Quasiperiodic Poincaré Persistence at High Degeneracy

Weichao Qian\(^a\), Yong Li\(^a,b\), Xue Yang\(^a,b\)

\(^a\)College of Mathematics, Jilin University, P. R. China
\(^b\)School of Mathematics and Statistics &
Center for Mathematics and Interdisciplinary Sciences,
Northeast Normal University, P. R. China

Abstract

For Hamiltonian systems with high-order degenerate perturbation, we study the persistence of resonant invariant tori, where the resonant tori might be elliptic, hyperbolic or mixed types. As a consequence, we prove a quasiperiodic Poincaré theorem at high degeneracy. This answers a long standing conjecture on the persistence of resonant invariant tori in general situations.

Keywords
Hamiltonian systems; high-order degenerate perturbation; KAM theory; resonant invariant tori; quasiperiodic Poincaré theorem.

1 Introduction

This paper concerns the persistence of resonant invariant tori for the following Hamiltonian system

\[
H(\theta, I) = H_0(I) + \varepsilon P(\theta, I, \varepsilon),
\]

where \( \theta \in T^d = \mathbb{R}^d / \mathbb{Z}^d, I \in G \) (\( G \) is a bounded closed region in \( \mathbb{R}^d \)), \( H_0(I) \) and \( P(\theta, I, \varepsilon) = P_0(I, \theta, 0) + \sum_{1 \leq i \in \mathbb{Z}^+} \varepsilon^i P_i(I, \theta, 0) \) are real analytic functions on a complex neighborhood of the bounded closed region \( T^d \times G \) and \( \varepsilon > 0 \) is a small parameter. Here the so-called resonant invariant tori mean the frequency...
\( \omega(I) = \frac{\partial H_0}{\partial I} \) is resonant for some I, i.e., there exists at least one \( k \in \mathbb{Z}^d \setminus \{0\} \) such that \( \langle k, \omega \rangle = 0 \).

The celebrated KAM theory due to Kolmogorov, Arnold and Moser asserts that, if an integrable system, \( H_0(I) \) in \( \mathbb{R}^d \), is nondegenerate, i.e. \( \det \frac{\partial^2 H_0}{\partial I^2} \neq 0 \), then, for the perturbed system \( H(\theta, I) = H_0(I) + \varepsilon P(\theta, I, \varepsilon) \), most of nonresonant invariant tori still survive (11, 15, 24). For some recent developments and applications related to KAM theory, refer to [12, 13, 16, 26, 29, 30, 31]. However, in the presence of resonance, the persistence problem becomes very complicated. Let us do a brief recall. The periodic case can go back to the work of Poincaré in nineteenth century, which does not involve the small divisor problem (28). There has been a long standing conjecture about resonant tori under a convexity assumption on \( H_0 \) (5, 8, 9, 11, 17), as written by Kappeler and Pöschel in [17]:

For \( m = 1 \) in particular, such a torus is foliated into identical closed orbits. Bernstein & Katok (3) showed that in a convex system at least \( d \) of them survive any sufficiently small perturbation, \( \cdots \). For the intermediate cases with \( 1 < m < d - 1 \), only partial results are known \( \cdots \). The long standing conjecture is that at least \( d - m + 1 \), and generically \( 2^{d-m} \), invariant \( m \)-tori always survive in a nondegenerate system \( \cdots \). That is, their number should be equal to the number of critical points of smooth functions on the torus \( T^{d-m} \).

In above description, \( m \) and \( d \) are dimensions of the lower-dimensional invariant tori and the degree of freedom, respectively.

The first breakthrough of the conjecture mentioned above was due to Treschëv (36) for the persistence of hyperbolic resonant tori in 1989, 35 years after the establishment of KAM theory, and such tori are called Treschëv’s tori today. For the persistence of general resonant tori, we refer readers to [8, 20, 37, 38]. In fact, for Hamiltonian system (1.1), when \( P_0(I, \theta, 0) \) in \( \varepsilon P(\theta, I, \varepsilon) = \varepsilon P_0(I, \theta, 0) + \sum_{1 \leq i \in \mathbb{Z}_+} \varepsilon^{i+1} P_i(I, \theta, 0) \) is nondegenerate (we will explain what ‘non-degenerate’ means later), the proof of the conjecture mentioned above has been completed, see [6, 8, 20, 39]. However,

What happens to the conjecture if \( P_0(I, \theta, 0) \) is degenerate?

In the present paper we will touch this essential problem.

In order to state our main result, first, let us introduce some notations. We say that a frequency vector \( \omega = \frac{\partial H_0}{\partial I} \) is nonresonant for some \( I \), if \( \langle k, \omega \rangle \neq 0 \) for any \( k \in \mathbb{Z}^d \setminus \{0\} \). Furthermore, if there is a subgroup \( g \) of \( \mathbb{Z}^d \) such that \( \langle k, \omega \rangle = 0 \) for all \( k \in g \) and \( \langle k, \omega \rangle \neq 0 \) for all \( k \in \mathbb{Z}^d/g \), then \( \omega \) is called multiplicity \( m_0 \) resonant frequency (\( g \)-resonant frequency), where \( g \) is generated by independent \( d \)-dimensional integer vectors \( \tau_1, \ldots, \tau_{m_0} \). For a given subgroup \( g \), the manifold

\[ \tilde{\Lambda}(g, G) = \{ I \in G : \langle k, \omega(I) \rangle = 0, k \in g \} \]
is called $g$–resonant surface. By group theory, there are integer vectors $\tau'_1, \ldots, \tau'_m \in \mathbb{Z}^d$, such that $\mathbb{Z}^d$ is generated by $\tau_1, \ldots, \tau_{m_0}, \tau'_1, \ldots, \tau'_m$, and $\det K_0 = 1$, where $K_0 = (K_*, K')$, $K_* = (\tau'_1, \ldots, \tau'_m)$, $K' = (\tau_1, \ldots, \tau_{m_0})$ are $d \times d$ matrix, $d \times m$ matrix, respectively, and $K_*$ generates the quotient group $\mathbb{Z}^d / g$, while $K'$ generates the group $g$ ([46]). If $\det K' \partial^2 H_0 K' \neq 0$ and $\det \partial^2 H_0 \neq 0$ for $I \in \Lambda(g, G)$, Hamiltonian system (1.1) is called $g$–nondegenerate.

The motion equation of the unperturbed Hamiltonian system $H_0(I)$ in (1.1) is

$$\begin{cases} \dot{\theta} = \omega(I), \\ \dot{I} = 0. \end{cases}$$

Denote $p = (y, v), q = (x, u)$, where $y = (p_1, \ldots, p_m)^T, v = (p_{m+1}, \ldots, p_d)^T$, $x = (q_1, \ldots, q_m)^T, u = (q_{m+1}, \ldots, q_d)^T$. When $\omega(I)$ is $g$–resonant, under the following symplectic transformation

$$\phi_g : (p, q) \to (I, \theta),$$

where $K_0 \theta = q, I - I_0 = K_0 p$, the equation of motion becomes

$$\begin{cases} \dot{x} = K_0^T \omega(I), \\ \dot{u} = 0, \\ \dot{y} = 0, \\ \dot{v} = 0, \end{cases}$$

where $K_0$ and $K_*$ are mentioned as above. (We place the verification that $\phi_g$ is symplectic on Appendix A.) We call such $(y, v, u)$ the relative critical point.

With transformation $\phi_g$, Hamiltonian system (1.1) could be transformed to

$$H(x, y, u, v) = H \circ \phi_g = \tilde{H}_0(y, v) + \varepsilon \tilde{P}(x, y, u, v, \varepsilon),$$

where

$$\tilde{P}(x, y, u, v, \varepsilon) = P((K_0)^{-1} \begin{pmatrix} x \\ u \end{pmatrix}, I_0 + K_0 \begin{pmatrix} y \\ v \end{pmatrix}, \varepsilon) = \sum \frac{\varepsilon^i}{i!} \tilde{P}_i(x, y, u, v, 0).$$

(For the normal form in detail, refer to section 3.) Let $[\tilde{P}](y, u, v, \varepsilon) = \int_{\mathbb{T}^m} \tilde{P}(x, y, u, v, \varepsilon) dx = \sum \tilde{P}_i(y_0, u_0, v_0, 0).$ When $\det \partial^2 H_0 \neq 0$ and no eigenvalue of $\hat{\partial}^2 \tilde{P} [\tilde{P}] K' \partial^2 H_0 K'$ is positive or zero, Treshchëv ([59]) dealt with the persistence of resonant tori. When $\det \hat{\partial}^2 \tilde{P}[\tilde{P}] \neq 0$, for $g$–nondegenerate Hamiltonian system (1.1), Cong, Kipper, Li and You ([8]) dealt with the persistence of resonant invariant tori. Li and Yi ([20]) further removed the $g$–nondegenerate condition. When $\det \hat{\partial}^2 \tilde{P}_0 = 0$, what happens to the persistence of resonant tori becomes very complicated. The conjecture says that the number of the survival resonant tori is at least $m_0 + 1$ and generically $2^{m_0}$ for nondegenerate systems. We call perturbation $\tilde{P}(x, y, u, v, \varepsilon)$ $\kappa$–order nondegenerate, if
Let \( \det \partial^2_\nu [\tilde{P}](y_0, u_0, v_0, 0) = 0 \) for \( 0 \leq \nu \leq \kappa - 1 \) and \( \det \partial^2_\nu [\tilde{P}_\kappa](y_0, u_0, v_0, 0) \neq 0 \), where \((y_0, u_0, v_0)\) is the critical point of \([\tilde{P}]\). Obviously, above results only deal with the persistence of resonant tori for Hamiltonian \((1.1)\) with 0-order nondegenerate perturbation. In the present paper, we prove that \(2^{m_0}\) families of invariant torus survive for Hamiltonian system \((1.1)\) with \(\kappa\)-order nondegenerate perturbation, where \(\kappa\) is a given integer.

Now we are in a position to state our main results. We call \(P(I, \theta, \varepsilon)\) in \((1.1)\) \(\kappa\)-order nondegenerate, if there is a symplectic transformation \(\phi_g\) as in \((1.2)\) such that \(\tilde{P}(x, y, u, v, \varepsilon) = P \circ \phi_g\) is \(\kappa\)-order nondegenerate. First, we show results about a simple case, a \(g\)-nondegenerate Hamiltonian system with \(\kappa\)-order nondegenerate perturbation.

**Theorem 1.1.** Let \(g\)-nondegenerate Hamiltonian system \((1.1)\) with \(\kappa\)-order nondegenerate perturbation \(P(I, \theta, \varepsilon)\) be real analytic on the complex neighborhood of \(T^d \times G\). We have:

i) There exists a \(\varepsilon_0 > 0\) and a family of Cantor sets \(\tilde{\Lambda}_\varepsilon(g, G) \subset \tilde{\Lambda}(g, G)\), \(0 < \varepsilon < \varepsilon_0\), such that for each \(I \in \tilde{\Lambda}_\varepsilon(g, G)\), system \((1.1)\) admits \(2^{m_0}\) families of invariant torus, possessing hyperbolic, elliptic or mixed types, associated to nondegenerate relative equilibria. All such perturbed tori corresponding to a same \(I \in \tilde{\Lambda}_\varepsilon(g, G)\) are symplectically conjugated to the standard quasiperiodic \(m\)-tori \(T^m\) with the Diophantine frequency vector \(\omega_* = K_T \partial_I H_0(I)\). Moreover, the relative Lebesgue measure \(|\tilde{\Lambda}(g, G) \setminus \tilde{\Lambda}_\varepsilon(g, G)|\) tends to 0 as \(\varepsilon \to 0\).

ii) Consider \(g\)-nondegenerate Hamiltonian system \((1.1)\) with \(\kappa\)-order nondegenerate perturbation \(P(I, \theta, \varepsilon)\) on \(\sum = \{ I : H_0(I) = c \}\). Assume

\[
\text{(S1).} \quad \text{rank } \begin{pmatrix} K_T^T \partial^2_I H_0(I)K_0 & \omega_*^\top \\ \omega_* & 0 \end{pmatrix} = m + m_0 + 1, \text{ where } \omega_* = \left( \begin{array}{c} \omega_* \\ 0 \end{array} \right) \in R^{m+1+m_0}, \omega_* = K_T^T \partial_I H_0(I).
\]

Then there exists a \(\varepsilon_0 > 0\) and a family of Cantor sets \(\tilde{\Xi}_\varepsilon \subset \tilde{\Xi} = \{ I \in G : H_0(I) = c, (k, \omega) = 0, k \in G \}, 0 < \varepsilon < \varepsilon_0\), such that for each \(I \in \tilde{\Xi}_\varepsilon\), on a given energy-level manifold system \((1.1)\) admits \(2^{m_0}\) families of invariant torus, possessing hyperbolic, elliptic or mixed types, associated to nondegenerate relative equilibria. The frequencies \(\hat{\omega}\) of the persistent tori satisfy that \(\hat{\omega} = \omega_*\), where \(t \to 1\) as \(\varepsilon \to 0\). Moreover, the relative Lebesgue measure \(|\tilde{\Xi} \setminus \tilde{\Xi}_\varepsilon|\) tends to 0 as \(\varepsilon \to 0\).

**Remark 1.** Here a map defined on a Cantor set is said to be smooth in Whitney’s sense if it has a smooth Whitney extension. For details, see [27].

Since \([\tilde{P}](y, u, v, \varepsilon)\) is \(T^m\)-periodic in \(u\), there are at least \(m_0 + 1\) critical points for \([\tilde{P}](y, u, v, \varepsilon)\) for given \(y_0, v_0\) and \(\varepsilon_0(23)\). Note that \([\tilde{P}]\) is \(\kappa\)-order nondegenerate, \(\det \partial^2_{\nu}[\tilde{P}_\kappa](y_0, u_0, v_0, 0) \neq 0\), where \((y_0, u_0, v_0)\) is relative critical point, which means that such perturbations are generic according to Morse theory \((14, 23)\). Therefore, Theorem \((14)\) shows the persistence of resonant tori for
a $g$–nondegenerate Hamiltonian system (1.1) with a generic perturbation in the sense of the $\kappa$-order nondegeneracy, where $\kappa$ is a given positive integer. Hence this positively verifies the conjecture mentioned above in a general situation for $g$-nondegenerate Hamiltonian system (1.1).

The $\kappa$–order nondegenerate perturbation in the present paper is different from the case given by Treschëv ([36]), where the corresponding Hamiltonian is the following:

$$H(x, y, \varepsilon) = H_0(y) + \varepsilon H_1(y) + \cdots + \varepsilon^k H_k(y) + \varepsilon^{k+1} H_{k+1}(x, y, \varepsilon). \quad (1.4)$$

If there is some condition on the $0$–order Taylor coefficient of the average of $H_{k+1}$ in (1.4), he obtained the persistence of resonant tori (hyperbolic), and for some recent developments of such system, refer to [13, 33, 37, 38]. Actually, for the nearly integrable Hamiltonian system with a resonant integrable part and a $\kappa$-order nondegenerate perturbation, with finite KAM steps Hamiltonian system (1.1) can be reduced to the following system:

$$H(x, y, u, v, \varepsilon) = H_0(y, u, v) + \varepsilon H_1(y, u, v) + \cdots + \varepsilon^\kappa H_\kappa(y, u, v) + \varepsilon^{\kappa+1} H_{\kappa+1}(x, y, u, v, \varepsilon),$$

where $y$ and $v$ come from $I$ of the original system (1.1), $x$ and $u$ come from $\theta$ of the original system (1.1). For detail definitions and the process of reduction, refer to Section 2. Moreover, $\kappa$–order nondegenerate perturbation ensures the relative equilibria of $\varepsilon H_1(y, u, v) + \cdots + \varepsilon^\kappa H_\kappa(y, u, v)$ is nondegenerate, which means there are $2^{m_0}$ relative critical points according to Morse theory.

Next, we will give a more general case, in which we remove the $g$–nondegeneracy and study the partial preservation of frequency and partial preservation of ratios of frequencies. Let us do some assumptions for Hamiltonian system (1.1) first:

(S2). For $H_0(I)$ in (1.1), $\omega_\ast(I) = K_0^T \partial_I H_0(I)$ satisfies Rüssmann non-degenerate condition, i.e., for some $N > 0$, $\text{rank}\{\partial_I^\alpha \omega_\ast(I), |\alpha| < N\} = m$ for every $I \in \tilde{\Lambda}(g, G)$, where $\partial_I^\alpha \omega_\ast(I) = \partial^{\alpha_1}_{I_1} \cdots \partial^{\alpha_d}_{I_d} \omega_\ast, \alpha = (\alpha_1, \cdots, \alpha_d) \in Z^d_+, |\alpha| = |\alpha_1| + \cdots + |\alpha_d| \leq N$;

(S3). $\text{rank}(K_0^T \partial_I^2 H_0 K_0) = n + m_0, 0 \leq n \leq m$, and $\text{rank}((K')^T \partial_I^2 H_0 K_\ast, (K')^T \partial_I^2 H_0 K' \ast) = m_0$, where $H_0(I)$ comes from (1.1), $K_0$ and $K_\ast$ are defined as above;

(S4). $\text{rank} \left( \begin{array}{c} K^T_0 \partial_I^2 H K_0 \\ \omega_\ast \\ 0 \end{array} \right) = n + m_0 + 1, 0 \leq n \leq m$, where $\omega_\ast = \left( \begin{array}{c} \omega_\ast \\ 0 \end{array} \right) \in R^{m+m_0}$, $\omega_\ast = K_0^T \partial_I H_0(I) \in R^m$, $H_0(I)$ comes from (1.1), $K_0$ and $K_\ast$ are defined as above.

Now, let us state these more general results.

**Theorem 1.2.** Let Hamiltonian system (1.1) with a $\kappa$-order nondegenerate perturbation $P(I, \theta, \varepsilon)$ be real analytic on the complex neighborhood of $T^d \times G$. We have:
Consider the following Hamiltonian system

\[ H(\tilde{x}, \tilde{y}) = (\tilde{\omega}, \tilde{y}) + \frac{\varepsilon}{2} (\tilde{y}, M \tilde{y}) + \varepsilon^3 \cos(2x_1 - x_2) \\
+ \varepsilon^2 \cos(2x_1 - x_2) \sin(-x_1) e^{-y_1 - 2y_2}, \tag{1.5} \]

where \( \tilde{x} = (x_1, x_2)^T, \ \tilde{y} = (y_1, y_2)^T, \ \tilde{\omega} = (\omega_1, 2\omega_1)^T, \ x_1, x_2 \in T^1, \ y_1, y_2 \in \mathbb{R}^1, \)

\( \omega_1 \in \mathbb{R} \setminus \{0\} \) and \( M = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/4 \end{pmatrix}. \) Let \( P = \varepsilon^3 \cos(2x_1 - x_2) + \varepsilon^2 \cos(2x_1 - x_2) \sin(-x_1) e^{-y_1 - 2y_2}. \) Denote \( \tilde{\phi}_g : \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} y \\ v \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix}. \) Obviously, previous works do not apply to this system, since \( \tilde{\phi}(x, u, y, v) = P \circ \tilde{\phi}_g \) is 2–order nondegenerate perturbation. Actually, under transformation \( \tilde{\phi}_g, (1.4) \) is changed to

\[ H(x, y, u, v) = -\omega_1 y + \frac{\varepsilon}{2} \begin{pmatrix} y \\ v \end{pmatrix}, \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{5}{4} \end{pmatrix} \begin{pmatrix} y \\ v \end{pmatrix} \\
+ \varepsilon^3 \cos u + \varepsilon^2 \cos u \sin x, \]

which implies that our Theorem 1.2 works. Moreover, with our results there are 2 families of invariant torus, possessing hyperbolic, elliptic or mixed types, associated to nondegenerate relative equilibria. And \( n \) coordinates of the frequency \( \tilde{\omega} \) on the persistent tori coincide with \( n \) coordinates of \( \omega_\ast \). Moreover, the relative Lebesgue measure \( |\tilde{\Lambda}(g, G) \setminus \tilde{\Lambda}_g(G)| \) tends to 0 as \( \varepsilon \to 0. \)

Remark 2. Consider the following Hamiltonian system

\[ H(\tilde{x}, \tilde{y}) = (\tilde{\omega}, \tilde{y}) + \frac{\varepsilon}{2} (\tilde{y}, M \tilde{y}) + \varepsilon^3 \cos(2x_1 - x_2) \\
+ \varepsilon^2 \cos(2x_1 - x_2) \sin(-x_1) e^{-y_1 - 2y_2}, \tag{1.5} \]

Consider Hamiltonian system (1.1) with a \( \kappa \)–order nondegenerate perturbation \( P(I, \theta, \varepsilon) \) on \( \sum = \{ I \in G : H_0(I) = c \}. \) Assume (S2), (S3) and (S4) hold on \( \sum, \) let \( \tilde{\Xi} = \{ I \in G : H_0(I) = c, \langle k, \omega \rangle = 0, k \in g \}. \) Then there exists a \( \varepsilon_0 > 0 \) and a family of Cantor sets \( \tilde{\Xi}_I \subset \tilde{\Xi}, 0 < \varepsilon < \varepsilon_0, \) such that for each \( I \in \tilde{\Xi}_I, \) on a given energy-level manifold, system (1.1) admits at least \( 2^{\kappa_0} \) families of invariant tori, possessing hyperbolic, elliptic or mixed types, associated to nondegenerate relative equilibria. And \( n \) coordinates of the frequency \( \tilde{\omega} \) on the persistent tori coincide with \( n \) coordinates of \( t \omega_\ast, \) where \( t \to 1 \) as \( \varepsilon \to 0. \) Moreover, the relative Lebesgue measure \( |\tilde{\Xi} \setminus \tilde{\Xi}_I| \) tends to 0 as \( \varepsilon \to 0. \)

Remark 3. Condition (S2) ensures the existence of the resonant tori for perturbed system.

Remark 4. If \( n = m \) and \( (K')^T \partial^2 H_0 K' \) is nondegenerate, condition (S3) is \( g \)–nondegenerate condition mentioned in [5, 39], which ensures the preservation
of frequency in the process of KAM iteration. When \(n = m\), condition (S3) is the condition mentioned in \([20]\). Obviously, condition (S3) is weaker than all of them if \(n < m\). Combining conditions (S2) and (S3), in the process of KAM iteration, we could show the partial preservation of frequencies, which is determined by \((K^T \partial^2 H_0 K, K^T \partial^2 HK')\). The details will be shown in Section 2.

**Remark 5.** Under the isoenergetic nondegenerate condition:

\[
\det \left( \begin{array}{cc} \partial^2 H_0 & \partial I H_0 \\ \partial I H_0^T & 0 \end{array} \right) \neq 0,
\]

for Hamiltonian system (1.1), Arnold (\([2]\)) proved that on each energy-level manifold, the invariant tori form majority, which means that the Lebesgue measure of the complement of their union is small and depends on the perturbation. Conditions (S2), (S3) and (S4) are isoenergetic nondegenerate conditions for resonant tori, where (S3) and (S4) are closely related to the preservation of ratios of frequencies on a given energy-level manifold. As is well-known, the Kolmogorov nondegenerate condition and the classical isoenergetic nondegenerate condition are independent \([35]\). Our conditions do not violate this fact and reveal a further fact on partial preservation of ratios of frequencies: (S3) is also essential for the preservation of energy.

**Remark 6.** When \(n = m\) in condition (S3), (S2) holds automatically. If \(n = m\) in condition (S3) and perturbation \(P(I, \theta, \epsilon)\) in Hamiltonian system (1.1) is 0-order nondegenerate, part 1) of Theorem 1.2 is the result of \([20]\).

Finally, we give the following corollary according to Theorem 1.2.

**Corollary 1.1.** Let Hamiltonian system (1.1) with a \(\kappa\)-order nondegenerate perturbation \(P(I, \theta, \epsilon)\) be real analytic on the complex neighborhood of \(T^d \times G\). Assume (S2), (S4) and

(S5). \(K^T \partial^2 H_0 K\) has a \((m_0 + n) \times (m_0 + n)\) nonsingular minor, \(0 \leq n \leq m\), and \(\det K^T \partial^2 H_0 K' \neq 0\).

Then the conclusions of Theorem 1.2 also hold.

**Remark 7.** (S5) is equivalent to the following (S5\'): (S5\'). rank \((K^T \partial^2 H K' K^T \partial^2 H K' K^T \partial^2 H K' + K^T \partial^2 H K') = n, n < m, and \(\det K^T \partial^2 H_0 K' \neq 0\) for \(I \in \tilde{\Lambda}(g, G)\), which follows from the following fact:

\[
\begin{pmatrix} I_r & 0 \\ -DB^{-1} & I_{m-r} \end{pmatrix} \begin{pmatrix} B & C \\ D & E \end{pmatrix} = \begin{pmatrix} B & C \\ 0 & -DB^{-1}C + E \end{pmatrix},
\]

where \(B\) is nonsingular.
The classical Birkhoff normal form theory provides a formal integrability to harmonic oscillators with perturbation. But it does not work for the persistence of resonant tori studied in present paper, due to the nonlinearity of the unperturbed system and the degeneracy of \( \tilde{P}_0(y, u, v, 0) \). To overcome these difficulties, besides using Treschëv’s reduction, we propose a quasilinear normal form program by introducing quasilinear KAM iteration, which is used for searching high nondegeneracy and keeping critical points that are related to certain quasiperiodicity of the perturbation. In particular, our KAM iteration is more suitable for problems with worse normal forms. Hence, this approach provides a thorough way to study the persistence of resonant invariant tori under high degenerate perturbations.

The paper is organized as follows. In Section 2, we give an abstract Hamiltonian system and show the persistence of invariant tori. In this section, we introduce modificatory KAM step, which is interesting in itself. With the results of the abstract Hamiltonian system we finish the proof of Theorem 1.2 in Section 3. Finally, in Section 4 we also give two examples to show the complexity resulting from the high degeneracy of the perturbation.

2 Abstract Hamiltonian systems

Throughout the paper, unless specified explanation, we shall use the same symbol \(| \cdot |\) to denote an equivalent (finite dimensional) vector norm and its induced matrix norm, absolute value of functions, and measure of sets, etc., and use \(|.|_D\) to denote the supremum norm of functions on a domain \( D \). Also, for any two complex column vectors \( \xi, \zeta \) of the same dimension, \( \langle \xi, \zeta \rangle \) always stands for \( \xi^T \zeta \), i.e., the transpose of \( \xi \) times \( \zeta \). For the sake of brevity, we shall not specify smoothness orders for functions having obvious orders of smoothness indicated by their derivatives taking. All constants below are positive and independent of the iteration process. Moreover, all Hamiltonian functions in the sequel are associated to the standard symplectic structure.

Let \( z = (u, v) \in \mathbb{R}^{2m_0} \). To prove Theorem 1.2, consider the following real analytic Hamiltonian system with more general normal form

\[
H(x, y, z, \lambda, \varepsilon) = N(y, z, \lambda, \varepsilon) + \varepsilon^2 P(x, y, z, \lambda, \varepsilon),
\]

\[
N(y, z, \lambda, \varepsilon) = \langle \omega(\lambda), y \rangle + \frac{\varepsilon}{2} \left( \begin{array}{c} y \\ z \end{array} \right)^T M(\lambda) \left( \begin{array}{c} y \\ z \end{array} \right) + \varepsilon h(y, z, \lambda, \varepsilon),
\]

defined on

\[
D(r, s) = \{(x, y, z) : |Im x| < r, |y| < s, |z| < s\},
\]

where \( x \in \mathbb{T}^m, y \in \mathbb{R}^m, \lambda \in \Lambda, M, \) a symmetric matrix, depends smoothly on \( \lambda, h = O(|\left( \begin{array}{c} y \\ z \end{array} \right)|^3) \) is smooth. Here, \( \Lambda \) is a bounded closed region in \( \mathbb{R}^m \).

Thorough the paper, all \( \lambda \)–dependence are of class \( C^{l_0} \) for some integer \( l_0 \geq d \).
Rewrite
\[ M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}, \]
where \( M_{11}, M_{12}, M_{21}, M_{22} \) are \( m \times m, m \times 2m_0, 2m_0 \times m, 2m_0 \times 2m_0 \) matrices, respectively.

### 2.1 A General Theorem

To show the persistence of invariant tori for Hamiltonian (2.1), assume:

1. **(A1)** \( \text{rank} \{ \frac{\partial \omega}{\partial \lambda} : 0 \leq |\alpha| \leq m-1 \} = m \) for all \( \lambda \in \Lambda \).
2. **(A2)** For given \( n, 0 \leq n \leq m \), \( \text{rank}(M) = n + 2m_0 \) and \( \text{rank}(M_{21}, M_{22}) = 2m_0 \) for all \( \lambda \in \Lambda \), where \( M = (m_{ij})(m+2m_0) \times (m+2m_0) \).
3. **(A3)** For given \( n, 0 \leq n \leq m \),
   \[ \text{rank} \begin{pmatrix} M(\lambda) & \tilde{\omega}_1(\lambda) \\ \tilde{\omega}_T(\lambda) & 0 \end{pmatrix} = n + 2m_0 + 1, \]
   where \( \tilde{\omega}_1 \in \mathbb{R}^{m+2m_0}, \tilde{\omega}_T \in \mathbb{R}^{m}, M = (m_{ij})(m+2m_0) \times (m+2m_0) \).

**Remark 8.** We call (A2) and (A3) sub-isoenergetically nondegenerate conditions for the persistence of lower dimensional invariant tori. Specifically, when \( n = m \) and \( m_0 = 0 \), they are isoenergetically nondegenerate condition introduced by Arnold (II). When \( m_0 = 0 \), they are similar to the isoenergetically nondegenerate condition contained in (7, 35). When \( M \) is a block diagonal matrix, refer to (32) for a similar condition.

We state our results for (2.1) as follows.

**Theorem 2.1.** Let \( H(x, y, z, \lambda) \) in (2.1) be real analytic on the complex neighborhood of \( T^d \times G \).

i) Assume (A1) and (A2) hold on \( \Lambda \). Then there exists a \( \varepsilon_0 > 0 \) and a family of Cantor sets \( \Lambda_\varepsilon \subset \Lambda, 0 < \varepsilon < \varepsilon_0 \), such that for each \( \lambda \in \Lambda_\varepsilon \), system (2.1) admits a family of invariant tori. And \( n \) coordinates of the frequency \( \tilde{\omega} \) on the persistent tori coincide with \( n \) coordinates of \( \omega \), which are determined by those rows of \( (M_{11}, M_{12}) \) that are linearly independent. Moreover, the relative Lebesgue measure \( |\Lambda \setminus \Lambda_\varepsilon| \) tends to 0 as \( \varepsilon \to 0 \).

ii) Assume (A1), (A2) and (A3) hold on \( \mathbb{E} = \{ \lambda \in \Lambda : N(y, z, \lambda) = c \} \). Then there exists a \( \varepsilon_0 > 0 \) and a family of Cantor sets \( \mathbb{E}_\varepsilon \subset \mathbb{E}, 0 < \varepsilon < \varepsilon_0 \), such that for each \( \lambda \in \mathbb{E}_\varepsilon \), on a given energy-level manifold, system (2.1) admits a family of invariant tori. And \( n \) coordinates of the frequency \( \tilde{\omega} \) on the persistent tori coincide with \( n \) coordinates of \( t\omega \), which are determined by those rows of \( (M_{11}, M_{12}) \) that are linearly independent, where \( t \to 1 \) as \( \varepsilon \to 0 \). Moreover, the relative Lebesgue measure \( |\mathbb{E} \setminus \mathbb{E}_\varepsilon| \) tends to 0 as \( \varepsilon \to 0 \).
The proof of Theorem 2.1 will proceed by quasilinear KAM iteration process, which consists of infinitely many KAM steps. Due to the existence of small parameter $\varepsilon$ in term $\varepsilon^2 \left\langle \begin{bmatrix} y \\ z \end{bmatrix}, M \left( \begin{bmatrix} y \\ z \end{bmatrix} \right) + \varepsilon h(y, z, \lambda, \varepsilon) \right\rangle$, we weaken nondegenerate condition for the persistence of lower dimensional invariant tori, which is interesting in itself. For the case that there is no small parameter in normal direction, refer to [4, 10, 19, 21, 22, 25, 27, 39]. Next, we show the detail of our KAM steps.

2.2 KAM step

We show first the $0$–th KAM step. For the sake of induction, let

$$ r_0 = r, \ s_0 = s, \ \Lambda_0 = \Lambda, \ H_0 = H, \ N_0 = N, \ P_0 = P, \ M_0 = M, \ h_0 = h, $$

where $0 < r, s \leq 1$, and denote

$$ M^* = \max_{|l| \leq l_0, |j| \leq 2} |\partial^j_\lambda \partial^l_\mu h_0(y, z, \lambda)|_{D(r_0, s_0) \times \Lambda_0}. $$

For $j \in \mathbb{Z}_m^+$, define

$$ a_j = 1 - \text{sgn}(|j| - 1) = \begin{cases} 
2, & |j| = 0, \\
1, & |j| = 1, \\
0, & |j| \geq 2.
\end{cases} $$

Denote the complex neighborhood of $\Lambda_0$ by $\bar{\Lambda}_0 = \{ \lambda \in \mathbb{C}^m, |\lambda - \Lambda_0| \leq \varrho_0 \}$ for given constant $\varrho_0$. Let $\varepsilon = \delta, \ \gamma_0 = \varepsilon \frac{10^{l_0 + 9}}{\eta_0}, \ s_0 = \varepsilon^{\frac{1}{3}}, \ \mu_0 = \varepsilon^{\frac{1}{3}}, \ \iota \in (0, \frac{1}{3})$ and $\eta_0 = \frac{1}{4} \varrho_0$. Therefore, by Cauchy’s estimate,

$$ |\partial^j_\lambda P_0|_{D(r_0, s_0) \times \bar{\Lambda}_0} \leq c \delta \gamma_0^{l_0 + 9} \frac{s_0^2 \mu_0}{\eta_0} \eta_0 \eta_0 \quad (2.2) $$

for all $l \in \mathbb{Z}_m^+, |l| \leq l_0$, where $c > 0$ is a constant.

Next we characterize the iteration scheme for Hamiltonian (2.1) in one KAM step, say, from the $\nu$–th KAM step to the $(\nu + 1)$–th step. Recall $M_\nu = \begin{pmatrix} M_{11,\nu} & M_{12,\nu} \\ M_{21,\nu} & M_{22,\nu} \end{pmatrix}$, for given $k \in \mathbb{Z}_m^+$, denote

$$ \tilde{L}_{k0,\nu} = \sqrt{-1} \langle k, \omega_\nu \rangle, $$

$$ \tilde{L}_{k1,\nu} = \begin{pmatrix} \tilde{L}_{k0,\nu} I_m & -\delta M_{12,\nu} J \\ 0 & \tilde{L}_{k0,\nu} I_{2m_0} - \delta M_{22,\nu} J \end{pmatrix}, $$

$$ \tilde{L}_{k2,\nu} = \begin{pmatrix} I_m \otimes \tilde{L}_{k0,\nu} I_m & (\delta J M_{21,\nu})^T \otimes I_m \\ 0 & \tilde{a}_{22,\nu} & \tilde{a}_{23,\nu} & 0 \\ 0 & 0 & \tilde{a}_{33,\nu} \end{pmatrix}, $$
where \( \bar{a}_{22, \nu} = I_{2m_0} \otimes \bar{L}_{k0, \nu} I_m - (\delta M_{22, \nu}) \otimes I_m, \bar{a}_{33, \nu} = \bar{L}_{k0, \nu} I_{4m_0^2} - (\delta M_{22, \nu}) \otimes I_{2m_0} - I_{2m_0} \otimes (\delta M_{22, \nu}) \). For given matrix \( A \), \( A^\dagger \) represents conjugate transpose of \( A \). Let

\[
\Lambda_\nu = \{ \lambda \in \Lambda_{\nu-1} : |\bar{L}_{k0, \nu}| > \frac{\gamma_{\nu}}{|k|}, \bar{L}_{k1, \nu} I_{m+2m_0}, \bar{L}_{k2, \nu} > \frac{\gamma_{\nu}}{|k|} I_{m^2+2m_{m_0}+4m_0^2}, \text{ for all } 0 < |k| \leq K_\nu \},
\]

\[
\tilde{\Lambda}_\nu = \{ \lambda \in \mathbb{C}^n, |\lambda - \Lambda_\nu| \leq 4\eta_\nu \}, \quad \eta_\nu = \mu_{\nu-1}.
\]

Now, suppose that after \( \nu \) KAM steps, we have arrived at the following real analytic Hamiltonian system

\[
H_\nu(x, y, z) = N_\nu(y, z) + P_\nu(x, y, z, \varepsilon), \quad (2.3)
\]

\[
N_\nu(y, z) = (\omega_\nu(\lambda), y) + \frac{\delta}{2} \left( \begin{array}{c} y \\ z \end{array} \right), M_\nu(\lambda) \left( \begin{array}{c} y \\ z \end{array} \right) + \delta h_\nu(y, z, \lambda, \varepsilon),
\]

\[
|\partial_\alpha P_{\nu}|_{D(r_+, s_+)} \times \tilde{\Lambda}_\nu \leq \frac{\delta_\nu g^{\nu+9} s_\nu^2 \mu_\nu}{\eta_\nu}, \quad |l| \leq l_0,
\]

where \( M_\nu(\lambda) = (m_{ij})_{(m+2m_0) \times (m+2m_0)} \) satisfies that \( \text{rank}(M_\nu) = n + 2m_0 \) and \( \text{rank}(M_{21, \nu}, M_{22, \nu}) = 2m_0 \) for positive integer \( n \in [0, m] \) and \( \lambda \in \Lambda_\nu \), \( h_\nu = O(\left| \begin{array}{c} y \\ z \end{array} \right| |^3) \). For convenience, we shall omit the index for all quantities of the \( \nu \)-th KAM step and use ‘+’ to index all quantities in the \((\nu + 1)\)-th KAM step. To simplify the notions, we shall suspend the \( \lambda \)-dependence in most terms of this section. By considering both averaging and translation, we shall find a symplectic transformation \( \Phi_+ \), which, on a small phase domain \( D(r_+, s_+) \) and a smaller parameter domain \( \Lambda_+ \), transforms Hamiltonian \((2.3)\) into the following form:

\[
H_+ = H \circ \Phi_+ = N_+ + P_+,
\]

where on \( D(r_+, s_+) \times \tilde{\Lambda}_+ \), \( N_+ \) and \( P_+ \) enjoy similar properties as \( N \) and \( P \), respectively.

Define

\[
s_+ = \frac{1}{8} \alpha s, \quad \mu_+ = 6\alpha \mu_+ ^{\frac{13}{12}}, \quad r_+ = r - \frac{r_0}{2^{\nu+1}}, \quad \gamma_+ = \gamma - \frac{\gamma_0}{2^{\nu+1}},
\]

\[
\eta_+ = \mu_+ ^{-\frac{13}{12}}, \quad K_+ = (\log \left| \lambda \right| + 1)^3 \eta_+, \quad \Gamma(r - r_+) = \sum _{0 < |k| \leq K_+} |k| e^{-|k|},
\]

\[
\Lambda_+ = \{ \lambda \in \Lambda : |\bar{L}_{k0}| > \frac{\gamma}{|k|^2}, \bar{L}_{k1} \bar{L}_{k1} > \frac{\gamma}{|k|^2} I_{m+2m_0}, \bar{L}_{k2} \bar{L}_{k2} > \frac{\gamma}{|k|^2} I_{m^2+2m_{m_0}+4m_0^2}, \text{ for all } 0 < |k| \leq K_+ \},
\]

\[
\tilde{\Lambda}_+ = \{ \lambda \in \mathbb{C}^n, |\lambda - \Lambda_+| \leq 4\eta_+ \},
\]

\[
\hat{D}(\lambda) = D(r_+ + \frac{7}{8}(r - r_+), \lambda), \quad \hat{D}(\lambda) = \{ y \in \mathbb{C}^n : |y| < \lambda \}.
\]
\[
D_{\alpha} = D(r_+ + \frac{i-1}{8}(r-r_+), \frac{i}{8}\alpha s), \quad i = 1, 2, \ldots, 8,
\]

where \(\alpha = \mu \frac{1}{m}, \chi = 3\chi_1 = 3(m^2 - 2mm_0 + 4m_0^2)((l_0 + 5)\tau + 5l_0 + 10 + m^2 + 2mm_0 + 4m_0^2), c_0\) is the maximal among all \(c'\)s mentioned in this paper and depends on \(r_0, \beta_0\).

2.2.1 Truncation of the perturbation

Consider the Taylor-Fourier series of \(P\):

\[
P = \sum_{i \in Z^m} \sum_{j \in Z^2_{m_0}} \sum_{k \in Z^m} p_{kij} y^i z^j e^{\sqrt{-1} \langle k, x \rangle},
\]

and let \(R\) be the truncation of \(P\) with the following form:

\[
R = \sum_{|k| \leq K_+} \left( p_{k00} + \langle p_{k10}, y \rangle + \langle p_{k01}, z \rangle + \langle y, p_{k20} \rangle + \langle y, p_{k11} \rangle 
+ \langle z, p_{k02} \rangle \right) e^{\sqrt{-1} \langle k, x \rangle},
\]

where \(K_+\) is defined as above.

**Lemma 2.1.** Assume that

(H1) \(K_+ \geq \frac{8(\alpha + l_0)}{r - r_+}\),

(H2) \(\int_{K_+} x^{m+l_0} e^{-x^{r-r_+}} dx \leq \mu\).

Then there is a constant \(c\) such that for all \(|l| \leq l_0, \lambda \in \Lambda\),

\[
|\partial_\lambda^{l}(P - R)|_{D_{\alpha} \times \tilde{\Lambda}} \leq c \frac{\delta \gamma \lambda l_0 + s^2 \mu^2}{\eta l_0}.
\]

**Proof.** The proof is standard. For detail, refer to, for example, **Lemma 3.1** of [20].

2.2.2 Homological equations

We want to average out all coefficients of \(R\) by constructing a symplectic transformation as the time-1 map \(\phi_1^F\) of the flow generated by a Hamiltonian \(F\) with the following form:

\[
F = \sum_{0 < |k| \leq K_+} \left( f_{k00} + \langle f_{k10}, y \rangle + \langle f_{k01}, z \rangle + \langle y, f_{k20} \rangle + \langle y, f_{k11} \rangle + \langle z, f_{k02} \rangle \right) e^{\sqrt{-1} \langle k, x \rangle},
\]

(2.4)
where \( f_{kij}, 0 \leq |i| + |j| \leq 2 \), are scalar, vectors or matrices with obvious dimensions, which are allowed to depend on \( y, z \) and \( \lambda \). Under the time-1 map \( \phi^1_F \), Hamiltonian (2.3) becomes

\[
H \circ \phi^1_F = (N + R) \circ \phi^1_F + (P - R) \circ \phi^1_F \\
= N + R + \{N, F\} + \int_0^1 \{R_t, F\} \circ \phi^1_F dt + (P - R) \circ \phi^1_F, \tag{2.5}
\]

where \( R_t = (1 - t)\{N, F\} + R \). Let

\[
\{N, F\} + R - [R] - R' = 0, \tag{2.6}
\]

where

\[
[R] = \int_T R(x, \cdot) dx, \\
R' = \partial_x hJ \partial_x F + \langle y, M_{12}J \Delta_0 \rangle + \langle z, M_{22}J \Delta_0 \rangle, \\
\Delta_0 = \langle y, \partial_x f_{k20}y \rangle + \langle y, \partial_z f_{k11}z \rangle + \langle z, \partial_z f_{k02}z \rangle, \\
\hat{h} = \frac{\delta}{2} \left( \begin{array}{c} y \\ z \end{array} \right), M \left( \begin{array}{c} y \\ z \end{array} \right) + \delta h(y, z, \lambda, \varepsilon). \tag{2.7}
\]

Then Hamiltonian (2.5) arrives at

\[
\tilde{H}_+ = \tilde{N}_+(y, z) + \tilde{P}_+(x, y, z), \tag{2.8}
\]

where \( \tilde{N}_+ = N + [R], \tilde{P}_+ = R' + \int_0^1 \{R_t, F\} \circ \phi^1_F dt + (P - R) \circ \phi^1_F \).

Consider the following symplectic translation:

\[
\phi : x \to x, \left( \begin{array}{c} y \\ z \end{array} \right) \to \left( \begin{array}{c} y + y_0 \\ z + z_0 \end{array} \right), \tag{2.9}
\]

where \((y_0, z_0)\) is determined by

\[
\delta \frac{M}{2} \left( \begin{array}{c} y_0 \\ z_0 \end{array} \right) + \delta \left( \begin{array}{c} \partial_y h(y_0, z_0, \lambda) \\ \partial_z h(y_0, z_0, \lambda) \end{array} \right) = - \left( \begin{array}{c} p_{010} \\ p_{001} \end{array} \right). \tag{2.10}
\]

Then Hamiltonian system (2.8) is changed to

\[
H_+ = \tilde{H}_+ \circ \phi \\
= e_+ + \langle \omega, y \rangle + \frac{\delta}{2} \left( \begin{array}{c} y_0 \\ z_0 \end{array} \right), M_+ \left( \begin{array}{c} y_0 \\ z_0 \end{array} \right) + \delta h_+(y, z, \lambda, \varepsilon) + P_+, 
\]

where

\[
e_+ = e + \langle \omega, y_0 \rangle + \frac{\delta}{2} \left( \begin{array}{c} y_0 \\ z_0 \end{array} \right), M_+ \left( \begin{array}{c} y_0 \\ z_0 \end{array} \right) + p_{000} + \left( \begin{array}{c} p_{010} \\ p_{001} \end{array} \right), \left( \begin{array}{c} y_0 \\ z_0 \end{array} \right) \\
+ \left( \begin{array}{c} y_0 \\ z_0 \end{array} \right), \left( \begin{array}{c} p_{010} \\ p_{001} \end{array} \right), \left( \begin{array}{c} y_0 \\ z_0 \end{array} \right) + \delta h(y_0, z_0, \lambda),
\]
\[ \omega_+ = \omega + \frac{\delta M}{2} \left( \begin{array}{c} y_0 \\ z_0 \end{array} \right) + \delta \left( \begin{array}{c} \partial_y h(y_0, z_0, \lambda) \\ \partial_z h(y_0, z_0, \lambda) \end{array} \right) + \left( \begin{array}{c} p_{010} \\ p_{001} \end{array} \right), \]

\[ M_+ = M + 2 \left( \begin{array}{c} \frac{p_{020}}{2p_{011}} \\ \frac{p_{010}}{2p_{002}} \end{array} \right) + \partial^2_{(y, z)} h(y_0, z_0, \lambda), \]  \quad (2.11)

\[ P_+ = \hat{P}_+ + \delta \left( \begin{array}{c} y \\ z \end{array} \right), \quad \left( \begin{array}{c} \frac{p_{010}}{2p_{011}} \\ \frac{p_{010}}{2p_{002}} \end{array} \right) \left( \begin{array}{c} y_0 \\ z_0 \end{array} \right), \]  \quad (2.12)

\[ h_+ = h(y, z, \lambda) - h(y_0, z_0, \lambda) - \left( \begin{array}{c} \partial_y h(y_0, z_0, \lambda) \\ \partial_z h(y_0, z_0, \lambda) \end{array} \right) \left( \begin{array}{c} y \\ z \end{array} \right) \]

\[ - \frac{1}{2} \left( \begin{array}{c} y \\ z \end{array} \right), \partial^2_{(y, z)} h(y_0, z_0, \lambda) \left( \begin{array}{c} y \\ z \end{array} \right). \]  \quad (2.13)

### 2.2.3 Estimate on the transformation

According to the definition of Poisson bracket on coordinate \((x, y, z) \in T^m \times R^m \times R^{2m_0}\),

\[ \{N, F\} = \partial_x N \partial_y F - \partial_y N \partial_x F + \partial_z N \partial_z F \]

\[ = -\partial_y N \partial_x F + \partial_z \hat{h} J \partial_z F, \]

where \(J = \left( \begin{array}{cc} 0 & I_{m_0 \times m_0} \\ -I_{m_0 \times m_0} & 0 \end{array} \right) \). Then (2.10) is changed to

\[ -\partial_y N \partial_x F + \partial_z \hat{h} J \partial_z F + R - [R] = 0. \]

(2.14)

Denote \(\Delta_1 = \partial_y \hat{h} \delta(M_{11} y + M_{12} z + \partial_y h(y, z, \lambda))\). Directly,

\[ \partial_y N \partial_x F = \sum_{0 \leq |k| \leq K_x} \sqrt{-1}(k, \omega + \Delta_1) \left( f_{k00} + \langle f_{k10}, y \rangle + \langle f_{k01}, z \rangle \right. \]

\[ + \langle y, f_{k20} \rangle y + \langle y, f_{k11} z \rangle + \langle z, f_{k02} \rangle z \rangle e^{\sqrt{-1}(k, x)}, \]

(2.15)

\[ R - [R] = \sum_{0 \leq |k| \leq K_x} \left( p_{k00} + \langle p_{k10}, y \rangle + \langle p_{k01}, z \rangle \right. \]

\[ + \langle y, p_{k20} \rangle y + \langle y, p_{k11} z \rangle + \langle z, p_{k02} \rangle z \rangle e^{\sqrt{-1}(k, x)}. \]

(2.16)

Substituting (2.15) and (2.16) into (2.14) yields:

\[ \sqrt{-1}(k, \omega + \Delta_1) f_{k00} = p_{k00}, \]

(2.17)

\[ \sqrt{-1}(k, \omega + \Delta_1) f_{k10} - \delta M_{12} J f_{k01} = p_{k10} + \delta M_{12} J \partial_z f_{k00}, \]

(2.18)

\[ \sqrt{-1}(k, \omega + \Delta_1) f_{k01} - \delta M_{22} J f_{k01} = p_{k01} + \delta M_{22} J \partial_z f_{k00}, \]

(2.19)

\[ \sqrt{-1}(k, \omega + \Delta_1) f_{k11} + \delta f_{k11} J M_{21} = p_{k20} + \delta M_{12} J \partial_z (f_{k10})^T, \]

(2.20)

\[ \sqrt{-1}(k, \omega + \Delta_1) f_{k11} - 2\delta M_{12} J f_{k02} - \delta f_{k11} (M_{22} J)^T \]

\[ = p_{k11} + \delta M_{12} J \partial_z (f_{k01})^T + (\delta M_{22} J \partial_z (f_{k10})^T)^T, \]

(2.21)

\[ \sqrt{-1}(k, \omega + \Delta_1) f_{k02} - \delta M_{22} J f_{k02} + \delta f_{k02} J M_{22} \]

\[ = p_{k02} + \delta M_{22} J \partial_z (f_{k10})^T. \]

(2.22)
For any matrix $A = (a_{ij})_{p \times q}$, denote $T(A) = (a_{11}, \cdots, a_{p1}, \cdots, a_{1q}, \cdots, a_{pq})^T$. Let

\[
L_{k0} = \sqrt{1} \langle k, \omega + \Delta \rangle,
L_{k1} = \begin{pmatrix} L_{k0} I_m & -\delta M_{12} J \\ 0 & L_{k0} I_{2m_0} - \delta M_{22} J \end{pmatrix},
L_{k2} = \begin{pmatrix} I_m \otimes (L_{k0} I_m) & (\delta J M_{21})^T \otimes I_m \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix},
\]

\[
a_{22} = I_{2m_0} \otimes (L_{k0} I_m) - (\delta M_{22} J) \otimes I_m,
\]

\[
a_{33} = L_{k0} I_{4m_0}^2 - (\delta M_{22} J) \otimes I_{2m_0} - I_{2m_0} \otimes (\delta M_{22} J).
\]

Rewrite (2.17) - (2.22) as follows

\[
L_{k0} f_{k00} = p_{k00},
L_{k1} \begin{pmatrix} f_{k10} \\ f_{k01} \end{pmatrix} = \begin{pmatrix} p_{k10} \\ p_{k01} \end{pmatrix} + \delta \begin{pmatrix} M_{12} J \partial_z f_{k00} \\ M_{22} J \partial_z f_{k00} \end{pmatrix},
L_{k2} \begin{pmatrix} T(f_{k10}) \\ T(f_{k11}) \\ T(f_{k02}) \end{pmatrix} = \begin{pmatrix} T(p_{k10}) \\ T(p_{k11}) \\ T(p_{k02}) \end{pmatrix} + \delta \begin{pmatrix} T(M_{12} J \partial_z (f_{k10})^T) \\ T(M) \\ T(M_{22} J \partial_z (f_{k10})^T) \end{pmatrix},
\]

where $\hat{M} = M_{12} J \partial_z (f_{k10})^T + (M_{22} J \partial_z (f_{k00})^T)^T$.

On pages 17-19,

**Lemma 2.2.** Assume that

\[(H3) \max_{|j| \leq 2} |\partial_y^j \hat{h}(y, z, \lambda) - \partial_y^j \hat{h}_0(y, z, \lambda)|_{D(r,s) \times \Lambda} \leq \mu_0^{\frac{1}{2}}.\]

Then there is a constant $c$ such that for all $|l| \leq l_0$,

\[
|\partial_{\lambda}^l e_+ - \partial_{\lambda}^l e|_{D(r,s) \times \Lambda} \leq c \frac{\mu_0^{l_0 + 9} s \mu}{\eta_0^{l_0}},
\]

\[
|\partial_{\lambda}^l M_+ - \partial_{\lambda}^l M|_{D(r,s) \times \Lambda} \leq c \frac{\mu_0^{l_0 + 9} \mu}{\eta_0^{l_0}},
\]

\[
|\partial_{\lambda}^l \omega_+ - \partial_{\lambda}^l \omega|_{D(r,s) \times \Lambda} \leq c \frac{\delta_0 (\gamma_0^{l_0 + 9} + \mu)}{\eta_0^{l_0}},
\]

\[
|\partial_{\lambda}^l y_0 (\partial_{\lambda}^l z_0)|_{D(r,s) \times \Lambda} \leq c \frac{\mu_0^{l_0 + 9} \mu}{\eta_0^{l_0}}.
\]

**Proof.** Obviously,

\[
|\partial_{\lambda}^l p_{000}|_{\Lambda} \leq c \frac{\delta_0^{l_0 + 9} \mu^{2} \mu}{\eta_0^{l_0}},
|\partial_{\lambda}^l p_{010}|_{\Lambda} + |\partial_{\lambda}^l p_{001}|_{\Lambda} \leq c \frac{\delta_0^{l_0 + 9} \mu^{2} \mu}{\eta_0^{l_0}}.
\]

Denote

\[
B = \frac{M}{2} + \left( \int_0^1 \partial_y^2 h(\theta y, z, \lambda) d\theta \right.
\]

\[
\left. + \int_0^1 \partial_y \partial_z h(\theta y, z, \lambda) d\theta \right) \int_0^1 \partial_y \partial_z h(y, \theta z, \lambda) d\theta.
\]
Then (2.10) becomes

\[
\begin{align*}
\delta B \begin{pmatrix} y \\ z \end{pmatrix} &= - \begin{pmatrix} p_{010} \\ p_{001} \end{pmatrix}.
\end{align*}
\]

(2.30)

For given matrix \(A = (a_{ij})_{n \times n}\), let \(||A||_1 = \frac{1}{n} \sum_{i,j=1}^{n} |a_{ij}(\lambda)|\), where \(|a_{ij}(\lambda)|\) is the absolute value of \(a_{ij}(\lambda), \lambda \in \Lambda\). According to assumption \((H3)\) and the definition of \(M^*\), we have \(||M - M_0||_1 \leq \mu_0^2, ||\partial^2_{(y,z)} h||_1 \leq (M^* + 1)s\), where \(M_0\) is \(M_\nu\) for \(\nu = 0\). Denote \(M_* = ||M_0^{-1}||_1\) for \(\lambda \in \Lambda\). Without loss of generality, let \(\mu_0\) and \(s_0\) be small enough such that \(s_0^2 M_* (M^* + 1) \leq \frac{1}{4}\) and \(\mu_0 M_* \leq \frac{1}{4}\).

Then

\[
||M_0 - B||_1 \leq ||M - M_0||_1 + ||B - M||_1
\leq \mu_0^2 + (M^* + 1)s^2
\leq \frac{1}{2M_*}.
\]

Let \(M_0\) be nonsingular. It follows that \(B\) is nonsingular and

\[
||B^{-1}||_1 = \| \frac{M_0^{-1}}{I - (M_0 - B)M_0^{-1}} \|_1
\leq \| M_0^{-1} \|_1
\leq \| I - (M_0 - B)M_0^{-1} \|_1
\leq \| M_0^{-1} \|_1
\leq \frac{M_*}{1 - ||(M_0 - B)||_1 ||M_0^{-1}||_1}
\leq \frac{M_*}{\frac{M_*}{2M_*}}
= 2M_*.
\]

Here, we use the fact that \(||(I - A)^{-1}||_1 \leq \frac{1}{1 - ||A||_1}||I||_1\), which is obvious if \(||I||_1 = 1\) and \(||A||_1 < 1\). Therefore,

\[
\left| \begin{pmatrix} y \\ z \end{pmatrix} \right|_{D(r,s) \times \tilde{\lambda}} = \left| \frac{1}{\delta} B^{-1} \begin{pmatrix} p_{010} \\ p_{001} \end{pmatrix} \right|
\leq \frac{m + 2m_0}{\delta} ||B^{-1}||_1 \left| \begin{pmatrix} p_{010} \\ p_{001} \end{pmatrix} \right|
\leq c\gamma^{b+9} s \mu.
\]

Consider the differential with respect to \(\lambda\) on both sides of (2.30)

\[
\partial_{(y,z)} B \left( \begin{pmatrix} \partial_\lambda y \\ \partial_\lambda z \end{pmatrix} \right) \begin{pmatrix} y \\ z \end{pmatrix} + \partial_\lambda B \begin{pmatrix} y \\ z \end{pmatrix} + B \begin{pmatrix} \partial_\lambda y \\ \partial_\lambda z \end{pmatrix} = - \begin{pmatrix} \partial_\lambda P_{010} \\ \partial_\lambda P_{001} \end{pmatrix}.
\]

16
Then
\[
\left| \left( \frac{\partial_{\lambda} y}{\partial_{\lambda} z} \right) \right|_{D(r,s) \times \hat{\Lambda}} = |B^{-1}(\frac{\partial_{\lambda} P_{010}}{\partial_{\lambda} P_{001}}) + \partial_{(y,z)} B \left( \frac{\partial_{\lambda} y}{\partial_{\lambda} z} \right) (y, z) |
\]
\[
+ \partial_{\lambda} B \left( \frac{y}{z} \right) |
\]
\[
\leq 2M_{*} \frac{\gamma l_{0} + 9s \mu}{\eta} + 4M_{*}^{2}(M^{*} + 1) \frac{\gamma l_{0} + 9s \mu}{\eta} \left| \left( \frac{\partial_{\lambda} y}{\partial_{\lambda} z} \right) \right| (y, z)
\]
\[
+ 4M_{*}^{2}(M^{*} + 1) \frac{\gamma l_{0} + 9s \mu}{\eta}
\]
\[
\leq c \frac{\gamma l_{0} + 9s \mu}{\eta}.
\]

Inductively, we get (2.29). According to the definition of \( e_+ \), \( \omega_+ \), and \( M_+ \), (2.26), (2.27) and (2.28) are obvious.

Recall \( \chi_1 = (m^2 + 2mm_0 + 4m_0^2)(l_0 + 5) \tau + 5l_0 + 10 + m^2 + 2mm_0 + 4m_0^2) \).

**Lemma 2.3.** Assume that

\( \text{(H4)} \max \{s, \mu^{\frac{1}{2}} \} K_{\gamma}^1 = o(\gamma) \).

The following hold for all \( 0 < |k| \leq K_+ \).

1. On \( D(s) \times \hat{\Lambda}_+ \), for \( l_0 \leq |l| \leq l_0 + |i| + |j| \leq 2 \),

\[
|\partial_{\lambda} \partial_{\eta} \partial_{\zeta} f_{kij} |_{D(s) \times \hat{\Lambda}_+} \leq c \frac{\delta |k|^{3} l_{0}^{2} - |j| - |i| \mu e^{-|k|r}}{\eta_+^{l_{0}}};
\]

2. On \( \hat{D}(s) \times \hat{\Lambda}_+ \),

\[
|\partial_{\lambda} \partial_{\eta} \partial_{\zeta} (y,z) F |_{\hat{D}(s) \times \hat{\Lambda}_+} \leq c \frac{\delta s^2 \mu \Gamma (r - r_+)}{\eta_+^{l_{0}}}, \quad |i| < l_0, \quad |j| \leq 2, \quad |l| < l_0.
\]

**Proof.** Denote \( \omega = \omega(\lambda) \) for \( \lambda \in \Lambda \) and \( \omega_0 = \omega(\lambda) \) for \( \lambda \in \hat{\Lambda} \). Recall \( \eta_+ = \mu^{\frac{1}{2}} \).

For any \( \lambda \in \hat{\Lambda}_+ \), \( 0 < |k| \leq K_+ \), with assumption (H4) we have

\[
|L_{k0}|_{D(s) \times \hat{\Lambda}_+} = |\sqrt{-1} \langle k, \omega \rangle + \sqrt{-1} \langle k, \Delta \rangle + \sqrt{-1} \langle k, \omega - \omega_0 \rangle| \geq \frac{\gamma}{|k|^2} - c \max \{s, \mu^{\frac{1}{2}} \} \delta K_+
\]
\[
\geq \frac{\gamma}{2|k|^2};
\]

(2.31)
and \( |\partial_{\lambda}^l \partial_{\eta}^j \partial_{\tau}^r L_{k_0}|_{D(s) \times \Lambda_+} \leq c|k| \). Applying the above and the following inequalities

\[
|\partial^l L_{k_0}^{-1}| \leq |L_{k_0}^{-1}| \sum_{|\tau'|=1}^{||l||} \left( \begin{array}{c} l \\
\tau'
\end{array} \right) |\partial^{l-l'} L_{k_0}^{-1}||\partial^{l'} L_{k_0}|,
\]

inductively, we deduce that

\[
|\partial_{\lambda}^l \partial_{\eta}^j \partial_{\tau}^r L_{k_0}^{-1}|_{D(s) \times \Lambda_+} \leq c|k|^{||l||+|j|+|r||} \frac{|L_{k_0}^{-1}|^{||l||+|j|+|r||+1}}{\gamma^{||l||+|j|+|r||+1}}.
\]

(2.32)

It follows from (2.32), (2.33) and Cauchy’s estimate that

\[
|\partial_{\lambda}^l \partial_{\eta}^j \partial_{\tau}^r f_{k_00}|_{D(s) \times \Lambda_+} \leq \delta|\partial_{\lambda}^l \partial_{\eta}^j \partial_{\tau}^r (L_{k_0}^{-1} p_{k_00})|_{(D(s) \times \Lambda_+)} \leq \frac{\delta|k|^{\chi_1}}{\gamma^{||l||+|j|+|r||+1}} \frac{\gamma^{\alpha}}{\eta^2} \frac{\delta \mu}{\eta^2} e^{-|k||r|},
\]

(2.33)

Recall \( L_{k_1} = L_{k_1} + \sqrt{-1}(k, \Delta_1)I_{m+2m_0} \). Then, according to the basic property of Hermitian matrix (157), on \( D(s) \times \Lambda_+ \),

\[
L_{k_1}^* L_{k_1} = L_{k_1}^* L_{k_1} + \sqrt{-1}(k, \Delta_1)(L_{k_1}^* L_{k_1} - L_{k_1} + (\sqrt{-1}(k, \Delta_1)I_{m+2m_0})^*)
\]

\[
\geq \frac{\gamma}{|k|^r} I_{m+2m_0} - c \max \{s, \mu, \frac{m_0}{s} \} K_+ I_{m+2m_0}
\]

\[
\geq \frac{\gamma}{2|k|^r} I_{m+2m_0}.
\]

(2.34)

Therefore,

\[
|\det L_{k_1}^* L_{k_1}|_{D(s) \times \Lambda_+} \leq \left( \frac{|\det L_{k_1}^* L_{k_1}|_{D(s) \times \Lambda_+}}{|\det L_{k_1}|_{D(s) \times \Lambda_+}} \right)^2 \leq \left( \frac{\gamma}{2|k|^r} \right)^{m+2m_0}.
\]

Inductively,

\[
|\partial_{\lambda}^l \partial_{\eta}^j \partial_{\tau}^r \left( \frac{1}{\det L_{k_1}} \right)|_{D(s) \times \Lambda_+} \leq \frac{|k|^{m+2m_0)(||l||+|j|+|r||)}{\det L_{k_1}} \leq \frac{|k|^{m+2m_0)(||l||+|j|+|r||+1)}{\gamma^{m+2m_0)(||l||+|j|+|r||+1)}}.
\]

Hence

\[
|\partial_{\lambda}^l \partial_{\eta}^j \partial_{\tau}^r \left( f_{k_10} \right)|_{D(s) \times \Lambda_+}
\]
\[
\frac{\partial^2}{\partial y^2} \left( L_{k1}^{-1} \left( \begin{array}{c} p_{k10} \\ p_{k01} \end{array} \right) + \delta \begin{pmatrix} M_{12} J \partial_z f_{k00} \\ M_{22} J \partial_z f_{k00} \end{pmatrix} \right) \right)_{D(s) \times \Lambda^+}
\]

\[
= \frac{\partial^2}{\partial y^2} \left( \frac{\text{adj} L_{k1}}{\det L_{k1}} \left( \begin{array}{c} p_{k10} \\ p_{k01} \end{array} \right) + \delta \begin{pmatrix} M_{12} J \partial_z f_{k00} \\ M_{22} J \partial_z f_{k00} \end{pmatrix} \right) \right)_{D(s) \times \Lambda^+}
\]

\[
\leq \frac{\gamma^{(m+2m_0)(\|l\|+\|i\|+\|j\|+1) + 1}}{\eta_0^l} \left( \frac{\delta^t \lambda \mu e^{-|k| r} + s^2 \mu |k| \chi_3 e^{-|k| r}}{\eta_0^l} \right)
\]

Similarly, on \( D(s) \times \Lambda^+ \),

\[
L_{k2}^* L_{k2} = \tilde{L}_{k2}^* \tilde{L}_{k2} + \sqrt{-1} (k, \Delta) (\tilde{L}_{k2}^* - \tilde{L}_{k2} + (\sqrt{-1} (k, \Delta) I_{m^2 + 2m_m + 4m_r^2})^*) \geq \frac{\gamma}{2|k|^r} I_{m^2 + 2m_m + 4m_r^2}.
\]

Hence

\[
\frac{\partial^2}{\partial y^2} \left( \begin{array}{c} f_{k20} \\ f_{k11} \\ f_{k02} \end{array} \right) - \frac{\partial^2}{\partial y^2} \left( \begin{array}{c} T(f_{k20}) \\ T(f_{k11}) \\ T(f_{k02}) \end{array} \right) \right)_{D(s) \times \Lambda^+}
\]

\[
= \frac{\partial^2}{\partial y^2} \left( \frac{\text{adj} L_{k2}}{\det L_{k2}} \left( \begin{array}{c} T(p_{k20}) \\ T(p_{k11}) \\ T(p_{k02}) \end{array} \right) \right)
\]

\[
\leq \frac{|k| \chi_1}{\gamma^{(m^2 + 2m_m + 4m_r^2)(\|l\|+\|i\|+\|j\|+1) + 1}} \left( \frac{\delta^t \lambda \mu e^{-|k| r} + s^2 \mu |k| \chi_3 e^{-|k| r}}{\eta_0^l} \right)
\]

Now, we finish the proof of part (1).

For part (2), by part (1) and directly differentiating to (2.4), we have, on \( \tilde{D}(s) \times \Lambda^+ \),

\[
\frac{\partial^2}{\partial y^2} \left( \begin{array}{c} \partial^i_{(y,z)} f_{(y,z)} \end{array} \right)_{\tilde{D}(s) \times \Lambda^+} \leq \sum_{0 < |k| \leq K^+} |k| |\partial^i_{(y,z)} f_{(y,z)}| s^{l - \text{sgn}(j)}
\]

\[
+ |\partial^i_{(y,z)} f_{(y,z)}| s^{l - \text{sgn}(j)}
\]

\[
+ |\partial^i_{(y,z)} f_{(y,z)}| s^{l - \text{sgn}(j)}
\]

\[
+ |\partial^i_{(y,z)} f_{(y,z)}| s^{l - \text{sgn}(j)}
\]

\[
+ |\partial^i_{(y,z)} f_{(y,z)}| s^{l - \text{sgn}(j)}
\]

\[
+ |\partial^i_{(y,z)} f_{(y,z)}| s^{l - \text{sgn}(j)}
\]

\[
+ \frac{1}{\eta_0^l} (r + \tau(r - r_+))
\]

19
\[
\begin{align*}
\leq c \frac{\delta \mu s^{a_i}}{\eta_+^{l_0}} \sum_{0 < |k| \leq K_+} |k|^{3\alpha_1} e^{-|k|/(r-r_+)} \\
= c \frac{\delta \mu s^{a} \Gamma(r-r_+)}{\eta_+^{l_0}}.
\end{align*}
\]

\[\square\]

Similar to Lemma 3.6 of [29], here, \( F \) can also be smoothly extended to functions of Hölder class \( C^{\alpha_0+\alpha_1, l_0-1+\alpha_0}(\bar{D}(\sigma_0) \times \Lambda_0) \), where \( 0 < \sigma_0 < 1 \) is fixed. Moreover, there is a constant \( c \) such that
\[|F|_{C^{\alpha_0+\alpha_1, l_0-1+\alpha_0}(\bar{D}(\sigma_0) \times \Lambda_0)} \leq c \delta \mu \Gamma(r-r_+).\]

**Lemma 2.4.** Assume

(H5) \( c_\mu \Gamma(r-r_+) < \frac{1}{2}(r-r_+) \),

(H6) \( c_\mu \Gamma(r-r_+) < \frac{1}{2}\alpha \).

Then the following hold:

1) For all \( 0 \leq t \leq 1 \),
\[
\begin{align*}
\phi_F^t : D_{\frac{1}{2}+} &\to D_{\frac{1}{2}+}, \\
\phi : D_{\frac{1}{2}+} &\to D_{\frac{1}{2}+}
\end{align*}
\]
are well defined, real analytic and depend smoothly on \( \lambda \in \Lambda_+ \);

2) There is a constant \( c \) such that for all \( 0 \leq t \leq 1 \), \( |l| \leq l_0 \), \( |j| \leq 2 \), \( |i| \leq l_0 \),
\[
|\partial_\lambda^i \partial_\phi^j \partial_\delta^l (\phi_F^t \circ \phi - id)|_{D_{\frac{1}{2}+} \times \Lambda_+} \leq c \frac{\mu \Gamma(r-r_+)}{\eta_+^{l_0}}.
\]

**Proof.** Let \( \phi_F^t = (\phi_F^t, \phi_F^t, \phi_F^t)^T \), where \( \phi_F^t \), \( \phi_F^t \) and \( \phi_F^t \) are components of \( \phi_F^t \) in \( x \), \( y \) and \( z \)-coordinate, respectively. Obviously, \( \phi_F^t = id + \int_0^t X_F \circ \phi_F^t \circ ds \), where \( X_F = (\partial_\phi F, -\partial_\phi F, J\phi_\phi F)^T \). Let \( (x, y, z) \) be any point in \( D_{\frac{1}{2}+} \) and let \( t_* = \sup\{t \in [0, 1] : \phi_F^t(x, y, z) \in D_{\frac{1}{2}+}\} \). Then, for \( t \in [0, t_*] \), \( \lambda \in \Lambda_+ \), with (H5) and (H6),
\[
\begin{align*}
|\phi_F^t(x, y, z) - x|_{D_{\frac{1}{2}+}} &\leq \int_0^t |F_y \circ \phi_F^t|_{D_{\frac{1}{2}+}} ds \leq |F_y|_{\bar{D}(\sigma_0)} \leq \delta \mu \Gamma \leq \frac{1}{8}(r-r_+), \\
|\phi_F^t(x, y, z) - y|_{D_{\frac{1}{2}+}} &\leq \int_0^t |F_x \circ \phi_F^t|_{D_{\frac{1}{2}+}} ds \leq |F_x|_{\bar{D}(\sigma_0)} \leq \delta \mu s^2 \Gamma \leq \frac{\alpha s}{8}, \\
|\phi_F^t(x, y, z) - z|_{D_{\frac{1}{2}+}} &\leq \int_0^t |F_z \circ \phi_F^t|_{D_{\frac{1}{2}+}} ds \leq |F_z|_{\bar{D}(\sigma_0)} \leq \delta \mu s \Gamma \leq \frac{\alpha s}{8},
\end{align*}
\]
which implies \( |\phi_F^t(x, y, z)| < r_* + \frac{3}{8}(r-r_+), |\phi_F^t(x, y, z)| < \frac{\alpha s}{8}, |\phi_F^t(x, y, z)| < \frac{\alpha s}{8} \), i.e. \( \phi_F^t(x, y, z) \in D_{\frac{1}{2}+} \). Using (2.23) and (H6), \( \phi : D_{\frac{1}{2}+} \to D_{\frac{1}{2}+} \) is obvious.

The proof of 2) follows from Lemma 2.3. \[\square\]
2.2.4 New perturbation

Here we will estimate the new perturbation $P_+$ on the domain $D_+ \times \Lambda_+$, where $D_+ = D\tilde{\mu}_+$. 

**Lemma 2.5. Assume**

(H7) $\mu \to \Gamma^3(r - r_+) \leq \gamma_{+}^{l_0+9}$. 

Then 

$$|\partial^j P_+|_{D_+ \times \Lambda_+} \leq c \frac{\delta \gamma_{+}^{l_0+9} s_2 \mu_+}{\eta_+^{l_0}}.$$ 

**Proof.** Directly, 

$$|R'|_{D_+ \times \Lambda_+} \leq c \delta^2 s^3 \mu \Gamma(r - r_+).$$ 

Denote $\partial^{i,j} = \partial_x^i \partial_y^j$ for $|j| \leq 2$, $|i| \leq l_0$. Then 

$$|\partial^{i,j} \{ R_t, F \} \circ \phi_{R} |_{D_+ \times \Lambda_+} \leq c \delta s^{s_0} \mu^2 \Gamma^3(r - r_+),$$ 

$$|\partial^{i,j} (P - R) \circ \phi_{R} |_{D_+ \times \Lambda_+} \leq c \delta \gamma_{+}^{l_0+9} s_0 \mu_+^2 \Gamma(r - r_+),$$ 

$$|\partial^{i,j} R' \circ \phi |_{D_+ \times \Lambda_+} \leq c \delta^2 s^{s_0+1} \mu \Gamma(r - r_+),$$ 

$$|\partial^{i,j} \langle \begin{pmatrix} y \\ z \end{pmatrix} , \begin{pmatrix} p_{002} \\ p_{011} \\ \frac{1}{2} p_{011} \\ p_{002} \end{pmatrix} \rangle |_{D_+ \times \Lambda_+} \leq c \delta \gamma_{+}^{l_0+9} s_0 \mu^2.$$ 

Further, by (2.12), we have 

$$|\partial^j P_+|_{D_+ \times \Lambda_+} \leq c \frac{\delta s^2 \mu^2 \Gamma^3(r - r_+)}{\eta_+^{l_0}}.$$ 

Here we use the fact that $s = c \mu \mu_0^{-2} s_0$ and $\delta \mu_0^{-2} s_0 = o(c)$. (According to the construction of $s_\nu$ and $\mu_\nu$, obviously, $s = c \mu \mu_0^{-2} s_0$.) Using assumption (H7), we finish the proof of this lemma. 

\[ \square \]

2.2.5 The preservation of frequencies

Combining the argument in subsections 2.2.2 and 2.2.3, if $M(\lambda)$ is nonsingular, there is a transformation (2.28) such that all the frequencies are preserved after a KAM step. However, when $M(\lambda)$ is singular, (2.10) is not solvable, i.e. there is no transformation such that all frequencies are preserved after a KAM step. To show the part preservation of frequency, we give a simple property.
Lemma 2.6. For an $n \times n$ symmetrical matrix $A$ with rank($A$) = $m$, there is an invertible matrix $T$ that corresponds to a linear transformation, under which only some rows of $A$ exchange, such that

$$T^{-1}AT = \begin{pmatrix} B & C \\ D & E \end{pmatrix},$$

where $B$ is an $m \times m$ nonsingular minor.

Proof. Rewrite

$$A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = (b_1, b_2, \cdots, b_n),$$

where $a_i$ is $i$–th row of $A$ and $b_i$ is $i$–th column of $A$, $i = 1, \cdots, n$. Since $A$ is symmetrical, $a_i = b_i^T$, $i = 1, \cdots, n$, which means that there is a same linear relation between $a_i$ and $b_i$, $i = 1, \cdots, n$. Because rank($A$) = $m$, there are $m$ linearly independent rows (columns) of $A$. Then there is an invertible matrix $T$, which corresponds to a linear transformation that exchange some rows of $A$, such that

$$T \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} a_1^1 \\ a_2^1 \\ \vdots \\ a_m^1 \end{pmatrix},$$

where $a_1^1, \cdots, a_m^1$ are linearly independent. Since $T^{-1} = T$ and $T^{-1}$ does not change the linear relation among $b_1, \cdots, b_m$, we get

$$T^{-1}AT = \begin{pmatrix} a_1^1 \\ a_2^1 \\ \vdots \\ a_m^1 \end{pmatrix} T = \begin{pmatrix} B & C \\ D & E \end{pmatrix},$$

where $B$ is an $m \times m$ nonsingular minor. \qed
Combining assumption (A2) and Lemma 2.6 there is an invertible matrix $T$, which corresponds to a transformation only exchanging columns or rows, such that

$$T^{-1} \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} T = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix},$$

where $(C_{11}, C_{12})_{(n+2m_0)\times(m+2m_0)}$ is a matrix with $\text{rank}(C_{11}, C_{12}) = n + 2m_0$ and $(C_{21}, C_{22})_{(m-n)\times(m+2m_0)}$ is the complements. Moreover, $(C_{11})_{(n+2m_0)\times(n+2m_0)}$ is nonsingular. Denote $p, z \in \mathbb{R}^{m+2m_0}$, which corresponds to a transformation only exchanging columns or rows,

$$\partial y_1 h(y_0, z_0, \lambda) + \partial y_2 h(y_0, z_0, \lambda) = - \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}. \quad (2.38)$$

Since $\text{rank}(C_{11}, C_{12}) = \text{rank} \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}$, there is an invertible matrix $T_1$ that only exchange columns or rows such that $T_1 \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} \\ 0 & 0 \end{pmatrix}$, which is equivalent to the fact that the rows of $(C_{21}, C_{22})$ is linearly dependent on the rows of $(C_{11}, C_{12})$. Obviously, $T_1$ is a matrix with the following form

$$\begin{pmatrix} I & 0 \\ D_1 & I \end{pmatrix},$$

where $D_1$ is determined by the linear relation among the rows of $(C_{21}, C_{22})$ and $(C_{11}, C_{12})$. Then

$$T_1 \begin{pmatrix} \partial y_1 h(y_0, z_0, \lambda) \\ \partial y_2 h(y_0, z_0, \lambda) \end{pmatrix} = \begin{pmatrix} \partial y_1 h(y_0, z_0, \lambda) \\ D_1 \partial y_1 h(y_0, z_0, \lambda) + \partial y_2 h(y_0, z_0, \lambda) \end{pmatrix},$$

$$T_1 \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} = \begin{pmatrix} p_1 \\ D_1 p_1 + p_2 \end{pmatrix}.$$

Consider the following equation

$$\frac{\delta}{2} \begin{pmatrix} C_{11} & C_{12} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} \partial y_1 h(y_0, z_0, \lambda) \\ 0 \end{pmatrix} = - \begin{pmatrix} p_1 \\ 0 \end{pmatrix}, \quad (2.39)$$

where $C_{11}$ is nonsingular. Obviously, $(y_1, y_2)^T = (y_1, 0)^T$ is a specific solution of (2.39), i.e., with assumption (A2) there is a symplectic transformation such that part of the frequencies are preserved.

**Remark 9.** If $M$ is singular, some of the frequencies are preserved and the others drift. Moreover, the drift depends on $D_1 p_1 + p_2$ and $D_1 \partial y_1 h(y_0, z_0, \lambda) + \partial y_2 h(y_0, z_0, \lambda)$ and the estimate on drift is showed by (??).

Consider :

$$\langle \omega, y_0 \rangle + \frac{\delta}{2} \left( y_0 \begin{pmatrix} y_0 \\ z_0 \end{pmatrix}, M \begin{pmatrix} y_0 \\ z_0 \end{pmatrix} \right) + p_{00} + \left( p_{010} \begin{pmatrix} y_0 \\ p_{001} \end{pmatrix} \right) \quad (2.40)$$
\[ + \left( \begin{array}{c} y_0 \\ z_0 \end{array} \right), \left( \begin{array}{c} p_{010} \\ \frac{1}{2} p_{011} \\ p_{002} \end{array} \right) \left( \begin{array}{c} y_0 \\ z_0 \end{array} \right) \right) + \delta h(y_0, z_0, \lambda) = 0, \]
\[
\delta M \left( \begin{array}{c} y_0 \\ z_0 \end{array} \right) + \delta \left( \frac{\partial y}{\partial y} h(y_0, z_0, \lambda) \right) + \left( \begin{array}{c} p_{010} \\ p_{001} \end{array} \right) - t \left( \begin{array}{c} \omega \\ 0 \end{array} \right) = 0. \tag{2.41}
\]

If \( M \) is nonsingular, according to (A3) and the continuity of determinant, we have
\[
\det \left( \begin{array}{c} \tilde{\omega}_1 \\ 0 \end{array} \right) \neq 0, \text{ where } \omega_1 = (\omega, 0)^T \in R^{n+2m_0}, \omega_2 = (p_{010} + \omega, p_{001}).
\]
Then, combining (2.40) and (2.41), with implicit theorem we get \((y_0, z_0, t), \text{i.e., we construct a transformation such that on the same energy surface the ratios of the frequencies are preserved after a KAM step.}

**Remark 10.** If \( M \) is nonsingular, the condition \( \det \left( \begin{array}{c} M \\ \tilde{\omega}_1 \\ 0 \end{array} \right) \neq 0 \) is a generalization of the isoenergetically nondegenerate condition given by V. I. Arnold (2) to the persistence of lower dimensional invariant tori on a given energy surface, where \( \omega_1 = (\omega, 0)^T \).

Assume \( M \) is singular and conditions (A2) and (A3) hold. Denote \( \tilde{\omega}_1 \) by the first \( n + 2m_0 \) components of \( T_1 T^{-1}(\omega, 0)^T \), which is equal to the first \( n + 2m_0 \) components of \( T^{-1}(\omega, 0)^T \). In fact,
\[
T_1 T^{-1} \left( \begin{array}{c} \omega \\ 0 \end{array} \right) = T_1 \left( \begin{array}{c} \tilde{\omega}_1 \\ \omega_4 \end{array} \right) = \left( \begin{array}{c} I \\ D_1 \end{array} \right) \left( \begin{array}{c} \tilde{\omega}_1 \\ \omega_4 \end{array} \right) = \left( \begin{array}{c} \tilde{\omega}_1 \\ \tilde{D}_1 \omega_1 + \omega_4 \end{array} \right),
\]
where \( \omega = (\omega_3, \omega_4)^T \in R^m, \omega_1 = (\omega_3, 0)^T \in R^{n+2m_0} \). Similarly, combining (2.39), we have
\[
\delta \left( \begin{array}{c} C_{11} \\ C_{12} \\ 0 \\ 0 \end{array} \right) \left( \begin{array}{c} y_1 \\ y_2 \end{array} \right) + \delta \left( \begin{array}{c} \partial y/h(y_0, z_0, \lambda) \\ 0 \end{array} \right) - t \left( \begin{array}{c} \tilde{\omega}_1 \\ 0 \end{array} \right) = - \left( \begin{array}{c} p_1 \\ 0 \end{array} \right). \tag{2.42}
\]
Assume
\[
\det \left( \begin{array}{c} C_{11} \\ C_{12} \\ \tilde{\omega}_2 \\ 0 \end{array} \right) \neq 0,
\]
where \( \tilde{\omega}_2 \) is the first \( n + 2m_0 \) components of \((p_{010} + \omega, p_{001})T \). Then there is a \((y_0^0, \cdots, y_0^n, 0, \cdots, 0, z_0^0, \cdots, z_0^{2m_0}, t) \) such that
\[
\langle \omega, y_0 \rangle + \delta \left( \begin{array}{c} y_0 \\ z_0 \end{array} \right), M \left( \begin{array}{c} y_0 \\ z_0 \end{array} \right) \right) + p_{000} + \left( \begin{array}{c} p_{010} \\ p_{001} \end{array} \right), \left( \begin{array}{c} y_0 \\ z_0 \end{array} \right) \right) + \left( \begin{array}{c} y_0 \\ z_0 \end{array} \right), \left( \begin{array}{c} \frac{1}{2} p_{011} \\ p_{002} \end{array} \right) \left( \begin{array}{c} y_0 \\ z_0 \end{array} \right) \right) + \delta h(y_0, z_0, \lambda) = 0,
\]
\[
\delta \left( \begin{array}{c} C_{11} \\ C_{12} \\ 0 \\ 0 \end{array} \right) \left( \begin{array}{c} y_1 \\ y_2 \end{array} \right) + \delta \left( \begin{array}{c} \partial y/h(y_0, z_0, \lambda) \\ 0 \end{array} \right) - t \left( \begin{array}{c} \tilde{\omega}_1 \\ 0 \end{array} \right) = - \left( \begin{array}{c} p_1 \\ 0 \end{array} \right). \tag{2.42}
\]
Finally, combining (A2), (A3), Property (2.6) and the continuity of determinant, assumption (2.42) holds. Therefore, on a given energy surface there is a transformation such that ratios of frequencies between the unperturbed torus and the perturbed are preserved.
Remark 11. Assume (A2) and (A3). For a given energy, n coordinates of the frequency $\omega_\epsilon$ coincide with n coordinates of $t\omega$, where $t \to 0$ as $\epsilon \to 0$. Simultaneously, the other frequencies slightly drift and the drift depend on $D_1 p_1 + p_2$ and $D_1 \partial_y h(y_0, z_0, \lambda) + \partial_y_h(y_0, z_0, \lambda)$.

2.3 Iteration Lemma

Let $r_0, \gamma_0, s_0, \eta_0, \Lambda_0, H_0, N_0, e_0, P_0$ be given as above and denote $D_0 = D(r_0, \beta_0)$. For any $\nu = 0, 1, \cdots$, denote

$$r_\nu = r_0(1 - \sum_{i=0}^{\nu} \frac{1}{2^{i+1}}), \quad \gamma_\nu = \gamma_0(1 - \sum_{i=0}^{\nu} \frac{1}{2^{i+1}}), \quad \alpha_\nu = \frac{1}{\nu},$$

$$\eta_\nu = \frac{\mu_\nu}{\mu_\nu^0}, \quad \mu_\nu = 64c_0\mu_{\nu-1}^0, \quad K_\nu = (\log \frac{1}{\mu_\nu-1}) + 1)^3\eta_\nu,$$

$$D_\nu = D(r_\nu, s_\nu), \quad \hat{D}_\nu = D(r_\nu + \frac{7}{8}(\nu - 1) - \nu), \quad s_\nu = \frac{1}{8}\alpha_{\nu-1}s_{\nu-1},$$

$$\Lambda_\nu = \{ \lambda \in \Lambda_{\nu-1} : |\hat{L}_{k0,\nu}| > \nu \frac{\mu_\nu}{|k|}, \hat{L}_{k1,\nu} > \frac{\gamma_\nu}{|k|^2} I_{m+2m_0},$$

$$\hat{L}_{k2,\nu} > \frac{\gamma_\nu}{|k|^2} I_{m^2+2m_{m_0}+4m_0^2}, \text{ for all } 0 < |k| \leq K_\nu \},$$

$$\hat{\Lambda}_\nu = \{ \lambda \in \mathbb{C}^m, |\lambda - \Lambda_\nu| \leq 4\eta_\nu \}.$$

We have the following Iteration Lemma.

Lemma 2.7. Assume (2.2) hold. Then the KAM step described in Section 2.2 is valid for all $\nu = 0, 1, \cdots$, and the following facts hold for all $\nu = 1, 2, \cdots$.

1. $P_\nu$ is real analytic in $(x, y, z) \in D_\nu$, smooth in $(x, y, z) \in \hat{D}_\nu$ and smooth in $\lambda \in \Lambda_\nu$, and moreover,

$$|\partial_x^l P_\nu|_{D_\nu \times \hat{\Lambda}_\nu} \leq c_0 \frac{\delta \mu_\nu^{3l+9}s_\nu^2}{\eta_\nu^{3l+9}}, \quad |l| \leq l_0;$$

2. $\Phi_\nu = \phi_\nu \circ \phi : \hat{D} \times \Lambda_0 \to \hat{D}_{\nu-1}, D_\nu \times \Lambda_\nu \to D_{\nu-1}$, is symplectic for each $\lambda \in \Lambda_0$, and is of class $C^{\alpha+1+\sigma_0, \alpha+2+\sigma_0}$, $C^{\alpha, \alpha+1}$, respectively, where $\alpha$ stands for real analyticity and $0 < \sigma_0 < 1$ is fixed. Moreover,

$$H_\nu = H_{\nu-1} \circ \Phi_\nu = N_\nu + P_\nu,$$

on $\hat{D} \times \Lambda_\nu$, and

$$|\Phi_\nu - \text{id}|_{C^{\alpha+1+\sigma_0, \alpha+2+\sigma_0}(\hat{D} \times \hat{\Lambda}_\nu)} \leq c_0 \frac{\delta \mu_\nu^{3l_0}}{2^\nu};$$

3. $\Lambda_\nu = \{ \lambda \in \Lambda_{\nu-1} : |\hat{L}_{k0,\nu}| > \frac{\gamma_\nu}{|k|}, \hat{L}_{k1,\nu} > \frac{\gamma_\nu}{|k|^2} I_{m+2m_0},$

$$\hat{L}_{k2,\nu} > \frac{\gamma_\nu}{|k|^2} I_{m^2+2m_{m_0}+4m_0^2}, \text{ for all } 0 < |k| \leq K_\nu \}.$$

Proof. The proof of this lemma is to verify conditions (H1) – (H7). Those are standard and we place the detail on Appendix [3].
2.4 Convergence and measure estimate

Let \( \Psi^\nu = \Phi_1 \circ \Phi_2 \circ \cdots \circ \Phi_\nu, \ \nu = 1, 2, \cdots \). Then \( \Psi^\nu : \tilde{D}_\nu \times \Lambda_0(g, G) \to \tilde{D}_0 \), and
\[
H_0 \circ \Psi^\nu = H_{\nu} = N_{\nu} + P_{\nu}, \\
N_{\nu} = e_{\nu} + \langle \omega_{\nu}, y \rangle + h_{\nu}(y, \omega), \ \nu = 0, 1, \cdots,
\]
where \( \Psi_0 = id \).

Standardly, \( N_{\nu} \) converges uniformly to \( N_{\infty} \), \( P_{\nu} \) converges uniformly to \( P_{\infty} \) and \( \partial_y^l \partial_x^k P_{\infty} = 0, |l| + |k| \leq 2 \).

Hence for each \( \lambda \in \Lambda_{\infty} \), \( T^d \times \{0\} \times \{0\} \) is an analytic invariant torus of \( H_{\infty} \) with the total frequency \( \omega_{\infty} \), which for all \( k \in \mathbb{Z}^m \setminus \{0\} \), \( 1 \leq q \leq n \), by the definition of \( \Lambda_{\nu} \) and Lemma 2.7 (2), satisfies the following facts

1. if (A1) holds and \( M \) is nonsingular, then \( \omega_{\infty} \equiv \omega_0, \ |\langle k, \omega_{\infty} \rangle| > \frac{\gamma}{2|k|}; \)
2. if (A1) and (A3) hold and \( M \) is nonsingular, then on a given energy surface \( \omega_{\infty} \equiv t_\omega, \ |\langle k, \omega_{\infty} \rangle| > \frac{\gamma}{2|k|}; \)
3. if (A1) and (A2) hold, then \( (\omega_{\infty})_{i_1} \equiv (\omega_0)_{i_1}, q = 1, \cdots, n, \ |\langle k, \omega_{\infty} \rangle| > \frac{\gamma}{2|k|}; \)
4. if (A1), (A2) and (A3) hold, then \( (\omega_{\infty})_{i_1} \equiv t(\omega_0)_{i_1}, q = 1, \cdots, n, \ |\langle k, \omega_{\infty} \rangle| > \frac{\gamma}{2|k|}. \)

Following the Whitney extension of \( \Psi^\nu \), all \( e_{\nu}, \omega_{\nu}, h_{\nu}, P_{\nu}, (\nu = 0, 1, \cdots) \) admit uniform \( C^{l_0-1+\varepsilon_0} \) extensions in \( \lambda \in \Lambda_{\nu} \) with derivatives in \( \lambda \) up to order \( l_0 - 1 \). Thus, \( e_{\nu}, \omega_{\infty}, h_{\nu}, P_{\nu} \) are \( C^{l_0-1} \) Whitney smooth in \( \lambda \in \Lambda_{\nu} \), and the derivatives of \( e_{\nu}, \omega_{\infty}, h_{\nu}, P_{\nu}, h_{\nu} - h_0 \) satisfy similar estimates. Consequently, the perturbed tori form a \( C^{l_0-1} \) Whitney smooth family on \( \Lambda_{\nu}(g, G) \).

The measure estimate is the same as ones in [7, 31, 33, 34] and for the sake of completeness we place details on Appendix C. Now we have finished the proof of Theorem 1.2.

3 Proof of Theorem 1.2

For \( d \)-dimensional manifold \( \mathcal{M} \) with a global coordinate, there is a bounded closed region \( \Lambda \in \mathbb{R}^n \) and a \( C^{l_0} \) diffeomorphism \( I : \Lambda \to \mathcal{M} \) such that \( \mathcal{M} = I(\Lambda) \). Under the transformation \( I \to I + I(\lambda) \), Hamiltonian system (1.1) is changed to
\[
H(I, \theta, \lambda, \varepsilon) = e + \langle \omega(\lambda), I \rangle + \frac{1}{2}(I, \partial^2_I H_0(\lambda)I) + O(|I|^3) + \varepsilon P(I, \theta, \lambda, \varepsilon), \quad (3.1)
\]
where \( e = H_0(I(\lambda)), \ \omega(\lambda) = \partial_t H(I(\lambda)) \). Let
\[
\Gamma = K_0^T \partial^2_I H_0(\lambda) K_0 = \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{pmatrix},
\]
and
where $\Gamma_{11}, \Gamma_{12}, \Gamma_{21}, \Gamma_{22}$ are $m \times m$, $m \times m_0$, $m_0 \times m$, $m_0 \times m_0$ matrices, respectively, $\Gamma_{12} = \Gamma_{21}^T$, $\Gamma_{22} = K^T \partial^2_{\lambda} H_0(\lambda) K'$, and $m_0 = d - m$. Denote $\omega^*(\lambda) = K^T \omega(\lambda) \in \Lambda(g, \Lambda)$, where $\Lambda(g, \Lambda) = \{ \lambda \in \Lambda : \langle k, \omega(\lambda) \rangle = 0, k \in g \}$ and $\tilde{\Lambda}(g, \Lambda) = \{ \omega^*(\lambda) = K^T \omega \in R^m, \lambda \in \Lambda(g, \Lambda) \}$. Recall $p = (y, v)$, $q = (x, u)$, where $y = (p_1, \ldots, p_m)^T$, $v = (p_{m+1}, \ldots, p_d)^T$, $x = (q_1, \ldots, q_m)^T$, $u = (q_{m+1}, \ldots, q_d)^T$. For any $\lambda \in \tilde{\Lambda}(g, \Lambda)$, with the following coordinate transformation $I = K_0p, q = K_0^T \theta$, Hamiltonian (3.1) is changed to

$$H(x, y, u, v) = \langle \omega^*, y \rangle + \frac{1}{2} \left( \begin{pmatrix} y \\ v \end{pmatrix}, \Gamma(\lambda) \left( \begin{pmatrix} y \\ v \end{pmatrix} \right) \right) + O(|K_0 \left( \begin{pmatrix} y \\ v \end{pmatrix} \right)|^3) + \varepsilon P(x, y, u, v, \varepsilon) \quad (3.2)$$

up to a constant, where $\tilde{P}(x, y, u, v, \varepsilon) = P(K_0 \left( \begin{pmatrix} y \\ v \end{pmatrix} \right), (K_0^T)^{-1} \left( \begin{pmatrix} x \\ u \end{pmatrix} \right), \varepsilon)$. By the following symplectic transformation:

$$\left( \begin{array}{c} y \\ v \end{array} \right) \rightarrow \varepsilon^{\frac{1}{2}} \left( \begin{array}{c} y \\ v \end{array} \right), \quad \left( \begin{array}{c} x \\ u \end{array} \right) \rightarrow \left( \begin{array}{c} x \\ u \end{array} \right), \quad H \rightarrow \varepsilon^{-\frac{3}{2}} H, \quad (3.3)$$

Hamiltonian (3.2) is changed to

$$H(x, y, u, v) = \langle \omega^*, y \rangle + \frac{\varepsilon^{\frac{3}{2}}}{2} \left( \begin{pmatrix} y \\ v \end{pmatrix}, \Gamma(\lambda) \left( \begin{pmatrix} y \\ v \end{pmatrix} \right) \right) + O(|K_0 \left( \begin{pmatrix} y \\ v \end{pmatrix} \right)|^3) + \varepsilon^{\frac{3}{2}} \tilde{P}(x, y, u, v, \varepsilon). \quad (3.4)$$

In order to use Theorem 2.1, we should reduce Hamiltonian system (3.3) to (2.1). But the traditional method fails due to high degeneracy of perturbation, which does not guarantee that the perturbation is sufficiently small. Hence we have to proceed a program, finite quasilinear KAM steps, to improve the order of the perturbation. To fix thought, we only give an outline.

Let $\epsilon = \varepsilon^{\frac{3}{2}}$. Rewrite Hamiltonian system (3.3) with the following form:

$$H_1(x, y, u, v) = N_1(y, v) + \epsilon^2 P_1(x, y, u, v, \epsilon), \quad (3.5)$$

where $N_1 = \langle \omega_1, y \rangle + \tilde{h}_1, \tilde{h}_1 = \frac{\varepsilon}{2} \left( \begin{pmatrix} y \\ v \end{pmatrix}, \tilde{M}_1 \left( \begin{pmatrix} y \\ v \end{pmatrix} \right) \right) + \epsilon^2 O(|K_0 \left( \begin{pmatrix} y \\ v \end{pmatrix} \right)|^3),

P_1(x, y, u, v) = \epsilon P(x, y, u, v).$ Rewrite $\tilde{M}_1 = \left( \begin{pmatrix} \tilde{M}_{11,1} & \tilde{M}_{12,1} \\ \tilde{M}_{21,1} & \tilde{M}_{22,1} \end{pmatrix} \right)$, where $\tilde{M}_{11,1}$, $\tilde{M}_{12,1}$, $\tilde{M}_{21,1}$, $\tilde{M}_{22,1}$ are $m \times m, m \times m_0, m_0 \times m, m_0 \times m_0$ matrices, respectively.

Let $z = (u, v)$ and $M_1 = \left( \begin{pmatrix} M_{11,1} & M_{12,1} \\ M_{21,1} & M_{22,1} \end{pmatrix} \right)$, where $M_{11,1} = 0, M_{21,1} = 0, M_{31,1} = 0, M_{41,1} = 0, M_{51} = 0$ with obvious dimension. Choose $\epsilon = \delta$, $\gamma = \delta^{\frac{3}{4}}$, $s = \delta^{\frac{3}{4}}, \mu = \delta^{\frac{3}{4}}$. Then (3.5) is changed to

$$H_1(x, y, u, v, \lambda) = N_1(y, v, \lambda) + \delta^2 P_1(x, y, u, v, \lambda, \epsilon), \quad (3.6)$$

27
where $N_1 = \langle \omega_1(\lambda), y \rangle + h_1$, $\hat{h}_1 = \frac{\delta^2}{2} \left( \begin{array}{c} y \\ z \end{array} \right), M_1(\lambda) \left( \begin{array}{c} y \\ z \end{array} \right), |P_1(x, y, u, v, \lambda)| \leq \gamma \lambda^{n+9} s^2 \mu$. Here, $h_1$ is a polynomial of $K_0 \left( \begin{array}{c} y \\ v \end{array} \right)$ from the third order term.

Let $M_{11,1} = \tilde{M}_{11,1}$, $M_{12,1} = (\tilde{M}_{1,1}, \tilde{M}_{11,1})$, $M_{21,1} = \left( \begin{array}{c} \tilde{M}_{2,1} \\ \tilde{M}_{22,1} \end{array} \right)$, $M_{22,1} = \left( \begin{array}{c} \tilde{M}_{3,1} \\ \tilde{M}_{4,1} \\ \tilde{M}_{5,1} \\ \tilde{M}_{22,1} \end{array} \right)$.

Write, for $|i| + |j| \leq 2$,

$$P_1 = \sum_k p_{kij} y^i z^j e^{\sqrt{T}(k,x)},$$

$$R_1 = \sum_{|k| \leq K_1} p_{kij} y^i z^j e^{\sqrt{T}(k,x)},$$

$$P_1 - R_1 = \sum_{|k| > K_1} p_{kij} y^i z^j e^{\sqrt{T}(k,x)},$$

where $K_1$ is specified in Section 2.

Next, we are going to improve the order of $P_1$ by the symplectic transformation $\Phi^1_{F_1}$, the time–1 map generated by the vector field $J \nabla F_1$ with $J = \left( \begin{array}{cc} 0 & I_m \\ -I_m & 0 \\ 0 & 0 & -I_{m_0} \end{array} \right)$, where $F_1(x, y, z, \lambda) = \sum_{0 < |k| \leq K_1} f_{kij} y^i z^j e^{\sqrt{T}(k,x)}$

that satisfies

$$\{N_1, F_1\} + \delta^2 (R_3 - [R_1]) - R_1' = 0,$$

$$R_1' = \partial_z h_1 J \partial_x F_1 + \langle y, M_{12,1} J \Delta_0 \rangle + \langle z, M_{22,1} J \Delta_0 \rangle,$$

$$\Delta_0 = \langle y, \partial_z f_{k20} y \rangle + \langle y, \partial_z f_{k11} z \rangle + \langle z, \partial_z f_{k02} z \rangle,$$

$$[R_1](y, z, \lambda, \varepsilon) = \int_{\Gamma_m} R_1(x, y, z, \lambda, \varepsilon) dx.$$

Using [67] and comparing coefficients, we obtain the following quasilinear homological equations

$$L_{k0,1} f_{k00} = p_{k00}, \quad (3.8)$$

$$L_{k1,1} \left( \begin{array}{c} f_{k10} \\ f_{k01} \end{array} \right) = \left( \begin{array}{c} p_{k10} \\ p_{k01} \end{array} \right) + \delta \left( \begin{array}{c} M_{12,1} J \partial_x f_{k00} \\ M_{22,1} J \partial_x f_{k00} \end{array} \right), \quad (3.9)$$

$$L_{k2,1} \left( \begin{array}{c} T(f_{k20}) \\ T(f_{k11}) \\ T(f_{k02}) \end{array} \right) = \left( \begin{array}{c} T(p_{k20}) \\ T(p_{k11}) \\ T(p_{k02}) \end{array} \right) + \delta \left( \begin{array}{c} T(M_{12,1} J \partial_x (f_{k10})^T) \\ T(M_{22,1} J \partial_x (f_{k01})^T) \end{array} \right), \quad (3.10)$$

where $\tilde{M}_1 = M_{12,1} J \partial_x (f_{k01})^T + (M_{22,1} J \partial_x (f_{k01})^T)^T$, which are uniquely solvable on the following domain

$$A_1 = \{ \lambda \in A_0 : |\tilde{L}_{k0,1}| > \frac{\gamma_1}{|k|^r}, \tilde{L}_{k1,1} > \frac{\gamma_1}{|k|^r} I_{m+2m_0} \}.$$
By (3.7), we have
\[ \bar{H}_2 = H_1 \circ \Phi^1_{F_1} = N_2(y, u, v, \lambda) + \delta^2 \bar{P}_2(x, y, u, v, \lambda, \varepsilon), \]
where
\[ N_2 = N_1 + \delta^2[R_1], \]
\[ \bar{P}_2 = \frac{1}{\delta^2}(R'_1 + \int_0^1 \{ R_{1,t}, F_1 \} \circ \Phi^1_{F_1} dt + \delta^2(\bar{P}_1 - R_1) \circ \Phi^1_{F_1}), \]
\[ R_{1,t} = t\delta^2 R_1 + (1 - t)R'_1 + (1 - t)\delta^2[R_1]. \]
It is easy to see that $[R_1]$ has critical point on $u$, due to the $T^{m_0}$—periodicity in $u$. Consider the following transformation
\[ \phi : x \rightarrow x, \quad y \rightarrow y + y_0, \quad v \rightarrow v + v_0, \quad u \rightarrow u, \]
where $y_0$ and $v_0$ are determined by the following equation:
\[ \delta \bar{M}_1 \left( \begin{array}{c} y_0 \\ v_0 \end{array} \right) + \delta^2 \left( \begin{array}{c} \frac{\partial_y h(y_0, v_0)}{\partial_v h(y_0, v_0)} \\ \frac{\partial_v h(y_0, v_0)}{\partial_v h(y_0, v_0)} \end{array} \right) = \delta^2 \left( \begin{array}{c} \frac{\partial_y [R_1]}{\partial_v [R_1]} \\ \frac{\partial_v [R_1]}{\partial_v [R_1]} \end{array} \right). \]
Here and below, denote $[R_i]_2 = O(\| (y, v) \|)^2). Then
\[ H_2 = N_2(y, u, v, \lambda) + \delta^2 P_2(x, y, u, v, \lambda, \varepsilon), \quad (3.11) \]
where
\[ N_2 = \langle \omega_2, y \rangle + \frac{\delta}{2} \left( \begin{array}{c} y \\ v \end{array} \right), M_2 \left( \begin{array}{c} y \\ v \end{array} \right) + \delta^2 h_2 + \delta^2[R_1], \]
\[ \omega_2 = \omega + \delta \bar{M}_1 \left( \begin{array}{c} y_0 \\ v_0 \end{array} \right) + \delta^2 \left( \begin{array}{c} \frac{\partial_y h(y_0, v_0)}{\partial_v h(y_0, v_0)} \\ \frac{\partial_v h(y_0, v_0)}{\partial_v h(y_0, v_0)} \end{array} \right) + \delta^2 \left( \begin{array}{c} \frac{\partial_y [R_1]}{\partial_v [R_1]} \\ \frac{\partial_v [R_1]}{\partial_v [R_1]} \end{array} \right), \]
\[ \bar{M}_2 = \bar{M}_1 + \delta^2 \left( y, v \right) h_1, \]
\[ h_2 = O(\| K_0 \left( \begin{array}{c} y \\ v \end{array} \right) \|^3), \]
\[ P_2 = \bar{P}_2 \circ \phi + \left( \begin{array}{c} y \\ v \end{array} \right), \partial^2_{(y,v)}[R_1] \left( \begin{array}{c} y_0 \\ v_0 \end{array} \right). \]
Moreover,
\[ |P_2| \leq c\delta^{\#}. \]

Here and below, we denote $c$ the positive constant independent of the iteration process. Generally, the $k$—th KAM step state as follows, where $k$ is a given constant. After $k$ KAM steps, we get
\[ H_k = N_k(y, u, v, \lambda) + \delta^2 P_k(x, y, u, v, \lambda, \varepsilon), \quad (3.12) \]
\[ N_\kappa = (\omega_\kappa, y) + \frac{\delta}{2} \left( \begin{array}{c} y \\ v \end{array} \right), M_\kappa \left( \begin{array}{c} y \\ v \end{array} \right) + \delta^2 h_\kappa + \delta^2 [R_1]_2 + \cdots + \delta^2 [R_\kappa]_2, \]

Denote \( \tilde{M}_\kappa = \left( \begin{array}{ccc} \tilde{M}_{11,\kappa} & \tilde{M}_{12,\kappa} \\ \tilde{M}_{21,\kappa} & \tilde{M}_{22,\kappa} \end{array} \right) \), where \( \tilde{M}_{11,\kappa}, \tilde{M}_{12,\kappa}, \tilde{M}_{21,\kappa}, \tilde{M}_{22,\kappa} \) are \( m \times m \), \( m \times m_0 \), \( m_0 \times m \), \( m_0 \times m_0 \) matrices, respectively. Let \( \bar{M}_\kappa = \left( \begin{array}{ccc} \bar{M}_{11,\kappa} & \bar{M}_{12,\kappa} \\ \bar{M}_{21,\kappa} & \bar{M}_{22,\kappa} \end{array} \right), \)

where \( \tilde{M}_{1,\kappa} = 0, \tilde{M}_{2,\kappa} = 0, \tilde{M}_{3,\kappa} = 0, \tilde{M}_{4,\kappa} = 0, \tilde{M}_{5,\kappa} = 0 \) with obvious dimension. Let \( M_\kappa = \left( \begin{array}{ccc} M_{11,\kappa} & M_{12,\kappa} \\ M_{21,\kappa} & M_{22,\kappa} \end{array} \right), \)

where \( M_{11,\kappa} = \tilde{M}_{11,\kappa} + \delta \partial_z \partial_y ([R_1]_2 + \cdots + [R_\kappa]_2), M_{12,\kappa} = \tilde{M}_{12,\kappa} + \delta \partial_z \partial_y ([R_1]_2 + \cdots + [R_\kappa]_2), M_{21,\kappa} = \tilde{M}_{21,\kappa} + \delta \partial_z \partial_y ([R_1]_2 + \cdots + [R_\kappa]_2), M_{22,\kappa} = \tilde{M}_{22,\kappa} + \delta \partial_z \partial_y ([R_1]_2 + \cdots + [R_\kappa]_2). \)

Rewrite (3.12) as follows:

\[ H_\kappa = N_\kappa (y, u, v, \lambda) + \delta^2 P_\kappa (x, y, u, v, \lambda) \]

\[ N_\kappa = (\omega_\kappa, y) + \frac{\delta}{2} \left( \begin{array}{c} y \\ z \end{array} \right), M_\kappa \left( \begin{array}{c} y \\ z \end{array} \right) + \delta^2 h_\kappa. \]

Write, for \(|i| + |j| \leq 2,\)

\[ P_\kappa = \sum_k p_{kij} y^j z^k e^{\sqrt{-1} \langle k, x \rangle}, \]

\[ R_\kappa = \sum_{|k| \leq K_\kappa} p_{kij} y^j z^k e^{\sqrt{-1} \langle k, x \rangle}, \]

\[ P_\kappa - R_\kappa = \sum_{|k| > K_\kappa} p_{kij} y^j z^k e^{\sqrt{-1} \langle k, x \rangle}. \]

Improve the order of \( P_\kappa \) by the symplectic transformation \( \Phi^1_{F_\kappa} \), where

\[ F_\kappa (x, y, z, \lambda) = \sum_{0 < |k| \leq K_\kappa \atop |i| + |j| \leq 2} f_{kij} y^j z^k e^{\sqrt{-1} \langle k, x \rangle} \]

that satisfies

\[ \{ N_\kappa, F_\kappa \} + \delta^2 (R_\kappa - [R_\kappa]) - R'_\kappa = 0, \]

\[ R'_\kappa = \partial_z h_\kappa J \partial_z F_\kappa + \langle y, M_{12,\kappa} J \Delta_0 \rangle + \langle z, M_{22,\kappa} J \Delta_0 \rangle, \]

\[ \Delta_0 = \langle y, \partial_z f_{k20} \rangle + \langle y, \partial_z f_{k11} \rangle + \langle z, \partial_z f_{k02} \rangle, \]

\[ [R_i] = \int_{T^m} R_i (x, y, z, \lambda, \varepsilon) dx, \quad 1 \leq i \leq \kappa. \]
Using (3.15) and comparing coefficients, we obtain the following nonlinear homological equations

\[ L_{k0,k}f_{k00} = p_{k00}, \quad (3.16) \]
\[ L_{k1,k} \begin{pmatrix} f_{k10} \\ f_{k01} \end{pmatrix} = \begin{pmatrix} p_{k10} \\ p_{k01} \end{pmatrix} + \delta \left( \begin{array}{c} M_{12,k} J \partial_z f_{k00} \\ M_{22,k} J \partial_z f_{k00} \end{array} \right), \quad (3.17) \]
\[ L_{k2,k} \begin{pmatrix} T(f_{k20}) \\ T(f_{k11}) \\ T(f_{k02}) \end{pmatrix} = \begin{pmatrix} T(p_{k20}) \\ T(p_{k11}) \\ T(p_{k02}) \end{pmatrix} + \delta \left( \begin{array}{c} T(M_{12,k}) J \partial_z (f_{k10})^T \\ T(M_{22,k}) J \partial_z (f_{k00})^T \end{array} \right), \quad (3.18) \]

where \( M_k = M_{12,k} J \partial_z (f_{k10})^T + (M_{22,k} J \partial_z (f_{k00})^T) \), which are uniquely solvable on the following domain

\[ \Lambda_k = \{ \lambda \in \Lambda_{k-1} : \| \bar{L}_{k0,k} \| > \frac{\gamma_u}{|k|^\tau}, \bar{L}_{k1,k}^* \bar{L}_{k1,k} > \gamma_u |k|^\tau I_{m+2m_0}, \]
\[ \bar{L}_{k2,k}^* \bar{L}_{k2,k} > \frac{\gamma_u}{|k|^\tau} I_{m^2+2mm_0+4m^2}, \text{ for all } 0 < |k| \leq K_k \}. \]

Let \( u_0 \) be the critical point of \( [R] = [R_1]_2 + \cdots + [R_{k+1}]_2 \). Consider the following transformation

\[ \phi : \ x \to x, \ y \to y + y_0, \ v \to v + v_0, \ u \to u, \]

where \( y_0 \) and \( v_0 \) are determined by the following equation:

\[ \delta \bar{M}_k \left( \begin{array}{c} y_0 \\ v_0 \end{array} \right) + \delta^2 \left( \begin{array}{c} \partial_y h(y_0, v_0) \\ \partial_v h(y_0, v_0) \end{array} \right) = \delta^2 \left( \begin{array}{c} \partial_y [R_1] \\ \partial_v [R_1] \end{array} \right). \]

Then

\[ H_{k+1} = H_k \circ \Phi_{F_k} \circ \phi = N_{k+1}(y, u, v, \lambda) + P_{k+1}(x, y, u, v, \lambda, \varepsilon), \]

where

\[ N_{k+1} = (\omega_{k+1}, y) + \frac{\delta}{2} \left( \begin{array}{c} y \\ v \end{array} \right), \]
\[ \omega_{k+1} = \omega_k + \delta \bar{M}_k \left( \begin{array}{c} y_0 \\ v_0 \end{array} \right) + \delta^2 \left( \begin{array}{c} \partial_y h_k(y_0, v_0) \\ \partial_v h_k(y_0, v_0) \end{array} \right) + \delta^2 \left( \begin{array}{c} \partial_y [R_k] \\ \partial_v [R_k] \end{array} \right), \]
\[ M_{k+1} = M_k + \delta^2 \partial_{(y,z)} \nabla_k + \delta^2 \partial_{(x,z)} [R], \]
\[ P_{k+1} = R'_k \circ \phi + \int_0^1 \{ R_{k,t}, F_k \} \circ \Phi_{F_k} \circ \phi dt + (P_k - R_k) \circ \Phi_{F_k} \circ \phi \]
\[ + \left( \begin{array}{c} y \\ u \\ v \end{array} \right), \partial_{(y,z)} [R] \left( \begin{array}{c} y_0 \\ 0 \\ v_0 \end{array} \right), \]
\[ R_{k,t} = tR_k + (1-t)R'_k + (1-t)[R_k]. \]
Hence

$$|P_{\kappa+1}| \leq c_0 \delta^{\frac{1}{2} + \frac{1}{N}} (\frac{1}{\delta})^{\kappa+1}, |l| \leq d.$$  

Therefore, after \( \kappa \) KAM steps, the new Hamiltonian reads as

$$H_{\kappa+1} = N_{\kappa+1} + \delta^2 P_{\kappa+1},$$  

(3.19)

where

$$N_{\kappa+1} = \langle \omega_{\kappa+1}, y \rangle + \frac{\delta}{2} \left( \begin{array}{c} y \\ v \end{array} \right), \tilde{M}_{\kappa+1} \left( \begin{array}{c} y \\ v \end{array} \right) + \delta^2 h_{\kappa+1}$$

$$+ \delta^2 [R_1]_2 + \delta^2 [R_2]_2 + \cdots + \delta^2 [R_\kappa]_2,$$

$$\omega_{\kappa+1} = \omega_\kappa + \delta M_{\kappa} \left( \begin{array}{c} y_0 \\ v_0 \end{array} \right) + \delta^2 \left( \frac{\partial_y h_\kappa(y_0, v_0)}{\partial_r h_\kappa(y_0, v_0)} \right) + \delta^2 \left( \frac{\partial_v [R_\kappa]}{\partial_r [R_\kappa]} \right),$$

$$\tilde{M}_{\kappa+1} = \tilde{M}_\kappa + \delta^2 \partial_{\langle y, v \rangle} h_{\kappa}.$$  

Let

$$\tilde{g} = \delta^2 [R_1]_2 + \delta^2 [R_2]_2 + \cdots + \delta^2 [R_\kappa]_2$$

$$= \delta^2 \sum_{j_1} \delta^{2\frac{1}{\delta} + j_1} [R_1]_2^{(j_1)} + \delta^2 \sum_{j_2} \delta^{2\frac{1}{\delta} + j_2} [R_2]_2^{(j_2)} + \cdots$$

$$+ \delta^2 \sum_{j_\kappa} \delta^{2\frac{1}{\delta} + (\frac{1}{N})^{\kappa+1} + j_\kappa} [R_\kappa]_2^{(j_\kappa)}.$$  

**Definition 3.1.** If the following two hold:

1. At critical points of \( \tilde{g} \), \( (y_0, u_0, v_0) \),

$$\det \partial_u^a \partial_v^{\kappa-a+1} \tilde{g} = 0;$$

2. At critical points of \( \tilde{g} \), \( (y_0, u_0, v_0) \), there is a constant \( \tilde{a}_0 > 0 \), such that

$$|\det \partial_u^a \partial_v^{\kappa-a} \tilde{g}| \geq \tilde{a}_0,$$

then \( \tilde{g} \) is called a--order nondegenerate at \( (y_0, u_0, v_0) \).

**Remark 12.** Since \( \tilde{P}(x, y, u, v) \) is \( \kappa \)-order nondegenerate, at relative critical point \( (y_0, u_0, v_0) \) det \( \partial_u^a \tilde{P}_\kappa \)(\( (y_0, u_0, v_0, 0) \) \neq 0, which implies that \( \tilde{g} \) is \( \kappa \)-order nondegenerate, where \( 0 < a \leq \kappa \). And since \( \tilde{g} \) is \( T^{\kappa_0} \) periodic in \( u \), it has \( 2^{\kappa_0} \) critical points via the high order nondegeneracy and Morse theory (3.19).

**Remark 13.** Assumption (2) in definition 3.1 is equivalent to the following (\( \Theta 1 \)).

(\( \Theta 1 \)) At critical point of \( \tilde{g} \), \( (y_0, u_0, v_0) \), there exists a constant \( c \) > 0 such that the minimum \( \lambda_{\min}^\kappa(\omega) \) among absolute values of all eigenvalues of \( \partial_u^a \tilde{g} \) satisfies \( |\lambda_{\min}^\kappa| \geq c e^a \) for all \( \omega \in \Lambda(g, G) \).
Denote $\delta$. Then the re-scaled Hamiltonian reads
\[ H(x, y, u, v) = N(y, u, v) + \delta^{\alpha+1} \tilde{P}(x, y, u, v, \varepsilon), \]
where
\[ N = \langle \omega_{\kappa+1}, y \rangle + \frac{\delta}{2} \left( \begin{array}{c} y \\ v \end{array} \right), \tilde{M}_{\kappa+1} \left( \begin{array}{c} y \\ v \end{array} \right) \rangle + \delta^2 h_{\kappa} + \frac{\delta^\alpha}{2} \langle u, Vu \rangle + \delta^\alpha O(|u|^{3}), \]
and
\[ \delta^{\alpha+1} \tilde{P} = \delta^{\alpha+1} P(x, y, u, v, \varepsilon) + O(\delta^{\alpha+1}), \]
where $x \in T^n, y \in R^m, u, v \in R^{m_0}, 1 \leq \alpha \leq \kappa$. In the above, all $\lambda-$dependence is of class $C^{l_0}$ for some $l_0 \geq d$.

Next we should raise the order of $\tilde{P}$ by performing finite times quasilinear KAM steps. Let $\tilde{\tau}$ be the smallest integer such that $\lceil \frac{\delta}{2} + \frac{3\alpha+1}{12} \rceil \geq \frac{3\alpha+1}{2}$, where $\alpha$ is a constant. After $\tilde{\tau}$ KAM steps mentioned as above, at each critical point, we obtain the following
\[ H_{\tilde{\tau}}(x, y, u, v) = \langle \omega_{\kappa}, y \rangle + \frac{\delta}{2} \left( \begin{array}{c} y \\ v \end{array} \right), \tilde{M}_{\tilde{\tau}} \left( \begin{array}{c} y \\ v \end{array} \right) \rangle + \delta^2 h_{\tilde{\tau}} + \frac{\delta^\alpha}{2} \langle u, V_{\tilde{\tau}}(\lambda)u \rangle + \delta^{\alpha+1} \tilde{P}(x, y, u, v), \]
up to a constant, where
\[ V_{\tilde{\tau}} = V + \delta^\alpha h_{\tilde{\tau}}, \quad \tilde{u}(u) = \tilde{u} + (\tilde{h} - \delta^\alpha h_{\alpha} u, u), \]
\[ \tilde{h} = \delta^{\alpha+1} \delta^{\alpha+1} \langle \tilde{h}, u \rangle \tilde{h}, \]
\[ \tilde{P} = \delta^{\alpha+1} P(x, y, u, v, \delta), \quad 1 \leq \alpha \leq \kappa, \]
with nonsingular $V_{\tilde{\tau}}$. But in each KAM step we have a similar hypothesis in form, $\delta K^{\alpha+1} = o(\gamma)$. And the assumption obviously holds for finite times KAM steps. Consider re-scaling $x \rightarrow x, y \rightarrow \delta^{\alpha+1} y, u \rightarrow u, v \rightarrow \delta^{\alpha+1} v, H \rightarrow \delta^{\alpha+1} H$. Then the re-scaled Hamiltonian reads
\[ H_{\alpha+1}(x, y, u, v) = \langle \omega_{\alpha+1}, y \rangle + \frac{\delta^{\alpha+1}}{2} \left( \begin{array}{c} y \\ v \end{array} \right), \tilde{M}_{\alpha+1} \left( \begin{array}{c} y \\ v \end{array} \right) \rangle + \delta^2 h_{\alpha+1} + \frac{\delta^{\alpha+1}}{2} \langle u, V_{\alpha+1}(\lambda)u \rangle + \delta^{\alpha+1} \tilde{P}(x, y, u, v). \]
Denote $\delta^{\alpha+1} = \delta$. Then we have
\[ H(x, y, u, v) = N(y, u, v) + P(x, y, u, v), \]
with
\[ N = \langle \omega_{\alpha}, y \rangle + \frac{\delta}{2} \left( \begin{array}{c} y \\ v \end{array} \right), \tilde{M}_{\alpha} \left( \begin{array}{c} y \\ v \end{array} \right) \rangle + \delta^2 h_{\alpha} + \frac{\delta}{2} \langle u, V_{\alpha}(\lambda)u \rangle + \delta \tilde{u}(u), \]
\[ P = \delta^2 \tilde{P}(x, y, u, v), \quad \tilde{u}(u) = O(|u|^{3}), \]
33
where \( x \in T^m, y \in R^m, u, v \in R^{m_0} \). In the above, all \( \lambda \)-dependence is of class \( C^{l_0} \) for some \( l_0 \geq d \).

Applying Theorem 2.1 to (3.22), the system admits a family of invariant tori. By Morse theory, there are \( 2^{m_0} \) critical points, and consequently it has \( 2^{m_0} \) families of resonant torus. This completes the proof of Theorem 1.2.

4 Example

Here we give two examples to show how the program mentioned in section 3 work.

Example 1. Consider the following Hamiltonian system

\[
H(\tilde{x}, \tilde{y}) = \langle \tilde{\omega}, \tilde{y} \rangle + \frac{\epsilon}{2} \langle \tilde{y}, M \tilde{y} \rangle + \epsilon^3 \cos(-\frac{x^2}{2}) + \epsilon^2 \cos(-\frac{x^2}{2}) \sin(-2x_1 + 2x_2) e^{-y_1 - 2y_2},
\]

(4.23)

where \( \tilde{x} = (x_1, x_2)^T, \tilde{y} = (y_1, y_2)^T, \tilde{\omega} = (\omega_1, 2\omega_1)^T, x_1, x_2 \in T^1, y_1, y_2 \in R^1, \omega_1 \in R \setminus \{0\} \) and \( M = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \).

Consider transformation \( \tilde{\phi}_g : (y_1, y_2) \mapsto \begin{pmatrix} y \\ v \end{pmatrix}, (x_1, x_2) \mapsto \begin{pmatrix} x \\ u \end{pmatrix}, \) where

\[
\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -2 & 0 \\ 1 & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} v \\ y \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & -1 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix}.
\]

Denote \( \omega = -\omega_1 \). Then Hamiltonian (4.23) is changed to

\[
H(x, y, u, v) = \omega y + \frac{\epsilon}{2} v^2 + \epsilon^3 \cos u + \epsilon^2 \cos \omega \sin x e^y,
\]

which means that previous works do not apply to this system, since, first, \( \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \) is degenerate and, second, the perturbation \( P_1 = \epsilon^3 \cos u + \epsilon^2 \cos \omega \sin x e^y \) is 2-order nondegenerate perturbation.

Next, we will improve the order of \( P_1 \) by the symplectic transformation \( \Phi_{F_1}^1 \), where \( F_1(x, y, u, v) = a_1(y, u, v) \sin x + b_1(y, u, v) \cos x \) satisfies

\[
\{ N, F_1 \} + P_1 - [P_1] - P_1' = 0,
\]

(4.24)

\[
P_1' = \partial_u N \partial_v F_1 - \partial_v N \partial_u F_1,
\]

\[
N = \omega y + \frac{\epsilon}{2} v^2.
\]

Take \( F_1(x, y, u) = -\frac{\omega^2 \cos u e^x \sin^2 x}{\omega} \). Then

\[
H_2(x, y, u, v) = N_2(y, u) + P_1'(x, y, u, v, \epsilon) + \int_0^1 \{ (1-t) \{ N, F_1 \} + P_1, F_1 \} \circ \phi_{F_1}^t dt,
\]

34
Consider the following Hamiltonian system

\[ N_2(y, u) = N(y, u) + \varepsilon^3 \cos u, \]
\[ P_1'(x, y, u, v) = -\varepsilon^3 v \sin u \frac{e^y \cos x}{\omega}, \]
\[ P_2 = \int_0^1 \{ (1 - t) \{ N, F_1 \} + P_1, F_1 \} \circ \phi_t dt = O(\varepsilon^4). \]

In fact,

\[ R_t = (1 - t) \{ N, F_1 \} + P_1 \]
\[ = (1 - t)(-\varepsilon^2 \cos u e^y \sin x - \frac{\varepsilon^3 v \sin u e^y \cos x}{\omega}) + \varepsilon^3 \cos u + \varepsilon^3 \cos u e^y, \]
\[ \{ R_t, F_1 \} = \frac{\partial R_t \partial F_1}{\partial x} \frac{\partial R_t \partial F_1}{\partial y} + \frac{\partial R_t \partial F_1}{\partial u} \frac{\partial R_t \partial F_1}{\partial v} - \frac{\partial R_t \partial F_1}{\partial u} \frac{\partial R_t \partial F_1}{\partial v} \]
\[ = \frac{\varepsilon^4 t \cos^2 y \sin^2 \cos^2 2x}{\omega} + \varepsilon^5 (1 - t) \sin^2 \sin^2 (\cos 2x + 1). \]

Let \( F_2 = \frac{\varepsilon^2 v \sin x e^y \sin x}{\omega} \). Then \( \{ N_2, F_2 \} + P_2 - [P_2] - P_2' = 0 \), where \( P_2' = \partial_u N_2 \partial_1 F_2 - \partial_u N_2 \partial_u F_2 \). With the help of \( \Phi_{F_2} \), we have
\[ H_3(x, y, u, v) = N_2(y, u) + P_3(x, y, u, v), \]
where \( N_2 = \omega y + \frac{\varepsilon^2 v^2}{4} + \varepsilon^3 \cos u \), \( P_3 = O(\varepsilon^4) \). Therefore, using Theorem 2.1, there are two families of invariant tori for the Hamiltonian \( \{ 4.23 \} \) associated with relative critical points \( (y, u, v) = (y_0, 0, 0), (y_0, \pi, 0) \).

**Example 2.** Consider the following Hamiltonian system
\[ H(x, y) = \langle \hat{\omega}, \hat{y} \rangle + \frac{\varepsilon}{2} \langle \hat{y}, M \hat{y} \rangle + \varepsilon^4 \cos \left( \frac{x_1}{2} + \frac{\pi}{4} \right) \]
\[ + \varepsilon^2 \sin \left( -\frac{x_1}{2} \right) \sin (-2x_1 + x_2)e^{-y_1 - 2y_2}, \]
where \( \hat{\omega} = (\omega_1, \omega_2)^T, \hat{x} = (x_1, x_2)^T, \hat{y} = (y_1, y_2)^T, x_1, x_2 \in T^1, y_1, y_2 \in R^1, \omega_1 \in R^1 \setminus \{0\} \) and \( M = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 0 \end{pmatrix} \).

Consider the following transformation:
\[ \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -2 & 0 \\ 1 & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix}, \]
\[ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & -1 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix}. \]
Denote \( \omega = -\omega_1 \). Hamiltonian system \( \{ 4.27 \} \) is transformed to
\[ H(x, y, u, v) = \omega y + \frac{\varepsilon}{2} v^2 + \varepsilon^4 \cos(u + \frac{\pi}{4}) + \varepsilon^2 \sin u \sin x e^y, \]
which means the perturbation \( P_1 = \varepsilon^4 \cos(u + \frac{\pi}{4}) + \varepsilon^2 \sin u \sin x e^y \) is 3-order nondegenerate, i.e. previous works do not apply to this system. Let us prove the persistence of resonant tori for Hamiltonian system \( \{ 4.27 \} \) using Theorem 2.1.
Denote $F_1(x, y, u) = \frac{-\varepsilon^2 \sin \nu e^y \cos x}{\omega}$. Then
\[
\{N_1, F_1\} + P_1 - [P_1] - P_1' = 0,
\]
where
\[
P_1' = \partial_x N_1 \partial_x F_1 - \partial_y N_1 \partial_y F_1,
\]
\[
[P_1] = \int_0^{2\pi} P_1(x, y, u, \varepsilon)dx,
\]
\[
N_1 = \omega y + \frac{\varepsilon}{2} \nu^2.
\]
Therefore, under the symplectic transformation $\Phi_{F_1}$, we have
\[
H_2(x, y, u, v) = N_2(y, u) + P_2'(x, u, v) + \bar{P}_3(x, y, u, v),
\]
where
\[
N_2(y, u) = \omega y + \frac{\varepsilon}{2} \nu^2 + \varepsilon^3 \cos(u + \frac{t\pi}{4}),
\]
\[
P_2'(x, u, v) = \frac{\varepsilon^3 \nu \cos u e^y \cos x}{\omega},
\]
\[
\bar{P}_3 = \int_0^1 \{R_t, F_1\} \circ \phi_t \, dt,
\]
\[
R_t = (1 - t)\{N, F_1\} + P_1.
\]
Moreover,
\[
R_t = (1 - t)\{N, F_1\} + P \frac{\varepsilon^2 \sin \nu e^y \cos x}{\omega} + \varepsilon^4 \cos(u + \frac{t\pi}{4}),
\]
\[
\{R_t, F_1\} = \frac{\partial R_t \partial F_1}{\partial x} \partial y - \frac{\partial R_t \partial F_1}{\partial y} \partial x + \frac{\partial R_t \partial F_1}{\partial u} \partial v - \frac{\partial R_t \partial F_1}{}\partial v \partial u
\]
\[
= -(t \varepsilon^2 \sin \nu e^y \cos x + (1 - t) \varepsilon^3 \nu \cos u e^y \sin x) \frac{\varepsilon^2 \sin \nu e^y \cos x}{\omega} + (1 - t) \varepsilon^3 \cos u e^y \cos x \frac{\varepsilon^2}{\omega} + (1 - t) \frac{\varepsilon^4 \sin \nu e^y \cos x}{\omega} + O(\varepsilon^5).
\]
Hence $|\bar{P}_3| = -\frac{\varepsilon^4 \sin^2 \nu e^y \cos x}{\omega^2} + O(\varepsilon^5)$. Set $F_2 = \frac{-\varepsilon^2 \nu \cos u e^y \sin x}{\omega^2}$. Then
\[
\{N_2, F_2\} + P_2 - [P_2] - P_2' = o(\varepsilon^3),
\]
where
\[
[P_2] = \int_0^{2\pi} P_2(x, u, v)dx.
\]
\[ P'_2 = \partial_u N_2 \partial_v F_2 - \partial_v N_2 \partial_u F_2 \]
\[ = -\frac{\varepsilon^7 \sin(u + \frac{\nu \pi}{4}) \cos \nu \sin x}{\omega^2} + \frac{\varepsilon^4 \nu^2 \sin \nu \sin x}{\omega^2}. \]

With the aid of \( \Phi_{P_2} \), we have
\[ H_3(x, y, u, v) = N_3(y, u) + P_3(x, y, u, v), \]
where
\[ N_3 = \omega y + \frac{\varepsilon}{2} \nu^2 + \varepsilon^4 \cos(u + \frac{\nu \pi}{4}) - \frac{\varepsilon^4 \sin^2 u \nu \sin x}{2\omega}, \]
\[ P_3 = \frac{\varepsilon^4 \nu^2 \sin \nu \sin x}{\omega^2} + O(\varepsilon^5). \]

Let \( F_3 = -\frac{\varepsilon^4 \nu^2 \sin \nu \cos x}{\omega^2} \). Then
\[ \{ N_3, F_3 \} + P_3 - [P_3] - P'_3 = 0, \]
where
\[ P'_3 = \partial_u N_3 \partial_v F_3 - \partial_v N_3 \partial_u F_3 \]
\[ = \frac{\varepsilon^5 \nu^3 \cos \nu \cos x}{\omega^3} + \frac{\varepsilon^8(2\nu \sin(u + \frac{\nu \pi}{4}) \sin \nu \cos x \nu \omega + 2\nu^2 \sin^2 \nu \cos \nu \omega x)}{\omega^4}. \]

With the help of \( \Phi_{P_3} \), we have
\[ H_4(x, y, u, v) = N_4(y, u) + P_4(x, y, u, v), \]
where
\[ N_4 = \omega y + \frac{\varepsilon}{2} \nu^2 + \varepsilon^4 \cos(u + \frac{\nu \pi}{4}) - \frac{\varepsilon^4 \sin^2 u \nu \sin x}{2\omega}, \quad P_4 = O(\varepsilon^5). \]

Therefore, using Theorem 2.1, there are two families of invariant tori for the Hamiltonian (4.25).

**Remark 14.** The survival resonant tori are closely related to relative critical points. Relative critical points maybe drift when we do KAM iteration. In detail, both \( \cos(u + \frac{\nu \pi}{4}) \) and \( g = -\cos(u + \frac{\nu \pi}{4}) - \frac{2\sin^2 \nu \cos x}{2\omega} \) have two relative critical points. These critical points of \( \cos(u + \frac{\nu \pi}{4}) \) are \( u = -\frac{\nu \pi + 4\pi}{4} \) and \( -\frac{\nu \pi + 8\pi}{4} \), which are not relative critical point of \( g \) when \( \nu = 1 \), since \( \partial_u g(-\frac{\pi}{4} + \pi) = \sqrt{2} \cos \pi = -1 \) and \( \partial_u g(-\frac{\pi}{4} + 2\pi) = \sqrt{2} \cos 2\pi = 1 \).

**Acknowledgement**

We sincerely thank the anonymous referee for their most useful comments, which allowed us to vastly improve the exposition of our result.
A Some Properties

Property A.1. Coordinate transformation $\phi_g: I - I_0 = K_0 p, q = K_0^T \theta$ is symplectic.

Proof. In fact,

$$
\begin{pmatrix}
\theta \\
I - I_0
\end{pmatrix} = \begin{pmatrix}
(K_0^T)^{-1} & 0 \\
0 & K_0
\end{pmatrix}
\begin{pmatrix}
q \\
p
\end{pmatrix}.
$$

Then

$$
\begin{pmatrix}
((K_0^T)^{-1})^T & 0 \\
0 & K_0^T
\end{pmatrix}
\begin{pmatrix}
0 & I \\
-I & 0
\end{pmatrix}
\begin{pmatrix}
(K_0^T)^{-1} & 0 \\
0 & K_0
\end{pmatrix}
$$

$$
= \begin{pmatrix}
K_0^{-1} & 0 \\
0 & K_0^T
\end{pmatrix}
\begin{pmatrix}
0 & I \\
-I & 0
\end{pmatrix}
\begin{pmatrix}
(K_0^T)^{-1} & 0 \\
0 & K_0
\end{pmatrix}
$$

$$
= \begin{pmatrix}
0 & I \\
-I & 0
\end{pmatrix},
$$

which means that the coordinate transformation is symplectic.

Property A.2. Transformation $[3.3]$ is symplectic.

Proof. Let $p = \begin{pmatrix} y \\ v \end{pmatrix}$ and $q = \begin{pmatrix} x \\ u \end{pmatrix}$. With $p \rightarrow \varepsilon^\pm p$, $q \rightarrow q$, the motion equation of Hamiltonian system $[3.2]$ is changed to

$$
\begin{cases}
\dot{p} = \frac{\partial H_1(\varepsilon^\pm p, q)}{\partial q}, \\
\dot{q} = -\frac{\partial H_1(\varepsilon^\pm p, q)}{\partial p}.
\end{cases}
$$

Since

$$
H(p, q) = \varepsilon^\pm H_1(\varepsilon^\pm p, q),
$$

i.e., the symplectic structure is preserved under transformation $[3.3]$.

B Proof of Lemma 2.7

Directly,

$$
\mu_\nu = 64C_0 \left(64C_0 \mu_{\nu-2}\right)^{\frac{12}{13}}
$$

$$
= 64C_0 \left(64C_0 \left(64C_0 \mu_{\nu-3}\right)^{\frac{12}{13}}\right)^{\frac{12}{13}}
$$

$$
\ldots
$$

$$
= (64C_0)^{1 + \frac{12}{13} + \cdots + (\frac{12}{13})^\nu} \mu_0 \left(\frac{12}{13}\right)^\nu
$$

$$
= (64C_0)^{12(\frac{12}{13})^\nu-1} \mu_0 \left(\frac{12}{13}\right)^\nu,
$$

(B.26)

$$
\alpha_\nu = (64C_0)^{3(\frac{12}{13})^\nu-1} \mu_0 \left(\frac{12}{13}\right)^\nu,
$$

(B.27)
\[ s_\nu = \frac{1}{8} \mu_{\nu-1} \frac{1}{8} \mu_{\nu-2} s_{\nu-2} \]
\[ = \frac{1}{8} \mu_{\nu-1} \frac{1}{8} \mu_{\nu-2} - \frac{1}{8} \mu_{\nu-3} s_{\nu-3} \]
\[ = \frac{1}{8} s^\nu (\mu_{\nu-1} \mu_{\nu-2} \cdots \mu_0) s_0 \]
\[ = \frac{1}{8^\nu} (64C_0 \frac{(\frac{13}{12})^{\nu-1} - 1}{12} \mu_0 (\frac{13}{12})^{\nu-1} + 64C_0 \frac{(\frac{13}{12})^{\nu-2} - 1}{12} \mu_0 (\frac{13}{12})^{\nu-2} \cdots \mu_0) s_0 \]
\[ = \frac{1}{8^\nu} (64C_0 \frac{(\frac{13}{12})^{\nu-1} - 1}{12} + \frac{13}{12} \nu - 1 - \frac{13}{12} \nu - 2 - 1 + \frac{13}{12} \nu - 1 + (\frac{13}{12})^{\nu-2} \cdots + 1) \right) s_0 \]
\[ = \frac{1}{8^\nu} (64C_0) s^3 (\nu - 1) \mu_0 (\frac{13}{12})^{\nu-1} s_0. \quad \text{(B.28)} \]

Then
\[ K_+ = \left[ \left( -\frac{13}{12} \nu \right) - \log(64C_0) - (\frac{13}{12})^\nu \log \mu_0 \right] + 3^\eta \]
\[ = \left( -12 \frac{13}{12} \nu \log(64C_0) + 12 \log(64C_0) - (\frac{13}{12})^\nu \log \mu_0 \right] + 1) + 3^\eta \]
\[ = \left( \left( \frac{13}{12} \nu \right) (-12 \log(64C_0) - \log \mu_0) + 12 \log(64C_0) \right] + 1) + 3^\eta \quad \text{(B.29)} \]
\[ \geq \frac{13}{12} \nu \log \frac{1}{\mu_0} \]
\[ \geq 8(m + l_0)2^{\nu+2}. \]

Combining
\[ r_\nu - r_{\nu+1} = r_0 \left( 1 - \sum_{i=0}^{\nu-1} \frac{1}{2^{i+1}} \right) - r_0 \left( 1 - \sum_{i=0}^{\nu} \frac{1}{2^{i+1}} \right) = \frac{r_0}{2^{\nu+1}}, \quad \text{(B.30)} \]

we finish the verification of \((H1)\) for all \( \nu = 1, 2, \cdots \).

According to \((H2, 29)\), we have
\[ K^{2\chi_1}_{\nu+1} \leq 2^{2\chi_1} \left( \frac{13}{12} \right)^{3(2\chi_1)} (\log \frac{1}{\mu_0})^{3(2\chi_1)}. \]

Then, for small enough \( \mu_0 \),
\[ s_\nu K^{2\chi_1}_{\nu+1} \leq \frac{1}{8^\nu} (64C_0)^{\nu-1} (64C_0) s_0 (2(\frac{13}{12})^\nu \log \frac{1}{\mu_0})^{3(2\chi_1)} \]
\[ \leq \frac{1}{8^\nu} \mu_0 (\frac{13}{12})^{\nu-1} s_0 (2(\frac{13}{12})^\nu \log \frac{1}{\mu_0})^{3(2\chi_1)} \]
\[ \leq 2^{(2\chi_1)} (\frac{13}{12})^{3(2\chi_1)} \frac{1}{2^\nu} \mu_0 (\frac{13}{12})^{\nu-1} (\frac{13}{12})^{\nu-1} s_0 (\log \frac{1}{\mu_0})^{3(2\chi_1)} \]
\[ \leq 2^{(2\chi_1)} (\frac{13}{12})^{3(2\chi_1)} \frac{1}{2^\nu} (\mu_0)^{\nu} (\mu_0^{\nu}) s_0 (\log \frac{1}{\mu_0})^{3(2\chi_1)} \]

39
\[ \gamma_{\nu} = \gamma_0 \left( 1 - \sum_{i=0}^{\nu-1} \frac{1}{2^{i+1}} \right) = \frac{\gamma_0}{2^\nu}, \]  

(B.31)

we verify (H4) for \( \nu = 1, 2, \cdots \).

Since

\[ \Gamma_{\nu} = \Gamma_{\nu}(r_{\nu} - r_{\nu-1}) \leq \int_1^\infty t^\nu e^{-\frac{(r_{\nu} - r_{\nu-1})}{s}} dt \leq \left( 2^{\nu+6}e^{-\frac{\nu+6}{\nu}} + 2^{2(\nu+6)}e^{-\frac{\nu+6}{\nu}} + \cdots + 2^{(\nu+6)}e^{-\frac{\nu+6}{\nu}} \right) \leq c(\nu+6)e^{-\frac{\nu+6}{\nu}}, \]

it is clear that

\[ \mu_{\nu} \Gamma_{\nu}^3 < (64C_0) \left( \frac{\nu}{\nu_0} \right)^{(\frac{\nu}{\nu_0})-1} \mu_0 \left( \frac{\nu}{\nu_0} \right)^\nu \left( 2^{(\nu+6)}e^{-\frac{\nu+6}{\nu}} \right)^3. \]

Combining (B.31) and (B.32), assumption (H5) holds for \( \nu = 1, 2, \cdots \). Using (B.27) and (B.32), we finish the proof of (H6) for \( \nu = 1, 2, \cdots \). With (B.31) and (B.32), we verify (H7) for \( \nu = 1, 2, \cdots \). Combining (2.7), (2.11), (2.13) and (B.26), yield

\[ |\partial^j_{(y,z)}(\hat{h}_\nu - \hat{h}_0)| \leq \sum_{\nu} \mu_{\nu} \leq \mu_0^{\frac{j}{2}}, \]

which implies (H3) hold for \( \nu = 1, 2, \cdots \).

\section*{C Measure Estimate}

Theorem C.1. Let \( \Lambda_\ast = \bigcap_{\nu=0}^\infty \Lambda_{\nu} \). Assume (A2) hold. Then, for sufficiently small \( \delta \),

\[ |\Lambda_0 \setminus \Lambda_\ast| \rightarrow 0 \text{ as } \gamma_0 \rightarrow 0. \]

Proof. Let

\[ R_{\nu+1} = \{ \lambda \in \Lambda_{\nu}(\lambda) : \left| \hat{L}_{k0,\nu} \right| \leq \frac{\gamma_{\nu}}{|k|^\nu}, \hat{L}_{k1,\nu} \hat{L}_{k1,\nu} \leq \frac{\gamma_{\nu}}{|k|^\nu} |I_{m+2m_0}, \]

which is a measure estimate.
\[ L_{k,\nu}^* L_{k,\nu} \leq \frac{\gamma_{\nu}}{|k|^p} I_{m^2 + 2mm_0 + 4m_0^2}, \text{for all } K_\nu < |k| \leq K_{\nu+1} \]
\[ \subset S_1 \cup S_2 \cup S_3, \]

where \( S_1 = \{ \lambda \in \Lambda_\nu : 0 \leq |L_{k,\nu}| \leq \frac{\gamma_{\nu}}{|k|^p}, K_\nu < |k| \leq K_{\nu+1} \}, \)
\( S_2 = \{ \lambda \in \Lambda_\nu : 0 \leq L_{k,\nu}^* L_{k,\nu} \leq \frac{\gamma_{\nu}}{|k|^p}, K_\nu < |k| \leq K_{\nu+1} \}, \)
\( S_3 = \{ \lambda \in \Lambda_\nu : 0 \leq L_{k,\nu}^* L_{k,\nu} \leq \frac{\gamma_{\nu}}{|k|^p}, K_\nu < |k| \leq K_{\nu+1} \}. \)

Let \( \zeta = \frac{k}{|k|} \in S^m, \) where \( S^m \) is a \( m \)-dimensional ball. For given \( \lambda_0 \in \Lambda_\nu, \)
denote \( \Omega_\nu(\lambda_0) = (\nu, \nu, \ldots, \nu) \), \( \Omega_\nu(\lambda_0 + t) = (\nu, \nu, \ldots, \nu) \). Using Taylor series,
\[ L_{k,\nu} = |k|^p \Omega_\nu(\lambda_0) \lambda. \]

Let \( Q_{\lambda_0, \nu} = (q_{ij})_{i \times \nu}, \) where \( q_{1,1} = q_{2,2} = \cdots = q_{m,m} = 1, \) \( \tau_p = \tau_0, 1 \leq p, q \leq \nu \) and other elements of \( Q_{\lambda_0, \nu} \) are equal to 0. Since \( \text{rank} \Omega_\nu(\lambda_0) = m \) for \( \lambda_0 \in \Lambda_\nu, \)
i.e. condition (42), there is an matrix \( Q_{\lambda_0, \nu} = (q_{ij})_{i \times \nu} \) such that \( \Omega_\nu(\lambda_0)Q_{\lambda_0, \nu} = (A_\nu(\lambda_0)Q_{\lambda_0, \nu}) \), where \( A_\nu(\lambda_0) = (a_{ij})_{m \times m} \) is nonsingular. Denote \( \Lambda_{\lambda_0, \nu} \) the neighborhood of \( \lambda_0 \) and \( \Lambda_{\lambda_0, \nu} \) the closure of \( \Lambda_{\lambda_0, \nu}. \) Then \( \text{det} \ A_\nu(\lambda) \neq 0 \) for \( \lambda \in \Lambda_{\lambda_0, \nu}. \) Therefore, there is an orthogonal matrix \( Q_{\lambda_0, \nu} \) such that \( \Omega_\nu(\lambda)Q_{\lambda_0, \nu} = (A_\nu(\lambda), B_\nu(\lambda)) \) for \( \lambda \in \Lambda_{\lambda_0, \nu}, \) where \( \text{det} \ A_\nu(\lambda) \neq 0 \) on \( \Lambda_{\lambda_0, \nu}. \) Denote the eigenvalues of \( (A_\nu(\lambda)A_\nu^*(\lambda) + B_\nu(\lambda)B_\nu^*(\lambda)) \) by \( \lambda_{1,\nu} \leq \cdots \leq \lambda_{n,\nu}. \) Since \( \text{rank} (A_\nu(\lambda)A_\nu^*(\lambda) + B_\nu(\lambda)B_\nu^*(\lambda)) = \text{rank} (A_\nu(\lambda), B_\nu(\lambda)) \) (45), there is a unitary \( U_\nu \) and a real diagonal \( V_\nu = \text{diag} (\lambda_{1,\nu}, \cdots, \lambda_{n,\nu}) \) such that \( (A_\nu(\lambda)A_\nu^*(\lambda) + B_\nu(\lambda)B_\nu^*(\lambda)) = U_\nu V_\nu U_\nu^*. \) Therefore, using Poincaré separation theorem,
\[ \zeta^*(A_\nu(\lambda)A_\nu^*(\lambda) + B_\nu(\lambda)B_\nu^*(\lambda))\zeta = \zeta^* U_\nu^* A_\nu(\lambda)A_\nu^*(\lambda) + B_\nu(\lambda)B_\nu^*(\lambda))U_\nu\zeta \]
\[ = \zeta^* U_\nu^* \text{diag}(\lambda_{1,\nu}, \cdots, \lambda_{n,\nu})U_\nu\zeta \]
\[ \geq \zeta^* U_\nu^* \hat{\lambda}_{1,\nu} U_\nu \zeta \]
\[ \geq \hat{\lambda}_{1,\nu}. \]

Since the nonzero eigenvalues of \( \begin{pmatrix} A_\nu^T(\lambda) \xi \xi^T A_\nu(\lambda) & A_\nu^T(\lambda) \xi \xi^T B_\nu(\lambda) \\ B_\nu^T(\lambda) \xi \xi^T A_\nu(\lambda) & B_\nu^T(\lambda) \xi \xi^T B_\nu(\lambda) \end{pmatrix} \) and \( \xi^T (A_\nu(\lambda)A_\nu^*(\lambda) + B_\nu(\lambda)B_\nu^*(\lambda)) \xi \) are the same, there is an unitary matrix \( U_\nu(\lambda) \) such that
\[ \begin{pmatrix} A_\nu^T(\lambda) \xi \xi^T A_\nu(\lambda) & A_\nu^T(\lambda) \xi \xi^T B_\nu(\lambda) \\ B_\nu^T(\lambda) \xi \xi^T A_\nu(\lambda) & B_\nu^T(\lambda) \xi \xi^T B_\nu(\lambda) \end{pmatrix} = U_\nu(\lambda) \text{diag}(0, \cdots, 0, \hat{\lambda}_\nu) U_\nu^*(\lambda), \]
where \( \hat{\lambda}_\nu = \zeta^* U_\nu^* \text{diag}(\lambda_{1,\nu}, \cdots, \lambda_{n,\nu})U_\nu\zeta \). Denote \( (U_\nu(\lambda)Q_{\lambda_0, \nu})_i \) the \( i \)-th row of \( U_\nu(\lambda)Q_{\lambda_0, \nu}. \) Therefore, \( |(U_\nu(\lambda)Q_{\lambda_0, \nu})_i| \geq \min_{1 \leq j \leq m} |\lambda_j|^{2N+2}. \)

Hence
\[ |L_{k,\nu}^* L_{k,\nu}| = |k|^p |\hat{\lambda}|^{2N+2} Q_{\lambda_0, \nu} Q_{\lambda_0, \nu}^{-1} \Omega \xi \xi^T \Omega Q_{\lambda_0, \nu} Q_{\lambda_0, \nu}^{-1} \hat{\lambda}| \]

41
\[ \begin{align*}
\hat{\lambda} &= |k|^2 \hat{\lambda}^T Q_{\lambda_0, \nu} \left( A^T(\lambda) \hat{A}(\lambda) A^T(\lambda) \hat{B}(\lambda) \right) Q^{-1}_{\lambda_0, \nu} \hat{\lambda} \\
&= |k|^2 \hat{\lambda}^T Q_{\lambda_0, \nu} U^*_\nu(\lambda) \text{diag}(0, \cdots, 0, \hat{\lambda}) U_\nu(\lambda) Q^{-1}_{\lambda_0, \nu} \hat{\lambda} \\
&\geq |k|^2 \hat{\lambda}_1 \left( \min_{1 \leq j \leq m} \hat{\lambda}_j \right)^{2N+2}.
\end{align*} \]

Then
\[ \{ \lambda \in \Lambda_\nu \cap \tilde{\Lambda}_{\lambda_0, \nu} : |\tilde{L}_{k_{0, \nu}}(\lambda)| \leq \frac{\gamma_\nu}{|k|^2}, K_\nu < |k| \leq K_{\nu+1} \} \]
\[ \leq \frac{1}{\lambda_1} D^{m-1} \frac{\gamma_\nu}{|k|^2}, \]
where \( D \) is the exterior diameter of \( \tilde{\Lambda}_{\lambda_0, \nu} \) with respect to the maximum norm, \( m \) is the dimension of \( \Lambda_\nu \). Further, there are finite sets, \( \tilde{\Lambda}_{\lambda_i, \nu}, 1 \leq i \leq \tilde{i} \), such that \( \Lambda_\nu \subseteq \bigcup_{i=1}^{\tilde{i}} \tilde{\Lambda}_{\lambda_i, \nu} \) and
\[ |\tilde{L}_{k_{0, \nu}}(\lambda)| > |k|^2 \hat{\lambda}_1^i(\lambda) \left( \min_j \hat{\lambda}_j \right)^{2N+2} \text{ for } \lambda \in \tilde{\Lambda}_{\lambda_i, \nu}, \]
where \( \hat{\lambda}_1^i(\lambda) \) the minimum eigenvalue of \( \Omega_\nu^*(\lambda) \Omega_\nu(\lambda) \) on \( \tilde{\Lambda}_{\lambda_i, \nu} \). Therefore,
\[ |S_1| = \{ \lambda \in \Lambda_\nu : |\tilde{L}_{k_{0, \nu}}(\lambda)| \leq \frac{\gamma_\nu}{|k|^2}, K_\nu < |k| \leq K_{\nu+1} \} < c D^{m-1} \frac{1}{|k|^2}, \]
where \( c \) depends on \( \Lambda, D, m \) and \( \hat{\lambda}_1^i, 1 \leq i \leq \tilde{i} \).

Denote
\[ B_\nu = \begin{pmatrix}
0 & (JM_{21, \nu})^T \otimes I_m \\
0 & -(M_{22, \nu}J) \otimes I_m \\
0 & 0 & -I_{2m_0} \otimes (2M_{12, \nu}J) \\
0 & 0 & -(M_{22, \nu}J) \otimes I_{2m_0} - I_{2m_0} \otimes (M_{22, \nu}J)
\end{pmatrix}. \]

Let \( \tilde{B}_\nu = -\sqrt{-1}(k, \omega_\nu) I_{m^2 + 2m m_0 + 4m_0^2} + B_\nu + B_\nu^* \sqrt{-1}(k, \omega_\nu) I_{m^2 + 2m m_0 + 4m_0^2} + \delta B_\nu B_\nu \).

Combining Poincaré separation theorem and eigenvalue perturbation theorem (1E), for sufficiently small \( \delta \), we have
\[ \tilde{L}_{k_{2, \nu}} \tilde{L}_{k_{2, \nu}} = |k|^2 (\langle \cdot, \omega_\nu \rangle)^2 I_{m^2 + 2m m_0 + 4m_0^2} + \delta \tilde{B}_\nu \]
\[ \geq \frac{|k|^2}{2} \hat{\lambda}_1^i \left( \min_j \hat{\lambda}_j \right)^{2N+2} I_{m^2 + 2m m_0 + 4m_0^2} + \delta \tilde{B}_\nu \]
\[ \geq \frac{|k|^2}{4} \hat{\lambda}_1^i \left( \min_j \hat{\lambda}_j \right)^{2N+2} I_{m^2 + 2m m_0 + 4m_0^2}, \]
\[ \geq \frac{|k|^2}{4} \hat{\lambda}_1^i \left( \min_j \hat{\lambda}_j \right)^{2N+2} I_{m^2 + 2m m_0 + 4m_0^2}. \]
Therefore,

\[ |S_3| = |\{ \lambda \in \Lambda_{\nu} : \tilde{L}_{k_2,\nu} \tilde{L}_{k_2,\nu} \leq \frac{\gamma_\nu}{|k|^2} I_{m^2 + 2m_m + m_0^2}, K_\nu < |k| \leq K_{\nu + 1} \}| \]

\[ < cD^{m-1} \frac{\gamma_0}{|k|^{|N+1|}}. \]

Similarly,

\[ |S_2| = |\{ \lambda \in \Lambda_{\nu} : \tilde{L}_{k_1,\nu} \tilde{L}_{k_1,\nu} \leq \frac{\gamma_\nu}{|k|^2} I_{m + 2m_0}, K_\nu < |k| \leq K_{\nu + 1} \}| \]

\[ < cD^{m-1} \frac{\gamma_0}{|k|^{|N+1|}}. \]

Obviously, \(|R_{\nu + 1}| \leq \frac{\gamma_0}{|k|^{|N+1|}}. \) Thus

\[ \left| \bigcup_{\nu = 0}^{\infty} \bigcup_{K_\nu \leq |k| \leq K_{\nu + 1}} R_{\nu + 1} \right| \leq c \sum_{\nu = 0}^{\infty} \sum_{K_\nu \leq |k| \leq K_{\nu + 1}} \frac{\gamma_0}{|k|^{|N+1|}} \to 0 \text{ as } \gamma_0 \to 0. \]

\[ \square \]

References

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

[2] Arnold, V. (1963). Small denominators and problems of stability of motion in classical and celestial mechanics. Uspehi Mat. Nauk. 18, 91 - 192.

[3] Bernstein, D. and Katok, A. (1987). Birkhoff periodic orbits for small perturbations of completely integrable Hamiltonian systems with convex Hamiltonians. Invent. Math. 88(2), 225 - 241.

[4] Bourgain, J. (1997). On Melnikov’s persistency problem. Math. Res. Lett. 4, 445 - 458.

[5] Broer, H., Huitema, G. and Sevryuk, M. (1996). Quasi-periodic motions in families of dynamical systems, Lect. Notes Math. 1645, Springer-Verlag.

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

[7] Chow, S., Li, Y. and Yi, Y. (2002). Persistence of invariant tori on sub-manifolds in Hamiltonian system, J. Nonl. Sci. 12, 585 - 617.
[8] Cong, F., Küpper, T., Li, Y. and You, J. (2000). KAM-type theorem on resonant surfaces for nearly integrable Hamiltonian systems, J. Nonl. Sci. 10, 49 - 68.

[9] Corsi, L. and Gentile, G. (2017). Resonant tori of arbitrary codimension for quasi-periodically forced systems, NoDEA Nonlinear Differential Equations Appl. 24(1), 3 - 24.

[10] Eliasson, L. (1988). Perturbations of stable invariant tori for Hamiltonian systems. Ann. Sc. Norm. Super. Pisa Cl. Sci. 15, 115 - 147.

[11] Gentile, G. (2007). Degenerate lower-dimensional tori under the Bryuno condition, Ergod. Theory Dyn. Syst. 27(2), 427 - 457.

[12] Guardia, M., Kaloshin, V. and Zhang, J. (2019). Asymptotic density of collision orbits in the restricted circular planar 3 body problem, Arch. Ration. Mech. Anal. 233, 799 - 836.

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

[14] Hirsch, M. (1976). Differential Topology, Springer-Verlag, New York.

[15] Horn, R. and Johnson, C. Matrix analysis, Cambridge University Press, New York, 2013.

[16] Kaloshin, V., and Zhang, K. (2018). Density of convex billiards with rational caustics, Nonlinearity 31, 5214 - 5234.

[17] Kappeler, T. and Pöschel, J. (2003). KdV & KAM, In: Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics, vol. 45. Springer, Berlin.

[18] Kolmogorov, A. (1954). On quasiperiodic motions under small perturbations of the Hamiltonian, Dokl. Akad. Nauk USSR 98, 527 - 530.

[19] Kuksin, S. Nearly Integrable Infinite-Dimensional Hamiltonian Systems, Springer, Berlin 1993.

[20] Li, Y. and Yi, Y. (2003). A quasiperiodic Poincaré’s theorem, Math. Ann. 326, 649 - 690.

[21] Li, Y. and Yi, Y. (2004). Persistence of lower dimensional tori of general types in Hamiltonian systems, Trans. Amer. Math. Soc. 357, 1565 - 1600.

[22] Melnikov, V. (1965). On some cases of the conservation of conditionally periodic motions under a small change of the Hamiltonian function, Sov. Math. Dokl. 6, 1592 - 1596.

[23] Milnor, J. (1963). Morse Theory, Princeton Univ. Press.
[24] Moser, J. (1962). On invariant curves of area preserving mappings of an annulus, Nathr. Akad. Wiss. Gott. Math. Phys. K1,2, 1 - 20.

[25] Moser, J. (1967) Convergent series expansions for quasiperiodic motions, Math. Ann. 169, 136 - 176.

[26] Meyer, K., Palacián, J. F. and Yanguas, P. (2011). Geometric Averaging of Hamiltonian Systems: Periodic Solutions, Stability, and KAM Tori, SIAM J. Appl. Dyn. Syst., 10(3), 817 - 856.

[27] Pöschel, J. (1989). On the elliptic lower dimensional tori in Hamiltonian systems, Math. Z. 202, 559 - 608.

[28] Poincaré, H. (1892, 1893, 1899). Les Méthodes Nouvelles de la Mécaniques Céleste, I - III, Gauthier-Villars. (The English translation: New methods of celestial mechanics, AIP Press,Williston, 1992.)

[29] Palacián, J., Sayas, F. and Yanguas, P. (2013). Regular and singular reductions in the spatial three-body problem, Qual. Theory Dyn. Syst. 12, 143 - 182.

[30] Palacián, J., Vidal, C., Vidarte, J. and Yanguas, P. (2017). Periodic solutions and KAM tori in a triaxial potential, SIAM J. Appl. Dyn. Syst. 16, no. 1, 159 - 187.

[31] Qian, W., Li, Y. and Yang, X. (2019). Multiscale KAM theorem for Hamiltonian systems, Journal of Differential Equations, 266, Issue 1, 70 - 86

[32] Qian, W., Li, Y. and Yang, X. (2019). The isoenergetic KAM-type theorem at resonant case for nearly integrable Hamiltonian systems, J. Appl. Anal. Comput., 9, Issue 5, 1616 - 1638

[33] Qian, W., Li, Y. and Yang, X. (2020). Persistence of Lagrange invariant tori at tangent degeneracy, Journal of Differential Equations, 268, Issue 9, 5078 - 5112

[34] Qian, W., Li, Y. and Yang, X. (2020). Melnikov’s conditions in matrices, J. Dynam. Differential Equations 32, Issue 4, 1779 - 1795.

[35] Sevryuk, M. (2006). Partial preservation of frequencies in KAM theory, Nonlinearity, 5, 1099 - 1140.

[36] Treshchëv, D. (1989). Mechanism for destroying resonance tori of Hamiltonian systems, Mat. USSR Sb. 180, 1325 - 1346.

[37] Xu, L., Li, Y. and Yi, Y. (2017). Lower-dimensional tori in multi-scale, nearly integrable Hamiltonian systems, Ann. Henri Poincaré 18 no.1, 53 - 83.

[38] Xu, L., Li, Y., Yi, Y. (2018). Poincaré-Treshchëv mechanism in multi-scale, nearly integrable Hamiltonian systems, J. Nonl. Sci. 28 no. 1, 337 - 369.
[39] You, J. (1999). Perturbations of lower-dimensional tori for Hamiltonian systems. J. Differ. Equ. 152, 1 - 29.