some positive constant \( c_5 \), where \( \text{diam} \Pi_0 \) represents the diameter of \( \Pi_0 \).

**Proof.** We only give the proof for the case \( n_{22} \neq 0 \), the proof for the case \( n_{22} = 0 \) is analogous and is omitted.

Due to the continuity of the derivatives and the compactness of \( \Pi_0 \) and \( S_{n_{21}, 1} \), the non-degenerate condition (5.1) implies that there exist finite covers \( \{ \Pi_i \}_{i=1}^{I_0} \) and \( \{ S^j_i \}_{i=1}^{I_0} \) of \( \Pi_0 \) and \( S_{n_{21}, 1} \), respectively, and \( \mu_{ij} : 1 \leq \mu_{ij} \leq \alpha, i = 1, \ldots, I_0; j = 1, \ldots, J_0, \Pi_i \) is chosen to be convex, such that
\[ \| D^{\mu_{ij}} (b, \omega^\nu (\xi, \varepsilon)) \| \geq \frac{c_2}{4} \varepsilon^{\rho_5} \quad \text{for all } \xi \in \Pi_i, b \in S^j. \]
Hence, for \( 0 \neq k \in \mathbb{Z}^{n_2}, |k|_2 \in S^j, \) we have
\[ \| D^{\mu_{ijk}} f^\nu_{km} (\xi) \| \geq |k|_2 \| D^{\mu_{ijk}} \left( \frac{k}{|k|_2}, \omega^\nu (\xi, \varepsilon) \right) \| \geq |k|_2 \| -2c_1 |m|_1 \varepsilon^{\rho_5} n_3^{-1} \| \geq \frac{c_2}{2} \varepsilon^{\rho_5} |k|_2 \]
for all \( \xi \in \Pi_i, i = 1, \ldots, I_0, \) admitted \( m \) and \( \nu \) if \( |k|_2 \geq \frac{16}{c_2} c_1 |m|_1 n_3^{-1/2} \).

Now we estimate the measure of \( \mathcal{R}^\nu_{km} (\gamma) \cap \Pi_i \). It follows by (5.3) and the definition of the norm (see Lemma 6.5) that there is a vector \( a \in S_{n_{31}} \) such that
\[ \| D^{\mu_{ijk}} f^\nu_{km} (\xi) a^{\otimes \mu_{ijk}} \| \geq \frac{c_2}{8} \varepsilon^{\rho_5} |k|_2 \quad \text{for all } \xi \in \Pi_i. \]
Write \( \xi = at + \zeta \) with \( t \in \mathbb{R}, \zeta \in a^\perp \) and let \( f(t) = f^\nu_{km} (at + \zeta), \ I_\zeta = \{ t \in \mathbb{R} : at + \zeta \in \Pi \} \).
The inequality (5.4) means
\[ \left| \frac{d^{\mu_{ijk}} f(t)}{d^{\mu_{ijk}}} \right| \geq \frac{c_2}{8} \varepsilon^{\rho_5} |k|_2 \quad \text{for all } t \in I_\zeta. \]

By Fubini’s theorem and Lemma 6.4, it implies
\[ \text{meas}(R^{\nu}_{km}(\gamma) \cap \Pi^t) \leq c_6 (\text{diam}\Pi_0)^{n_3-1} (\gamma |k|^t_{2^{-1}})^{\frac{1}{2}}, \]
where \( c_6 = 4 \max \{1, (4\alpha/\epsilon_2)^{1/\alpha} \} \). Therefore
\[ \text{meas} R^{\nu}_{km}(\gamma) \leq \sum_{i=1}^{i_0} \text{meas} (R^{\nu}_{km}(\gamma) \cap \Pi^t) \leq i_0 c_6 (\text{diam}\Pi_0)^{n_3-1} (\gamma |k|^t_{2^{-1}})^{\frac{1}{2}}. \]
The estimate (5.2) is proved by setting \( c_5 = i_0 c_6 \).

Now, let
\[ \mathcal{R} = \{(k, m) \in \mathbb{Z}^{n_2} \times \mathbb{Z}^{n_1} : 0 < |k|_2 < K^*, 0 \neq m \in \mathbb{m} \}. \]
By an analogous proof to Lemma 5.1, we also have

**Lemma 5.2.** If \((k, m) \in \mathcal{R} \) and the condition \((ii)^j \) in Theorem 2.4 holds, then there is a constant \( c_7 > 0 \) such that
\[ \text{meas} R^{\nu}_{km}(\gamma) \leq c_7 (\text{diam}\Pi_0)^{n_3-1} (\gamma |k|^t_{2^{-1}})^{\frac{1}{2}}. \] (5.5)

**Remark 5.1.** If \( q_3 > q_5 \geq 0 \), then without the conditions \((ii)\) and \((ii)^j\), we can obtain
\[ \text{meas} R^{\nu}_{km}(\gamma) \leq c_5 (\text{diam}\Pi_0)^{n_3-1} (2\gamma |k|^t_{2^{-1}})^{\frac{1}{2}} \text{ for all } (k, m) \in \mathcal{R}. \]

In fact, for sufficiently small \( \gamma \) (equivalently, sufficiently small \( \epsilon \)), we have
\[ R^{\nu}_{km}(\gamma) \subseteq \{ \xi \in \Pi_0 : |\langle k, \omega^\nu(\xi, \epsilon) \rangle| < 2\gamma \epsilon^m |k|^t_{2^{-1}} \} \]
for all \((k, m) \in \mathcal{R}\). From Lemma 5.1 with \( |m|_1 = 0 \), it follows
\[ \text{meas} R^{\nu}_{km}(\gamma) \leq \text{meas} \{ \xi \in \Pi_0 : |\langle k, \omega^\nu(\xi, \epsilon) \rangle| < 2\gamma \epsilon^m |k|^t_{2^{-1}} \} \leq c_5 (\text{diam}\Pi_0)^{n_3-1} (2\gamma |k|^t_{2^{-1}})^{\frac{1}{2}}. \]

**Proof of Theorem 2.4.** By Lemmas 5.1 and 5.2, we obtain
\[
\begin{align*}
\text{meas}(\Pi_0 \setminus \Pi^t) & \leq \sum_{\nu=1}^{\infty} \sum_{K_{\nu-1} < |k|_2 \leq K_{\nu}} \left( \text{meas} R^{\nu}_{k0} + \sum_{|m|_1 = 1} \text{meas} R^{\nu}_{km} + \sum_{|m|_1 = 2} \text{meas} R^{\nu}_{km} \right) \\
& \leq \sum_{(k, m) \in \mathcal{R}} \text{meas} R^{\nu}_{km} + c_5 (\text{diam}\Pi_0)^{n_3-1} \gamma^{\frac{1}{2}} \left[ \sum_{0 \neq k \in \mathbb{Z}^{n_2}} |k|_{2^{-1}}^{-\frac{\alpha+1}{\alpha}} \right. \\
& \left. + n_1 \sum_{|k|_2 \geq K^*} |k|_{2^{-1}}^{-\frac{\alpha+1}{\alpha}} + n_1 (n_1 - 1) \sum_{|k|_2 \geq K^*} |k|_{2^{-1}}^{-\frac{\alpha+1}{\alpha}} \right] \\
& \leq c_8 \gamma^{\frac{1}{2}},
\end{align*}
\]
where \( c_8 \) is a positive constant depending on \( n_1, n_2, \text{diam}\Pi, \omega_1 \) and \( \Lambda_i (i = 1, 2) \), and \( \text{meas}(\Pi_0 \setminus \Pi_0) = O(\gamma) \). This proves Theorem 2.4.

**Proof of Theorem 2.3.** From the proof of Theorem 2.4, it is easy to see
\[
\text{meas} \left( \bigcup_{\nu=1}^{\infty} \bigcup_{(k, m) \in \mathcal{R}} R^{\nu}_{km}(\gamma) \right) \leq c_9 \gamma^{\frac{1}{2}} \to 0 \text{ as } \gamma \to 0.
\]
On the other hand, there is a \( \nu_0 \) such that \( K_{\nu_0} \leq K^* \), hence
\[
\bigcup_{\nu=1}^{\infty} \bigcup_{(k,m) \in \mathbb{R}} \mathcal{N}_K^\nu(\gamma) \subset \bigcup_{\nu=1}^{\nu_0} \bigcup_{(k,m) \in \mathbb{R}} \mathcal{N}_K^\nu(\gamma).
\]
By the condition (ii) in Theorem 2.3 and the boundedness of \( \Pi \), we have \( \text{meas} \mathcal{N}_K^\nu(\gamma) \to 0 \) as \( \gamma \to 0 \). Since \( \mathbb{R} \) is finite, we also have
\[
\text{meas} \left( \bigcup_{\nu=1}^{\nu_0} \bigcup_{(k,m) \in \mathbb{R}} \mathcal{N}_K^\nu(\gamma) \right) \to 0 \quad \text{as} \quad \gamma \to 0.
\]

Appendix.

A.1. Whitney extension theorem. Let \( \bar{\Omega} \subset \mathbb{R}^n \) be a closed set, \( p \) be a non-negative integer, \( p < l \leq p + 1 \). \( C^{l}_W(\bar{\Omega}) \) is the class of all collections \( f = \{ f^{(k)} \}_{|k| \leq p} \) of functions defined on \( \bar{\Omega} \) which satisfy, for some finite \( M \),
\[
|f^{(k)}(x)| \leq M, \quad |f^{(k)}(x) - P_k(x,y)| \leq M|x - y|^{l-|k|},
\]
for all \( x, y \in \Omega \) and \( |k| \leq p \), where
\[
P_k(x,y) = \sum_{|k+j| \leq p} \frac{1}{j!} f^{(k+j)}(y)(x-y)^j
\]
is the analogue of the \( k \)-th Taylor polynomial. \( f \) is called \( C^{l}_W \) Whitney in \( \bar{\Omega} \) with Whitney derivatives \( D^k f = f^{(k)} \) for \( |k| \leq p \). Define a norm
\[
||f||_{C^{l}_W(\bar{\Omega})} = \inf M
\]
is the smallest \( M \) for which both inequalities in (6.1) hold. Then \( C^{l}_W(\bar{\Omega}) \) with the norm is a Banach space.

The following extension theorem indicates that a Whitney differentiable function has an extension to \( \mathbb{R}^n \) which is differentiable in the standard sense.

**Lemma 6.3.** (Whitney extension theorem, [39, 37, 27]) Let \( \bar{\Omega} \) be a closed set in \( \mathbb{R}^n \), \( p \in \mathbb{Z}_+ \) and \( p < l \leq p + 1 \). Then there exists a linear extension operator
\[
\mathcal{E} : C^{l}_W(\bar{\Omega}) \to C^{l}(\mathbb{R}^n), \quad f = \{ f^{(k)} \}_{|k| \leq p} \to F = \mathcal{E}f
\]
such that
\[
D^k F \big|_{\bar{\Omega}} = f^{(k)}, \quad |k| \leq p
\]
and
\[
||F||_{C^l(\mathbb{R}^n)} \leq C ||f||_{C^{l}_W(\bar{\Omega})},
\]
where the constant \( C \) depends only on \( l \) and the dimension \( n \), but not on \( \bar{\Omega} \). Moreover, if \( \bar{\Omega} = \bar{\Omega}_1 \times \mathbb{T}^{n_2} \subset \mathbb{R}^{n_1} \times \mathbb{T}^{n_2} \), then the extension can be chosen to be defined on \( \mathbb{R}^{n_1} \times \mathbb{T}^{n_2} \), so that the periodicity is preserved.

A.2. Measure estimate lemmas.

**Lemma 6.4.** [34] Let \( f : [a, b] \to \mathbb{R} \) with \( a < b \) be an \( \alpha \)-times continuously differentiable function satisfying
\[
\left| \frac{d^\alpha f(x)}{dx^\alpha} \right| \geq c, \quad x \in [a, b]
\]
for some \( \alpha \in \mathbb{N} \) and a constant \( c > 0 \). Then we have the measure estimate
Remark A.2. From the proof of Lemma 6.5, it is easy to see that if the condition (6.2) is replaced by
\[ \text{rank } \left( \frac{\partial^\beta f(\xi)}{\partial \xi^\beta} : 1 \leq |\beta|_1 \leq \alpha \right) = q \quad \text{and} \quad p + p^2 + \cdots + p^\alpha \geq q, \]
then we also have
\[ \max_{1 \leq \mu \leq \alpha} ||D^\mu \langle b, f(\xi) \rangle|| \geq c \quad \text{for all } b \in S_{q,1}, \xi \in \Pi. \]
A.3. Properties of analytic smoothing operator. Let $l > 0, m \in \mathbb{N}$ and $C^j(\mathbb{R}^m)$ be the Hölder space defined in Definition 2.1 without parameter variables, $u_0 \in C^\infty_0(\mathbb{R})$ be an even function, vanishing outside the interval $[-1, 1]$ and identically equal to 1 in a neighbourhood of 0, $u(x) = u_0(|x|^2)$ for $x \in \mathbb{R}^m$ and

$$\tilde{u}(z) = \int_{\mathbb{R}^m} u(x)e^{\sqrt{-1}(z,x)}dx \quad \text{for } z \in \mathbb{C}^m,$$

$$f_r(x) \equiv (S_r f)(x) \equiv r^{-m} \int_{\mathbb{R}^m} \tilde{u}((x-y)/r)f(y)dy$$

for $x \in \mathbb{C}^m$ and $r \in (0, 1)$.

**Lemma 6.6.** The following assertions are valid

(a) $\int_{\mathbb{R}^m} \tilde{u}(x)dx = u(0) = 1$;

(b) $(S_r f)(x) = \int_{\mathbb{R}^m} \tilde{u}(y)f(x-ry)dy$ for $x \in \mathbb{R}^m$;

(c) $\int_{\mathbb{R}^m} x^k \tilde{u}(x)dx = 0$ for $0 \neq k \in \mathbb{Z}^m_+$;

(d) for any $p \in \mathbb{N}$, there is a constant $C_p > 0$ such that

$$|D^k \tilde{u}(z)| \leq \frac{C_p}{(1 + |z|^2)^{|k|}}e^{C_2|\text{Im} z|}$$

for all $|k| \leq p, k \in \mathbb{Z}^m_+$,

where $D^k = D^k_1 \circ D^k_2 \circ \cdots \circ D^k_m$, and $D^k_j = \partial^{|k|_j} / \partial x^{|k|_j}$;

(e) if $P$ is a polynomial, then $(S_r P)(x) = P(x)$;

(f) there exists a constant $C_1 > 0$ such that

$$|D^k f_r(x) - \sum_{|\beta| \leq |k|} D^{k+\beta} f(\text{Re} x) \frac{(\sqrt{-1} \text{Im} x)^\beta}{\beta!}| \leq C_1 r^{l-|k|_1} ||f||_{l; \mathbb{R}^m}, ||\text{Im} x|| \leq r \leq 1$$

for all $k \in \mathbb{Z}^m_+$ with $|k|_1 \leq l$. In particular, for $x \in \mathbb{R}^m$ and $p \in \mathbb{Z}^+$,

$$||f_r| - f||_{p; \mathbb{R}^m} \leq C_1 r^{l-p} ||f||_{l; \mathbb{R}^m}, \quad p \leq l$$

for a suitable constant $C_1$, depending on $l, p$ and $m$.

**Proof.** The definitions of $\tilde{u}$ and $S_r$ imply (a) and (b), respectively. Noting that $\tilde{u}$ is a Schwartz function (see (d)), and the Fourier transformation and differentiation can be exchanged, we have

$$\int_{\mathbb{R}^m} x^k \tilde{u}(x)dx = \left((\sqrt{-1})^{k_1} D^k_y \int_{\mathbb{R}^m} \tilde{u}(x)e^{-\sqrt{-1}(y,x)}dx\right)_{y=0} = \left((\sqrt{-1})^{k_1} D^k u(y)\right)_{y=0} = 0,$$

which verifies (c). See Lemma 9, Proposition 8 and Remark 15 (i) in [11] for (d), (e) and (f), respectively, also see the proof of Lemma 2.1 in Part I of [40] for (e).

**Lemma 6.7.** Let $K$ be a positive integer and $f$ be a bounded and analytic function in the strip $\{x : ||\text{Im} x|| < r\}$ of $\mathbb{T}^n$, $f(x) = \sum_{k \in \mathbb{Z}^n} \hat{f}(k)e^{\sqrt{-1}(k,x)}$. Define the truncation operator $\Gamma_K$ as follows

$$\Gamma_K f = \sum_{|k|_2 \leq K} \hat{f}(k)e^{\sqrt{-1}(k,x)}.$$

If $K > (2\rho)^{-1}$, then we have

$$||\text{Id} - \Gamma_K||_{r^{-2}\rho} \leq C(n)||f||_r \rho^{-n}e^{-\rho K}, \quad 0 < 2\rho \leq r,$$

where $C(n) = (6n!)^{n}e^{-n}$. 




Proof. Set $\sigma = 2\rho$. Based on the fact that the number of all $k$ with $|k|_1 = m$ is bounded by $2mn^{n-1}$, we have
\[
|(\text{Id} - \Gamma_K)f|_{r-\sigma} \leq \sum_{|k|_2 > K} |f| |e^{-\sigma|k|_1}| \leq \sum_{|k|_2 > K} |f| |e^{-\sigma|k|_1}| \leq |f| r \sum_{m > K} 2mn^{n-1}e^{-\sigma m}.
\]
(6.3)
Here we use Lemma A.1 in [29]. Since the function $y^{n-1}e^{-\sigma y}$ is monotonically decreasing in the interval $\left[\frac{n-1}{\sigma}, +\infty\right)$ and $K > \sigma^{-1}$, therefore,
\[
\sum_{m > K} m^{n-1}e^{-\sigma m} < \int_{K}^{+\infty} y^{n-1}e^{-\sigma y} dy = \left(\frac{1}{\sigma}K^{n-1} + \frac{n-1}{\sigma^2}K^{n-2} + \cdots + \frac{(n-1)!}{\sigma^n}\right)e^{-\sigma K} < 3(n-1)!K^n e^{-\sigma K}.
\]
Hence, by (6.3) we obtain
\[
|(\text{Id} - \Gamma_K)f|_{r-\sigma} \leq 6n!|f|_r K^n e^{-\sigma K}.
\]
(6.4) Noting that the maximum of the function $y^n e^{-y}$ on the interval $(0, +\infty)$ is $n^n e^{-n}$, (6.4) implies
\[
|(\text{Id} - \Gamma_K)f|_{r-2\rho} \leq 6(n!)n^n e^{-n}|f|_r \rho^n e^{-\rho K}.
\]

Let $\Omega_1$ and $\Omega_2$ be domains in $\mathbb{C}^n$, $\Pi$ be an open set in $\mathbb{R}^m$, $f(x, \xi)$ and $g(x, \xi)$ be analytic in $x \in (\Omega_1 + r)$ and in $x \in \Omega_2$ respectively, and continuously differential up to order $\alpha$ in $\xi \in \Pi$, $g : \Omega_2 \times \Pi \to \Omega_1$, where $r > 0$, $\Omega_1 + r = \{x \in \mathbb{C}^n : \text{dist}(x, \Omega_1) < r\}$.

We introduce the notation for $1 \leq \mu \leq \alpha$,
\[
|Df|_{\mu} \equiv \max_{1 \leq |\beta| \leq \mu} |\partial_\xi^\beta Df|_{\Omega_{1+r,0;\Pi}}, \quad |g|_{\mu} \equiv \max_{1 \leq |\beta| \leq \mu} |\partial_\xi^\beta g|_{\Omega_{2,0;\Pi}},
\]
where $Df$ represents the differential operator with respect to the coordinate variable $x$. Using the Chain Rule on differentiation of a composition of mappings and Cauchy inequality, we easily prove the following lemma.

**Lemma 6.8.** Let $\beta \in \mathbb{Z}_+^m$ and $|\beta|_1 = \mu, 1 \leq \mu \leq \alpha$. Then
\[
\text{(i) } |\partial_\xi^\beta(f \circ g - f)|_{\Omega_{2,0;\Pi}} \leq \left\{ \begin{array}{ll}
|\partial_\xi^\beta f|_{\Omega_{2,0;\Pi}} + |Df|_{\Omega_{1+r,0;\Pi}}|g|_1 & \text{for } \mu = 1 \\
|\partial_\xi^\beta f|_{\Omega_{2,0;\Pi}} + \sum_{j=1}^{\mu-1} C_j |Df|_{\mu-1} |g|_{\mu-1} \\
+ |Df|_{\Omega_{1+r,0;\Pi}}(|g|_{\mu} + \frac{(\mu-1)!}{\mu!} |g|_1^2) & \text{for } 2 \leq \mu \leq \alpha;
\end{array} \right.
\]
\[
\text{(ii) } |\partial_\xi^\beta(f \circ g)|_{\Omega_{2,0;\Pi}} \leq \left( |\partial_\xi^\beta f|_{\Omega_{2,0;\Pi}} + \sum_{j=1}^{\mu} C_j |f|_{\mu-1} |g|_\mu^j \right., \text{ where } C_j (j = 1, \cdots, \mu) \text{ are nonnegative constants only depending on } \beta.
\]

**A.4.** An estimate lemma for small divisors.

**Lemma 6.9.** Assume the frequency vector $\omega = (\omega_1, \cdots, \omega_n)$ satisfies the inequalities
\[
|\langle k, \omega \rangle| \geq \frac{\gamma}{|k|_2} \quad \text{and} \quad |\langle k, \omega \rangle + \lambda| \geq \frac{\gamma}{|k|_2}
\]
(6.5)
for all integer vectors $0 \neq k \in \mathbb{Z}^n$ with $|k|_2 \leq K \leq \infty$, and some constants $\tau > n - 1 \geq 1$, $K > 0$, $\gamma > 0$ and $\lambda \in \mathbb{R}$. Then the following inequalities hold

$$\sum_{0 \neq |k|_2 \leq K} |k|_1^n |\langle k, \omega \rangle|^{-b} e^{-\sigma |k|_1} \leq C \gamma^{-b} \sigma^{-(\tau b + v + 1)} \tag{6.6}$$

and

$$\sum_{0 \neq |k|_2 \leq K} |k|_1^n |\langle k, \omega \rangle + \lambda|^{-b} e^{-\sigma |k|_1} \leq C \gamma^{-b} \sigma^{-(\tau b + v + 1)} \tag{6.7}$$

with

$$C = 15 \tau \sqrt{\tau b + v} 2^{(n+b)-3} \tau^b v^{n-1}(\tau b - n + 1)^{-1} \left( \frac{\tau b + v}{e} \right)^{\tau b + v},$$

where $v \geq 0, b \geq 1$ and $\sigma \in (0, 1)$ are constants.

**Proof.** The proof is based on the fact that only a few of the denominators $|\langle k, \omega \rangle|$ and $|\langle k, \omega \rangle + \lambda|$ are small, which was used by Siegel [36], Arnol’d [3] and Moser [25].

For the sake of completeness, we present the proof for our situation and make the involved constants explicit in estimates. We only prove the inequality (6.7) with $K = \infty$. The proof of (6.6) and the case $K < \infty$ is analogous and is omitted.

Set

$$K(m, j) = \{ k = (k_1, \cdots, k_n) \in \mathbb{Z}^n : |k| = m, \gamma^{-1} 2^j < |\langle k, \omega \rangle| \leq \gamma^{-1} 2^{j+1} \}$$

where $|k| \equiv \max_{1 \leq i \leq n} |k_i|$, and let $K(m, j)^\#$ denote the number of points in $K(m, j)$. Then we have

$$K(m, j)^\# \leq (2n)^n (2m)^{-\frac{1}{2}} (n-1)(j-1). \tag{6.8}$$

In fact, if $k, k' \in K(m, j)$ are different points, then

$$\gamma |k - k'|_2^{-\tau} \leq |\langle k, \omega \rangle| \leq |\langle k, \omega \rangle + \lambda| + |\langle k', \omega \rangle + \lambda| < \gamma 2^{1-j},$$

which implies

$$|k - k'| \geq n^{-\frac{1}{2}} |k - k'|_2 > n^{-1} 2^{j-1} = 2 \rho_j.$$  

Noting $|k - k'| \leq 2m$ we get $\rho_j \leq m$. If we encircle every point $k \in K(m, j)$ by a cube $\mathcal{C}_k : |x - k| \leq \rho_j$, then these cubes are mutually disjoint. The intersections of these cubes $\mathcal{C}_k$ with the curved surface $|x| = m$ are disjoint $n-1$ dimensional sets with $n-1$ dimensional volume $\geq \rho_j^{n-1}$. As the $n-1$ dimensional volume of the curved surface $|x| = m$ is $2n(2m)^{n-1}$, we obtain

$$K(m, j)^\# \leq \frac{2n(2m)^{n-1}}{\rho_j^{n-1}},$$

which verifies the inequality (6.8). Thus we have

$$\sum_{K(m, j)} |\langle k, \omega \rangle + \lambda|^{-b} \leq \gamma^{-b} \sigma^{j+1} K(m, j)^\#$$

$$\leq 2^{2(n+b)-1} n^m m^{n-1} \gamma^{-b} \sigma^{(b - \frac{n-1}{2}) (j-1)} \equiv C_1 2^{(b - \frac{n}{2}) (j-1)},$$

Let $j^*$ be the greatest occurring $j$ for which $K(m, j) \neq \emptyset$. Then the facts that

$$\gamma^{-1} 2^{j^*} < |\langle k, \omega \rangle + \lambda|^{-1} \leq \gamma^{-1} |k|_2 \leq \gamma^{-1} (nm)^{\tau}$$

and

$$\{ k \in \mathbb{Z}^n : |k| = m \}^\# = (2m + 1)^n - (2m - 1)^n < 2n(4m)^{n-1}$$
implies
\[ \sum_{|k|=m} |\langle k, \omega \rangle + \lambda|^{-b} \leq \sum_{|k|=m, |\langle k, \omega \rangle + \lambda|^{-1} \leq 2\gamma^{-1}} |\langle k, \omega \rangle + \lambda|^{-b} + \sum_{j=1}^{j^*} \sum_{K(m,j)} |\langle k, \omega \rangle + \lambda|^{-b} \leq \sum_{|k|=m} |\langle k, \omega \rangle + \lambda|^{-b} \]
\[ \leq n 2^{2n+b-1} m^{n-1} \gamma^{-b} + C_1 \sum_{j=1}^{j^*} 2(\frac{n}{e}\tau-1)(j-1) \]
\[ \leq \tau 2^{(n+b)-1} \gamma^{-b} \leq \gamma^{-b} m^{\tau b} \equiv C_2 m^{\tau b}. \]

Therefore,
\[ \sum_{0 \neq k \in \mathbb{Z}^n} |k|_1^v |\langle k, \omega \rangle + \lambda|^{-b} e^{-\sigma|k|_1} \leq \sum_{m=1}^{\infty} \sum_{|k|=m} (nm)^v |\langle k, \omega \rangle + \lambda|^{-b} e^{-\sigma m} \]
\[ \leq C_2 n^v \sum_{m=1}^{\infty} m^{\tau b+v} e^{-\sigma m}. \quad (6.9) \]

Noting that the function \( g(x) = x^{\tau b+v} e^{-\sigma x} \) on the interval \([1, \infty)\) gets its maximum at \( x_0 = \frac{\tau b+v}{\sigma} \), moreover is strictly increasing and decreasing on \([1, x_0)\) and \((x_0, \infty)\), respectively. Denote the integer part of \( \frac{\tau b+v}{\sigma} \) by \( m_0 \). Then \( m_0 \geq 1 \) and
\[ \sum_{m=1}^{\infty} m^{\tau b+v} e^{-\sigma m} \leq \int_1^{m_0} x^{\tau b+v} e^{-\sigma x} dx + g(\frac{\tau b+v}{\sigma}) + \int_{m_0}^{\infty} x^{\tau b+v} e^{-\sigma x} dx \]
\[ \leq g(\frac{\tau b+v}{\sigma}) + \sigma^{-(\tau b+v+1)} \int_0^{\infty} y^{\tau b+v} e^{-y} dy \]
\[ = \left( \frac{\tau b+v}{e\sigma} \right)^{\tau b+v} + \sigma^{-(\tau b+v+1)} \Gamma(\tau b+v+1). \quad (6.10) \]

By the Stirling formula of the gamma function, we have
\[ \Gamma(\tau b+v+1) < \frac{11}{4} \sqrt{\tau b+v} \left( \frac{\tau b+v}{e} \right)^{\tau b+v}. \quad (6.11) \]
Combining (6.9)-(6.11), we obtain the estimate (6.7). The proof of the lemma is complete.

**Remark A.3.** From the proof of Lemma 6.9 it is easy to see that if the norm \( |k|_2 \) in the condition (6.5) is replaced by the norm \( |k|_1 \), then the estimates (6.6) and (6.7) are still valid.

**REFERENCES**

[1] J. Albrecht, On the existence of invariant tori in nearly integrable Hamiltonian systems with finitely differentiable perturbations, Regul. Chaotic Dyn., 12 (2007), 281–320.

[2] V. I. Arnol’d, Proof of a theorem by A. N. Kolmogorov on the invariance of quasi periodic motions under small perturbations of the Hamiltonian, Uspehi. Math. Nauk., 18 (1963), 13–40, Russian Math. Survey, 18 (1963), 9–36.

[3] V. I. Arnol’d, Small divisor problems in classical and celestial mechanics, Russian Math. Survey, 18 (1963), 85–191.

[4] D. Bambusi, M. Berti and E. Magistrelli, Degenerate KAM theory for partial differential equations, J. Differential Equations, 250 (2011), 3379–3397.
[5] D. Bambusi and G. Gaeta, Invariant tori for non-conservative perturbations of integrable systems, *NoDEA Nonlinear Differ. Equ. Appl.*, **8** (2001), 99–116.

[6] N. N. Bogoljubov, Ju. A. Mitropolskii and A. M. Samoilenko, *Methods of Accelerated Convergence in Nonlinear Mechanics*, Springer, Berlin, 1976.

[7] H. W. Broer, G. B. Huitema and M. B. Sevryuk, *Quasi-periodic Motions in Families of Dynamical Systems: Order amidst Chaos*, Lecture Notes in Math., Springer, Berlin, 1996.

[8] A. D. Bruno, On conditions for nondegeneracy in Kolmogorov’s theorem, *Soviet Math. Dokl.*, **45** (1992), 221–225.

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

[10] C.-Q. Cheng and S. Wang, The surviving of lower dimensional tori from a resonant torus of Hamiltonian systems, *J. Differential Equations*, **155** (1999), 311–326.

[11] L. Chierchia, KAM Lectures, Dynamical Systems, *Pubbl. Cent. Ric. Mat. Ennio Giorgi, Sc. Norm. Sup.*, Pisa, Part I: (2003), 1–55.

[12] L. Chierchia and G. Pinzari, Properly degenerate KAM theory (following V. I. Arnold), *Discrete. Contin. Dyn. Syst. S*, **3** (2010), 545–578.

[13] L. Chierchia and D. Qian, Moser’s theorem for lower dimensional tori, *J. Differential Equations*, **206** (2004), 55–93.

[14] J. Féjoz, Démonstration du ‘théorème d’Arnold’ sur la stabilité du système planétaire (d’après Herman), *Ergodic Theory Dyn. Syst.*, **24** (2004), 1521–1582.

[15] G. Gentile, Degenerate lower-dimensional tori under the Bryuno condition, *Ergodic Theory Dyn. Syst.*, **27** (2007), 427–457.

[16] G. Gentile and G. Gallavotti, Degenerate elliptic resonances, *Comm. Math. Phys.*, **257** (2005), 319–362.

[17] Y. Han, Y. Li and Y. Yi, Degenerate lower-dimensional tori in Hamiltonian systems, *J. Differential Equations*, **227** (2006), 670–691.

[18] Y. Han, Y. Li and Y. Yi, Invariant tori in Hamiltonian systems with high order proper degeneracy, *Ann. Henri Poincaré*, **10** (2010), 1419–1436.

[19] M. R. Herman, Sur les courbes invariantes par les difféomorphismes de l’anneau, *Vol. I. Astérisque*, **103** (1983), i+221pp.

[20] X. Li, On the persistence of quasi-periodic invariant tori for double Hopf bifurcation of vector fields, *J. Differential Equations*, **260** (2016), 7320–7357.

[21] X. Li and R. de la Llave, Convergence of differentiable functions on closed sets and remarks on the proofs of the “converse approximation lemmas”, *Discrete Contin. Dyn. Syst. S*, **3** (2010), 623–641.

[22] X. Li and X. Yuan, Quasi-periodic solutions for perturbed autonomous delay differential equations, *J. Differential Equations*, **252** (2012), 3752–3796.

[23] Y. Li and Y. Yi, A quasi-periodic Poincaré’s theorem, *Math. Ann.*, **326** (2003), 649–690.

[24] J. Moser, On invariant curves of area-preserving mappings of an annulus, *Nach. Akad. Wiss. Göttingen, II Math. Wiss.*, **1962** (1962), 1–20.

[25] J. Moser, A rapidly convergent iteration method and nonlinear partial differential equations I and II, *Ann. Scuola Norm. Sup. Pisa*, **3** (1966), 265–315.

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

[27] J. Pöschel, Integrability of Hamiltonian systems on Cantor sets, *Commun. Pure Appl. Math.*, **35** (1982), 653–696.

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

[29] J. Pöschel, A lecture on the classical KAM theorem, *Proc. Symp. Pure Math.*, **69** (2001), 707–732.

[30] H. Rüssmann, Kleine Nenner I: Über invariante Kurven differenzierbarer Abbildungen eines Kreisringes, *Nach. Akad. Wiss. Göttingen, II Math. Phys. Kl.*, **1970** (1970), 67–105.

[31] H. Rüssmann, On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus, *Lecture Notes in Phys.*, Springer, Berlin, **38** (1975), 598–624.

[32] H. Rüssmann, On the existence of invariant curves of twist mappings of an annulus, *Lecture Notes in Math.*, Springer, Berlin, **1007** (1983), 677–718.

[33] H. Rüssmann, Nondegeneracy in the perturbation theory of integrable dynamical systems, *London Math. Soc. Lecture Note Ser.*, **134** (1989), 5–18.
[34] H. Rüssmann, Invariant tori in non-degenerate nearly integrable Hamiltonian systems, *Regul. Chaotic Dyn.*, 6 (2001), 119–204.

[35] Z. Shang, A note on the KAM theorem for symplectic mappings, *J. Dyn. Diff. Eqs.*, 12 (2000), 357–383.

[36] C. L. Siegel, *Verlesungen über Himmelsmechanik*, Springer, 1956.

[37] E. M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton Math. Series, No. 30, Princeton University Press, 1970.

[38] F. Wagener, A parameterised version of Moser’s modifying terms theorem, *Discrete Contin. Dyn. Syst. S*, 3 (2010), 719–768.

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

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

Received for publication November 2018.

E-mail address: lixuemei_l@sina.com

E-mail address: zaijiu@amss.ac.cn