Persistence of Hyperbolic Tori in Generalized Hamiltonian Systems*

Zhenxin Liu†, Dalai Yihe and Qingdao Huang
College of Mathematics, Jilin University, Changchun 130012, P. R. China

Abstract

In this paper we prove the persistence of hyperbolic invariant tori in generalized Hamiltonian systems, which may admit a distinct number of action and angle variables. The systems under consideration can be odd dimensional in tangent direction. Our results generalize the well-known results of Graff and Zehnder in standard Hamiltonians. In our case the unperturbed Hamiltonian systems may be degenerate. We also consider the persistence problem of hyperbolic tori on sub-manifolds.

Keywords: hyperbolic invariant tori; KAM theorem; generalized Hamiltonian systems

1 Introduction and Main Result

According to the celebrated KAM (Kolmogorov-Arnold-Moser) theory, we know that most of invariant tori of integrable Hamiltonians persist under a small perturbation. In their case, the Hamiltonian is standard and the dimension of invariant tori equals the degree of freedom, i.e., the "highest dimensional tori". However, some Hamiltonian systems' highest dimensional tori cannot survive the perturbations, but some tori which have lower dimension can be persisted under small perturbations, which are called lower dimensional invariant tori. In 1965, Melnikov[1] formulated a KAM type persistence result for elliptic lower dimensional tori of integrable Hamiltonian systems under so-called Melnikov's non-resonance condition. But the complete proof of his result was carried out more than twenty years later by Eliasson, Kuksin, and Pöschel (see [2]–[4]). For the persistence of hyperbolic lower dimensional tori, in 1974, Graff[5] considered the following Hamiltonian system:

$$H = e + \langle \omega_0, y\rangle + \frac{1}{2}\langle y, Ay\rangle + \frac{1}{2}\langle z, Mz\rangle + P(x, y, z), \tag{1.1}$$

where $(x, y, z) \in T^n \times R^n \times R^{2m}$, $M = \begin{pmatrix} O & B \\ B^\top & O \end{pmatrix}$, $\omega_0 \in R^n$ is a fixed Diophantine toral frequency, and P is a small perturbation. The persistence of the unperturbed Diophantine hyperbolic torus $T^n \times \{0\} \times \{0\}$ was shown as well as the preservation of the toral frequency ω_0. Zehnder in [6], using generalized implicit function theorem, proved the same result. More recently, Li and Yi[7] generalized the results of Graff and Zehnder on the persistence of hyperbolic invariant tori in Hamiltonian systems by allowing the degeneracy of the unperturbed Hamiltonians and they obtain the preservation of part or full components of frequencies. They adopted the Fourier series expansion for normal form N, which is a new technique.

Due to important technical reasons, the development of KAM theory for odd dimensional systems has been considered as a challenging problem. Li and Yi[8] solve the delicate problem by considering generalized Hamiltonian systems which preserve a prescribed Poisson

*This paper has been published in: *Northeast. Math. J.* **21** (4) (2005), 447-464.
†Corresponding author: zxliu@email.jlu.edu.cn (Zhenxin Liu).

structure instead of volume. In their case, the Hamiltonians considered may admit distinct number of action and angle variables and more important, which can be odd dimensional. Motivated by their work, in this paper, we show that the Graff-Zehnder result also holds in generalized Hamiltonian systems.

We consider the following parameter-dependent Hamiltonian system:

$$H = e(\lambda) + \langle \Omega(\lambda), y \rangle + \frac{1}{2} \left\langle \begin{pmatrix} y \\ z \end{pmatrix}, \mathcal{M}(x,\lambda) \begin{pmatrix} y \\ z \end{pmatrix} \right\rangle + h(x,y,z,\lambda) + P(x,y,z,\lambda), \tag{1.2}$$

where $(x, y, z) \in T^n \times R^l \times R^{2m}$, λ is a parameter in a bounded, closed, connected region $\Lambda \subset R^k$, $\mathcal{M}$ is symmetric, real analytic in $x \in \mathcal{D}(r) = \{x \in C^n/Z^n : \ |\mathrm{Im}x| < r\}$, $h(x,y,z,\lambda) = O(|(y,z)|^3)$ is real analytic, and, the perturbation P is real analytic in a complex neighborhood $D(r,s) = \{(x,y,z) : \ |\mathrm{Im}x| < r, |y| < s, |z| < s\}$ of $T^n \times \{0\} \times \{0\}$. In the above, all λ dependence are of class C^{l_0} for some $l_0 \geq n$.

Write $\mathcal{M}$ in (1.2) into blocks:

$$\mathcal{M} = \begin{pmatrix} A & B \\ B^\top & M \end{pmatrix}, \tag{1.3}$$

where $A = A(x,\lambda)$, $B = B(x,\lambda)$, $M = M(x,\lambda)$ are $l \times l$, $l \times 2m$, $2m \times 2m$ minors of $\mathcal{M} = \mathcal{M}(x,\lambda)$ respectively.

A so-called generalized Hamiltonian system is defined on a Poisson manifold which can be odd dimensional and structurally degenerate. Consider the manifold $G \times T^n \times R^{2m}$, where $G \subset R^l$ is a bounded, connected and closed region, T^n is the standard n-torus and l, n, m are positive integers. Let I be the structure matrix in tangent direction, and J be the $2m \times 2m$ standard symplectic matrix in norm direction. As in [8], assume $I = I(\lambda)$ be real analytic. Then the structure matrix $\tilde{I}$ on $G \times T^n \times R^{2m}$ has the following form:

$$\tilde{I}(\lambda) = \begin{pmatrix} I(\lambda) & O \\ O & J \end{pmatrix}, \quad I(\lambda) = \begin{pmatrix} O & E(\lambda) \\ -E^\top(\lambda) & C(\lambda) \end{pmatrix},$$

where O denotes zero matrix, $E = E_{l,n}, C = C_{n,n}$ with $C^\top = -C$. Let ∇ denote the standard Euclidean gradient on $R^l \times T^n \times R^{2m}$. Then $\tilde{I}$ defines a Poisson structure or a 2-form ω^2 in the following way:

$$\{f_1, f_2\} = df_2(\tilde{I}df_1) = \langle \nabla f_1, \tilde{I}\nabla f_2 \rangle = \omega^2(\tilde{I}df_1, \tilde{I}df_2),$$

for all smooth functions f_1, f_2 defined on $G \times T^n \times R^{2m}$, where $\{\cdot,\cdot\}$ denotes the Poisson bracket.

Then the equation of motion associated with (1.2) reads

$$\begin{pmatrix} \dot{y} \\ \dot{x} \\ \dot{z} \end{pmatrix} = \tilde{I}\nabla H.$$

Thus, the unperturbed system associated with (1.2) admits a smooth family of invariant n-tori $T_\lambda = T^n \times \{0\} \times \{0\}$ with toral frequencies $\omega(\lambda) = -E^\top(\lambda)\Omega(\lambda)$ parameterized by $\lambda \in \Lambda$. As in [7], we first assume that $J[M]$ is hyperbolic on Λ, i.e., if $\lambda_i(\lambda)$, $i = 1, 2, \cdots, 2m$, are eigenvalues of $J[M](\lambda)$, then

H) there exists a constant $\sigma_0 > 0$ such that

$$|\mathrm{Re}\lambda_i(\lambda)| \geq \sigma_0,$$

for all $\lambda \in \Lambda$ and $i = 1, \cdots, 2m$.

Next, we assume the Rüssmann condition

R)

$$\max_{\lambda\in\Lambda} \operatorname{rank}\{\partial^\alpha \omega(\lambda): \quad \forall |\alpha| \le n-1\} = n.$$

To make a difference between $\Omega(\lambda)$ and toral frequency $\omega(\lambda) = -E^\top(\lambda)\Omega(\lambda)$, we call $\Omega(\lambda)$ as pseudo-frequency. For general structure matrix I, we cannot obtain the persistence of part frequency components by the associate persistence of part pseudo-frequency components. But for some special structure matrix I, we can even obtain the unchanged toral frequency in spite that only part pseudo-frequency components are preserved (see Example 5.2), which of course depends closely on the specific form of $E(\lambda)$ and $\Omega(\lambda)$. So it is necessary to study the preservation of part or full toral pseudo-frequency components in connection with the degree of non-degeneracy of the matrix $[A]$. As in [7], we assume that

ND) there is a $1 \le n_0 \le l$ such that both the $n_0 \times n_0$ ordered principal minor U of $[A]$ and $Y \equiv [M] - [B]^\top \operatorname{diag}(U^{-1}, O)[B]$ are non-singular on Λ, where O denotes the zero matrix.

It is clear that ND) holds automatically if $[A]$ is non-singular on Λ and $|[B]|_\Lambda$ is sufficiently small (in particular, when $[B] \equiv 0$).

Define

$$\eta = \frac{2}{\sqrt{\rho_0^2 + 4\alpha\rho_0} + \rho_0}, \tag{1.4}$$

where

$$\alpha = (1+2m)(|Y^{-1}| + |U^{-1}| + (|Y^{-1}||U^{-1}|)(2|[B]| + |[B]|^2|U^{-1}|))_\Lambda, \tag{1.5}$$

$$\rho_0 = \frac{4m}{\sigma_0}\left(1 + \frac{2m}{\sigma_0}|[M]|_\Lambda\right)^{2m-1}. \tag{1.6}$$

The main result of this paper is the following.

Theorem 1.1 *Consider* (1.2) *and assume the conditions* H), R), ND) *and*

$$|M - [M]|_{\mathcal{D}(r)\times\Lambda}, \quad |B - [B]|_{\mathcal{D}(r)\times\Lambda} < \eta. \tag{1.7}$$

Then there is an $\varepsilon = \varepsilon(r, s, l_0, \sigma_0, U) > 0$ *sufficiently small such that if*

$$|\partial_\lambda^l P|_{D(r,s)\times\Lambda} < \gamma^{n+1}s^2\varepsilon, \qquad |l| \le l_0, \tag{1.8}$$

then

1) there is a $0 < r_0 = r_0(r, \sigma_0, U) \le r$ *and a Cantor-like set* $\Lambda_\gamma \subset \Lambda$, *with* $|\Lambda \setminus \Lambda_\gamma| = O(\gamma^{\frac{1}{n_*-1}})$, *where* $n_* = \max\{2, n\}$, *for which there is a* C^{l_0-1} *Whitney smooth family of real analytic, symplectic transformations*

$$\Psi_\lambda : D\left(\frac{r_0}{2}, \frac{s}{2}\right) \to D(r_0, s), \quad \lambda \in \Lambda_\gamma,$$

which are C^{l_0} *uniformly close to the identity such that*

$$H \circ \Psi_\lambda = e_* + \langle \Omega_*(\lambda), y\rangle + \frac{1}{2}\left\langle \begin{pmatrix} y \\ z \end{pmatrix}, \mathcal{M}_*(x,\lambda)\begin{pmatrix} y \\ z \end{pmatrix}\right\rangle + h(x,y,z,\lambda) + P_*(x,y,z,\lambda),$$

where

$$|\partial_\lambda^l e_* - \partial_\lambda^l e|_{\Lambda_\gamma} = O(\gamma^{n+1}s\varepsilon\zeta),$$
$$|\partial_\lambda^l \Omega_* - \partial_\lambda^l \Omega|_{\Lambda_\gamma} = O(\gamma^{n+1}s\varepsilon\zeta),$$
$$|\partial_\lambda^l \mathcal{M}_* - \partial_\lambda^l \mathcal{M}|_{\mathcal{D}(r_0)\times\Lambda_\gamma} = O(\gamma^{n+1}\varepsilon\zeta).$$

Thus, all unperturbed tori T_λ with $\lambda \in \Lambda_\gamma$ will persist and give rise to a C^{l_0-1} Whitney smooth family of slightly deformed, analytic, quasi-periodic, invariant n-tori of the perturbed system;

2)

$$(\Omega_*(\lambda))_i = (\Omega_0(\lambda))_i, \quad \lambda \in \Lambda_\gamma,\ i = 1, 2, \cdots, n_0,$$

i.e., the first n_0 components of the perturbed toral pseudo-frequency remain unchanged. In particular, if $n_0 = l$, i.e., $U = [A]$ is non-singular on Λ, then every Diophantine tori T_λ with Diophantine type (γ, τ) for a fixed $\tau > n-1$ will persist with unchanged toral frequencies.

2 KAM Step

In this section, we describe the linear iterative scheme with respect to (1.2) for one KAM step, say, form a νth step to the $(\nu+1)$th step. Below, let $\tau > \max\{n(n-1)-1, l(l-1)-1, 0\}$ be fixed.

Consider (1.2) and define $e_0 = e$, $\Omega_0 = \Omega$, $\mathcal{M}^0 = \mathcal{M}$, $A^0 = A$, $B^0 = B$, $M^0 = M$, $h_0 = h$, $P_0 = P$, $\Lambda_0 = \Lambda$, $\gamma_0 = \gamma$, $r_* = r$, $s_0 = \left(\frac{\gamma_0}{2}\right)^{n+1} \varepsilon_0^{\frac{5}{9}}$. We rewrite $[A^0](= [A])$ into blocks:

$$[A^0] = \begin{pmatrix} U^0 & D^0 \\ (D^0)^\top & V^0 \end{pmatrix},$$

where $U^0 = U$. Without loss of generality, assume that $0 < s_0, r_0, \varepsilon_0 \le 1$. By (1.8), we have

$$|\partial_\lambda^l P_0|_{D(r_0,s_0)} \le \gamma_0^{n+1} s_0^2 \varepsilon_0, \quad |l| \le n. \tag{2.1}$$

In what follows, quantities (domains, normal form, perturbation, etc.) without subscripts denotes the Hamiltonian in ν-th step, while those with subscript "+" denotes the Hamiltonian of $(\nu+1)$-th step. And we shall use "$< \cdot$" to denote "$< c$" with a constant c which is independent of the iteration step. For simplicity, we set $l_0 = n$.

Suppose that at the ν-th step, we have arrived at the following Hamiltonian:

$$\begin{aligned} H &= N + P, \\ N &= e + \langle \Omega(\lambda), y \rangle + \frac{1}{2} \left\langle \begin{pmatrix} y \\ z \end{pmatrix}, \mathcal{M}(x, \lambda) \begin{pmatrix} y \\ z \end{pmatrix} \right\rangle + h_0(x, y, z, \lambda), \end{aligned} \tag{2.2}$$

where $(x, y, z) \in D = D(r, s)$, $\lambda \in \Lambda$, $e(\lambda), \Omega(\lambda)$ are smooth on Λ, $\mathcal{M}(x, \lambda) = \begin{pmatrix} A & B \\ B^\top & M \end{pmatrix}$ is real symmetric over $\mathcal{D} \times \Lambda = \{x : |\mathrm{Im} x| < r\} \times \Lambda$ which is smooth in $\lambda \in \Lambda$ and real analytic in $x \in \mathcal{D} = \mathcal{D}(r)$, P is real analytic in $(x, y, z) \in D$, smooth in $\lambda \in \Lambda$, and moreover,

$$|\partial_\lambda^l P|_{D \times \Lambda} \le \gamma^{n+1} s^2 \varepsilon, \quad |l| \le n.$$

We shall construct a symplectic transformation $\Phi = \Phi_{\nu+1}$ which transforms the Hamiltonian (2.2), in smaller phase and frequency domains, to the desired Hamiltonian in the next KAM cycle (the $(\nu+1)$th KAM step).

Define

$$\begin{aligned} \varepsilon_+ &= \varepsilon^{\frac{10}{9}}, \\ \gamma_+ &= \frac{\gamma_0}{4} + \frac{\gamma}{2}, \\ r_+ &= \frac{r_0}{4} + \frac{r}{2}, \\ s_+ &= \frac{1}{8}\alpha s,\ \alpha = \varepsilon^{\frac{1}{3}}, \end{aligned}$$

$$
\begin{aligned}
K_+ &= ([\log \frac{1}{s}] + 1)^{a^*+2}, \\
D(a) &= D(r_+ + \frac{6}{8}(r - r_+), a), \quad a > 0, \\
\mathcal{D}(a) &= \{x : |\mathrm{Im} x| < a\}, \quad a > 0, \\
\Gamma(a) &= \sum_{0<|k|\le K_+} |k|^{3n+(n+1)\tau} e^{-|k|\frac{a}{8}}, \quad a > 0, \\
D_+ &= D(r_+, s_+), \\
\mathcal{D}_+ &= \mathcal{D}(r_+) = \{x : |\mathrm{Im} x| < r_+\}, \\
D_i &= D(r_+ + \frac{i-1}{8}(r - r_+), i s_+), \quad i = 1, 2, \cdots, 8,
\end{aligned}
$$

where a^* is a constant such that $(\frac{10}{9})^{a^*} > 2$.

2.1 Truncating perturbations

Consider the Taylor-Fourier series of P:

$$
P = \sum_{i \in Z_+^l, j \in Z_+^{2m}, k \in Z^n} p_{kij} y^i z^j e^{\sqrt{-1}\langle k, x\rangle}
$$

and consider the truncation

$$
\begin{aligned}
R &= \sum_{|i|+|j|\le 2, |k|\le K_+} p_{kij} y^i z^j e^{\sqrt{-1}\langle k,x\rangle} = \sum_{|k|\le K_+} (P_{k00} + \langle P_{k10}, y\rangle \\
&\quad + \langle P_{k01}, z\rangle + \langle y, P_{k20} y\rangle + \langle y, P_{k11} z\rangle + \langle z, P_{k02} z\rangle) e^{\sqrt{-1}\langle k,x\rangle}. \qquad (2.3)
\end{aligned}
$$

Lemma 2.1 *Assume that*

H1)

$$
\int_{K_+}^{\infty} \lambda^n e^{-\lambda \frac{r-r_+}{8}} \mathrm{d}\lambda \le \varepsilon.
$$

Then we have

$$
|\partial_\lambda^l (P - R)|_{D_8} \le \cdot \gamma^{n+1} s^2 \varepsilon^2, \quad |\partial_\lambda^l R|_{D_8} \le \cdot \gamma^{n+1} s^2 \varepsilon, \qquad |l| \le n.
$$

Proof. Let

$$
\begin{aligned}
I &= \sum_{|k|>K_+} p_{kij} y^i z^j e^{\sqrt{-1}\langle k,x\rangle}, \\
II &= \sum_{|k|\le K_+, |i|+|j|>2} p_{kij} y^i z^j e^{\sqrt{-1}\langle k,x\rangle} \\
&= \int \frac{\partial^{(p,q)}}{\partial y^p \partial z^q} \sum_{|k|\le K_+, |i|+|j|>2} p_{kij} e^{\sqrt{-1}\langle k,x\rangle} y^i z^j \mathrm{d}y \mathrm{d}z,
\end{aligned}
$$

where $\int$ is the obvious anti-derivative of $\frac{\partial^{(p,q)}}{\partial y^p \partial z^q}$ for $|p| + |q| = 3$. Clearly,

$$
P - R = I + II.
$$

Since, by Cauchy's estimate,

$$
|\sum_{i \in Z_+^l, j \in Z_+^{2m}} \partial_\lambda^l p_{kij} y^i z^j| \le |\partial_\lambda^l P|_{D(r,s)} e^{-|k|r} \le \gamma^{n+1} s^2 \varepsilon e^{-|k|r}, \quad |l| \le n,
$$

from H1) we get that

$$\begin{aligned}|\partial_\lambda^l I|_{D_8} &\leq \sum_{|k|>K_+}\gamma^{n+1}s^2\varepsilon e^{-|k|r}e^{|k|(r_++\frac{7}{8}(r-r_+))}\\ &\leq \gamma^{n+1}s^2\varepsilon\sum_{\kappa=K_+}^{\infty}\kappa^n e^{-\kappa\frac{r-r_+}{8}} \leq \gamma^{n+1}s^2\varepsilon\int_{K_+}^{\infty}\lambda^n e^{-\lambda\frac{r-r_+}{8}}\mathrm{d}\lambda\\ &\leq \gamma^{n+1}s^2\varepsilon^2,\quad |l|\leq n.\end{aligned}$$

It follows that

$$|\partial_\lambda^l(P-I)|_{D_8}\leq|\partial_\lambda^l P|_{D(r,s)}+|\partial_\lambda^l I|_{D_8}\leq\cdot\gamma^{n+1}s^2\varepsilon,\quad |l|\leq n.$$

By Cauchy's estimate we obtain

$$\begin{aligned}|\partial_\lambda^l II|_{D_8} &\leq \left|\int\frac{\partial^{(p,q)}}{\partial y^p\partial z^q}\sum_{|k|\leq K_+,|i|+|j|>2}\partial_\lambda^l p_{kij}e^{\sqrt{-1}\langle k,x\rangle}y^i z^j\mathrm{d}y\mathrm{d}z\right|_{D_8}\\ &\leq \left|\int\left|\frac{\partial^{(p,q)}}{\partial y^p\partial z^q}\partial_\lambda^l(P-I-R)\right|_{D_*}\mathrm{d}y\mathrm{d}z\right|_{D_8}\\ &\leq \cdot\frac{1}{s^3}\gamma^{n+1}s^2\varepsilon\left|\int\mathrm{d}y\mathrm{d}z\right|_{D_8}\leq\cdot\frac{1}{s^3}\gamma^{n+1}s^2\varepsilon s_+^3\leq\cdot\gamma^{n+1}s^2\varepsilon^2,\quad |l|\leq n.\end{aligned}$$

Thus,

$$|\partial_\lambda^l(P-R)|_{D_8}\leq c\gamma^{n+1}s^2\varepsilon^2,$$

and therefore,

$$|\partial_\lambda^l R|_{D_8}\leq|\partial_\lambda^l(P-R)|_{D_8}+|\partial_\lambda^l P|_{D_8}\leq\cdot\gamma^{n+1}s^2\varepsilon,\quad |l|\leq n.$$

2.2 Transformation and homogeneous equation

Write $\mathcal{M}$ into blocks

$$\mathcal{M}(x,\lambda)=\begin{pmatrix}A & B\\ B^\top & M\end{pmatrix},$$

where

$$A(x,\lambda)=\sum_{k\in Z^n}A_k e^{\sqrt{-1}\langle k,x\rangle},\quad B(x,\lambda)=\sum_{k\in Z^n}B_k e^{\sqrt{-1}\langle k,x\rangle},\quad M(x,\lambda)=\sum_{k\in Z^n}M_k e^{\sqrt{-1}\langle k,x\rangle}$$

are $l\times l$, $l\times 2m$, $2m\times 2m$ minors of $\mathcal{M}$ respectively.

To transform (2.2) into the Hamiltonian in the next KAM cycle, we will construct the averaging transformation as the time 1-map ϕ_F^1 of the flow generated by a Hamiltonian F. To this end, suppose F has the following form:

$$F=\sum_{0<|k|\leq K_+}(f_{k0}+\langle f_{k1},y\rangle+\langle F_{k1},z\rangle)e^{\sqrt{-1}\langle k,x\rangle}+\langle F_{01},z\rangle. \tag{2.4}$$

As in [7], to be able to keep the first n_0 components of the toral pseudo-frequencies, we shall also find a $Y_*\in R^{n_0}$ so that the translation of coordinate

$$\phi: x\to x,\quad y\to y+\begin{pmatrix}Y_*\\0\end{pmatrix},\quad z\to z$$

removes all possible drifts among the first n_0 components of the new toral pseudo-frequencies.

We introduce the following notations:

$$
\begin{aligned}
[A] &= \begin{pmatrix} U & D \\ D^\top & V \end{pmatrix}, \\
R' &= \sum_{0<|k|\le K_+} (\langle y, P_{k20} y\rangle + \langle y, P_{k11} z\rangle + \langle z, P_{k02} z\rangle) e^{\sqrt{-1}\langle k,x\rangle} \\
&\quad + [R] - \langle P_{001}, z\rangle + \sum_{|k|\le K_+} \langle B_{-k} J F_{k1}, y\rangle, && (2.5) \\
R_t &= (1-t)\{N, F\} + R, && (2.6) \\
y_* &= \begin{pmatrix} Y_* \\ 0 \end{pmatrix},
\end{aligned}
$$

where U, D, V are the $n_0 \times n_0$, $n_0 \times (l-n_0)$, $(l-n_0)\times(l-n_0)$ minors of $[A]$ respectively.

Denote

$$\Phi_+ = \phi_F^1 \circ \phi.$$

Then we have

$$
\begin{aligned}
H_+ &= H \circ \Phi_+ = H \circ \phi_F^1 \circ \phi = (N+R)\circ \phi_F^1 \circ \phi + (P-R)\circ \phi_F^1 \circ \phi \\
&= (N+R')\circ\phi - \langle y_*, (A-[A])y\rangle - \langle y_*, Bz\rangle \\
&\quad + (\{N,F\} + R - R')\circ \phi + \langle y_*, (A-[A])y\rangle + \langle y_*, Bz\rangle - Q \\
&\quad + \int_0^1 \{R_t, F\}\circ \phi_F^t \circ \phi \mathrm{d}t + (P-R)\circ\phi_F^1\circ\phi + Q,
\end{aligned}
$$

where Q is to be determined in the following.

As in [7], we need to choose a function Q such that both equations

$$(\{N,F\} + R - R')\circ\phi - Q + \langle y_*, (A-[A])y\rangle + \langle y_*, Bz\rangle = 0, \qquad (2.7)$$

$$\mathrm{diag}(U, O) y_* = \mathrm{diag}(I_{n_0}, O)(-P_{010} - \sum_{|j|\le K_+} B_{-j} J F_{j1}) \qquad (2.8)$$

are solvable. If this is the case, we then arrive at that

$$
\begin{aligned}
H_+ &= N_+ + P_+, \\
N_+ &= e_+ + \langle \Omega_+(\lambda), y\rangle + \frac{1}{2}\left\langle \begin{pmatrix} y \\ z\end{pmatrix}, \mathcal{M}^+ \begin{pmatrix} y \\ z\end{pmatrix}\right\rangle + h_0(x,y,z,\lambda) \\
&= e_+ + \langle \Omega_+(\lambda), y\rangle + \frac{1}{2}\left\langle \begin{pmatrix} y \\ z\end{pmatrix}, \begin{pmatrix} A^+ & B^+ \\ B^{+\top} & M^+ \end{pmatrix} \begin{pmatrix} y \\ z\end{pmatrix}\right\rangle + h_0(x,y,z,\lambda),
\end{aligned}
$$

where

$$e_+ = e + P_{000} + \langle \Omega, y_*\rangle + \frac{1}{2}\langle y_*, [A] y_*\rangle, \qquad (2.9)$$

$$\Omega_+ = \Omega + \mathrm{diag}(O, I_{n-n_0})([A]y_* + P_{010} + \sum_{|k|\le K_+} B_{-k} J F_{k1}), \qquad (2.10)$$

$$\omega_+ = -E^\top \Omega_+, \qquad (2.11)$$

$$A^+ = A + \sum_{|k|\le K_+} 2P_{k20} e^{\sqrt{-1}\langle k,x\rangle}, \qquad (2.12)$$

$$B^+ = B + \sum_{|k|\le K_+} P_{k11} e^{\sqrt{-1}\langle k,x\rangle}, \qquad (2.13)$$

$$M^+ = M + \sum_{|k|\le K_+} 2P_{k02} e^{\sqrt{-1}\langle k,x\rangle}, \tag{2.14}$$

$$\begin{aligned} P_+ = & \int_0^1 \{R_t, F\} \circ \phi_F^t \circ \phi \mathrm{d}t + (P-R)\circ \phi_F^1 \circ \phi \\ & + \frac{1}{2}\langle y_*, (A-[A])y_*\rangle + h_0(x, y+y_*, z, \lambda) - h_0(x,y,z,\lambda) \\ & + \sum_{|k|\le K_+} (\langle y_*, P_{k20} y_*\rangle + \langle y_*, 2P_{k20} y\rangle + \langle y_*, P_{k11} z\rangle) e^{\sqrt{-1}\langle k,x\rangle} + Q. \end{aligned} \tag{2.15}$$

We now consider the equations (2.7) and (2.8). By careful observation of (2.7), we suppose that Q has the following form:

$$\begin{aligned} Q =& \Big(\sum_{0<|k|\le K_+} \big(-\langle \frac{1}{2}\partial_x \langle y, A(x,\lambda) y\rangle + \partial_x \langle y, B(x,\lambda) z\rangle \\ & + \frac{1}{2}\partial_x\langle z, M(x,\lambda) z\rangle + \partial_x h_0(x,y,z,\lambda), E^\top f_{k1}\rangle \\ & + \sqrt{-1}\langle Ek, A(x,\lambda) y + B(x,\lambda) z \\ & + \partial_y h_0(x,y,z,\lambda)\rangle (f_{k0} + \langle f_{k1}, y\rangle + \langle F_{k1}, z\rangle) \\ & + \sqrt{-1}\langle \frac{1}{2}\partial_x\langle y, A(x,\lambda) y\rangle + \partial_x\langle y, B(x,\lambda) z\rangle \\ & + \frac{1}{2}\partial_x\langle z, M(x,\lambda) z\rangle + \partial_x h_0(x,y,z,\lambda), Ck\rangle (f_{k0} + \langle f_{k1}, y\rangle + \langle F_{k1}, z\rangle)\big) e^{\sqrt{-1}\langle k,x\rangle} \\ & + \sum_{|k|>K_+} (\langle B_k J F_{01}, y\rangle + \langle M_k J F_{01}, z\rangle) e^{\sqrt{-1}\langle k,x\rangle} \\ & + \sum_{|k|>K_+, 0<|j|\le K_+} (\langle B_{k-j} J F_{j1}, y\rangle + \langle M_{k-j} J F_{j1}, z\rangle) e^{\sqrt{-1}\langle k,x\rangle} \\ & + \sum_{0<|k|\le K_+} \langle \partial_z h_0(x,y,z,\lambda) J F_{k1}, z\rangle e^{\sqrt{-1}\langle k,x\rangle} \Big) \circ \phi \\ & + \sum_{0<|k|\le K_+} (\sqrt{-1}\langle k, E^\top \Omega(\lambda)\rangle \langle f_{k1}, y_*\rangle + \langle P_{k10}, y_*\rangle) e^{\sqrt{-1}\langle k,x\rangle} \\ & + \sum_{|k|>K_+} (\langle y_*, A_k y\rangle + \langle y_*, B_k z\rangle) e^{\sqrt{-1}\langle k,x\rangle} \\ & + \sum_{0<|k|\le K_+, 0\le|j|\le K_+} \langle y_*, B_{k-j} J F_{j1}\rangle e^{\sqrt{-1}\langle k,x\rangle}. \end{aligned} \tag{2.16}$$

Substituting (2.3)–(2.5) and (2.16) into (2.7) and comparing coefficients, from equations (2.7) and (2.8) we obtain the following linear equations for all $0 < |k| \le K_+$:

$$\sqrt{-1}\langle k, \omega(\lambda)\rangle f_{k0} = P_{k00}, \tag{2.17}$$

$$\sqrt{-1}\langle k, \omega(\lambda)\rangle f_{k1} = P_{k10} + A_k y_* + \sum_{|j|\le K_+} B_{k-j} J F_{j1}, \tag{2.18}$$

$$\begin{aligned} \sqrt{-1}\langle k, \omega(\lambda)\rangle F_{k1} - [M] J F_{k1} = & \sum_{0<|j|\le K_+, j\ne k} M_{k-j} J F_{j1} \\ & + P_{k01} + B_k^\top y_* + M_k J F_{01}, \end{aligned} \tag{2.19}$$

$$[M] J F_{01} = -P_{001} - \sum_{0<|j|\le K_+} M_{-j} J F_{j1} - [B]^\top y_*, \tag{2.20}$$

$$\mathrm{diag}(U,O)y_* = \mathrm{diag}(I_{n_0},O)(-P_{010} - \sum_{0<|j|\le K_+} B_{-j}JF_{j1} - [B]JF_{01}), \qquad (2.21)$$

where $\omega(\lambda) = -E^\top(\lambda)\Omega(\lambda)$.

Denote

$$\Lambda_+ = \{\lambda \in \Lambda : |\langle k, \omega(\lambda)\rangle| > \frac{\gamma}{|k|^\tau}, \quad 0 < |k| \le K_+\}. \qquad (2.22)$$

If we assume that

H2)

$$|\partial_\lambda^l(\mathcal{M} - \mathcal{M}^0)|_{\mathcal{D}(r)\times\Lambda} \le \varepsilon_0^{\frac{1}{4}}, \quad |l| \le n,$$

then as in [7], (2.17)–(2.21) can be equivalently written into the following system form:

$$(\Lambda - \mathcal{A})\mathcal{F} = \mathcal{P}, \qquad (2.23)$$

where Λ and $\mathcal{A}$ are defined as in [7].

2.3 Estimate on $(\Lambda - \mathcal{A})^{-1}$

As in [7], by the hyperbolicity of $J[M^0]$ and the definition of η, we can prove that

$$|(\Lambda^0)^{-1}|_{\Lambda_0} \le \frac{\rho_0}{2}, \quad |\mathcal{A}^0|_{\Lambda_0} \le \frac{1}{\rho_0}, \quad |(\Lambda^0 - \mathcal{A}^0)^{-1}|_{\Lambda_0} < 2\rho_0. \qquad (2.24)$$

Lemma 2.2 *Assume* H2) *and also that*

H3)

$$|\partial_\lambda^l \mathcal{A} - \partial_\lambda^l \mathcal{A}^0|_\Lambda < \varepsilon_0^{\frac{1}{4}}.$$

Then for ε_0 sufficiently small, $\mathcal{L} = \Lambda - \mathcal{A}$ is non-singular on Λ, and moreover, the following holds:

$$|\partial_\lambda^l \mathcal{L}^{-1}|_\Lambda \le \cdot K_+^n, \quad |l| \le n.$$

Proof. Similar to Lemma 3.2 of [7], we have $|\partial_\lambda \mathcal{L}^{-1}|_\Lambda \le \cdot K_+$. By induction,

$$|\partial_\lambda^l \mathcal{L}^{-1}|_\Lambda \le \cdot K_+^n, \quad |l| \le n.$$

Above all, by the hypotheses H2) and H3), the linear system (2.23) can be uniquely solved on Λ_+ to yield smooth functions $f_{k0}, f_{k1}, F_{k1}, F_{01}, y_*$, $0 < |k| \le K_+$.

2.4 Estimates on the transformation

Denote

$$\zeta = K_+^{n+2}\Gamma(r - r_+)^2.$$

Lemma 2.3 *Assume* H2)*. Then the following holds for all $|l| \le n$:*

1) $|\partial_\lambda^l y_*|_{\Lambda_+} \le \cdot \gamma^{n+1} s\varepsilon\zeta$;

2) *On* $D(s) \times \Lambda_+$,

$$|\partial_\lambda^l F|, \ |\partial_\lambda^l F_x|, \ s|\partial_\lambda^l F_y|, \ s|\partial_\lambda^l F_z| \le \cdot s^2\varepsilon\zeta;$$

3) *On* $D(s) \times \Lambda_+$,

$$|\partial_\lambda^l D^i F| \le \cdot \varepsilon\zeta, \quad |i| \ge 2.$$

Proof. The proof is similar to that in [7].

Lemma 2.4 *Assume* H2), H3) *and also that*

H4)

$$s\varepsilon\zeta < \frac{1}{8}(r - r_+), \quad s\varepsilon\zeta < s_+.$$

Let ϕ_F^t be the flow generated by F. Then the following holds:

1) *For all $0 \le t \le 1$, $\phi_F^t : D_2 \to D_3$, $\phi : D_1 \to D_2$ are well defined, real analytic and depend smoothly on $\lambda \in \Lambda_+$, i.e., $\Phi_+ = \phi_F^1 \circ \phi : D_+ \to D$;*

2) *$|\partial_\lambda^l(\phi_F^t - id)|_{D(s)\times\Lambda_+} \le \cdot s\varepsilon\zeta$, $|\partial_\lambda^l D^i(\Phi_+ - id)|_{\tilde{D}_+\times\Lambda_+} \le \cdot\varepsilon\zeta$, for all $|l| \le n$, $i \ge 0$, $0 \le t \le 1$, where $D = \partial_{(x,y,z)}$.*

Proof. Let $\lambda \in \Lambda_+$.

1) It is easy to see that $\phi : D_1 \to D_2$ holds by Lemma 2.3 1) and H4).

We note that

$$\phi_F^t = \mathrm{id} + \int_0^t X_F \circ \phi_F^\xi \mathrm{d}\xi, \tag{2.25}$$

where

$$X_F = \tilde{I}(\lambda)\nabla F = (E(\lambda)F_x, -E^\top(\lambda)F_y + C(\lambda)F_x, JF_z)^\top.$$

Denote $\phi_{F1}^t, \phi_{F2}^t, \phi_{F3}^t$ as components of ϕ_F^t in y, x, z planes respectively. For any $(x, y, z) \in D_2$, let $t_* = \sup\{t \in [0, 1] : \phi_F^t(x, y, z) \in D_3$. By making ε_0 small, we have that $D_3 \subset D(s)$. It follows from H4) and Lemma 2.3 that

$$\begin{aligned}
|\phi_{F1}^t(x,y,z)| &\le |y| + |\int_0^t E(\lambda)F_x \circ \phi_F^\xi \mathrm{d}\xi| \le |y| + \cdot|F_x|_{D(s)} \le 2s_+ + \cdot s^2\varepsilon\zeta < 3s_+,\\
|\phi_{F2}^t(x,y,z)| &\le |x| + |\int_0^t (-E^\top(\lambda)F_y + C(\lambda)F_x) \circ \phi_F^\xi \mathrm{d}\xi| \le |x| + \cdot(|F_x| + |F_y|)_{D(s)}\\
&\le r_+ + \frac{1}{8}(r - r_+) + \cdot s\varepsilon\zeta\\
&< r_+ + \frac{2}{8}(r - r_+),\\
|\phi_{F3}^t(x,y,z)| &\le |z| + |\int_0^t JF_z \circ \phi_F^\xi \mathrm{d}\xi| \le |z| + |F_z|_{D(s)} \le 2s_+ + \cdot s\varepsilon\zeta < 3s_+,
\end{aligned}$$

i.e., $\phi_F^t(x, y, z) \in D_3$ for all $0 \le t \le t_*$. Thus, $t_* = 1$ and 1) holds.

2) By Lemma 2.3 and (2.25), we immediately have

$$|\phi_F^t - id|_{D(s)} \le \cdot s\varepsilon\zeta.$$

Differentiating (2.25) with respect to λ yields

$$\begin{aligned}
\partial_\lambda\phi_F^t &= \int_0^t X_F \circ \phi_F^\xi \partial_\lambda\phi_F^\xi \mathrm{d}\xi + \int_0^t (\partial_\lambda X_F) \circ \phi_F^\xi \mathrm{d}\xi\\
&= \int_0^t (E(\lambda)F_x, -E^\top(\lambda)F_y + C(\lambda)F_x, JF_z)^\top \circ \phi_F^\xi \partial_\lambda\phi_F^\xi \mathrm{d}\xi\\
&\quad + \int_0^t \partial_\lambda(E(\lambda)F_x, -E^\top(\lambda)F_y + C(\lambda)F_x, JF_z)^\top \circ \phi_F^\xi \mathrm{d}\xi.
\end{aligned}$$

It follows from Lemma 2.3 and Gronwall's inequality that

$$|\partial_\lambda\phi_F^t|_{D(s)} \le \cdot s\varepsilon\zeta.$$

By induction, we have

$$|\partial_\lambda^l\phi_F^t|_{D(s)} \le \cdot s\varepsilon\zeta, \quad |l| \le n.$$

The estimates for Φ_+ follow from a similar application of Lemma 2.3 and Gronwall's inequality, and the identity

$$\Phi_+ - id = (\phi_F^1 - id) \circ \phi + \begin{pmatrix} 0 \\ y_* \\ 0 \end{pmatrix}.$$

We omit the details.

2.5 Estimate on N_+

We first estimate the new normal form.

Lemma 2.5 *For the new normal form, we have the following holds for all $|l| \leq n$:*

$$\begin{aligned} |\partial_\lambda^l (e_+ - e)|_{\Lambda_+} &\leq \cdot\, \gamma^{n+1} s\varepsilon\zeta, \\ |\partial_\lambda^l (\Omega_+ - \Omega)|_{\Lambda_+} &\leq \cdot\, \gamma^{n+1} s\varepsilon\zeta, \\ |\partial_\lambda^l (\omega_+ - \omega)|_{\Lambda_+} &\leq \cdot\, \gamma^{n+1} s\varepsilon\zeta, \\ |\partial_\lambda^l (\mathcal{M}^+ - \mathcal{M})|_{\mathcal{D}_+ \times \Lambda_+} &\leq \cdot\, \gamma^{n+1} \varepsilon\Gamma(r - r_+). \end{aligned}$$

Proof. First, by Cauchy's estimate we have

$$\begin{aligned} |\partial_\lambda^l P_{kij}|_{\mathcal{O}} &\leq \cdot s^{-(i+j)} |\partial_\lambda^l P|_{D(r,s)\times\mathcal{O}} e^{-|k|r} \\ &\leq \cdot \gamma^{n+1} s^{2-i-j} \varepsilon e^{-|k|r}, \quad |k| \geq 0,\ i,j = 0,1,2. \end{aligned} \tag{2.26}$$

Then from (2.9)–(2.14) and (2.26) the Lemma immediately follows.

2.6 Frequency property

Lemma 2.6 *Assume that*

H5)

$$\gamma^{n+1} s\varepsilon\zeta K_+^{\tau+1} < \gamma - \gamma_+.$$

Then

$$|\langle k, \omega_+(\lambda)\rangle| > \frac{\gamma_+}{|k|^\tau},$$

for all $\lambda \in \Lambda_+$ and $0 < |k| \leq K_+$.

Proof. By H5) and Lemma 2.5, one has

$$\begin{aligned} |\langle k, \omega_+(\lambda)\rangle| &= |\langle k, \omega(\lambda)\rangle + \langle k, \omega_+(\lambda) - \omega(\lambda)\rangle| \\ &\geq |\langle k, \omega(\lambda)| - \gamma^{n+1} s\varepsilon\zeta K_+ \\ &\geq \frac{\gamma}{|k|^\tau} - \frac{\gamma - \gamma_+}{|k|^\tau} = \frac{\gamma_+}{|k|^\tau}, \end{aligned} \tag{2.27}$$

as desired.

2.7 Estimate on the new perturbation

Denote

$$\Delta = \cdot s^3 \varepsilon^2 \zeta^2 + \cdot \gamma^{n+1} s^2 \varepsilon^2 \zeta^2 + \cdot s_+ s^2 \varepsilon\zeta. \tag{2.28}$$

Lemma 2.7 *Assume* H1)–H4). *Then* $|\partial_\lambda^l P_+|_{D_+} \le \Delta,\ |l| \le n$. *Thus, if* **H6)**

$$\Delta \le \gamma_+^{n+1} s_+^2 \varepsilon_+,$$

then

$$|\partial_\lambda^l P_+|_{D_+} \le \gamma_+^{n+1} s_+^2 \varepsilon_+. \tag{2.29}$$

Proof. Let $|l| \le n$, $\lambda \in \Lambda_+$. By (2.15), we have that

$$P_+ = W_0 \circ \phi + W_1 + Q + q + (P - R) \circ \Phi_+, \tag{2.30}$$

where

$$\begin{aligned} W_0 &= \int_0^1 \{R_t, F\} \circ \phi_F^t \mathrm{d}t, \\ W_1 &= \frac{1}{2}\langle y_*, (A - [A]) y_*\rangle + \sum_{|k| \le K_+} (\langle y_*, P_{k20} y_*\rangle + \langle y_*, 2P_{k20} y\rangle + \langle y_*, P_{k11} z\rangle) e^{\sqrt{-1}\langle k, x\rangle}, \\ q &= h_0(x, y + y_*, z, \lambda) - h_0(x, y, z, \lambda). \end{aligned}$$

1) We first estimate $(P - R) \circ \Phi_+$.

By Lemma 2.4 1) and Lemma 2.1, we have

$$|\partial_\lambda^l (P - R) \circ \Phi_+|_{D_+} \le |\partial_\lambda^l (P - R)|_{D_3} \le \cdot \gamma^{n+1} s^2 \varepsilon^2. \tag{2.31}$$

2) Then we give the estimate of q.

Following the Taylor series expansion, H4) and Lemma 2.3 1), we obtain

$$\begin{aligned} |\partial_\lambda^l q|_{D_+} &= |\partial_\lambda^l (h_0'(y) y_*) + \frac{1}{2!} y_* h_0^{(2)} y_* + \frac{1}{3!} h_0^{(3)} y_*^3 + o(y_*^3)|_{D_+} \\ &\le s_+^2 |y_*| + s_+ |y_*|^2 + |y_*|^3 \le \cdot s_+^2 |y_*| \le \cdot \gamma^{n+1} s_+^2 s \varepsilon \zeta. \end{aligned} \tag{2.32}$$

3) Then we estimate W_1.

By Lemma 2.3 1), (2.15) and H4), we have that

$$\begin{aligned} |\partial_\lambda^l W_1|_{D_+} &\le \cdot |y_*|^2 + \sum_{|k| \le K_+} (|y_*|^2 \gamma^{n+1} \varepsilon + s_+ |y_*| \gamma^{n+1} \varepsilon) e^{-|k| \frac{r - r_+}{2}} \\ &\le \cdot |y_*|^2 + \sum_{|k| \le K_+} \cdot s_+ |y_*| \gamma^{n+1} \varepsilon e^{-|k| \frac{r - r_+}{2}} \\ &\le \cdot \gamma^{n+1} s^2 \varepsilon^2 \zeta^2 + \cdot s_+ \gamma^{n+1} s \varepsilon \zeta \gamma^{n+1} \varepsilon \Gamma \\ &\le \cdot \gamma^{n+1} s^2 \varepsilon^2 \zeta^2. \end{aligned} \tag{2.33}$$

4) Next, we give the estimate of Q.

By a similar computation to [7], and noting that $|E(\lambda)|, |C(\lambda)| \le c$ for some constant c, we obtain that

$$|\partial_\lambda^l Q|_{D_+} \le \cdot s_+ s^2 \varepsilon \zeta + \cdot \gamma^{n+1} s^2 \varepsilon^2 \zeta^2. \tag{2.34}$$

5) Now we can estimate $W_0 \circ \phi$.

We can obtain the estimate of $W_0 \circ \phi$ as in [7]:

$$|\partial_\lambda^l W_0 \circ \phi|_{D_+} \le \cdot s^3 \varepsilon^2 \zeta^2 + \cdot \gamma^{n+1} s^2 \varepsilon^3 \zeta^3 + \cdot \gamma^{n+1} s^2 \varepsilon^2 \zeta^2.$$

It will be proved that $\varepsilon\zeta \le 1$ later, so we have

$$|\partial_\lambda^l W_0 \circ \phi|_{D_+} \le \cdot s^3 \varepsilon^2 \zeta^2 + \cdot \gamma^{n+1} s^2 \varepsilon^2 \zeta^2. \tag{2.35}$$

Above all, it follows from (2.31), (2.32), (2.34), (2.30), (2.35), (2.33) that

$$|\partial_\lambda^l P_+|_{D_+} \le \cdot s^3 \varepsilon^2 \zeta^2 + \cdot \gamma^{n+1} s^2 \varepsilon^2 \zeta^2 + \cdot s_+ s^2 \varepsilon \zeta.$$

So by H6), (2.29) holds. This completes the proof of the Lemma.

This completes one cycle of KAM steps.

3 Iteration Lemma

Consider (1.2) and let $r_0, s_0, \varepsilon_0, \gamma_0, \Lambda_0, H_0, N_0, e_0, \Omega_0, \mathcal{M}^0, A^0, B^0, M^0, \mathcal{A}^0, h_0, P_0$ be defined in Section 2 and let $D_0 = D(r_0, s_0)$, $\mathcal{D}_0 = \{x : |\mathrm{Im}x| < r_0\}$, $K_0 = 0$, $\Phi_0 = id$. We define the following sequences inductively for all $\nu = 1, 2, \cdots$:

$$
\begin{aligned}
H_\nu &= H_\nu(x,y,z,\lambda) = N_\nu + P_\nu,\\
N_\nu &= e_\nu + \langle \Omega_\nu, y\rangle + \frac{1}{2}\left\langle \begin{pmatrix} y\\ z\end{pmatrix}, \mathcal{M}^\nu \begin{pmatrix} y\\ z\end{pmatrix}\right\rangle + h_0(x,y,z,\lambda),\\
\mathcal{M}^\nu &= \begin{pmatrix} A^\nu & B^\nu\\ (B^\nu)^\top & M^\nu\end{pmatrix},\\
\varepsilon_\nu &= \varepsilon_{\nu-1}^{\frac{10}{9}},\\
r_\nu &= r_0\left(1 - \sum_{i=1}^{\nu}\frac{1}{2^{i+1}}\right),\\
s_\nu &= \frac{1}{8}\alpha s_{\nu-1}, \quad \alpha_{\nu-1} = \varepsilon_{\nu-1}^{\frac{1}{3}},\\
\gamma_\nu &= \gamma_0\left(1 - \sum_{i=1}^{\nu}\frac{1}{2^{i+1}}\right),\\
K_\nu &= \left(\left[\log\frac{1}{s_{\nu-1}}\right] + 1\right)^3, \quad \nu \geq 1,\\
\Lambda_\nu &= \{\lambda \in \Lambda_{\nu-1} : |\langle k, \Omega_{\nu-1}(\lambda)\rangle| > \frac{\gamma_{\nu-1}}{|k|^\tau}, 0 < |k| \leq K_\nu\},\ \nu \geq 1,\\
D_\nu &= D(r_\nu, s_\nu),\\
\mathcal{D}_\nu &= \{x : |\mathrm{Im}x| < r_\nu\}.
\end{aligned}
$$

Lemma 3.1(Iteration Lemma) *If $\varepsilon_0 = \varepsilon_0(r_*, \sigma_0, U^0)$ is sufficiently small, then the following holds for all $|l| \leq n; \nu = 1, 2, \cdots$.*

1) *There is a transformation $\Phi_\nu : D_\nu \times \Lambda_\nu \longrightarrow D_{\nu-1}$, which is symplectic and analytic in $(x,y,z) \in D_{\nu+1}$, and smooth in $\lambda \in \Lambda_{\nu+1}$, such that $H_\nu = H_{\nu-1} \circ \Phi_\nu$ and*

$$|\partial_\lambda^l D^i(\Phi_\nu - id)|_{D_\nu \times \Lambda_\nu} \leq \cdot\varepsilon_{\nu-1}\zeta_{\nu-1},\ i \geq 0. \tag{3.1}$$

2)

$$
\begin{aligned}
&|\partial_\lambda^l (e_\nu - e_0)|_{\Lambda_\nu} \leq \cdot\gamma_0^{n+1} s_0\varepsilon_0\zeta_0, &(3.2)\\
&|\partial_\lambda^l (e_\nu - e_{\nu-1})|_{\Lambda_\nu} \leq \cdot\gamma_{\nu-1}^{n+1} s_{\nu-1}\varepsilon_{\nu-1}\zeta_{\nu-1}, &(3.3)\\
&|\partial_\lambda^l (\Omega_\nu - \Omega_0)|_{\Lambda_\nu} \leq \cdot\gamma_0^{n+1} s_0\varepsilon_0\zeta_0, &(3.4)\\
&|\partial_\lambda^l (\Omega_\nu - \Omega_{\nu-1})|_{\Lambda_\nu} \leq \cdot\gamma_{\nu-1}^{n+1} s_{\nu-1}\varepsilon_{\nu-1}\zeta_{\nu-1}, &(3.5)\\
&|\partial_\lambda^l (\omega_\nu - \omega_0)|_{\Lambda_\nu} \leq \cdot\gamma_0^{n+1} s_0\varepsilon_0\zeta_0, &(3.6)\\
&|\partial_\lambda^l (\omega_\nu - \omega_{\nu-1})|_{\Lambda_\nu} \leq \cdot\gamma_{\nu-1}^{n+1} s_{\nu-1}\varepsilon_{\nu-1}\zeta_{\nu-1}, &(3.7)\\
&|\partial_\lambda^l (\mathcal{M}^\nu - \mathcal{M}^0)|_{\mathcal{D}_\nu \times \Lambda_\nu} \leq \cdot\gamma_0^{n+1}\varepsilon_0\zeta_0, &(3.8)\\
&|\partial_\lambda^l (\mathcal{M}^\nu - \mathcal{M}^{\nu-1})|_{\mathcal{D}_\nu \times \Lambda_\nu} \leq \cdot\gamma_{\nu-1}^{n+1}\varepsilon_{\nu-1}\zeta_{\nu-1}, &(3.9)\\
&|\partial_\lambda^l P_\nu|_{D_\nu \times \Lambda_\nu} \leq \gamma_\nu^{n+1} s_\nu^2\varepsilon_\nu. &(3.10)
\end{aligned}
$$

3) $(\Omega_\nu(\lambda))_i = \Omega_i(\lambda), \quad i = 1, 2, \cdots, n_0.$

4)

$$\Lambda_{\nu+1} = \{\lambda \in \Lambda_\nu : |\langle k, \Omega_\nu(\lambda)\rangle| > \frac{\gamma_\nu}{|k|^\tau},\ K_\nu < |k| \leq K_{\nu+1}\}.$$

Proof. The proof amounts to the verification of H1)–H6) for all ν. The lemma will be proved by induction.

By definitions of ε_ν, s_ν, we clearly have

$$\varepsilon_\nu = \varepsilon_0^{(\frac{10}{9})^\nu}, \tag{3.11}$$
$$s_\nu = \left(\frac{1}{8}\right)^\nu \varepsilon_0^{3((\frac{10}{9})^\nu - 1)} s_0. \tag{3.12}$$

By making ε_0 small, we obtain that

$$\begin{aligned}
&\log(n+1)! + n(a^*+2)\log\left(\left[\log\frac{1}{\varepsilon_\nu}\right]+1\right) - \frac{1}{2^{\nu+5}}\left(\left[\log\frac{1}{\varepsilon_\nu}\right]+1\right)^{a^*+2} r_0 \\
&+(n+1)((\nu+5)\log 2 - \log r_0) \\
\le\ & \log(n+1)! + n(a^*+2)\log\left(\log\frac{1}{\varepsilon_\nu}+2\right) - \left(\log\frac{1}{\varepsilon_\nu}\right)^2 r_0 \\
&+(n+1)((\nu+5)\log 2 - \log r_0) \\
\le\ & -\log\frac{1}{\varepsilon_\nu},
\end{aligned}$$

where the first '$\le$' holds because of the choice of a^* satisfying $\frac{1}{2^{\nu+5}}\left(\log\frac{1}{\varepsilon_\nu}\right)^{a^*} \ge 1$. Hence,

$$\int_{K_{\nu+1}}^{\infty} \lambda^n e^{-\lambda\frac{r-r_+}{8}} \mathrm{d}\lambda \le (n+1)! K_{\nu+1}^n \left(\frac{2^{\nu+5}}{r_0}\right)^{n+1} e^{-K_{\nu+1}\frac{r_0}{2^{\nu+5}}} \le s_\nu.$$

This verifies H1).

The verification of H2)–H3) is similar to that in [7], the reader can refer to [7] for details.

To verify H4)–H6), we will prove $\varepsilon^{\frac{1}{9}}\zeta^2 \le 1$ at first. By making ε_0 sufficiently small, it follows that

$$\begin{aligned}
\varepsilon^{\frac{1}{9}}\zeta^2 &= \varepsilon_0^{\frac{1}{9}(\frac{10}{9})^\nu}(K_+^{n+2}\Gamma(r-r_+)^2)^2 \\
&\le \varepsilon_0^{\frac{1}{9}(\frac{10}{9})^\nu}\left(\log\frac{1}{\varepsilon}+1\right)^{2(n+2)(a^*+2)} ((3n+(n+1)\tau+1)!)^4 \left(\frac{2^{\nu+5}}{r_0}\right)^{4(3n+(n+1)\tau+1)} \\
&\le \cdot\varepsilon_0^{\frac{1}{9}(\frac{10}{9})^\nu}\left(\log\frac{1}{\varepsilon}+1\right)^{2(n+2)(a^*+2)} 2^{4\nu(3n+(n+1)\tau+1)} \\
&\le \cdot[\varepsilon_0^{\frac{1}{18}(\frac{10}{9})^\nu}\left(\log\frac{1}{\varepsilon}+1\right)^{2(n+2)(a^*+2)}][\varepsilon_0^{\frac{1}{18}(\frac{10}{9})^\nu} 2^{4\nu(3n+(n+1)\tau+1)}] \\
&\le 1.
\end{aligned} \tag{3.13}$$

Now, we can prove H4)–H6) easily. When ε is sufficiently small, the following hold:

$$\begin{aligned}
&s\varepsilon\zeta \le s\varepsilon^{\frac{8}{9}} = \left(\frac{1}{8}\right)^\nu \varepsilon_0^{3((\frac{10}{9})^\nu-1)} s_0 \varepsilon_0^{\frac{8}{9}(\frac{10}{9})^\nu} < \left(\frac{1}{2}\right)^{\nu+5} r_0 = \frac{r-r_+}{8}, \\
&s\varepsilon\zeta \le s\varepsilon^{\frac{8}{9}} < \frac{1}{8}\varepsilon^{\frac{1}{3}} s = s_+, \\
&\gamma^{n+1} s\varepsilon\zeta K_+^{\tau+1} \le^{(*)} \gamma_0^{n+1} s\varepsilon^{\frac{1}{2}} < \frac{\gamma_0}{2^{\nu+2}} = \gamma - \gamma_+,
\end{aligned}$$

i.e., H4, H5) hold, where $(*)$ holds similar to (3.13).

At last, we give the proof of H6). By the smallness of ε_0 and the choice of $s_0 = \left(\frac{\gamma_0}{2}\right)^{n+1} \varepsilon_0^{\frac{5}{9}}$, we have the following estimates:

$$
\begin{aligned}
& s^3\varepsilon^2\zeta^2 = 8^2\left(\frac{1}{8}\right)^2 \alpha^2 s^2(\varepsilon^{\frac{1}{9}}\zeta^2)s\varepsilon^{\frac{1}{9}}\varepsilon^{\frac{10}{9}} \le 8^2 s\varepsilon^{\frac{1}{9}} s_+^2\varepsilon_+ \le \gamma_+^{n+1}s_+^2\varepsilon_+, \\
& \gamma^{n+1}s^2\varepsilon^2\zeta^2 \le \gamma^{n+1}s^2\varepsilon\varepsilon^{\frac{8}{9}} = 8^2\gamma^{n+1}\varepsilon^{\frac{1}{9}}s_+^2\varepsilon_+ \le \gamma_+^{n+1}s_+^2\varepsilon_+, \\
& s_+s^2\varepsilon\zeta \le s_+s^2\varepsilon^{\frac{8}{9}} = 8s_+^2\varepsilon_+ s\varepsilon^{-\frac{5}{9}} \\
& \qquad = 8s_+^2\varepsilon_+\left(\frac{1}{8}\right)^{\nu}\varepsilon_0^{3((\frac{10}{9})^{\nu}-1)}(\frac{\gamma_0}{2})^{n+1}\varepsilon_0^{\frac{5}{9}}\varepsilon_0^{-\frac{5}{9}(\frac{10}{9})^{\nu}} \\
& \qquad \le \gamma_+^{n+1}s_+^2\varepsilon_+.
\end{aligned}
$$

This verifies H6).

Above all, H1)–H6) hold for all $\nu = 0, 1, \cdots$, i.e., the KAM step described in Section 2 is valid for all $\nu = 0, 1, \cdots$. Now, (3.3), (3.5), (3.7) and (3.9) follow from Lemma 2.5; (3.2), (3.4), (3.6) and (3.8) follow from (3.3), (3.5), (3.7) and (3.9) respectively; (3.10) follows from Lemma 2.7; part 2) of the lemma follows from Lemma 2.4; part 3) of the lemma follows from an inductive application of (2.10); part 4) of the lemma easily follows from Lemma 2.6. This completes the proof of the lemma.

4 Proof of Main Result

Let

$$\Psi^{\nu} = \Phi_0 \circ \Phi_1 \circ \cdots \circ \Phi_{\nu}, \quad \nu = 0, 1, \cdots$$

Then $\Psi^{\nu} : D_{\nu} \times \Lambda_{\nu} \to D_0$, and

$$H \circ \Psi^{\nu} = H_{\nu} = N_{\nu} + P_{\nu}, \quad \nu = 0, 1, \cdots$$

where $\Psi^0 = id$.

Denote

$$\Lambda_* = \bigcap_{\nu=0}^{\infty} \Lambda_{\nu}, \quad G_* = D\left(\frac{r_0}{2}, \frac{s_0}{2}\right) \times \Lambda_*.$$

Then Λ_* is a Cantor-like set consisting of non-resonant frequencies, and moreover, a measure estimate similar to that in [9] (also [8], [10]) yields that $|\Lambda \setminus \Lambda_*| = O(\gamma_0^{\frac{1}{n_*-1}})$.

By Lemma 3.1 2), it is easy to see that N_{ν} converges uniformly on G_* to

$$N_{\infty} = e_{\infty} + \langle \Omega_{\infty}, y \rangle + \frac{1}{2}\left\langle \begin{pmatrix} y \\ z \end{pmatrix}, \mathcal{M}^{\infty}\begin{pmatrix} y \\ z \end{pmatrix} \right\rangle + h_0(x, y, z, \lambda)$$

with

$$
\begin{aligned}
|e_{\infty} - e_0|_{\Lambda_*} &= O(\gamma_0^{n+1}s_0\varepsilon_0\zeta_0), \\
|\Omega_{\infty} - \Omega_0|_{\Lambda_*} &= O(\gamma_0^{n+1}s_0\varepsilon_0\zeta_0), \\
|\omega_{\infty} - \omega_0|_{\Lambda_*} &= O(\gamma_0^{n+1}s_0\varepsilon_0\zeta_0), \\
|\mathcal{M}^{\infty} - \mathcal{M}^0|_{\mathcal{D}(\frac{r_0}{2})\times\Lambda_*} &= O(\gamma_0^{n+1}\varepsilon_0\zeta_0).
\end{aligned}
$$

And as in [7], we get the convergence of Ψ^{ν} on G_* with the estimate

$$|\Psi^{\infty} - id|_{G_*} = O(\varepsilon_0\zeta_0) = O(\varepsilon_0^{\frac{8}{9}}).$$

Thus, we obtain that the perturbed system (1.2) possesses an analytic, quasi-periodic, invariant torus with the Diophantine toral frequency $\omega_\infty(\lambda) = -E^\top(\lambda) \cdot \Omega_\infty(\lambda)$ for each $\lambda \in \Lambda_*$. By Lemma 3.1 3), we have

$$(\Omega_\infty(\lambda))_i = (\Omega_0(\lambda))_i, \quad \lambda \in \Lambda_*, \ i = 1, 2, \cdots, n_0,$$

i.e., the perturbed pseudo-frequencies preserve the first n_0 components of their corresponding ones.

In particular, when $n_0 = l$, it is clear that $U^0 = [A^0]$, $U = [A]$, $\mathrm{diag}(I_{n_0}, O) = I_l$, and, $\mathrm{diag}(O, I_{l-n_0}) = O$. Hence, $\Omega_\nu \equiv \Omega_0$ for all $\nu = 0, 1, \cdots$, i.e., $\Omega_\infty \equiv \Omega_0, \omega_\infty \equiv \omega_0$. So we obtain that when $[A]$ is nonsingular, the Diophantine frequencies remain unchanged under small perturbations.

5 Some Examples

In this section we give some examples to illustrate our results. At first, we give an example for the persistence of invariant tori in generalized Hamiltonian systems.

Example 1. We consider the following unperturbed system:

$$N(y, u) = y + \frac{1}{2}y^2 + \frac{1}{2}(u^2 - v^2),$$

where $y, u, v \in R^1, x = (x_1, x_2)^\top \in T^2$, that is, $l = 2, n = 1, m = 1$, i.e., the system is an odd dimensional generalized Hamiltonian. The structure matrix in tangent direction I is assumed to be

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

where α, β, γ are arbitrary real numbers with $|\alpha| + |\beta| + |\gamma| \neq 0$. It is easy to see that

$$\Omega = 1 + y, \quad \omega = \begin{pmatrix} -\alpha(1+y) \\ -\beta(1+y) \end{pmatrix}, \quad A = (1).$$

It is easy to verify that the Rüssmann condition is not satisfied, but A is always nonsingular. So by Theorem 1.1 we obtain that the majority 2-tori will persist with unchanged toral frequency.

Then we give an example to illustrate the persistence of invariant tori on sub-manifolds in generalized Hamiltonian systems. For the persistence of elliptic invariant tori and mixed type of invariant tori on sub-manifolds, the reader can refer to [11] for details.

Example 2. We consider the following unperturbed system:

$$N(y, u) = \frac{1}{2}y_1^2 + \frac{1}{2}y_2^2 + \frac{1}{2}(u_1^2 - v_1^2) + \frac{1}{2}(u_2^2 - 2v_2^2),$$

where $u_i, v_i \in R^1, i = 1, 2$, and $x = (x_1, x_2)^\top \in T^2, y = (y_1, y_2, y_3)^\top \in R^3$, that is $l = 3, n = 2, m = 2$, i.e., the system is an odd dimensional generalized Hamiltonian. The structure matrix in tangent direction I is assumed to be

$$I = \begin{pmatrix} 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 & 0 \end{pmatrix}.$$

We consider the persistence of invariant tori on sub-manifold $M : y_3 = a, a \in R$. It is easy to see that

$$\Omega = \begin{pmatrix} y_1 \\ y_2 \\ 0 \end{pmatrix}, \quad \omega = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

By simple verification we see that the Rüssmann condition holds and A is always singular on the sub-manifold M. So by Theorem 1.1 we have that the first two components of Ω remain unchanged. And by the form of $\omega(\lambda)$ we obtain that the majority 2-tori will persist with unchanged toral frequency.

Acknowledgements The authors express their sincere thanks to Professor Yong Li for his instructions and encouragement.

References

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

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

[3] Kuksin, S. B., Hamiltonian perturbations of infinite dimensional linear systems with an imaginary spectrum, *Funct. Anal. Appl.,* **21**(1987), 192–205.

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

[5] Graff, S. M., On the continuation of hyperbolic invariant tori for Hamiltonian systems, *J. Differential Equations,* **15**(1974), 1–69.

[6] Zehnder, E., Generalized implicit function theorem with applications to some small divisors, I and II, *Comm. Pure Appl. Math.,* **28**(1975), 91–140; **29**(1976), 49–111.

[7] Li, Y. and Yi, Y., Persistence of hyperbolic tori in Hamiltonian systems, *J. Differential Equations,* **208**(2005), 344–387.

[8] Li, Y. and Yi, Y., Persistence of invariant tori in generalized Hamiltonian systems, *Ergodic Theory Dynam. Systems,* **22**(2002), 1233–1261,

[9] Xu, J., You, J. and Qiu, Q., Invariant tori for nearly integrable Hamiltonian systems with degeneracy, *Math. Z.,* **226**(1997), 375–387.

[10] Chow, S.-N., Li, Y. and Yi, Y., Persistence of invariant tori on submanifolds in Hamiltonian systems, *J. Nonlinear Sci.*, **12**(2002), 585–617.

[11] Liu, Z., Persistence of lower dimensional invariant tori on sub-manifolds in Hamiltonian systems, *Nonlinear Anal.,* **61**(2005), 1319–1342.