Journal of Applied Analysis and Computation
Volume 12, Number 6, December 2022, 2621–2639
Website:http://www.jaac-online.com
DOI:10.11948/20220236

ARNOLD-TYPE THEOREM ABOUT LOWER-DIMENSIONAL INVARIANT TORI IN GENERALIZED HAMILTONIAN SYSTEMS

Jiayin Du

Abstract In this paper, we consider the lower dimensional invariant tori for generalized Hamiltonian system. A so-called generalized Hamiltonian system is defined on a Poisson manifold which can be odd dimensional and structurally degenerate. The existence of quasi-periodic invariant tori for Hamiltonian with standard symplectic structure was first shown by Arnold [2] under a degeneracy-removing condition. We prove the persistence of lower dimensional tori for Hamiltonian with Poisson structure instead of standard symplectic structure, when the tangential and normal frequencies satisfy some certain conditions.

Keywords KAM theory, generalized Hamiltonian system, Arnold-type theorem, lower-dimensional tori.

MSC(2010) 37J40, 70H08.

1. Introduction and main result

The Hamiltonian system is frequently studied to describe models arising in celestial mechanics or the motion of charged particles in magnetic fields, see [4,19,30]. KAM theory, established by Kolmogorov, Arnold and Moser, is a landmark of the development of Hamiltonian system. It gave a reasonable explanation for the stability of solar system and brought a new method into the study of dynamical systems. The classical KAM theory [1, 13, 21, 25] pointed out that with respect to the standard symplectic form on $2n$-dimensional smooth manifold $(\mathbb{R}^n \times \mathbb{T}^n, \omega^2)$, most invariant n-tori of an integrable Hamiltonian system will persist to small perturbations under certain non-degenerate conditions. A generalization to $(l + n)$-dimensional symplectic manifold $(\mathbb{R}^l \times \mathbb{T}^n, \omega^2)$, with $l < n$, $(l+n)$ even, is in [5,12,23,24,33]. The corresponding Hamiltonian system is odd-dimensional while $(l + n)$ is odd. Since there is no symplectic structure for odd-dimensional systems, some results in classical Hamiltonian systems no longer hold, which makes it very difficult to develop KAM theory for odd-dimensional Hamiltonian systems, as pointed out in [6,20,28]. Later, paper [15] proved some KAM types of results for generalized Hamiltonian systems defined on Poisson manifolds $(\mathbb{R}^l \times \mathbb{T}^n, \omega^2)$, where the Poisson structure or 2-form ω^2 is defined by structure matrix $I(y)$ with

$$I = \begin{pmatrix} O & B \\ -B^\top & C \end{pmatrix}, \quad C^\top = -C.$$

Email: dujy20@mails.jlu.edu.cn(J. Du)
College of Mathematics,Jilin University,Changchun 130012, China

There the system can be odd dimension and permit more action than angle variables. Moreover, the 2-form is not necessary non-degenerate.

The persistence of lower dimensional tori for Hamiltonian systems is an essential problem and has been proved under many different conditions, see [3,7–10,14,16,18, 22,27,34,35] for usual Hamiltonian systems, see [31,32] for multi-scale Hamiltonian systems, see [17] for generalized Hamiltonian systems, which is defined on Poisson manifold $(\mathbb{R}^l \times \mathbb{T}^n \times \mathbb{R}^{2m}, \omega^2)$ with structure matrix

$$I(y) = \begin{pmatrix} O & B & O_1 \\ -B^\top & C & O_2 \\ O_3 & O_4 & J \end{pmatrix},$$

where O, O_1, O_2, O_3, O_4 are zero matrices, $C^\top = -C$, J is the standard symplectic matrix.

In this paper, we investigate the persistence of lower dimensional tori for generalized Hamiltonian systems with two-scales. It was Arnold in [2] proved that Hamiltonian system with two-scales admits full dimensional invariant tori under a degeneracy-removing condition. Similar results for multi-scale Hamiltonian systems have also been obtained in [11,26,29,32,36].

Consider the Poisson manifold $(G \times \mathbb{T}^n \times \mathbb{R}^d, \omega^2)$, where $G \subset \mathbb{R}^l$ is a bounded, connected and closed region, $\mathbb{T}^n$ is the standard n-torus, l, n, d are given positive integers. The structure matrix

$$I = (A_{ij}) : G \times \mathbb{T}^n \times \mathbb{R}^d \to \mathbb{R}^{l+n+d}$$

associated with 2-form ω^2 is a real analytic, anti-symmetric, matrix valued function and satisfies the following two conditions:

(i) rank $I > 0$,

(ii) Jacobi identity

$$\sum_{m=1}^{l+n+2d} (A_{im} \frac{\partial A_{jk}}{\partial w_m} + A_{jm} \frac{\partial A_{ki}}{\partial w_m} + A_{km} \frac{\partial A_{ij}}{\partial w_m}) = 0, \tag{1.1}$$

for all $w = (y, x, z) \in G \times \mathbb{T}^n \times \mathbb{R}^d, i, j, k = 1, 2, \cdots, l + n + d.$

The 2-form ω^2 is required to be invariant relative to $\mathbb{T}^n$. Hence its coefficients structure matrix I is independent of $x \in \mathbb{T}^n$, i.e. $I = I(y, z), y \in G, z \in \mathbb{R}^d$.

On the Poisson manifold $(G \times \mathbb{T}^n \times \mathbb{R}^d, \omega^2)$, we consider the generalized Hamiltonian

$$H(y, x, z, \xi) = e(\xi) + \langle \omega_\varepsilon(\xi), y \rangle + \frac{1}{2}\varepsilon \langle M(\xi) z, z \rangle + \varepsilon^2 P(y, x, z, \xi, \varepsilon), \tag{1.2}$$

where $(y, x, z) \in G \times \mathbb{T}^n \times \mathbb{R}^d$, $\xi \in G$, $\omega_\varepsilon(\xi) = (\omega^0(\xi)^\top, \varepsilon\omega^1(\xi)^\top)^\top$, $\omega^0(\xi) \in \mathbb{R}^m$, $\omega^1(\xi) \in \mathbb{R}^{l-m}$, $m < l$, $M(\xi)$ is a symmetric matrix valued function, P is a real analytic function.

Then the equation of motion of (1.2) associated to the 2-form ω^2 reads

$$\begin{pmatrix} \dot{y} \\ \dot{x} \\ \dot{z} \end{pmatrix} = I(y,z)\nabla H. \tag{1.3}$$

We further require that the unperturbed system associated to (1.2) is completely integrable, i.e., $y = (y_1, \cdots, y_l)^\top \in G$ need to satisfy the involution conditions:

$$\{y_i, y_j\} = 0, \quad i, j = 1, 2, \cdots, l.$$

Hence the structure matrix I must have the form

$$I = \begin{pmatrix} O & B & O_1 \\ -B^\top & C & D \\ O_2 & -D^\top & E \end{pmatrix}, \tag{1.4}$$

where $O = O_{l,l}$, $O_1 = O_{l,d}$, and $O_2 = O_{d,l}$ are zero matrices, $B = B_{l,n}$, $C = C_{n,n}$ with $C^\top = -C$, $D = D_{n,d}$, $E = E_{d,d}$ with $E^\top = -E$.

Consider the integrable part of (1.2):

$$N := e(\xi) + \langle \omega_\varepsilon(\xi), y\rangle + \frac{1}{2}\varepsilon\langle M(\xi)z, z\rangle.$$

Then, the associated Hamiltonian system reads

$$\begin{cases} \dot{y} = 0 \\ \dot{x} = -B^\top \partial_y N + D\partial_z N \\ \dot{z} = E\partial_z N. \end{cases} \tag{1.5}$$

Obviously the system (1.5) admits n-dimensional invariant tori $\mathbb{T}^n_\xi = \{(0, x, 0) : x \in \mathbb{T}^n\}$ carrying quasi-periodic linear flow $\{x_0 + \omega(\xi)t\}$, where $\omega(\xi) = -B^\top(0,0)\partial_y N(0,0) = -B^\top(0,0)\omega_\varepsilon(\xi)$.

Like the case for a standard nearly integrable Hamiltonian system, the problem of the persistence of lower dimensional, quasi-periodic, invariant tori of (1.2) then concerns the persistence of the majority of these n-tori $\{\mathbb{T}^n_\xi\}$ as $\varepsilon \to 0$.

Denote $\omega(\xi) =: (\Omega^0(\xi)^\top, \varepsilon\Omega^1(\xi)^\top)^\top$, where $\Omega^0(\xi) \in \mathbb{R}^{n_0}$, $\Omega^1(\xi) \in \mathbb{R}^{n-n_0}$ with $0 \le n_0 \le n$.

We can now state our conditions:

(A0) Rank $\{\frac{\partial^i \Omega}{\partial \xi^i} : i \in \mathbb{Z}^n_+, |i| \le n-1\} = n$, $\Omega = (\Omega^0(\xi)^\top, \Omega^1(\xi)^\top)^\top$.

(A1) The set $\{\xi \in G : \sqrt{-1}\langle k, \Omega^1(\xi)\rangle - \lambda_p(\xi) - \lambda_q(\xi) \ne 0,\ \textit{for all } k \in \mathbb{Z}^{n-m} \setminus \{0\}, 1 \le p, q \le d\}$ admits full Lebesgue measure relative to G , where $\lambda_1(\xi), \cdots, \lambda_d(\xi)$ are the eigenvalues of $M(\xi)E(\xi, 0)$.

Our main result is the following:

Theorem 1.1. *Consider the Hamiltonian (1.2) and assume **(A0)-(A1)**.*

Then there is an $\varepsilon_0 > 0$ such that when $0 < \varepsilon < \varepsilon_0$, there exists a Cantor set $G_\varepsilon \subset G$ with $|G \setminus G_\varepsilon| \to 0$, as $\varepsilon \to 0$, such that for any $\xi \in G_\varepsilon$ the unperturbed tori $\{\mathbb{T}^n_\xi\}$ persist and give rise to a Whitney smooth family of quasi-periodic, invariant n-tori with slightly deformed Diophantine frequencies.

Remark 1.1. For simplicity, the generalized Hamiltonian system with only two-scales has been considered in Theorem 1.1. For more general case, multi-scale generalized Hamiltonian systems will be studied in our forthcoming paper.

The proof of our theorem uses KAM procedure. It should be emphasized that for the generalized Hamiltonian (1.2), not only does the y, z dependence of a structure matrix need to be taken into consideration in all KAM steps, but also the degeneracy of the structure matrix. So, a correction term will be introduced to modify the iterative scheme typically used in standard KAM theory, see subsection 2.2 for details.

The paper is organized in the following way. In Section 2, we give the detailed construction and estimates for one cycle of KAM steps, and by finite-times KAM steps we get a new normal form. In Section 3, using the new normal form mentioned in Section 2 and scale transformation, we complete the proof of Theorem 1.1.

Throughout the paper, unless specified explanation, we shall use the same symbol $|\cdot|$ to denote an equivalent vector norm and its induced matrix norm, absolute value of functions, and measure of sets, etc., and denote by $|\cdot|_D$ the sup-norm of functions on a domain D. Also, for any two column vectors η, ζ of the same dimension, $\langle \eta, \zeta \rangle$ always means $\eta^\top \zeta$, i.e. the transpose of η times ζ.

2. KAM Step

In this section, we will show the detailed construction and estimates for one cycle of KAM steps, which is essential to study the KAM theory, see [11,15,16,26,31,32].

Denote $0 < \delta < \frac{1}{12}$. Let $0 < N < (d^2 - 1)(\tau + 1)$ be a positive integer and $b = d^2(N + 1)$. We first define the following 0-th KAM step parameters:

$$\begin{aligned} &r_0 = r, \qquad \gamma_0 = \varepsilon^{\frac{\delta}{b}}, \qquad s_0 = \varepsilon^{\frac{1-3\delta}{2}}, \qquad \mu_0 = \varepsilon^{2\delta}, \\ &e_0 = e(\xi), \quad \omega_0 = \omega_\varepsilon(\xi), \quad M_0 = M(\xi), \quad G_0 = G, \\ &D(s_0, r_0) := \{(y, x, z) : |y| < s_0^2, |z| < s_0, |\mathrm{Im} x| < r_0\}. \end{aligned} \tag{2.1}$$

Therefore, we have that

$$\begin{aligned} H_0 &=: H(y, x, z, \xi) = N_0 + \varepsilon P_0, \\ N_0 &=: N_0(y, z, \xi) = e_0 + \langle \omega_0(\xi, \varepsilon), y \rangle + \frac{1}{2} \varepsilon \langle M_0(\xi) z, z \rangle, \\ P_0 &=: \varepsilon P(y, x, z, \xi, \varepsilon) = \varepsilon P, \end{aligned}$$

and

$$| \, \partial_\xi^l P_0 \, |_{D(s_0, r_0) \times G_0} \leq \gamma_0^b s_0^2 \mu_0, \quad |l| \leq N.$$

We now define the ν-th KAM step parameters:

$$r_\nu = \frac{r_{\nu-1}}{2} + \frac{r_0}{4}, \quad s_\nu = \frac{1}{8} s_{\nu-1}^{1+\frac{1}{3}}, \quad \mu_\nu = \mu_{\nu-1} s_{\nu-1}^\delta, \quad G_\nu \subset G.$$

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

$$H_\nu = N_\nu + \varepsilon P_\nu, \tag{2.2}$$

defined on $D(s_\nu, r_\nu)$ and

$$|\partial_\xi^l P_\nu|_{D(s_\nu,r_\nu)\times G_\nu} \le \gamma_0^b s_\nu^2 \mu_\nu, \quad |l| \le N.$$

For simplicity, we will omit the index for all quantities of the ν-th KAM step and use $+$ to index all quantities in the $(\nu+1)$-th KAM step. To simplify the notations, we will not specify the dependence of P, P_+ etc.

In the following, we will construct a generalized canonical transformation Φ_+ (preserving the 2-form ω^2 invariant), which, on a smaller phase domain D_+ and a smaller parameter domain G_+, transforms (2.2) into the Hamiltonian with the following form

$$H_+ = H \circ \Phi_+ = N_+ + \varepsilon P_+,$$

enjoying the similar properties to (2.2) but with a much smaller perturbation P_+.

All the constants $c_1 - c_5$ below are positive and independent of the iteration process, and we will also use c to denote any intermediate positive constant which is independent of the iteration process. Define

$$\begin{aligned}
&r_+ = \frac{r}{2} + \frac{r_0}{4},\\
&s_+ = \frac{1}{8}\alpha s, \qquad \alpha = \mu^{2\rho} = s^{\frac{1}{3}},\\
&\mu_+ = c_0 \mu s^\delta, \qquad c_0 = \max\{1, c_1, c_2, \cdots, c_5\},\\
&K_+ = ([\log \frac{1}{s}] + 1)^3,\\
&D(s) = \{(y,z) : |y| < s^2, |z| < s\},\\
&\hat{D} = D(s, r_+ + \frac{7}{8}(r - r_+)),\\
&D_i = D(\frac{i}{8}\alpha s, r_+ + \frac{i-1}{8}(r - r_+)), \quad i = 1, 2, \cdots, 8,\\
&D_+ = D_{\frac{1}{8}\alpha} = D(s_+, r_+),\\
&\Gamma(r - r_+) = \sum_{0<|k|\le K_+} |k|^{N+(N+1)4d^2\tau} e^{-|k|\frac{r-r_+}{8}}.
\end{aligned}$$

2.1. Truncation

Consider the Taylor-Fourier series of P:

$$P = \sum_{k\in\mathbb{Z}^n,\ \imath,\jmath\in\mathbb{Z}_+^n} p_{k\imath\jmath} y^\imath z^\jmath e^{\sqrt{-1}\langle k,x\rangle},$$

and let R be the truncation of P of the form

$$\begin{aligned}
R &= \sum_{|k|\le K_+,\ 2|\imath|+|\jmath|\le 2} p_{k\imath\jmath} y^\imath z^\jmath e^{\sqrt{-1}\langle k,x\rangle}\\
&= \sum_{|k|\le K_+} (p_{k00} + \langle p_{k10}, y\rangle + \langle p_{k01}, z\rangle + \langle z, p_{k02} z\rangle) e^{\sqrt{-1}\langle k,x\rangle}.
\end{aligned} \tag{2.3}$$

Lemma 2.1. *Assume that*

$$\textbf{(H1)}: \int_{K_+}^{\infty} t^n e^{-t^{\frac{r-r_+}{16}}} dt \le s.$$

Then there is a constant c_1 such that for all $\xi \in G$, $|l| \le N$,

$$|\partial_\xi^l (P-R)|_{D_7\times G} \le c_1 \gamma_0^b s^3 \mu, \tag{2.4}$$

$$|\partial_\xi^l R|_{D_7\times G} \le c_1 \gamma_0^b s^2 \mu. \tag{2.5}$$

Proof. Denote

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

where

$$I = \sum_{|k|>K_+,\ \imath,\jmath\in\mathbb{Z}_+^n} p_{k\imath\jmath} y^\imath z^\jmath e^{\sqrt{-1}\langle k,x\rangle}, \quad II = \sum_{|k|\le K_+,\ 2|\imath|+|\jmath|>2} p_{k\imath\jmath} y^\imath z^\jmath e^{\sqrt{-1}\langle k,x\rangle}.$$

By a similar argument to [10,32], we have for any $|l| \le N$

$$|\partial_\xi^l I|_{\hat{D}(s)\times G} \le c\gamma_0^b s^3 \mu,$$

and

$$|\partial_\xi^l II|_{D_7\times G} \le \frac{c}{s^3} \Big| \int (P-I)_{\hat{D}(s)} dydz \Big|_{D_7\times G} \le c\gamma_0^b s^3 \mu,$$

where $\int$ is the obvious anti-derivative of $\partial_{(y,z)}^{(p,q)}$ with $2|p|+|q|=3$. Therefore,

$$|\partial_\xi^l (P-R)|_{D_7\times G} \le c\gamma_0^b s^3 \mu$$

and

$$|\partial_\xi^l R|_{D_7\times G} \le c\gamma_0^b s^2 \mu.$$

$\square$

2.2. The modified homological equation

In the following, we will find a generalized Hamiltonian F such that, under the transformation of the time-1 map ϕ_F^1 of the flow generated by F , we can eliminate all resonant terms in R:

$$p_{k\imath\jmath} y^\imath z^\jmath e^{\sqrt{-1}\langle k,x\rangle}, \qquad 0 < |k| \le K_+, 2|\imath| + |\jmath| \le 2.$$

We first construct a generalized Hamiltonian F of the form

$$F = \sum_{0<|k|\le K_+,\ 2|\imath|+|\jmath|\le 2} F_{k\imath\jmath} y^\imath z^\jmath e^{\sqrt{-1}\langle k,x\rangle} + \langle F_{001}, z\rangle, \tag{2.6}$$

which satisfies the equation

$$\{N, F\} + \varepsilon(R - [R] + \langle p_{001}, z\rangle) - Q = 0, \tag{2.7}$$

where $[R] = \frac{1}{(2\pi)^n}\int_{\mathbb{T}^n} R(y,x,z)dx$ is the average of the truncation R, and the correction term

$$Q = \sum_{0<|k|\le K_+,\ 2|\imath|+|\jmath|\le 2} \sqrt{-1}\langle k, \omega(\xi)(B^\top(y,z) - B^\top(0,0))\rangle F_{k\imath\jmath} y^\imath z^\jmath e^{\sqrt{-1}\langle k,x\rangle}$$

$$
\begin{aligned}
&+\sum_{0<|k|\le K_+,\ 2|\imath|+|\jmath|\le 2}\sqrt{-1}\langle z,\varepsilon M(\xi)(D^\top(y,z)-D^\top(0,0))k\rangle F_{k\imath\jmath}y^\imath z^\jmath e^{\sqrt{-1}\langle k,x\rangle}\\
&+\sum_{0<|k|\le K_+}\langle z,\varepsilon M(\xi)(E^\top(y,z)-E^\top(0,0))(F_{k01}+F_{k02}z+F_{k02}^\top z)\rangle e^{\sqrt{-1}\langle k,x\rangle}\\
&+\sum_{0<|k|\le K_+}\sqrt{-1}\langle z,\varepsilon M(\xi)D^\top(0,0)k(F_{k01}^\top y+z^\top F_{k02}z)\rangle e^{\sqrt{-1}\langle k,x\rangle}\\
&+\langle z,\varepsilon M(\xi)(E(y,z)-E(0,0)F_{001})\rangle. \qquad (2.8)
\end{aligned}
$$

Substituting (2.3), (2.6) and (2.8) into (2.7) and comparing the coefficients, we then obtain the following equations for all $0<|k|\le K+$,

$$
\begin{aligned}
&\sqrt{-1}\langle k,(-B^\top(0,0)\omega(\xi))\rangle F_{k00}=\varepsilon p_{k00}, &(2.9)\\
&\sqrt{-1}\langle k,(-B^\top(0,0)\omega(\xi))\rangle F_{k10}=\varepsilon p_{k10}, &(2.10)\\
&\quad\sqrt{-1}\langle k,(-B^\top(0,0)\omega(\xi))\rangle F_{k01}-\varepsilon M(\xi)E(0,0)F_{k01}\\
&=\varepsilon p_{k01}-\sqrt{-1}\varepsilon M(\xi)D^\top(0,0)kF_{k00}, &(2.11)\\
&\quad\sqrt{-1}\langle k,(-B^\top(\xi,0)\omega(\xi))\rangle F_{k02}-\varepsilon M(\xi)E(0,0)F_{k02}-F_{k02}\varepsilon M(\xi)E(0,0)\\
&=\varepsilon p_{k02}-\sqrt{-1}\varepsilon M(\xi)D^\top(0,0)kF_{k01}^\top, &(2.12)\\
&\varepsilon M(\xi)E(0,0)F_{001}=\varepsilon p_{001}. &(2.13)
\end{aligned}
$$

Denote

$$
\begin{aligned}
L_{k0}&=\frac{1}{\varepsilon}(\sqrt{-1}\langle k,(-B^\top(0,0)\omega(\xi))\rangle),\\
L_{k1}&=\frac{1}{\varepsilon}(\sqrt{-1}\langle k,(-B^\top(0,0)\omega(\xi))\rangle I_d-\varepsilon M(\xi)E(0,0)),\\
L_{k2}&=\frac{1}{\varepsilon}(\sqrt{-1}\langle k,(-B^\top(0,0)\omega(\xi))\rangle I_{d^2}-\varepsilon M(\xi)E(0,0)\otimes I_d-I_d\otimes\varepsilon M(\xi)E(0,0)),
\end{aligned}
$$

where $\otimes$ stands for the tensor product of matrices and I_d is a d order identity matrix. Then (2.9)-(2.12) become

$$
\begin{aligned}
&L_{k0}F_{k00}=p_{k00}, &(2.14)\\
&L_{k0}F_{k10}=p_{k10}, &(2.15)\\
&L_{k1}F_{k01}=p_{k01}-\sqrt{-1}M(\xi)D^\top(0,0)kF_{k00}, &(2.16)\\
&L_{k2}F_{k02}=p_{k02}-\sqrt{-1}M(\xi)D^\top(0,0)kF_{k01}^\top, &(2.17)
\end{aligned}
$$

which are clearly solvable as long as all L_{k0},L_{k1},L_{k2} are invertible.

Consider the set

$$
G_+=\{\xi\in G:|L_{k0}^0|>\frac{\gamma_0}{|k|^\tau},|\det L_{k1}^0|>\frac{\gamma_0^d}{|k|^{d\tau}},|\det L_{k2}^0|>\frac{\gamma_0^{d^2}}{|k|^{d^2\tau}},\forall 0<|k|\le K_+\},\tag{2.18}
$$

where

$$
\begin{aligned}
L_{k0}^0&=\frac{1}{\varepsilon}\sqrt{-1}\langle k,(-B^\top(0,0)\omega_0(\xi))\rangle,\\
L_{k1}^0&=\frac{1}{\varepsilon}(\sqrt{-1}\langle k,(-B^\top(0,0)\omega_0(\xi))\rangle I_d-\varepsilon M_0(\xi)E(0,0)),
\end{aligned}
$$

$$L_{k2}^0 = \frac{1}{\varepsilon}(\sqrt{-1}\langle k, (-B^\top(0,0)\omega_0(\xi))\rangle I_{d^2} - \varepsilon M_0(\xi)E(0,0)\otimes I_d - I_d\otimes \varepsilon M_0(\xi)E(0,0)).$$

Then, we can solve equations (2.9)-(2.13) on G_+. The details can be seen in the following lemma:

Lemma 2.2. *Assume that*

($\boldsymbol{H2}$) : $\max\{s^{\frac{1}{3}}, \varepsilon\}K_+^{d^2(\tau+1)} = o(\gamma_0)$,

($\boldsymbol{H3}$) : $|\partial_\xi^l(M(\xi) - M_0(\xi))|_G = O(\mu_0^{\frac{1}{2}}), |\partial_\xi^l(\omega(\xi) - \omega_0(\xi))|_G = O(\varepsilon), \forall |l| \le N.$

Then the equations (2.9)-(2.13) can be uniquely solved on $\hat{D}(s)\times G_+$ to obtain F which satisfy the following equality:

$$|F|, |F_x|, s^2|F_y|, s|F_z| \le c_2 s^2\mu\Gamma(r - r_+), \tag{2.19}$$

and

$$|\partial_\xi^l\partial_x^\ell\partial_y^i\partial_z^j F| \le c_2\mu\Gamma(r - r_+), \tag{2.20}$$

where $|l| \le N$, $2|i| + |j| \le 2$, c_2 is a constant.

Proof. For any $0 < |k| \le K_+$, denote

$$L = \frac{1}{\varepsilon}(\sqrt{-1}\langle k, -B^\top(0,0)(\omega - \omega_0)\rangle I_d - \varepsilon(M(\xi) - M_0(\xi))E(0,0)).$$

It follows from (**H3**) that L is bounded from the above by a constant. Notice that

$$L_{k1} = L_{k1}^0 + \varepsilon L,$$

we have by (**H2**) that

$$|\det L_{k1}|_{G_+} \ge |\det L_{k1}^0| - O(\varepsilon)K_+^d \ge \frac{\gamma_0^d}{2|k|^{d\tau}}, \tag{2.21}$$

and

$$|L_{k1}^{-1}|_{G_+} = |(L_{k1}^0 + \varepsilon L)^{-1}| \le c\frac{|(L_{k1}^0)^{-1}|}{1 - |(L_{k1}^0)^{-1}||\varepsilon L|} \le c\frac{|k|^{d\tau+d-1}}{\gamma_0^d}. \tag{2.22}$$

Similarly,

$$|L_{k0}|_{G_+} \ge \frac{\gamma_0}{2|k|^\tau}, \qquad |L_{k2}^{-1}|_{G_+} \le c\frac{|k|^\tau}{\gamma_0}, \tag{2.23}$$

$$|\det L_{k2}|_{G_+} \ge \frac{\gamma_0^{d^2}}{2|k|^{d^2\tau}}, \quad |L_{k2}^{-1}|_{G_+} \le c\frac{|k|^{d^2\tau+d^2-1}}{\gamma_0^{d^2}}. \tag{2.24}$$

Using (2.21), (2.23), (2.22), and (2.24) and applying the identity

$$\partial_\xi^l L_{kq}^{-1} = -\sum_{|j'|=1}^{|j|}\begin{pmatrix} j \\ j' \end{pmatrix}(\partial_\xi^{j-j'}L_{kq}^{-1}\partial_\xi^{j'}L_{kq})L_{kq}^{-1}$$

inductively, it is easy to see that

$$|\partial_\xi^l L_{kq}^{-1}| \leq c\frac{|k|^{|l|+(|l|+1)(d)^q\tau}}{\gamma_0^{(d)^q(|l|+1)}}, \quad |l| \leq N, \quad q = 0,1,2. \tag{2.25}$$

By Cauchy's estimate, we also have

$$|\partial_\xi^l p_{k\imath\jmath}|_G \leq |\partial_\xi^l P|_{D(s,r)\times G} s^{-(2\imath+\jmath)} e^{-|k|r} \leq c\gamma_0^b s^{2-(2\imath+\jmath)}\mu e^{-|k|r}. \tag{2.26}$$

Notice by (2.9)-(2.10) that

$$F_{k\imath 0} = L_{k0}^{-1} p_{k\imath 0}, \quad \imath = 0,1,$$

then by (2.25)

$$|\partial_\xi^l F_{k\imath 0}| \leq c\frac{|k|^{|l|+\tau}}{\gamma_0^{|l|+1}}\gamma_0^b s^{2-(2\imath)}\mu e^{-|k|r}, \quad \imath = 0,1. \tag{2.27}$$

Recall from (2.16) that

$$F_{k01} = L_{k1}^{-1}(p_{k01} - \sqrt{-1}M(\xi)D^\top(0,0)kF_{k00}).$$

Note by **(H2)** that

$$|\partial_\xi^l (M(\xi)D^\top(0,0)kF_{k00})|_{G_+} \leq c\frac{K_+^{|l|+\tau+1}}{\gamma_0^{|l|+1}}\gamma_0^b s^2\mu e^{-|k|r} \leq c\gamma_0^b s\mu e^{-|k|r}. \tag{2.28}$$

This together with (2.25) yields

$$|\partial_\xi^l F_{k01}|_{G_+} \leq c\frac{|k|^{|l|+(|l|+1)d\tau}}{\gamma_0^{(d)(|l|+1)}}\gamma_0^b s\mu e^{-|k|r}. \tag{2.29}$$

By a similar argument, we can prove

$$|\partial_\xi^l F_{k02}|_{G_+} \leq c\frac{|k|^{|l|+(|l|+1)d^2\tau}}{\gamma_0^{(d^2)(|l|+1)}}\gamma_0^b \mu e^{-|k|r}. \tag{2.30}$$

It follows from (2.26), (2.27), (2.29), and (2.30) that

$$|\partial_\xi^l F_{001}|_{G_+} \leq cs\mu, \tag{2.31}$$

$$|\partial_\xi^l F_{k\imath\jmath}|_{G_+} \leq c\frac{|k|^{|l|+(|l|+1)d^2\tau}}{\gamma_0^{(d^2)(|l|+1)}}\gamma_0^b s^{2-(2\imath+\jmath)}\mu e^{-|k|r}, \; 0<|k|<K_+. \tag{2.32}$$

Thus

$$|\partial_\xi^l\partial_x^\ell\partial_y^i\partial_z^j F|_{\hat{D}(s)\times G_+} \leq c\sum_{0<|k|\leq K_+}|k|^{|\ell|}(\partial_\xi^l F_{k00} + \partial_\xi^l F_{k10}s^{2-2i} + \partial_\xi^l F_{k01}s^{1-j} \tag{2.33}$$

$$+ \partial_\xi^l F_{k02}s^{2-j})e^{|k|(r_+ + \frac{7}{8}(r-r_+))} + \partial_\xi^l\partial_x^\ell\partial_y^i\partial_z^j\langle F_{001}, z\rangle$$

$$\leq c(s^2 + s^{2-2|i|} + s^{2-|j|})\mu \tag{2.34}$$

$$\sum_{0<|k|\le K_+} |k|^{|\ell|+|l|+(|l|+1)d^2\tau} e^{-|k|\frac{r-r_+}{8}}$$
$$\le c\mu\Gamma(r-r_+),$$

which implies (2.19) and (2.20). □

Let $\Phi_+ = \phi_F^1$ be the time-1 map of the equation of motion associated to F, i.e.,

$$\begin{pmatrix} \dot{y} \\ \dot{x} \\ \dot{z} \end{pmatrix} = I(y,z)\nabla F(y,x,z). \tag{2.35}$$

Then Φ_+ is a canonical transformation, and

$$\begin{aligned} H\circ\Phi_+ &= H\circ\phi_F^1 = (N+\varepsilon R)\circ\phi_F^1 + \varepsilon(P-R)\circ\phi_F^1 \\ &= (N+\varepsilon R) + \{N,F\} + \int_0^1 \{(1-t)\{N,F\}+\varepsilon R, F\}\circ\phi_F^t dt \\ &\quad + \varepsilon(P-R)\circ\phi_F^1 \\ &= N + \varepsilon([R]-\langle p_{001}, z\rangle) + \varepsilon\int_0^1 \{R_t, F\}\circ\phi_F^t dt + \varepsilon(P-R)\circ\phi_F^1 + Q \\ &=: N_+ + \varepsilon P_+, \end{aligned}$$

and

$$N_+ = N + \varepsilon([R]-\langle p_{001}, z\rangle), \tag{2.36}$$
$$P_+ = \int_0^1 \{R_t, F\}\circ\phi_F^t dt + (P-R)\circ\phi_F^1 + \frac{1}{\varepsilon}Q, \tag{2.37}$$
$$R_t = \frac{1}{\varepsilon}(1-t)\{N,F\} + R.$$

It should be pointed that it is due to the Jacobi identity (1.1) that the structure matrix $I(y,z)$ is kept unchanged at each KAM step.

2.3. Estimates of the new Hamiltonian

Now, we give the estimate of N+.

Lemma 2.3. *There is a constant c_3 such that for all $|l|\le N$:*

$$|\partial_\xi^l(e_+ - e)|_{G_+} \le c_3\varepsilon\gamma_0^b s^2\mu, \tag{2.38}$$
$$|\partial_\xi^l(\omega_+ - \omega)|_{G_+} \le c_3\varepsilon\gamma_0^b\mu, \tag{2.39}$$
$$|\partial_\xi^l(M_+ - M)|_{G_+} \le c_3\gamma_0^b\mu. \tag{2.40}$$

Proof. The lemma follows easily from (2.26) and (2.36). □

Next, we estimate Φ_+.

Lemma 2.4. *Assume that*

$$(\boldsymbol{H4}): c\mu\Gamma(r-r_+) < \frac{1}{8}(r-r_+),$$

$$(\boldsymbol{H5}): cs^2\mu\Gamma(r-r_+) < s_+^2.$$

Then the followings hold.

(1) Let ϕ_F^t *be the flow generated by equation (2.35). Then*

$$\phi_F^t : D_3 \to D_4, \quad \textit{for all } 0 \le t \le 1.$$

(2) Let $\Phi_+ = \phi_F^1$*. Then for each* $\xi \in G_+$*,*

$$\Phi_+ : D_+ \to D(s,r).$$

(3) There is a constant c_4 *such that*

$$|\partial_\xi^l(\phi_F^t - id)|_{D_+\times G_+} \le c_4\mu\Gamma(r-r_+),\ \forall t\in[0,1],\ |l|\le N, \tag{2.41}$$

$$|\partial_\xi^l\partial_x^\ell\partial_y^i\partial_z^j(\Phi_+ - id)|_{D_+\times G_+} \le c_4\mu\Gamma(r-r_+),\ \forall |l|\le N,\ |\ell|,|i|,|j| = 0,1. \tag{2.42}$$

Proof. Let X_F be the vector field on the right-hand side of (2.35). Then

$$\phi_F^t = id + \int_0^t X_F\circ\phi_F^\lambda d\lambda := (\phi_{F1}^t, \phi_{F2}^t, \phi_{F3}^t)^\top, \qquad 0\le t\le 1. \tag{2.43}$$

For any $(y,x,z)\in D_3$, we let $t_* = \sup\{t\in[0,1] : \phi_F^t(x,y,z)\in D_4\}$. It follows from (**H4**), (**H5**) and Lemma 2.2 that

$$\begin{aligned}
|\phi_{F_2}^t(y,x,z)|_{D_3} &\le |y| + \int_0^t |B(\phi_{F1}^\lambda,\phi_{F3}^\lambda)F_x\circ\phi_F^\lambda|_{D_3}d\lambda\\
&\le (3s_+)^2 + cs^2\mu\Gamma(r-r_+)\\
&< (4s_+)^2,\\
|\phi_{F_2}^t(y,x)|_{D_3} &\le |x| + \int_0^t |-B^\top(\phi_{F1}^\lambda,\phi_{F3}^\lambda)F_y\circ\phi_F^\lambda + C(\phi_{F1}^\lambda,\phi_{F3}^\lambda)F_x\circ\phi_F^\lambda\\
&\quad + D(\phi_{F1}^\lambda,\phi_{F3}^\lambda)F_z\circ\phi_F^\lambda|_{D_3}d\lambda\\
&\le r_+ + \frac{2}{8}(r-r_+) + c\mu\Gamma(r-r_+) + cs^2\mu\Gamma(r-r_+) + cs\mu\Gamma(r-r_+)\\
&< r_+ + \frac{2}{8}(r-r_+) + c\mu\Gamma(r-r_+)\\
&< r_+ + \frac{3}{8}(r-r_+),\\
|\phi_{F_3}^t(x,y)|_{D_3} &\le |z| + \int_0^t |-D^\top(\phi_{F1}^\lambda,\phi_{F3}^\lambda)F_x\circ\phi_F^\lambda + E(\phi_{F1}^\lambda,\phi_{F3}^\lambda)F_z\circ\phi_F^\lambda|_{D_3}d\lambda\\
&\le 3s_+ + c_4s^2\mu\Gamma(r-r_+) + c_4s\mu\Gamma(r-r_+)\\
&\le 3s_+ + \frac{s_+^2}{s}\\
&< 4s_+.
\end{aligned}$$

Thus, $\phi_F^t\in D_4$, i.e. $t_*=1$ and (1) holds.

(2) follows from (1).

(3) Using Lemma 2.2 and the argument above, we immediately have

$$|\phi_F^t - id| \leq c\mu\Gamma(r - r_+).$$

By Lemma 2.2 and the Gronwall Inequality to the identity

$$\begin{aligned}\nabla\phi_F^t &= Id + \int_0^t (\nabla(I\nabla F))\nabla\phi_F^\lambda d\lambda \\ &= Id + \int_0^t (\nabla I \cdot \nabla F) \circ \phi_F^\lambda \cdot \nabla\phi_F^\lambda + (I\nabla^2 F) \circ \phi_F^\lambda \cdot \nabla\phi_F^\lambda d\lambda,\end{aligned}$$

we have

$$\begin{aligned}|\nabla\phi_F^t - Id| &\leq \int_0^t (|\nabla I||\nabla F| + |I||\nabla^2 F|)|\nabla\phi_F^\lambda - Id|d\lambda + (|\nabla I||\nabla F| + |I||\nabla^2 F|) \\ &\leq c\mu\Gamma(r - r_+).\end{aligned}$$

□

For the new perturbation P_+ , we have the following estimate.

Lemma 2.5. *Let* $\Delta := s^3\mu^2\Gamma^2(r - r_+) + \gamma_0^b s^3\mu + s^3\mu\Gamma(r - r_+)$. *Then there is a constant* c_5 *such that for all* $|l| \leq N$, $|\partial_\xi^l P_+| \leq c_5\Delta$. *Moreover, if*

$$(\boldsymbol{H6}) : c_5\Delta \leq \gamma_0^b s_+^2\mu_+,$$

then

$$|\partial_\xi^l P_+|_{D_+\times G_+} \leq \gamma_0^b s_+^2\mu_+, \quad |l| \leq N. \tag{2.44}$$

Proof. Recall the definition of Q as in (2.8), then

$$\begin{aligned}|\partial_\xi^l(\frac{1}{\varepsilon}Q)|_{D_+\times G_+} &\leq \alpha s^3\mu\Gamma(r - r_+) + K_+\alpha s^3\mu\Gamma(r - r_+) + \alpha s^4\mu\Gamma(r - r_+) \\ &\quad + \alpha s^3\mu\Gamma(r - r_+) + \alpha K_+ s^3\mu\Gamma(r - r_+) + \alpha s^3\mu \\ &\leq cs^3\mu\Gamma(r - r_+).\end{aligned}$$

Let

$$W = \int_0^1 \{R_t, F\} \circ \phi_F^t dt.$$

By a similar argument to [15], we have

$$|\partial_\xi^l W|_{D_+\times G_+} \leq cs^3\mu^2\Gamma^2(r - r_+).$$

Recall that

$$P_+ = W + (P - R) \circ \phi_F^1 + \frac{1}{\varepsilon}Q.$$

The above estimates together with Lemma 2.1 imply that

$$|\partial_\xi^l P_+|_{D_+\times G_+} \leq cs^3\mu^2\Gamma^2(r - r_+) + \gamma_0^b s^3\mu + s^3\mu\Gamma(r - r_+) = c\Delta.$$

□

2.4. Finite-times KAM steps

Following the standard KAM theory, what we should do is to proceed KAM steps. To this end, we shall find a family of iteration sequences, r_ν, s_ν, α_ν, μ_ν, K_ν, Φ_ν, D_ν, such that all assumptions **(H1)**-**(H6)** hold. Particularly, let

$$\begin{aligned}
r_\nu &= r_0(\frac{1}{2}+\frac{1}{2^{\nu+1}}),\\
s_\nu &= \frac{1}{8}\alpha_{\nu-1}s_{\nu-1},\\
\alpha_\nu &= s_\nu^{2\rho} = s_\nu^{\frac{1}{3}},\\
\mu_\nu &= c_0\mu_{\nu-1}s_{\nu-1}^{\delta},\\
K_\nu &= ([\log(\frac{1}{s_{\nu-1}})]+1)^3,\\
D_\nu &= D(s_\nu, r_\nu),\\
G_\nu &= \{\xi\in G_{\nu-1}: |L_{k0}^0| > \frac{\gamma_0}{|k|^\tau}, |\det L_{k1}^0| > \frac{\gamma_0^d}{|k|^{d\tau}}, |\det L_{k2}^0| > \frac{\gamma_0^{d^2}}{|k|^{d^2\tau}},\\
&\qquad \forall 0<|k|\le K_\nu\}.
\end{aligned}$$

Note that

$$r_{\nu-1}-r_\nu = \frac{r_0}{2^{\nu+1}},$$

and

$$s_\nu = (\frac{1}{8})^{3((1+\frac{1}{3})^\nu-1)} s_0^{(1+\frac{1}{3})^\nu}. \tag{2.45}$$

Similar to [11,26,32], hypotheses **(H1)**, **(H2)**, **(H4)** and **(H5)** hold for all ν and sufficiently small ε besides the part of hypotheses (**H2**), i.e.,

$$\varepsilon K_+^{d^2(\tau+1)} = o(\gamma_0), \tag{2.46}$$

which will only holds if the number of the iterations is finite. Indeed, if we take

$$\nu_* = [\frac{\log(2l(N+6)+1)-\log(1-3\delta)}{\log(\frac{4}{3})}]+1,$$

then it is easy to see that (2.46) holds as $\varepsilon\ll 1$ for all $\nu=1,2,\cdots,\nu_*$.

It follows from (2.39) and (2.40) in Lemma 2.3 that

$$\begin{aligned}
|\partial_\xi^l(\omega_\nu-\omega_0)| &\le c\varepsilon(\mu_{\nu-1}+\mu_{\nu-2}+\cdots+\mu_0) = O(\varepsilon),\\
|\partial_\xi^l(M_\nu-M_0)| &\le c(\mu_{\nu-1}+\mu_{\nu-2}+\cdots+\mu_0) = O(\mu_0^{\frac{1}{2}}),
\end{aligned}$$

which implies **(H3)**.

We now prove **(H6)**. Recall the definition of Δ, we have

$$|\Delta|\le c\gamma_0^b s^{2(1+\frac{1}{3})}\mu s^\delta (s^{\frac{1}{3}-\delta}\varepsilon^{-\delta}\mu\Gamma^2(r-r_+)+s^{\frac{1}{3}-\delta}+s^{\frac{1}{3}-\delta}\varepsilon^{-\delta}\Gamma(r-r_+))\le\gamma_0^b s_+^2\mu_+.$$

Let

$$r_*=r_{\nu_*},\quad s_*=\mu_{\nu_*},\quad \gamma_*=\varepsilon^l,\quad \mu_*=\varepsilon^{\frac{1+\delta}{2}},\quad G_*=G_{\nu_*},$$

$$H_* = H_{\nu_*}, \quad N_* = N_{\nu_*}, \quad P_* = \varepsilon P_{\nu_*}, \quad \Psi_* = \Phi_1 \circ \Phi_2 \circ \cdots \circ \Phi_{\nu_*}.$$

Hence after ν_* times KAM steps, the Hamiltonian can be written as follows:

$$H_* = H_0 \circ \Psi_* = N_*(y,z,\xi) + P_*(y,x,z,\varepsilon), \tag{2.47}$$

$$N_*(y,z,\xi) = e_{\nu_*} + \langle \omega_{\nu_*}(\xi,\varepsilon), y\rangle + \frac{1}{2}\varepsilon\langle M_{\nu_*}(\xi)z, z\rangle + g(y,z,\varepsilon). \tag{2.48}$$

By (2.45), we can choose $\varepsilon \ll 1$ independent of ν such that

$$s_\nu^2 < \varepsilon^{(1-3\delta)(1+\frac{1}{3})^\nu}, \nu = 1, 2, \cdots, \nu_*.$$

It follows that

$$s_{\nu_*}^2 \le \varepsilon^{2l(N+6)+1} = \varepsilon\gamma_*^{2(N+6)},$$

and hence

$$|P_*|_{D(s_*,r_*)} \le \varepsilon\gamma_0^b s_{\nu_*}^2 \mu_{\nu_*} \le \varepsilon\gamma_*^{2(N+6)} s_* \mu_*^2.$$

2.5. Measure estimates

We now estimate the excluded measure of G_0 after ν_*-th iteration. The first lemma shows the connection between the condition **(A1)** with the small divisor condition involved in (2.18) for $\nu = 0$. The second lemma deals with measure estimates. For simplicity, we denote $\Omega_0(\xi) =: -B^\top(0,0)\omega_0(\xi) =: (\Omega_0^0(\xi)^\top, \varepsilon\Omega_0^1(\xi)^\top)^\top$, where $\Omega_0^0 \in \mathbb{R}^{n_0}$, $\Omega_0^1 \in \mathbb{R}^{n-n_0}$, $0 \le n_0 \le n$.

Lemma 2.6. *For all $k \in \mathbb{Z}^n$*

$$\det L_{k1}^0 = \prod_{p=1}^{d} \frac{1}{\varepsilon}(\sqrt{-1}\langle k, \Omega_0(\xi)\rangle - \varepsilon\lambda_p(\xi)),$$

$$\det L_{k2}^0 = \prod_{p,q=1}^{d} \frac{1}{\varepsilon}(\sqrt{-1}\langle k, \Omega_0(\xi)\rangle - \varepsilon\lambda_p(\xi) - \varepsilon\lambda_q(\xi)).$$

Proof. Let $\mathcal{J}$ be the Jordan canonical form of $M_0(\xi)E(0,0)$ and let T be the non-singular matrix such that $T^{-1}M_0(\xi)E(0,0)T = \mathcal{J}$. Then

$$\begin{aligned}\det L_{k1}^0 &= \det \frac{1}{\varepsilon}(\sqrt{-1}\langle k, \Omega_0(\xi)\rangle I_d - \varepsilon M_0(\xi)E(0,0)) = \det \frac{1}{\varepsilon}(\sqrt{-1}\langle k, \Omega_0(\xi)\rangle I_d - \varepsilon\mathcal{J}) \\ &= \prod_{p=1}^{d} \frac{1}{\varepsilon}(\sqrt{-1}\langle k, \Omega_0(\xi)\rangle - \varepsilon\lambda_p(\xi)).\end{aligned}$$

Since

$$T^{-1} \otimes T^{-1}(M_0E \otimes I_d + I_d \otimes M_0E)(T \otimes T) = \mathcal{J} \otimes I_d + I_d \otimes \mathcal{J},$$

it follows that

$$\begin{aligned}\det L_{k2}^0 &= \det \frac{1}{\varepsilon}(\sqrt{-1}\langle k, \Omega_0(\xi)\rangle I_{d^2} - \varepsilon M_0(\xi)E(0,0) \otimes I_d - I_d \otimes \varepsilon M_0(\xi)E(0,0)) \\ &= \det \frac{1}{\varepsilon}(\sqrt{-1}\langle k, \Omega_0(\xi)\rangle I_{d^2} - \varepsilon\mathcal{J} \otimes I_d - \varepsilon I_d \otimes \mathcal{J})\end{aligned}$$

$$= \prod_{p,q=1}^{d} \frac{1}{\varepsilon}(\sqrt{-1}\langle k, \Omega_0(\xi)\rangle - \varepsilon\lambda_p(\xi) - \varepsilon\lambda_q(\xi)).$$

□

Lemma 2.7. $|G_0 \setminus G_*| \to 0$ *as* $\varepsilon \to 0$.

Proof. Recall that

$$G_* = G_{\nu_*} = \{\xi \in G_0 : |L^0_{k0}| > \frac{\gamma_0}{|k|^\tau}, |\det L^0_{k1}| > \frac{\gamma_0^d}{|k|^{d\tau}}, |\det L^0_{k2}| > \frac{\gamma_0^{d^2}}{|k|^{d^2\tau}},$$
$$\forall 0 < |k| \le K_{\nu_*}\}.$$

Consider the set $G_{**} = G^1 \cap G^2 \cap G^3$, where

$$G^1 = \{\xi \in G_0 : |\langle k, \Omega(\xi)\rangle| > \frac{2\gamma_0}{|k|^\tau}, \Omega(\xi) = (\Omega^0_0(\xi)^\top, \Omega^1_0(\xi)^\top)^\top, k \in \mathbb{Z}^n,$$
$$\forall 0 < |k| \le K_{\nu_*}\},$$
$$G^2 = \{\xi \in G_0 : |\sqrt{-1}\langle k^1, \Omega^1_0(\xi)\rangle - \lambda_p(\xi)| > \frac{2\gamma_0}{|k|^\tau}, k^1 \in \mathbb{Z}^{n-n_0}, \forall 0 < |k^1| \le K_{\nu_*},$$
$$0 < p \le d\},$$
$$G^3 = \{\xi \in G_0 : |\sqrt{-1}\langle k^1, \Omega^1_0(\xi)\rangle - \lambda_p(\xi) - \lambda_q(\xi)| > \frac{2\gamma_0}{|k|^\tau}, k^1 \in \mathbb{Z}^{n-n_0},$$
$$\forall 0 < |k^1| \le K_{\nu_*}, 0 < p, q \le d\}.$$

Let $\xi \in G_{**}$. For given $k = (k^0, k^1)^\top$ with $0 < |k| \le K_{\nu_*}$, let k^i, for $i = 0, 1$, be the first nonzero components of k. Since $\xi \in G^1$ and **(H2)**, we have

$$|L^0_{k0}| = \frac{1}{\varepsilon}|\langle k^0, \Omega^0_0(\xi)\rangle + \langle k^1, \varepsilon\Omega^1_0(\xi)\rangle| > \frac{1}{2\varepsilon}|\langle k^i, \varepsilon^i\Omega^i_0(\xi)\rangle| > \frac{1}{2\varepsilon^{1-i}}\frac{2\gamma_0}{|k|^\tau} > \frac{\gamma_0}{|k|^\tau}.$$

We first consider a typical term $l_{k1} := \frac{1}{\varepsilon}(\sqrt{-1}\langle k^0, \Omega^0_0\rangle + \sqrt{-1}\langle k^1, \varepsilon\Omega^1_0\rangle - \varepsilon\lambda_p)$ of L_{k1}. There are the following two cases:

(i) $|k^0| \ne 0$. It follows from $\xi \in G^1$ and **(H2)** that

$$|l_{k1}| \ge \frac{1}{\varepsilon}|(\sqrt{-1}\langle k^0, \Omega^0_0\rangle + \sqrt{-1}\langle k^1, \varepsilon\Omega^1_0\rangle - \varepsilon\lambda_p)| > \frac{1}{2\varepsilon}\frac{2\gamma_0}{|k|^\tau} > \frac{\gamma_0}{|k|^\tau}.$$

(ii) $|k^0| = 0, |k^1| \ne 0$. It follows from $\xi \in G^2$ and **(H2)** that

$$|l_{k1}| \ge \frac{1}{\varepsilon}|(\sqrt{-1}\langle k^1, \varepsilon\Omega^1_0\rangle - \varepsilon\lambda_p)| > \frac{2\gamma_0}{|k|^\tau} > \frac{\gamma_0}{|k|^\tau}.$$

Then we have

$$|\det L^0_{k1}| > \frac{\gamma_0^d}{|k|^{d\tau}}.$$

By a similar argument, we also conclude that

$$|\det L^0_{k2}| > \frac{\gamma_0^{d^2}}{|k|^{d^2\tau}}.$$

Above all, it implies that $G_{**} \subset G_*$. We note that **(A0)** implies that $|G_0 \backslash G^1| \to 0$ as $\varepsilon \to 0$. As a similar argument in [16], we have

$$\begin{aligned} &\{\xi \in G_0 : \sqrt{-1}\langle k^1, \Omega_0^1\rangle - \lambda_p - \lambda_q \neq 0, k^1 \in \mathbb{Z}^{l-m}\} \\ \subset &\{\xi \in G_0 : \sqrt{-1}\langle k^1, \Omega_0^1\rangle - \lambda_p \neq 0, k^1 \in \mathbb{Z}^{l-m}\}. \end{aligned}$$

Then, it follows from **(A1)** that $|G_0 \setminus G^3| \to 0$ as $\varepsilon \to 0$. Consequently,

$$|G_0 \setminus G_*| \to 0$$

as $\varepsilon \to 0$. □

3. Proof of Main Result

In this section, we will perform infinite steps of standard KAM iterations to the normal form (2.47) to prove the main theorem.

Consider the rescaling

$$y \to \varepsilon\gamma_*^{N+6}\mu_* y, \qquad z \to \varepsilon\gamma_*^{N+6}\mu_* z, \qquad H_* \to \frac{H_*}{\varepsilon\gamma_*^{N+6}\mu_*},$$

then the Hamiltonian reads

$$H_0 := \frac{H_*}{\varepsilon\gamma_*^{N+6}\mu_*} = e_0 + \langle\omega_0, y\rangle + \frac{1}{2}\langle z, M_0 z\rangle + P_0$$

defined on $D(s_0, r_0) \times G_0$, where $r_0 =: r_*$, $s_0 =: s_*$, $\mu_0 = \mu_*$, $G_0 =: G_*$, $e_0 =: \frac{e_*}{\varepsilon\gamma_*^{N+6}\mu_*}$, $\omega_0 =: \omega_*$, $M_0 = \varepsilon^2\gamma_*^{N+6}\mu_* M_*$, and

$$P_0 =: \frac{P_*}{\varepsilon\gamma_*^{N+6}\mu_*}$$

with

$$|\partial_\xi^l P_0|_{D(r_0,s_0)\times G_0} \leq \gamma_0^{N+6} s_0\mu_0, \quad |l| \leq N.$$

Let $r_0, s_0, \gamma_0, \mu_0, H_0, N_0, P_0$ be given at the beginning of this section and let $D_0 = D(s_0, r_0)$, $K_0 = 0$, $\Phi_0 = id$. We define the following sequence inductively for all $\nu = 1, 2, \cdots$.

$$\begin{aligned} &r_\nu = r_0\left(\frac{1}{2} + \frac{1}{2^{\nu+1}}\right), \\ &\beta_\nu = \beta_0\left(\frac{1}{2} + \frac{1}{2^{\nu+1}}\right), \quad \beta_0 = s_0, \\ &s_\nu = \frac{1}{8}\alpha_{\nu-1}s_{\nu-1}, \\ &\alpha_\nu = \mu_\nu^{2\rho} = \mu_\nu^{\frac{1}{3}}, \\ &\mu_\nu = c\mu_{\nu-1}^{1+\rho}, \\ &D_\nu = D(s_\nu, r_\nu), \\ &\tilde{D}_\nu = D\left(\beta_\nu, r_\nu + \frac{6}{8}(r_{\nu-1} - r_\nu)\right), \end{aligned}$$

$$K_\nu = ([\log(\frac{1}{\mu_{\nu-1}})]+1)^3,$$
$$L_{k1}^{\nu-1} = \sqrt{-1}\langle k, \omega_{\nu-1}\rangle I_d - M_{\nu-1}E, \quad 0<|k|\le K_\nu,$$
$$L_{k2}^{\nu-1} = \sqrt{-1}\langle k, \omega_{\nu-1}\rangle I_{d^2} - M_{\nu-1}E\otimes I_d - I_d\otimes M_{\nu-1}E, \quad 0<|k|\le K_\nu,$$
$$G_\nu = \{\xi\in G_{\nu-1} : |\langle k,\omega_{\nu-1}\rangle| > \frac{\gamma_{\nu-1}}{|k|^\tau}, |\det L_{k1}^{\nu-1}| > \frac{\gamma_{\nu-1}^d}{|k|^{d\tau}}, |\det L_{k2}^{\nu-1}| > \frac{\gamma_{\nu-1}^{d^2}}{|k|^{d^2\tau}},$$
$$\forall 0<|k|\le K_\nu\}.$$

The following iteration lemma is a special case contained in [10, 32].

Lemma 3.1. *Let ε be sufficiently small. Then the following hold for all $\nu = 1, 2, \cdots$.*

(1) There is a symplectic, real analytic, near identity transformation

$$\Phi_\nu = \phi_F^1 : \tilde{D}_\nu \to \tilde{D}_{\nu-1}, \qquad \forall \xi\in G_\nu,$$

such that

$$H_\nu = H_{\nu-1}\circ\Phi_\nu = e_\nu + \langle\omega_\nu, y\rangle + \frac{1}{2}\langle z, M_\nu z\rangle + P_\nu,$$

where

$$G_\nu = \{\xi\in G_{\nu-1} : |\langle k,\omega_{\nu-1}\rangle| > \frac{\gamma_{\nu-1}}{|k|^\tau}, |\det L_{k1}^{\nu-1}| > \frac{\gamma_{\nu-1}^d}{|k|^{d\tau}},$$
$$|\det L_{k2}^{\nu-1}| > \frac{\gamma_{\nu-1}^{d^2}}{|k|^{d^2\tau}}, \forall 0<|k|\le K_\nu\}, \tag{3.1}$$
$$|\partial_\xi^l(\omega_\nu-\omega_0)|_{G_\nu} \le \gamma_0^{N+6}\mu_0^{\frac{1}{4}}, \tag{3.2}$$
$$|\partial_\xi^l(M_\nu-M_0)|_{G_\nu} \le \gamma_0^{N+6}\mu_0^{\frac{1}{4}}, \tag{3.3}$$
$$|\partial_\xi^l P_\nu|_{D_\nu\times G_\nu} < \gamma_0^{N+6}s_\nu\mu_\nu. \tag{3.4}$$

(2) The Whitney extensions of

$$\Psi^\nu = \Phi_1\circ\Phi_2\circ\cdots\circ\Phi_\nu, \quad \nu=1,2,\cdots$$

converge C^N uniformly to a smooth symplectic map, say, Ψ^∞, on $D(\frac{s_0}{2},\frac{r_0}{2})\times G_\infty$ with $G_\infty = \bigcap_{\nu\ge 0} G_\nu$ such that

$$H_\nu = H_0\circ\Psi^{\nu-1} \to H_\infty =: H_0\circ\Psi^\infty = e_\infty + \langle\omega_\infty, y\rangle + \frac{1}{2}\langle z, M_\infty z\rangle + P_\infty$$

with $e_\infty = \lim_{\nu\to} e_\nu$, $\omega_\infty = \lim_{\nu\to}\omega_\nu$, $M_\infty = \lim_{\nu\to} M_\nu$, $P_\infty = \lim_{\nu\to\infty} P_\nu$ and moreover

$$\partial_{(y,z)}^i P_\infty|_{D(0,\frac{r_0}{2})\times G_\infty} = 0, \quad |i|\le 2.$$

It follows from the same argument as in [16] that

$$|G_0\setminus G_\infty| \to 0, \text{ as } \varepsilon\to 0.$$

Acknowledgments

The author sincerely thank Professor Li Yong for instructions and many invaluable suggestions.

References

[1] V. I. Arnold, *Proof of a theorem of A. N. Kolmogorov on the preservation of conditionally periodic motions under a small perturbation of the Hamiltonian*, Uspehi Mat. Nauk, 1963, 18(5), 13–40.

[2] V. I. Arnold, *Small denominators and problems of stability of motion in classical and celestial mechanics*, Uspehi Mat. Nauk, 1963, 18(6), 91–192.

[3] J. Bourgain, *Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE*, Internat. Math. Res. Notices, 1994, 11, 475–497.

[4] L. Chierchia, *Periodic solutions of the planetary N-body problem*, XVIIth International Congress on Mathematical Physics, World Sci. Publ., Hackensack, NJ, 2014, 269–280.

[5] F. Cong and Y. Li, *Existence of higher dimensional invariant tori for Hamiltonian systems*, J. Math. Anal. Appl., 1998, 222(1), 255–267.

[6] R. de la Llave, *Recent progress in classical mechanics*, Mathematical physics, X(Leipzig, 1991), 3–19, Springer, Berlin, 1992.

[7] A. Delshams, M. Gonchenko and P. Gutiérrez, *Exponentially small splitting of separatrices associated to 3D whiskered tori with cubic frequencies*, Comm. Math. Phys., 2020, 378(3), 1931–1976.

[8] L. H. Eliasson, *Perturbations of stable invariant tori for Hamiltonian systems*, Ann. Scuola Norm. Sup. Pisa Cl. Sci., 1988, 15(1), 115–147.

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

[10] Y. Han, Y. Li and Y. Yi, *Degenerate lower-dimensional tori in Hamiltonian systems*, J. Differential Equations, 2006, 227(2), 670–691.

[11] Y. Han, Y. Li and Y. Yi, *Invariant tori in Hamiltonian systems with high order proper degeneracy*, Ann. Henri Poincaré, 2010, 10(8), 1419–1436.

[12] M. R. Herman, *Exemples de flots hamiltoniens dont aucune perturbation en topologie $\mathbb{C}^\infty$ n'a d'orbites périodiques sur un ouvert de surfaces d'énergies*, C. R. Acad. Sci. Paris Sér. I Math., 1991, 312(13), 989–994.

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

[14] S. B. Kuksin, *Hamiltonian perturbations of infinite dimensional linear systems with an imaginary spectrum*, Funktsional. Anal. i Prilozhen., 1987, 21(3), 22–37.

[15] Y. Li and Y. Yi, *Persistence of invariant tori in generalized Hamiltonian systems*, Ergodic Theory Dynam. Systems, 2002, 22(4), 1233–1261.

[16] Y. Li and Y. Yi, *Persistence of lower dimensional tori of general types in Hamiltonian systems*, Trans. Amer. Math. Soc., 2005, 357(4), 1565–1600.

[17] B. Liu, W. Zhu and Y. Han, *Persistence of lower-dimensional hyperbolic invariant tori for generalized Hamiltonian systems*, J. Math. Anal. Appl., 2006, 322(1), 251–275.

[18] V. K. Melnikov, *On certain cases of conservation of almost periodic motions with a small change of the Hamiltonian function*, Dokl. Akad. Nauk SSSR, 1965, 165, 1245–1248.

[19] K. R. Meyer, *Periodic solutions of the N-body problem*, Lecture Notes in Mathematics, 1719, Springer-Verlag, Berlin, 1999.

[20] I. Mezić and S. Wiggins, *On the integrability and perturbation of three-dimensional fluid flows with symmetry*, J. Nonlinear Sci., 1994, 4(2), 157–194.

[21] J. Moser, *On invariant curves of area-preserving mapping of an annulus*, Nachr. Akad. Wiss. Göttingen Math. Phys. Kl. II, 1962, 1962, 1–20.

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

[23] J. Moser, *Old and new applications of KAM theory, Hamiltonian systems with three or more degrees of freedom (S'Agaró, 1995)*, 184–192, NATO Adv. Sci. Inst. Ser. C: Math. Phys. Sci., 533, Kluwer Acad. Publ., Dordrecht, 1999.

[24] I. O. Parasyuk, *Preservation of multidimensional invariant tori of Hamiltonian systems*, Ukrain. Mat. Zh., 1984, 36(4), 467–473.

[25] J. Pöschel, *A lecture on the classical KAM theorem, Smooth ergodic theory and its applications (Seattle, WA, 1999)*, 707–732, Proc. Sympos. Pure Math., 69, Amer. Math. Soc., Providence, RI, 2001.

[26] W. Qian, Y. Li and X. Yang, *Multiscale KAM theorem for Hamiltonian systems*, J. Differential Equations, 2019, 266(1), 70–86.

[27] W. Qian, Y. Li and X. Yang, *Melnikov's conditions in matrices*, J. Dynam. Differential Equations, 2020, 32(4), 1779–1795.

[28] M. B. Sevryuk, *Some problems in KAM theory: conditionally periodic motions in generic systems*, Russian Math. Surveys, 1995, 50(2), 341–353.

[29] B. Sommer, *A KAM theorem for the spatial lunar problem*, Ph. D Thesis, 2003.

[30] C. E. Wayne, *An introduction to KAM theory, Dynamical systems and probabilistic methods in partial differential equations (Berkeley, CA, 1994)*, 3–29, Lectures in Appl. Math., 31, Amer. Math. Soc., Providence, RI, 1996.

[31] L. Xu, Y. Li and Y. Yi, *Lower-dimensional tori in multi-scale nearly integrable Hamiltonian systems*, Ann. Henri Poincaré, 2017, 18(1), 53–83.

[32] L. Xu and Y. Yi, *Lower dimension tori of general types in multi-scale Hamiltonian systems*, Nonlinearity, 2019, 32(6), 2226–2245.

[33] J. C. Yoccoz, *Travaux de Herman sur les tores invariants*, Astérisque, 1992, 206, 311–344.

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

[35] E. Zehnder, *Generalized implicit function theorems with applications to some small divisor problems. I and II*, Comm. Pure Appl. Math., 1975, 28, 91–140.

[36] X. Zhao and Y. Li, *A moser theorem for multi-scale mappings*, Discrete Contin. Dyn. Syst., 2022. DOI: 10.3934/dcds.2022037.